

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

Estudos sobre o duplo direcionamento de proteínas de plantas

Carolina Vianna Morgante

**Tese apresentada para obtenção do título de Doutor
em Agronomia. Área de concentração: Genética e
Melhoramento de Plantas**

**Piracicaba
2008**

Carolina Vianna Morgante
Bacharel em Ciências Biológicas

Estudos sobre o duplo direcionamento de proteínas de plantas

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

Tese apresentada para obtenção do título de Doutor em
Agronomia. Área de concentração: Genética e Melhoramento
de Plantas

Piracicaba
2008

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

Morgante, Carolina Vianna
Estudos sobre o duplo direcionamento de proteínas de plantas / Carolina Vianna
Morgante. - - Piracicaba, 2008.
138 p. : il.

Tese (Doutorado) - - Escola Superior de Agricultura Luiz de Queiroz, 2008.
Bibliografia.

1. Arroz 2. Cloroplastos 3. Mitocôndrias vegetais 4. Proteínas de plantas 5.
Papaverales I. Título

CDD 581.87

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

DEDICATÓRIA

Aos leitores, na intenção de compartilhar com vocês os prazeres que tive como “detetive da Natureza”.

AGRADECIMENTOS

Ao Departamento de Genética da Escola Superior de Agricultura “Luiz de Queiroz”, pelo apoio acadêmico de excelência.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pelo suporte financeiro no país e pelo privilégio da bolsa sanduíche.

Ao Prof. Dr. Marcio de Castro Silva Filho, pela orientação, pelas boas oportunidades concedidas e pela disponibilização de excelente infra-estrutura para a pesquisa.

Ao Dr. José Manuel Gualberto pela acolhedora estadia em Estrasburgo, pela orientação e por sua dedicação ao meu projeto.

Ao Prof. Dr. Alberto Ribeiro e Waldir Caldeira (IB-USP), pelo auxílio na realização das análises de microscopia confocal.

Ao Prof. Dr. Paulo Alberto Otto (IB - USP) e ao Prof. Dr. Antônio Augusto Franco Garcia (ESALQ - USP), pelo auxílio nas análises estatísticas.

À Profa. Dra. Margarida Aguiar Perecin, pela disponibilização do uso do microscópio de fluorescência e ao amigo Mateus Modin, pelo apoio técnico e pelas conversas científicas.

Aos funcionários e alunos do Departamento de Genética, ESALQ-USP, em especial a Rafael Colombi, pelo carinho com que sempre tratou os alunos do nosso laboratório.

Aos colegas, veteranos e novatos, do Laboratório de Biologia Molecular de Plantas, ESALQ-USP, pela amizade, companheirismo e boa convivência. Um agradecimento especial para Camila Motta Borgonove, Daniel Scherer de Moura, Mariana Cicarelli Cia e Phellipe Arthur dos Santos Marbach, co-autores de parte deste trabalho. Torcerei sempre para o sucesso profissional de todos vocês.

Aos colegas do *Institut de Biologie Moléculaire des Plantes*, pela atenção recebida durante minha estadia na França, em especial a Claire Pujol, Anne-Marie Duchêne, Sarah Holec e Monique Le Ret.

Aos meus amigos G8, Ana Bonna, Éder Jorge Oliveira, José Francisco Farias, Juliano de Pádua, Karem Xavier, Sybelle Barreira e Tassiano Maxwell Câmara. Guardo com muito carinho as lembranças dos momentos divertidos que convivemos em Piracicaba. Apesar da distância, terei sempre um sentimento muito especial por todos vocês. Lembro e repito as palavras da nossa amiga Karem: “nenhum de nós é tão bom quanto todos nós juntos”.

A minha sempre amiga Karina Martins, pela amizade sincera, pela convivência saudável nos anos de república e pelo exemplo de dedicação e coragem.

À Sybelle Barreira, pela eterna amizade e pelas longas conversas e gargalhadas. Muito obrigada por seu apoio emocional durante esses anos. Guardo nossa amizade no meu coração.

À Jeannette Voelcker, minha mãe francesa, pessoa muito especial por quem tenho imenso carinho. O meu muito obrigada pelo acolhimento tão caloroso e pelo amor materno e filial que cresceu entre nós em tão pouco tempo.

Aos meus pais, principais incentivadores da minha vida científica e exemplos de dedicação para o progresso da ciência no Brasil. Agradeço o apoio e as oportunidades que me deram durante todo o período da minha formação. O meu amor por vocês é infinito.

SUMÁRIO

RESUMO	9
ABSTRACT	10
LISTA DE ABREVIATURAS E SÍMBOLOS	11
1 INTRODUÇÃO	12
2 REVISÃO BIBLIOGRÁFICA	14
2.1 A transferência gênica entre os endossimbiontes e o núcleo	14
2.2 O direcionamento de proteínas codificadas no núcleo	17
2.2.1 O transporte de proteínas para mitocôndrias	17
2.2.2 O transporte de proteínas para cloroplastos	18
2.2.3 Translocases de membranas de mitocôndrias e cloroplastos	19
2.3 Predição da localização celular de proteínas	22
2.4 O duplo direcionamento de proteínas	23
3 MATERIAL E MÉTODOS	30
3.1 Análises de seqüências de proteínas de <i>Oryza sativa</i> e <i>Arabidopsis thaliana</i> ..	30
3.2 Clonagem de seqüências de cDNA no vetor pPMA4-GFP para expressão transiente	31
3.3 Clonagem de seqüências de cDNA no vetor pUCAP-GFP para expressão transiente	34
3.4 Clonagem de seqüências no vetor binário pBinPLUS	36
3.5 Clonagem de seqüências para transcrição e tradução <i>in vitro</i>	38
3.6 Isolamento de plasmídeos	38
3.7 Mutações sítio-dirigidas em seqüências de nucleotídeos	39
3.8 Transformação de <i>Agrobacterium tumefaciens</i>	40
3.9 Transformação transiente via biobalística	40
3.9.1 Preparo de partículas	40
3.9.2 Precipitação de DNA sobre as partículas	40
3.9.3 Bombardeamento do material vegetal	41
3.10 Transformação de <i>Nicotina benthamiana</i> por infiltração	41
3.11 Preparo de protoplastos	42

3.12 Análises ao microscópio confocal	42
3.13 Análise quantitativa das imagens obtidas no microscópio confocal	43
3.14 Análise estatística dos dados obtidos no programa ImageJ.....	45
3.15 Transcrição e tradução <i>in vitro</i>	45
3.16 Isolamento de cloroplastos	46
3.17 Isolamento de mitocôndrias	47
3.18 Importação <i>in vitro</i> de proteínas para organelas	47
3.18.1 Importação <i>in vitro</i> de proteínas para mitocôndrias	47
3.18.2 Importação <i>in vitro</i> de proteínas para cloroplastos	49
4 RESULTADOS E DISCUSSÃO	50
4.1 A conservação do duplo direcionamento de proteínas em <i>Arabidopsis thaliana</i> e <i>Oryza sativa</i> sugere um padrão similar na evolução de famílias gênicas	50
4.1.1 Busca por homólogos de proteínas com duplo direcionamento	50
4.1.2 Predição da localização celular de proteínas	51
4.1.3 Um método alternativo para a identificação de proteínas com duplo direcionamento?	57
4.1.4 O tamanho das famílias gênicas de membros com duplo direcionamento	58
4.1.5 Análise filogenética das famílias gênicas de membros com duplo direcionamento	60
4.1.6 Localização celular em <i>Oryza sativa</i> de ortólogos de proteínas com duplo direcionamento em <i>Arabidopsis thaliana</i>	62
4.1.7 O duplo direcionamento como um processo conservado	67
4.2 O duplo direcionamento de proteínas de ligação ao RNA (RBP)	71
4.2.1 Localização celular dos membros da subfamília RBP1	72
4.2.2 Mutações na seqüência de RBP1b	75
4.2.3 Análise estatística dos dados gerados pelo programa ImageJ	81
4.2.4 Conclusão das análises dos efeitos das mutações sítio-dirigidas realizadas na seqüência de direcionamento de RBP1b	93
4.2.5 Quantificação do direcionamento de RBP1b, RBP1a e RPS19 para mitocôndrias e cloroplastos	98
4.3 O duplo direcionamento da proteína THI1 de <i>Arabidopsis thaliana</i>	101

5 CONCLUSÕES 111
REFERÊNCIAS 113
ANEXOS..... 131

RESUMO

Estudos sobre o duplo direcionamento de proteínas de plantas

Na célula eucariota, os processos metabólicos estão compartimentalizados em organelas e proteínas sintetizadas no citosol são endereçadas para elas por meio de sistema celular específico. Devido a sobreposições funcionais entre organelas, uma dada proteína pode ser requerida em mais de um compartimento. É o caso de proteínas com duplo direcionamento, em que um gene nuclear é capaz de gerar produtos protéicos direcionados para mais de uma organela. Cerca de 60 proteínas de plantas já tiveram seu duplo direcionamento demonstrado, a maioria para mitocôndrias e cloroplastos, portanto um fenômeno não tão raro como imaginado. Estudos recentes procuram esclarecer os mecanismos que permitem esse duplo direcionamento, mesmo existindo fatores celulares que garantem a especificidade do transporte para cada organela. O presente trabalho está dividido em três partes. Na primeira, foram investigados aspectos evolutivos do duplo direcionamento. Na análise comparativa de famílias gênicas com membros cujos produtos protéicos apresentam duplo direcionamento, em *Arabidopsis thaliana* e *Oryza sativa*, foi demonstrada a conservação do duplo direcionamento de monodeidroascorbato redutase, metionina aminopeptidase e, provavelmente, de THI1 (enzima da biossíntese de tiazol), entre as duas espécies. Os dados sugeriram um mesmo padrão de evolução em famílias incluindo membros com duplo direcionamento. Na segunda parte, o foco foi a seqüência de duplo direcionamento. Documentado o duplo direcionamento, para mitocôndrias e cloroplastos, de proteínas de ligação ao RNA, RBP1a, RBP1b e RPS19, mutações sítio-dirigidas foram introduzidas na seqüência de direcionamento de RBP1b. A importância de aminoácidos positivos para o direcionamento de proteínas para mitocôndrias foi confirmada. Demonstrou-se que a mutação da alanina na posição 2, conservada em seqüências de direcionamento ambíguas para mitocôndrias e cloroplastos, não afeta o duplo direcionamento de RBP1b. A informação para o duplo direcionamento foi localizada entre os 17 primeiros aminoácidos da região amino-terminal. Os sinais de direcionamento para cloroplastos se distribuíram ao longo da seqüência. Enquanto a metade amino-terminal da seqüência foi suficiente para determinar o duplo direcionamento, a seqüência compreendendo os 13 aminoácidos seguintes afetaram a eficiência do transporte. Nessas análises, mostrou-se a adequação de método quantitativo na medida dos sinais de fluorescência da GFP, para ser aplicado em estudos quantitativos do direcionamento de proteínas para mitocôndrias e cloroplastos, *in vivo*. Na terceira parte, o foco foi o mecanismo traducional do duplo direcionamento de THI1, em *A. thaliana*. Entretanto, não foi possível testar a hipótese da presença de um sítio interno de entrada do ribossomo (IRES) no mRNA de *thi1*, pois não se encontrou um sistema de expressão transiente capaz de reproduzir dados da literatura que mostraram o duplo direcionamento de THI1. Nos sistemas testados, a proteína foi encontrada somente em cloroplastos, o que inviabilizou o prosseguimento da investigação.

Palavras-chave: Duplo direcionamento de proteínas; Mitocôndria; Cloroplasto; *Arabidopsis thaliana*; *Oryza sativa*

ABSTRACT

Studies on the dual targeting of plant proteins

Compartmentalization of the metabolic processes in organelles, each one having a characteristic protein pool and distinct functions, is a property of eukaryote cells. A highly specific cellular system directs proteins, which are synthesized in the cytosol, to the proper organelles. However, due to functional overlaps between organelles, a given protein may be needed in different compartments. This is the case of dual-targeted proteins, which are the product of single nuclear genes, but are somehow directed to different organelles. About 60 plant proteins have had their dual targeting demonstrated, most of them to mitochondria and chloroplasts, the phenomenon being not so rare as previously supposed. Investigations have focused on the mechanisms, which enable protein dual targeting, even in the presence of other cell mechanisms that guarantee the specific protein transport to each organelle. The present work on the subject can be divided in three parts. In the first part, evolutionary aspects of dual targeting were investigated. A comparative analysis of gene families that included members encoding dual-targeted proteins in *Arabidopsis thaliana* and *Oryza sativa* demonstrated that the dual targeting of monodehydro-ascorbate reductase, methionine aminopeptidase and, probably, of the thiazole biosynthetic enzyme THI1 was evolutionary conserved between the two species. In addition, the data suggested the same pattern of evolution for families with members presenting dual targeting. The focus of the second part was the ambiguous sequence for dual targeting. After showing that the RNA-binding proteins RBP1a, RBP1b and RPS19 were dual-targeted to mitochondria and chloroplasts, site-directed mutations were introduced in the targeting sequence of RBP1b. The importance of positive-charged amino acids for directing the protein to mitochondria was confirmed. Mutation of alanine at position 2, which is conserved in ambiguous sequences, was shown not to affect RBP1b dual targeting. Information for dual targeting was localized among the 17 first amino acids in the amino-terminal region. The signals for directing the protein to chloroplasts appeared distributed along the targeting sequence. While the amino-terminal half of the sequence was sufficient for RBP1b dual targeting, the sequence comprising the next 13 amino acids appeared to affect the efficiency of the transport. In these analyses, a quantitative method to measure the intensity of fluorescent signals of GFP had its efficacy demonstrated to be adopted for *in vivo* quantitative analysis of dual targeting to mitochondria and chloroplasts. In the third part, the focus was the translational mechanism enabling THI1 dual targeting to mitochondria and chloroplasts, in *A. thaliana*. It was hypothesized that an internal ribosomal entry site (IRES) was present in *thi1* mRNA. However, a transient expression system could not be found that reproduced the literature data demonstrating THI1 dual targeting. In the tested systems the protein was addressed solely to chloroplasts, thus preventing the objective to be pursued.

Key words: Protein dual targeting; Mitochondria; Chloroplast; *Arabidopsis thaliana*; *Oryza sativa*

LISTA DE ABREVIATURAS E SÍMBOLOS

ADP - difosfato de adenosina
ATP - trifosfato de adenosina
BLAST - *Basic Local Alignment and Search Tool*
BSA – albumina de soro bovino
cDNA - DNA complementar ao RNA
DTT - ditionitrosol
EDTA - ácido etilendiamino tetra-acético
GFP - proteína fluorescente verde
Hepes - ácido 4-(2-hidrossietil)-1-piperazinil-etanosulfônico
MAP - metionina aminopeptidase
MDAR - monodeidroascorbato redutase
MES - ácido 2-(N-morfolino)etano-sulfônico
mRNA - RNA mensageiro
NADH - forma reduzida da nicotinamida adenina dinucleotídeo
ORF - *open reading frame* ou quadro aberto de leitura
PCR - reação em cadeia da polimerase
PMSF - fluoreto de fenil-metano-sulfonil
PVP – polivinilpirrolidona
RBP - proteína de ligação ao RNA
RPS19 - proteína ribossomal S19
RT-PCR - PCR após transcrição reversa
SDS - dodecil sulfato de sódio
THI1 - proteína da biossíntese do tiazol
TIC - translocase da membrana interna de cloroplastos
TIM - translocase da membrana interna de mitocôndrias
TOC - translocase da membrana externa de cloroplastos
TOM - translocase da membrana externa de mitocôndrias

1 INTRODUÇÃO

A célula eucariota é caracterizada pela compartimentalização de processos metabólicos em organelas, estruturas especializadas envoltas por membranas. Cada uma delas apresenta seu próprio conjunto de proteínas, o que lhes confere funções distintas na célula. Uma vez que a maioria das proteínas é sintetizada no citosol, existem sistemas eficientes para o endereçamento específico de proteínas para as organelas, os quais compreendem fatores citosólicos, especificidade dos receptores encontrados nas membranas das organelas e características da seqüência primária das proteínas transportadas.

Devido à sobreposição de funções entre organelas, uma mesma proteína pode ser encontrada em mais de um compartimento celular. Nessas situações, a duplicação gênica é a estratégia mais comum, de forma que existe um gene nuclear para cada isoforma da proteína. Uma outra estratégia é a de proteínas que apresentam duplo direcionamento e, nesses casos, apenas um gene nuclear é capaz de gerar produtos que são endereçados para mais de um compartimento celular (SILVA-FILHO, 2003).

O duplo direcionamento de proteínas é um fenômeno comum entre eucariotos e já foi documentado em leveduras, plantas e animais. Em plantas, foi primeiramente demonstrado para a glutathione redutase de ervilha, a qual é endereçada simultaneamente para mitocôndrias e cloroplastos (CREISSEN et al., 1995). A partir de então, quase 60 proteínas tiveram seu duplo direcionamento evidenciado em plantas, a maioria para mitocôndrias e cloroplastos, mostrando que se trata de um fenômeno não tão raro como se imaginava. Atualmente, os principais estudos no assunto visam esclarecer os mecanismos que permitem o duplo direcionamento de proteínas, mesmo na presença de fatores celulares que garantem a especificidade do transporte de proteínas para cada organela.

O presente trabalho está dividido em três partes com objetivos distintos, porém todas com o duplo direcionamento como principal questão.

Na primeira parte (item 4.1), foi realizada uma análise comparativa de famílias gênicas com membros cujo produto protéico apresenta duplo direcionamento, em *Arabidopsis thaliana* e *Oryza sativa*. O objetivo foi pesquisar se essas famílias seguem

um mesmo padrão de evolução nas duas espécies, sendo o duplo direcionamento um processo conservado.

Na segunda parte (item 4.2), o objetivo foi identificar aminoácidos na seqüência de direcionamento, importantes para o endereçamento de proteínas para mitocôndrias e cloroplastos, contribuindo, dessa forma, para o entendimento do processo do duplo direcionamento para essas organelas. Outro objetivo foi desenvolver um sistema quantitativo para a análise *in vivo* de proteínas direcionadas para cloroplastos e mitocôndrias. Uma vez documentado o direcionamento de proteínas mitocondriais de ligação ao RNA (RBP) também para cloroplastos, as análises dos determinantes do duplo direcionamento foram realizadas em um de seus membros, RBP1b, o qual tem duplo direcionamento.

Na terceira parte (item 4.3), o objetivo foi investigar o duplo direcionamento da THI1, proteína da biossíntese do tiazol, a qual é direcionada para mitocôndrias e cloroplastos por um mecanismo pós-transcricional, em *A. thaliana*. Em seu mRNA, existem dois códons AUG, posicionados *in frame*, nos quais a tradução pode ser iniciada. O início em um ou em outro códon gera produtos protéicos distintos, cada um específico para uma das duas organelas (CHABREGAS et al., 2003). A proposta do trabalho foi investigar os mecanismos celulares que permitem o início da tradução no segundo AUG. Como existe entre os dois AUG uma região complementar ao rRNA 18S, a hipótese foi que essa região constituísse um sítio interno de entrada do ribossomo (IRES). A tradução no segundo AUG seria independente da extremidade 5' e não ocorreria pelo relaxamento do início da tradução no primeiro AUG.

2 REVISÃO BIBLIOGRÁFICA

A evolução da célula eucariota compreendeu a segregação de processos metabólicos por meio do desenvolvimento de um complexo sistema de compartimentos envoltos por membranas, representados pelas organelas, cada um com seu conjunto específico de proteínas. Apesar de a compartimentalização da célula eucariota ter permitido funções distintas a cada organela, sobreposições funcionais existem entre elas e uma mesma proteína pode ser requerida em mais de um compartimento.

A aquisição de mitocôndrias e, no caso de algas e plantas, de cloroplastos, foi fundamental ao desenvolvimento do metabolismo energético da célula eucariota. Ambas as organelas originaram-se por eventos endossimbióticos há cerca de 1,5 bilhões de anos, o que foi seguido pelo estabelecimento de um sistema de beneficiamento mútuo. Mitocôndrias originaram-se pela associação entre uma célula hospedeira e um ancestral de α -proteobactéria. Os cloroplastos surgiram a partir da adição de um segundo endossimbionte, uma cianobactéria. Diversos são os indicativos desses eventos, como a presença, nessas organelas, de genoma próprio semelhante ao de bactérias e o fato de serem envoltas por um sistema de dupla membrana (MARGULIS, 1970).

2.1 A transferência gênica entre os endossimbiontes e o núcleo

A presença de três compartimentos com genomas próprios em uma mesma célula foi acompanhada de massiva troca de material genético entre eles. Como consequência, o genoma dos endossimbiontes foi drasticamente reduzido e estes perderam sua autonomia. O genoma mitocondrial de *Arabidopsis thaliana*, por exemplo, contém 57 genes que codificam proteínas (UNSELD et al., 1997) que participam, principalmente, dos mecanismos de transporte de elétrons, replicação do DNA e expressão gênica. No entanto, o tamanho estimado de seu proteoma mitocondrial é de 2.000 proteínas (MILLAR et al., 2005). Cloroplastos possuem genoma com 50-150 genes que codificam proteínas e proteoma estimado em 2.700 proteínas (ABDALLAH; SALAMINI; LEISTER, 2000; MILLAR; WHELAN; SMALL, 2006). Essas informações revelam a magnitude da importação de proteínas por parte dessas organelas.

A transferência de genes de organelas para o núcleo, seguida da perda da cópia no genoma simbiote, foi crucial na evolução de células eucariotas e é um fenômeno ainda freqüente em plantas (HUANG; AYLIFFE; TIMMIS, 2003; STEGEMANN et al., 2003). Grandes segmentos de DNA foram transferidos, independentemente, para o núcleo durante a evolução, formando blocos multigênicos de origem organelar (LEISTER, 2005; NOUTSOS; RICHLI; LEISTER, 2005). Por exemplo, no cromossomo 2 de *A. thaliana*, foi identificado um fragmento de 270 kb de origem mitocondrial (LIN et al., 1999). Já no braço longo do cromossomo 10 de arroz, há uma inserção de aproximadamente 33 kb de DNA de origem cloroplastidial, derivados de duas regiões diferentes do genoma dessa organela. O DNA, agora nuclear, apresenta alta identidade ao do cloroplasto, o que indica uma transferência recente (YUAN et al., 2002).

Em plantas, a transferência de DNA também ocorreu a partir dos cloroplastos para mitocôndrias, compreendendo principalmente, pseudogenes (CUMMINGS et al., 2003).

Para se tornarem um gene nuclear funcional, as seqüências derivadas das organelas adquiriram uma região promotora que permitiu sua expressão no núcleo e uma seqüência de direcionamento capaz de endereçar a proteína ao compartimento celular correto. Nesse sistema, foi necessário o estabelecimento de uma maquinaria para a importação específica de proteínas nas membranas das organelas (TIMMIS et al., 2004).

Genes nucleares de origem organelar podem ter adquirido seqüências de direcionamento por diferentes mecanismos: recombinação e duplicação de seqüências pré-existentes no núcleo (KADOWAKI et al., 1996; MARTIN; HERRMANN, 1998; SANDOVAL et al., 2004; UEDA et al., 2006); acúmulo de mutações de ponto na seqüência de genes transferidos (KUBO et al., 2000); embaralhamento de exons (exon *shuffling*) (LONG et al., 1996) e *splicing* alternativo (KUBO et al., 1999). Além disso, alguns genes transferidos de organelas para o núcleo poderiam já possuir seqüências com características que facilitaram o direcionamento de seus produtos para as organelas. A evidência vem de dados que mostram que aproximadamente 5% das proteínas codificadas pelo genoma de *Escherichia coli* possuem características típicas de seqüências de direcionamento de mitocôndrias. Uma explicação para isso poderia

ser o fato de que os aminoácidos mais abundantes em seqüências de direcionamento mitocondriais estão representados com maior freqüência no código genético universal (LUCATTINI; LIKI; LITHGOW, 2004; MACKENZIE, 2005).

O redirecionamento de proteínas para as organelas e o citosol é um processo dinâmico. Quando um gene é transferido da organela para o núcleo, não há um programa específico para guiar seu produto à organela de origem. As seqüências de origem organelar podem sofrer, no núcleo, uma série de adaptações, como mudança do quadro de leitura, conservação parcial da ORF (*open reading frame*) original ou diversificação promovida por mutações não sinônimas e, assim, a proteína inicial pode adquirir nova função em novos compartimentos (NOUTSOS et al., 2007). Martin et al. (2002) demonstraram que cerca de 18% das proteínas codificadas pelo genoma nuclear em *A. thaliana* originaram-se em cianobactérias e incluem proteínas de todas as categorias funcionais, inclusive algumas funções atípicas em cianobactérias, como resistência a doenças e transporte intracelular de proteínas. Mais da metade dessas proteínas não é direcionada para cloroplastos e muitas seguem a via secretória.

Outro exemplo de redirecionamento de proteínas de origem organelar é o caso das aminoacil-tRNA sintetases. Essas enzimas catalisam a adição do aminoácido ao seu tRNA correspondente e estão presentes no citosol, cloroplastos e mitocôndrias. Análises filogenéticas demonstraram que não existem casos em que os três genes originados dos genomas ancestrais coexistem, provavelmente porque seus produtos desempenhavam funções semelhantes. Durante a evolução, ocorreu uma intensa troca dessas enzimas entre os compartimentos e hoje a maioria é compartilhada, sendo o duplo direcionamento uma regra entre elas. Das 45 aminoacil-tRNA sintetases de *A. thaliana*, 21 são exclusivamente citosólicas, 17 são compartilhadas entre mitocôndrias e cloroplastos, quatro entre mitocôndrias e citosol, duas são exclusivas de cloroplastos e uma ocorre nos três compartimentos. Suas origens são variadas, havendo enzimas mitocôndriais semelhantes às de cianobactérias, cloroplastidiais semelhantes às de Archaea e citosólicas semelhantes a enzimas de eubactérias, além de outras combinações. Apesar da coevolução entre aminoacil-tRNA sintetases e seus respectivos tRNA, não há correspondência entre eles quanto à origem, podendo uma

enzima de origem bacteriana interagir, em mitocôndrias, com um tRNA codificado no núcleo, além de outras possibilidades (DUCHÊNE et al., 2005).

2.2 O direcionamento de proteínas codificadas no núcleo

A maioria das proteínas da célula é sintetizada nos ribossomos citosólicos. Seu destino final depende da seqüência de aminoácidos, a qual pode conter a informação para o endereçamento da proteína para diferentes organelas. Proteínas que não possuem seqüência de direcionamento permanecem no citosol, enquanto aquelas transportadas para as organelas possuem características específicas que permitem seu reconhecimento e correto endereçamento. Proteínas direcionadas ao núcleo apresentam sinal de localização nuclear (NLS), com seqüência rica em aminoácidos carregados positivamente, como lisina e arginina (TERRY; SHOWS; WENTE, 2007). Já as proteínas destinadas aos peroxissomos possuem três aminoácidos característicos em sua região carboxi-terminal, SKL (Ser-Lys-Leu) (SPARKES; BAKER, 2002). O transporte de proteínas para o retículo endoplasmático é co-traducional. O sinal para o direcionamento está na região amino-terminal da proteína e é composto por cinco a dez aminoácidos hidrofóbicos (JOHNSON; VAN WAES, 1999). As proteínas direcionadas para mitocôndrias e cloroplastos também possuem os sinais para direcionamento localizados na região amino-terminal e são denominados pré-seqüência e peptídeo de trânsito, respectivamente.

2.2.1 O transporte de proteínas para mitocôndrias

As proteínas destinadas às mitocôndrias são sintetizadas no citosol. Em alguns casos, o transporte pode ser co-traducional. A pré-seqüência é fundamental para a ligação da proteína a receptores e é clivada por peptidases dentro da organela, originando a proteína madura. Apesar da função em comum, as pré-seqüências não apresentam consenso quanto à seqüência primária. No entanto, compartilham algumas características: são ricas em aminoácidos positivos e hidroxilados, possuem poucos aminoácidos ácidos e têm tendência a formar uma α -hélice anfifílica (VON HEIJNE; STEPPUHN; HERRMANN, 1989; DUBY; BOUTRY, 2002). Algumas proteínas mitocondriais são guiadas à organela por sinais de direcionamento internos que não

são clivados. É o caso de proteínas destinadas à membrana, como a proteína carreadora ADP/ATP de levedura (POLCICOVÁ et al., 2003) e a proteína Tim23 de levedura (DAVIS; RYAN; JENSEN, 1998). A região carboxi-terminal pode mediar a inserção de algumas proteínas na membrana externa (LAN; ISENMANN; WATTENBERG, 2000).

Fatores citosólicos também participam no endereçamento de proteínas para mitocôndrias. Chaperonas ligam-se a proteínas recém-sintetizadas no citosol, mantendo-as em conformação adequada para a translocação através das membranas mitocondriais. Previnem a agregação de proteínas ou dobramentos que possam permitir sua atividade biológica em compartimento inapropriado. Além disso, as chaperonas auxiliam no direcionamento de proteínas precursoras, aproximando-as das mitocôndrias e promovendo sua interação com a maquinaria de importação. Entre essas chaperonas, destaca-se a família das Hsp70, que, apesar de não serem específicas para o transporte mitocondrial, são essenciais ao processo. Outros fatores citosólicos incluem o complexo associado ao polipeptídeo nascente (NAC) e o complexo associado ao ribossomo (RAC), ambos com a função de manter a conformação da proteína precursora compatível para sua ligação com a membrana externa de mitocôndrias (BEDDOE; LITHGOW, 2002).

2.2.2 O transporte de proteínas para cloroplastos

O transporte de proteínas em plantas tem complexidade aumentada pela presença dos cloroplastos. Um sistema de importação de proteínas similar, em diversos aspectos, ao de mitocôndrias, exige mecanismos que garantam a especificidade do transporte para os cloroplastos.

O peptídeo de trânsito tem tamanho que pode variar de 20 a 120 aminoácidos. Assim como em mitocôndrias, não há consenso quanto à seqüência primária, que é altamente variável. No entanto os peptídeos de trânsito apresentam algumas características comuns: são ricos em resíduos básicos e hidrofóbicos e assumem uma estrutura secundária do tipo "espiral ao acaso". Sua região amino-terminal, não carregada, geralmente se inicia com uma metionina seguida de uma alanina e termina com glicina e prolina. O domínio central do peptídeo de trânsito é rico em serina e

treonina, com poucos resíduos ácidos. Já a região carboxi-terminal é rica em argininas e tem o potencial de formar uma estrutura β -pregueada anfifílica com aproximadamente 10 resíduos (VON HEIJNE; STEPPUHN; HERRMANN 1989; BRUCE, 2000; BRUCE, 2001).

As proteínas precursoras importadas para os tilacóides seguem a mesma via de importação que as proteínas do estroma, porém são sintetizadas com um sinal de direcionamento adicional localizado na região carboxi-terminal do peptídeo de trânsito. No estroma, a primeira parte da seqüência de direcionamento é clivada e a proteína é direcionada aos tilacóides (ROBINSON; WOOLHEAD; EDWARDS, 2000).

A regulação do processo de importação de proteínas para os cloroplastos inicia-se no citosol, onde as proteínas precursoras sofrem fosforilações antes de passarem através da membrana externa da organela. Nos peptídeos de trânsito, próximos aos sítios de fosforilação, são encontrados motivos de ligação a quinases semelhantes ao das proteínas 14-3-3. A proteína precursora fosforilada forma um complexo guia heterooligomérico que inclui um dímero de 14-3-3, Hsp70 e outras proteínas, sendo encaminhado para os receptores na membrana dos cloroplastos (MAY; SOLL, 2000).

2.2.3 Translocases de membranas de mitocôndrias e cloroplastos

A importação de proteínas para mitocôndrias e cloroplastos envolve complexos protéicos formados por diversas subunidades localizadas nas membranas das organelas. Em mitocôndrias, esses complexos são designados TOM e TIM (translocases da membrana externa e interna, respectivamente) e, em cloroplastos, TOC e TIC.

O complexo TOM é formado por dois grupos de proteínas bem definidos funcionalmente que formam o poro de importação geral ou agem como receptores que encaminham proteínas ao poro. Os componentes do complexo TOM de fungos e animais são conservados entre si quanto a estrutura e função (HOOGENRAAD; WARD; RYAN, 2002), diferindo, entretanto, dos de plantas. Apenas três subunidades do complexo TOM são comuns a esses organismos, Tom7, Tom9 (Tom22 em animais e fungos) e Tom40, que também possuem homólogos em diatomáceas, amebas e em protistas parasitas, como *Plasmodium*. Isso sugere que essas subunidades foram

estabelecidas cedo na evolução como translocases de membranas. Tom40 é a principal subunidade que forma o poro da membrana externa e há evidências de que foi o primeiro componente de translocase. As subunidades Tom7 e Tom9 promovem a interação do poro com os receptores (MAĆASEV et al., 2004).

O receptor Tom20 está bem caracterizado em plantas e, em sua estrutura, há uma região hidrofóbica capaz de interagir com pré-seqüências. Essa subunidade é estrutural e funcionalmente similar a Tom20 de fungos e animais, com algumas diferenças. Em plantas, Tom20 está ancorada na membrana externa pela sua região carboxi-terminal, enquanto a proteína de mamíferos e leveduras ancora-se pela região amino-terminal. Tom20 de plantas não apresenta homologia com proteínas de leveduras e mamíferos, tendo a similaridade entre elas resultado de processo de evolução convergente (PERRY et al., 2006).

Ao contrário das diferenças entre o complexo TOM de plantas e o de leveduras e animais, o complexo TIM apresenta homologia entre esses organismos. Todos os componentes até agora caracterizados em leveduras têm ortólogos em plantas, diferindo, no entanto, em tamanho e presença de certos motivos (MILLAR; WHELAN; SMALL, 2006).

Após passarem pela membrana externa da mitocôndria, as proteínas podem interagir com um dos dois complexos localizados na membrana interna, TIM22 e TIM23, cujas subunidades são semelhantes quanto a estrutura e função. O complexo TIM23 pode existir em duas formas modulares alternativas, associado ou não ao complexo PAM (*presequence translocase-associated motor*). PAM, cujo principal componente é uma chaperona Hsp70, é responsável por translocar a proteína para a matriz mitocondrial. TIM23 não associado ao PAM posiciona a proteína na membrana interna. Os precursores hidrofóbicos da membrana interna são também reconhecidos e nela inseridos pelo complexo TIM22, cujo funcionamento é dependente do potencial eletroquímico da membrana e independente de ATP (DOLEZAL et al., 2006; MILENKOVIC et al., 2007).

A função de diversos componentes do complexo TOC já foi bem elucidada por análises em *A. thaliana*. Assim como em mitocôndrias, existem receptores na membrana externa capazes de reconhecer peptídeos de trânsito presentes em

proteínas precursoras destinadas aos cloroplastos. Esses receptores sofrem ciclos de fosforilação-desfosforilação, que determinam o reconhecimento de peptídeos de trânsito. A identificação de proteínas precursoras é específica para aquelas relacionadas ou não à fotossíntese. Toc159, que apresenta também uma isoforma citosólica, e Toc33 reconhecem preferencialmente proteínas da via fotossintética, ao passo que Toc132, Toc120 e Toc34 reconhecem as demais proteínas. A proteína Toc64 também desempenha papel importante no reconhecimento de peptídeos de trânsito, uma vez que é capaz de interagir com chaperonas citosólicas, principalmente Hsp70 e Hsp90 ligadas às proteínas precursoras. A proteína Toc12, associada à Toc64, assiste a transferência da proteína precursora através do espaço entre membranas. No centro do complexo, associada aos receptores, está a proteína Toc75, a qual forma o canal de translocação de TOC (STENGEL; SOLL; BÖLTER, 2007).

Na maioria dos casos, a passagem da proteína pelas translocases TOC e TIC ocorre simultaneamente em um processo que requer ATP e é independente do potencial eletroquímico da membrana interna, ao contrário do que ocorre em mitocôndrias. A subunidade central desse complexo é Tic110, a qual forma o canal de translocação da membrana interna. Estudando a estrutura do receptor Tic62, Balsera et al. (2007) mostraram que sua estrutura, na forma em que é encontrada em *A. thaliana*, é exclusiva de plantas vasculares. *Physcomitrella patens*, algas, cianobactérias e bactérias sulfurosas verdes contêm uma forma truncada dessa proteína em que falta a região carboxi-terminal, apesar de a região amino-terminal de Tic62 ser conservada entre os organismos estudados. Isso mostra a origem antiga de Tic62 como receptor da membrana interna de cloroplastos e que essa subunidade em angiospermas originou-se parcialmente daquela de cianobactérias, tendo as modificações sofridas lhe conferido novas funções.

Em algumas situações, uma dada proteína pode compor translocases de mitocôndrias e cloroplastos. Uma isoforma de Toc64, denominada OM64, é encontrada na membrana externa de mitocôndrias. A função dessa proteína ainda não foi identificada, porém algumas hipóteses foram formuladas. OM64 poderia ser a substituta do receptor Tom70, ausente em plantas, ou então ser um receptor relacionado ao reconhecimento de proteínas com duplo direcionamento (CHEW et al., 2004). Além

disso, Murcha et al. (2007) demonstraram que uma proteína da membrana interna de mitocôndrias, codificada pelo gene no loco At5g24650, tem duplo direcionamento e também está na membrana externa de cloroplastos.

As semelhanças entre as translocases de mitocôndrias e cloroplastos, assim como relatos que mostram que uma proteína de cloroplasto pode ser encontrada em mitocôndria (HURT et al., 1986) ou vice-versa (HUANG et al., 1990) geram a questão de como é mantida a especificidade de direcionamento entre essas organelas. O fato de o direcionamento incorreto de proteínas cloroplastidiais para mitocôndrias ocorrer com certa frequência em sistemas *in vitro* (SILVA-FILHO et al., 1997; CLEARY et al., 2002) reforça a hipótese de que fatores citosólicos, especialmente chaperonas, podem atuar na prevenção do direcionamento incorreto (LITHGOW; HØJ; HOOGENRAAD, 1993). A forma citosólica do receptor Toc159 poderia também representar um fator citosólico que garante a especificidade do direcionamento, de forma que, ao se ligar com proteínas de cloroplastos, preveniria o reconhecimento pelos receptores mitocondriais. Um outro mecanismo poderia ser o direcionamento do mRNA para as proximidades das mitocôndrias. As proteínas recém-sintetizadas encontrariam facilmente os receptores da membrana mitocôndrial, evitando-se, dessa forma, o encontro das proteínas precursoras com cloroplastos (CHEW; WHELAN, 2004).

2.3 Predição da localização celular de proteínas

Proteínas direcionadas para uma mesma organela têm seqüências de direcionamento com características comuns. Em alguns casos, não há conservação da seqüência primária, mas de sua composição, de acordo com a natureza dos aminoácidos. Sendo assim, levando em conta as semelhanças entre seqüências de direcionamento, alguns programas computacionais sugerem a predição da localização celular de proteínas a partir de informações contidas em sua estrutura primária. Os principais programas disponíveis são Predotar (SMALL et al., 2004), TargetP (EMANUELSSON et al., 2000), iPSORT (BANNAI et al., 2002) e TargetLoc (HÖGLUND et al., 2006).

No entanto, esses programas deixam de reconhecer a correta localização de metade a 3/4 das proteínas organelares, além de apresentarem altas taxas de falso-

positivos (MILLAR; WHELAN; SMALL, 2006). Quando a proteína possui duplo direcionamento celular, a imprecisão é ainda maior a ponto de tornar as predições não confiáveis (vide item 4.1.2). Apesar das imprecisões, esses programas ainda são uma boa ferramenta para guiar estudos de localização celular de proteínas. No entanto, a confirmação por dados experimentais é mandatória.

As proteínas podem ter sua localização celular determinada por ensaios de importação *in vitro* para organelas, pela fusão de sua seqüência a uma proteína repórter como a GFP (*green fluorescent protein*, proteína fluorescente verde) ou por técnicas de imunolocalização, por microscopia ou *Western Blotting*. Além disso, o destino final de proteínas pode ser determinado, detectando-se sua atividade ou de proteína repórter nos compartimentos celulares. Análises proteômicas de organelas ou de seus subcompartimentos (tilacóide, por exemplo) isolados também fornecem informação importante sobre a localização de proteínas nas células. O banco de dados SUBA (*The SubCellular Proteomic Database*) (HEAZLEWOOD et al., 2007) reúne as informações geradas sobre a localização celular de proteínas em *A. thaliana*.

2.4 O duplo direcionamento de proteínas

Devido à sobreposição de funções entre organelas, principalmente entre cloroplastos e mitocôndrias, algumas proteínas são necessárias em mais de um compartimento celular. Essas participam de processos como a replicação e reparo de DNA, transcrição, tradução, metabolismo anti-oxidativo, entre outros. Nesses casos, a duplicação gênica é a estratégia mais comum: para diferentes isoformas da proteína há um gene nuclear equipado com uma seqüência de direcionamento específica. Exceções a essa estratégia são as proteínas que apresentam duplo direcionamento: há apenas um gene nuclear capaz de gerar produtos protéicos que são direcionados para locais diferentes na célula ou para subcompartimentos de uma mesma organela (SILVA-FILHO, 2003).

O duplo direcionamento parece ser um fenômeno universal entre eucariotos, já tendo sido relatado em fungos (SASS; KARNIELY; PINES, 2003), musgos (RICHTER et al., 2002), animais (COLOMBO et al., 2005) e plantas superiores (PEETERS; SMALL, 2001). A primeira proteína de plantas a ter seu duplo direcionamento demonstrado foi a

glutaciona redutase de ervilha, a qual é endereçada simultaneamente para mitocôndrias e cloroplastos (CREISSEN et al., 1995). Anteriormente, tinha sido descrita a capacidade da seqüência de direcionamento mitocondrial da proteína citocromo oxidase subunidade Va de levedura de direcionar uma proteína repórter para mitocôndrias e cloroplastos, em células de tabaco (HUANG et al., 1990). Atualmente, quase 60 proteínas já tiveram seu duplo direcionamento documentado em *A. thaliana* (vide Material e Métodos, item 4.1) ou em outras espécies (Tabela 1).

Tabela 1 – Proteínas com duplo direcionamento em espécies de plantas, exceto *A. thaliana*

Proteína	Espécie	Localização	Referência
Fator sigma (Sig2B)	<i>Zea mays</i>	M/C	(BEARDSLEE et al., 2002)
Fosforibosil aminoimidazol sintase	<i>Vigna unguiculata</i>	M/C	(GOGGIN et al., 2003)
Glutaciona preoxidase	<i>Populus trichocarpa</i>	M/C	(NAVROT et al., 2006)
Citocromo b5	<i>Brassica oleracea</i>	M/RE	(ZHAO et al., 2003)
Protoporfirinogenio oxidase	<i>Spinacea oleracea</i>	M/C	(WATANABE et al., 2001)
Lectina	<i>Nicotiana tabacum</i>	Cit/Nuc	(CHEN et al., 2002)
Proteína quinase ativada por mitógenos	<i>Nicotiana tabacum</i>	Cit/Nuc	(CORONADO et al., 2002)
Calmodulina (CaM53)	<i>Petunia hybrida</i>	MP/Nuc	(RODRÍGUEZ-CONCEPCIÓN et al., 1999)
Ricina	<i>Ricinus communis</i>	RE/Cit	(DI COLA et al., 2001)
Ascorbato peroxidase	<i>Gossypium hirsutum</i>	RE/P	(MULLEN; TRELEASE, 2000)

Nota: Abreviações, C, cloroplasto; Cit, citosol; M, mitocôndria; MP, membrana plasmática; Nuc, núcleo; P, peroxissomo; RE, retículo endoplasmático.

Os mecanismos que possibilitam o processo de duplo direcionamento de proteínas para mitocôndrias e cloroplastos podem ser de três tipos: regulação transcricional, em que ocorrem múltiplos sítios de iniciação da transcrição ou processamento alternativo do mRNA; controle traducional, em que um único transcrito contém múltiplos sítios de iniciação da tradução e estratégias pós-traducionais, que incluem proteínas com direcionamento ambíguo, modificações pós-traducionais e direcionamento competitivo (SILVA-FILHO, 2003; MACKENZIE, 2005).

A maioria das proteínas com duplo direcionamento é endereçada para mitocôndrias e cloroplastos e o mecanismo mais comum é a presença de uma seqüência de direcionamento ambígua, isto é, uma mesma seqüência tem a informação para o direcionamento a essas duas organelas. Esse é o caso das aminoacil-tRNA sintetases (DUCHÊNE et al., 2005), glutamina sintetase (TAIRA et al., 2004), ascorbato peroxidase (CHEW; WHELAN; MILLAR, 2003), entre outras proteínas.

A proteína monodeidroascorbato redutase de *A. thaliana* constitui um exemplo de duplo direcionamento regulado por mecanismo transcricional. As isoformas mitocondrial e cloroplastidial são codificadas por um único gene nuclear, porém existem dois transcritos, que diferem na extensão em 21 pb, gerados pela existência de dois sítios de início da transcrição. A proteína mais longa é mitocondrial (OBARA; SUMI; FUKUDA, 2002).

As proteínas THI1 (enzima da biossíntese de tiazol) e DNA girase são exemplos de proteína com duplo direcionamento determinado por mecanismo pós-transcricional (CHABREGAS et al., 2003; WALL; MITCHENAL; MAXWELL, 2004). Nesses casos, há dois códon *in frame* no mRNA nos quais a tradução pode ser iniciada. O início no primeiro ou no segundo AUG determina a localização das proteínas em cloroplastos ou em mitocôndrias. A proteína γ -type DNA polimerase de *A. thaliana* também tem duplo direcionamento pelo mecanismo de múltiplos sítios de início da tradução. O primeiro códon de início está localizado na região 5' não traduzida do mRNA (5' UTR), na posição -7, e sua seqüência é CUG, ao invés de AUG. Quando a tradução se inicia nesse códon, a proteína apresenta duplo direcionamento para mitocôndrias e cloroplastos, enquanto o início da tradução na posição +1, em um códon AUG, gera a proteína destinada a cloroplastos. Esse exemplo mostra o potencial da 5' UTR como contenedora da informação para a importação da proteína e a possibilidade de esse ser um mecanismo de controle da eficiência do direcionamento aos cloroplastos (CHRISTENSEN et al., 2005). Um mecanismo similar de início da tradução na 5' UTR promove o duplo direcionamento da RNA polimerase *phage-type* de *Physcomitrella patens*. Quando a tradução se inicia no primeiro AUG, na região 5' UTR, a proteína tem duplo direcionamento para mitocôndrias e cloroplastos, ao passo que o início no

segundo AUG direciona a proteína exclusivamente para mitocôndrias (RICHTER et al., 2002).

No mecanismo de regulação do duplo direcionamento pós-traducional, fatores metabólicos e do meio externo podem influenciar no direcionamento de proteínas. Na célula eucariota, sinais internos ou externos, como estresse, mudanças metabólicas, prenitação e luz podem determinar o destino final de proteínas. A localização do fotoreceptor fototropina 1 de *A. thaliana*, por exemplo, é regulada por luz. Normalmente essa proteína está localizada na membrana plasmática, porém move-se para o citosol em presença de luz azul (SAKAMOTO; BRIGGS, 2002).

O duplo direcionamento da calmodulina CaM53 de *Petunia hybrida* também ocorre por um mecanismo pós-traducional e depende de seu estado metabólico. A adição de grupos isoprenil na região carboxi-terminal da proteína determina a localização celular. A proteína em estado prenitado encontra-se na membrana plasmática, porém, em situações de bloqueio de síntese de isoprenóides, a proteína localiza-se no núcleo, em estado não-prenitado (RODRÍGUEZ-CONCEPCIÓN et al., 1999).

A seqüência de direcionamento ambígua tem características semelhantes daquelas exclusivamente mitocondriais ou cloroplastidiais e sua correta localização celular dificilmente é indicada por programas de predição *in silico* de localização de proteínas. Estudos com o objetivo de caracterizar a seqüência de direcionamento ambígua e determinar os fatores responsáveis pelo direcionamento para uma ou outra organela foram realizados, porém ainda são muitas as dúvidas nesse assunto.

A seqüência de direcionamento ambígua contém características comuns a pré-seqüências e peptídeos de trânsito: são ricas em serina e aminoácidos positivos e pobres em aminoácidos ácidos e glicina, quando comparadas a proteínas citosólicas. No entanto, seqüências de direcionamento ambíguas possuem mais leucina e fenilalanina e menos alanina (PEETERS; SMALL, 2001). Além disso, pré-seqüências apresentam abundância de arginina entre os primeiros 20 aminoácidos, quando comparadas à mesma região de peptídeos de trânsito. Seqüências de duplo direcionamento contêm quantidade de arginina intermediária (PUJOL; MARECHAL-DROUARD; DUCHÊNE, 2007). Aminoácidos carregados positivamente, como arginina,

são importantes para a interação da proteína com chaperonas citosólicas e com as translocases das membranas mitocondriais (PFANNER; GEISLER, 2001).

Rudhe et al. (2002a), com o objetivo de compreender os sinais de direcionamento da proteína glutathione reductase de ervilha, introduziram deleções na região amino-terminal de sua seqüência de direcionamento ambígua. Foi verificado que a remoção dos 30 primeiros aminoácidos reduziu a importação para mitocôndrias e cloroplastos em 85 e 65%, respectivamente, em comparação com a proteína selvagem. Os autores ainda identificaram uma região, a qual compreende, pelo menos, os 16 primeiros aminoácidos, com efeitos inibitórios no direcionamento para mitocôndrias. Os dados indicaram que a primeira metade da seqüência de direcionamento da glutathione reductase não contém informação essencial ao transporte da proteína, mas aumenta a eficiência de sua importação.

Experimentos de mutações sítio-dirigidas também foram realizados com o objetivo de caracterizar a natureza do sinal de direcionamento da glutathione reductase de ervilha. Chew et al. (2003) substituíram, ao longo da seqüência de direcionamento da proteína, aminoácidos positivos, como alanina, por outros negativos e também grupos de três aminoácidos hidrofóbicos por outros não hidrofóbicos. A substituição de aminoácidos positivos teve maior efeito inibitório no transporte para mitocôndrias, enquanto o transporte para cloroplastos foi mais afetado por mutações em aminoácidos hidrofóbicos na região amino-terminal da seqüência de direcionamento. Observou-se também que os sinais para direcionamento a cloroplastos e mitocôndrias encontram-se sobrepostos. Os autores concluíram que mutações em resíduos positivos e hidrofóbicos na região amino-terminal ou mediana da seqüência de direcionamento têm efeitos inibitórios significativos no transporte da proteína para mitocôndrias e cloroplastos e que mutações localizadas na região carboxi-terminal inibem o direcionamento para mitocôndrias.

Rudhe et al. (2004) introduziram mutações de ponto e deleções na seqüência de direcionamento da glutathione reductase, com o objetivo de compreender o processamento de peptídeos com duplo direcionamento. As proteínas endereçadas para mitocôndrias e cloroplastos, ao adentrarem na organela, sofrem processamento por peptidases localizadas no estroma ou na matriz mitocondrial. Essas enzimas são

capazes de reconhecer sítios de clivagem de diversas proteínas, mesmo que não tenham similaridade entre as seqüências primárias. Apesar de existirem semelhanças entre as peptidases mitocondriais e cloroplastidiais, assim como entre as seqüências de direcionamento, o processamento de proteínas precursoras é específico em cloroplastos e mitocôndrias. Peptidases mitocôndriais são incapazes de processar proteínas direcionadas aos cloroplastos e vice-versa. No caso da glutathione redutase, o sítio de reconhecimento para o processamento difere entre essas duas organelas. Além disso, o processamento por peptidases mitocôndriais é mais sensível a mutações de ponto nas proximidades do sítio de clivagem, assim como a substituições de aminoácidos positivos na região carboxi-terminal da seqüência de direcionamento. Essas alterações não influenciam a atividade de peptidases cloroplastidiais. Os resultados indicam que a especificidade do processamento de proteínas precursoras parece ser mantida entre aquelas com duplo direcionamento.

Também com o objetivo de investigar as características de seqüências de duplo direcionamento ambíguas, Pujol; Maréchal-Drouard; Duchêne (2007) introduziram uma série de mutações sítio-dirigidas nas seqüências de direcionamento de metionil, glutamil e fenilalanil-tRNA sintetases de *A. thaliana* e analisaram seus efeitos pela metodologia de transporte *in vitro* de proteínas para organelas. Diferentes sítios de processamento das proteínas foram reconhecidos por peptidases mitocondriais e cloroplastidiais. Foi confirmada a importância de argininas para o endereçamento das proteínas às mitocôndrias e investigado o papel da alanina na posição 2. A substituição da alanina na posição 2 não gerou efeitos no transporte aos cloroplastos e, para as mitocôndrias, o efeito foi variável, pois o direcionamento não foi afetado ou foi drasticamente reduzido. Além disso, a substituição da treonina na posição 2, na metionil-tRNA sintetase de *A. thaliana*, por uma alanina, comum em outras seqüências de metionil-tRNA sintetase de plantas, não interferiu no direcionamento para mitocôndrias, porém reduziu a quantidade de proteínas nos cloroplastos. Os efeitos das mutações na posição 2 parecem ser dependentes do contexto protéico, o que sugere uma cooperação do segundo aminoácido com outros determinantes na região amino-terminal.

Diversos trabalhos focados no estudo de aspectos genéticos e celulares do duplo direcionamento de proteínas forneceram informações importantes sobre um tópico que

ainda não está totalmente compreendido. A crescente quantidade de dados gerados sobre o assunto nos últimos dez anos permitiu o entendimento de diversos aspectos da determinação do duplo direcionamento, mostrando que esse é um mecanismo complexo e não tão raro como se imaginava.

No entanto, ainda permanece uma importante questão: por que as células desenvolveram mecanismos peculiares capazes de promover o duplo direcionamento de proteínas para diferentes organelas? Poder-se-ia pensar que se trata de um mecanismo de “economia genômica”, principalmente quando se observa que a maioria das proteínas de plantas com duplo direcionamento são de *A. thaliana*, a qual apresenta um genoma pequeno. No entanto, o duplo direcionamento já foi relatado em espécies com genomas maiores e, apesar do tamanho do genoma de plantas ser variável, o número de genes funcionais é bem semelhante (SILVA-FILHO, 2003). Além disso, algumas proteínas que apresentam duplo direcionamento em *A. thaliana*, também apresentam esse mecanismo em outras espécies (KOBAYASHI et al., 2002; ONO et al., 2007), o que mostra que o estabelecimento do duplo direcionamento deve estar relacionado principalmente a aspectos funcionais da proteína e não ao tamanho do genoma. O maior número de proteínas com duplo direcionamento em *A. thaliana* poderia ser explicado pelo fato de se tratar de planta modelo para a genética e a biologia celular de plantas e, dessa forma, é a mais utilizada em estudos desse tipo.

A rápida sofisticação que as técnicas de bioinformática, microscopia e biologia molecular e celular vêm adquirindo certamente contribuirá, em breve, para as respostas a diversas das questões relacionadas ao duplo direcionamento de proteínas.

3 MATERIAL E MÉTODOS

3.1 Análises de seqüências de proteínas de *Oryza sativa* e *Arabidopsis thaliana*

Foi realizada uma busca por seqüências de proteínas de *A. thaliana* cujo duplo direcionamento foi descrito na literatura. Essas seqüências foram utilizadas como *input* em uma pesquisa BLAST (*Basic Local Alignment and Search Tool*) no banco de dados TAIR (*The Arabidopsis Information Resource database*) (2006). As isoformas de *A. thaliana* recuperadas foram usadas para uma busca no banco de dados de arroz do TIGR (*The Institute for Genomic Research*) (2006). Os *hits* com valor *e* (*e-value*) inferiores a 1.0 e-65 e seqüências cujos domínios funcionais das proteínas estavam incompletos foram descartados. Exceções foram feitas para aminoacil-tRNA sintetases (DUCHÊNE et al., 2005), ascorbato peroxidase (CHEW; WHELAN; MILLAR, 2003; TEIXEIRA et al., 2006), proteínas AP2/EREBP (SCHWACKE et al., 2007) e a família das deformilases (GIGLIONE et al., 2000; DINKINS et al., 2003), em que foram consideradas as isoformas previamente descritas. Uma busca do tipo TBLASTN no banco de dados de EST (*expressed sequence tag*) do NCBI (Centro Nacional para a Informação Biotecnológica) (ALTSCHUL et al., 1997) foi feita com o objetivo de verificar se as isoformas recuperadas eram expressas. Os resultados foram confirmados por pesquisas realizadas nos bancos de dados de microarranjos do TAIR e do TIGR.

As predições de direcionamento das proteínas foram testadas com os programas Predotar (SMALL et al., 2004), TargetP (EMANUELSSON et al., 2000), iPSORT (BANNAI et al., 2002) e TargetLoc (HÖGLUND et al., 2006). Para análises filogenéticas, as seqüências de aminoácidos das famílias estudadas foram alinhadas no programa CLUSTAL W (THOMPSON; HIGGINS; GIBSON, 1994). Árvores filogenéticas foram construídas, utilizando-se o programa SplitsTree (HUSON; BRYANT, 2006), pelo método de *neighbor joining* (SAITOU; NEI, 1987). O número de *bootstrap* usado foi 100.

3.2 Clonagem de seqüências de cDNA no vetor pPMA4-GFP para expressão transiente

O segmento de cDNA correspondente às regiões amino-terminais das proteínas metionina aminopeptidase, MAP (Os02g52420) e monodeidroascorbato redutase, MDAR (Os08g05570) de *O. sativa* foi amplificado pela técnica de PCR (reação em cadeia da polimerase) a partir dos clones OSJNEc05E04 e OSJNEb11D10, respectivamente, obtidos do *Arizona Genomics Institute* (2005). Foram utilizados o iniciador (*primer*) universal *forward* M13F e os *reverse* OsMET-R e OsMDAR-R, ambos com sítios de restrição da enzima *Kpn* I na extremidade 5' (Tabela 2). Utilizou-se a enzima *Taq* DNA polimerase recombinante (Invitrogen), seguindo um ciclo de 94° C por 1 minuto; 58° C por 30 segundos e 72° C por 40 segundos, repetido por 30 vezes, no termociclador PTC - 100 (MJ Research). O produto da amplificação foi analisado após eletroforese (5V/cm, 20 minutos) em gel de agarose 0,8% com tampão TBE 0,5 x. Os fragmentos amplificados, incluindo os primeiros 336 pb, para MAP, e 510 pb, para MDAR, foram digeridos com as enzimas de restrição *Bam*H I e *Kpn* I (Invitrogen), de acordo com instruções do fabricante, e purificados com o kit QIAquick PCR Purification (Qiagen), seguindo o manual do produto). Utilizando a enzima T4 DNA ligase (Invitrogen) de acordo com o manual do produto, os fragmentos foram ligados *in frame* com a seqüência da GFP (proteína fluorescente verde) no vetor pTZ19U-EN50PMA-β-GFP (DUBY; DEGAND; BOUTRY, 2001), aqui denominado pPMA4-GFP (Figura 1), previamente digerido com as enzimas *Bgl* II e *Kpn* I (Invitrogen). A construção foi utilizada para a transformação de células DH5α eletrocompetentes, preparadas conforme descrito por Sambrook; Russel (2001), e plaqueadas em meio LB (Luria-Bertani) com ampicilina 100 mg/L. O resultado da clonagem foi conferido por seqüenciamento (realizado no Centro de Estudos do Genoma Humano, Instituto de Biociências, USP, São Paulo).

O cDNA de *thi1* de *Oryza sativa* foi obtido pela técnica de RT-PCR. Para tanto, sementes de arroz cultivar IAC 202, gentilmente cedidas pelo Prof. Dr. J. Baldin (ESALQ/USP), foram semeadas em vaso, contendo substrato vegetal (Plantmax Horticultural) e vermiculita na proporção 2:1, e cultivadas em casa de vegetação. O RNA total de plantas com 15 dias de idade foi extraído com o reagente TRIZOL

(Invitrogen), seguindo instruções do fabricante, a partir de 300 mg de folha macerada em nitrogênio líquido. O cDNA foi obtido a partir do RNA mensageiro total, utilizando-se a enzima transcriptase reversa *SuperScript II RT* (Invitrogen), e o oligo dT, de acordo com as recomendações do fabricante. A amplificação específica do cDNA de *thi1* foi realizada com os iniciadores OsTHI1-F e OsTHI1-R, contendo, na região 5', os sítios das enzimas *Bgl* II e *Kpn* I, respectivamente (Tabela 2). Foi seguido o ciclo de 94° C, por 1 minuto, 59° C por 45 segundos e 72° C por 50 segundos, repetido por 35 vezes no termociclador PTC – 100 (MJ Research). O produto da amplificação foi analisado por eletroforese, conforme descrito acima.

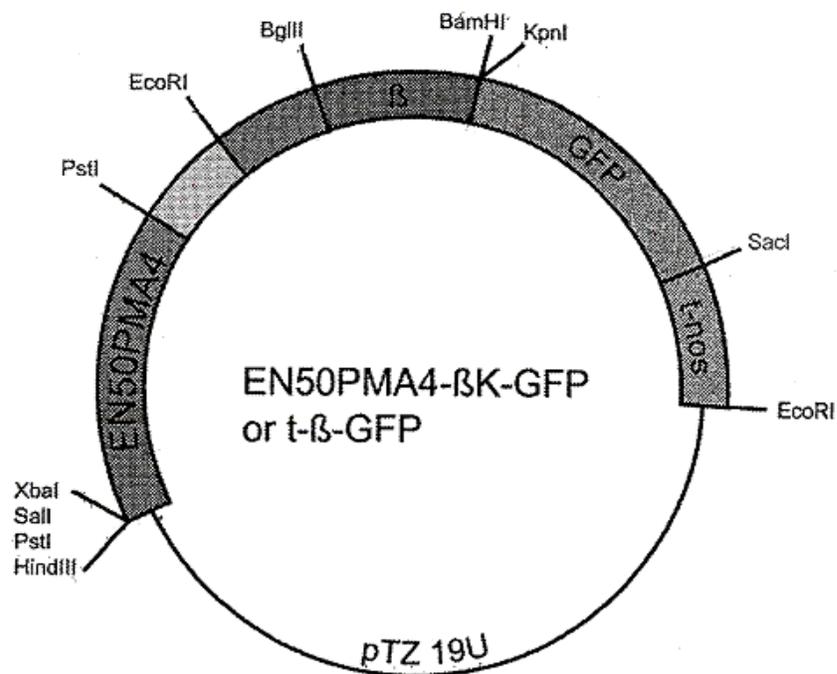


Figura 1 - Mapa do plasmídeo pPMA4-GFP (DUBY; DEGAND; BOUTRY, 2001), usado para clonagem de seqüências de cDNA para expressão transiente

Tabela 2 – Sequência de nucleotídeos dos iniciadores utilizados para a realização de clonagens, mutações sítio dirigidas e deleções

Iniciadores	Seqüências (5'-3')
<i>Clonagens de seqüências no vetor pPMA4-GFP</i>	
OsMET-R	CATGGTACCTGGTCGCTCATATGGCGGATAC
OsMDAR-R	TAGGTACCTCAAGGTCTTTGACGAGCATTG
OsTHI1-F	CAAGATCTACTCTCCTCATCCCAGAGCA
OsTHI1-R	CATGGTACCCGACTGCGTGTGCGTGGTTCAT
THI15'-F	CAAGATCTAATGGCTGCCATA
THI13'-R	TAGGTACCGTTCTTACTGATCT
THI1full-R	TAGGTACCAGCATCTACGGT
<i>Clonagens nos vetores pUCAP-GFP ou pCRII-TOPO</i>	
RBP1b-F	CAGGTACCGAAATGGCTTTCTGCAACA
RBP1b-R	ATGGATCCAAGGAGCTGTCATCAGTTC
RBP1a-F	GAGGTACCGAAATGGCTTTCTGTAACA
RBP1a-R	ATGGATCCAAGGATGCGTCATCGGTTC
RPS19-F	CAGGTACCGATATGGCTTTCTGCACTAA
RPS19-R	ATGGATCCAAGGAGTGTGTCAGTT
RPS19full-R	ATGGATCCGGCAAACCTTGGCCAATC
RBP1bfull-R	ATGGATCCTCACCACCACCAAATGATG
THI1WT-F	TAGGTACCATGGCTGCCATAGCTTC
THI1SD-R	CTGGATCCGCAACTTGAACATTAGG
THI1mito-F	CTGGTACCATGACGAGAAGGTACATGAC
THI1full3'-R	TTAAGCATCTACGGTTTCAGC
<i>Mutações sítio-dirigidas e deleções</i>	
A2N-F	GCATAGCCATGAATTTCTGCAAC
A2N-R	GTTGCAGAAATTCATGGCTATCG
K6Q-F	GCTTTCTGCAACCACTTAGCGGTATC
K6Q-R	GATACCGCTAAGTTGGTTGCAGAAAGC
R12N-F	CTTAGCGGTATCTTGAATCAAGGGGTTTCTCAG
R12N-R	CTGAGAAACCCCTTGATTCAAGATACCGCTAAG
V15NS16G-F	GAGACAAGGGAAATGGACAGAGCTCAAATGG
V15NS16G-R	CCATTTGAGCTCTGTCCATTCCTTGTCTC
S18NS19G-F	GGTTTCTCAGAAATGGTAATGGTCCAGTTAC
S18NS19G-R	GTAAGTGGACCATACCATTCTGAGAAACC
Δ(18-19)-F	GACAAGGGGTTTCTCAGAATGGTCCAGTTAC
Δ(18-19)-R	GTAAGTGGACCATCTGAGAAACCCCTTGTC
R31Q-F	GCTTGGCTCTCTTCAATACATGTCC
R31Q-R	GGACATGTATTGAAGAGAGCCAAGC
Δ(2-17)-F	TAGGTACCATGAGCTCAAATGGTCC
Δ18-30-F	CGTTACATGTCCTCTAAGC
Δ18-30-R	GCTTAGAGGACATGTAACGCTGAGAAACCCCTTGTCT
PUCAP-F	GCATTCTACTTCTATTGCAGC
18S1A-F	ATAGTAGTTTTAACTTTCAGCCATCTCCGCAGC
18S1B-R	GAAAGTTTTAACTACTATCGAAAAGTCTCTGAGGT
18S2A-F	CATCTCCGCAATTGCTATTAGTATCGGTCTCAAAC
18S2B-R	GTTTGAGACCGATACTAATAGCAATTGCGGAGATG

Nota: em azul, estão indicados sítios de enzimas de restrição e, em vermelho, regiões em que a seqüência original foi mutada.

O fragmento amplificado, com aproximadamente 680 pb, foi clonado, utilizando-se o *kit TOPO-TA Cloning*, no vetor pCR 2.1 (Invitrogen), de acordo com o manual do produto. Em seguida, os vetores pPMA4-GFP e pCR2.1+thi1 foram digeridos com as enzimas *Bgl* II e *Kpn* I (Invitrogen). Os fragmentos de interesse foram purificados do gel de agarose, usando o *kit Qiaex II Gel Extraction* (Qiagen), conforme recomendações do fabricante. Foi então realizada a ligação do fragmento do cDNA de *thi1* no vetor pPMA4-GFP, *in frame* com a GFP, seguindo procedimentos descritos acima.

Para a clonagem da seqüência completa do cDNA, exceto o códon de parada, de *thi1* de *A. thaliana* no vetor pPMA4-GFP, foram utilizados os iniciadores THI15'-F e THI1full-R, os quais contêm os sítios de restrição de *Bgl* II e *Kpn* I na extremidade 5', respectivamente. Para a clonagem somente da seqüência de direcionamento, que inclui os primeiros 315 pb da seqüência do cDNA de *thi1*, foram utilizados os iniciadores THI15'-F e THI13'-R, este último com o sítio *Kpn* I na extremidade 5'. As amplificações foram realizadas a partir do cDNA de *thi1* clonado no plasmídeo Bluescript SK(-) (ALMEIDA, 2004). Para a clonagem dos fragmentos de DNA amplificados, foram utilizadas técnicas descritas acima.

3.3 Clonagem de seqüências de cDNA no vetor pUCAP-GFP para expressão transiente

Os fragmentos de DNA a serem clonados foram amplificados pela técnica de PCR com oligonucleotídeos específicos (Tabela 2), contendo o sítio da enzima de restrição *Kpn* I ou *Bam*H I. Utilizou-se a enzima *Taq* DNA polimerase recombinante (Invitrogen), seguindo um ciclo de 94° C por 1 minuto, 58° C por 30 segundos e 72° C por 40 segundos, repetido por 30 vezes no termociclador Bio-Rad *Gene Cycler*. Utilizaram-se os seguintes iniciadores:

- RBP1b-F e RBP1b-R, para a clonagem dos primeiros 152 pb do cDNA de RBP1b, os quais compõem sua seqüência de direcionamento;
- RBP1a-F e RBP1a-R, para a clonagem dos primeiros 152 pb do cDNA de RBP1a, os quais compõem sua seqüência de direcionamento;
- RPS19-F e RPS19-R, para a clonagem dos primeiros 140 pb do cDNA de RPS19, os quais compõem sua seqüência de direcionamento;

- RPS19-F e RPS19full-R, para a clonagem dos primeiros 341 pb do cDNA de RPS19;
- THI1WT-F e THI1SD-R, para a clonagem dos primeiros 332 pb do cDNA de THI1 de *A. thalina*, referentes a sua seqüência de direcionamento;
- THI1mito-F e THI1SD-R, para a clonagem de parte da seqüência de direcionamento de THI1. São amplificados 125 pb do cDNA de *thi1*, a partir do segundo ATG.

Os fragmentos amplificados foram analisados por eletroforese a 5,5 V/cm, em gel de agarose 1%, e purificados com o *kit NucleoSpin Extract II* (Macherey-Nagel, de acordo com instruções do fabricante). Os fragmentos e o vetor pUCAP-GFP (Figura 2) (gentilmente cedido por J.M. Gualberto, Université Louis Pasteur, Estrasburgo, França) foram digeridos com as enzimas *Kpn I* e *BamH I* (Fermentas), de acordo com o manual do fabricante. O DNA foi quantificado e a ligação dos fragmentos ao vetor foi feita com o *kit Rapid DNA Ligation* (Fermentas), de acordo com instruções do fabricante, obedecendo a proporção de uma quantidade do vetor para três do fragmento. Terminado o período de incubação, o DNA foi purificado com fenol/clorofórmio e, em seguida, precipitado com 5 mg de glicogênio (Fermentas), 1/10 V de acetato de potássio 3 M e 2,5 V de etanol 100%. Após a secagem, o DNA foi diluído em 5 µL de água destilada estéril e 1 µL foi utilizado para a eletroporação de células competentes de *Escherichia coli* XI1-blue, preparadas segundo Sambrook; Russel (2001), utilizando-se o eletroporador *Gene Pulser II* (BioRad). Após a incubação das células em 1 mL de meio LB (Luria-Bertani), por uma hora a 37° C, procedeu-se à centrifugação, a 6 krpm, por 2 minutos; parte do meio foi descartado e a suspensão concentrada de bactérias foi plaqueada em LB sólido contendo ampicilina 100 mg/L. As placas foram mantidas por aproximadamente 14h à 37° C. As colônias obtidas foram analisadas pela técnica de PCR para detectar a inserção do fragmento no vetor. Para tanto, foram utilizados um iniciador *forward* complementar a uma seqüência do vetor e um oligonucleotídeo *reverse* complementar a uma seqüência do fragmento de interesse. A PCR foi realizada conforme descrito acima. O produto da amplificação foi analisado por eletroforese a 5,5 V/cm em gel de agarose 1%.

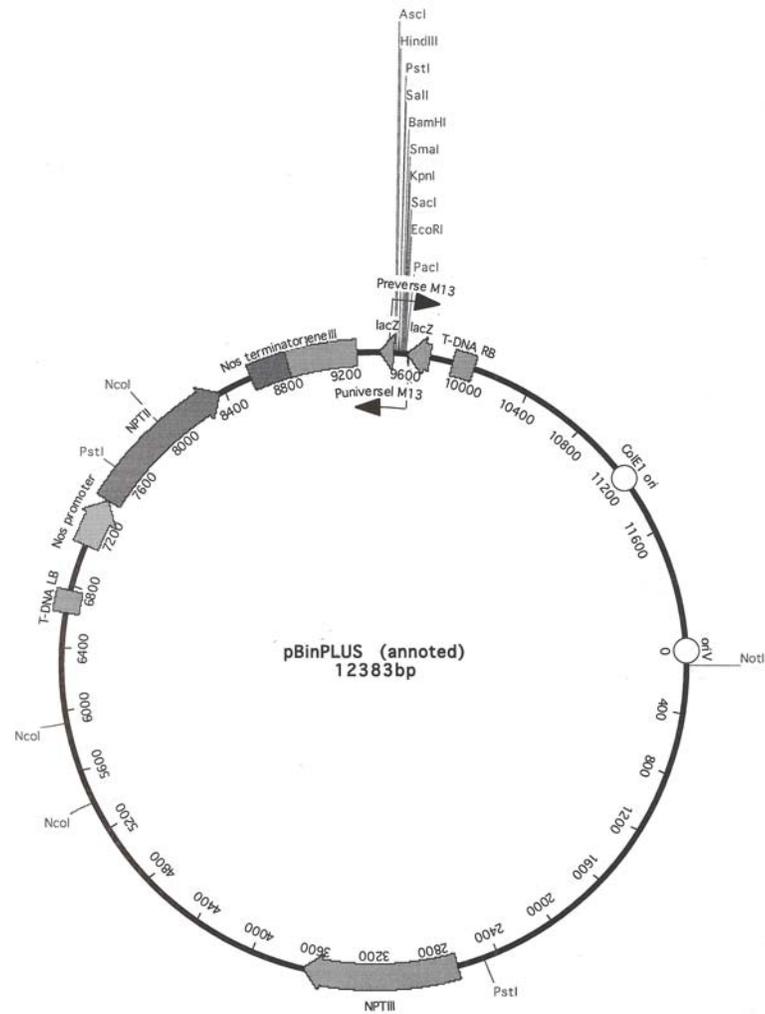


Figura 3 - Mapa do plasmídeo pBinPLUS, usado para clonagem de seqüências de DNA para transformação de plantas via *Agrobacterium tumefaciens*

3.5 Clonagem de seqüências para transcrição e tradução *in vitro*

Algumas construções foram testadas pelo sistema de transporte *in vitro* de proteínas para cloroplastos e mitocôndrias. Para tanto, os fragmentos de DNA foram amplificados por PCR, seguindo um ciclo de 94 ° C por 1 minuto, 58 ° C por 30 segundos e 72° C por 40 segundos, repetido por 30 vezes, no termociclador Bio-Rad *Gene Cycler*. Foram utilizados os iniciadores:

- RBP1b-F e RBP1bfull-R, para a amplificação da seqüência completa de RBP1b, com ou sem mutações e;
- THI1WT-F e THI1full3'-R, para a amplificação da seqüência completa de THI1, com ou sem mutações de ponto ou deleções.

Os produtos das amplificações foram visualizados após eletroforese, a 5,5 v/cm, em gel de agarose 1%. Os fragmentos amplificados foram ligados a vetores de clonagem de produto de PCR, pCRII-TOPO (Invitrogen) ou pSC-A Strataclone (Stratagene), de acordo com instruções dos fabricantes. O sentido da clonagem e a seqüência do fragmento clonado foram conferidos por seqüenciamento, realizado no centro de seqüenciamento do Institut de Biologie Moléculaire des Plantes, Estrasburgo, França, no aparelho ABI Prism 3100 *Genetic Analyzer* (Applied Biosystems).

3.6 Isolamento de plasmídeos

O isolamento de plasmídeos em pequena escala foi realizado com o kit NucleoSpin Plasmid QuickPure Mini (Macherey-Nagel), de acordo com instruções do fabricante ou pela metodologia de lise alcalina (SAMBROOK; RUSSEL, 2001). No final, foi adicionada água de forma a se ter uma concentração de plasmídeo de 200 ng/μL. O isolamento de plasmídeos em média escala foi realizado com o *kit NucleoBond PC 100* (Macherey-Nagel) ou com o *Plasmid Midi Kit* (Qiagen), de acordo com instruções dos fabricantes, adicionando-se água para uma concentração final do plasmídeo de 1 μg/μL.

3.7 Mutações sítio-dirigidas em seqüências de nucleotídeos

As mutações sítio-dirigidas, de ponto ou deleções de até 2 pb do interior das seqüências de nucleotídeos, foram obtidas, utilizando-se o *kit Quick Change II Site-Directed Mutagenesis* (Stratagene), de acordo com instruções do fabricante. Algumas mutações de ponto foram obtidas pela técnica de PCR, utilizando-se iniciadores contendo as mutações de interesse. Deleções maiores que 2 pb foram introduzidas pela técnica de PCR.

Para a introdução das mutações de ponto e deleções na seqüência completa do cDNA de RBP1b, foi levado em consideração o uso de códon de *A. thaliana* e utilizaram-se os seguintes iniciadores (seqüências na Tabela 2):

- A2N-F e A2N-R, para a substituição do resíduo 2, alanina, por asparagina;
- K6Q-F e K6Q-R, para a substituição do resíduo 6, lisina, por glutamina;
- R12N-F e R12N-R, para a substituição do resíduo 6, arginina, por asparagina;
- V15NS16G-F e V15NS16G-R, para a substituição dos resíduos 15 e 16, valina e serina, por asparagina e glutamina, respectivamente;
- S18NS19G-F e S18NS19G-R, para a substituição dos resíduos 18 e 19, serinas, por asparagina e glutamina, respectivamente;
- $\Delta(18-19)$ -F e $\Delta(18-19)$ -R, para a deleção dos resíduos 18 e 19, duas serinas;
- R31Q-F e R31Q-R, para a substituição do resíduo 31, arginina, por glutamina;
- $\Delta(2-17)$ -F e RBP1bfull-R, para a deleção dos primeiros 17 resíduos, exceto a primeira metionina;
- $\Delta(18-30)$ -F, $\Delta(18-30)$ -R, RBP1bfull-R e pUCAP-F, para a deleção dos resíduos 18 a 30;
- 1b s/TS-F e RBP1bfull-R, para a deleção da seqüência de direcionamento.

Para a introdução das mutações de ponto na seqüência completa do cDNA de *thi1*, foi levado em consideração o uso de códon de *A. thaliana* e utilizaram-se os seguintes iniciadores (seqüências na Tabela 2):

- 18S1A-F e 18S1B-R, para as mutações da região 18S1 na seqüência do cDNA *thi1* complementar ao rRNA 18S (item 4.3)
- 18S2A-F e 18S2B-R, para as mutações da região 18S2 na seqüência do cDNA de *thi1* complementar ao rRNA 18S (item 4.3)

3.8 Transformação de *Agrobacterium tumefaciens*

Células competentes de *A. tumefaciens* GV3101 (preparadas de acordo com Sambrook; Russel, 2001) foram retiradas do freezer -80° C e, antes que descongelassem, a elas foram adicionados 200 ng do plasmídeo binário. Procedeu-se a incubação a 37° C durante 5 minutos e, em seguida, em gelo por pelo menos um minuto. Foram adicionados 500 μ L de meio LB e os tubos foram levados a um agitador a 28° C por duas horas. Terminado esse período, as células foram plaqueadas em meio LB contendo canamicina (30 mg/L), rifampicina (100mg/L) e gentamicina (50mg/L) e as placas foram incubadas a 28° C durante 48 horas.

3.9 Transformação transiente via biobalística

3.9.1 Preparo de partículas

A 60 mg de partículas de tungstênio foi adicionado 1 mL de etanol 70% e os tubos foram mantidos em um agitador por 15 minutos. Em seguida, o material foi centrifugado a 6 krpm, durante 5 minutos. O sobrenadante foi descartado e as partículas foram lavadas por duas vezes em água destilada. Ao final, as partículas foram ressuspensas em 1 mL de glicerol 50%.

3.9.2 Precipitação de DNA sobre as partículas

A 30 μ L de partículas de tungstênio adicionaram-se 5 μ g do DNA, 30 μ L CaCl_2 2,5 M e 8 μ L de espermidina 0,1 M. A mistura foi mantida sob agitação contínua durante 20 minutos. Em seguida, adicionaram-se 150 μ L de etanol 100% e os tubos foram submetidos a centrifugação a 3 krpm, durante 3 segundos. O sobrenadante foi descartado e adicionaram-se mais 150 μ L de etanol 100%. As partículas foram ressuspensas e os tubos submetidos a centrifugação. Repetiu-se o procedimento mais uma vez e, no final, as partículas foram ressuspensas em 20 μ L de etanol 100%.

3.9.3 Bombardeamento do material vegetal

Para o bombardeamento de *Nicotiana benthamiana*, uma folha de uma planta, com 45 dias de idade, cultivada em casa de vegetação, foi destacada e posicionada, com a face abaxial voltada para cima, em uma placa de Petri contendo 15 mL de meio MS (*Murashige and Skoog Plant Salt Mix*, ICN Biomedicale). Para o bombardeamento de cebola, as camadas dos catáfilos, cortadas em quadrados de 2 cm², foram posicionadas no centro de uma placa de Petri, com a face abaxial voltada para cima.

Utilizou-se o equipamento Biolistic PDS-100/He *Particle Delivery System* (BioRad). Os seguintes parâmetros foram adotados: vácuo de 26 mm Hg, distância de 6 cm entre as partículas e o alvo e pressão de He de 1.100 psi, para *N. benthamiana* e 900 psi, para cebola.

3.10 Transformação de *Nicotina benthamiana* por infiltração

Células de *A. tumefaciens* contendo o plasmídeo de interesse foram multiplicadas em meio LB adicionado de canamicina (50 mg/L), rifampicina (100 mg/L) e gentamicina (50 mg/L), a 28° C, sob agitação constante, durante 16 h. Em seguida, a cultura bacteriana foi centrifugada a 6 krpm, durante 10 minutos e o sobrenadante foi descartado. As células foram ressuspensas em MgCl₂ 10 mM, em volume suficiente para permitir uma densidade óptica final de 0,6, medida a 600 nm. À suspensão bacteriana, foi adicionada aceto-seringona na concentração final de 150 µM, mantendo-se por três horas à temperatura ambiente. Em seguida, foi realizada a infiltração em folhas de plantas de *N. benthamiana* com aproximadamente 40 dias de idade. Para tanto, foram feitas pequenas ranhuras na face abaxial das folhas e, com a ajuda de uma seringa de plástico de 2 mL sem agulha, a suspensão bacteriana foi infiltrada nas folhas. As plantas foram mantidas a 25° C, sob regime de 16 h de luz e 8 h de escuro e a fluorescência da GFP foi visualizada após 48 h.

3.11 Preparo de protoplastos

Após a realização da infiltração de folhas de *N. benthamiana*, as plantas foram mantidas a 25° C durante um dia. As folhas infiltradas foram destacadas e imersas, em placa de Petri, em uma solução enzimática 1 x (para 100 mL de solução estoque 5 x: 0,2 g de macerozima, 0,5 g de driselase, 1 g de celulase Onuzuka) diluída em meio para protoplasto (para 1 L: 4,3 g de meio MS, *Murashige and Skoog Plant Salt Mix*, ICN Biomedicale; 0,5 g MES; 20 g de sacarose; 80 g de manitol; KOH até pH 5,6). As folhas foram mantidas embebidas à 30° C, durante 6 h, período suficiente para sua digestão total. Não houve necessidade de filtrar o material. Em pequenas alíquotas da solução, contendo as folhas digeridas, os protoplastos puderem ser visualizados diretamente ao microscópio confocal.

3.12 Análises ao microscópio confocal

Os materiais transformados via biobalística e infiltração foram visualizados utilizando-se microscópio confocal Zeiss modelo SM 410 (*Laser Scan Microscopy*) ou microscópio de fluorescência Zeiss Axiophot. A região da folha de *N. benthamiana* atingida pelas partículas de tungstênio foi cortada e posicionada sobre uma lâmina de vidro embebida com água, coberta por uma lamínula, selando-se com fita adesiva. Para a visualização dos protoplastos, cerca de 300 µL da suspensão foram aplicados sobre uma lâmina e cobertos com lamínula. No caso da cebola, uma fina camada da epiderme dos catáfilos, a mais superior e atingida pelas partículas, foi removida com o auxílio de uma pinça de ponta fina, posicionada sobre uma lâmina de vidro e coberta com uma lamínula. A excitação da GFP foi feita a 488 nm e a fluorescência emitida foi visualizada através de um filtro “band-pass” 530±15 nm. A clorofila foi também excitada a 488 nm e sua autofluorescência foi visualizada através de um filtro “long-pass” 530 nm. Já o marcador *mitotracker* e a RFP (proteína vermelha fluorescente) foram excitadas com laser a 565 nm e visualizadas com um filtro “band-pass” 630±15 nm.

3.13 Análise quantitativa das imagens obtidas no microscópio confocal

As imagens referentes aos bombardeamentos de células do mesófilo de *N. benthamiana* foram obtidas ao microscópio confocal de modo a não terem um sinal saturado da fluorescência da GFP. Cada imagem incluiu apenas uma célula. Foram analisados planos isolados e não imagens sobrepostas (*stack*). As imagens foram analisadas com o programa ImageJ (COLLINS, 2007). O programa permite a leitura do sinal da GFP em números, medido em pixels, de acordo com sua intensidade. O canal verde (referente à fluorescência da GFP) e o vermelho (referente à autofluorescência da clorofila) podem ser visualizados separadamente (Figura 4, A e B).

No canal vermelho, o sinal da autofluorescência foi marcado e selecionado (linhas amarelas) (Figura 4C). A seleção do sinal dos cloroplastos foi sobreposta à imagem no canal verde (Figura 4D). Dessa forma, foi marcado, no canal verde, a fluorescência da GFP localizada nos cloroplastos. A fluorescência que não se encontra nessa seleção corresponde à encontrada em mitocôndrias. Foi então medida, em números, a intensidade da fluorescência total da GFP e da fluorescência presente somente na seleção (referente aos cloroplastos). A medida da intensidade do sinal da GFP em mitocôndrias foi obtida subtraindo-se, do valor total, o valor referente aos cloroplastos.

Um exemplo de medidas obtidas, utilizando o programa ImageJ, é mostrado abaixo (Tabela 3). Os valores em preto são os calculados pelo programa. São calculadas a área total ocupada pela fluorescência verde (A_t) e a área relativa à fluorescência dos cloroplastos (A_c , a área da seleção representada pelos círculos em amarelo). O programa também fornece o valor médio da fluorescência do sinal e a densidade, que é o valor da área multiplicado pelo valor da média ($A \times M$). Os valores em azul são os calculados a partir dos números fornecidos pelo programa. A área e a densidade para mitocôndrias são calculadas pela subtração: total – cloroplasto ($A_m = A_t - A_c$; $D_m = D_t - D_c$). A média para mitocôndrias é obtida dividindo-se o valor da densidade pela área ($M_m = D_m/A_m$). Os valores relativos de densidade foram utilizados nas análises comparativas entre as diversas células.

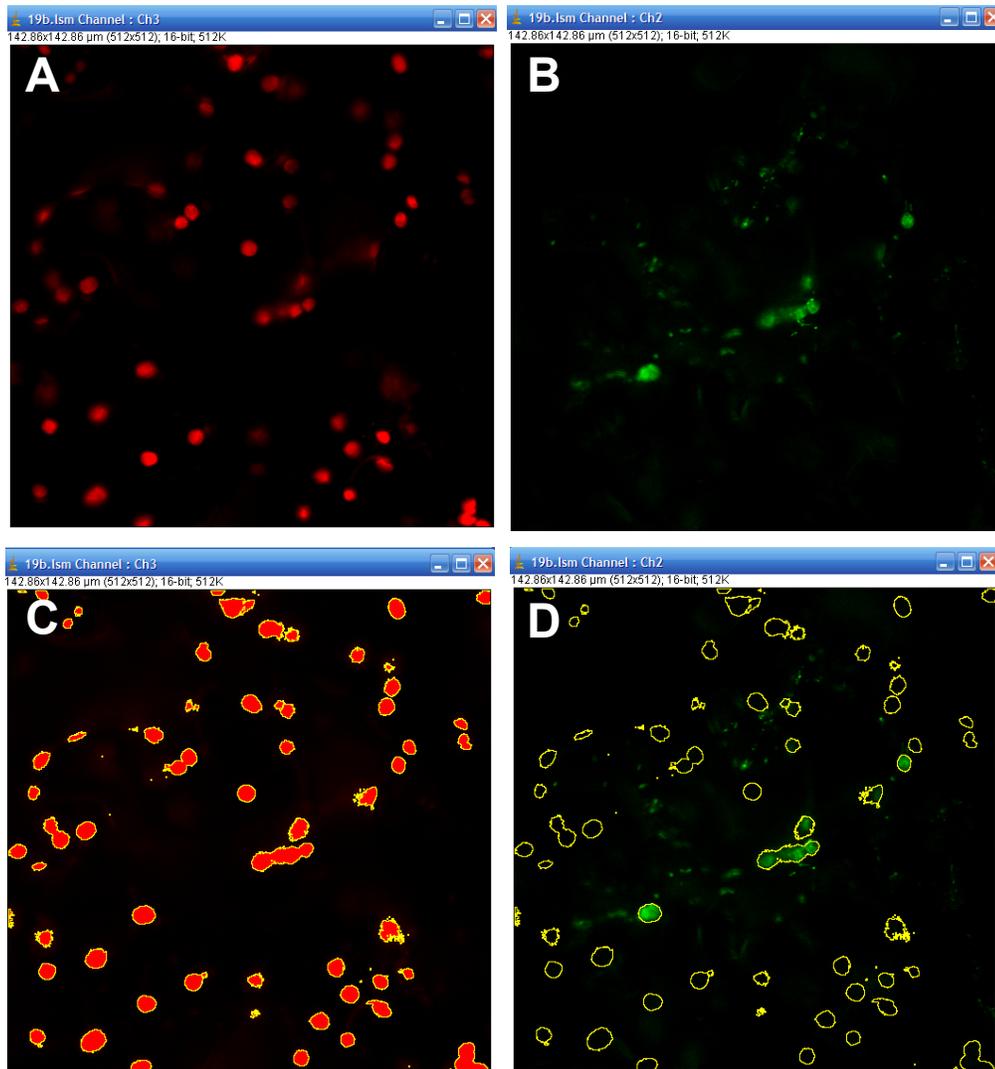


Figura 4 - Análise das imagens de microscopia confocal, utilizando o programa ImageJ. (A) canal vermelho, referente à autofluorescência da clorofila; (B) canal verde, referente à fluorescência da GFP; (C) canal vermelho com a seleção da autofluorescência de cloroplastos; (D) seleção da autofluorescência de cloroplastos (círculos amarelos) sobreposta à imagem do canal verde

Tabela 3 - Exemplo de medidas obtidas na análise de imagens, utilizando o programa ImageJ

	Área	Média	Densidade	Valores Relativos
Cloroplasto	237,215 (A _c)	1.006,173 (M _c)	238.679,328 (D _c)	0,7670 (D _c /D _t)
Total da célula	383,363 (A _t)	811,744 (M _t)	311.192, 615 (D _t)	
Mitocôndria	146,148 (A _m)	496 (M _m)	72.513,287 (D _m)	0,2330 (D _m /D _t)

3.14 Análise estatística dos dados obtidos no programa ImageJ

Os métodos estatísticos usados estão detalhados no livro-texto de estatística, Zar (1974) e em edições posteriores (1984, 1999) ou em manuais de pacotes avançados como o SAS. Foram empregados também programas comerciais de estatística (Minitab * MiniTab, Inc., 1993 e InStat * GraphPad Software, 1993) e programas desenvolvidos em ambiente DOS ou em interface gráfica para Windows pelo grupo do Dr. Paulo A. Otto, do Departamento de Genética e Biologia Evolutiva, Instituto de Biociências, USP.

3.15 Transcrição e tradução *in vitro*

As construções realizadas nos plasmídeos pCRII-TOPO (Invitrogen) ou pSC-A Strataclone (Stratagene) foram utilizadas para transcrição e tradução *in vitro*, com o *kit TnT Coupled Reticulocyte Lysate Systems* (Promega), seguindo as instruções do fabricante. Foram utilizadas as polimerases T7 ou SP6, dependendo do vetor e da orientação de clonagem do fragmento, e metionina radioativa (Redivue L-[³⁵S] methionine, Amersham Bioscience). A análise foi realizada em gel de acrilimida desnaturante (SDS) 12 ou 15% (LAEMMLI, 1970), dependendo do peso molecular da proteína.

3.16 Isolamento de cloroplastos

Os cloroplastos foram isolados a partir de 60 g de folhas de plantas de ervilha com 10 dias de idade, crescidas em casa de vegetação. As folhas foram levemente trituradas em um liquidificador com 180 mL de tampão GB 1 x (Hepes 50 mM, pH 7,3; sorbitol 330 mM; 0,1% de BSA; MgCl₂ 1 mM; MnCl₂ 1 mM; EDTA 2 mM). O extrato foi filtrado através de gaze, distribuído em quatro tubos e centrifugado durante 5 minutos, a 4 krpm, a 4° C. O sobrenadante foi removido e o precipitado, ressuspensão em 1 mL do mesmo tampão.

Para a preparação do gradiente, foram misturados 60 mL de Percoll 100%, 60 mL de tampão GB 2 x e 120 mg de glutatona redutase. A solução foi distribuída em tubos de 50 mL que foram centrifugados durante 35 minutos, a 18 krpm, a 4° C, sem frenagem.

O extrato das folhas foi submetido a gradiente de Percoll. Os tubos foram centrifugados durante 30 minutos, a 7 krpm, a 4° C, sem frenagem. A banda referente aos cloroplastos foi recuperada e diluída em tampão GB 1 x. O material foi novamente centrifugado por 4 minutos, a 4 krpm, a 4° C. O sobrenadante foi descartado e o precipitado ressuspensão em pequeno volume de tampão GB 1 x, dependendo da quantidade do precipitado. Em seguida, procedeu-se à quantificação dos cloroplastos. Para tal, 2,5, 5 e 10 µL do extrato foram diluídos, separadamente, em acetona 80%. A absorbância foi medida em um espectrofotômetro nos comprimentos de onda 663 e 645 nm e os valores foram aplicados às formulas (1) e (2):

$$[\text{clorofila A}] = 20,2 \times A_{645} + 8,02 \times A_{663} \quad (1)$$

$$[\text{clorofila B}] = 8,14 \times A_{663} + 20,27 \times A_{645} \quad (2)$$

Os valores são obtidos em µg de cloroplastos, ajustando-se para uma concentração final de 1 µg/µL.

3.17 Isolamento de mitocôndrias

Para o isolamento de mitocôndrias, foram utilizadas cinco batatas médias deixadas no escuro por 12 h. As batatas foram descascadas e passadas em um processador juntamente com o tampão de trituração 3 x (sacarose 0,9 M; $\text{Na}_4\text{P}_2\text{O}_7$ 90 mM; EDTA 6 mM; 2,4% de PVP; 0,15% de cisteína; glicina 15 mM; BSA 6 mM, pH 7,5). O suco obtido foi filtrado em uma membrana de nylon de 45 μm e distribuído em tubos de 50 mL que foram centrifugados durante 10 minutos, a 3 krpm, a 4° C. Em seguida, o sobrenadante foi distribuído em tubos de 30 mL que foram centrifugados por 12 minutos, a 9 krpm, a 4° C. O sobrenadante foi descartado e o precipitado foi ressuspenso em 6 mL de tampão de lavagem com BSA (sacarose 0,3 M; tampão fosfato Gomori 10 mM, pH 7,5; EDTA 1 mM; 0,1% BSA; glicina 5 mM). O material foi transferido para microtubos de 2 mL e centrifugado por 10 minutos, a 3 krpm, a 4° C. Em seguida, o sobrenadante foi transferido para novos tubos de 2 mL e centrifugados durante 10 minutos, a 10 krpm, a 4° C. O sobrenadante foi descartado e o precipitado ressuspenso em 400 μL de tampão de lavagem com BSA. O material foi submetido a gradiente de Percoll (sacarose 0,3 M; tampão fosfato Gomori 10 mM, pH 7,5; EDTA 1 mM; 0,1% de BSA; 30% de Percoll) e centrifugado por 15 minutos, a 25 krpm, a 4° C. A banda referente às mitocôndrias foi isolada e lavada com o tampão com BSA. O material foi centrifugado por 10 minutos, a 12 krpm, a 4° C. O sobrenadante foi removido e o precipitado foi ressuspenso em 2 mL de tampão de lavagem sem BSA. O procedimento foi repetido por mais uma vez e, ao final, as mitocôndrias foram ressuspenso em 150 μL de tampão de lavagem sem BSA. Para quantificá-las, foi utilizado o reagente Bradford (BRADFORD, 1976) e a densidade óptica foi medida a um comprimento de onda de 595 nm.

3.18 Importação *in vitro* de proteínas para organelas

3.18.1 Importação *in vitro* de proteínas para mitocôndrias

Inicialmente foi feita uma mistura de importação contendo, para cada reação, 50 μg de mitocôndria, 25 μL de tampão de importação para mitocôndrias 2 x (manitol 600 mM; tampão fosfato Gomori 400 mM, pH 7,5; malato 2 mM; NADH 2 mM; DTT 2 mM; Hepes 20 mM, KOH até pH 7,5), 0,5 μL de ADP 4 mM, 1 μL de ATP 100mM pH 7,2 e

água destilada para o volume final de 45 μ L. Em seguida, foram montadas as reações. Para cada proteína a ser testada, foram realizadas quatro reações, conforme esquematizado abaixo (Tabela 4).

Tabela 4 - Montagem das reações para o teste de importação *in vitro* de proteínas para mitocôndrias

Tubos	1	2	3	4
Mistura de importação	45 μ L	45 μ L	45 μ L	45 μ L
Valinomicina (100 μ M)	-	-	1 μ L	1 μ L
Proteína ⁽¹⁾	5 μ L	5 μ L	5 μ L	5 μ L
Proteinase K (10 μ g/ μ L)	-	0,5 μ L	-	0,5 μ L
PMSF ⁽²⁾ 0,1 M	-	4 μ L	-	4 μ L

⁽¹⁾Proteínas obtidas pela tradução *in vitro*

⁽²⁾fluoreto de fenil-metano-sulfonil

A mistura de importação foi adicionada aos quatro tubos e a valinomicina, um inibidor do transporte para mitocôndrias, aos tubos 3 e 4. Após incubação à temperatura ambiente durante 10 minutos, adicionou-se a proteína (produto da tradução *in vitro*) e o material foi incubado à temperatura ambiente durante 30 minutos, sob agitação leve. Terminado o tempo, adicionou-se proteinase K aos tubos 2 e 4, para a digestão das proteínas que não foram transportadas para o interior das mitocôndrias, e os tubos foram incubados no gelo, durante 10 minutos. Adicionou-se, então, PMSF, um inibidor da atividade da proteinase K.

Os produtos das reações foram submetidos a gradiente de sacarose 27% (27% de sacarose; EDTA 1 mM; tampão fosfato Gomori 10 mM, pH 7,5; 0,1% de BSA) e centrifugados a 10 krpm, durante 15 minutos. O sobrenadante foi descartado e o precipitado foi ressuspenso em 10 a 15 μ L de tampão Laemmli 1 x (2% de SDS, 10% de glicerol, 5% de 2-mercaptoetanol, 0,002% de azul de bromofenol e Tris-HCl 0,075 M). A análise foi realizada após eletroforese a 25 mA em gel desnaturante de acrilamida (LAEMMLI, 1970), 12 ou 15%, dependendo do peso molecular da proteína.

3.18.2 Importação *in vitro* de proteínas para cloroplastos

Inicialmente foram preparadas duas misturas de importação, com ou sem GTP, contendo, cada reação: 25 μL de tampão de importação para cloroplastos 2 x (Hepes 100 mM pH 7,3 e sorbitol 660 mM), metionina 200 mM, GTP 50 mM e água destilada para o volume de 45 μL . Em seguida, foram montadas as reações. Para cada proteína testada, foram realizadas quatro reações, conforme esquematizado abaixo (Tabela 5):

Tabela 5 - Montagem das reações para o teste *in vitro* de importação de proteína para cloroplastos

Tubos	1	2	3	4
Mistura de importação	45 μL	45 μL	45 μL	45 μL
Proteína ⁽¹⁾	5 μL	5 μL	5 μL	5 μL
Cloroplastos (1 $\mu\text{g}/\mu\text{L}$)	100 μL	100 μL	100 μL	100 μL
Proteinase K (10 $\mu\text{g}/\mu\text{L}$)	-	0,5 μL	-	0,5 μL
PMSF	-	4 μL	-	4 μL

⁽¹⁾Proteínas obtidas pela tradução *in vitro*

Primeiramente, os cloroplastos foram incubados à temperatura ambiente durante 10 minutos. Em seguida, os tubos 3 e 4 foram envolvidos em papel alumínio, para inibição do transporte. Adicionaram-se aos tubos a mistura de importação para cloroplastos, a proteína e os cloroplastos. O material foi incubado à temperatura ambiente, durante 30 minutos, sob agitação leve. Em seguida, adicionou-se proteinase K aos tubos 2 e 4, incubando-se em gelo, durante 10 minutos, quando foi adicionado PMSF. Os produtos das reações foram submetidos a gradiente de Percoll (0,1% de Percoll e tampão de importação para cloroplastos 1 x). Para cada amostra foram utilizados 900 μL do gradiente. Os tubos foram centrifugados a 10 krpm, durante 5 minutos. O sobrenadante foi descartado e o precipitado foi ressuspensão em 1 mL de tampão de importação para cloroplastos 1 x. As amostras foram novamente centrifugadas, o sobrenadante foi descartado e o precipitado foi ressuspensão em 4 μL de tampão Laemmli 5 x. A análise foi realizada após eletroforese, a 25 mA, em gel desnaturante de acrilamida (LAEMMLI, 1970), 12 ou 15%, dependendo do peso molecular da proteína.

4 RESULTADOS E DISCUSSÃO

4.1 A conservação do duplo direcionamento de proteínas em *Arabidopsis thaliana* e *Oryza sativa* sugere um padrão similar na evolução de famílias gênicas

4.1.1 Busca por homólogos de proteínas com duplo direcionamento

As duplicações gênicas representam um mecanismo central para mudanças evolutivas. Como consequência, mais de um terço do genoma eucarioto consiste de duplicações gênicas ou de famílias gênicas (LYNCH; CONERY, 2000). Duplicações em grande escala e até mesmo duplicações genômicas ocorreram na história evolutiva de eucariotos (VANDEPOELE; SIMILLION; VAN DE PEER, 2003). Com o objetivo de compreender a evolução das famílias gênicas que codificam proteínas com duplo direcionamento e sua possível correlação com duplicações de genes, foi feita uma busca, no genoma de *Arabidopsis* e de arroz, por ortólogos e parálogos de proteínas de *A. thaliana* previamente relatadas como tendo duplo direcionamento. Foram considerados os casos em que o duplo direcionamento foi confirmado por dados experimentais, como importação *in vitro* de proteínas para organelas, *Western Blotting* ou fusão da proteína com um gene repórter. Até o momento, 49 proteínas de *A. thaliana* foram demonstradas, por dados experimentais, ser endereçadas simultaneamente para mais de um compartimento celular, a maioria para mitocôndrias e cloroplastos. Essas proteínas estão distribuídas em 42 famílias gênicas (Tabela 6). Todos os genes encontrados em *Arabidopsis* têm ortólogos no genoma de arroz. A combinação de dados de similaridade de seqüências, presença de domínios funcionais e de expressão gênica permitiu o estabelecimento de ortologia entre as seqüências de *Arabidopsis* e de arroz. Das 42 famílias gênicas isoladas, 20 apresentam o mesmo número de parálogos em *Arabidopsis* e em arroz, o que sugere um padrão similar de expansão gênica das famílias que apresentam membros com duplo direcionamento, nas duas espécies. As 22 famílias restantes podem ser divididas em dois grupos: um com dez famílias, as quais apresentam maior número de parálogos em *Arabidopsis*, quando comparado ao de arroz, e outro com 12 famílias, apresentando maior número de parálogos em arroz do que em *Arabidopsis*. Uma vez que o genoma de arroz é cerca de quatro vezes maior

que o de Arabidopsis e que arroz tem um número 40% maior de seqüências únicas expressas que Arabidopsis (THE INSTITUTE FOR GENOMIC RESEARCH) (2006), pode-se considerar que os números de parálogos encontrados em Arabidopsis e arroz não devem diferir de modo significativo.

4.1.2 Predição da localização celular de proteínas

A predição da localização celular de membros de famílias com duplo direcionamento foi realizada com os programas computacionais Predotar, TargetP, iPSORT e MultiLoc/TargetLoc (Tabela 6). A comparação entre os resultados dos quatro programas mostra que eles são incongruentes.

A predição da localização de proteínas com duplo direcionamento é um problema para todos os programas utilizados. Das 52 proteínas com duplo direcionamento testadas, 45 localizam-se em cloroplastos e mitocôndrias. Se forem consideradas somente essas 45 situações, em apenas onze, pelo menos um programa indicou não haver consenso (indicado na Tabela 6 pela abreviação SC) quanto ao direcionamento para cloroplastos ou mitocôndrias. Nesses casos, os valores de probabilidade de direcionamento para essas duas organelas foram próximos, o que pode ser uma indicativa de duplo direcionamento. Os programas TargetP e TargetLoc foram os que mais apontaram essa dúvida, com cinco cada um. O programa Predotar indicou essa incerteza apenas uma vez e o iPSORT, nenhuma vez.

Para a predição de proteínas localizadas exclusivamente em mitocôndrias ou em cloroplastos, a eficiência dos programas testados é melhor. Na Tabela 6, existem 11 proteínas cuja localização foi experimentalmente confirmada em uma dessas duas organelas. Em dois casos, nenhum dos programas indicou corretamente a localização das proteínas. Os programas TargetLoc, iPSORT, TargetP e Predotar acertaram nove, oito, sete e seis vezes, respectivamente, a localização das outras nove proteínas.

Tabela 6 - Proteínas de Arabidopsis com duplo direcionamento codificadas por um único gene, seus homólogos em arroz e predição de sua localização celular

(continua)

	Loci	Predotar (1)	TargetLoc (1)	Target P (1)	IPSORT (1)	Localização Experimental	Referências
Alanil-tRNA sintetase	At1g50200	M	C	SC	M	M/C	(DUCHÊNE et al., 2005)
	At5g22800	OL	C	C	M	M/C	
	Os10g10244	M	M	M	M	-	
	Os06g13660	C	C	C	M	-	
	Os07g15440	OL	OL	OL	M	-	
Asparaginil-tRNA sintetase	At4g17300	C	C	C	M	M/C	(DUCHÊNE et al., 2005)
	At1g70980	OL	OL	OL	OL	-	
	At5g56680	OL	OL	OL	OL	-	
	Os07g30200	M	M	SC	M	-	
	Os01g27520	OL	OL	OL	OL	-	
Aspartil-tRNA sintetase	Os12g22600	OL	C	C	OL	-	(DUCHÊNE et al., 2005)
	At4g33760	M	M	SC	M	M/C	
	At4g31180	OL	OL	C	OL	-	
	At4g26870	OL	OL	OL	OL	-	
	Os01g06020	OL	OL	OL	OL	-	
Cisteinil-tRNA sintetase	Os02g46130	OL	OL	SC	OL	-	(DUCHÊNE et al., 2005)
	Os02g04700	OL	OL	OL	OL	-	
	At2g31170	C	SC	C	M	M/C	
	At3g56300	OL	OL	OL	OL	-	
	At5g38830	OL	OL	OL	OL	-	
Glicil-tRNA sintetase	Os09g38420	RE	M	M	M	-	(DUCHÊNE et al., 2005)
	Os10g32570	OL	OL	OL	OL	-	
	Os03g04960	OL	OL	OL	OL	-	
	At1g29880	OL	C	M	M	Cit/M	
	At3g48110	RE	C	C	C	M/C	
Glutamil-tRNA syntetase	Os08g42560	OL	OL	OL	OL	-	(DUCHÊNE et al., 2005)
	Os04g32650	M	C	M	M	-	
	At5g64050	OL	SC	M	M	M/C	
	At5g26710	OL	C	C	S	-	
	Os10g22380	C	C	C	OL	-	
Histidil-tRNA sintetase	Os01g16520	OL	OL	OL	OL	-	(DUCHÊNE et al., 2005)
	Os02g02860	OL	SC	SC	M	-	
	At3g46100	M	M	M	S	M/C	
	At3g02760	OL	OL	OL	OL	-	
	Os02g51830	M	M	SC	M	-	
Lisil-tRNA sintetase	Os05g05840	C	C	OL	S	-	(DUCHÊNE et al., 2005)
	At3g13490	C	C	M	M	M/C	
	At3g11710	OL	OL	OL	OL	-	
	Os02g41470	OL	M	SC	M	-	
	Os03g38980	OL	OL	OL	OL	-	

Tabela 6 - Proteínas de Arabidopsis com duplo direcionamento codificadas por um único gene, seus homólogos em arroz e predição de sua localização celular

							(continuação)
	Loci	Predotar (1)	TargetLoc (1)	Target P (1)	IPSORT (1)	Localização Experimental	Referências
Metionil-tRNA sintetase	At3g55400	C	C	C	C	M/C	(DUCHÊNE et al., 2005)
	At4g13780	OL	OL	OL	OL	-	
	Os03g11120	M	M	M	M	-	
	Os10g26050	OL	OL	OL	M	-	
	Os06g31210	OL	C	OL	C	-	
Fenilalanil-tRNA sintetase	At3g58140	OL	SC	C	C	M/C	(DUCHÊNE et al., 2005)
	At4g39280	OL	OL	OL	OL	-	
	At1g72550	OL	OL	OL	OL	-	
	Os12g34860	M	SC	SC	M	-	
	Os10g26130	OL	OL	OL	OL	-	
Prolil-tRNA sintetase	Os05g48510	OL	OL	OL	OL	-	(DUCHÊNE et al., 2005)
	At5g52520	OL	C	SC	M	M/C	
	At3g62120	OL	OL	OL	OL	-	
	Os07g07060	OL	OL	M	M	-	
Seril- tRNA sintetase	Os12g25710	OL	OL	OL	OL	-	(DUCHÊNE et al., 2005)
	At1g11870	OL	C	M	S	M/C	
	At5g27470	OL	OL	OL	OL	-	
	Os03g10190	OL	OL	OL	OL	-	
	Os11g39670	OL	C	SC	M	-	
Treonil- tRNA sintetase	Os01g37837	OL	OL	OL	OL	-	(DUCHÊNE et al., 2005)
	At2g04842	C	C	C	C	M/C	
	At5g26830	M	M	M	M	-	
	At1g17960	OL	OL	OL	OL	-	
	Os08g19850	M	M	M	M	-	
Tryptofanil-tRNA sintetase	Os02g33500	SC	M	SC	M	-	(DUCHÊNE et al., 2005)
	At2g25840	OL	SC	C	C	M/C	
	At3g04600	OL	OL	OL	OL	-	
	Os01g54020	OL	SC	SC	M	-	
	Os12g35570	OL	OL	SC	OL	-	
Tirosil-tRNA sintetase	Os07g17770	OL	M	M	OL	-	(DUCHÊNE et al., 2005)
	At3g02660	C	C	C	M	M/C	
	At2g33840	OL	OL	C	OL	-	
	At1g28350	OL	OL	OL	OL	-	
	Os01g31610	M	SC	SC	M	-	
	Os08g05490	OL	C	C	OL	-	
	Os08g23110	OL	OL	C	OL	-	
AtPreP metaloprotease	Os08g09260	OL	OL	OL	OL	-	(BHUSHAN et al., 2005)
	At3g19170	M	M	C	M	M/C	
	At1g49630	M	SC	C	M	M/C	
AtSufE	Os02g52390	M	C	M	M	-	(XU; MOLLER, 2006)
	At4g26500	C	C	C	C	M/C	
	Os09g09790	C	C	C	M	-	

Tabela 6 - Proteínas de Arabidopsis com duplo direcionamento codificadas por um único gene, seus homólogos em arroz e predição de sua localização celular

	Loci	Predotar (1)	TargetLoc (1)	Target P (1)	IPSORT (1)	Localização Experimental	Referências
Glutamina sintetase	At5g35630	C	C	C	M	M/C	(TAIRA et al., 2004)
	At5g37600	OL	OL	OL	OL	Cit	
	At1g66200	OL	OL	OL	OL	Cit	
	At3g17820	OL	OL	OL	OL	Cit	
	At5g16570	OL	OL	OL	OL	Cit	
	At1g48470	OL	OL	OL	OL	Cit	
	Os04g56400	C	C	C	C	-	
	Os02g50240	OL	OL	OL	OL	-	
	Os03g12290	OL	OL	OL	OL	-	
	Os03g50490	OL	OL	OL	OL	-	
Tiazol biosynthetic enzyme	At5g54770	C	C	C	C	M/C	(CHABREGAS et al., 2003)
	Os07g34570	C	C	C	M	M/C	
<i>Phage-type</i> RNA polimerase	At5g15700	OL	C	C	C	M/C	(HEDTKE et al., 1999; HEDTKE; BORNER; WEIHE, 2000; KUSUMI et al., 2004)
	At2g24120	C	C	C	C	C	
	At1g68990	M	M	M	OL	M	
	Os09g07120	C	C	C	C	M	
	Os06g44230	C	C	C	C	C	
γ -type DNA polimerases	At3g20540	OL	C	SC	M	M/C	(ELO et al., 2003; CHRISTENSEN et al., 2005)
	At1g50840	C	C	C	C	C	
	Os08g07840	C	C	C	C	-	
	Os08g07850	OL	C	SC	C	-	
	Os04g54500	C	SC	C	C	-	
DNA girase A	At3g10690	C	C	C	C	M/C	(WALL; MITCHENAL; MAXWELL, 2004)
	Os03g59750	M	M	M	M	-	
DNA ligase I	At1g08130	M	M	C	M	M/Núcleo	(SUNDERLAND et al., 2006)
	At1g49250	OL	OL	M	OL	-	
	At1g66730	OL	C	C	OL	-	
	Os10g34750	OL	OL	OL	OL	-	
	Os01g49180	C	C	C	M	-	
Proteínas OSB	At5g44785	OL	M	M	M	M/C	(ZAEGEL et al., 2006)
	At1g47720	OL	M	C	M	M	
	At4g20010	OL	C	OL	C	C	
	At1g31010	OL	M	C	M	-	
	Os01g72049	M	M	M	M	-	
	Os03g43420	C	C	C	C	-	
Peptideo deformilase	Os03g41530	OL	M	M	M	-	(GIGLIONE et al., 2000; DINKINS et al., 2003)
	At1g15390	OL	C	C	OL	M/C	
	At5g14660	OL	C	C	C	M/C	
	Os01g45070	M	M	M	M	-	
	Os01g44980	M	M	SC	M	-	
Pseudouridine sintetase	Os01g37510	RE	M	C	OL	-	(PEETERS; SMALL, 2001)
	At2g30320	OL	C	M	C	M/C	
	Os03g21980	M	C	M	M	-	

Tabela 6 - Proteínas de Arabidopsis com duplo direcionamento codificadas por um único gene, seus homólogos em arroz e predição de sua localização celular

							(continuação)	
	Loci	Predotar (1)	TargetLoc (1)	Target P (1)	IPSORT (1)	Localização Experimental	Referências	
Metionina aminopeptidase	At3g25740	OL	C	C	C	M/C	(GIGLIONE et al., 2000)	
	At4g37040	OL	C	M	M	M/C		
	At2g45240	OL	OL	SC	OL	Cit		
	At1g13270	C	C	C	C	C		
	Os02g52420	C	M	M	M	M/C		
	Os04g52100	OL	C	C	C	-		
	Os07g25410	OL	OL	OL	OL	-		
	Os10g36470	OL	OL	OL	M	-		
	Os07g32590	C	C	OL	M	-		
Criptocrome	At5g24850	OL	OL	OL	OL	M/C	(KLEINE; LOCKHART; BATSCHAUER, 2003)	
	Os06g45100	OL	C	OL	M	-		
Ferroquelatase	At5g26030	OL	C	C	C	M/C	(CHOW et al., 1998)	
	At2g30390	OL	C	C	C	C		
	Os09g12560	OL	M	OL	M	-		
	Os05g29760	OL	C	C	M	-		
<i>Preprotein and Amino Acid Transporter</i>	At5g24650	OL	OL	OL	OL	M/C	(MURCHA et al., 2007)	
	At3g49560	OL	OL	OL	OL	C		
	Os04g33220	OL	M	M	M	-		
Formato desidrogenase	At5g14780	OL	M	M	M	M/C	(HERMAN et al., 2002)	
	Os06g29180	M	M	M	M	-		
	Os06g29220	M	M	M	M	-		
Ascorbato peroxidase	At4g08390	C	C	C	M	M/C	(CHEW; WHELAN; MILLAR, 2003)	
	At1g77490	C	C	C	C	-		
	At4g35000	OL	OL	OL	OL	-		
	At4g35970	OL	OL	OL	OL	-		
	At3g09640	OL	OL	OL	OL	-		
	At1g07890	OL	OL	OL	OL	-		
	At4g32320	OL	M	C	M	-		
	At4g09010	C	C	C	C	-		
	Os04g35520	C	M	C	M	-		(TEIXEIRA et al., 2006)
	Os02g34810	C	C	C	C	-		
	Os12g07820	M	M	M	M	M		
	Os12g07830	M	C	M	M	-		
	Os04g14680	OL	OL	OL	OL	P		
	Os08g43560	OL	M	SC	OL	-		
	Os03g17690	OL	OL	OL	OL	-		
	Os07g49400	OL	OL	OL	OL	-		
Os08g41090	C	C	C	M	-			
Glutaciona redutase	At3g54660	C	C	C	C	M/C	(CHEW; WHELAN; MILLAR, 2003)	
	At3g24170	OL	OL	OL	OL	-		
	Os03g06740	C	C	C	C	-		
	Os10g28000	OL	OL	OL	M	-		
	Os02g56850	OL	OL	OL	OL	-		

Tabela 6 - Proteínas de Arabidopsis com duplo direcionamento codificadas por um único gene, seus homólogos em arroz e predição de sua localização celular

							(conclusão)
	Loci	Predotar (1)	TargetLoc (1)	Target P (1)	IPSORT (1)	Localização Experimental	Referências
Monodeidroascorbato redutase	At1g63940	C	C	C	C	M/C	(OBARA; SUMI; FUKUDA, 2002; LISENBEE; LINGARD; TRELEASE, 2005)
	At3g52880	OL	OL	OL	OL	P	
	At3g27820	RE	C	OL	S	P	
	At5g03630	OL	OL	OL	OL	Cit	
	At3g09940	OL	OL	OL	OL	Cit	
	Os08g05570	C	C	SC	C	M/C	
	Os09g39380	OL	OL	OL	OL	-	
	Os08g44340	OL	OL	OL	OL	-	
	Os02g47800	OL	OL	M	S	-	
	Os02g47790	OL	OL	OL	S	-	
Mercaptopiruvato sulfotransferase	At1g79230	SC	C	C	M	M/C	(NAKAMURA; YAMAGUCHI; SANO, 2000)
	At1g16460	OL	SC	OL	M	Cit	
	Os12g41500	OL	OL	OL	OL	-	
	Os02g07040	OL	OL	SC	OL	-	
Fosfatidilglicerofosfato sintetase I	At2g39290	OL	M	C	M	M/C	(BABYCHUK et al., 2003)
	At3g55030	OL	SC	OL	OL	-	
	Os03g17520	C	SC	C	M	-	
Proteínas de ligação ao RNA	At4g13850	OL	S	M	M	M/C	(MORGANTE; GUALBERTO, 2007)
	At3g23830	OL	M	M	M	M/C	
	At5g47320	OL	M	M	M	M/C	
	Os01g68790	OL	SC	C	M	-	
Proteínas AP2/EREBP	At2g44940	C	Núcleo	C	M	C/Núcleo	(SCHWACKE et al., 2007)
	Os04g46400	OL	Cit	C	OL	-	
Holocarboxilase sintetase	At2g25710	OL	OL	M	M	C/Cit	(PUYAUBERT; DENIS; CLAUDE, in press)
	At1g37150	OL	OL	OL	OL	-	
	Os02g07040	M	C	M	OL	-	
<i>Vacuoleless 1</i>	At2g38020	OL	OL	OL	OL	T/CPV	(ROJO et al., 2003)
	Os01g47650	OL	C	OL	OL	-	
<i>Vacuolar protein sorting11</i>	At2g05170	OL	OL	OL	OL	T/CPV	(ROJO et al., 2003)
	Os04g31390	OL	OL	OL	OL	-	
<i>Vacuolar protein sorting33</i>	At3g54860	OL	OL	OL	OL	T/CPV	(ROJO et al., 2003)
	Os04g14650	OL	OL	OL	OL	-	

Nota: Abreviações, C, cloroplasto; Cit, citosol; OL, outra localização (o programa não foi capaz de determinar com precisão a localização celular da proteína); RE, retículo endoplasmático; M, mitocôndria; SC, sem consenso (o programa indica duas prováveis localizações celulares com valores iguais ou menores que 0,5); P, Peroxissomo; CPV, compartimento pré-vacuolar; S, via secretória; T, tonoplasto.

(¹) Programas computacionais de predição da localização celular de proteínas

Esses dados revelam a imprecisão desses programas. No entanto, constituem ferramenta importante para guiar o início de estudos de localização celular de proteínas, pois fornecem uma inferência que deve ser checada experimentalmente. Schwacke et al. (2007) desenvolveram uma metodologia para obter um consenso entre os resultados gerados por esses programas. Seus testes com fatores de transcrição mostram que a imprecisão desses programas foi reduzida.

A análise combinada entre bioinformática, genômica comparativa e dados experimentais para a determinação da real localização celular de uma proteína é fundamental (HEAZLEWOOD et al., 2005).

4.1.3 Um método alternativo para a identificação de proteínas com duplo direcionamento?

Atualmente, existem muitas informações sobre a localização celular de proteínas, geradas por análises do proteoma de organelas isoladas de *A. thaliana*. Nas situações em que a organela pode ser eficientemente isolada, há dados numerosos e consistentes. Esse é o caso de mitocôndrias e cloroplastos, para os quais existem muitas informações organizadas em banco de dados (HEAZLEWOOD et al., 2004, entre outros). O subproteoma de cloroplastos foi também estudado, incluindo proteínas de tilacóide, estroma e envelope (SUN; EMANUELSSON; VAN WIJK, 2004; PELTIER et al., 2004, 2006; FERRO et al., 2003). Outros trabalhos ainda determinaram o proteoma de compartimentos celulares como o vácuolo (JAQUINOD et al., 2007), o nucléolo (PENDLE et al., 2005), os peroxissomos (REUMANN et al., 2007), entre outros.

Diante da grande quantidade de informações geradas por esses trabalhos, é até mesmo simples pensar que elas possam ser utilizadas para a confirmação experimental do duplo direcionamento de proteínas. Para testar essa possibilidade, os locos gênicos que codificam proteínas de *A. thaliana* com duplo direcionamento, listadas na Tabela 6, foram utilizados para uma busca no programa SUBA Database (*The SubCellular Proteomic Database*) (HEAZLEWOOD et al., 2007). Dos 49 locos consultados, o duplo direcionamento foi confirmado para apenas três: At3g10690, At5g14780 e At3g54660. Em 12 casos foi possível a confirmação do direcionamento para apenas um dos dois

compartimentos e em 32 situações não havia dados de espectrofotometria de massa para a confirmação de qualquer localização celular das proteínas. Em dois casos, o resultado obtido para a localização das proteínas não estava de acordo com o previamente demonstrado por outras técnicas.

Os resultados obtidos demonstram a imprecisão desse tipo de análise para a detecção da dupla localização de proteínas na célula. Uma limitação a análises proteômicas é a detecção de proteínas menos abundantes e proteínas presentes somente em um estágio único do desenvolvimento ou em um tipo celular específico (falsos negativos). Há também o problema de contaminantes, fazendo com que a origem de uma proteína seja erroneamente indicada (falsos positivos), (HEAZLEWOOD; MILLAR, 2005).

Possivelmente, com o avanço das técnicas de purificação de organelas e detecção de proteínas, futuramente esse tipo de análise poderá ser viável.

4.1.4 O tamanho das famílias gênicas de membros com duplo direcionamento

Análises genômicas em grande escala revelaram a existência de mais de 800 famílias gênicas em *A. thaliana* (TAIR). Foi realizada uma análise comparativa entre o tamanho de famílias gênicas que apresentam ou não membros com duplo direcionamento em *A. thaliana*. Foram tomadas, ao acaso, 674 famílias cujos membros não têm duplo direcionamento e 26, com membros que têm duplo direcionamento. A Figura 5 mostra que 85% das famílias gênicas de membros com duplo direcionamento apresentam dois ou três parálogos, enquanto isso ocorre em 47% daquelas sem membros com duplo direcionamento. Por outro lado, 21% das famílias gênicas sem membros com duplo direcionamento possuem mais de 11 parálogos, enquanto naquelas com membros com duplo direcionamento, o número máximo de parálogos é oito.

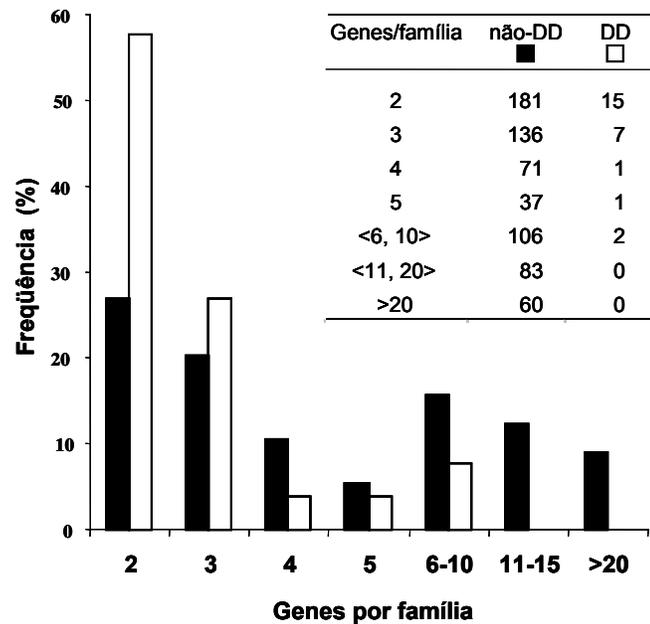


Figura 5 - Comparação entre o número de genes por família gênica de Arabidopsis de acordo com a presença (DD) ou ausência (não-DD) de membros com duplo direcionamento (DD). Os números absolutos estão listados acima do histograma

Apesar de números similares de parálogos já terem sido demonstrados no caso de proteínas sem duplo direcionamento, como para os genes de Arabidopsis e arroz que codificam fatores de transcrição (SHIU et al., 2005), os dados aqui obtidos, em que foram consideradas ao acaso famílias gênicas sem membros com duplo direcionamento, indicam que o número de parálogos nas famílias gênicas sem membros com duplo direcionamento é maior do encontrado naquelas que possuem membros com duplo direcionamento. Isso sugere que pressões de seleção distintas atuaram na evolução dessas famílias.

4.1.5 Análise filogenética das famílias gênicas de membros com duplo direcionamento

Para a validação da determinação de parálogos, foram construídas árvores filogenéticas para as famílias que possuem mais de uma isoforma em *Arabidopsis* ou em arroz. Quando as famílias gênicas das monodeidroascorbato redutases (MDAR) e das metionina aminopeptidases (MAP) são analisadas, por exemplo, pode-se observar que as seqüências de proteínas com duplo direcionamento em *Arabidopsis* agrupam-se com isoformas de arroz (Figura 6).

A proteína MDAR codificada pelo gene no loco At1g63940 é transportada para mitocôndrias e cloroplastos por um mecanismo de múltiplos sítios de início da transcrição (OBARA; SUMI; FUKUDA, 2002). Essa proteína pertence a uma família gênica que contém cinco membros em *Arabidopsis* e em arroz. A análise filogenética demonstrou que o ortólogo de arroz Os08g05570 agrupa-se unicamente com At1g63940 (Figura 6a). Isso sugere que a isoforma de arroz também apresente duplo direcionamento.

O mesmo fato se repete na família das MAP. A proteína codificada pelo gene no loco At4g37040 é transportada para cloroplastos e mitocôndrias (GIGLIONE et al., 2000) e agrupa-se unicamente com a isoforma de arroz, Os02g52420 (Figura 6b), o que também sugere seu duplo direcionamento.

A enzima da via de biossíntese do tiazol, THI1, é codificada por um gene de cópia única, tanto em *Arabidopsis* como em arroz. Chabregas et al. (2003) demonstraram que a isoforma de *Arabidopsis* é direcionada para mitocôndrias e cloroplastos. A isoforma de arroz, Os07g34570, pode também seguir esse padrão de direcionamento.

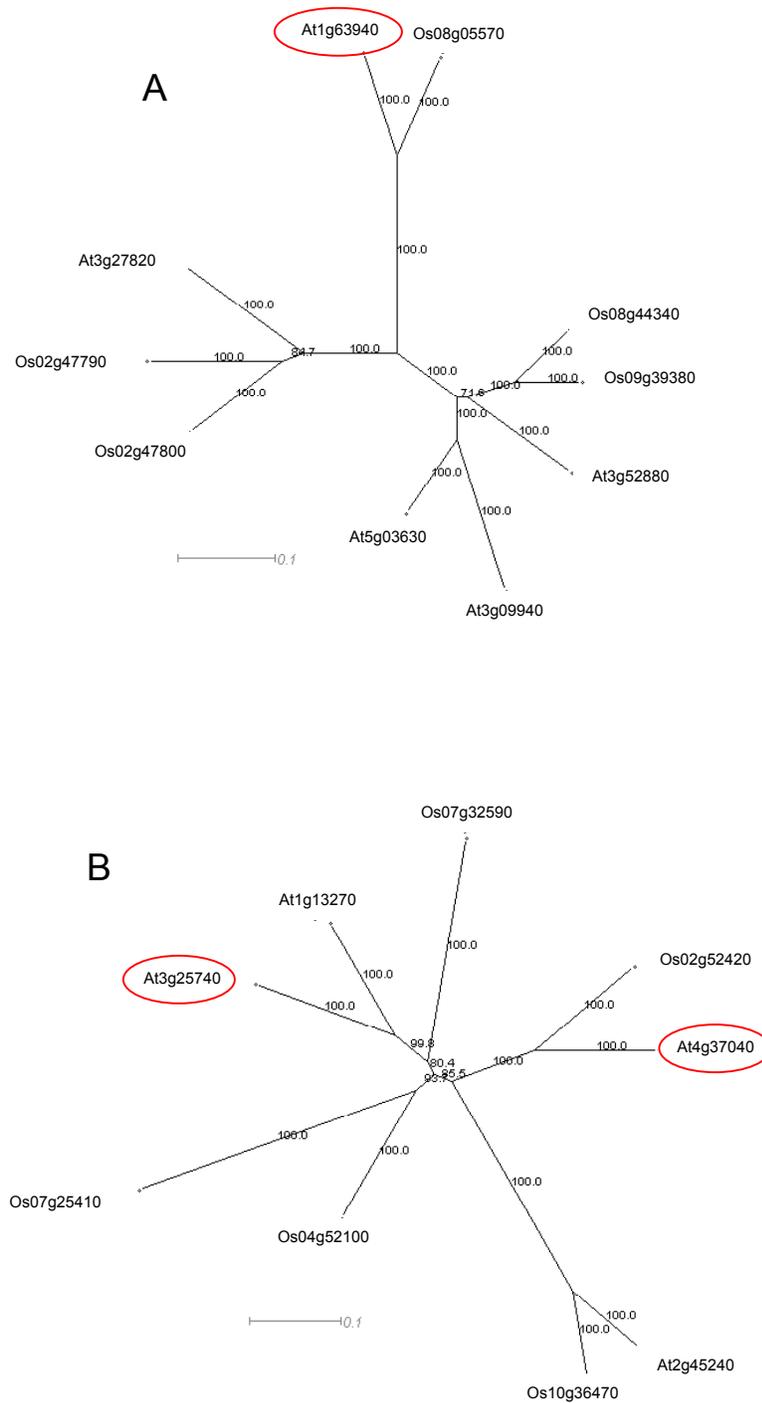


Figura 6 – Análise filogenética das famílias gênicas de membros com duplo direcionamento. Árvores filogenéticas de membros das famílias das MDAR (A) e MAP (B) de *A. thaliana* e de *O. sativa*. Os círculos vermelhos indicam as isoformas com duplo direcionamento

4.1.6 Localização celular em *Oryza sativa* de ortólogos de proteínas com duplo direcionamento em *Arabidopsis thaliana*

A localização celular das proteínas de arroz MDAR (Os08g05570), MAP (Os02g52420) e THI1 (Os07g34570) foi averiguada experimentalmente. A região amino-terminal dessas proteínas foi fusionada com a seqüência da GFP e as construções foram utilizadas para a transformação transiente de células do epitélio de cebola. Os resultados foram analisados por microscopia confocal. Foram utilizados como controles mitocondrial e cloroplastial os plasmídeos RecA-GFP (KÖHLER et al., 1997) e PMA4- β -GFP (DUBY; DEGAND; BOUTRY, 2001), respectivamente (Figura 7), e os plasmídeos PMD1d, L-MDAR e S-MDAR, gentilmente doados por Hiroo Fukuda (Universidade de Tóquio, Japão) e Thierry Meinel (CNRS, França), referentes às fusões, com a GFP, da região amino-terminal das proteínas MAP e MDAR de *A. thaliana*.

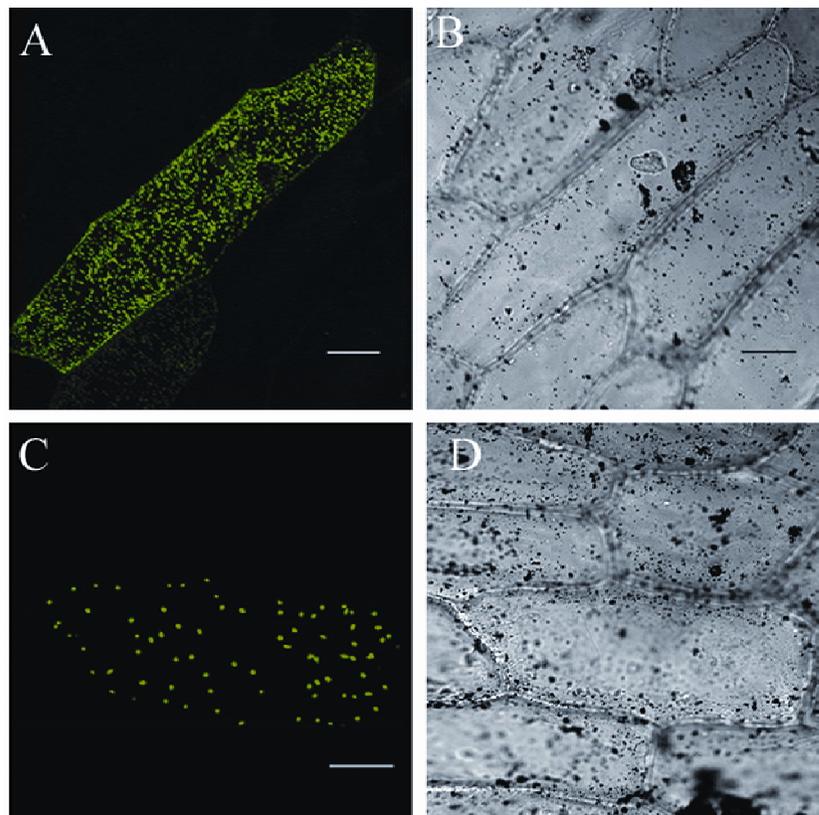


Figura 7 - Células do epitélio de cebola transformadas com as construções: controle mitocondrial PMA4- β -GFP (A e B) e controle plastidial RecA-GFP (C e D). A barra representa 50 μ m

O duplo direcionamento das proteínas MAP e MDAR de *A. thaliana* foi confirmado no sistema de expressão heteróloga utilizado (Figura 8, E e F e Figura 9, I a L), o que demonstra a adequação do sistema escolhido.

Conforme sugerido pelas análises de filogenia realizadas, observou-se que os ortólogos em arroz das proteínas MAP e MDAR são também direcionados para mitocôndrias e cloroplastos (Figura 8, G e H e Figura 9, M e N). No caso das proteínas MDAR, o duplo direcionamento foi conservado, mas não o mecanismo que permite esse processo. Em *A. thaliana*, a proteína MDAR codificada pelo gene no loco At1g63940 possui duplo direcionamento por um mecanismo de múltiplos sítios de início de transcrição. Existem dois transcritos que diferem em 21 pb de extensão na extremidade 5' e geram produtos traducionais distintos em tamanho. A proteína maior, L-MDAR, é direcionada às mitocôndrias, enquanto que a menor, S-MDAR, é direcionada aos cloroplastos (OBARA; SUMI; FUKUDA, 2002) (Figura 9, I a L). Em arroz, observou-se que a fusão da região amino-terminal da MDAR à GFP direciona a proteína para mitocôndrias e cloroplastos (Figura 9, M e N), o que indica um outro mecanismo de duplo direcionamento, possivelmente, pela presença de uma seqüência de direcionamento ambígua.

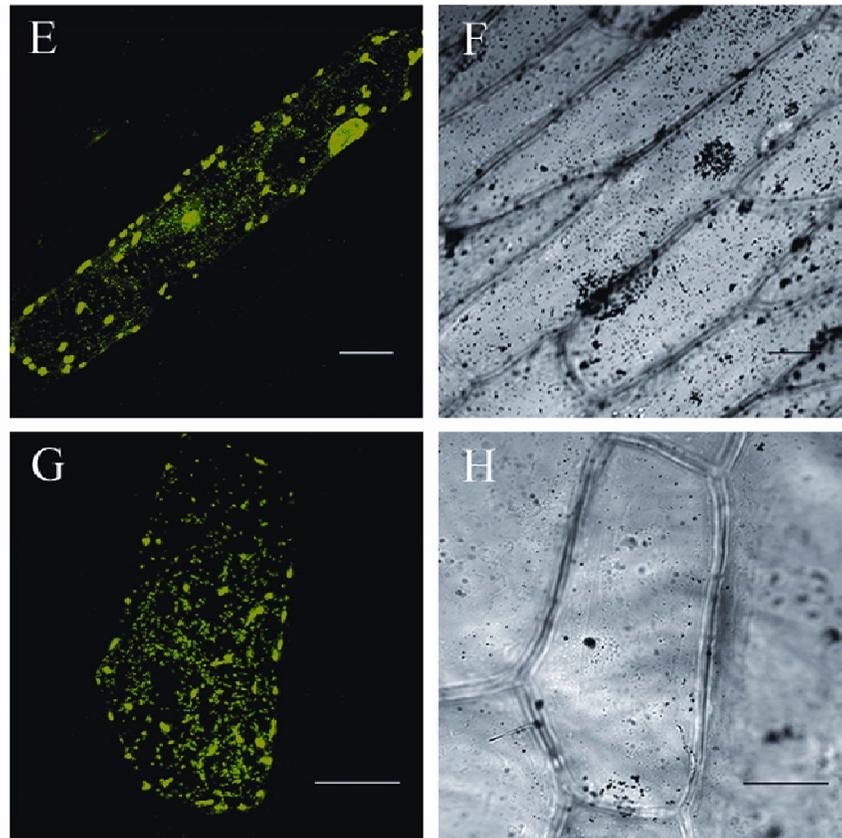


Figura 8 - O duplo direcionamento das proteínas MAP. Células do epitélio de cebola transformadas com as construções: PMG 1d, MAP de *A. thaliana* (E e F) e; PMA4-OsMAP, MAP de *O. sativa* (G e H). As imagens à esquerda correspondem ao canal verde, em que foi detectado o sinal da GFP, e as imagens à direita representam o contraste de fase. As barras correspondem 50 μm

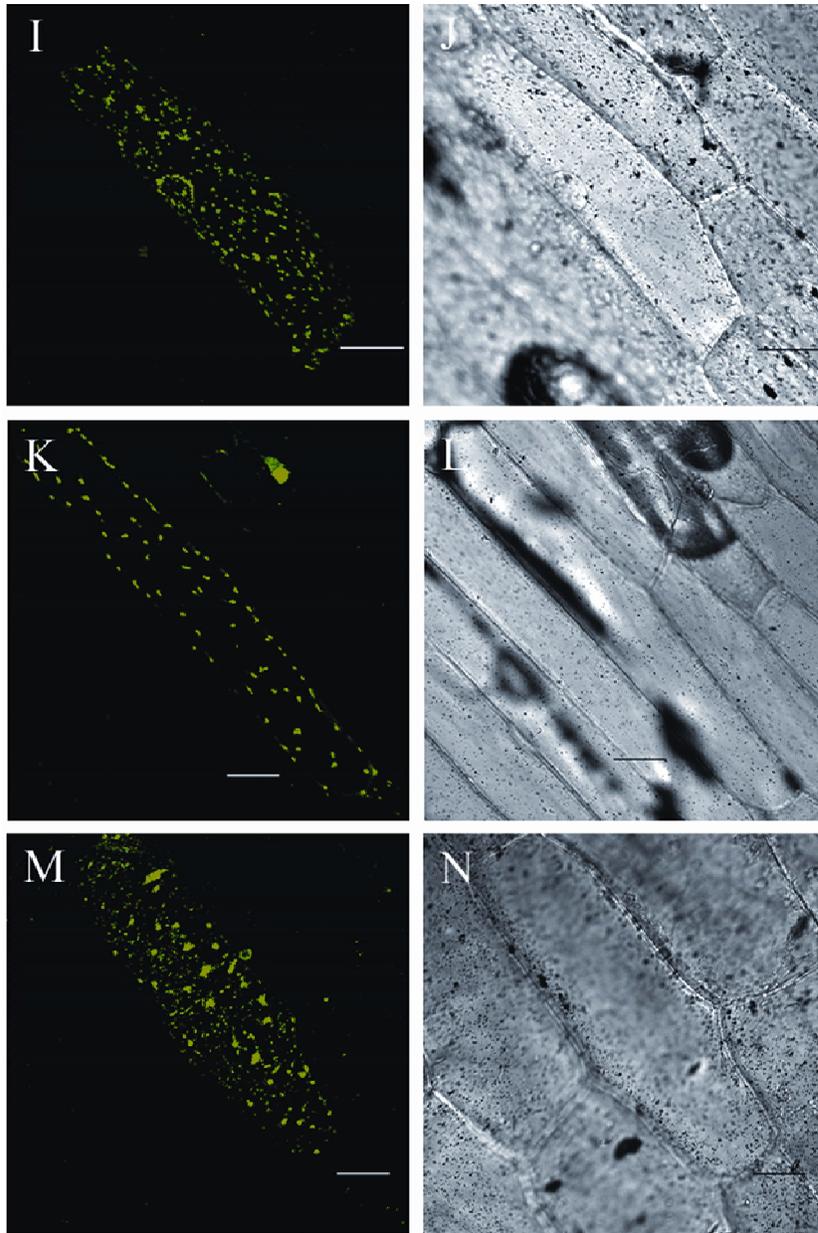


Figura 9 - O duplo direcionamento das proteínas MDAR. Células do epitélio de cebola transformadas com as construções: L-MDAR (I - J); S-MDAR (K - L) e; PMA4 - OsMDAR. As imagens à esquerda correspondem ao canal verde, em que foi detectado o sinal da GFP, e as imagens à direita representam o contraste de fase. As barras correspondem 50 μm

O duplo direcionamento da proteína THI1 de arroz também foi investigado experimentalmente. Apesar das imagens obtidas por microscopia confocal não estarem tão claras como as anteriormente apresentadas, os dados sugerem o duplo direcionamento de THI1 de arroz. Para a confirmação desse resultado, foi também realizada a co-transformação do epitélio de cebola com as construções PMA4-OsTHI1 e mRFP (gentilmente doada por I. Small) (Figura 10). Esta última é um controle mitocondrial o qual contém a seqüência de direcionamento da proteína CoxIV fusionada *in frame* à seqüência da RFP. Os resultados mostram uma certa sobreposição dos sinais da GFP e RFP, o que indica a localização de THI1 em mitocôndrias (Figura 10).

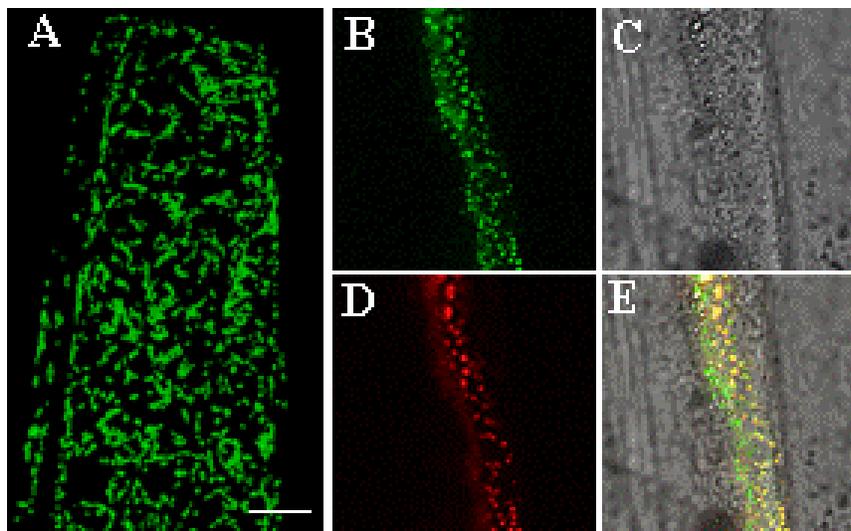


Figura 10 - Localização celular de THI1. (A e B) THI1 de arroz fusionada à GFP e direcionada para mitocôndrias e cloroplastos; (C) Imagem obtida com a luz visível; (D) fluorescência da RFP nas mitocôndrias; (E) Imagens B, C e D sobrepostas. Os cloroplastos seriam os pontos maiores verdes na sobreposição das imagens. A barra corresponde 50 μm

4.1.7 O duplo direcionamento como um processo conservado

Apesar de a evolução das famílias gênicas de membros com duplo direcionamento ter ocorrido independentemente em arroz e *Arabidopsis* pela duplicação de parálogos, os dados aqui apresentados mostram a conservação da estratégia de duplo direcionamento entre proteínas ortólogas dessas duas espécies. Além disso, mesmo tendo essas duplicações ocorrido independentemente, observa-se um número equivalente de parálogos em uma mesma família nas duas espécies estudadas. Isso sugere que essas duplicações não foram aleatórias e que a evolução das famílias com membros com duplo direcionamento segue um mesmo padrão.

A família das *phage-type* RNA polimerases é uma exceção. Em *Arabidopsis*, existem três membros que compõem essa família, enquanto em arroz há apenas dois. As proteínas de *Arabidopsis* são direcionadas para mitocôndrias e cloroplastos (At5g15700), cloroplastos (At2g24120) e mitocôndrias (At1g68990). Já em arroz, nenhum dos dois membros da família apresenta duplo direcionamento. A proteína codificada pelo gene no loco Os09g07120 é transportada para cloroplastos e a codificada pelo gene no loco Os06g44230, para mitocôndrias.

Neste trabalho, a conservação do mecanismo de duplo direcionamento entre mono e dicotiledôneas foi investigada e confirmada para as proteínas MDAR, MAP e THI1. Analisando-se a literatura sobre o assunto, observam-se outros casos de conservação do duplo direcionamento, a maioria entre dicotiledôneas.

O duplo direcionamento da *phage-type* RNA polimerase não foi conservado em arroz, mas foi em *Nicotiana tabacum* e em *Nicotiana glauca*. Nesta última, existem três isoformas da proteína: RpoT-A (Q93Y94), direcionada para mitocôndrias; RpoT-B (Q8VWF8), presente em mitocôndrias e cloroplastos e; RpoT-C (P69242), cloroplastidial (KOBAYASHI; DOKIYA; SUGITA, 2001; KOBAYASHI et al., 2002). As proteínas homólogas em *A. thaliana* (At1g68990, At5g15700 e At2g24120) assim como as de *N. tabacum* (RpoT1, RpoT2 e RpoT3), respectivamente, seguem o mesmo padrão de direcionamento (HEDTKE et al., 1999; HEDTKE; BORNER; WEIHE, 2000; HEDTKE et al., 2002). Em *N. glauca*, o mecanismo de duplo direcionamento é o de controle pós-transcricional por meio de múltiplos sítios de início de tradução. Em *A. thaliana*, a seqüência de direcionamento é ambígua.

Interessantemente, o duplo direcionamento dessas proteínas foi também relatado em uma briófito. Em *Physcomitrella patens* existem dois genes que codificam *phage-type* RNA polimerases. As duas proteínas têm duplo direcionamento para cloroplastos e mitocôndrias (RICHTER et al., 2002). Os autores, baseados em análises filogenéticas, sugerem que a duplicação gênica que deu origem a *phage-type* RNA Polimerases com duplo direcionamento em angiospermas ocorreu após a divergência entre mono e dicotiledôneas, uma vez que as isoformas de arroz e milho não apresentam duplo direcionamento.

Outro exemplo de conservação do duplo direcionamento ocorre entre as γ -type DNA polimerases. Em *A. thaliana*, existem dois membros que compõem essa família gênica: At3g20540 e At1g50840, cujos produtos são direcionados para mitocôndrias/cloroplastos e cloroplastos, respectivamente (ELO et al., 2003; CHRISTENSEN et al., 2005). Em *N. tabacum*, há também dois membros: *NtPoll-like1* (AB174897) e *NtPoll-like2* (AB174899). As seqüências de direcionamento dessas duas proteínas são idênticas e ambas têm duplo direcionamento (ONO et al., 2007), o que indica origem por duplicação recente. O produto de At3g20540 apresenta duplo direcionamento por um mecanismo de múltiplos inícios de tradução, enquanto as proteínas de *N. tabacum* apresentam seqüência de direcionamento ambígua.

A proteína glutationa redutase é um outro exemplo. Seu duplo direcionamento para mitocôndrias e cloroplastos foi primeiramente demonstrado em ervilha (CREISSEN et al., 1995; RUDHE et al., 2002b; CHEW et al., 2003) e confirmado em *A. thaliana* (CHEW; WHELAN; MILLAR, 2003). Nessas duas espécies, a seqüência de direcionamento é ambígua.

Outro caso é o da DNA girase A, cujo duplo direcionamento, também para mitocôndrias e cloroplastos, foi demonstrado em *A. thaliana* (WALL; MITCHENAL; MAXWELL, 2004) e em *N. benthamiana* (CHO et al., 2004). No entanto, o mecanismo de duplo direcionamento não foi conservado. A girase A de *A. thaliana* é transportada para as duas organelas por um mecanismo pós-traducional de múltiplos sítios de início da tradução, enquanto que a proteína de *N. benthamiana* possui uma seqüência de direcionamento ambígua.

O mesmo resultado não foi encontrado para a família das DNA girases B. Em *A. thaliana*, existem dois locos que codificam girases B, At3g10270 e At5g04130. Essas proteínas são direcionadas para cloroplastos e mitocôndrias, respectivamente (WALL; MITCHENAL; MAXWELL, 2004). Em *N. benthamiana*, há dois genes que codificam a proteína girase B, porém, ao contrário do encontrado em *A. thaliana*, uma delas (AAR07943) é direcionada para mitocôndrias e cloroplastos. Não há dados de localização celular para a segunda isoforma.

Os exemplos acima reforçam a idéia de que o duplo direcionamento é conservado entre espécies vegetais, porém, em muitos casos, via mecanismos diferentes. Essa observação torna-se ainda mais intrigante quando se observa que a seqüência de direcionamento presente na região amino-terminal da proteína não é conservada entre as espécies, apesar de os domínios funcionais de proteínas homólogas apresentarem alta similaridade. Uma explicação para isso é que a conservação do duplo direcionamento é funcional. As seqüências de direcionamento não teriam uma origem comum, contrariamente ao restante da proteína. Ao contrário, teriam se originado por eventos independentes e evoluído funcionalmente por convergência. Uma outra explicação para o fato é supor que a taxa de mutação difere entre a região funcional da proteína e sua seqüência de direcionamento. Esta última teria origem em um ancestral comum e seria mais tolerante a mutações, sem alterar, entretanto, sua função de endereçamento. Essa última hipótese é até plausível, quando se considera que não há motivos específicos na estrutura primária da proteína que determinam seu direcionamento para mitocôndrias e cloroplastos, por exemplo. No entanto, há estudos que revelaram que a seqüência de direcionamento não é muito tolerante a mutações. Alterações em um ou dois resíduos podem interferir em suas propriedades de endereçamento de proteínas (item 4.2) (PUJOL; MARÉCHAL-DROUARD; DUCHÊNE, 2007).

O fato de *Arabidopsis* e arroz possuírem genomas relativamente pequenos, quando comparados aos das demais espécies vegetais, levou à procura por homólogos de proteínas duplo direcionadas em *Zea mays*. O genoma do milho tem 2.365 Mb, expandiu rapidamente por eventos de retrotransposição e teve duas origens: por um evento de duplicação genômica, alotetraploidização e; pela duplicação de segmentos

gênicos (JORGENSEN, 2004). Embora exista quantidade razoável de seqüências do tipo EST (*expressed sequence tag*) em bancos de dados públicos, as análises não puderam ser realizadas completamente porque ainda faltam muitas informações. Nessa situação, o fato de não ser possível encontrar um homólogo em milho pode ser decorrente de sua inexistência no genoma ou então porque essa informação ainda não foi armazenada nos bancos. Esta incerteza certamente prejudicaria as análises. Apesar da dificuldade, ainda foi possível encontrar homólogos no genoma de milho e, em alguns casos, o número de isoformas encontrado foi similar ao observado para *Arabidopsis* e arroz (dados não mostrados). Isso sugere que o número de proteínas com duplo direcionamento não se expande em genomas maiores.

O entendimento das forças evolutivas que definem o tamanho das famílias gênicas ainda não está bem definido. Com base na compilação das informações conhecidas sobre as famílias gênicas contendo membros com duplo direcionamento, nos dados experimentais aqui obtidos e na dupla função que uma proteína pode exercer, o duplo direcionamento poderia ser considerado uma força evolutiva que controla o tamanho de famílias gênicas. A maior parte dos genes de *A. thaliana* que codificam proteínas com duplo direcionamento estão em regiões do genoma que não sofreram duplicações gênicas. O duplo direcionamento poderia ser uma restrição à duplicação gênica. Apesar do número restrito de famílias conhecidas com membros com duplo direcionamento representar uma limitação a conclusões, os resultados apresentados permitem sugerir uma tendência na evolução dessas famílias gênicas. Análises mais detalhadas das famílias de membros com duplo direcionamento e a caracterização de novas proteínas com duplo direcionamento se fazem necessária para testar a hipótese aqui formulada.

4.2 O duplo direcionamento de proteínas de ligação ao RNA (RBP)

A família das RBP mitocondriais de *A. thaliana* foi identificada por meio da purificação, via cromatografia de afinidade, de proteínas localizadas na mitocôndria e que apresentam capacidade de se ligar a ácidos nucleicos de fita única (VERMEL et al., 2002). Os membros dessa família são caracterizados pela presença de um motivo de reconhecimento ao RNA (RRM) na região amino-terminal e são estrutural e filogeneticamente relacionadas a proteínas de ligação ao RNA ricas em glicina. Ainda não se conhece a função dessas RBP mitocondriais, mas proteínas que contêm o motivo RRM estão envolvidas em processos pós-transcricionais e na regulação da expressão gênica.

As RBP de *A. thaliana* formam uma família gênica composta por 11 membros, todos direcionados para mitocôndria (Figura 11).

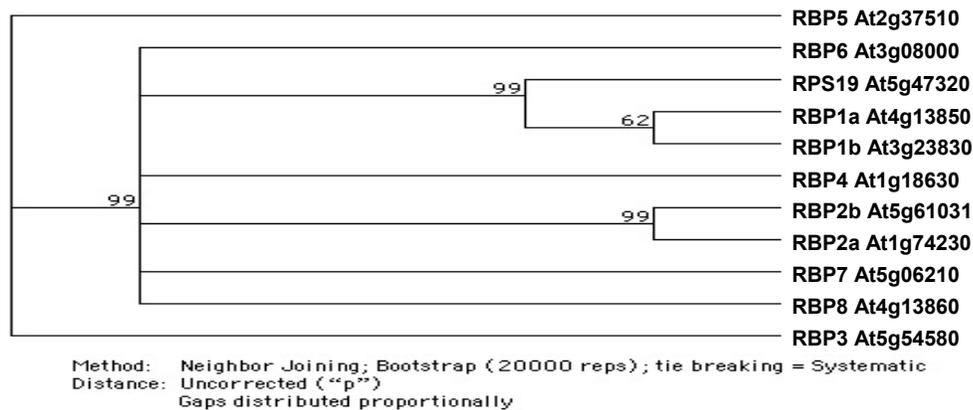


Figura 11 - Árvore filogenética dos membros da família das RBP de *A. thaliana*. A subfamília RBP1 é composta pelos membros RBP1a, RBP1b e RPS19 (VERMEL et al., 2002)

Gualberto¹ observou que um membro na subfamília RBP1, RBP1b, era direcionado tanto para mitocôndrias como para cloroplastos, enquanto os demais, RBP1a e RPS19 (proteína ribossomal S19) eram exclusivamente direcionados para mitocôndrias (informação verbal). Isso tornava essa subfamília um bom material para o estudo dos determinantes do duplo direcionamento, uma vez que as seqüências de

¹ GUALBERTO, J.M. *Université Louis Pasteur*, Estrasburgo, França.

direcionamento de seus membros são conservadas, provavelmente por se terem originado como resultado de duplicações recentes (Figura 12). Sendo assim, seria possível identificar as diferenças entre as seqüências e associá-las ao duplo direcionamento. A proposta foi realizar mutações em seqüências correspondentes a aminoácidos não conservados de RBP1b de forma a abolir o direcionamento para uma das duas organelas.

RBP1b	M	A	F	C	N	K	L	S	G	I	L	R	Q	G	V	S	Q	S	S	N	G	P	V	T	S	M	L	G	S	L	R	V	M	S	S	K	L
RBP1a	M	A	F	C	N	K	L	G	G	L	L	R	Q	N	I	S	S	N	G	N	V	P	V	T	S	M	L	G	S	L	R	L	M	S	T	K	L
RPS19	M	A	F	C	T	K	L	G	G	H	W	K	Q	G	V	-	-	-	N	V	P	V	S	S	M	L	G	S	L	R	V	M	S	T	K	L	

Figura 12 - Alinhamento das seqüências de direcionamento dos membros da subfamília RBP1 (1a, 1b e RPS19)

4.2.1 Localização celular dos membros da subfamília RBP1

A primeira etapa do trabalho foi a de confirmação da localização celular dos três membros da subfamília RBP1. Inicialmente, em experimentos de expressão transiente, foram utilizadas construções contendo a seqüência de direcionamento das proteínas, incluindo os primeiros 152 pb da seqüência do cDNA, fusionada *in frame* à seqüência da GFP. Foi realizado o bombardeamento de folhas de *N. benthamiana* e o material foi visualizado ao microscópio confocal. Nos experimentos controle, as transformações foram realizadas com as construções pUCAP-RecA (gentilmente cedida por J.M.Gualberto, *Université Louis Pasteur*, Estrasburgo, França) e pPMA4-GFP (DUBY; DEGAND; BOUTRY, 2001), que comprovadamente são direcionadas para cloroplastos e mitocôndrias, respectivamente (Figura 13). Nos experimentos em que as construções continham a seqüência de direcionamento das proteínas RBP1a, RBP1b e RPS19, inesperadamente, formaram-se agregados, estruturas no citosol maiores que mitocôndrias e estáticas, contrastando com a movimentação dessas organelas (Figura 14). Não foi possível, assim, determinar e a localização celular das proteínas RBP1a, RBP1b e RPS19.

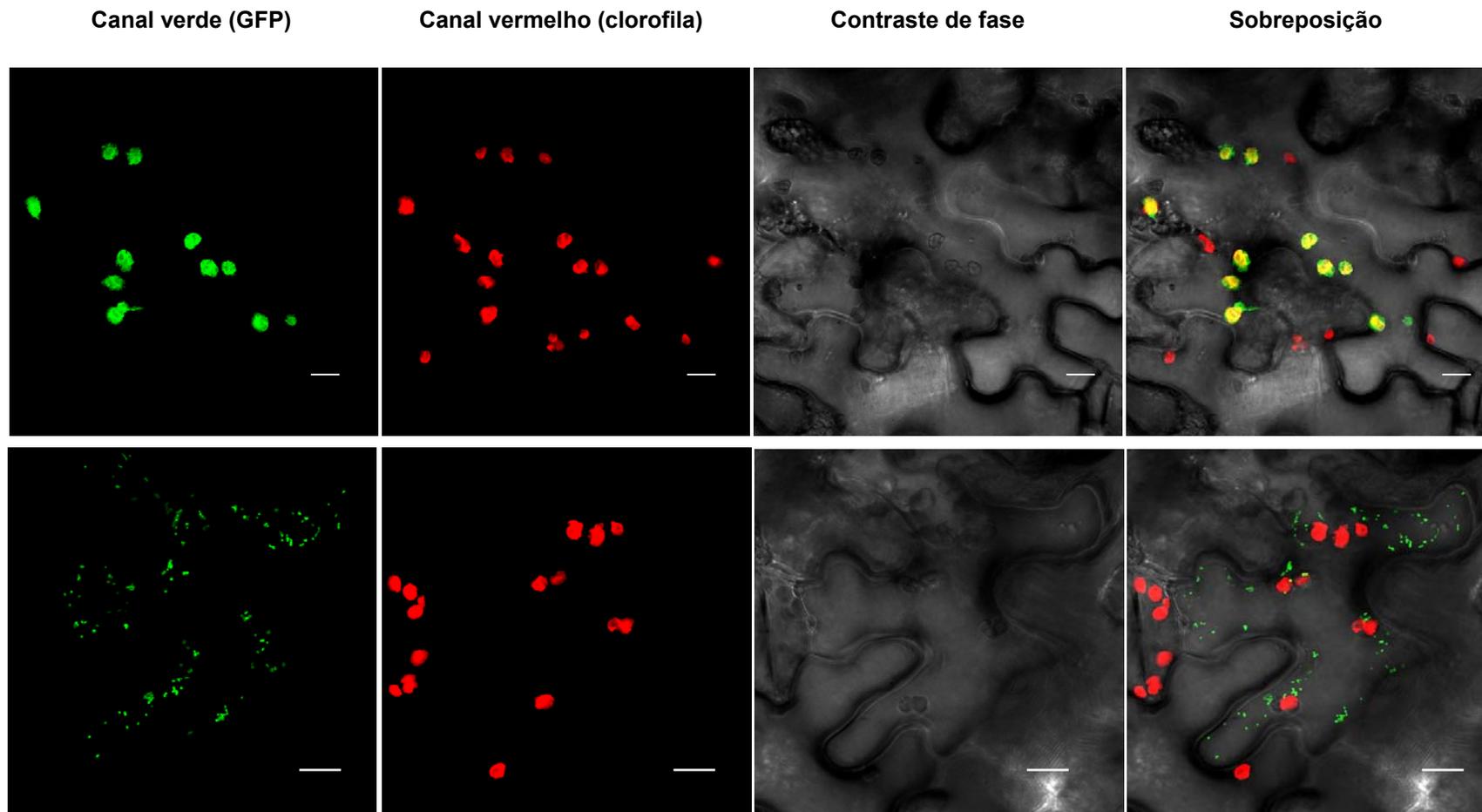


Figura 13 - Células do mesófilo de *N.benthamiana* transformadas com as construções controle pUCAP-RecA, direcionada para cloroplasto (acima) e, pPMA4-GFP, direcionada para mitocôndrias (abaixo). A barra representa 10 μm

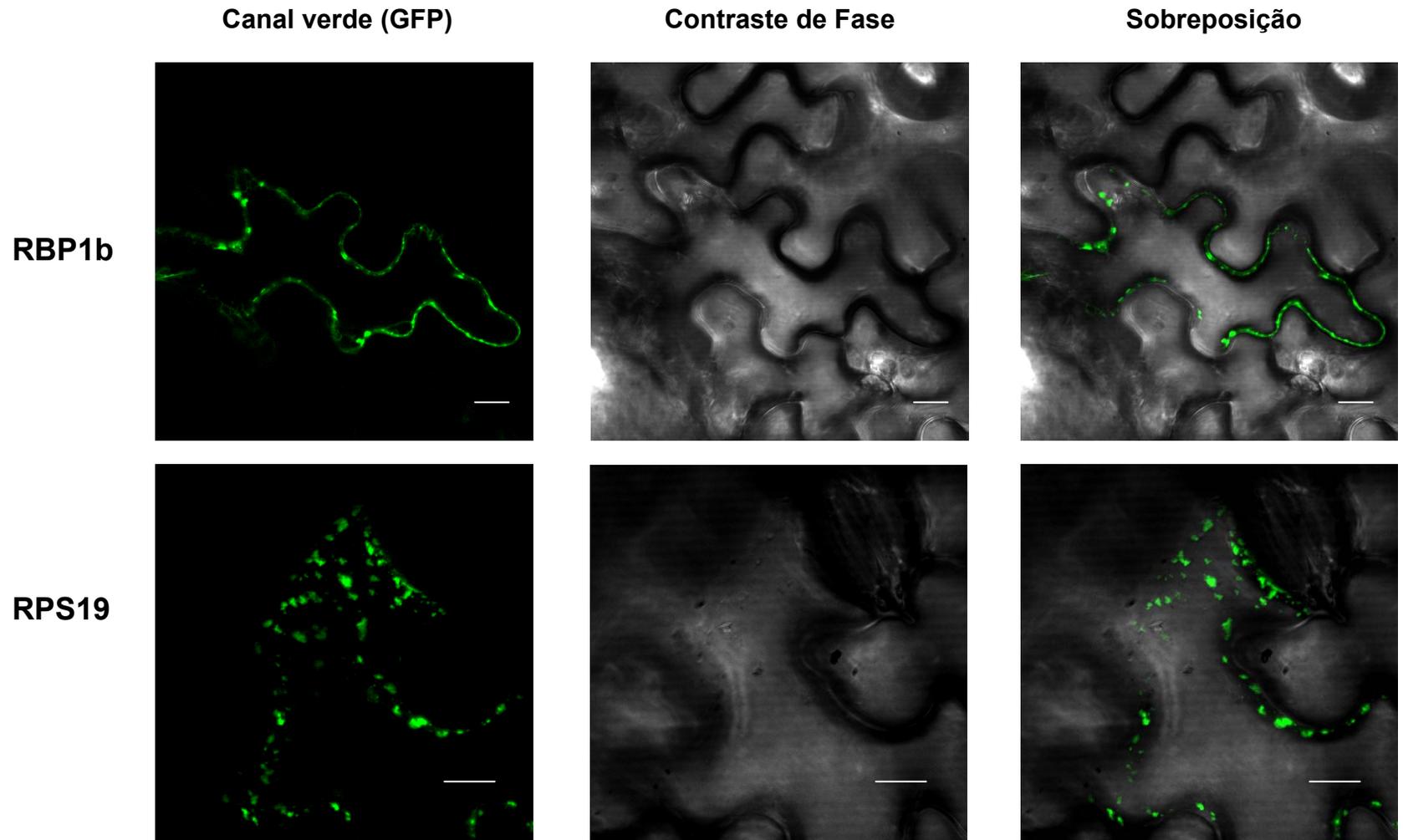


Figura 14 - Células do mesófilo de *N. benthamiana* transformadas com a seqüência de direcionamento de RBP1b e RPS19 fusionadas à GFP. Observou-se a formação de agregados. Os mesmos resultados foram obtidos para a RBP1a (dado não mostrado). A barra representa 10 μm

Diante desses resultados, a alternativa encontrada foi utilizar não somente a seqüência de direcionamento das proteínas, mas sua seqüência completa. Foram realizadas novas construções que continham 450, 408 e 341 pb da seqüência do cDNA de RBP1a, RBP1b e RPS19, respectivamente. Essas construções foram utilizadas para a transformação transiente por meio do bombardeamento de folhas de *N. benthamina* e o material foi analisado ao microscópio confocal. O problema da formação de agregados foi solucionado e determinou-se a localização celular das três proteínas (Figura 15 e 16). Ao contrário dos resultados prévios, foi constatado que as três proteínas, RBP1a, RBP1b e RPS19 possuem duplo direcionamento para mitocôndrias e cloroplastos. O duplo direcionamento de RBP1b foi também confirmado pela transformação transiente de *N. benthamiana* via *A. tumefaciens* seguida do preparo de protoplastos (Figura 17). A confirmação do direcionamento de RBP1a e RBP1b para cloroplastos pela técnica de importação *in vitro* para organelas não foi possível em decorrência de problemas metodológicos.

4.2.2 Mutações na seqüência de RBP1b

O trabalho foi então continuado com outro enfoque: mutar especificamente aminoácidos da seqüência de direcionamento de RBP1b e analisar seus efeitos no endereçamento da proteína para as organelas. Para a escolha dos aminoácidos a serem mutados, buscaram-se características compartilhadas por seqüências de direcionamento ambíguas de proteínas endereçadas para mitocôndrias e cloroplastos. Os primeiros 80 aminoácidos de cada proteína (Anexo A) foram alinhados, utilizando-se o programa Weblogo (CROOKS et al., 2004), o qual fornece informações sobre a freqüência dos aminoácidos em cada posição (Figura 18). Conforme documentado por Pujol; Maréchal-Drouard; Duchêne (2007), observa-se, principalmente, a predominância de serinas ao longo da seqüência. A presença de leucina (L) e arginina (R) também chama atenção, assim como a conservação da alanina (A) na posição 2.

Diante dessas informações, para a mudança de aminoácidos da seqüência de RBP1b, foram escolhidos resíduos positivos (lisinas e argininas), valina e serina. Trabalhos preliminares mostraram a importância de resíduos positivos na região amino-terminal da seqüência da proteína para o direcionamento para mitocôndrias e que

seqüências de direcionamento de cloroplastos são ricas em serina (VON HEIJNE, 1989; CHEW et al., 2003).

Resultados anteriores mostram que uma mutação de ponto na seqüência de direcionamento de uma proteína raramente abole totalmente seu transporte para uma organela. O que geralmente acontece é uma redução ou aumento no direcionamento da proteína para uma ou outra organela (CHEW et al., 2003; PUJOL; MARÉCHAL-DROUARD; DUCHÊNE, 2007). Sendo assim, a proposta do trabalho, de observar os efeitos de mutações na seqüência de direcionamento de RBP1b, necessitava de um sistema quantitativo para análise. As metodologias mais utilizadas são o transporte *in vitro* de proteínas para mitocôndrias e cloroplastos e *Western blotting*. No entanto, essas metodologias têm algumas limitações. O direcionamento inespecífico é comum em experimentos de importação *in vitro* (CLEARY et al., 2002; SILVA-FILHO et al., 1997), além de se tratar de metodologia bastante trabalhosa, o que dificulta quantidade razoável de repetições. No caso da técnica de *Western blotting*, a principal limitação é a necessidade de desenvolvimento de anticorpos com alta especificidade. Quando se trabalha com uma subfamília como a RBP1, em que seus membros provavelmente originaram-se por duplicações gênicas recentes e seus parálogos apresentam seqüências extremamente conservadas, o problema é agravado, pois o anticorpo deve ter especificidade que impeça o reconhecimento também das outras isoformas.

Dessa forma, foi testado um sistema de medida do sinal da GFP para a quantificação do transporte de proteínas para cloroplastos e mitocôndrias. As imagens obtidas por microscopia confocal foram analisadas no programa ImageJ (COLLINS, 2007). Nesse sistema, o sinal da GFP é lido em pixels e transformado em números. A principal vantagem dessa metodologia é a possibilidade de um grande número de repetições para uma boa análise estatística.

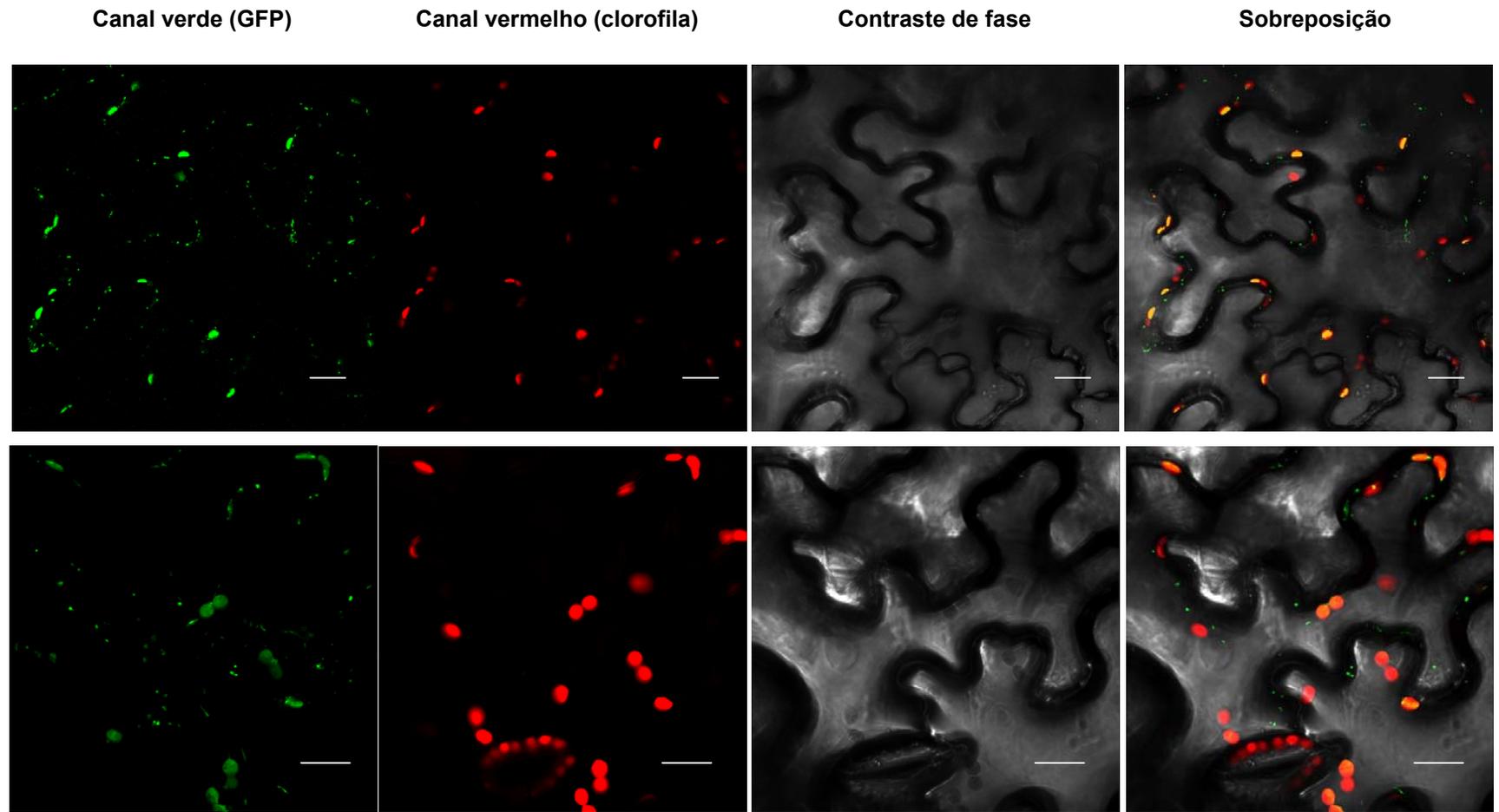


Figura 15 – Células do mesófilo de *N.benthamiana* transformadas com a construção controle de duplo direcionamento, CySRS (DUCHÊNE et al., 2005) (acima) e com a construção pUCAP-RBP1a (abaixo), onde o sinal da GFP pode ser observado em mitocôndrias e cloroplastos. A barra representa 20 μ m

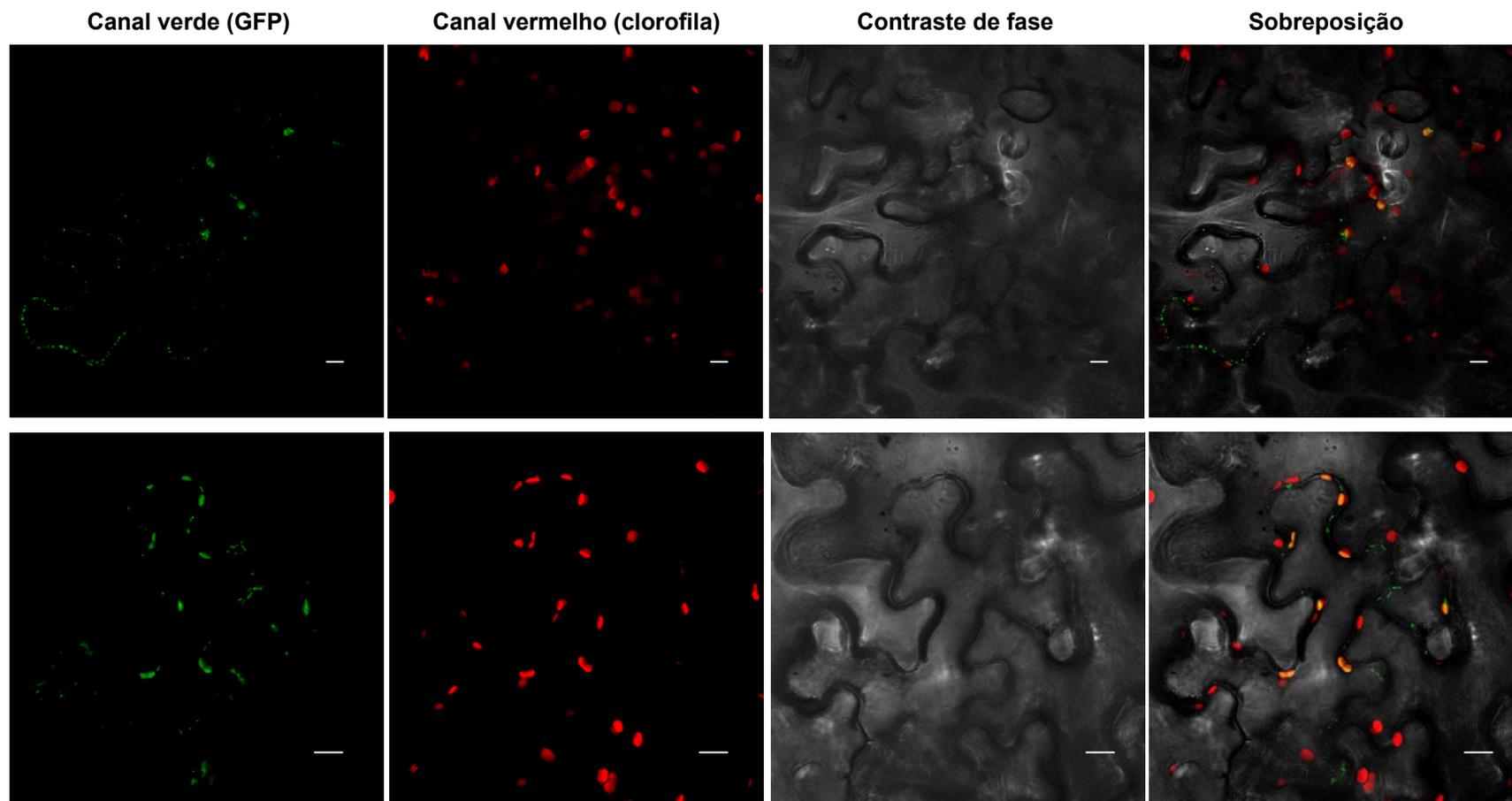


Figura 16 - Duplo direcionamento das proteínas RBP1b e RPS19 para mitocôndrias e cloroplastos. Células do mesófilo de *N.benthamiana* transformadas com as construções: pUCAP-RBP1b (acima) e; pUCAP-RPS19 (abaixo). A barra representa 10 μ m

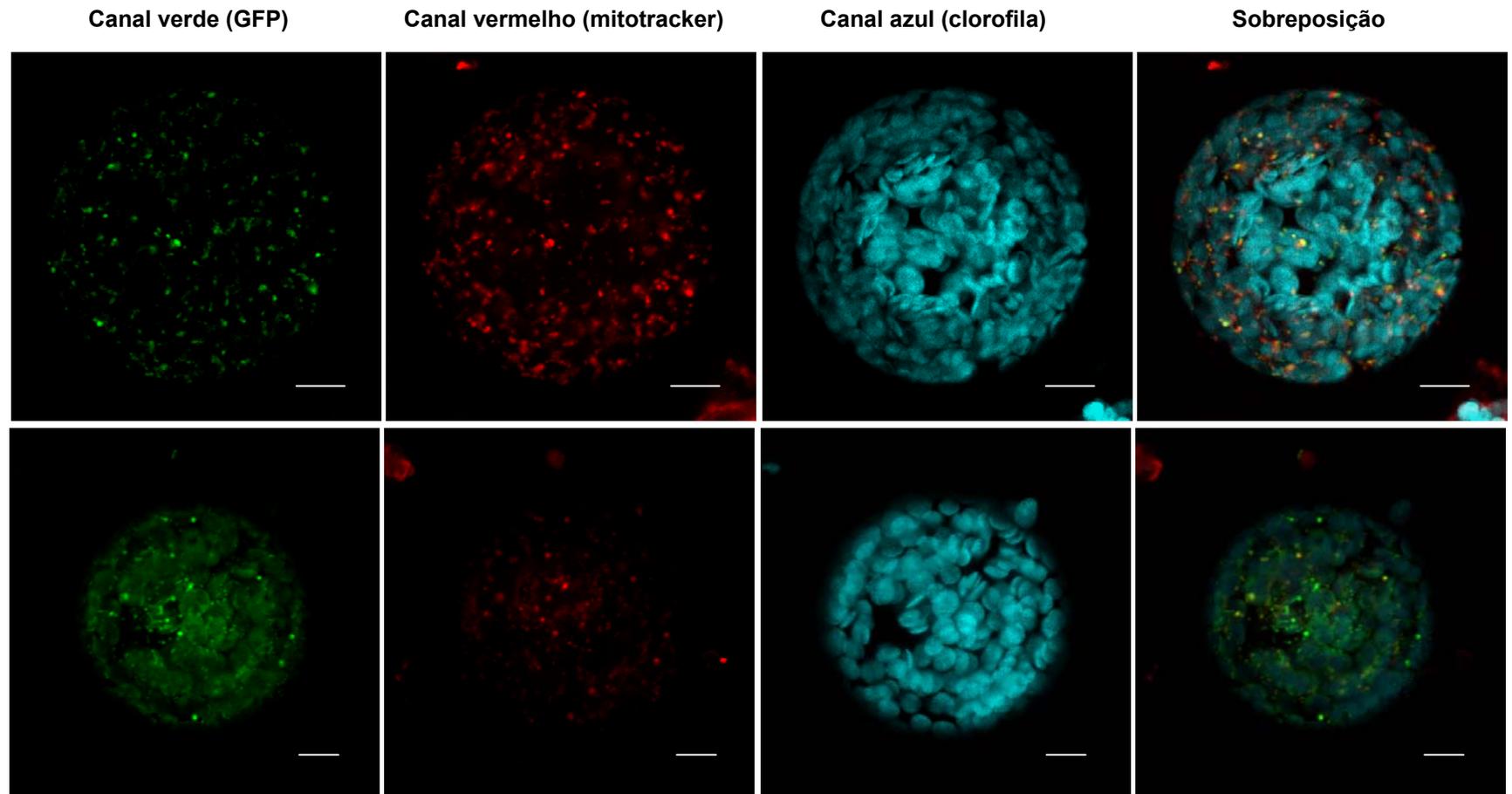


Figura 17 – Duplo direcionamento da proteína RBP1b demonstrado em protoplastos de células do mesófilo de *N. benthamiana*. Acima, protoplasto transformado com a construção controle mitocondrial, pNP5CoxVI (gentilmente doada por A.M. Duchêne, Université Louis Pasteur, Estrasburgo, França). Abaixo, protoplasto transformado com a construção pBIN-RBP1b. Observa-se a fluorescência da GFP em cloroplastos e mitocôndrias. A barra representa 10 μm

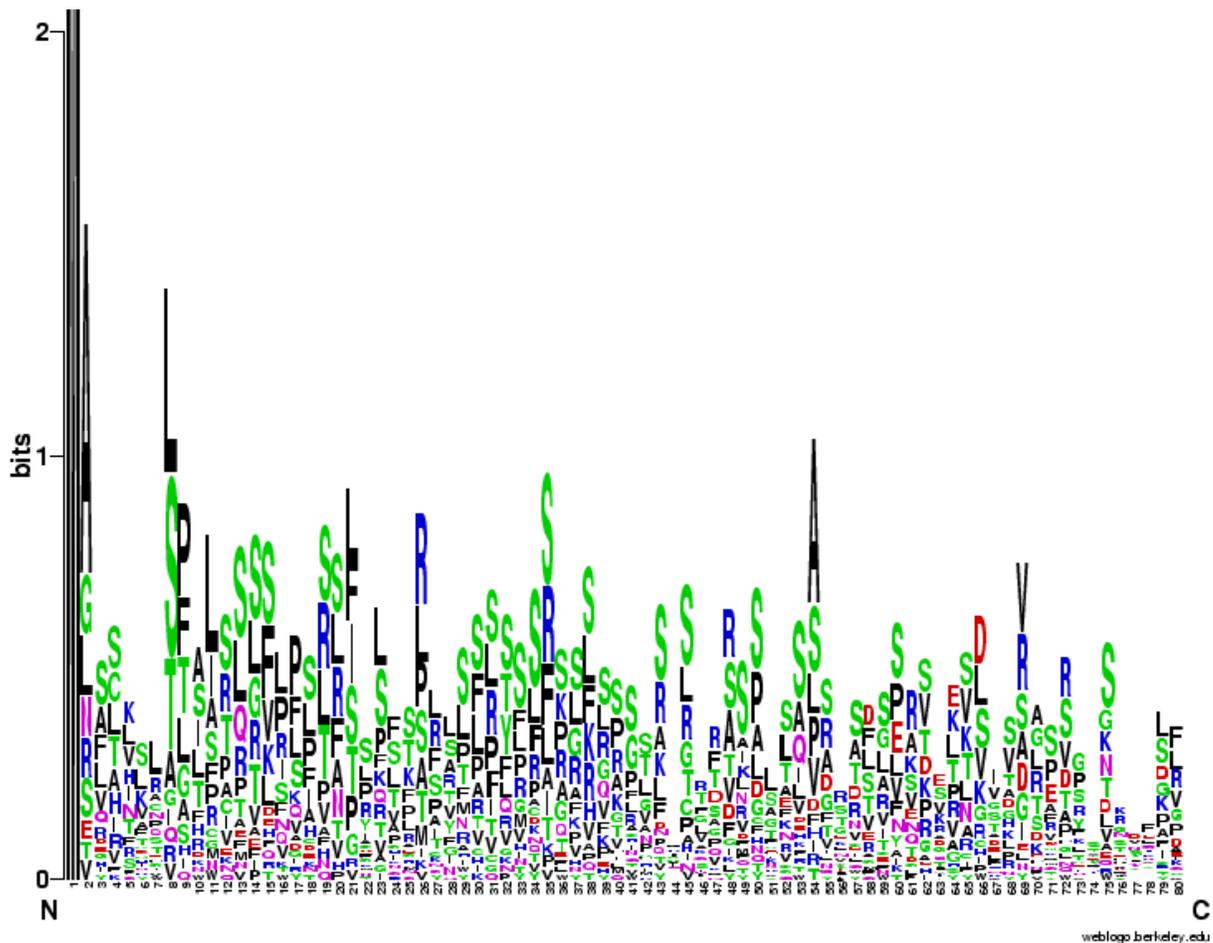


Figura 18 - Representação gráfica gerada pelo programa Weblogo do alinhamento das seqüências de direcionamento ambíguas de proteínas endereçadas para mitocôndrias e cloroplastos. A altura da pilha de letras indica a conservação da seqüência em cada posição. A altura das letras indica a freqüência relativa dos aminoácidos em cada posição

Para o estudo dos determinantes do duplo direcionamento da RBP1b, foram realizadas construções, no vetor pUCAP-GFP, contendo a seqüência completa da proteína com mutações de ponto ou deleções em sua seqüência de direcionamento. Estas estão indicadas abaixo:

- RBP1b sem seqüência de direcionamento, em que os primeiros 39 aminoácidos foram deletados, exceto a primeira metionina;
- $\Delta(2-17)$, em que os 17 primeiros aminoácidos foram deletados, exceto a primeira metionina;

- **Δ(18-30)**, em que a segunda parte da seqüência de direcionamento, do aminoácido 18 ao 30, foi deletada;
- **A2N**, em que o aminoácido na posição 2, uma alanina, foi substituído por uma asparagina;
- **K6Q**, em que o aminoácido na posição 6, uma lisina, foi substituído por uma glutamina;
- **R12N**, em que o aminoácido na posição 12, uma arginina, foi substituído por uma asparagina;
- **K6QR12N**, em que foram substituídos os aminoácidos nas posições 6 e 12, por glutamina e asparagina, respectivamente;
- **V15NS16G**, em que foram substituídos os aminoácidos nas posições 15 (valina) e 16 (serina), por asparagina e glicina, respectivamente;
- **S18NS19G**, em que foram substituídos os aminoácidos 18 e 19 (serinas), por asparagina e glicina, respectivamente;
- **Δ(18-19)**, em que os aminoácidos nas posições 18 e 19 (duas serinas) foram deletados;
- **R31Q**, em que o aminoácido na posição 31, arginina, foi substituído por uma glutamina.

Foram também averiguadas as diferenças nas proporções de direcionamento para mitocôndrias e cloroplastos das proteínas selvagens RBP1a, RBP1b e RPS19.

4.2.3 Análise estatística dos dados gerados pelo programa ImageJ

Os dados gerados pelo programa ImageJ e utilizados nas análises estatísticas estão organizados no Anexo B. A Tabela 7 mostra os elementos de estatística descritiva da amostra (proporção relativa x de proteínas localizadas nas mitocôndrias). A quantidade complementar, $y = 1-x$, é a proporção de proteínas localizadas nos cloroplastos. Na tabela correspondente ao parâmetro y , a média e a mediana são exatamente complementares, o desvio padrão (s.d.) e o erro padrão da média (s.e.) permanecem inalterados e os limites inferior e superior observados (V.MÍN. e V.MÁX.) e os limites superiores dos quartis primeiro e terceiro (Q1 e Q3) correspondem às

quantidades complementares dos limites não respectivos. Assim, os elementos y correspondentes ao parâmetro WT1b serão respectivamente: média (y) = 1 – média (x) = 0,4048, mediana (y) = 1 – mediana (x) = 0,4143, S.D.(y) = S.D.(x) = 0,0719, S.E.(y) = S.E.(x) = 0,0123, V.MÍN.(y) = 1 - V.MÁX.(x) = 0,2810, V.MÁX.(y) = 1 - V.MÍN.(x) = 0,5323, Q1(y) = 1 - Q3(x) = 0,3321 e Q3(y) = 1 - Q1(x) = 0,4609.

Tabela 7 - Elementos de estatística descritiva das amostras analisadas

	N	MÉDIA	MEDIANA	S.D.	S.E.	V.MÍN.	V.MÁX	Q1	Q3
WT1b	34	0,5952	0,5857	0,0719	0,0123	0,4677	0,7190	0,5391	0,6679
WT1a	22	0,3131	0,3019	0,0880	0,0188	0,1925	0,6331	0,2755	0,3595
RPS19	29	0,4056	0,4071	0,0696	0,0129	0,2539	0,5579	0,3574	0,4592
A2N	15	0,5852	0,5759	0,0703	0,0182	0,4725	0,7255	0,5342	0,6545
K6Q	47	0,3279	0,3340	0,1037	0,0151	0,1365	0,6080	0,2405	0,3945
R12N	32	0,3040	0,3047	0,0934	0,0165	0,1648	0,5116	0,2293	0,3624
V15NS16G	42	0,4342	0,3877	0,2031	0,0313	0,1490	0,8478	0,2753	0,5383
S18NS19G	25	0,3765	0,3909	0,0888	0,0178	0,2381	0,5514	0,3108	0,4393
D1819	32	0,4350	0,4128	0,1093	0,0193	0,2656	0,6668	0,3433	0,5445
R31Q	15	0,3209	0,3095	0,0696	0,0180	0,2261	0,4500	0,2728	0,3790

Nota: Abreviações, N, número de observações; S.D., desvio padrão; S.E., erro padrão; V.MÍN, limite inferior observado; V.MÁX, limite superior observado; Q1, limite superior do quartil primeiro; Q3, limite superior do quartil terceiro.

A análise das distribuições desses dez parâmetros (variáveis sem transformação e transformadas) mostrou que:

a) duas das variáveis não transformadas (WT1a e V15NS16G) possuem distribuições muito discrepantes da normal, com valores de teste de Anderson-Darling 1,280 e 1,045, correspondentes a probabilidades de 0,002 e 0,009, favorecendo a hipótese de normalidade. O valor correspondente à variável WT1a é atribuível à ocorrência de um valor anômalo ou *outlier*,

b) a transformação $\text{sen}^{-1} \sqrt{x}$, que seria a transformação de eleição para a variável estudada (proporções), não melhora significativamente a aderência dos valores das distribuições a normais correspondentes;

c) a transformação \sqrt{x} , no entanto (apesar de, *a priori*, não ter indicação formal para ser aplicada a proporções), melhora a aderência das distribuições a normais de

mesma média e desvio. No entanto, a premissa de homogeneidade de variâncias é violada nos três casos (valores respectivos de qui-quadrado com nove graus de liberdade) correspondentes a testes de Bartlett de 79,06, 82,02 e 76,77 para variáveis não transformadas, com transformação arco-seno e com transformação raiz quadrada, respectivamente. Esses valores muito elevados do teste são provavelmente atribuíveis à contribuição da variável V15NS16G, que possui um nível de dispersão ou variabilidade dos valores de ordem de grandeza muito superior ao dos demais grupos. Com a retirada desse grupo, o valor do qui-quadrado do teste de Bartlett cai abruptamente para 26,19 (oito graus de liberdade), apesar de indicar ainda um resíduo significativo de heterogeneidade de variâncias entre os grupos restantes.

Devido ao fato de que com a transformação raiz quadrada as distribuições se normalizaram razoavelmente (apesar de violada a condição de homogeneidade de variâncias entre os grupos), numa primeira aproximação, os dados assim transformados foram submetidos a um modelo unifatorial de análise de variância, conforme o resumo abaixo.

A análise prévia dos dados não transformados mostrou que a variação média entre as médias dos grupos (WT1b, ..., R31Q) é 19,92 vezes maior que a variação intra-grupos. Admitindo-se que não existem diferenças entre as médias verdadeiras, a chance de ocorrência de uma razão no mínimo tão grande como essa é menor que 1%. Isso constitui evidência de que existem diferenças entre as médias dos grupos. Uma análise de máxima verossimilhança mostrou que a transformação raiz quadrada seria apropriada para normalizar os dados. Tomou-se o cuidado de retirar da amostra dados classificados como *outliers* por apresentarem valores residuais studentizados correspondentes a probabilidades menores que 0,05. O modelo de análise de variância com as variáveis transformadas e com a correção para a desigualdade entre as variâncias dos grupos forneceu o valor de $F(9;48) = 40,29$ ($P < 0,0001$), concluindo-se, portanto, que a variação entre as médias dos grupos é 40,292 vezes a variação média dentro dos grupos. Testes de comparações múltiplas permitiram a construção dos intervalos aproximados de credibilidade a 95% mostrados na Tabela 8.

Tabela 8 - Intervalos de confiança a 95% construídos com base em testes de comparações múltiplas

Grupo	S.D.	Intervalo de Confiança
WT1a	0,0461	0,508 - 0,615
WT1b	0,0424	0,747 - 0,806
RPS19	0,0521	0,594 - 0,670
A2N	0,0417	0,727 - 0,812
K6Q	0,0732	0,541 - 0,597
R12N	0,0705	0,509 - 0,578
V15NS16G	0,1344	0,593 - 0,660
S18NS19G	0,0574	0,581 - 0,654
D1819	0,0597	0,620 - 0,689
R31Q	0,0551	0,518 - 0,618

Conclui-se, portanto, que existem dois conjuntos distintos de médias: (WT1b e A2N) e (WT1a, RPS19, K6Q, R12N, V15NS16G, S18NS19G, D1819, R31Q). Isso indica de maneira direta que o grupo mutante A2N não difere de WT1b, ao contrário dos demais grupos mutantes (D1819, R31Q, K6Q, R12N, S18NS19G, V15NS16G).

Essa análise preliminar já permitiu esclarecer alguns pontos levantados no estudo:

a) existem diferenças no direcionamento intracelular das proteínas para as mitocôndrias e os cloroplastos entre o tipo selvagem WT1b e os mutantes produzidos (A2N, D1819, R31Q, K6Q, R12N, S18NS19G, V15NS16G)?

b) existem diferenças entre os três grupos selvagens (WT1b, WT1a, RPS19)?

c) é válida a utilização de dados obtidos a partir de dois ou mais planos de uma mesma célula, obtidos por meio da microscopia confocal, ou os dados assim obtidos são de certa maneira enviesados?

d) existem diferenças dependentes do tipo celular em que foi inferida a proporção de proteína (mesófilo ou estômato)?

A questão também importante da variabilidade em certos grupos de mutantes (em especial o V15NS16G) ser maior que em outros já foi comentada anteriormente, diante dos indícios de que a variância do grupo V15NS16G era maior que as dos demais grupos. Para resolver quantitativamente esse problema, realizou-se um teste de

comparações múltiplas entre as variâncias dos grupo, listadas abaixo em ordem crescente de variância (Tabela 9)

Tabela 9 - Variância, organizada em ordem crescente, dos grupos analisados

Grupo	Variância	N (total)
RPS19	0,0048	29
R31Q	0,0048	15
A2N	0,0049	15
WT1b	0,0052	34
WT1a	0,0077	22
S18NS19G	0,0079	25
R12N	0,0087	32
K6Q	0,0108	47
D1819	0,0119	32
V15NS16G	0,0412	42

Um teste de Bartlett aplicado a essas quantidades para verificar a presença de heterogeneidade forneceu um valor de $\chi^2 = 79,06$ (g.l. = 9; $P < 0,001$), descartando-se, portanto, a hipótese nula de homogeneidade de variâncias. Foi realizado um teste de comparações múltiplas, baseado na distribuição studentizada q, capaz de comparar todos os pares possíveis de valores de variância. A Tabela 10 mostra os resultados dos testes que forneceram valores de q superiores ao valor crítico $q(cr) = 4,474$, aplicável ao caso de 10 grupos de comparação (total de $9 \times 10 / 2 = 45$ testes) e g.l. = ∞ , tirado de tabelas da distribuição q.

Tabela 10 – Resultados do teste de comparação múltipla dos grupos analisados (são mostrados apenas os valores significativos)

Comparação	Dif*	s.e.	q	p
RPS19 vs V15NS16G	2,1419	0,245	8,737	10
R31Q vs V15NS16G	2,1419	0,310	6,919	9
A2N vs V15NS16G	2,1219	0,310	6,855	8
WT1b vs V15NS16G	2,0768	0,234	8,880	7
WT1a vs V15NS16G	1,6727	0,268	6,233	6
S18NS19G vs V15NS16G	1,6546	0,257	6,438	5
R12N vs V15NS16G	1,5536	0,238	6,528	4
K6Q vs V15NS16G	1,3444	0,215	6,259	3
D1819 vs V15NS16G	1,2392	0,238	5,207	2

Conclui-se que apenas a variável V15NS16G possui variabilidade significativamente superior à de todas as demais variáveis, as quais possuem, por sua vez, variâncias semelhantes.

Uma vez que a variável proporção relativa de proteínas localizadas numa determinada organela (mitocôndria ou cloroplasto) é obtida da aplicação de dados numéricos (área A e intensidade I obtidos por aparelhagem de leitura) e manipulados algebricamente segundo, $\text{prop}(\text{mit}) = \frac{A(\text{mit}) \times I(\text{mit})}{[A(\text{mit}) \times I(\text{mit}) + A(\text{clo}) \times I(\text{clo})]}$ e $\text{prop}(\text{clo}) = 1 - \text{prop}(\text{mit})$, a distribuição dessa variável segue a distribuição indeterminada de Cauchy, justificando-se o uso de métodos de estatística não-paramétrica aos dados.

Inicialmente, verificou-se se os dados devem ser analisados levando-se em conta ou não os parâmetros obtidos por microscopia confocal (repetições dentro de uma mesma célula em planos diferentes). Para isso, foram comparados os dados de cada grupo com todas as informações e com todas as leituras confocais retiradas, com exceção da primeira. Não foram retiradas todas as leituras confocais porque isso, que seria estatisticamente mais apropriado, por não incluir elementos comuns nos dois grupos de comparação, iria reduzir drasticamente o tamanho amostral da maioria dos grupos, tornando desnecessários quaisquer testes estatísticos devido ao seu fraco poder *a priori*. Os resultados obtidos com a aplicação de testes de Mann-Whitney estão

sumarizados na Tabela 11, mostrando que as percentagens medianas obtidas por meio de ambos os métodos são idênticas, o que valida, portanto, a utilização de parâmetros inferidos a partir da observação de vários planos de uma mesma célula.

Tabela 11 – Comparação entre os dados das análises em microscopia confocal, incluindo vários planos ou um plano por célula

Variável	U	U'	$\Sigma r(\text{cpc})$	$\Sigma r(\text{spc})$	P
WT1b	335,50	446,50	930,50	772,50	0,3709 n.s.
WT1a	69,50	84,50	337,50	97,50	0,7086 n.s.
RPS19	189,00	217,00	652,00	294,00	0,7264 n.s.
A2N	81,50	83,50	201,50	149,50	0,9594 n.s.
K6Q	580,00	642,00	1708,00	993,00	0,7253 n.s.
R12N	269,50	274,50	797,50	427,50	0,9665 n.s.
V15NS16G	394,50	403,50	1306,50	584,50	0,9503 n.s.
S18NS19G	184,50	190,50	509,50	310,50	0,9443 n.s.
D1819	251,50	292,50	820,50	404,50	0,6744 n.s.
R31Q	51,00	69,00	189,00	87,00	0,5910 n.s.

Nota: Abreviações, U e U', estatística de Mann-Whitney; cpc, vários planos por célula; spc, um único plano por célula; $\Sigma r(\text{cpc})$ e $\Sigma r(\text{spc})$, soma dos postos dos dados cpc e spc.

Devido a esses resultados, as análises que se seguem foram realizadas utilizando-se os dados completos relativos a cada variável, que incluem parâmetros estimados pelo exame de células em vários planos confocais. Para verificar se existem diferenças no direcionamento intracelular das proteínas para as mitocôndrias e os cloroplastos entre o tipo selvagem WT1b e os mutantes (A2N, D1819, R31Q, K6Q, R12N, S18NS19G, V15NS16G) e entre os três grupos selvagens (WT1b, WT1a, RPS19), utilizou-se o teste não paramétrico de Kruskal-Wallis, que é o correspondente não-paramétrico à análise de variância unifatorial tradicional da estatística paramétrica.

Os resultados referentes à verificação da existência de eventuais diferenças entre o tipo selvagem WT1b e os mutantes estão sumarizados na Tabela 12.

Tabela 12- Análise do direcionamento das proteínas WT1b e mutantes para mitocôndrias, aplicando-se o teste de Kruskal-Wallis

Grupo	n	Σr	m(r)	z
WT1b	34	6917	203,44	7,36
A2N	15	2997	199,80	4,47
K6Q	47	3872	82,38	-4,27
R12N	32	2221	69,41	-4,52
V15NS16G	42	5145	122,50	0,10
S18NS19G	25	2735	109,40	-0,91
D1819	32	4346	135,81	1,24
R31Q	15	1170	78	-2,48
TOTAL	242		121,50	

Nota: Abreviações, n, número de pontos; Σr , soma dos postos; m(r), média dos postos; z, desvio reduzido normal.

O valor do teste de Kruskal-Wallis $H = \chi^2 = 105,63$ (7 graus de liberdade), corresponde a um valor de probabilidade $P < 0.0001$. Para localizar as diferenças, usou-se o teste de comparações múltiplas de Dunn, que leva em conta, apenas as comparações desejadas, no caso, um controle contrastado com sete grupos mutantes diferentes (Tabela 13). Como esse teste emprega aproximações normais, os valores de probabilidade correspondentes a cada comparação são exatos apenas para amostras muito grandes.

Tabela 13 – Resultados da aplicação do teste de comparações múltiplas de Dunn, entre WT1b e seus mutantes, no direcionamento para mitocôndrias

Comparação (i vs. j)	$m_i(r) - m_j(r)$	P
WT1b vs. A2N	3,641	>0,05 n.s.
WT1b vs. K6Q	121,06	<0,001 ***
WT1b vs. R12N	134,03	<0,001 ***
WT1b vs. V15NS16G	80,941	<0,001 ***
WT1b vs. S18NS19G	94,041	<0,001 ***
WT1b vs. D1819	67,629	<0,01 **
WT1b vs. R31Q	125,44	<0,001 ***

As conclusões, portanto, não diferem das obtidas com o modelo de análise de variância unifatorial aplicado ao conjunto de dados transformados. O grupo mutante A2N não difere de WT1b, ao contrário dos demais grupos mutantes (D1819, R31Q, K6Q, R12N, S18NS19G, V15NS16G).

Em seguida, foram realizadas análises que permitem detectar eventuais diferenças entre os três tipos selvagens WT1b, WT1a e RPS19. Os resultados encontram-se sumarizados na Tabela 14, na qual foram usadas as mesmas abreviaturas das duas tabelas imediatamente anteriores.

Tabela 14 - Análise do direcionamento das proteínas da subfamília RPB1 para mitocôndrias, aplicando-se o teste de Kruskal-Wallis

Grupo	n	Σr	m(r)	z
WT1b	34	2285	67,21	7,38
WT1a	22	375	17,05	-5,73
RPS19	29	995	34,31	-2,34
TOTAL	85		43	

O valor do teste de Kruskal-Wallis $H = \chi^2 = 60,63$ (2 graus de liberdade) corresponde a um valor de probabilidade $P < 0,0001$. Para localizar as diferenças, usou-se, como no caso anterior, o teste de comparações múltiplas de Dunn. Os resultados indicaram que as medianas dos três grupos selvagens são diferentes (Tabela 15). A comparação WT1a vs. RPS19 revelou resultado fracamente significativo ao nível de 5%, mas não ao nível de 1%. Levando-se em conta o número de testes realizados, seria prudente não considerar esse resultado significativo.

Tabela 15 - Resultados da aplicação do teste de comparações múltiplas de Dunn, entre proteínas da subfamília RBP1 WT1b, no direcionamento para mitocôndrias

Comparação (i vs. j)	$m_i(r) - m_j(r)$	P
WT1b vs. WT1a	50,160	<0,001 ***
WT1b vs. RPS19	32,896	<0,001 ***
WT1a vs. RPS19	-17,265	<0,05 *

Os resultados comentados acima podem ser visualizados nos dois gráficos abaixo: gráfico de dispersão dos valores observados (Figura 19) e gráfico *boxplot*, mostrando as medianas, os limites superiores dos quartis 1 e 3 e a amplitude de variação das variáveis, incluindo os valores *outlyers*, assinalados por asterisco (Figura 20). Nesses dois gráficos o grupo 0 corresponde a todos os grupos aglutinados, os números 1 a 10 identificam respectivamente os grupos WT1b, WT1a, RPS19, A2N, K6Q, R12N, V15NS16G, S18NS19G, D1819 e R31Q. O grupo 11 corresponde ao conjunto dos dois grupos WT1b e A2N e o grupo 12 aglutina os grupos WT1a, RPS19, K6Q, R12N, V15NS16G, S18NS19G, D1819 e R31Q.

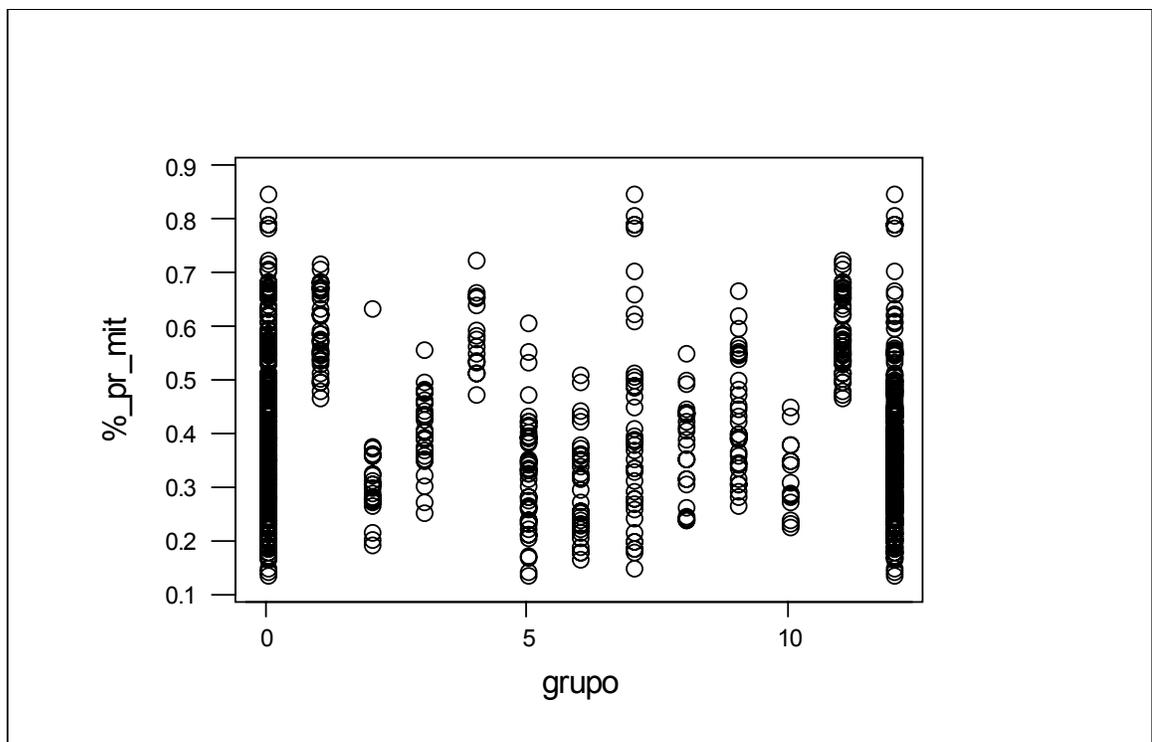


Figura 19 - Gráfico de dispersão dos valores observados para todos os valores (0), WT1b (1), WT1a (2), RPS19 (3), A2N (4), K6Q (5), R12N (6), V15NS16G (7), S18NS19G (8), D1819 (9) e R31Q (10), A2N e WT1b (11) e WT1a, RPS19, K6Q, R12N, V15NS16G, S18NS19G, D1819 e R31Q juntos (12). %_pr_mit = proporção, em percentagem, da quantidade da proteína encontrada em mitocôndrias

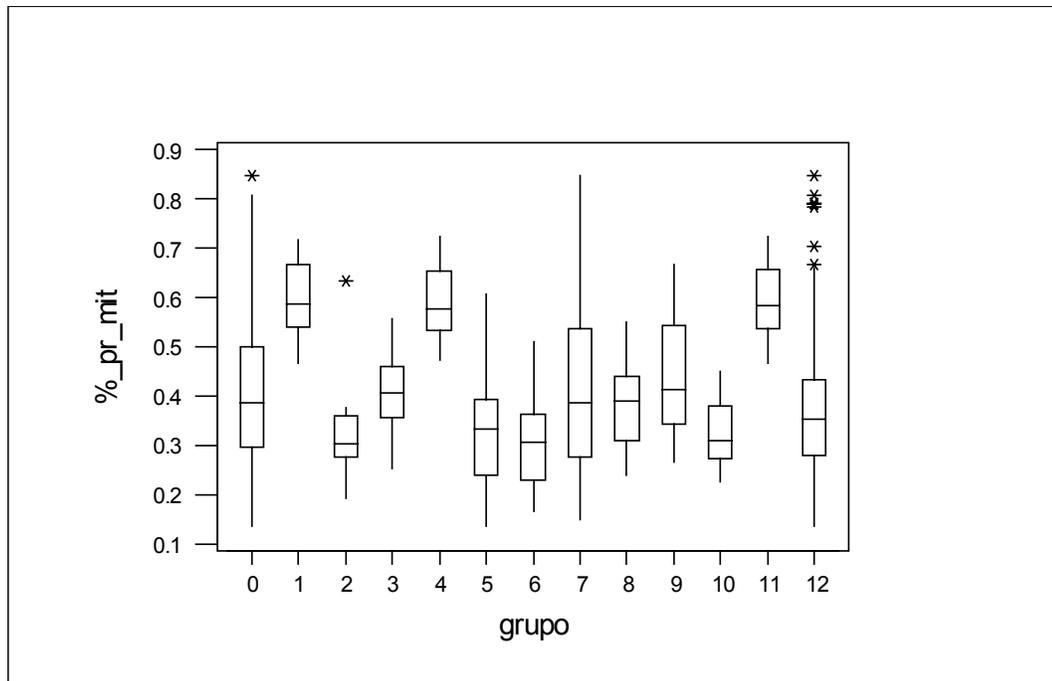


Figura 20 - Gráfico *boxplot* com os valores obtidos para WT1b (1), WT1a (2), RPS19 (3), A2N (4), K6Q (5), R12N (6), V15NS16G (7), S18NS19G (8), D1819 (9) e R31Q (10), A2N e WT1b (11) e WT1a, RPS19, K6Q, R12N, V15NS16G, S18NS19G, D1819 e R31Q juntos (12). Representações: medianas (traço horizontal no interior do retângulo), os limites superiores dos quartis 1 e 3 (extremidades inferior e superior do retângulo) e amplitude de variação das variáveis (barra vertical acima e abaixo do retângulo). Os valores *outliers* estão assinalados por asterisco. %_pr_mit = proporção, em porcentagem, da quantidade da proteína encontrada em mitocôndrias

O último ponto foi verificar se existiam diferenças dependentes do tipo celular em que foi inferida a proporção de proteína: célula do mesófilo (mes) ou estômato (est). Empregou-se o teste de Mann-Whitney, em relação aos dois grupos nos quais foi possível a aplicação do teste (WT1b e K6Q), pois eram os únicos que possuíam quantidade razoável de medidas realizadas em estômatos.

Os dados de estatística descritiva dessas duas variáveis apresentam-se distribuídos pelo mesófilo e pelos estômatos, na Tabela 16.

Tabela 16 – Dados de estatística descritiva para a comparação WT1b (mes x est) e K6Q (mes x est)

Parâmetro	WT1b(mes)	WT1b(est)	K6Q(mes)	K6Q(est)
média	0,6037	0,5627	0,3290	0,3238
n	27	7	37	10
s.d.	0,0672	0,0856	0,1051	0,1037
s.e.	0,0129	0,0324	0,0173	0,0328
Valor mínimo	0,4966	0,4677	0,1365	0,2134
Valor máximo	0,7190	0,6820	0,6080	0,5541
Mediana	0,5949	0,5465	0,3342	0,2783
l.inf.ic95%(2)	0,5771	0,4835	0,2940	0,2497
l.sup.ic95%(2)	0,6303	0,6419	0,3641	0,3980

Nota: Abreviações, mes, mesófilo; est, estômato; s.d., desvio padrão; s.e. erro padrão; l.inf.ic95% e l.sup.ic95%, limite inferior e superior no intervalo de confiança a 95%, respectivamente.

A Tabela 17 mostra a comparação entre as proporções de proteína mitocondrial inferidas a partir de observações no mesófilo e em estômatos. Os resultados acima indicam que as proporções de proteína mitocondrial ou cloroplastidial distribuem-se igualmente no mesófilo e nos estômatos.

Tabela 17 - Resultados da estatística de Mann-Whitney e da soma de pontos dos dados referentes a mesófilo e estômato, no direcionamento mitocondrial de Wt1b e K6Q

Variável	U	U'	$\Sigma r(\text{mes})$	$\Sigma r(\text{est})$	P
WT1b	63	126	504	91	0,1912 n.s.
K6Q	173	197	900	228	0,7650 n.s.

Nota: Abreviações, mes, mesófilo; est, estômato; U e U', estatística de Mann-Whitney; $\Sigma r(\text{mes})$ e $\Sigma r(\text{est})$, soma dos postos dos dados mesófilo e estômato, respectivamente.

4.2.4 Conclusão das análises dos efeitos das mutações sítio-dirigidas realizadas na seqüência de direcionamento de RBP1b

Quando RBP1b, sem sua seqüência de direcionamento, foi fusionada à GFP, observou-se o sinal da GFP apenas no citosol. Esse dado confirmou que a informação essencial necessária ao direcionamento de RBP1b está localizada na região amino-terminal da proteína, entre os primeiros 39 aminoácidos de sua seqüência (Figura 21).

Observou-se que a mutação da alanina na posição 2 por asparagina, A2N, não teve efeitos no direcionamento de RBP1b para mitocôndrias e cloroplastos (os dados obtidos para A2N não diferem significativamente dos obtidos para RBP1b, vide item 4.2.3, Tabela 13). Trata-se de um aminoácido conservado entre as proteínas com duplo direcionamento e o efeito de mutações nessa posição já foi investigado anteriormente. Os efeitos de mutações nos primeiros aminoácidos no direcionamento da formato desidrogenase de batata foram analisados *in vivo* (AMBARD-BRETTEVILLE et al., 2003). A proteína, exclusivamente mitocondrial, teve seu direcionamento abolido quando a metionina e a alanina nas posições 1 e 2 foram deletadas. Essa modificação gera cargas positivas na porção hidrofóbica da α -hélice anfifílica, o que pode comprometer a interação da proteína com chaperonas citosólicas ou translocases das membranas da organela. A substituição da alanina na posição 2 por arginina não teve efeitos no direcionamento, mas quando a troca foi feita por serina, o direcionamento foi menos eficiente e a proteína pôde ser detectadas em mitocôndrias e no citosol.

Já Pujol; Maréchal-Drouard; Duchêne (2007) observaram, em experimentos *in vitro*, que mutações no segundo aminoácido das proteínas metionil, glutamil e fenilalanil-tRNA sintetases tiveram efeitos variados em seu direcionamento. A mutação da alanina na posição 2, para glicina ou serina, na seqüências de metionil-tRNA sintetase reduziu significativamente a quantidade da proteína em mitocôndrias, ao passo que a mudança da alanina 2 por serina aumentou a quantidade de glutamil-tRNA sintetase em mitocôndrias. Apesar de a alanina na posição 2 ser freqüente em proteínas direcionadas aos cloroplastos, a troca da treonina na posição 2 por alanina na seqüência da fenilalanil-tRNA sintetase reduziu a quantidade da proteína em cloroplastos.

Os efeitos variados das mutações na posição 2 até agora descritos sugerem elas sejam dependentes do contexto protéico e que há cooperação do segundo aminoácido com outros determinantes na região amino-terminal.

Já as mutações K6Q e R12N, em aminoácidos positivos na região amino-terminal da seqüência de direcionamento, afetam negativamente, de modo significativo (Tabela 13), o transporte de RBP1b para mitocôndrias. A dupla mutação, K6QR12N abole totalmente o transporte de RBP1b para mitocôndrias (Figura 22), provavelmente por alterar cargas positivas na estrutura da α -hélice anfifílica. Esse resultado foi confirmado pela importação *in vitro* de proteínas para mitocôndrias (dado não mostrado). Os dados obtidos para as mutações K6Q e R12N estão de acordo com aqueles obtidos anteriormente (CHEW et al., 2003; PUJOL; MARÉCHAL-DROUARD; DUCHÊNE, 2007), em que foi apontada a importância de aminoácidos carregados positivamente na região amino-terminal da seqüência de direcionamento para o endereçamento de proteínas para mitocôndrias.

A mutação R31Q, em um aminoácido positivo na região carboxi-terminal, próxima ao sítio de clivagem, também afeta negativamente o transporte de RBP1b para mitocôndrias. Resultados similares foram obtidos por Rudhe et al. (2004) que substituíram aminoácidos positivos próximos ao sítio de processamento da proteína e observaram que o transporte para mitocôndrias foi significativamente reduzido.

Os resultados obtidos para as construções S18NS19G e Δ (18-19) diferem significativamente dos de RBP1b (Tabela 13). Apesar de a presença de aminoácidos hidroxilados, como serina, já ter sido associada ao direcionamento para cloroplastos, o efeito dessas mutações reduziu o endereçamento de RBP1b para mitocôndrias e, conseqüentemente, aumentou o direcionamento para cloroplastos.

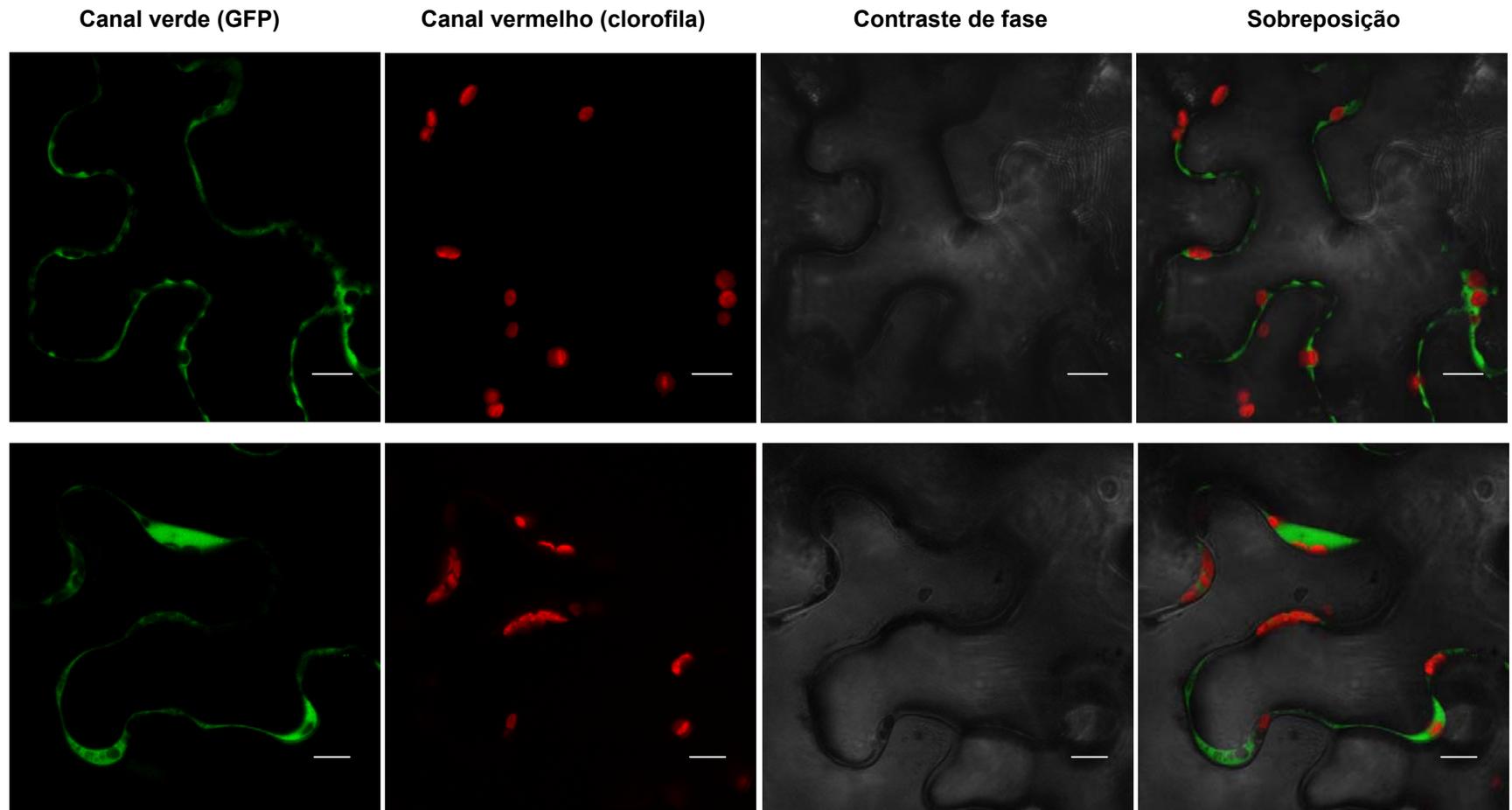


Figura 21 - Localização citosólica da proteína RBP1b sem seqüência de direcionamento. Células do mesófilo de *N. benthamiana* transformadas com as construções pUCAP-GFP, controle citosol (acima) e pUCAP-RBP1b sem seqüência de direcionamento (abaixo). A barra representa 10 μ m

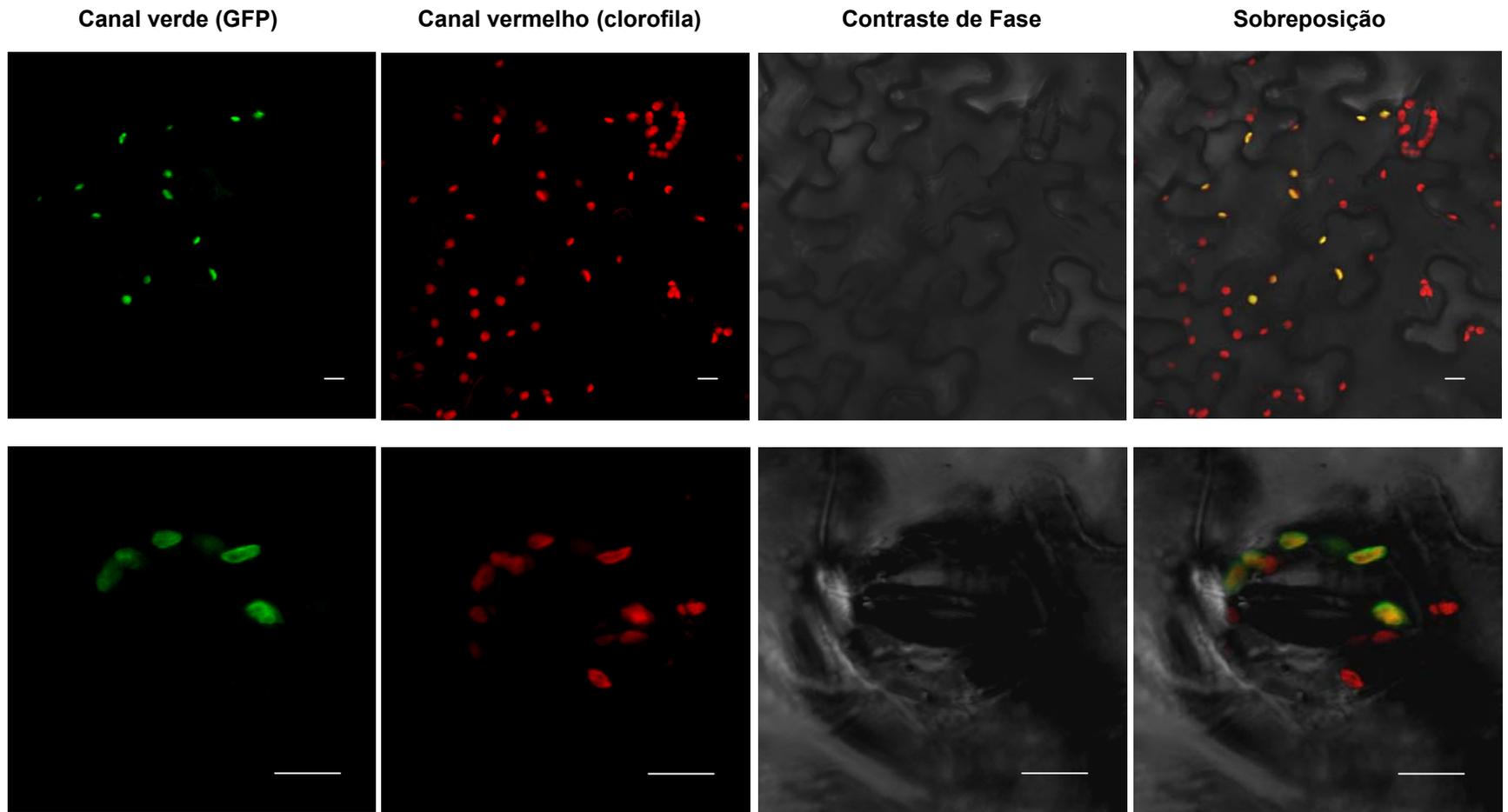


Figura 22 - Localização celular da proteína mutada, K6QR12N exclusivamente em cloroplastos. Célula do mesófilo de *N. benthamiana* (acima) e estômato (abaixo) transformados com a construção pK6QR12N. Observa-se a fluorescência da GFP nos cloroplastos. A barra representa 10 μm

Os resultados do duplo mutante V15NS16G, apesar de significativamente diferentes do selvagem RBP1b (Tabela 13), são inconclusivos. Os dados obtidos para essa mutação apresentam ampla distribuição e desvio padrão e variância elevados (Tabela 7, Figura 19). Não foi observada uma tendência, o que sugere que o efeito da mutação foi variável nas células analisadas, podendo as mutações inibir drasticamente o transporte para mitocôndrias ou não.

Proteínas transportadas para mitocôndrias e cloroplastos, mesmo com seqüências de direcionamento ambíguas, parecem ter sinais para a importação para uma ou outra organela localizados em regiões específicas da seqüência de direcionamento. Geralmente, o sinal de direcionamento para mitocôndrias está localizado na região amino-terminal (CHEW et al., 2003). Diante dessas informações, foram realizadas as deleções $\Delta(2-17)$ e $\Delta(18-30)$ na seqüência de direcionamento de RBP1b. A deleção da primeira metade da seqüência de direcionamento, $\Delta(2-17)$, aboliu o transporte da proteína para mitocôndrias e o sinal da GFP foi observado em cloroplastos e no citosol (Figura 23). Já quando a segunda metade da seqüência de direcionamento foi deletada, o sinal da GFP pôde ser observado com menos intensidade em mitocôndrias e cloroplastos e também no citosol (Figura 24). Esses dados demonstram que a informação essencial para o direcionamento de RBP1b está localizada na região amino-terminal, entre os primeiros 17 aminoácidos da seqüência de direcionamento. O sinal para endereçamento aos cloroplastos encontra-se ao longo de toda a seqüência, uma vez que não foi abolido por qualquer das duas mutações. Além disso, a primeira metade da seqüência de direcionamento de RBP1b tem informação suficiente para promover o duplo direcionamento de RBP1b e a segunda metade parece ser responsável pela eficiência da importação.

Pujol; Maréchal-Drouard; Duchêne (2007) demonstraram que deleções na região amino-terminal da glutamil-tRNA sintetase de *A. thaliana* reduzem a quantidade da proteína em mitocôndrias e cloroplastos. A deleção dos primeiros 29 aminoácidos, o correspondente a 2/3 da proteína, reduz significativamente sua ligação a cloroplastos, mas não a mitocôndrias.

No entanto, os dados aqui obtidos estão em desacordo com os obtidos por Rhude et al. (2002b) em análises *in vitro* dos efeitos de deleções na seqüência de

direcionamento da glutathione redutase de ervilha. A deleção dos primeiros 16 aminoácidos resultou em aumento do transporte para mitocôndrias; os autores concluíram que a primeira metade da seqüência de direcionamento da glutathione redutase não continha informação essencial ao transporte da proteína, mas aumentava a eficiência de sua importação.

Essas diferenças mostram como as seqüências de direcionamento ambíguas são diversas em sua organização do sinal de direcionamento, não parecendo possível estabelecer consenso.

4.2.5 Quantificação do direcionamento de RBP1b, RBP1a e RPS19 para mitocôndrias e cloroplastos

Membros da subfamília RBP1 têm seqüências conservadas, o que sugere uma origem recente por duplicação gênica e uma sobreposição funcional. No entanto, pode haver, entre os membros dessa subfamília, uma regulação gênica diferencial nos níveis temporal e espacial ou até mesmo uma isoforma pode ser requerida em maior quantidade em mitocôndrias do que em cloroplastos ou vice-versa. Para investigar a última hipótese, foram feitas análises quantitativas com imagens obtidas pela transformação transiente de *N. benthamiana* com as construções pUCAP-RBP1a, pUCAP-RBP1b e pUCAP-RPS19. Apesar da conservação entre as seqüências de direcionamento dos membros da subfamília RBP1, nas condições usadas, observou-se que as proteínas RBP1a e RPS19 são mais direcionadas para cloroplastos, enquanto RBP1b está presente em maior quantidade em mitocôndrias. Não foram registradas variações significativas entre os dados de RBP1a e RPS19 (Tabela 15). Esses dados representam um indício sobre a diferença de distribuição dessas proteínas entre cloroplastos e mitocôndrias. Certamente, dados mais precisos seriam obtidos com construções guiadas pela região promotora de cada um desses três genes.

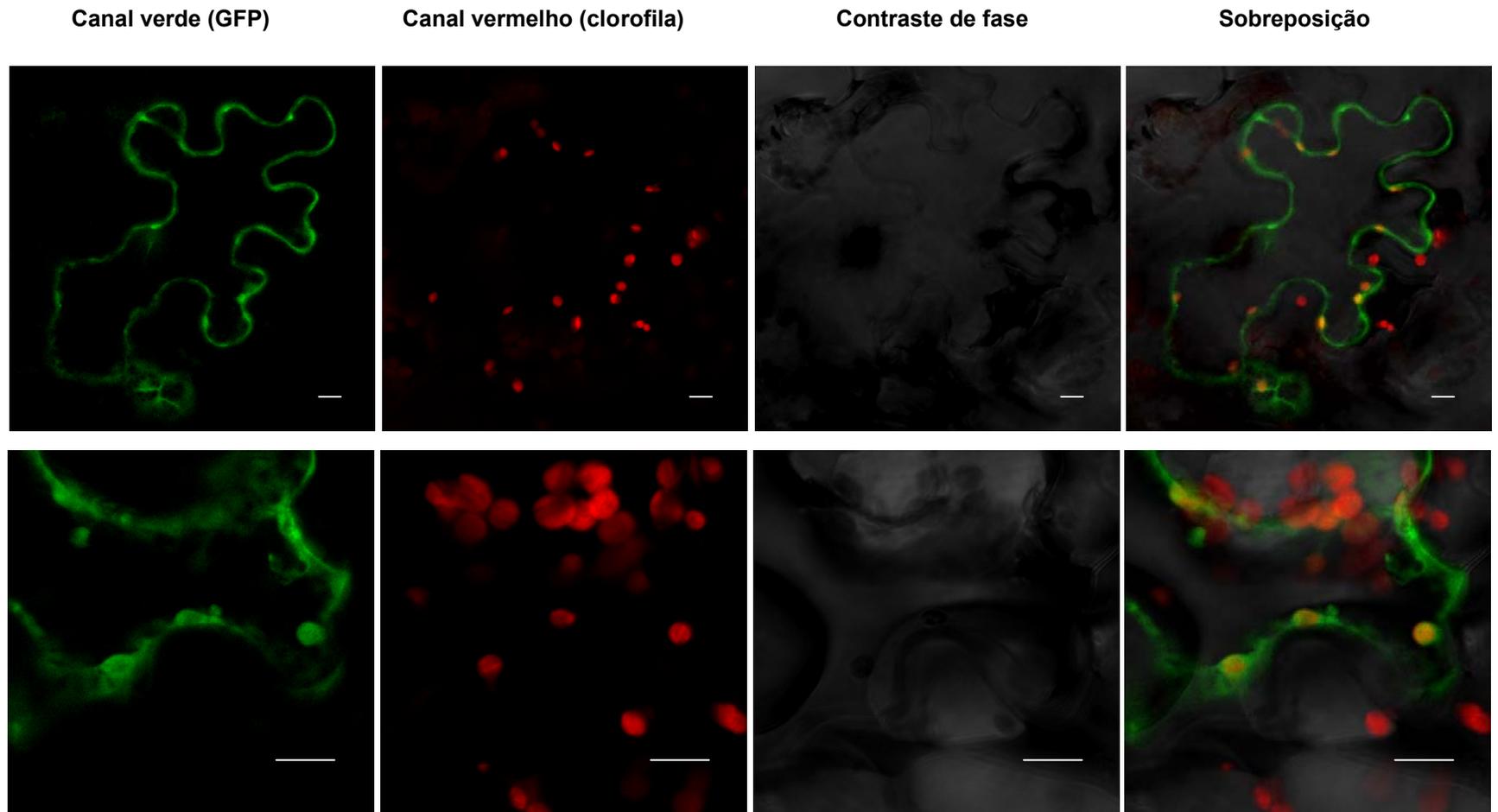


Figura 23 – Efeito da deleção $\Delta(2-17)$ no direcionamento de RBP1b. Células do mesófilo de *N. benthamiana* transformadas com a construção pUCAP- $\Delta(2-17)$. Observa-se a fluorescência da GFP no citosol e nos cloroplastos. A barra representa 10 μm

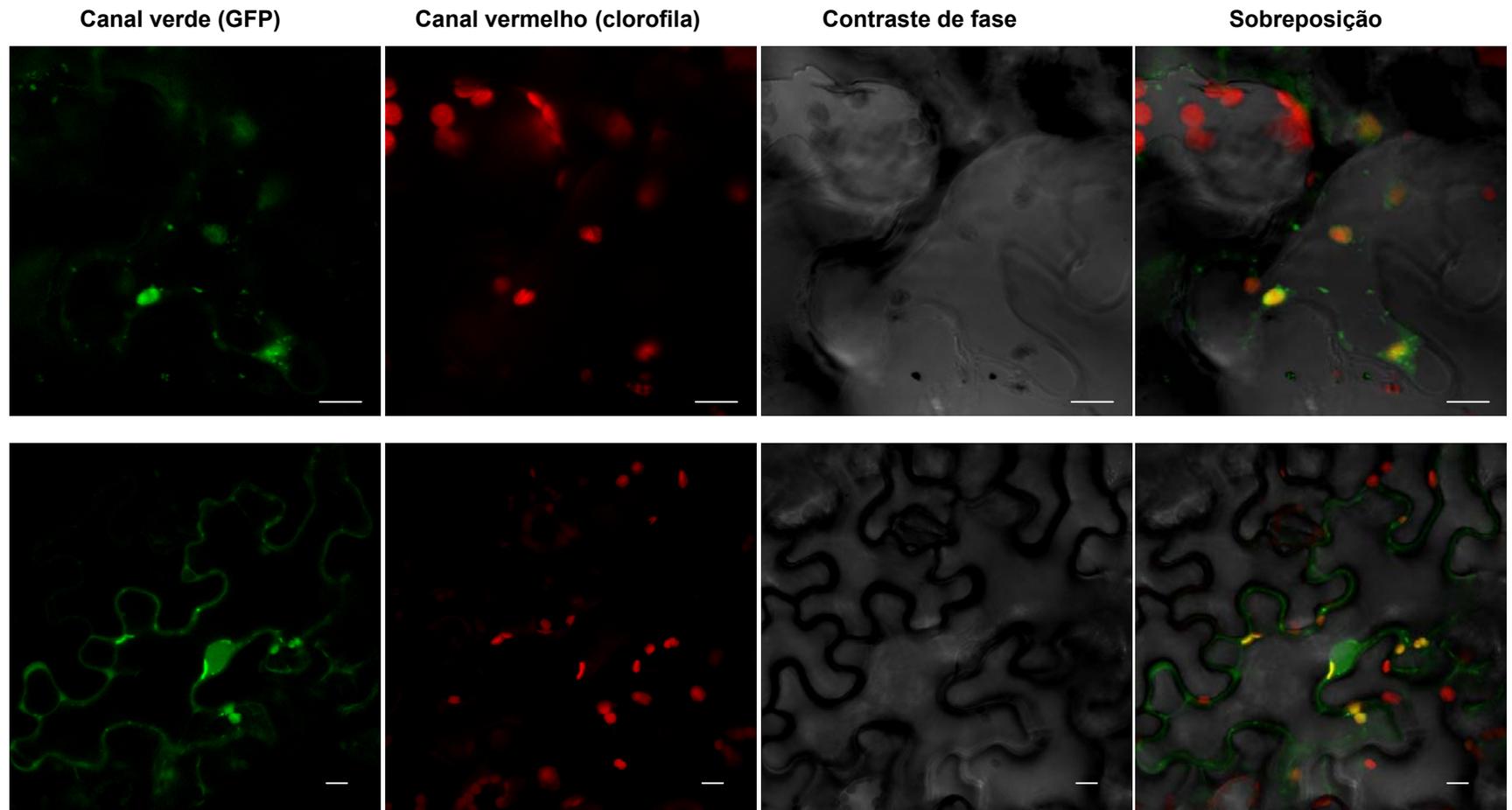


Figura 24 – Efeito da deleção $\Delta(18-30)$ no direcionamento de RBP1b Células do mesófilo de *N. benthamiana* transformadas com a construção pUCAP- $\Delta(18-30)$. Observa-se a fluorescência da GFP em mitocôndrias e cloroplastos (acima) e também no citosol (abaixo). A barra representa 10 μm

4.3 O duplo direcionamento da proteína THI1 de *Arabidopsis thaliana*

A THI1, enzima da biossíntese do tiazol, está envolvida na biossíntese de tiamina (vitamina B1) que, em plantas, ocorre nos cloroplastos. Essa proteína apresenta ortólogos em Archaea, fungos e diversas espécies de plantas. A tiamina pirofosfato, a forma ativa, é um cofator essencial requerido para a atividade de várias enzimas associadas a processos metabólicos importantes, incluindo o ciclo do ácido cítrico, a glicólise, a biossíntese de pigmentos, o ciclo da pentose fosfato entre outros (BELANGER et al., 1995; KIM et al, 1998).

O gene *thi1* de *A. thaliana* foi originalmente clonado pela sua capacidade de complementação em bactérias deficientes no reparo de DNA (MACHADO et al., 1996). Além disso, o homólogo de *thi1* em *Saccharomyces cerevisiae*, *thi4*, está envolvido com a tolerância a danos no DNA mitocondrial. Leveduras com o gene *thi4* silenciado reverterem ao fenótipo selvagem quando transformadas com os homólogos de milho, *Arabidopsis* e *Alnus glutinosa* (BELANGER et al., 2005; MACHADO et al., 1997; RIBEIRO et al., 1996). No último caso, a complementação ocorre somente quando a região amino-terminal da proteína é substituída pela seqüência de direcionamento de Thi4. Quando a região amino-terminal de THI1 de *A. thaliana* é deletada, a complementação é ineficiente tanto para a auxotrofia a tiamina, como para a instabilidade mitocondrial. Dessa forma, ficou evidenciada a bifuncionalidade de Thi4 e, provavelmente, de THI1 (MACHADO et al., 1997; MEDINA-SILVA et al., 2006).

A análise da expressão de *thi1* em *Arabidopsis thaliana* foi realizada em plantas selvagens e mutantes deficientes na biossíntese de tiamina. Esses mutantes contêm uma mutação pontual do tipo transição, que resultou na substituição de uma alanina por valina em uma região conservada de THI1. Não foram detectadas diferenças significativas na expressão gênica entre plantas selvagens e mutantes na presença ou ausência de tiamina. A expressão de *thi1* é dependente de luz, independente de tiamina e está reduzida em raízes (PAPINI-TERZI et al., 2003).

Ribeiro et al. (2005) realizaram a análise funcional do promotor de *thi1* de *A. thaliana* com o objetivo de determinar o padrão de expressão do gene e sua resposta a diferentes condições de estresse. Os autores mostraram que o promotor de *thi1* possui um amplo padrão de expressão, detectada em todos os órgãos em diferentes estágios

do desenvolvimento da planta. Sua expressão é mais acentuada na parte aérea do que em raízes, onde é restrita ao sistema vascular. A expressão de *thi1* é modulada pela luz e afetada positivamente em situações de alta salinidade, privação de açúcar e alagamento.

A proteína THI1 está relacionada à resistência a doenças em arroz. Plantas cujo gene *thi1* foi silenciado demonstraram susceptibilidade a *Xanthomonas oryzae* pv. *oryzae* e *Magnaporthe grisea*. Foram detectados baixos níveis de tiamina nessas plantas e a aplicação exógena de tiamina restabeleceu a resistência às bactérias. Esses dados reforçam o papel da tiamina na indução de genes de defesa em plantas (WANG et al., 2006).

Analisando a seqüência de aminoácidos na região amino terminal da THI1 de *Arabidopsis thaliana*, Chabregas et al. (2001) identificaram um peptídeo de trânsito cloroplastidial adjacente a uma pré-seqüência mitocondrial putativa (Figura 25). Os autores confirmaram o duplo direcionamento da proteína por meio de ensaios de imunolocalização e evidenciaram a presença de apenas um transcrito codificando as duas isoformas da THI1, o que sugeriu um mecanismo pós-transcricional de duplo direcionamento.

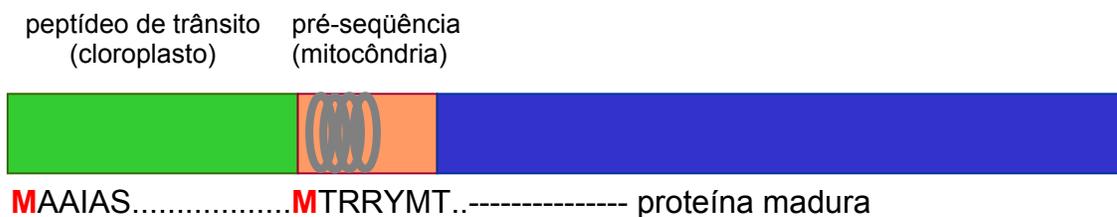


Figura 25 - Representação esquemática da THI1 de *A. thaliana*. Na região amino terminal da proteína há um peptídeo de trânsito (em verde) e uma pré-seqüência (em laranja), apresentando uma α -hélice anfifílica (em cinza). Em vermelho, estão destacadas as metioninas correspondentes aos dois inícios de tradução

A confirmação da existência de um mecanismo de controle traducional para o duplo direcionamento de THI1 foi obtida por Chabregas et al. (2003). Ensaios de transcrição e tradução *in vitro* revelaram a presença de dois produtos traducionais gerados pelo uso diferencial de dois códons de início de tradução situados no mesmo quadro de leitura. O segundo sítio de início da tradução encontra-se 204 pb após o primeiro. Esses dados foram corroborados por ensaios *in vivo*. Os primeiros 155 aminoácidos de THI1 foram fusionados à GFP e a construção foi utilizada para a transformação transiente de protoplastos de *Nicotiana tabacum*. Os resultados demonstraram a localização de THI1 em mitocôndrias e cloroplastos. Foram introduzidas mutações sítio dirigidas nos dois códons de início de tradução e os resultados revelaram que o início da tradução no primeiro AUG faz com que a proteína seja direcionada para os cloroplastos enquanto que o início da tradução no segundo AUG direciona-a para as mitocôndrias. Análises da eficiência de tradução demonstraram que o melhor contexto para seu início ocorre com a utilização do primeiro AUG, o que sugere que essa proteína é requerida em maior quantidade nos cloroplastos. Resta identificar as regiões da seqüência de direcionamento da proteína que conferem especificidade a seu direcionamento para cloroplastos e mitocôndrias.

O fato de o primeiro AUG possuir o melhor contexto para início da tradução, e, dessa forma, ser o preferencialmente utilizado, está de acordo com o modelo proposto por Kozak (1991), em que o sítio de iniciação em mRNAs eucarióticos são alcançados via o mecanismo de escaneamento e a tradução inicia-se no códon AUG mais próximo da extremidade 5' do mRNA (KOZAK, 2002). No caso da THI1, o início da tradução no segundo AUG poderia ser explicado por um mecanismo de *leaky scanning*, uma “falha” do escaneamento do mRNA em que a subunidade pequena do ribossomo não reconheceria o primeiro AUG e a tradução se iniciaria no códon AUG seguinte. Uma outra explicação para o início da tradução no segundo AUG seria um mecanismo do tipo IRES, sítio interno de entrada do ribossomo (*internal ribosomal entry site*) (KOMAR; HATZOGLU, 2005). O posicionamento da subunidade pequena do ribossomo ocorreria pelo pareamento de seqüências complementares entre o mRNA e o rRNA 18S, um mecanismo ainda não descrito em plantas. Chabregas et al. (2003) identificaram duas regiões na seqüência de direcionamento de THI1, aqui nomeadas

como 18S1 e 18S2, as quais apresentam certa complementaridade com a seqüência do rRNA 18S (Figura 26). Essas regiões não são encontradas na seqüência de mRNA de *thi1* de arroz, apesar de existirem evidências de que a THI1 de arroz também apresenta duplo direcionamento.

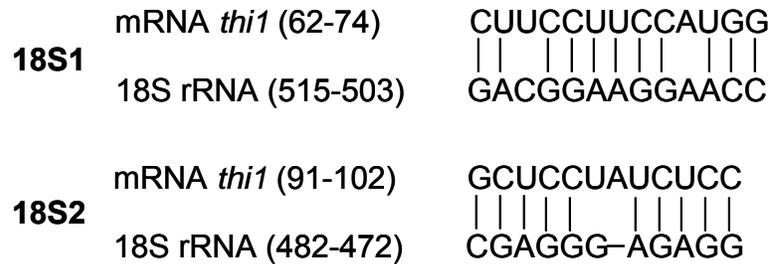


Figura 26 - Complementaridade entre regiões da seqüência de direcionamento de THI1 (18S1 e 18S2) e a seqüência do 18S rRNA detectada por Chabregas et al. (2003)

Diante desses fatos, a proposta deste trabalho foi investigar os mecanismos celulares que permitem o início da tradução no segundo AUG, bem como aqueles que conferem especificidade ao direcionamento de THI1 para cloroplastos e mitocôndrias.

A metodologia escolhida para essa investigação foi a de mutação sítio dirigida nas regiões do cDNA de *thi1* que apresentam complementaridade com o rRNA 18S. Essas regiões estão entre os dois códons de início da tradução e podem representar possíveis sítios IRES na seqüência de *thi1*. As mutações foram realizadas por PCR, utilizando-se iniciadores específicos contendo as mutações a serem inseridas na seqüência de *thi1* (Figura 27). Para a realização das mudanças, foi consultado o uso de códons em *A. thaliana* e utilizado o códon mais freqüente. Foram realizadas mutações na região 18S1, 18S2 e nas duas simultaneamente. As seqüências mutadas foram clonadas no vetor PMA4-GFP, *in frame* com a GFP. Como controle, a seqüência de THI1, sem mutações, foi também clonada no mesmo vetor. Foram feitas construções com a seqüência completa do cDNA de *thi1*, abrangendo seus 1.047 pb sem o códon de terminação, ou somente com sua seqüência de direcionamento, incluindo os primeiros 315 pb. As construções foram utilizadas para a transformação transiente de epitélio de cebola, via biobalística.

18S1

S S F H G
 Seqüência original: UCU UCC UUC CAU GGC

S S F K L
 Seqüência mutada: AGU AGU UUU AAA CUU

18S2

A P I S
 Seqüência original: GCU CCU AUC UCC

I A I S
 Seqüência mutada: AUU GCU AUU AGU

Figura 27 - Mutações realizadas nas regiões 18S1 e 18S2 na seqüência de direcionamento de THI1. Acima da seqüência de nucleotídeos está representada a respectiva seqüência de aminoácidos. Os nucleotídeos mutados estão em vermelho e aqueles que apresentam complementaridade ao 18S rRNA estão em azul

Quando a transformação de epitélio de cebola foi realizada com as construções contendo a seqüência completa de THI1, não foi possível observar a fluorescência da GFP. Já com as construções contendo apenas sua seqüência de direcionamento, a fluorescência da GFP foi detectada somente nos cloroplastos. O mesmo resultado foi observado quando foram realizadas as transformações com as construções contendo as mutações 18S1 e 18S2 (dados não mostrados). Diante dos problemas, a tentativa para uma solução foi mudar o sistema de transformação transiente. Células do mesófilo de *N. benthamiana* foram transformadas via biobalística com a construção contendo a seqüência de direcionamento de THI1, sem mutações. A fluorescência da GFP foi novamente detectada exclusivamente nos cloroplastos (Figura 28).

Os dados de Chabregas et al. (2003) demonstram que o códon preferencialmente utilizado para o início da tradução é o primeiro AUG, situação em que

é produzida a proteína direcionada para cloroplastos. Poder-se-ia supor que o início da tradução a partir do segundo AUG fosse pouco freqüente, a ponto de não ter sido possível visualizar o sinal da GFP em mitocôndrias no sistema utilizado. Dessa maneira, decidiu-se mudar o vetor de expressão e a seqüência de direcionamento de THI1 foi transferida para o vetor pUCAP-GFP, o qual contém a seqüência do promotor do vírus do mosaico da couve-flor CaMV35S duplicada *in tandem*. Com essa construção, folhas de *N. benthamiana* foram transformadas por bombardeamento. O resultado anterior repetiu-se e novamente o sinal da GFP foi detectado apenas nos cloroplastos (Figura 28). Quando as células são transformadas com a construção pUCAP-THI1ATC2, a qual contém a seqüência de direcionamento de THI1, com o segundo ATG mutado para ATC, fusionada à GFP, a proteína é encontrada somente em cloroplastos. Já quando a primeira parte da seqüência de direcionamento é deletada, até o segundo ATG (construção pUCAP-THI1mito), o sinal da GFP é observado no citosol (Figura 29).

Os resultados demonstram que deve haver um problema com o direcionamento de THI1 quando a proteína está fusionada à GFP. Essa fusão deve causar uma mudança em sua estrutura secundária a ponto de interferir no direcionamento. Resultados semelhantes foram obtidos por Almeida (2004) em uma tentativa de observar o duplo direcionamento de THI1 em plantas de *A. thaliana* transformadas estavelmente com a seqüência de direcionamento de THI1 fusionada *in frame* à da GFP. Nesta ocasião, o sinal da GFP foi detectado somente em cloroplastos, independentemente da região promotora utilizada, CamV35S ou o próprio promotor de *thi1*. A fluorescência da GFP também não foi observada quando a transformação foi realizada com construções contendo a seqüência completa do cDNA de *thi1* fusionada a da GFP.

Apesar da discordância entre os dados da localização de THI1, novas evidências confirmam sua presença em mitocôndrias. Stock², utilizando a técnica de duplo-híbrido, identificou a interação de THI1 de *A. thaliana* com a proteína ISC1 mitocondrial, a qual também atua na via de biossíntese de tiamina (informação verbal).

² Stock, C.P. Escola Superior de Agricultura Luiz de Queiroz, USP.

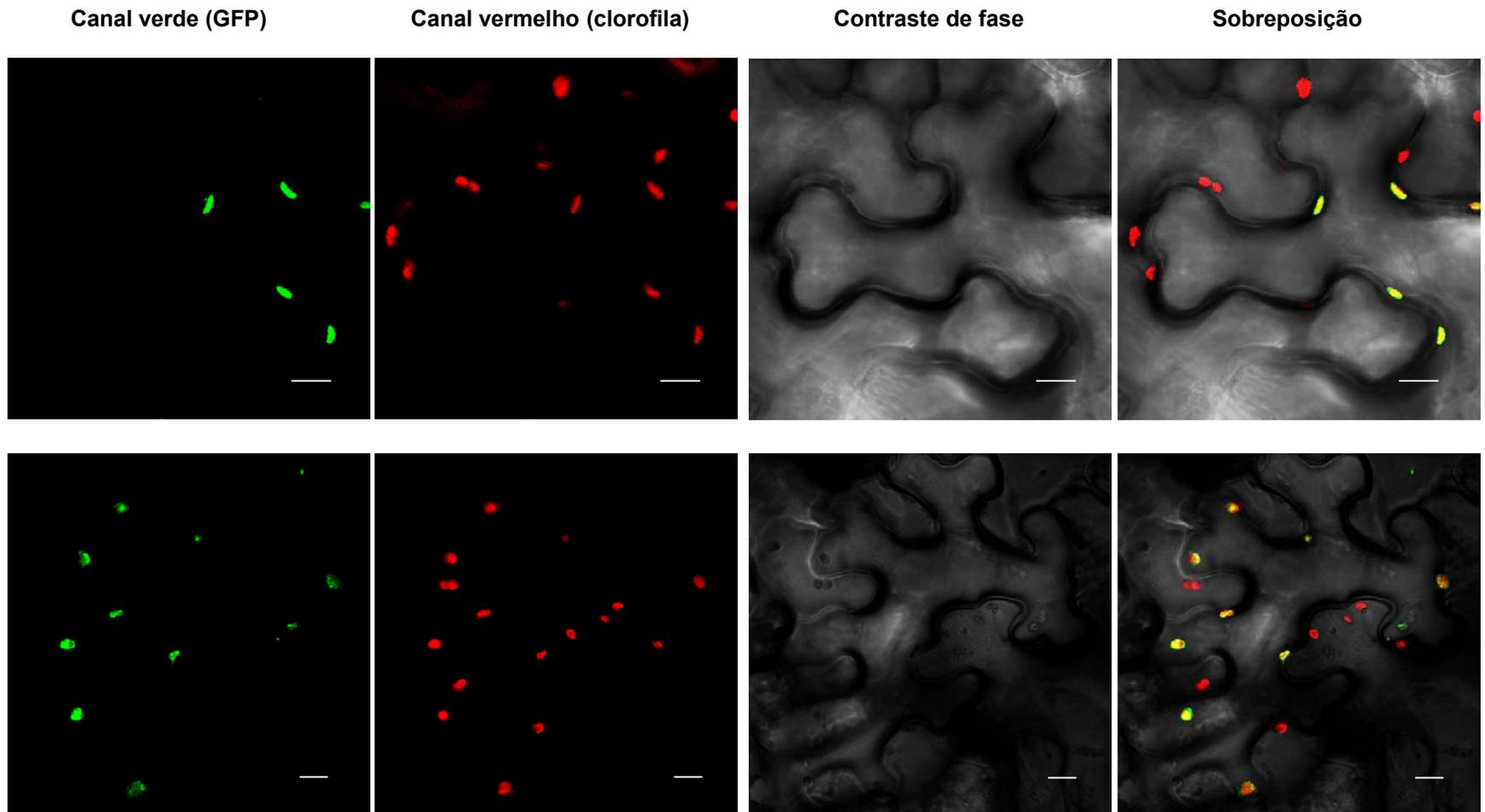


Figura 28 – O direcionamento da THI1 selvagem. Células do mesófilo de *N. benthamiana* transformadas com as construções pPMA4- THI1SD (acima) e pUCAP-THI1SD (abaixo). Nos dois casos, observa-se a fluorescência da GFP em cloroplastos. A barra representa 10 μ m

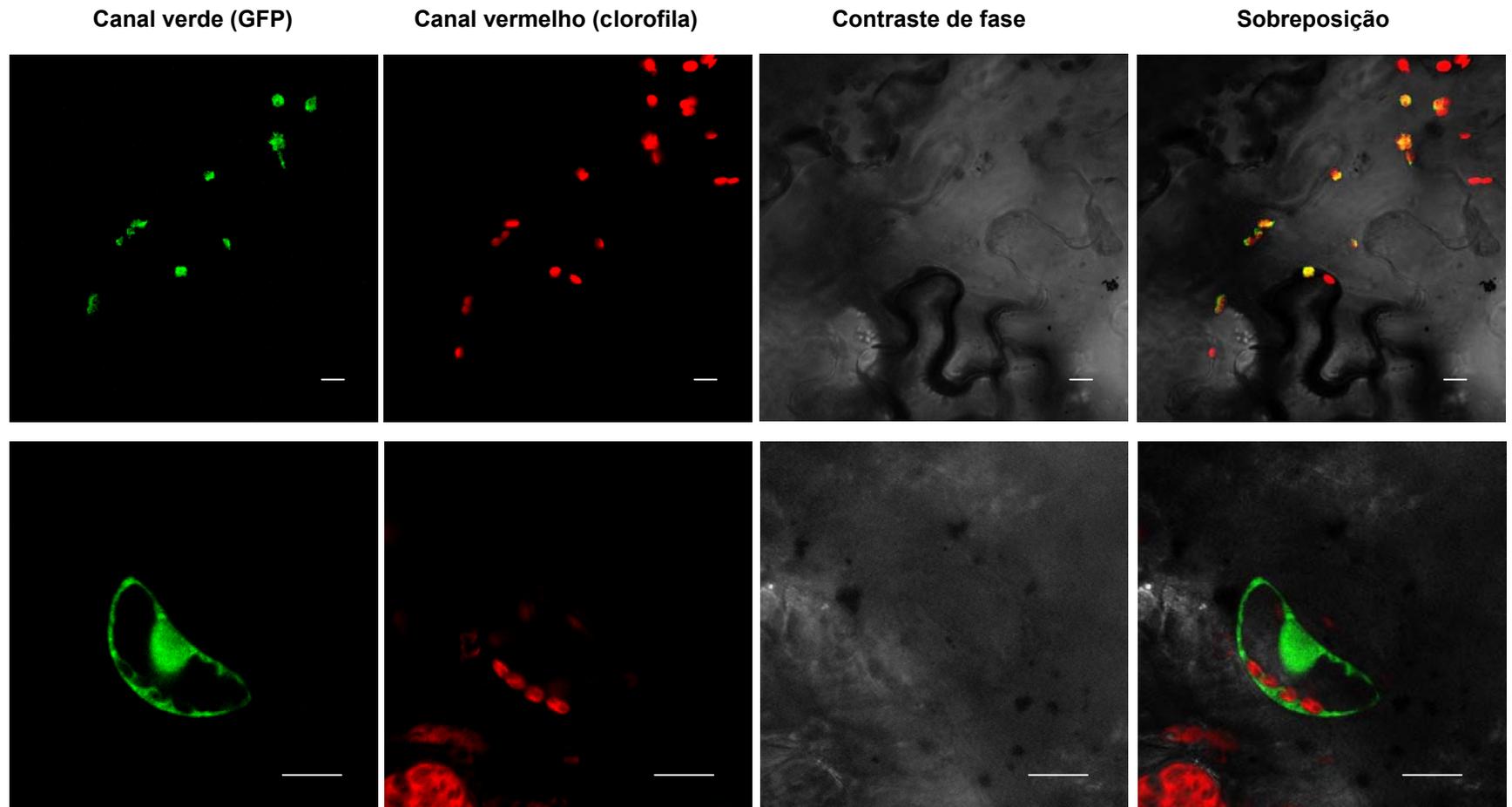


Figura 29 – Direcionamento da proteína THI1 contendo mutação de ponto na seqüência de direcionamento. Células do mesófilo de *N. benthamiana* transformadas com as construções pUCAP-THI1ATC2 (acima) e pUCAP-THI1mito (abaixo). A fluorescência da GFP foi detectada em cloroplastos e no citosol, respectivamente. A barra representa 10 μm

Além disso, relatos na literatura descrevem problemas com a localização de proteínas mitocondriais quando estão fusionadas à da GFP. A seqüência completa da glutathione redutase, proteína direcionada para mitocôndrias e cloroplastos, fusionada à GFP, é direcionada somente para cloroplastos (CHEW et al., 2003). No entanto, sua dupla localização celular foi confirmada por outros métodos e a atividade dessa enzima já foi detectada nessas duas organelas. Sua seqüência de direcionamento tem a capacidade de endereçar a proteína fosfinotricina-acetil-transferase bacteriana para mitocôndrias e cloroplastos, em células de *N. tabacum* (CREISSEN et al., 1995). Outro caso semelhante é o da amidoimidazol-nucleotídeo sintetase cujo direcionamento para mitocôndrias não pôde ser detectado quando a proteína estava fusionada à GFP (GOGGIN et al., 2003). Também para a proteína holocarboxilase sintetase 1, que teve seu duplo direcionamento detectado para mitocôndrias e cloroplastos por análises de *Western Blot*, não foi possível a confirmação de seu direcionamento para mitocôndrias em análises *in vivo* em que parte da proteína foi fusionada à GFP (PUYAUBERT; DENIS; ALBAN, in press).

Na tentativa de achar um outro sistema em que o duplo direcionamento da THI1 pudesse ser confirmado, foi realizado o transporte *in vitro* da proteína para mitocôndrias. A seqüência completa do cDNA de *thi1* foi clonada no vetor TOPO (Invitrogen) e a construção foi utilizada para ensaios de transcrição e tradução *in vitro*. Observou-se, como resultado da tradução, um produto com 37 kDa, semelhante ao peso molecular de THI1. Sendo assim, prosseguiu-se com os experimentos de isolamento de mitocôndrias de batata e transporte *in vitro* da proteína. Surpreendentemente, não foi possível detectar o transporte de THI1 para mitocôndrias, apesar do controle mitocondrial ter gerado o resultado esperado. Quando a transcrição e a tradução *in vitro* foram realizadas com a construção que continha a seqüência do cDNA de *thi1*, apresentando uma mutação no primeiro ATG para ATC, não foi possível detectar um produto traducional. No entanto, esses dados *in vitro*, apesar de estarem de acordo com os dados obtidos com os experimentos com a GFP, não podem ser utilizados para conclusões. Faltam repetições e um resultado negativo como esse, em que não se observam bandas no gel, pode ser devido a outros fatores que não o do fato

de a proteína não ser direcionada para mitocôndrias. Um problema metodológico interferindo com o transporte *in vitro* de THI1, por exemplo, geraria o mesmo resultado.

A falta de um sistema que reproduzisse o duplo direcionamento de THI1 impossibilitou a continuidade do projeto de investigação dos mecanismos que permitem o início da tradução no segundo AUG da seqüência do mRNA de *thi1*. Ficam aqui registradas algumas dificuldades no estudo da localização celular de proteínas em plantas.

5 CONCLUSÕES

(i) A estratégia de duplo direcionamento é conservada entre proteínas ortólogas de *Arabidopsis thaliana* e *Oryza sativa*. A conservação do direcionamento para mitocôndrias e cloroplastos foi demonstrada experimentalmente para as proteínas monodeidroascorbato redutase, metionina aminopeptidase e, provavelmente, para THI1. Os dados obtidos sugerem que, nas famílias gênicas com membros que apresentam duplo direcionamento, as duplicações que deram origem aos parálogos não foram aleatórias e que a evolução dessas famílias deve ter seguido um mesmo padrão em *A. thaliana* e *O. sativa*.

(ii) As proteínas RBP1a, RBP1b e RPS19 de *A. thaliana* têm duplo direcionamento para mitocôndrias e cloroplastos. A análise de mutações sítio-dirigidas na seqüência de direcionamento de RBP1b mostrou que a mutação do aminoácido na posição 2, uma alanina, não teve efeito no duplo direcionamento da proteína. Já a mutação dos aminoácidos positivos na região amino-terminal, K6Q e R12N, reduziu significativamente o transporte de RBP1b para mitocôndrias. Quando foi introduzida a dupla mutação, K6QR12N, o direcionamento de RBP1b foi totalmente abolido e a proteína foi encontrada somente em cloroplastos. A mutação de uma arginina na posição 31, próxima ao sítio de clivagem, também reduz significativamente o direcionamento para mitocôndrias. Confirmou-se, assim, a importância de aminoácidos positivos na determinação do direcionamento de proteínas para mitocôndrias. Mutações de ponto ou deleções de serinas nas posições 18 e 19 parecem reduzir o endereçamento de RBP1b para mitocôndrias, apesar de aminoácidos hidroxilados, como serina, estarem associados ao transporte para cloroplastos. Os resultados obtidos para o mutante V15NS16G não foram conclusivos, pois os efeitos das mutações variaram, de redução a aumento, no direcionamento da proteína para mitocôndrias. A informação para o duplo direcionamento foi localizada entre os 17 primeiros aminoácidos da região amino-terminal da seqüência de direcionamento de RBP1b. Os sinais de direcionamento para cloroplastos se distribuíram ao longo da seqüência de direcionamento. Enquanto a metade amino-terminal da seqüência foi suficiente para determinar o duplo direcionamento, a seqüência compreendendo os 13 aminoácidos

seguintes afetaram a eficiência do transporte. O método aplicado para a medida dos sinais de fluorescência da GFP, na análise das imagens obtidas em microscopia confocal, mostrou-se eficiente para a análise quantitativa *in vivo* do direcionamento de proteínas para mitocôndrias e cloroplastos, constituindo nova alternativa para esses estudos.

(iii) A investigação da existência de um sítio interno de entrada do ribossomo (IRES), na seqüência do mRNA de THI1 em *A. thaliana*, o que possibilitaria a explicação para o início da tradução no segundo AUG, não pôde ser concluída. Apesar de inúmeras tentativas, não foi possível encontrar um sistema de expressão transiente capaz de reproduzir dados da literatura, que mostraram o direcionamento da proteína para mitocôndrias e cloroplastos. Nos sistemas testados, THI1 foi encontrada somente em cloroplastos.

REFERÊNCIAS

ABDALLAH, F.; SALAMINI, F.; LEISTER, D. A prediction of the size and evolutionary origin of the proteome of chloroplasts of Arabidopsis. **Trends Plant Science**, Cambridge, v. 5, n. 4, p. 141-142, 2000.

ALMEIDA, J.D. Origem, evolução e direcionamento da proteína TH11 em plantas. 2004. 94p. Tese (Doutorado em Genética e Melhoramento de Plantas) – Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, 2004.

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

AMBARD-BRETTEVILLE, F.; SMALL, I.; GRANDJEAN, O.; COLAS DES FRANCS-SMALL, C. Discrete mutations in the presequence of potato formate dehydrogenase inhibit the in vivo targeting of GFP fusions into mitochondria. **Biochemical and Biophysical Research Communications**, New York, v. 311, n. 4, p. 966-971, 2003.

THE ARABIDOPSIS INFORMATION RESOURCE DATABASE (TAIR). Disponível em: <<http://www.arabidopsis.org>>. Acesso em: 2 jan. 2006.

ARIZONA GENOMICS INSTITUTE. Disponível em: <<http://genome.arizona.edu>>. Acesso em: 2 mar. 2005.

BABIYCHUK, E.; MÜLLER, F.; EUBEL, H.; BRAUN, H.P.; FRENTZEN, M.; KUSHNIR, S. Arabidopsis phosphatidylglycerophosphate synthase 1 is essential for chloroplast differentiation, but is dispensable for mitochondrial function. **The Plant Journal**, Oxford, v. 33, n. 5, p. 899-909, 2003.

BALSERA, M.; STENGEL, A.; SOLL, J.; BÖLTER, B. Tic62: a protein family from metabolism to protein translocation. **BMC Evolutionary Biology**, London, v. 7, n. 43, 2007. Disponível em: <<http://www.biomedcentral.com/1471-2148/7/43>>. Acesso em: 16 dez. 2007.

BANNAI, H.; TAMADA, Y.; MARUYAMA, O.; NAKAI, K.; MIYANO, S. Extensive feature detection of N-terminal protein sorting signals. **Bioinformatics**, Oxford, v. 18, n. 2, p. 298-305, 2002.

BEARDSLEE, T.A.; ROY-CHOWDHURY, S.; JAISWAL, P.; BUHOT, L.; LERBS-MACHE, S.; STERN, D.B.; ALLISON, L.A. A nucleus-encoded maize protein with sigma factor activity accumulates in mitochondria and chloroplasts. **The Plant Journal**, Oxford, v. 31, n. 2, p.199-209, 2002.

BEDDOE, T.; LITHGOW, T. Delivery of nascent polypeptides to the mitochondrial surface. **Biochimica Biophysica Acta**, v. 1592, n. 1, p. 35-39, 2002.

BELANGER, F.C.; LEUSTEK, T.; CHU, B.; KRIS, A.L. Evidence for thiamine biosynthetic pathway in higher plant plastids and its development regulation. **Plant Molecular Biology**, Dordrecht, v. 29, n. 4, p. 340-342, 1995.

BHUSHAN, S.; STAHL, A.; NILSSON, S.; LEFEBVRE, B.; SEKI, M.; ROTH, C.; MCWILLIAM, D.; WRIGHT, S.J.; LIBERLES, D.A.; SHINOZAKI, K.; BRUCE, B.D.; BOUTRY, M.; GLASER, E. Catalysis, subcellular localization, expression and evolution of the targeting peptides degrading protease, AtPreP2. **Plant Cell Physiology**, Oxford, v. 46, n. 6, p. 985-996, 2005.

BRADFORD, M.M. A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. **Analytical Biochemistry**, New York, v. 72, p. 248-254, 1976.

BRUCE, B. Chloroplast transit peptides: structure, function and evolution. **Trends in Cell Biology**, Cambridge, v.10, n. 10, p.440-447, 2000.

BRUCE, B. The paradox of plastids transit peptides: conservation of function despite divergence in primary structure. **Biochimica et Biophysica Acta**, Amsterdam, v.1541, n. 1-2, p.2-21, 2001.

CHABREGAS, M.S.; LUCHE, D.D.; FARIAS, L.P.; RIBEIRO, A.F.; VAN SLUYS, M.-A.; MENCK, C.F.M.M.; 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, n.6, p. 639-650, 2001.

CHABREGAS, M.S.; LUCHE, D.D.; VAN SLUYS, M.-A.; MENCK, C.F.M.; 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, n. 2, p. 285-291, 2003.

CHEN, Y.; PEUMANS, W.J.; HAUSE, B.; BRAS, J.; KUMAR, M.; PROOST, P.; BARRE, A.; ROUGÉ, P.; VAN DAMME, E.J. Jasmonic acid methyl ester induces the synthesis of a cytoplasmic/nuclear chito-oligosaccharide binding lectin in tobacco leaves. **Federation of American Societies for Experimental Biology Journal**, Bethesda, v. 16, n. 8, p. 905-907, 2002.

CHEW, O.; RUDHE, C.; GLASER, E.; WHELAN, J. Characterization of the targeting signal of dual-targeted pea glutathione reductase. **Plant Molecular Biology**, Dordrecht, v. 53, n. 3, p. 341-356, 2003.

CHEW, O.; WHELAN, J.; MILLAR, A.H. Molecular definition of the ascorbate-glutathione cycle in Arabidopsis mitochondria reveals dual targeting of antioxidant defenses in plants. **Journal of Biological Chemistry**, Baltimore, v. 278, n. 47, p. 46869-46877, 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. **Federation of European Biochemical Societies Letters**, Amsterdam, v. 557, n. 1-3, p. 109-114, 2004.

CHEW, O.; WHELAN, J. Just read the message: a model for sorting of proteins between mitochondria and chloroplasts. **Trends in Plant Science**, Cambridge, v. 9, n. 7, p. 318-319, 2004.

CHO, H.S.; LEE, S.S.; KIM, K.D.; HWANG, I.; LIM, J.S.; PARK, Y.I.; PAI, H.S. DNA gyrase is involved in chloroplast nucleoid partitioning. **The Plant Cell**, Baltimore, v. 16, n. 10, p. 2665-2682, 2004.

CHOW, K.S.; SINGH, D.P.; WALKER, A.R.; SMITH, A.G. Two different genes encode ferrochelatase in Arabidopsis: mapping, expression and subcellular targeting of the precursor proteins. **The Plant Journal**, Oxford, v. 15, n. 4, p. 531-541, 1998.

CHRISTENSEN, A.C.; LYZNIK, A.; MOHAMMED, S.; ELOWSKY, C.G.; ELO, A.; YULE, R.; MACKENZIE, S.A. Dual-domain, dual-targeting organellar protein presequences in Arabidopsis can use non-AUG start codons. **The Plant Cell**, Baltimore, v. 17, n. 10, p. 2805-2816, 2005.

CLEARY, S.P.; TAN, F.C.; NAKRIEKO, K.A.; THOMPSON, S.J.; MULLINEAUX, P.M.; CREISSEN, G.P.; VON STEDINGK, E.; GLASER, E.; SMITH, A.G.; ROBINSON, C. Isolated plant mitochondria import chloroplast precursor proteins in vitro with the same efficiency as chloroplasts. **Journal of Biological Chemistry**, Baltimore, v. 277, n. 7, p. 5562-5569, 2002.

COLLINS, T.J. ImageJ for microscopy. **Biotechniques**, Natick, v. 43, n. 1, p. 25-30, 2007.

COLOMBO, S.; LONGHI, R.; ALCARO, S.; ORTUSO, F.; SPROCATI, T.; FLORA, A.; BORGESSE, N. N-myristoylation determines dual targeting of mammalian NADH-cytochrome b5 reductase to ER and mitochondrial outer membranes by a mechanism of kinetic partitioning. **Journal of Cell Biology**, New York, v. 168, n. 5, p. 735-745, 2005.

CORONADO, M.J.; GONZÁLEZ-MELENDI, P.; SEGUÍ, J.M.; RAMÍREZ, C.; BÁRÁNY, I.; TESTILLANO, P.S.; RISUEÑO, M.C. MAPKs entry into the nucleus at specific interchromatin domains in plant differentiation and proliferation processes. **Journal of Structural Biology**, Duluth, v. 140, n. 1-3, p. 200-213, 2002.

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**, Oxford, v. 8, n. 2, p. 167-175, 1995.

CROOKS, G.E.; HON, G.; CHANDONIA, J.M.; BRENNER, S.E. WebLogo: a sequence logo generator. **Genome Research**, Woodbury, v. 14, n. 6, p. 1188-1190, 2004.

CUMMINGS, M.P.; NUGENT, J.M.; OLMSTEAD, R.G.; PALMER, J.D. Phylogenetic analysis reveals five independent transfers of the chloroplast gene *rbcl* to the mitochondrial genome in angiosperms. **Current Genetics**, Berlin, v. 43, n. 2, p. 131-138, 2003.

DAVIS, A.J.; RYAN, K.R.; JENSEN, R.E. Tim23p contains separate and distinct signals for targeting to mitochondria and insertion into the inner membrane. **Molecular Biology of the Cell**, Bethesda, v. 9, n. 9, p. 2577-2593, 1998.

DI COLA, A.; FRIGERIO, L.; LORD, J.M.; CERIOTTI, A.; ROBERTS, L.M. Ricin A chain without its partner B chain is degraded after retrotranslocation from the endoplasmic reticulum to the cytosol in plant cells. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 98, n. 25, p. 14726-14731, 2001.

DINKINS, R.D.; CONN, H.M.; DIRK, L.M.A.; WILLIAMS, M.A.; HOUTZ, R.L. The *Arabidopsis thaliana* peptide deformylase 1 protein is localized to both mitochondria and chloroplasts. **Plant Science**, Amsterdam, v. 165, n. 4, p. 751-758, 2003.

DOLEZAL, P.; LIKIC, V.; TACHEZY, J.; LITHGOW, T. Evolution of the molecular machines for protein import into mitochondria. **Science**, v. 313, n. 5785, p. 314-318, 2006.

DUBY, G.; DEGAND, H.; BOUTRY, M. Structure requirement and identification of a cryptic cleavage site in the mitochondrial processing of a plant F1-ATPase beta-subunit presequence. **Federation of European Biochemical Societies Letters**, Amsterdam, v. 505, n. 3, p. 409-413, 2001.

DUBY, G.; BOUTRY, M. Mitochondrial protein import machinery and targeting information. **Plant Science**, Amsterdam, v. 162, n. 4, p. 477-490, 2002.

DUCHÊNE, A.M.; GIRITCH, A.; HOFFMANN, B.; COGNAT, V.; LANCELIN, D.; PEETERS, N.M.; ZAEPFEL, M.; MARECHAL-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, n. 45, p. 16484-16489, 2005.

ELO, A.; LYZNIK, A.; GONZALEZ, D.O.; KACHMAN, S.D.; MACKENZIE, S.A. Nuclear genes that encode mitochondrial proteins for DNA and RNA metabolism are clustered in the *Arabidopsis* genome. **The Plant Cell**, Baltimore, v. 15, n. 7, p. 1619-1631, 2003.

EMANUELSSON, O.; NIELSEN, H.; BRUNAK, S.; VON HEIJNE, G. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. **Journal of Molecular Biology**, London, v. 300, n. 4, p. 1005-1016, 2000.

FERRO, M.; SALVI, D.; BRUGIÈRE, S.; MIRAS, S.; KOWALSKI, S.; LOUWAGIE, M.; GARIN, J.; JOYARD, J.; ROLLAND, N. Proteomics of the chloroplast envelope membranes from *Arabidopsis thaliana*. **Molecular and Cellular Proteomics**, Rockville, v. 2, n. 5, p. 325-345, 2003.

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

GOGGIN, D.E.; LIPSCOMBE, R.; FEDOROVA, E.; MILLAR, A.H.; MANN, A.; ATKINS, C.A.; SMITH, P.M. Dual intracellular localization and targeting of aminoimidazole ribonucleotide synthetase in cowpea. **Plant Physiology**, Rockville, v. 131, n. 3, p. 1033-1041, 2003.

HEAZLEWOOD, J.L.; TONTI-FILIPPINI, J.S.; GOUT, A.M.; DAY, D.A.; WHELAN, J.; MILLAR, A.H. Experimental analysis of the Arabidopsis mitochondrial proteome highlights signaling and regulatory components, provides assessment of targeting prediction programs, and indicates plant-specific mitochondrial proteins. **The Plant Cell**, Baltimore, v. 16, n. 1, p. 241–256, 2004.

HEAZLEWOOD, J.L.; MILLAR, A.H. AMPDB: the Arabidopsis Mitochondrial Protein Database. **Nucleic Acids Research**, London, v. 33, p. D605-D610, 2005.

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**, Rockville, v. 139, p. 598-609, 2005.

HEAZLEWOOD, J.L.; VERBOOM, R.E.; TONTI-FILIPPINI, J.; SMALL, I.; MILLAR, A.H. SUBA: the Arabidopsis Subcellular Database. **Nucleic Acids Research**, London, v. 35, p. D213 – D218, 2007.

HEDTKE, B.; MEIXNER, M.; GILLANDT, S.; RICHTER, E.; BORNER, T.; WEIHE, A. Green fluorescent protein as a marker to investigate targeting of organellar RNA polymerases of higher plants *in vivo*. **The Plant Journal**, Oxford, v. 17, n. 5, p. 557-561, 1999.

HEDTKE, B.; BORNER, T.; WEIHE, A. One RNA polymerase serving two genomes. **European Molecular Biology Organization Reports**, Oxford, v. 1, n. 5, p. 435-440, 2000.

HEDTKE, B.; LEGEN, J.; WEIHE, A.; HERRMANN, R.G.; BÖRNER, T. Six active phage-type RNA polymerase genes in *Nicotiana tabacum*. **The Plant Journal**, Oxford, v. 30, n. 6, p. 625-637, 2002.

HERMAN, P.L.; RAMBERG, H.; BAACK, R.D.; MARKWELL, J.; OSTERMAN, J.C. Formate dehydrogenase in *Arabidopsis thaliana*: overexpression and subcellular localization in leaves. **Plant Science**, Amsterdam, v. 163, p. 1137-1145, 2002.

HÖGLUND, A.; DÖNNES, P.; BLUM, T.; ADOLPH, H.W.; KOHLBACHER, O. MultiLoc: prediction of protein subcellular localization using N-terminal targeting sequences, sequence motifs and amino acid composition. **Bioinformatics**, Oxford, v. 22, n. 10, p. 1158-1165, 2006.

HOOGENRAAD, N.J.; WARD, L.A.; RYAN, M.T. Import and assembly of proteins into mitochondria of mammalian cells. **Biochimica et Biophysica Acta**, Amsterdam, v. 592, n. 1, p. 97-105, 2002.

HUANG, C.Y.; AYLIFFE, M.A.; TIMMIS, J.N. Direct measurement of the transfer rate of chloroplast DNA into the nucleus. **Nature**, London, v. 422, n. 6927, p. 72-76, 2003.

HUANG, J.; HACK, E.; THORNBURG, R.W.; MYERS, A.M. A yeast mitochondrial leader peptide functions *in vivo* as a dual targeting signal for both chloroplasts and mitochondria. **The Plant Cell**, Baltimore, v. 2, n. 12, p. 1249-1260, 1990.

HURT, E.C.; GOLDSCHMIDT-CLERMONT, M.; PESOLD-HURT, B.; ROCHAIX, J.D.; SCHATZ, G. A mitochondrial presequence can transport a chloroplast-encoded protein into yeast mitochondria. **Journal of Biological Chemistry**, Baltimore, v. 261, n. 25, p. 11440-11443, 1986.

HUSON, D.H.; BRYANT, D. Application of phylogenetic networks in evolutionary studies. **Molecular Biology and Evolution**, Chicago, v. 23, n. 2, p. 254-267, 2006.

THE INSTITUTE FOR GENOMIC RESEARCH (TIGR). Disponível em:
<<http://tigrblast.tigr.org>>. Acesso em: 1 jan. 2006.

JAQUINOD, M.; VILLIERS, F.; KIEFFER-JAQUINO, D.S.; HUGOUVIEUX, V.; BRULEY, C.; GARIN, J.; BOURGUIGNON, J. A proteomics dissection of *Arabidopsis thaliana* vacuoles isolated from cell culture. **Molecular and Cellular Proteomics**, Rockville, v. 6, n. 3, p. 394-412, 2007.

JOHNSON, A.E.; VAN WAES, M.A. The translocon: a dynamic gateway at the ER membrane. **Annual Review of Cell and Developmental Biology**, Palo Alto, v. 15, p. 799-842, 1999.

JORGENSEN, R. Sequencing maize: just sample the salsa or go for the whole enchilada. **The Plant Cell**, Baltimore, v. 16, n. 4, p. 787-788, 2004.

KADOWAKI, K.; KUBO, N.; OZAWA, K.; HIRAI, A. Targeting presequence acquisition after mitochondrial gene transfer to the nucleus occurs by duplication of existing targeting signals. **European Molecular Biology Organization Journal**, Oxford v.15, n. 23, p. 6652-6661, 1996.

KIM, Y.S.; NASAKA, K.; DOWNS, D.M.; KWAK, J.M.; PARK, D.; CHUNG, I.K.; NAM, H.G. A Brassica cDNA clone encoding a bifunctional hydroxymethylpyrimidine kinase/thiamin-phosphate pyrophosphorylase involved in thiamin biosynthesis. **Plant Molecular Biology**, Dordrecht, v. 37, n. 6, p. 955-966, 1998.

KLEINE, T.; LOCKHART, P.; BATSCHAUER, A. An Arabidopsis protein closely related to Synechocystis cryptochrome is targeted to organelles. **The Plant Journal**, Oxford, v. 35, n. 1, p. 93-103, 2003.

KOBAYASHI, Y.; DOKIYA, Y.; SUGITA, M. Dual targeting of phage-type RNA polymerase to both mitochondria and plastids is due to alternative translation initiation in single transcripts. **Biochemical and Biophysical Research Communication**, New York, v. 289, n. 5, p. 1106-1113, 2001.

KOBAYASHI, Y.; DOKIYA, Y.; KUMAZAWA, Y.; SUGITA, M. Non-AUG translation initiation of mRNA encoding plastid-targeted phage-type RNA polymerase in *Nicotiana sylvestris*. **Biochemical and Biophysical Research Communication**, New York, v. 299, n. 1, p. 57-61, 2002.

KÖHLER, R.H.; CAO, J.; ZIPFEL, W.R.; WEBB, W.W.; HANSON, M.R. Exchange of protein molecules through connections between higher plant plastids. **Science**, Washington, n. 5321, v. 276, p. 2039-2042, 1997.

KOMAR, A.A.; HATZOGLOU, M. Internal ribosome entry sites in cellular mRNAs: mystery of their existence. **Journal of Biological Chemistry**, Baltimore, v. 280, n. 25, 23425-23428, 2005.

KOZAK, M. Structural features in eukaryotic mRNAs that modulate the initiation of translation. **Journal of Biological Chemistry**, Baltimore, v. 266, p. 19867-19870, 1991.

KOZAK, M. Pushing the limits of the scanning mechanism for initiation of translation. **Gene**, Amsterdam, v. 299, n. 1-2, 1-34, 2002.

KUBO, N.; HARADA, K.; HIRAI, A.; KADOWAKI, K. A single nuclear transcript encoding mitochondrial RPS14 and SDHB of rice is processed by alternative splicing: common use of the same mitochondrial targeting signal for different proteins. **Proceedings of the National Academy of Sciences of the United States of America**, Washington v. 96, n. 16, p. 9207-9211, 1999.

KUBO, N.; JORDANA, X.; OZAWA, K.; ZANLUNGO, S.; HARADA, K.; SASAKI, T.; KADOWAKI, K. Transfer of the mitochondrial rps10 gene to the nucleus in rice: acquisition of the 5' untranslated region followed by gene duplication. **Molecular and General Genetics**, Berlin, v. 263, n. 4, p. 733-739, 2000.

KUSUMI, K.; YARA, A.; MITSUI, N.; TOZAWA, Y.; IBA, K. Characterization of a rice nuclear-encoded plastid RNA polymerase gene OsRpoTp. **Plant Cell Physiology**, Oxford, v. 45, p. 1194-1201, 2004.

LAEMMLI, U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. **Nature**, London, v. 227, n. 5259, p. 680-685, 1970.

LAN L.; ISENMANN, S.; WATTENBERG, B.W. Targeting and insertion of C-terminally anchored proteins to the mitochondrial outer membrane is specific and saturable but does not strictly require ATP or molecular chaperones. **Biochemical Journal**, London, v. 349, n. 2, p. 611-621, 2000.

LEISTER, D. Origin, evolution and genetic effects of nuclear insertions of organelle DNA. **Trends in Genetics**, Cambridge, v. 21, n. 12, p. 655-663, 2005.

LIN, X.; KAUL, S.; ROUNSLEY, S.; SHEA, T.P.; BENITO, M.I.; TOWN, C.D.; FUJII, C.Y.; MASON, T.; BOWMAN, C.L.; BARNSTEAD, M.; FELDBLYUM, T.V.; BUELL, C.R.; KETCHUM, K.A.; LEE, J.; RONNING, C.M.; KOO, H.L.; MOFFAT, K.S.; CRONIN, L.A.; SHEN, M.; PAI, G.; VAN AKEN, S.; UMayAM, L.; TALLON, L.J.; GILL, J.E.; ADAMS, M.D.; CARRERA, A.J.; CREASY, T.H.; GOODMAN, H.M.; SOMERVILLE, C.R.; COPENHAVER, G.P.; PREUSS, D.; NIERMAN, W.C.; WHITE, O.; EISEN, J.A.; SALZBERG, S.L.; FRASER, C.M.; VENTER, J.C. Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. **Nature**, London, v. 402, n. 6763, p. 761-780, 1999.

LISENBEE, C.S.; LINGARD, M.J.; TRELEASE, R.N. Arabidopsis peroxisomes possess functionally redundant membrane and matrix isoforms of monodehydroascorbate reductase. **The Plant Journal**, Oxford, v. 43, n. 6, p. 900-914, 2005.

LITHGOW, T.; HØJ, P.B.; HOOGENRAAD, N.J. Do cytosolic factors prevent promiscuity at the membrane surface? **Federation of European Biochemical Societies Letters**, Amsterdam, v. 329, n. 1-2, p. 1-4, 1993.

LONG, M.; DE SOUZA, S.J.; ROSENBERG, C.; GILBERT, W. Exon shuffling and the origin of the mitochondrial targeting function in plant cytochrome c1 precursor. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 93, n. 15, p. 7727-7731, 1996.

LUCATTINI, R.; LIKIC, V.A.; LITHGOW, T. Bacterial proteins predisposed for targeting to mitochondria. **Molecular Biology and Evolution**, Chicago, v. 21, n. 4, p. 652-658, 2004.

LYNCH, M.; CONERY, J.S. The evolutionary fate and consequences of duplicate genes. **Science**, Washington DC, v. 290, n. 5494, p. 1151-1155, 2000.

MAĆASEV, D.; WHELAN, J.; NEWBIGIN, E.; SILVA-FILHO, M.C.; MULHERN, T.D.; LITHGOW, T. Tom22', 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, n. 8, p. 1557-1564, 2004.

MACHADO, C.R.; COSTA DE OLIVEIRA, R.L.; BOITEUX, S.; PRAEKELT, U.M.; MEACOCK, P.A.; MENCK, C.F.M. *Thi1*, a thiamine biosynthetic gene in *Arabidopsis thaliana*, complements defects in DNA repair. **Plant Molecular Biology**, Dordrecht, v.31, n. 3, p.585-593, 1996.

MACKENZIE, S.A. Plant organellar protein targeting: a traffic plan still under construction. **Trends in Cell Biology**, Cambridge, v. 15, n. 10, p. 548-554, 2005.

MACHADO, C.R.; PRAEKELT, U.M.; COSTA DE OLIVEIRA, R.L.; BARBOSA, A.C.C.; BYRNE, K.L.; MEACOCK, P.A.; MENCK, C.F.M. Dual role for the yeast TH14 gene in thiamine biosynthesis and DNA damage tolerance. **Journal of Molecular Biology**, London, v. 273, n. 1, p.114-121, 1997.

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

MARTIN, W.; HERRMANN, R.G. Gene transfer from organelles to the nucleus: how much, what happens, and why? **Plant Physiology**, Rockville, v. 118, n. 1, p. 9-17, 1998.

MARTIN, W.; RUJAN, T.; RICHLI, E.; HANSEN, A.; CORNELSEN, S.; LINS, T.; LEISTER, D.; STOEBE, B.; HASEGAWA, M.; PENNY, D. Evolutionary analysis of Arabidopsis, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 99, n. 19, p.12246-12251, 2002.

MAY, T.; SOLL, J. 14-3-3 proteins form a guidance complex with chloroplast precursor proteins in plants. **Plant Cell**, Baltimore, v. 12, n. 1, p. 53-64, 2000.

MEDINA-SILVA, R.; BARROS, M.P.; GALHARDO, R.S.; NETTO, L.E.; COLEPICOLO, P.; MENCK, C.F. Heat stress promotes mitochondrial instability and oxidative responses in yeast deficient in thiazole biosynthesis. **Research in Microbiology**, Paris, v. 157, n. 3, p. 275-281, 2006.

MILENKOVIC, D.; MÜLLER, J.; STOJANOVSKI, D.; PFANNER, N.; CHACINSKA, A. Diverse mechanisms and machineries for import of mitochondrial proteins. **Biological Chemistry**, Berlin, v. 388, n. 9, p. 891-897, 2007.

MILLAR, A.H.; HEAZLEWOOD, J.L.; KRISTENSEN, B.K.; BRAUN, H.P.; MØLLER, I.M. The plant mitochondrial proteome. **Trends in Plant Science**, Kidlington, v. 10, n. 1, p. 36-43, 2005.

MILLAR, A.H.; WHELAN, J.; SMALL, I. Recent surprises in protein targeting to mitochondria and plastids. **Current Opinion in Plant Biology**, London, v. 9, n. 6, p. 610-615, 2006.

MORGANTE, C.V.; GUALBERTO, J.M. Dual targeting of RNA-binding proteins to mitochondria and chloroplasts. In: TRI-NATIONAL ARABIDOPSIS MEETING, 4., Viena. **Resumos...** Viena: BOKU, 2007. Resumo 131.

MULLEN, R.T.; TRELEASE, R.N. The sorting signals for peroxisomal membrane-bound ascorbate peroxidase are within its C-terminal tail. **Journal of Biological Chemistry**, Baltimore, v. 275, n. 21, p. 16337-16344, 2000.

MURCHA, M.W.; ELHAFEZ, D.; LISTER, R.; TONTI-FILIPPINI, J.; BAUMGARTNER, M.; PHILIPPAR, K.; CARRIE, C.; MOKRANJAC, D.; SOLL, J.; WHELAN, J. Characterization of the preprotein and amino acid transporter gene family in Arabidopsis. **Plant Physiology**, Rockville, v. 143, n. 1, p. 199-212, 2007.

NAKAMURA, T.; YAMAGUCHI, Y.; SANO, H. Plant mercaptopyruvate sulfurtransferases: molecular cloning, subcellular localization and enzymatic activities. **European Journal of Biochemistry**, Oxford, v. 267, p. 5621-5630, 2000.

NAVROT, N.; COLLIN, V.; GUALBERTO, J.; GELHAYE, E.; HIRASAWA, M.; REY, P.; KNAFF, D.B.; ISSAKIDIS, E.; JACQUOT, J.P.; ROUHIER, N. Plant glutathione peroxidases are functional peroxiredoxins distributed in several subcellular compartments and regulated during biotic and abiotic stresses. **Plant Physiology**, Rockville, v. 142, n. 4, p. 1364-1379, 2006.

NOOTSOS, C.; RICHLY, E.; LEISTER, D. Generation and evolutionary fate of insertions of organelle DNA in the nuclear genomes of flowering plants. **Genome Research**, Woodbury, v. 15, n. 5, p. 616-628, 2005.

NOOTSOS, C.; KLEINE, T.; ARMBRUSTER, U.; DALCORSO, G.; LEISTER, D. Nuclear insertions of organellar DNA can create novel patches of functional exon sequences. **Trends in Genetics**, Cambridge, v. 23, n. 12, p. 597-601, 2007.

OBARA, K.; SUMI, K.; FUKUDA, H. The use of multiple transcription starts causes the dual targeting of Arabidopsis putative monodehydroascorbate reductase to both mitochondria and chloroplasts. **Plant Cell Physiology**, Oxford, v. 43, n. 7, p. 697-705, 2002.

ONO, Y.; SAKAI, A.; TAKECHI, K.; TAKIO, S.; TAKUSAGAWA, M.; TAKANO, H. *NtPoll-like1* and *NtPoll-like2*, bacterial DNA polymerase I homologues isolated from BY-2 cultured tobacco cells, encode DNA polymerases engaged in DNA replication in both plastids and mitochondria. **Plant and Cell Physiology**, Tokyo, v. 48, n. 12, p. 1679-1692, 2007.

PAPINI-TERZI, F.S.; GALHARDO, R.S.; FARIAS, L.P.; MENCK, C.F.M.; VAN SLUYS, M.-A. Point Mutation is responsible for Arabidopsis tz-201 mutant phenotype affecting thiamin biosynthesis. **Plant Cell Physiology**, Oxford, v.44, n.8, p.856-860, 2003.

PEETERS, N.; SMALL, I. Dual targeting to mitochondria and chloroplasts. **Biochimica et Biophysica Acta**, Amsterdam, v. 54, p. 154-163, 2001.

PELTIER, J.B.; YTTERBERG, A.J.; SUN, Q.; VAN WIJK, K.J. New functions of the thylakoid membrane proteome of *Arabidopsis thaliana* revealed by a simple, fast, and versatile fractionation strategy. **Journal of Biological Chemistry**, Baltimore, v. 279, n. 47, 2004.

PELTIER, J.B.; CAI, Y.; SUN, Q.; ZABROUSKOV, V.; GIACOMELLI, L.; RUDELLA, A.; YTTERBERG, A.J.; RUTSCHOW, H.; VAN WIJK, K.J. The oligomeric stromal proteome of *Arabidopsis thaliana* chloroplasts. **Molecular and Cellular Proteomics**, Rockville, v. 5, n. 1, p.114-133, 2006.

PENDLE, A.F.; CLARK, G.P.; BOON, R.; LEWANDOWSKA, D.; LAM, Y.W.; ANDERSEN, J.; MANN, M.; LAMOND, A.I.; BROWN, J.W.; SHAW, P.J. Proteomic analysis of the Arabidopsis nucleolus suggests novel nucleolar functions. **Molecular Biology of the Cell**. Bethesda, v. 16, n. 1, p. 260-269, 2005.

PERRY, A.J.; HULETT, J.M.; LIKIĆ, V.A.; LITHGOW, T.; GOOLEY, P.R. Convergent evolution of receptors for protein import into mitochondria. **Current Biology**, London, v. 16, n. 3, p. 221-229, 2006.

PFANNER, N.; GEISSLER, A. Versatility of the mitochondrial protein import machinery. **Nature Reviews Molecular Cell Biology**, London, v. 2, n. 5, p. 339-349, 2001.

POLCICOVÁ, K.; KEMPNÁ, P.; SABOVÁ, L.; GAVURNÍKOVÁ, G.; POLCIC, P.; KOLAROV, J. The delivery of ADP/ATP carrier protein to mitochondria probed by fusions with green fluorescent protein and beta-galactosidase. **FEMS Yeast Research**, Oxford, v. 4, n. 3, p. 315-321, 2003.

PUJOL, C.; MARÉCHAL-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, n. 2, p. 356-367, 2007.

PUYAUBERT, J.; DENIS, L.; ALBAN, C. Dual-targeting of *Arabidopsis thaliana* Holocarboxylase Synthetase 1: a small upstream open reading frame (uORF) regulates translation initiation and protein targeting. **Plant Physiology**, Rockville. In press.

REUMANN, S.; BABUJEE, L.; MA, C.; WIENKOOP, S.; SIEMSEN, T.; ANTONICELLI, G.E.; RASCHE, N.; LÜDER, F.; WECKWERTH, W.; JAHN, O. Proteome Analysis of Arabidopsis Leaf Peroxisomes Reveals Novel Targeting Peptides, Metabolic Pathways, and Defense Mechanisms. **The Plant Cell**, Baltimore, v. 19, n. 10, p. 3170-3193, 2007.

RIBEIRO, A.; PRAEKELT, U.; AKKERMANS, A.D.L.; MEACOCK, P.A.; VAN KAMMEN, A.; BISSELING, T.; PAWLOWSKI, K. Identification of *agthi1*, whose product is involved in biosynthesis of the thiamine precursor thiazole, in actinorhizal nodules of *Alnus glutinosa*. **The Plant Journal**, Oxford, v. 10, n. 2, p. 361-368, 1996.

RIBEIRO, D.T.; FARIAS, L.P.; DE ALMEIDA, J.D.; KASHIWABARA, P.M.; RIBEIRO, A. F.; SILVA-FILHO, M.C.; MENCK, C.F.; VAN SLUYS, M.A. Functional characterization of the *thi1* promoter region from *Arabidopsis thaliana*. **Journal of Experimental Botany**, Oxford, v. 56, n. 417, p. 1797-1804, 2005.

RICHTER, U.; KIESSLING, J.; HEDTKE, B.; DECKER, E.; RESKI, R.; BÖRNER, T.; WEIHE, A. Two RpoT genes of *Physcomitrella patens* encode phage-type RNA polymerases with dual targeting to mitochondria and plastids. **Gene**, Amsterdam, v. 290, n. 1-2, p. 95-105, 2002.

ROBINSON, C.; WOOLHEAD, C.; EDWARDS, W. Transport of proteins into and across the thylakoid membrane. **Journal of Experimental Botany**, Oxford, v. 51, Spec No, p. 369-374, 2000.

RODRÍGUEZ-CONCEPCIÓN, M.; YALOVSKY, S.; ZIK, M.; FROMM, H.; GRUISSEM, W. The prenylation status of a novel plant calmodulin directs plasma membrane or nuclear localization of the protein. **European Molecular Biology Organization Journal**, Oxford, v. 18, n. 7, p. 1996-2007, 1999.

ROJO, E.; ZOUHAR, J.; KOVALEVA, V.; HONG, S.; RAIKHEL, N.V. The AtC-VPS protein complex is localized to the tonoplast and the prevacuolar compartment in Arabidopsis. **Molecular Biology of the Cell**, Bethesda, v. 14, n. 2, p. 361-369, 2003.

RUDHE, C.; CLIFTON, R.; WHELAN, J.; GLASER, E. N-terminal domain of the dual-targeted pea glutathione reductase signal peptide controls organellar targeting efficiency. **Journal of Molecular Biology**, London, v. 324, n. 4, p. 577-585, 2002a.

RUDHE, C.; CHEW, O.; WHELAN, J.; GLASER, E. A novel *in vitro* system for simultaneous import of precursor proteins into mitochondria and chloroplasts. **The Plant Journal**, Oxford, v. 30, n. 2, p. 213-220, 2002b.

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, n. 3, p. 639-647, 2004.

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

SAKAMOTO, K.; BRIGGS, W.R. Cellular and subcellular localization of phototropin 1. **The Plant Cell**, Baltimore, v. 14, n. 8, p. 1723-1735, 2002.

SAMBROOK, J.; RUSSEL, D.W. **Molecular cloning: a laboratory manual**. 3.ed. New York: Cold Spring Harbor Laboratory, 2001, 3v.

SANDOVAL, P.; LEÓN, G.; GÓMEZ, I.; CARMONA, R.; FIGUEROA, P.; HOLUIGUE, L.; ARAYA, A.; JORDANA, X. Transfer of RPS14 and RPL5 from the mitochondrion to the nucleus in grasses. **Gene**, Amsterdam, v. 324, p. 139-147, 2004.

SASS, E.; KARNIELY, S.; PINES, O. Folding of fumarase during mitochondrial import determines its dual targeting in yeast. **Journal of Biological Chemistry**, Baltimore, v. 278, n. 46, p. 45109-45116, 2003.

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, n. 6, p. 631-646, 2007.

SHIU, S.H.; SHIH, M.C.; LI, W.H. Transcription factor families have much higher expansion rates in plants than in animals. **Plant Physiology**, Rockville, v. 139, n. 1, p. 18-26, 2005.

SILVA-FILHO, M.D.; WIEËRS, M.C.; FLÜGGE, U.I.; CHAUMONT, F.; BOUTRY, M. Different *in vitro* and *in vivo* targeting properties of the transit peptide of a chloroplast envelope inner membrane protein. **Journal of Biological Chemistry**, Baltimore, v. 272, n. 24, p. 15264-15269, 1997

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

SMALL, I.; PEETERS, N.; LEGEAI, F.; LURIN, C. Predotar: A tool for rapidly screening proteomes for N-terminal targeting sequences. **Proteomics**, Weinheim, v. 4, n. 6, p. 1581-1590, 2004.

SPARKES, I.A.; BAKER, A. Peroxisome biogenesis and protein import in plants, animals and yeasts: enigma and variations? **Molecular Membrane Biology**, Abingdon, v. 19, n. 3, p. 171-185, 2002.

STEGEMANN, S.; HARTMANN, S.; RUF, S.; BOCK, R. High-frequency gene transfer from the chloroplast genome to the nucleus. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 100, n. 15, p. 8828-8833, 2003.

STENGEL, A.; SOLL, J.; BÖLTER, B. Protein import into chloroplasts: new aspects of a well-known topic. **Biological Chemistry**, Berlin, v. 388, n. 8, p. 765-772, 2007.

SUN, Q.; EMANUELSSON, O.; VAN WIJK, K.J. Analysis of curated and predicted plastid subproteomes of *Arabidopsis*. Subcellular compartmentalization leads to distinctive proteome properties. **Plant Physiology**, Rockville, v. 135, n. 2, p. 723-734, 2004.

SUNDERLAND, P.A.; WEST, C.E.; WATERWORTH, W.M.; BRAY, C.M. An evolutionarily conserved translation initiation mechanism regulates nuclear or mitochondrial targeting of DNA ligase 1 in *Arabidopsis thaliana*. **The Plant Journal**, Oxford, v. 47, n. 3, p. 356-367, 2006.

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**, Baltimore, v. 16, n. 8, p. 2048-2058, 2004.

TEIXEIRA, F.K.; MENEZES-BENAVENTE, L.; GALVÃO, V.C.; MARGIS, R.; MARGIS-PINHEIRO, M. Rice ascorbate peroxidase gene family encodes functionally diverse isoforms localized in different subcellular compartments. **Planta**, Berlin, v. 224, n. 2, p. 300-314, 2006.

TERRY, L.J.; SHOWS, E.B.; WENTE, S.R. Crossing the nuclear envelope: hierarchical regulation of nucleocytoplasmic transport. **Science**, Cambridge, v. 318, n. 5855, p. 1412-1416, 2007.

THOMPSON, J.D.; HIGGINS, D.G.; GIBSON, T.J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. **Nucleic Acids Research**, London, v. 22, n. 22, p. 4673-4680, 1994.

TIMMIS, J.N.; AYLIFFE, M.A.; HUANG, C.Y.; MARTIN, W. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. **Nature Reviews Genetics**, London, v. 5, n. 2, p. 123-135, 2004.

UEDA, M.; FUJIMOTO, M.; ARIMURA, S.-I.; TSUTSUMI, N.; KADOWAKI, K.-I. Evidence for transit peptide acquisition through duplication and subsequent frameshift mutation of a preexisting protein gene in rice. **Molecular Biology and Evolution**, Chicago, v. 23, n. 12, p. 2405-2412, 2006.

UNSELD, M.; MARIENFELD, J.R.; BRANDT, P.; BRENNICKE, A. The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. **Nature Genetics**, London, v. 15, n. 1, p. 57-61, 1997.

VANDEPOELE, K.; SIMILLION, C.; VAN DE PEER, Y. Evidence that rice and other cereals are ancient aneuploids. **The Plant Cell**, Baltimore, v. 15, n. 9, p. 2192-2202, 2003.

VERMEL, M.; GUERMANN, B.; DELAGE, L.; GRIENENBERGER, J.-M.; MARÉCHAL-DROUARD, L.; GUALBERTO, J.M. A family of RRM-type RNA-binding proteins specific to plant mitochondria. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 99, n. 9, p. 5866-5871, 2002.

VON HEIJNE, G.; STEPPUHN, J.; HERRMANN, R.G. Domain structure of mitochondrial and chloroplast targeting peptides. **European Journal of Biochemistry**, Oxford, v.180, n. 3, p.535-545, 1989.

XU, X.M.; MOLLER, S.G. AtSufE is an essential activator of plastidic and mitochondrial desulfurases in Arabidopsis. **European Molecular Biology Organization Journal**, Oxford, v. 25, n. 4, p. 900-909, 2006.

WALL, M.K.; MITCHENALL, L.A.; MAXWELL, A. Arabidopsis thaliana DNA gyrase is targeted to chloroplasts and mitochondria. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 101, n. 20, p. 7821-7826, 2004.

WANG, G.; DING, X.; YUAN, M.; QIU, D.; LI, X.; XU, C.; WANG, S. Dual function of rice OsDR8 gene in disease resistance and thiamine accumulation. **Plant Molecular Biology**, Dordrecht, v. 60, n. 3, p. 437-449, 2006.

WATANABE, N.; CHE, F.S.; IWANO, M.; TAKAYAMA, S.; YOSHIDA, S.; ISOGAI, A. Dual targeting of spinach protoporphyrinogen oxidase II to mitochondria and chloroplasts by alternative use of two in-frame initiation codons. **Journal of Biological Chemistry**, Baltimore, v. 276, n. 23, p. 20474-20481, 2001.

YUAN, Q.; HILL, J.; HSIAO, J.; MOFFAT, K.; OUYANG, S.; CHENG, Z.; JIANG, J.; BUELL, C.R. Genome sequencing of a 239-kb region of rice chromosome 10L reveals a high frequency of gene duplication and a large chloroplast DNA insertion. **Molecular Genetics and Genomics**, Berlin, v. 267, n. 6, p. 713-720, 2002.

ZAEGEL, V.; GUERMANN, B.; LE RET, M.; ANDRÉS, C.; MEYER, D.; ERHARDT, M.; CANADAY, J.; GUALBERTO, J.M.; IMBAULT, P. The plant-specific ssDNA binding protein OSB1 is involved in the stoichiometric transmission of mitochondrial DNA in Arabidopsis. **The Plant Cell**, Baltimore, v. 18, n. 12, p. 3548-3563, 2006.

ZAR, J.H. **Biostatistical Analysis**. Englewood Cliffs: Prentice-Hall International, 1974. 620 p.

ZHAO, J.; ONDUKA, T.; KINOSHITA, J.Y.; HONSHO, M.; KINOSHITA, T.; SHIMAZAKI, K.; ITO, A. Dual subcellular distribution of cytochrome b5 in plant, cauliflower, cells. **Journal of Biochemistry**, Tokyo, v. 133, n. 1, p. 115-121, 2003.

ANEXOS

ANEXO A - Sequências de aminoácidos utilizadas para a confecção do gráfico (Figura 18) no programa Weblogo

>At1g50200

MRLVKAASLLISSTKPPSRVFYSSHLRRPFFSHFRFSSSSSTSSSVAVMPGSEPSETQWPAKRVRDITYVDFFRG
KGHKFW

>At5g22800

MNFSRVNLFDFPLRPILLSHPSSIFVSTRFVTRTSAGVSPSILLPRSTQSPQIIAKSSSVSVQPVEDAKEDYQSKD
VSG

>At4g17300

MAATFLPATSLRLTQNSTLRFLSFFTISNPSYSLFRPLRRRVLPFDAPANSRRRCFCTAVSESLGSGDGNKVE
SYEKR

>At4g33760

MSLLRRTLPLRPTRFLSATAISISNATNFFVVPKRTNPLPGTRRTFSSSPVAAASGDVVVKVPSPPSVLRWVSRT
ELCG

>At2g31170

MASSVLNLFKSCRPFPIRFSSLPKSQFRIQFPLRPGKETQLRRCFTTLSSLTDGGAPISGGKELWLHNTMSRKK
ELFKP

>At3g48110

MAILHFSPLIVSFLRPHASPRFFLLPRSLSQSPFLSRRRFHRTSAVSSAAVHHQSYRNPDDDVTRAVSVPTFQQ
AIQRL

>At5g64050

MASLVYGPWLRVRSLELAPAFLLRRRQSSLFYCSRRSFAVVACSTPVNNGGSRVRFAPSPGNLHVGGART
ALFNLYF

>At3g46100

MRAIHIVTTRLSSSFRPILLDLVSSCSPPRQFSIPRRLICAAANGGGRSGSIVAPLVTEEDFHKKIDVNPPKGTRD
FPP

>At3g13490

MEALKVWSLTATPLKQLLRLSSSSTRLATTIYGRRSYHLSPALRCASAASSSSSSATTAETSKPSGRNRRSASS
NSTSD

>At3g55400

MAARINTSLHNALSFLKPFNTPLNTKPFSTRNSFRFSKLPYYSQFSSGKRALYCTSSSQESTVDEGETFVLTT
PLYVV

>At3g58140

MTVFSVQSTIFSRASVALLSSNGFKRFSFVSSFSSSAAYSPPKMRKRRYPIVSAVDIGGVAIARNDVVREDDPTN
NVPDS

>At5g52520

MVSSSLRPLSLTSLFPATTRYPATLRRTVCLNRPLSGFATAPSGTASPETKSSEVDRLRSDRAVTPRSQDFNA
WYLDV

>At1g11870

MGLHKLRLAAAVPVTFSSRLFLKPFTNTLTGLFSRHLQPRNKPLLVRAFSASAAVQDIPATQTS DSSVAARPLW
KAAID

>At2g04842

MASSHLLFSSSFLSKPSSFTSSLRRFVYLPTRQFWPRQRHGFSTVFAVATEPAISSSGPKKAEPSTVVLP SNES
SDKLL

>At2g25840

MGHATSLSHFLILSSSRFSRLGSLTRLLSKPTSLSGSFSSISVTGQGFRCCCSVATDDTSPSVKKRVVSGVQPTG
SVHLG

>At3g02660

MAYATGITFASRSILPICSRTFLSPLRVASLLVFPEKSSATFFRRVQVPHLFSTSTTTLFSSVKCSIHSTSSPETENQ
AV

>At3g19170

MLRTVSCLASRSSSSLFFRFFRQFPRSYMSLTSSTAALRVPSRNLRRISSPSVAGRRLLLRGLRIPSAAVRSVN
QQFSR

>At1g49630

MLRSLTCSSTITSTSLFFRSFRQLPRSYLSPSSSTTVVGASGRNIRRLSTLEAAGRRLFLRRGLKLLSAASRGLNG
QFSR

>At3g54660

MASTPKLTSTISSSSPSLQFLCKKLPIAIHLPSSSSSSFLSLPKLTSLYSLRPRIALLSNHRYYHSRRFSVCASTDN
GA

>At5g15700

MSSAQTPFLANQTKVFDHLIPLHKPFISSPNPVSQSFPMWARNIAKQAISSAARLNVSQTRGLLVSSPESIFSK
NLSF

>At1g79230

MASTLFSRTFLAASHRLITPSLPQKIFNPATFLSRSLHSQLGSASTAYKSTTWARRAMASTGVETKAGYSTSSVS
TSEPV

>At3g25740

MLQKISQSISLCNGDQFKPLIYLAGAPTNFISSPLSGKKKSSSLRIKRIQQLQSTLED RINPPLVCGTVSPRLSVPD
HIL

>At4g37040

MAGVKSLQPRLISSFLGNNISIRSTQPLIHLFRFDLGRRHVSMQLSRTFSGLTDLLFNRRNEDEVIDGKRKRLRPG
NVSPR

>At1g15390

MGLHRDEATAMETLFRVSLRLLPVSAAVTCRSIRFPVSRPGSSHLLNRKLYNLPTSSSSSLSTKAGWLLGLGEKK
KKVDL

>At5g14660

MAVCNCFLQAPPLSRILLPVLSRRATTLSAGYGRKSTVTF CSTVNRTSPLTSSVRAEVKRVSRKDDKVASATDV
QFETP

>At2g30320

MAVSFLRIPISRVFWYPTRLIRVTPALLRFSSSAAFQPSSTSLSPSSDNFLADKWESYRKKKVVIRIGYVGTDY
RGLQ

>At4g13850

MAFCNKLSGILRQGVSQSSNGPVTSMGLSLRYMSSKLFVGGLSWGTDSSSLKQAFSTFGEVTEATVIADRETG
RSRGFGF

>At5g35630

MAQILAASPTCQMRVPKHSSVIASSSKLWSSVVLKQKKQSNNKVRGFRVLALQSDNSTVNRVETLLNLDTKPYS
DRIIAE

>At5g24850

MNDHIHRVPALTEEEIDSVAIKTFERYALPSSSSVKRKGKGVITLWFRNDLRVLDNDALYKAWSSSDTILPVYCLD
PRLF

>At4g08390

MAERVSLTLNGTLLSPPPTTTTTTMSSSLRSTTAASLLLRSSSSSSRSTLTLSSSSLSFVRSLVSSPRLSSSSSL
SQKK

>At3g23830

MAFCNKLSGILRQGVSQSSNGPVTSMGLSLRYMSSKLFVGGLSWGTDSSSLKQAFSTFGEVTEATVIADRETG
RSRGFGF

>At5g47320

MAFCTKLGHWKQGVNVPVSSMLGSLRYMSTKLYIGGLSPGTDEHSLKDAFSSFNGVTEARVMTNKVTGRSR
GYGFVNFI

ANEXO B - Dados numéricos obtidos pelo programa ImageJ que foram utilizados nas análises estatísticas com as proteínas RBP

(continua)

Construção	Célula	Tipo celular	Valores relativos para mitocôndrias	Construção	Célula	Tipo celular	Valores relativos para mitocôndrias
WT-1b	1	N	0.575882924	WT-1a	34-2	N	0.362538645
WT-1b	2	N	0.672418121	WT-1a	34-3	N	0.633085683
WT-1b	7	N	0.685516659	WT-1a	35-1	N	0.298738358
WT-1b	8-1	N	0.594948243	WT-1a	35-2	N	0.266849206
WT-1b	8-2	N	0.619227078	WT-1a	35-3	N	0.373619015
WT-1b	8-3	N	0.624486795	WT-1a	35-4	N	0.323631401
WT-1b	9-1	N	0.655192145	WT-1a	37-1	N	0.360569506
WT-1b	9-2	N	0.502822214	WT-1a	37-2	N	0.305062651
WT-1b	11-1	N	0.554066769	WT-1a	37-3	N	0.305942802
WT-1b	11-2	N	0.497978849	WT-1a	37-4	N	0.325891843
WT-1b	12-1	N	0.719016171	WT-1a	38	N	0.273614828
WT-1b	12-2	N	0.670422309	WT-1a	42	N	0.288790653
WT-1b	13	N	0.684697751	WT-1a	45-1	N	0.279121832
WT-1b	14-1	N	0.585703233	WT-1a	45-2	N	0.375796288
WT-1b	14-2	N	0.634275738	RPS19	1-1	N	0.442303259
WT-1b	16	N	0.62345099	RPS19	1-2	N	0.349565645
WT-1b	18-1	N	0.539646741	RPS19	1-3	N	0.387865998
WT-1b	18-2	N	0.49663112	RPS19	1-4	N	0.407090159
WT-1b	19	N	0.556628905	RPS19	2-1	N	0.43185704
WT-1b	20-1	N	0.585687619	RPS19	2-2	N	0.402938593
WT-1b	20-2	N	0.57421459	RPS19	3-1	N	0.302048617
WT-1b	22-1	N	0.659539494	RPS19	3-2	N	0.432329936
WT-1b	22-2	N	0.512087296	RPS19	4	N	0.272479719
WT-1b	25	N	0.667041743	RPS19	5-1	N	0.392926991
WT-1b	27	N	0.568689802	RPS19	5-2	N	0.455324948
WT-1b	42	N	0.708323505	RPS19	6-1	N	0.368885288
WT-1b	32	N	0.530871853	RPS19	6-2	N	0.496464719
WT-1b	10	S	0.68201956	RPS19	8-1	N	0.351138287
WT-1b	26	S	0.537261626	RPS19	8-2	N	0.425126841
WT-1b	28	S	0.551259914	RPS19	9-1	N	0.481123178
WT-1b	29-1	S	0.467736786	RPS19	9-2	N	0.360844621
WT-1b	29-2	S	0.478697102	RPS19	10-1	N	0.463041342
WT-1b	29-3	S	0.546531366	RPS19	10-2	N	0.434778616
WT-1b	40	S	0.675162452	RPS19	11-1	N	0.476448677
WT-1a	33-1	N	0.310942117	RPS19	11-2	N	0.407064735
WT-1a	33-2	N	0.192472966	RPS19	12-1	N	0.391362686
WT-1a	33-3	N	0.215872717	RPS19	12-2	N	0.483753643
WT-1a	33-4	N	0.280829899	RPS19	13-1	N	0.5579304
WT-1a	33-5	N	0.276186999	RPS19	13-2	N	0.48119641
WT-1a	33-5	N	0.35913965	RPS19	14-1	N	0.375969317
WT-1a	33-6	N	0.202696372	RPS19	14-2	N	0.353905784
WT-1a	34-1	N	0.277375368	RPS19	15-1	N	0.253874961

ANEXO B - Dados numéricos obtidos pelo programa ImageJ que foram utilizados nas análises estatísticas com as proteínas RBP

				(continuação)			
Construção	Célula	Tipo celular	Valores relativos para mitocôndrias	Construção	Célula	Tipo celular	Valores relativos para mitocôndrias
RPS19	15-2	N	0.324053062	K6Q	32-2	N	0.33486118
A2N	4-1	N	0.53417477	K6Q	32-3	N	0.302787234
A2N	4-2	N	0.564672988	K6Q	34-1	N	0.169376152
A2N	6	N	0.641248341	K6Q	34-2	N	0.213290825
A2N	9	N	0.549110704	K6Q	34-3	N	0.35197367
A2N	25-1	N	0.511907325	K6Q	35-1	N	0.317489306
A2N	25-2	N	0.472545939	K6Q	35-2	N	0.221943196
A2N	26	N	0.654464803	K6Q	35-3	N	0.136535258
A2N	30	N	0.536077051	K6Q	36-1	N	0.394466148
A2N	33-1	N	0.657296149	K6Q	36-2	N	0.393204136
A2N	33-2	N	0.582399831	K6Q	36-3	N	0.392532369
A2N	35	N	0.575884768	K6Q	3	S	0.280864932
A2N	37-1	N	0.514520565	K6Q	16-1	S	0.554144667
A2N	37-2	N	0.665405766	K6Q	16-2	S	0.240485987
A2N	39	N	0.725512785	K6Q	19-1	S	0.275753016
A2N	3	S	0.593460761	K6Q	19-2	S	0.213370547
K6Q	1-1	N	0.324330804	K6Q	20	S	0.259918341
K6Q	1-2	N	0.334012471	K6Q	22	S	0.349996655
K6Q	5	N	0.349073613	K6Q	24	S	0.381910569
K6Q	7-1	N	0.232556527	K6Q	25	S	0.417107458
K6Q	7-2	N	0.331213358	K6Q	27	S	0.264630845
K6Q	8-1	N	0.608008141	R12N	1-1	N	0.443053956
K6Q	8-2	N	0.532586707	R12N	1-2	N	0.497574884
K6Q	10	N	0.385301508	R12N	1-3	N	0.511578581
K6Q	13-1	N	0.341112943	R12N	2	N	0.363382163
K6Q	13-2	N	0.422372979	R12N	3-1	N	0.164794408
K6Q	14-1	N	0.17331913	R12N	3-2	N	0.256933848
K6Q	14-2	N	0.344659327	R12N	5	N	0.349226399
K6Q	15	N	0.20531173	R12N	6-1	N	0.205110963
K6Q	17-1	N	0.262871973	R12N	6-2	N	0.232731632
K6Q	17-2	N	0.143334711	R12N	13	N	0.338207396
K6Q	21	N	0.327031532	R12N	17-1	N	0.255970605
K6Q	26	N	0.431443536	R12N	17-2	N	0.317622612
K6Q	28-1	N	0.402357608	R12N	19	N	0.294667473
K6Q	28-2	N	0.282805307	R12N	20	N	0.351195566
K6Q	28-3	N	0.239047403	R12N	21-1	N	0.380129617
K6Q	30-1	N	0.232470894	R12N	21-2	N	0.243965138
K6Q	30-2	N	0.334202902	R12N	22-1	N	0.423874022
K6Q	31-1	N	0.474605498	R12N	22-2	N	0.35945304
K6Q	31-2	N	0.412660137	R12N	23-1	N	0.318115055
K6Q	31-3	N	0.397598251	R12N	23-2	N	0.228104195
K6Q	32-1	N	0.421415401	R12N	23-3	N	0.324740189

ANEXO B - Dados numéricos obtidos pelo programa ImageJ que foram utilizados nas análises estatísticas com as proteínas RBP

				(continuação)			
Construção	Célula	Tipo celular	Valores relativos para mitocôndrias	Construção	Célula	Tipo celular	Valores relativos para mitocôndrias
R12N	23-4	N	0.433682412	V15NS16G	18-2	N	0.469311908
R12N	24-1	N	0.374158007	V15NS16G	18-3	N	0.277711195
R12N	24-2	N	0.188775353	V15NS16G	19-1	N	0.337063072
R12N	25-1	N	0.253120844	V15NS16G	19-2	N	0.661242028
R12N	25-2	N	0.178565498	V15NS16G	20-1	N	0.485735959
R12N	25-3	N	0.238581557	V15NS16G	20-2	N	0.378328725
R12N	26	N	0.177396911	V15NS16G	20-3	N	0.385306325
R12N	27	N	0.221414923	V15NS16G	21-1	N	0.352979611
R12N	28-1	N	0.216611122	V15NS16G	21-2	N	0.368304553
R12N	28-2	N	0.271888817	V15NS16G	22-1	N	0.216733577
R12N	28-3	N	0.314788666	V15NS16G	22-2	N	0.185657697
V15NS16G	2	N	0.313019761	S18NS19G	2	N	0.494821171
V15NS16G	3-1	N	0.334601431	S18NS19G	6-1	N	0.351155438
V15NS16G	3-2	N	0.291338859	S18NS19G	6-2	N	0.317262296
V15NS16G	3-3	N	0.240869399	S18NS19G	8-1	N	0.24081589
V15NS16G	4-1	N	0.490336309	S18NS19G	8-2	N	0.440466192
V15NS16G	4-2	N	0.622161697	S18NS19G	10-1	N	0.390875046
V15NS16G	5-1	N	0.704255432	S18NS19G	10-2	N	0.240441159
V15NS16G	5-2	N	0.791181608	S18NS19G	10-3	N	0.261396703
V15NS16G	6-1	N	0.197914011	S18NS19G	13-1	N	0.551439587
V15NS16G	6-2	N	0.148958377	S18NS19G	13-2	N	0.421490388
V15NS16G	6-3	N	0.279775667	S18NS19G	14-1	N	0.238057207
V15NS16G	8-1	N	0.791815399	S18NS19G	14-2	N	0.244302762
V15NS16G	8-2	N	0.790708075	S18NS19G	17	N	0.435242628
V15NS16G	9-1	N	0.4914233	S18NS19G	19-1	N	0.43562085
V15NS16G	9-2	N	0.409729723	S18NS19G	19-2	N	0.307236407
V15NS16G	10-1	N	0.18494523	S18NS19G	20-1	N	0.31443857
V15NS16G	10-2	N	0.268242984	S18NS19G	20-2	N	0.445507538
V15NS16G	11	N	0.199010839	S18NS19G	22	N	0.352200527
V15NS16G	12-1	N	0.514776193	S18NS19G	26	N	0.351567421
V15NS16G	12-2	N	0.608976016	S18NS19G	28	N	0.40508999
V15NS16G	13-1	N	0.179406206	S18NS19G	29-1	N	0.413943931
V15NS16G	13-2	N	0.395164802	S18NS19G	29-2	N	0.438449201
V15NS16G	13	N	0.25795959	S18NS19G	29-3	N	0.501171317
V15NS16G	14-1	N	0.806347965	S18NS19G	31	N	0.440074087
V15NS16G	14-2	N	0.785840183	S18NS19G	32	N	0.378534498
V15NS16G	14-3	N	0.847849072	D(18-19)	4-1	N	0.347244453
V15NS16G	16-1	N	0.505436196	D(18-19)	4-2	N	0.546296665
V15NS16G	16-2	N	0.44810001	D(18-19)	4-3	N	0.619332912
V15NS16G	17-1	N	0.500020582	D(18-19)	5-1	N	0.666813621
V15NS16G	17-2	N	0.328971794	D(18-19)	5-2	N	0.389565944
V15NS16G	18-1	N	0.389995488	D(18-19)	5-3	N	0.548818981

ANEXO B - Dados numéricos obtidos pelo programa ImageJ que foram utilizados nas análises estatísticas com as proteínas RBP

			(conclusão)
Construção	Célula	Tipo celular	Valores relativos para mitocôndrias
D(18-19)	5-4	N	0.393591463
D(18-19)	7	N	0.306951223
D(18-19)	8-1	N	0.500971979
D(18-19)	8-2	N	0.53899062
D(18-19)	9-1	N	0.446742885
D(18-19)	9-2	N	0.283824094
D(18-19)	11	N	0.394379127
D(18-19)	12-1	N	0.360306252
D(18-19)	12-2	N	0.424528051
D(18-19)	13-1	N	0.34201723
D(18-19)	13-2	N	0.554462497
D(18-19)	14-1	N	0.29366997
D(18-19)	14-2	N	0.401068724
D(18-19)	14-3	N	0.598366372
D(18-19)	15	N	0.432935495
D(18-19)	16	N	0.483454776
D(18-19)	19	N	0.567501732
D(18-19)	21	N	0.47287335
D(18-19)	22-1	N	0.265611339
D(18-19)	22-2	N	0.401015296
D(18-19)	23-1	N	0.561656832
D(18-19)	23-2	N	0.454051852
D(18-19)	24-1	N	0.335583703
D(18-19)	24-2	N	0.316733095
D(18-19)	24-3	N	0.3062902
D(18-19)	25	N	0.36494248
R31Q	14	N	0.284772121
R31Q	15-1	N	0.233016536
R31Q	15-2	N	0.38036366
R31Q	16-1	N	0.239119968
R31Q	16-2	N	0.287480002
R31Q	16-3	N	0.348249974
R31Q	17	N	0.450018683
R31Q	18	N	0.281156398
R31Q	19-1	N	0.309547317
R31Q	19-2	N	0.340971486
R31Q	19-3	N	0.226076719
R31Q	19-4	N	0.348336508
R31Q	20-1	N	0.378959708
R31Q	20-2	N	0.433094348
R31Q	22	N	0.272821922

Notas: Abreviações, N (célula do mesófilo), S (estômato). X-1...X-4, indicam medidas de planos diferentes de uma mesma célula.