UNIVERSITÉ DE STRASBOURG

UNIVERSITY OF WESTERN AUSTRALIA

2010

THÈSE EN COTUTELLE

THESIS

présentée pour obtenir le grade de

DOCTORAT EN BIOLOGIE MOLÉCULAIRE ET CELLULAIRE

is presented for the degree of

DOCTOR OF PHILOSOPHY IN BIOCHEMISTRY

Kamel HAMMANI

Études fonctionnelles des protéines PPR dans les organites d'Arabidopsis thaliana

Functional studies of PPR proteins in Arabidopsis organelles

THE UNIVERSITY OF WESTERN AUSTRALIA

Institut de Biologie Moléculaire des Plantes, CNRS UPR 2357
ARC Centre of Excellence in Plant Energy Biology

Soutenue le 20.07.2010 devant la commission d'examen:

Directeur de thèse:	Dr Philippe GIEGÉ	Université de Strasbourg
Directeur de thèse:	Prof. Ian SMALL	University of Western Australia
Rapporteur externe:	Prof. Axel BRENNICKE	Universität Ulm
Rapporteur externe:	Prof. Claire REMACLE	Université de Liège
Rapporteur interne:	Prof. Catherine FLORENTZ	Université de Strasbourg

You create your opportunities by asking for them.

REMERCIEMENTS

Je tiens à remercier les professeurs Catherine Florentz, Claire Remacle et Axel Brennicke d'avoir accepté de juger mes travaux de thèse.

Je remercie profondément mes directeurs de thèse, Docteur Philippe Giegé et Professeur Ian Small, de m'avoir accueilli dans leur laboratoire respectif et de m'avoir offert cette opportunité unique de mener ma thèse en cotutelle entre l'Université de Strasbourg et l'Université de Western Australia.

Egalement, je souhaite exprimer ma reconnaissance aux partenaires financiers qui ont soutenu mes travaux de recherche: la Région Alsace, The University of Western Australia, l'Université de Strasbourg, l'Agence Nationale de la Recherche (07-JCJC-0123).

Mes remerciements iront également à tous mes collègues et amis de l'Institut de Biologie Moléculaire des Plantes à Strasbourg en France et l'ARC Centre of Excellence in Plant Energy Biology à Perth en Australie avec qui j'ai partagé des moments agréables.

Enfin, je remercie ma famille et mes parents pour leur soutien et surtout pour l'éducation qu'ils m'ont donné: ne jamais se décourager et redoubler d'effort dans les moments difficiles.

RÉSUMÉ

Dans les organites de plantes, les étapes post-transcriptionnelles sont régulées par des protéines codées par le noyau et adressées aux organites. Chez les plantes, les protéines de la vaste famille PPR (PentatricoPeptide Repeat), sont des acteurs majeurs de cette régulation. Nous avons utilisé une stratégie de génétique inverse afin d'identifier les rôles de protéines PPR chez *Arabidopsis thaliana*. Dans les organites de plantes supérieures, l'édition de l'ARN convertit spécifiquement certains résidus cytidine en uridine dans les transcrits. La spécificité de reconnaissance de ces sites est assurée par des *trans*-facteurs protéiques qui lient l'ARN en amont du site d'édition. Nous avons identifié six nouveaux facteurs d'édition PPR dans les chloroplastes d'*Arabidopsis thaliana* (OTP80, OTP81, OTP82, OTP84, OTP85 et OTP86). Ces facteurs sont impliqués dans l'édition de 9 cytidines distinctes et certains reconnaissent plusieurs sites. Une analyse bio-informatique conduite sur les séquences cibles des *trans*-facteurs PPR reconnaissant plusieurs sites d'édition, suggère que leur spécificité de reconnaissance peut être expliquée si le facteur d'édition reconnait certains nucléotides conservés et distingue les résidus pyrimidines et purines.

Lors d'une seconde étude, nous avons identifié un gène essentiel pour le développement de l'embryon chez *Arabidopsis*. Ce gène code pour une nouvelle protéine PPR, PPR9. Des expériences de fusion à la GFP, d'immunodétection et d'immunohistochimie ont montré que PPR9 est à la fois localisée dans les mitochondries et le noyau des cellules. Des expériences de complémentation génétique prouvent que seule la localisation mitochondriale est essentielle au développement de l'embryon. PPR9 lie l'ARN *in vitro* et est associée avec les polysomes des mitochondries de plantes. La protéine interagit avec la protéine NAP1 (Nucleosome Assembly Protein1) et le facteur de transcription TCP8. L'ensemble des données suggère que PPR9 régule l'expression des gènes mitochondriaux et nucléaires.

ABSTRACT

In plant organelles, nuclear encoded proteins targeted to the organelles control post-transcriptional mechanisms. The PentatricoPeptide Repeat proteins (PPR) whose family has greatly expanded in land plants, are good candidates to play such regulatory roles. We used a reverse genetic screen in *Arabidopsis thaliana* to assign new functions to *PPR* genes. RNA editing in higher plant organelles results in the conversion of specific cytidine residues to uridine residues in RNA. The recognition of a specific target C site by the editing machinery involves *trans*-acting factors that bind to the RNA upstream of the C to be edited. We identified six new PPR chloroplast editing factors in *Arabidopsis thaliana* (OTP80, OTP81, OTP82, OTP84, OTP85, and OTP86). These six factors account for nine editing sites not previously assigned to an editing factor and, together with the nine PPR editing proteins previously described, explain more than half of the 34 editing events in *Arabidopsis* chloroplasts. An analysis of the target sites requiring the editing factors involved in editing of multiple sites suggests that editing factors can generally distinguish pyrimidines from purines and, at some positions, must be able to recognize specific bases.

Secondly, we report that loss-of-function mutations of *PPR9* are lethal for the embryo in *Arabidopsis thaliana*. *PPR9* encodes a novel PPR protein dual localized to mitochondria and nuclei as observed by GFP fusions, immunoblots using anti-PPR9 antibodies on sub-cellular fractions and immunohistochemistry experiments. Genetic complementation showed that loss of PPR9 function in mitochondria is lethal for the embryo but not in nuclei. PPR9 is able to bind RNA. In mitochondria, it is associated with polysomes and may play a role in translation, whereas, in the nucleus, PPR9 interacts with the nucleosome assembly protein NAP1 and the transcription factor TCP8. Altogether, our findings suggest that PPR9 might be involved in gene expression regulation between mitochondria and the nucleus.

TABLE DES MATIERES

1. PUBLICATIONS RESULTANTES DE CETTE THESE 2

2. INTRODUCTION 3

2.1. LES MITOCHONDRIES ET LES CHLOROPLASTES 3

2.1.1. ORIGINE 3

2.1.2. FONCTIONS BIOLOGIQUES 3

2.1.3. LES GENOMES DES ORGANITES CHEZ ARABIDOPSIS 4

2.2. EXPRESSION DES GENOMES DES ORGANITES 5

2.2.1. LA TRANSCRIPTION DES GENES 6

2.2.2. LA LIBERATION DES EXTREMITES 5' ET 3' DES MESSAGERS 8

2.2.3. EPISSAGE DES INTRONS 10

2.2.4. L'EDITION DE L'ARN 12

2.3. LA TRADUCTION 17

2.4. COORDINATION DE L'EXPRESSION DES GENOMES DES ORGANITES ET DU NOYAU 18

2.5. LES PROTEINES PPR 20

2.5.1. DEFINITION 20

2.5.2. FONCTIONS BIOLOGIQUES DES PROTEINES PPR 22

3. OBJECTIFS DES TRAVAUX DE THESE 23

4. RESULTATS 23

4.1. IMPLICATION DE PROTEINES PPR DANS L'EDITION D'ARN CHLOROPLASTIQUES 23

4.1.1. PUBLICATION 1: A STUDY OF NEW ARABIDOPSIS CHLOROPLAST RNA EDITING MUTANTS REVEALS GENERAL FEATURES OF EDITING FACTORS AND THEIR TARGET SITES, THE PLANT CELL (2009) 24

4.1.2. PUBLICATION 2: THE PENTATRICOPEPTIDE REPEAT PROTEIN OTP82 IS REQUIRED FOR RNA EDITING OF PLASTID NDHB AND NDHG TRANSCRIPTS, THE PLANT JOURNAL (2010) 25

4.2. IMPLICATION D'UNE PROTEINE PPR DANS LA REGULATION DE L'EXPRESSION DES GENOMES DU NOYAU ET DES MITOCHONDRIES 26

4.2.1. PUBLICATION 3: ARABIDOPSIS PPR9 IS DUAL LOCALIZED IN MITOCHONDRIA AND THE NUCLEUS WHERE IT INTERACTS WITH TRANSCRIPTION FACTORS. (EN PREPARATION) 27

4.3. CONTRIBUTIONS MINEURES A PUBLICATION 28

4.3.1. PUBLICATION 4: THE ARABIDOPSIS GENE YS1 ENCODING A DYW PROTEIN IS REQUIRED FOR EDITING OF rpoB TRANSCRIPTS AND THE RAPID DEVELOPMENT OF CHLOROPLASTS DURING EARLY GROWTH. THE PLANT JOURNAL (2009) 28

4.3.2. PUBLICATION 5: PPR336 IS ASSOCIATED WITH RIBOSOMES IN PLANT MITOCHONDRIA. JOURNAL OF MOLECULAR BIOLOGY (2009) 29

5. DISCUSSION 30

5.1. NECESSITE DE L'EDITION DE L'ARN DANS LES ORGANITES 30

5.2. SPECIFICITE DE RECONNAISSANCE MULTIPLE D'UN *TRANS*-FACTEUR PPR 31

5.3. LES PROTEINES PPR, FACTEURS PUTATIFS D'EDITION DANS LES ORGANITES 32

5.4. L'ADRESSAGE DE PROTEINES PPR EN DEHORS DES ORGANITES 33

5.5. INTERACTIONS PROTEIQUES DES PROTEINES PPR 34

5.6. LE DOMAINE PPR, UNE PLATEFORME MODULABLE D'INTERACTION AVEC L'ARN 34

6. BIBLIOGRAPHIE 37

1. PUBLICATIONS RESULTANTES DE CETTE THESE

1. Hammani K, Gobert A, Hleibieh K, Choulier L, Small I, Giegé P (en préparation) *Arabidopsis* PPR9 is dual localized in mitochondria and the nucleus where it interacts with transcription factors.

2. Okuda K*, Hammani K*, Tanz, S, Peng L, Fukao Y, Myouga F, Motohashi R, Shinozaki K, Small I, Shikanai T (2010) The pentatricopeptide repeat protein OTP82 is required for RNA editing of plastid ndhB and ndhG transcripts. The Plant Journal 61:339-349.

*co-first authors

3. Hammani K*, Okuda K*, Tanz SK, Chateigner-Boutin AL, Shikanai T, Small I (2009) A study of new Arabidopsis chloroplast RNA editing mutants reveals general features of editing factors and their target sites. The Plant Cell 21:3686-3699

*co-first authors

4. Zhou W, Cheng Y, Yap A, Chateigner-Boutin AL, Delannoy E, Hammani K, Small I, Huang J (2009) The Arabidopsis gene YS1 encoding a DYW protein is required for editing of rpoB transcripts and the rapid development of chloroplasts during early growth. The Plant Journal 58:82-96

5. Uyttewaal M, Mireau H, Rurek M, Hammani K, Arnal N, Quadrado M, Giegé P (2008) PPR336 is associated with ribosomes in plant mitochondria. Journal of Molecular Biology 375:626-636, 2008

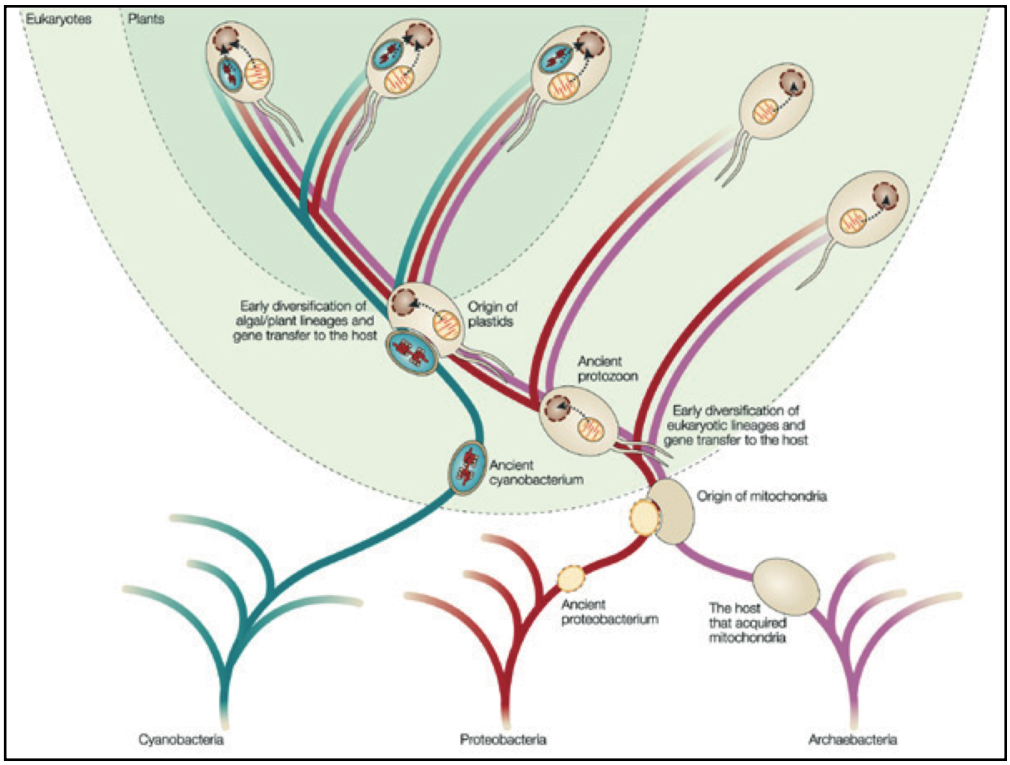


Figure 1: Origine endosymbiotique des organites et évolution des génomes.

La mitochondrie et le chloroplaste originent respectivement d'une α-protéobactérie et d'une cyanobactérie successivement endocytées par une cellule eucaryote primitive de type archaebatérie. L'arbre représente l'évolution de la cellule eucaryote et souligne particulièrement la co-évolution du génome nucléaire et de celui des organites. A l'heure actuelle, la nature de la cellule hôte ancestrale fait toujours débat entre les évolutionnistes (d'après Timmis et *al.*, 2004).

2. INTRODUCTION

2.1. LES MITOCHONDRIES ET LES CHLOROPLASTES

La mitochondrie et le chloroplaste sont deux organites qui se distinguent du reste de la cellule végétale par leur structure compartimentée et par leur semi-autonomie grâce à l'existence d'un génome au sein de leur structure.

2.1.1. ORIGINE

Il est communément admis que la mitochondrie et le chloroplaste ont une origine endosymbiotique (Timmis et *al.*, 2004). La mitochondrie serait issue d'une endosymbiose primaire entre une α-protéobactérie de type aérobie et une cellule eucaryote primitive de type archaebactérie il y a environ 1,5 milliards d'années (Gray et *al.*, 1999). Cette première endosymbiose aurait permis à la cellule hôte d'acquérir le mécanisme moléculaire de la respiration. Le chloroplaste quant à lui, résulterait d'une endosymbiose secondaire entre une cellule eucaryote issue de l'endosymbiose primaire et une cyanobactérie (Figure 1) (Raven et *al.*, 2003). Cet événement secondaire a donné lieu à l'émergence du règne végétal. Durant l'évolution des symbiotes en organites, la grande majorité des gènes ont été perdus ou transférés au noyau de la cellule hôte permettant à cette dernière un meilleur contrôle du fonctionnement des organites (Timmis et *al.*, 2004 ; Martin et *al.*, 2003). Des transferts génétiques du chloroplaste vers la mitochondrie ont été également fréquemment rapportés alors que les transferts en sens inverse sembleraient plus rares (Stern et *al.*, 1982, 1984 ; Martin et *al.*, 1998 ; Stegemann et *al.*, 2003 ; Hao et *al.*, 2009).

2.1.2. FONCTIONS BIOLOGIQUES

De manière générale, la fonction primaire de la mitochondrie au sein de la cellule eucaryote est son approvisionnement en énergie produit sous forme d'ATP lors de la respiration ou phosphorylation oxydative mitochondriale.

Ce mécanisme utilise les précurseurs réduits NADH ou $FADH_2$ comme donneurs d'électrons de la chaîne respiratoire localisée au sein de la membrane interne de la mitochondrie. Le transfert d'électrons le long des différents complexes respiratoires génère un gradient transmembranaire de protons. Ce gradient servira à fabriquer l'ATP lors du passage des protons dans la matrice à travers l'ATP synthase.

La mitochondrie abrite également le cycle de Krebs ou cycle des acides tricarboxyliques ou encore cycle de l'acide citrique (citrate). Ce cycle est une série de réactions biochimiques dont la finalité est de produire les intermédiaires énergétiques NADH et $FADH_2$ servant à la production d'ATP dans la chaîne respiratoire.

La mitochondrie est également impliquée dans la mort cellulaire programmée ou apoptose et dans l'homéostasie de certains ions tels que le calcium (Scott and Logan, 2008 ; Logan and Night, 2003).

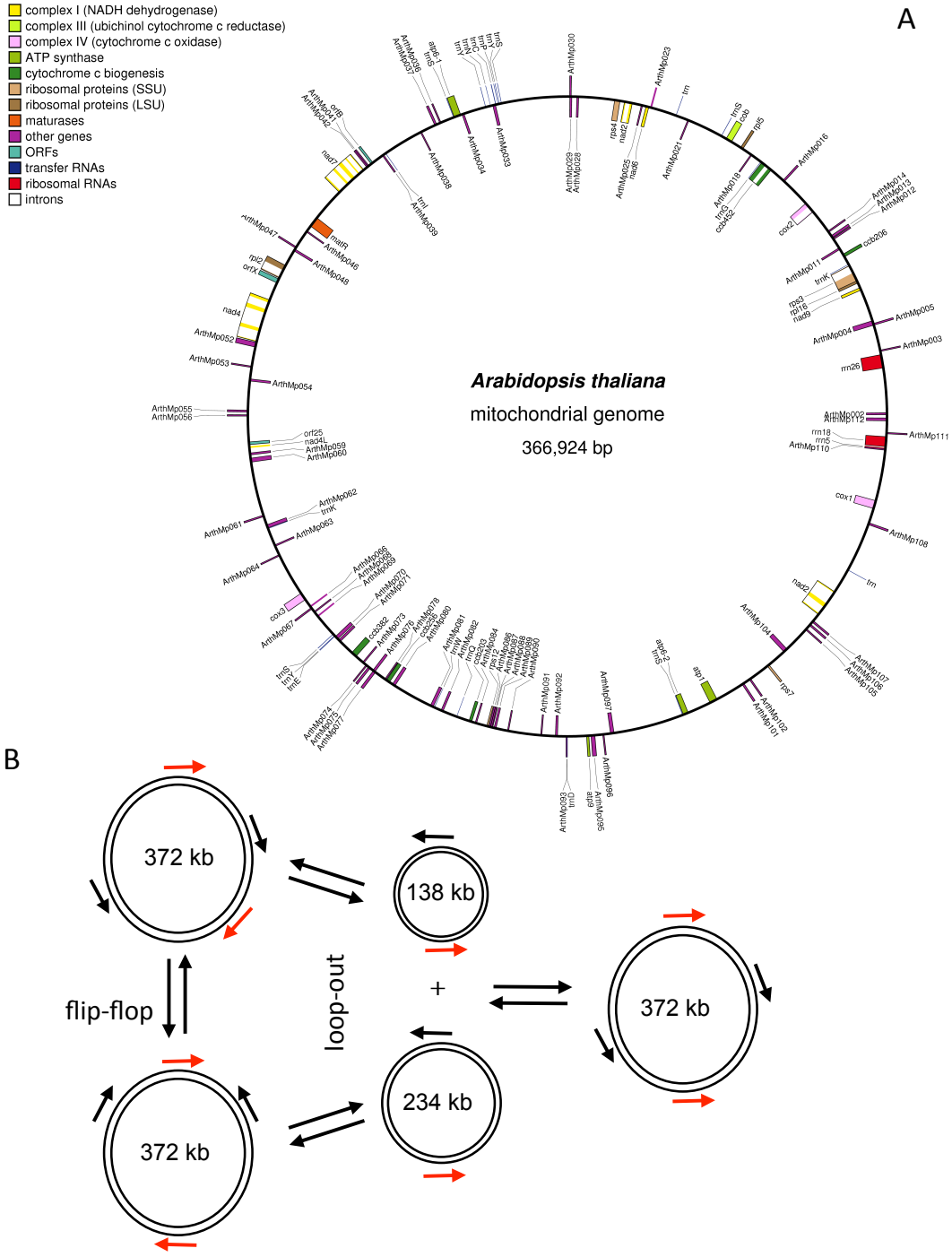


Figure 2: Carte du génome mitochondrial d'*Arabidopsis* (A) et son modèle d'organisation multipartite (B).

A) Le génome mitochondrial d'*Arabidopsis* code principalement des protéines impliquées dans le métabolisme respiratoire et des composants de la machinerie traductionnelle. Il est intéressant de noter sa faible densité génique et la présence de nombreuses ORFs de fonctions inconnues.
B) Le génome mitochondrial d'*Arabidopsis* sous forme de chromosome circulaire ou "cercle maitre" de 372 kb peut être présent sous différentes formes (trois sont représentées sur le schéma). Celles-ci sont générées par des recombinaisons de types "flip-flop" ou "loop-out" entre deux lots de séquences répétées (représentées par des flèches rouges et noires). Après recombinaison, le chromosome circulaire est en équilibre avec deux molécules circulaires subgénomiques de 138 kb et 234 kb (Klein et *al*., 1994).

Le chloroplaste est le siège de la photosynthèse, processus bioénergétique qui permet à la plante de synthétiser de la matière organique en exploitant l'énergie lumineuse captée par les pigments photosynthétiques des antennes réactionnelles localisées dans les membranes des thylakoïdes. Ce mécanisme assure l'autotrophie de la cellule végétale. Lors de la photosynthèse, une molécule d'eau est réduite en dioxygène libérant un électron transféré entre les différents photosystèmes permettant la libération de protons dans le lumen des thylakoïdes et la formation d'un gradient électrochimique. Tout comme dans la mitochondrie, ce dernier servira à la production d'ATP lors du passage des protons dans le stroma à travers l'ATP synthase. Lors du transfert d'électrons le NADP est également réduit en NADPH.

Dans un deuxième temps, l'ATP et le NADPH seront utilisés par le cycle de Calvin qui permet de fixer le carbone contenu dans le dioxyde de carbone atmosphérique en le liant aux atomes d'hydrogène des molécules d'eau. Au cours de ce cycle, le carbone sera réduit en composés organiques.

De plus, ces organites jouent un rôle dans le métabolisme des acides aminés, la biosynthèse des lipides, de vitamines, des bases puriques et pyrimidines et de groupements fer-soufre (Neuhaus et Emes, 2000 ; Millar et *al.*, 2004).

2.1.3. LES GENOMES DES ORGANITES CHEZ ARABIDOPSIS

Les génomes des organites sont polyploïdes et sont présents en un grand nombre de copies dans la mitochondrie et le chloroplaste.

Alors que le génome mitochondrial de l'humain est réduit à une taille d'environ 16 kb (Ingman et *al.*, 2000), celui des plantes supérieures est caractérisé par une grande taille de l'ordre de 490 kb chez le riz (Notsu et *al.*, 2002) et 367 kb chez *Arabidopsis thaliana* (Unseld et *al.*, 1997). Le grand génome mitochondrial des plantes supérieures représente un tiers du génome ancestral de l'α-protéobactérie *Rickettsia* mais ne code que 4% de la quantité de protéines codées par le génome de la bactérie. Ceci atteste de la très faible densité génique du génome mitochondrial des plantes supérieures caractérisé par de nombreuses régions non codantes.

Les 366,924 paires de bases qui composent le génome mitochondrial d'*Arabidopsis* comportent au total 57 gènes couvrant seulement 10% du génome (Figure 2A). Ces gènes codent pour 32 protéines, 3 ARN ribosomiques et 22 ARN de transfert (Unseld et *al.*, 1997 ; Marienfield et *al.*, 1999). Deux classes majeures de gènes peuvent être distinguées. La première classe regroupe les gènes codant des protéines impliquées dans le métabolisme mitochondrial et composant les cinq complexes de la chaine respiratoire et les enzymes de maturation du cytochrome *c*. La seconde classe de gènes code les composants de la machinerie traductionnelle mitochondriale : protéines de la petite et grande sous unités du ribosome, ARN ribosomiques et ARN de transfert.

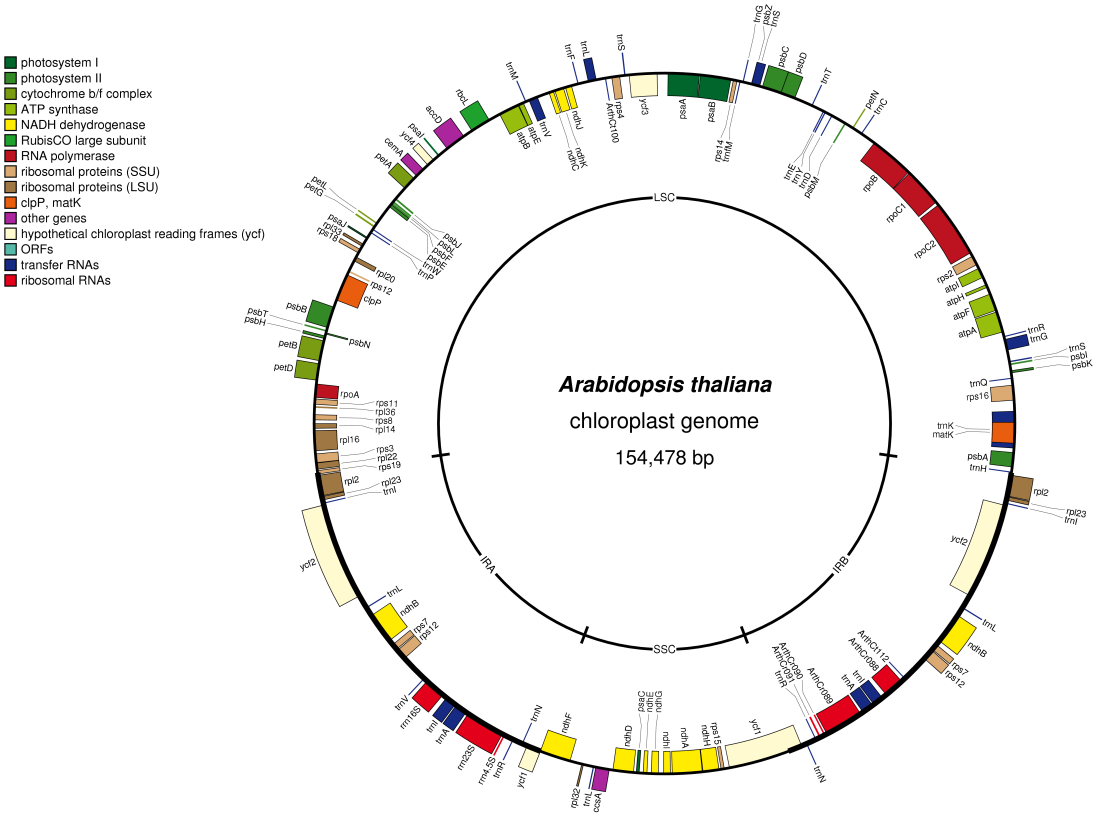


Figure 3: Carte du génome chloroplastique d'*Arabidopsis*.

Le génome mitochondrial d'*Arabidopsis* codent principalement des protéines impliquées dans le métabolisme photosynthétique, la transcription et la traduction. Deux régions inversées répétées (IRB et IRA) séparent une grande région simple copie (LSC pour "large single copy region") d'une petite région simple copie (SSC pour "small single copy region").

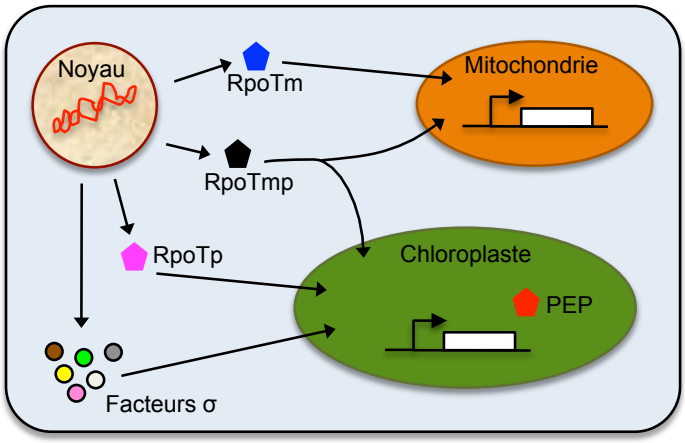


Figure 4: Machinerie transcriptionelle dans les organites d'*Arabidopsis*.

Le noyau d'*Arabidopsis* codent 3 polymérases RpoTm, RpoTmp et RpoTp adressées aux organites. Dans le chloroplaste, les polymérases sont de double origine: le génome code une polymérase PEP (pour Plastid Encoded Polymerase") dont la reconnaissance des promoteurs nécessite toutefois des facteurs σ codés par le noyau.

En plus de ces 57 gènes, 10% du génome est constitué d'ORFs de plus de 300 paires de bases sans homologie particulière pour des gènes identifiés et dont la fonctionnalité est encore sujette à débat (Unseld et *al.*, 1997). Certaines de ces ORFs sont toutefois associées au phénomène de stérilité mâle cytoplasmique caractérisé par l'incapacité de la plante à produire du pollen fertile (Budar et Pelletier, 2001).

L'ADN mitochondrial peut être représenté sous la forme théorique d'un chromosome circulaire appelé "cercle maître". Toutefois, la structure détectée au sein de la mitochondrie de la plupart des espèces végétales est de type multipartite. L'ADN mitochondrial des plantes apparaît sous des formes complexes de molécules circulaires et linéaires subgénomiques. Ce multipartisme est du à des événements fréquents de recombinaisons intra et intermoléculaires au sein du "cercle maître" (Mackenzie et *al.*, 2007). Ces recombinaisons impliquent de larges séquences répétées présentes au sein du génome mitochondrial. Ces recombinaisons peuvent se faire par "flip-flop" entre séquences répétées orientées en sens inverse ou, par "loop-out" entre séquences répétées orientées en sens direct donnant naissance à deux molécules sub-génomiques (Figure 2B) (Klein et *al.*, 1994).

Tout comme la mitochondrie, le génome du chloroplaste des plantes a subit une forte réduction de taille par rapport à celui de la cyanobactérie ancestrale. Le génome chloroplastique d'*Arabidopsis* est formé par 154,478 paires de bases (Sato et *al.*, 1999) (Figure 3) et comporte une paire de régions répétées inversées (IR) délimitant deux régions simple copie de grande et petite taille. Comparé au génome de la mitochondrie la densité génique du génome chloroplastique est plus élevée; les 154,478 paires de bases sont couvertes par 129 gènes chez *Arabidopsis*. Après décompte des gènes répétées, ces gènes codent pour 77 protéines différentes, 4 ARN ribosomiques et 30 ARN de transfert (Sato et *al.*, 1999). Les protéines codées par ce génome peuvent être regroupées en 3 classes principales. La première inclut les gènes codant les protéines impliquées dans le métabolisme photosynthétique (photosystèmes I et II, NADH déshydrogénase, complexe *b6/f* et ATP synthase). La seconde classe de gènes code les composants de la machinerie traductionnelle chloroplastique : protéines de la petite et grande sous unités du ribosome, ARN ribosomiques et ARN de transfert. La troisième classe regroupe les gènes codant des protéines impliquées dans la transcription du génome chloroplastique : sous-unités α, β, β', β'' de l'ARN polymérase PEP (pour plastid encoded polymerase). Il est intéressant de noter que les homologues de ces derniers sont absents du génome mitochondrial d'*Arabidopsis*. Enfin, d'autres gènes ont une fonction dans la maturation des ARN et des protéines ou encore dans la fixation du carbone lors du cycle de Calvin et dans la synthèse des acides gras.

2.2. EXPRESSION DES GENOMES DES ORGANITES

L'organisation des gènes des organites de plantes ressemble particulièrement à celle des procaryotes avec toutefois des caractéristiques eucaryotes. L'expression des gènes des organites nécessite différentes étapes : la transcription des gènes en ARN, la maturation

des extrémités 5' et 3' et le clivage des ARN polycistroniques, l'édition de certains ARN, l'épissage et la dégradation des ARN. Une fois fonctionnels, les ARNm peuvent être traduits par les ribosomes des organites. Ces différentes étapes sont régulées par de nombreuses protéines codées par le noyau et adressées aux organites.

2.2.1. LA TRANSCRIPTION DES GENES

La transcription des gènes des organites des plantes requiert des enzymes de type ARN polymérase. Celles-ci sont exclusivement importées du noyau dans la mitochondrie et sont de double origine dans le chloroplaste (Figure 4).

Les gènes *rpoA*, *rpoB*, *rpoC1* et *C2* présents dans le génome chloroplastique d'*Arabidopsis* codent pour les quatre sous-unités α, β, β', β'' de l'ARN polymérase PEP (Sato et *al.*, 1999). La PEP s'apparente aux polymérases eubactériennes et possède l'activité transcriptionnelle principale dans les chloroplastes matures (Pfannschmidt et *al.*, 2000). La spécificité de reconnaissance des promoteurs par la PEP, nécessite des facteurs additionnels σ codés par le génome nucléaire et importés dans le chloroplaste (Isono et *al.*, 1997 ; Allison, 2000 ; Suzuki et *al.*, 2004). Les promoteurs reconnus par la PEP sont de type bactérien σ^{70} et présentent deux séquences consensus TTGACA et TATAAT trouvées respectivement aux positions -35 et -10 nucléotides de la plupart des unités transcriptionnelles du chloroplaste (Bruce Cahoon et Stern, 2001).

Contrairement au génome chloroplastique, le génome mitochondrial d'*Arabidopsis* ne présente pas de tels gènes et la transcription est entièrement dépendante d'ARN polymérases nucléaires importées ou NEP (nuclear encoded polymerase). Trois gènes nucléaires *rpoTm*, *rpoTp* et *rpoTmp* codent pour de telles polymérases chez *Arabidopsis*. Les produits de ces gènes, RpoTm et RpoTp sont respectivement adressés aux mitochondries et aux chloroplastes alors que RpoTmp est doublement adressé aux deux organites (Chang et *al.*, 1999 ; Hedtke et *al.*, 2000). Contrairement à la PEP, les NEP sont constituées d'une seule sous-unité et s'apparentent aux ARN polymérases T3/T7 (Chang et *al.*, 1999, Shutt et Gray, 2006). Les NEP d'*Arabidopsis* pourraient nécessiter la présence de facteurs protéiques pour reconnaître les séquences promotrices des organites comme c'est le cas dans les mitochondries de levure et de l'humain (Shadel et *al.*, 1995 ; Tracy et *al.*, 1995). Les séquences promotrices du génome mitochondrial d'*Arabidopsis* comportent souvent le motif consensus YRTA (Y : base pyrimidique, R : base purique) alors que celles du génome chloroplastique comportent un motif de trois nucléotides YRT (Liere et Maliga, 1999). Néanmoins, ceci n'est pas toujours le cas et la transcription peut être initiée au niveau de motifs non canoniques (Binder et *al.*, 1994 ; Remacle et Maréchal-Drouard, 1996 ; Fey et Maréchal-Drouard, 1999 ; Bligny et *al.*, 2000 ; Kuhn et *al.*, 2005).

Le rôle de différentes polymérases (2 NEP et 1 PEP) dans le chloroplaste est encore méconnu, mais il apparaît de manière générale que RpoTp joue un rôle dans les tissus non photosynthétiques en transcrivant les gènes ménagers ; ARNr, ARNt, certains gènes métaboliques comme *clpP* et *accD* (impliquée dans la synthèse lipidique) et surtout les

gènes codant les sous-unités de la PEP. En retour, la PEP est active dans les chloroplastes en transcrivant les gènes photosynthétiques (Silhavy et *al.*, 1998 ; Shiina et *al.*, 2005). Toutefois, ce modèle simpliste est limité par la présence de gènes possédant des promoteurs à la fois NEP et PEP (Silhavy et *al.*, 1998 ; Chang et *al.*, 1999). Récemment, le rôle de RpoTmp dans la transcription spécifique de certains gènes mitochondriaux a été établie mais sa fonction dans le chloroplaste reste non élucidée (Kuhn et *al.*, 2009).

A l'inverse des vertébrés où, l'initiation de la transcription des gènes du génome mitochondrial se fait au niveau d'un seul site par brin d'ADN, les organites de plantes possèdent de multiples sites d'initiation. De plus, la transcription de chaque gène peut être initiée au niveau de plusieurs promoteurs (Sturm et *al.*, 1994 ; Lupold et *al.*, 1999 ; Kuhn et *al.*, 2005).

Dans le chloroplaste, les gènes de mêmes fonctions sont souvent organisés sous forme d'opéron et co-transcrits à partir d'un même promoteur (Mullet, 1993 ; Quesada-Vargas et *al.*, 2005). Les nombreux réarrangements géniques du gènome mitochondrial des plantes au cours de l'évolution ont conduit à la perte d'une organisation des gènes en unité fonctionnelle. En conséquence, la co-transcription dans la mitochondrie n'implique pas toujours des gènes de fonction apparentée et est moins fréquente que dans le chloroplaste (Schuster, 1993).

La terminaison de la transcription dans les organites de plante est encore méconnue. Alors que les transcrits de la bactérie et de la mitochondrie humaine possèdent des séquences répétées en orientation inverse formant une structure secondaire tige boucle en 3' qui stoppe la transcription, ces structures sont souvent absentes des transcrits des organites de plantes. Les tiges boucles présentes en 3' de certains transcrits mitochondriaux et chloroplastiques seraient des signaux de maturation de l'ARN plutôt que d'arrêt de la transcription (Hoffman et *al.*, 1999 ; Dombrowski et *al.*, 1997).

Des facteurs en *trans* pourraient également influencer la transcription de manière gène spécifique. Chez *Arabidopsis*, un mutant du gène codant la protéine pTAC2 de la famille PPR (pour pentatricopeptide repeat) montre un profile d'expression des gènes chloroplastiques similaire à celui d'un mutant déficient en ARN polymérase PEP (Pfalz et *al.*, 2006). Cette protéine a été identifié par co-purification de fractions chloroplastiques transcriptionnellement actives et semble être un élément important de la machine transcriptionnelle du chloroplaste.

La transcription dans les organites végétaux serait un mécanisme relâché et peu contrôlé où la taille des transcrits précurseurs serait déterminée par le choix du promoteur et la processivité de l'ARN polymérase.

Avant d'être traduits, les longs précurseurs d'ARN transcrits sont maturés au cours de différentes étapes post-transcriptionnelles. Des études menées sur les activités transcriptionnelles de certains gènes mitochondriaux et des taux relatifs d'ARNm codés par le génome mitochondrial d'*Arabidopsis* n'ont pas révélé de corrélation entre l'activité des

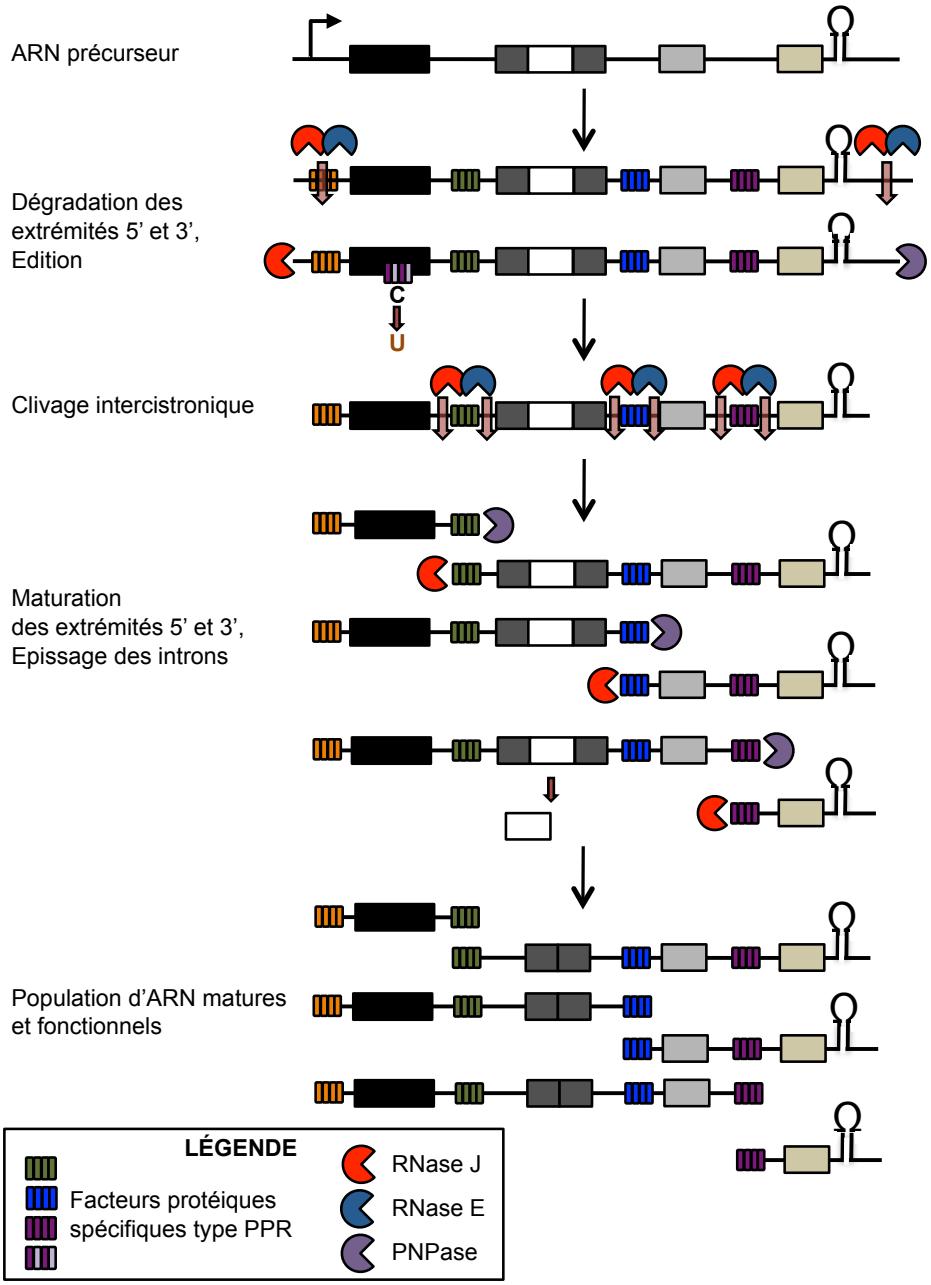


Figure 5: Modèle de maturation des ARN messagers dans les chloroplastes.

Un groupe de gènes est cotranscrit par un même promoteur en ARN polycistronique terminé par une structure tige boucle formée par des séquences inversées et répétées. Des protéines (par exemple de la famille PPR) lient l'ARN au niveau de sites spécifiques. La RNase J possède une activité endoribonucléase et exoribonucléase et aurait un rôle premier dans la définition de l'extrémité 5'. La RNase E à activité endoribonucléase cliverait les régions intergéniques et les extrémités 5' afin de les rendre accessible à la PNPase. Des structures secondaires ou des facteurs protéiques empêcheraient la progression des ribonucléases et définiraient les extrémités 5' et 3' des ARN matures (modifié d'après Stern et *al*., 2010). Dans les mitochondries, la maturation des précurseurs est plus obscure; la RNase E étant absente, l'activité endoribonucléique serait portée par la RNase P qui reconnaitrait les éléments t nombreux dans les régions 5' de ses gènes. Pour être fonctionnels, les ARNm nécessitent également d'être édités et épissés.

promoteurs et l'abondance de leurs transcrits respectifs (Giegé et *al.*, 2000; Holec et al., 2006). Ceci suggère que la stabilité et les taux des transcrits mitochondriaux seraient régulés par des mécanismes post-transcriptionnels.

2.2.2. LA LIBERATION DES EXTREMITES 5' ET 3' DES MESSAGERS

Les tailles des transcrits matures dans les organites diffèrent de celles des précurseurs transcrits suite à des étapes de réduction par dégradation des extrémités 5' et 3' couplées aux clivages intergéniques des ARN polycistroniques. Ces étapes requièrent trois types d'enzymes : des endoribonucléases, des 5'-3' exoribonucléases et 3'-5' exoribonucléases (Figure 5). Le génome des organites des plantes supérieurs ne code pas pour de telles enzymes qui sont donc d'origine nucléaire.

Deux mécanismes peuvent générer les extrémités 5' ou 3' : une activité exoribonucléase ou un clivage spécifique de type endoribonucléique. Chez les plantes supérieures, le mécanisme précis de dégradation 5'-3' des transcrits n'est pas totalement résolu. Deux protéines possédant une activité 3'-5' exoribonucléasique ont été caractérisées dans la mitochondrie : la RNase II qui est doublement adressée aux mitochondries et aux chloroplastes et une Polynucléotide Phosphoryplase ou PNPase (Hayes et *al.*, 1996 ; Perrin et *al.*, 2004 ; Kishine et *al.*, 2004).

Le chloroplaste des plantes supérieures est quand à lui doté de trois protéines ayant une activité endoribonucléasique : la RNase E, RNase J et CSP41 (Bollenbach et *al.*, 2007 ; Schein et *al.*, 2008). L'homologue bactérien de la RNase J présente une activité additionnelle 5'-3' exoribonucléasique (Mathy et *al.*, 2007). Celle-ci pourrait également jouer un rôle dans la maturation de l'extrémité 5' des transcrits chloroplastiques. Dans le chloroplaste, l'activité 3'-5' exoribonucléasique serait joué par une RNase II et une PNPase (Perrin et *al.*, 2004 ; Walter et *al.*, 2002).

Un modèle raisonnable de maturation des extrémités 5' et 3' des précurseurs dans les organites de plantes propose que l'extrémité 5' du transcrit précurseur soit soumise à des clivages endoribonucléiques successifs par la RNAse E ou J suivies par une dégradation 5'-3' par la RNAse J (Figure 5) (Stern et *al.*, 2010) des extrémités 5' sortantes et une dégradation des extrémités 3' sortantes par la PNPase et RNase II.

Dans les mitochondries, l'absence d'homologues aux RNase E/J rend le mécanisme de maturation des extrémités 5' incompris. Toutefois, les précurseurs formés par la co-transcription d'ARNm et d'ARNt ou de gènes possédant à leurs extrémités des structures mimant celle des ARNt (éléments t) peuvent être reconnus par les protéines de type RNase P et RNase Z impliquées dans la maturation des ARNt. Ces protéines clivent respectivement en 5' et 3' de la structure en trèfle ; ces éléments t étant nombreux dans le transcriptome mitochondrial (Forner et *al.*, 2007). Ce mécanisme de maturation est corroboré par une extension de type CCA communément additionnée à l'extrémité 3' des ARNt matures par une nucléotidyl-transférase qui peut être trouvée en 3' de certains

transcrits matures de la mitochondrie (Kunzmann et *al.*, 1998 ; Williams et *al.*, 2000 ; Forner et *al.*, 2007). De plus, une étude récente a montré qu'une protéine de la famille PPR, PRORP1 est adressée aux organites et possède une activité *in vitro* de type RNAse P en clivant les précurseurs ARNt mais également certains éléments t presents dans les transcrits mitochondriaux (Gobert et *al.*, 2010).

La fonction d'endoribonucléase dans la mitochondrie et les chloroplastes pourrait également être portée par certaines des protéines de la famille PPR qui possèdent une extension C-terminale (domaine DYW) ayant une activité endoribonucléasique *in vitro* (Nakamura et *al.*, 2008 ; Okuda et *al.*, 2009). Pour exemple, la protéine CRR2 est impliquée dans le clivage entre les gènes *rps7* et *ndhB* d'un précurseur tricistronique dans le chloroplaste d'*Arabidopsis* (Hashimoto et *al.*, 2003).

Dans le chloroplaste, la dégradation des extrémités 3' des transcrits précurseurs par les exoribonucléases (PNPase et RNase II) peut être stoppée par des séquences répétées en orientation inverse formant des structures en tige boucle. Ces structures seraient éliminées par clivages endoribonucléiques libérant l'extrémité 3' et permettant sa polyadénylation par la PNPase qui stimule la dégradation par cette enzyme ou par la RNase II. Ces structures secondaires sont souvent absentes du transcriptome mitochondrial et ne peuvent donc être considérées comme les signaux d'arrêt de la dégradation qui déterminent les extrémités 5' et 3' des transcrits matures.

Ces extrémités peuvent être définies par des facteurs protéiques d'origine nucléaire et adressés aux organites. Ces protéines lient l'ARN au niveau de séquences spécifiques et empêchent la progression des exoribonucléases. Plusieurs de ces protéines ont été identifiées dans les organites végétaux et la grande majorité de ceux-ci appartient à la famille des protéines PPR. Pour exemples :

Chez *Chlamydomonas*, les protéines MCA1 et MRL1 defininissent les extrémités 5' des transcrits chloroplastique *petA* et *rbcL* respectivement (Loiselay et *al.*, 2008 ; Johnson et *al.*, 2010). La protéine orthologue à MRL1 chez *Arabidopsis* joue le même rôle (Johnson et *al.*, 2010).

Chez *Arabidopsis*, la protéine chloroplastique HCF152 est impliquée dans la maturation du transcrit polycistronique *psbB-psbT-psbH-petB-petD* en empêchant la dégradation des extrémités 5' et 3' générées après le clivage endoribonucléique entre *psbH* et *petB* (Meierhoff et *al.*, 2003).

Tout comme HCF152, CRP1 chez le maïs définit les extrémités 3' et 5' respectives des gènes *petB* et *petD* (Barkan et *al.*, 1994) alors que la protéine PPR10 définit les extrémités 3' et 5' respectives des gènes *psaJ* et *atpH* (Pfalz et *al.*, 2010).

Dans la mitochondrie d'*Arabidopsis*, le facteur RPF2 définit les extrémités 5' des transcrits *nad9* et *cox3* (Jonietz et *al.*, 2010).

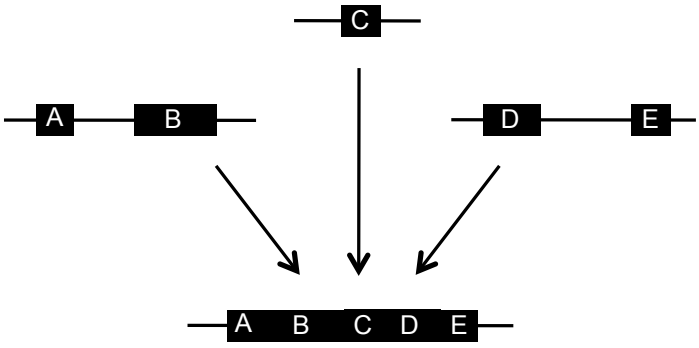


Figure 6: Représentation schématique de l'épissage en *cis* et *trans* des transcrits mitochondriaux *nad5* chez *Arabidopsis*.

Le transcrit mature *nad5* est généré à partir de trois précurseurs ARN via deux épissages en *cis* et en *trans*. A, B, C, D, E représentent les différents exons.

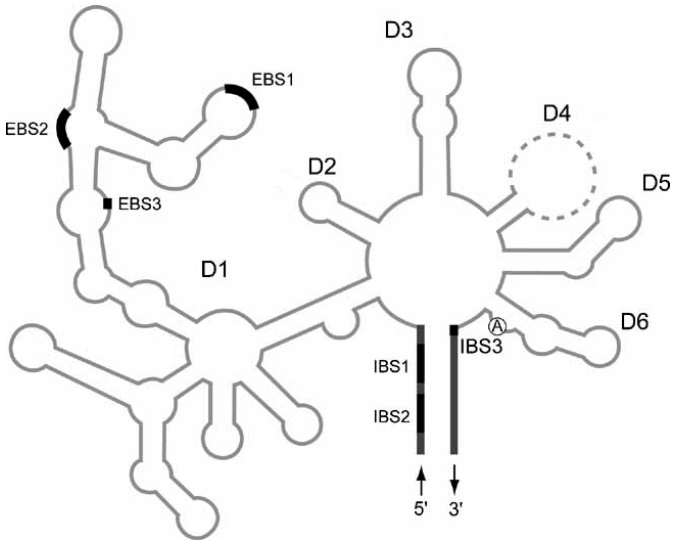


Figure 7: Structure secondaire typique d'un intron de groupe II.

Les séquences de l'intron et des exons sont représentées par des traits fins et épais respectivement. Les six domaines conservés sont représentés de D1 à D6. L'adénine du domaine 6 impliquée dans la première attaque nucléophile est entourée. EBS (pour « Exon Binding Site") et IBS (« Intron Binding Site") représentent les sites de contact complémentaires de l'exon et de l'intron permettant un appariement de bases nécessaire lors du mécanisme d'épissage (Glanz et al., 2009).

Ces facteurs de liaison à l'ARN codés par le noyau semblent donc être les acteurs majeurs du contrôle de la maturation des transcrits dans la mitochondrie et le chloroplaste des plantes.

2.2.3. EPISSAGE DES INTRONS

Tout comme les génomes eucaryotes, les génomes des organites de plante possèdent de nombreux introns qui doivent être excisés après transcription. Les introns sont classés en deux groupes selon leur structure secondaire et leur mode d'épissage.

Les génomes des organites des plantes supérieures possèdent deux types d'introns : les introns de groupe I et ceux de groupe II (Unseld et *al.*, 1997 ; Sato et *al.*, 1999). Cependant, ces introns sont principalement du groupe II (Bonen et Vogel, 2001 ; Barkan, 2004). Dans les organites de plantes, un intron de groupe I probablement acquis par transfert horizontal est trouvé dans le géne *coxI* de la mitochondrie de *Peperomia* et deux introns de groupe I sont présents dans le gène chloroplastique *trnL*-UAA des plantes terrestres (Cho et *al.*, 1998 ; Suguira, 1992).

Dans les mitochondries des plantes, les introns sont principalement localisés dans des gènes codant des protéines (Bonen et Vogel, 2001 ; Lambowitz et Zimmerly, 2004) alors qu'ils interrompent également les gènes codants les ARNt dans le chloroplaste (Tillich et Krause, 2010). Au total, on dénombre une quarantaine d'introns de groupe II dans les organites des plantes supérieures.

L'arrangement des introns d'un même gène peut être sous différentes configurations. La séquence intronique est flanquée par les exons du gène et un seul précurseur ARN est transcrit, on parle alors de configuration en *cis* ou alors le gène est fragmenté en plusieurs exons codés à différents loci du génome. Dans ce dernier cas, plusieurs précurseurs ARN sont transcrits et doivent être joints ; on parle de configuration en *trans*. Ces deux types d'arrangements donnent lieu à deux mécanismes d'épissage des introns dits en *cis* ou en *trans*, les deux pouvant être combinés afin d'obtenir un ARN fonctionnel (Chapdelaine et Bonen, 1991 ; Knoop et *al.*, 1991 ; Glanz et Kück, 2009). Pour exemple, le gène mitochondrial *nad5* est codé par 5 exons (A-E) (Figure 6): les exons A/B, C et D/E sont codés au niveau de différents loci donnant naissance à 3 précurseurs indépendants. Les exons A, B et D, E sont séparés par un intron en *cis*. La formation du transcrit *nad5* fonctionnel nécessite donc deux épissages en *cis* et deux épissages en *trans*.

Les introns de groupe II ont une structure secondaire conservée comportant 6 domaines hélicoïdaux connectés à un noyau central (Figure 7) (Fedorova et Zingler, 2007).

Ces introns sont généralement excisés en deux étapes de *trans*-estérification. L'extrémité 5' de l'intron subit une première attaque nucléophile par un groupement 2'-OH d'une adénosine non appariée du domaine VI établissant une liaison temporaire. Une seconde attaque nucléophile du groupement 3'-OH libre du premier exon sur l'extrémité 3' de

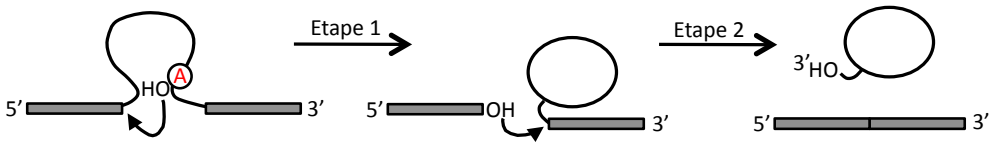


Figure 8: Mécanisme d'épissage des introns de groupe II.

Deux étapes successives de transestérification représentées par des flèches permettent la libération de l'intron sous forme de lasso.

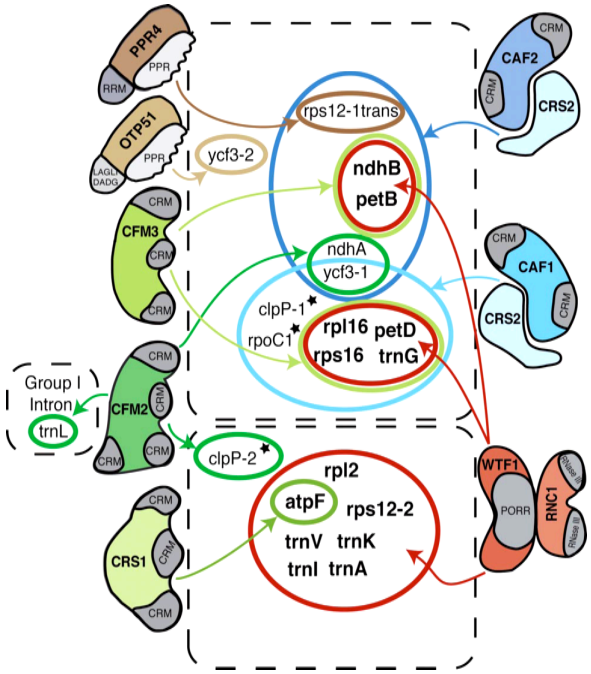


Figure 9: Facteurs nucléaires impliqués dans l'épissage des introns chloroplastiques de groupe II chez les plantes supérieures.

Les introns présents chez *Arabidopsis* mais absents chez le maïs sont marqués d'une étoile. Les facteurs sont annotés avec leur domaine protéique. Les maturases et le facteur ZmWHY1 ne sont pas indiqués. Le schéma reflète la complexité et l'aspect coopératif de la machinerie d'épissage des introns dans le chloroplaste des plantes supérieures (d'après Kroeger et *al*., 2009).

l'intron libère ce dernier sous forme de lasso (Figure 8) (Peebles et *al.*, 1986; Schmelzer et Schweyen, 1986; van der Veen et *al.*, 1986).

Contrairement aux introns de groupe II des bactéries et des levures qui sont capables de s'autoépisser *in vitro*, l'activité autocatalytique des introns des organites des plantes terrestres n'a pu être démontrée *in vitro* (Michel et Ferat, 1995 ; Vogel et Borner, 2002). Ceci suggère que des facteurs protéiques soient nécessaires pour l'épissage des introns probablement en facilitant leur repliement en une structure catalytiquement active (Lambowitz et Zimmerly, 2004 ; Barkan 2004). Certains de ces facteurs appelés maturases sont codés par des ORFs présents dans les introns de groupe II des bactéries (Wank et *al.*, 1999 ; Meng et *al.*, 2005). Chez les plantes, les protéines MatR et MatK sont codées respectivement par le quatrième intron du gène *nad1* de la mitochondrie et par l'intron du gène *trnK* du chloroplaste. La fonction de ces protéines est suggérée par des motifs conservés avec les maturases de bactérie et de levure impliquées dans l'épissage des introns de groupe II (Lambowitz et Zimmerly, 2004). Des expériences d'immunoprécipitation de l'ARN ont récemment montré que la protéine MatK est associée *in vivo* avec la plupart des introns de groupe II du chloroplaste du tabac (Zoschke et *al.*, 2009).

Chez *Arabidopsis*, quatres gènes nucléaires ; AtnMat-1a, AtnMat-1b, AtnMat-2a et AtnMat-2b codent des protéines présentant des similarités avec les maturases et possédant des séquences d'adressage aux organites (Mohr et Lambowitz, 2003). Des études génétiques ont confirmées le rôle des protéines AtnMat-1a et AtnMat-1b dans l'épissage des transcrits mitochondriaux *nad4* et *cox2*, *nad1*, *nad7* respectivement (Nakagawa et *al.*, 2006 ; Keren et *al.*, 2009).

Des approches de génétique directe chez le maïs et *Arabidopsis*, ont permis d'identifier de nombreux facteurs d'épissage des introns de groupe II chloroplastiques (Figure 9) (Jenkins et *al.*, 1997 ; Jenkins et Barkan 2001 ; Ostheimer et *al.*, 2003 ; Ostersetzer et *al.*, 2005 ; Asakura and Barkan 2006 ; Schmitz-Linneweber et *al.*, 2006 ; Asakura et Barkan, 2007 ; Watkins et *al.*, 2007 ; Asakura et *al.*, 2008 ; Keren et *al.*, 2008 ; Prikryl et *al.*, 2008 ; Williams-Carrier et *al.*, 2008 ; Kroeger et *al.*, 2009).

Certains de ces facteurs appartiennent à une famille de protéines présentant des domaines CRM (pour chloroplast RNA splicing and ribosome maturation) de liaison à l'ARN (Ostheimer et *al.*, 2003). Chez les procaryotes, les protéines à domaines CRM sont impliquées dans la maturation des ARN ribosomiques (Barkan et *al.*, 2007).

Chez le maïs, des expériences de co-immunoprécipitation de certaines protéines CRM ont permis l'identification des nouveaux facteurs d'épissage RNC1, ZmWHY1 et WTF1. Ceci suggère un rôle coopératif de ces facteurs dans l'épissage des introns (Watkins et *al.*, 2007; Prikryl et *al.*, 2008; Kroeger et *al.*, 2009)

Plusieurs protéines des familles CRM et WTF possèdent des séquences d'adressage aux mitochondries et leurs rôles dans l'épissage des introns pourraient être conservés dans ces dernières (Barkan et *al.*, 2007 ; Kroeger et *al.*, 2009).

Les protéines PPR4, PPR5, OTP51 et OTP43 de la famille PPR sont également nécessaires à l'épissage de certains introns chloroplastiques et mitochondriaux (Schmitz-Linneweber et *al.*, 2006 ; Beick et *al.*, 2008 ; de Longevialle et *al.*, 2008 ; de Longevialle et *al.*, 2007).

2.2.4. L'EDITION DE L'ARN

Le terme édition de l'ARN désigne un ensemble de mécanismes qui modifient la séquence d'un ARN par rapport à celle codée par son gène. Ce terme "édition de l'ARN" a été pour la première fois utilisé pour décrire des insertions et plus rarement des délétions d'uridines dans certains ARNm des mitochondries de trypanosome (Benne et *al.*, 1986 ; Stuart et *al.*, 2005). Plus tard, cette désignation fut utilisée pour décrire des processus post-transcriptionnels dans les mitochondries et les chloroplastes des plantes (Covello et Gray, 1989 ; Gualberto et *al.*, 1989 ; Hiesel et *al.*, 1989 ; Hoch et *al.*, 1991) et des modifications post-transcriptionnelles de certains transcrits chez les mammifères et des précurseurs de micro ARN (Nishikura, 2010).

Chez les plantes terrestres, l'édition des transcrits des organites correspond principalement à la modification spécifique de cytidines en uridines. La conversion inverse d'uridines en cytidines est toutefois observée dans les organites de certaines espèces de mousses et fougères (Duff et *al.*, 2005 ; Wolf et *al.*, 2004 ; Knoop, 2004 ; Malek et *al.*, 1996). Les plantes hépatiques de la sous classe Marchantiidae sont les seules plantes terrestres chez qui l'édition de transcrits dans les organites n'a pas été observé.

L'édition de l'ARN affecte 34 cytidines dans le chloroplaste d'*Arabidopsis* et plus de 500 dans la mitochondrie (Chateigner-Boutin et Small, 2007 ; Giegé et Brennicke, 1999 ; Bentolila et *al.*, 2008 ; Zehrmann et *al.*, 2008)

La plupart des sites d'édition identifiés sont localisés dans les ARN codant des protéines où ils ont été identifiés par comparaison de la séquence des génes à celles de leur ADN complémentaires. Certains sites d'édition sont présents dans les ARNt, les séquences non codantes transcrites et les introns. Aucun site n'a été identifié dans les ARN ribosomiques des organites.

L'édition est un événement post-transcriptionnel qui précéderait l'épissage des introns et le clivage intercistronique des précurseurs ARN (Del Campo et *al.*, 2002 ; Freyer et *al.*, 1993 ; Ruf et *al.*, 1994 ; Carillo et Bonen, 1997). L'édition du précurseur polycistronique préviendrait sa traduction précoce en protéine non fonctionnelle.

Ce mécanisme post-transcriptionnel n'est pas toujours total. Certains sites d'édition sont partiellement édités conduisant à l'existence de deux populations de transcrits au sein de l'organite (Chateigner-Boutin et Hanson, 2003 ; Zehrmann et *al.*, 2008 ; Bentolila et *al.*,

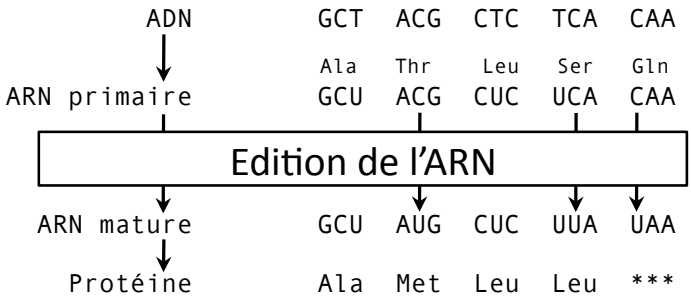


Figure 10: Edition de l'ARN dans les organites des plantes.

L'édition est un mécanisme post-transcriptionnel qui convertit des cytidines spécifiques (C) en uridine (U) de certains transcrits. La séquence protéique codée par le transcrit édité est différente de celle du transcrit non édité. Dans l'exemple présenté, l'édition altère l'identité de deux acides aminés et crée un codon terminateur (modifié d'après Takenaka et *al.*, 2008).

Figure 11: Alignement partiel des protéiques RPS14 (protéine ribosomale) codées par les génomes chloroplastiques de plantes et bactéries.

La séquence protéique RPS14 d'*Arabidopsis* a été alignée avec les homologue des autres espèces. L'acide aminé en position 27 et affecté par l'édition est indiqué par une flèche. L'édition du transcrit permet la restauration de l'acide aminé Leucine conservé au cours de l'évolution.

2008). Lorsque les transcrits présentent plusieurs sites d'édition, de multiples formes du transcrit peuvent également coexister dans l'organite (Schuster et *al.*, 1990).

Le nombre de sites d'édition par gènes est variable. Pour exemples, dans le chloroplaste d'*Arabidopsis*, un quart des sites d'édition sont présents dans le gène *ndhB* et les 11 gènes de la famille *ndh* (NADH déshydrogénase) regroupent presque la moitié des sites d'édition (Chateigner-Boutin et Small, 2007). Dans la mitochondrie d'*Arabidopsis*, les transcrits codant des protéines du complexe I (NADH ubiquinone oxydo-réductase) et des complexes CCM (impliqués dans la maturation du cytochrome *c*) ont une fréquence d'édition plus élevée que les autres gènes (Giegé et Brennicke, 1999).

L'édition des cytidines au niveau des ARNm a souvent pour effet de changer l'identité de l'acide aminé codé par le gène (Figure 10). Cependant, l'édition de la troisième base du codon a toujours un effet silencieux sur l'acide aminé codé.

Dans la majorité des cas, l'édition des ARNm permet de restorer un acide-aminé qui est conservé dans les protéines homologues des autres espèces de plantes ou des bactéries (Figure 11). Cet acide aminé hautement conservé semble être essentiel pour l'activité biochimique de la protéine (Bock et *al.*, 1994 ; Sasaki et *al.*, 2001).

L'édition peut être une étape post-transcriptionnelle nécessaire à la traduction en créant le codon initiateur AUG à partir du codon ACG ou le codon terminateur de certains transcrits des organites (Chapdelaine et Bonen, 1991 ; Hoch et *al.*, 1991, Wintz et Hanson, 1991).

Chez le maïs, l'édition d'un intron du transcrit mitochondrial *nad7*, permet son repliement correct et la formation d'une structure secondaire nécessaire à un épissage efficace (Carrillo et Bonen, 1997). De même, l'édition au niveau des précurseurs d'ARNs de transfert permet un appariement de bases nécessaire à leur repliement et leur maturation (Maréchal-Drouard et *al.*, 1996a ; Maréchal-Drouard et *al.*, 1996b ; Kunzmann et *al.*, 1998).

L'édition partielle de certains sites génère l'existence des formes éditées et non éditées du transcrit. Ces transcrits conduisent à la synthèse de différentes formes de la protéine. Cependant, des études protéomiques et immunologiques ont montré que seules les protéines résultant des transcrits édités s'accumulent au sein des complexes fonctionnels dans l'organite (Grohman et *al.*, 1994 ; Lu et Hanson, 1994 ; Phreaner et *al.*, 1996). Ces résultats suggèrent que la traduction des transcrits non édités puisse être inhibée et/ou que les protéines résultant de la traduction de ces transcrits soient instables et rapidement dégradées.

2.2.4.1. LE MECANISME BIOCHIMIQUE D'EDITION

Des expériences d'édition *in vitro* qui consistent à incuber des transcrits marqués radioactivement en présence d'un extrait mitochondrial de pomme de terre ont révélées l'absence de libération de groupement Phosphate [alpha] durant la réaction d'édition. Ce résultat élimine la possible excision de ribonucléotides au cours de la réaction d'édition

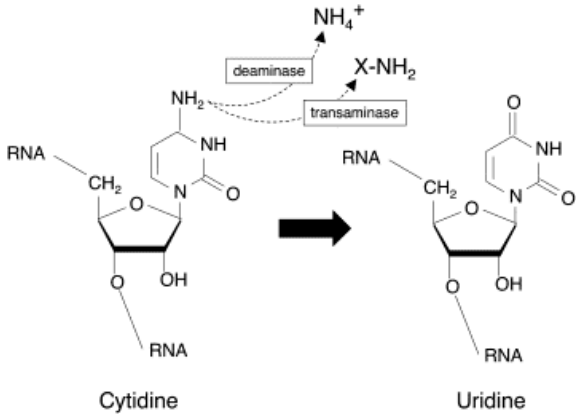


Figure 12: Mécanismes biochimiques possibles de l'édition dans les organites de plantes.

L'édition est un mécanisme de désamination des cytidines en uridines. Le groupement amine de la cytidine peut être supprimé par une enzyme de type cytidine désaminase ("deaminase") ou transféré à un accepteur par une transaminase (d'après Takenaka et *al*., 2008).

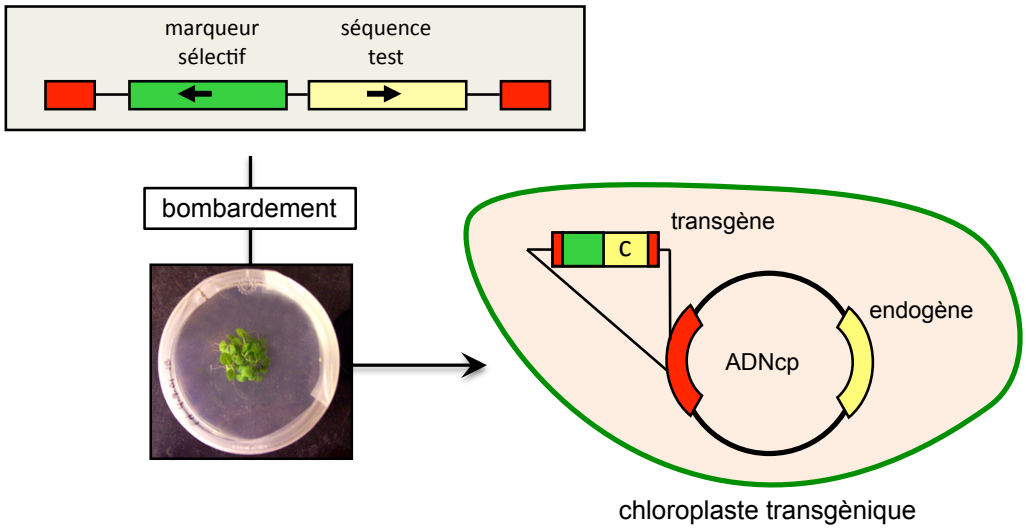


Figure 13: Méthode biolistique de transformation des chloroplastes de feuilles de tabac.

Après bombardement des tissus végétaux avec la construction d'intérêt, les feuilles sont placées sur un milieu sélectif afin de régénérer un calle puis une plantule transgènique. Cette méthode fut développée par Svab et Maliga, 1990.

dans les organites de plantes (Rajasekhar et Mulligan, 1993 ; Yu est Schuster, 1995 ; Hirose et Suguira, 2001). Deux mécanismes plus simples peuvent convertir une cytidine en uridine : une réaction de désamination ou de transamination. Dans ces cas, la molécule d'ARN reste intacte (Figure 12). Ces deux réactions impliquent respectivement une enzyme de type cytidine désaminase et transaminase. L'hypothèse de l'implication d'une cytidine désaminase dans le chloroplaste est privilégiée. Ces enzymes possèdent un domaine typique de liaison au zinc influençant son activité. Des réactions d'édition *in vitro* à l'aide d'extraits chloroplastiques ont montrées que l'activité d'édition est sensible à la présence de zinc (Hegeman et *al.*, 2005). Néanmoins, ceci n'est pas le cas dans les mitochondries et suggère que les enzymes catalysant la réaction d'édition dans les mitochondries et chloroplastes puissent être différentes (Takenaka et Brennicke, 2003).

Les génomes des organites des plantes ne codent pas pour de telles enzymes. L'activité de déamination ou transamination serait donc portée par une ou plusieurs protéines codées par le noyau et adressées aux chloroplastes et/ou mitochondries.

Une famille de neuf gènes nucléaires code pour des cytidine désaminases chez *Arabidopsis*. L'étude d'un de ces membres AtCDA1, a confirmé son activité désaminase *in vitro* mais la localisation cytoplasmique *in vivo* de la protéine n'a pu permettre de lui assigner un rôle dans les organites (Faivre-Nitschke et *al.*, 1999).

Jusqu'à ce jour, les tentatives d'identifier l'enzyme catalysant la réaction d'édition dans les organites de plantes se sont toutes révélées infructueuses. L'identité de cette enzyme est donc toujours un mystère (Shikanai, 2006 ; Takenaka et *al.*, 2007).

2.2.4.2. LES ELEMENTS *CIS*

34 sites et plus de 500 sites d'édition sont présents dans le génome mitochondrial et chloroplastique d'*Arabidopsis* respectivement (Chateigner-Boutin et Small, 2007 ; Giegé et Brennicke, 1999 ; Bentolila et *al.*, 2008 ; Zehrmann et *al.*, 2008). Un signal dans chaque transcrit doit donc permettre à la machinerie d'édition de cibler ces cytidines de manière spécifique. L'analyse de l'ensemble des sites d'édition chloroplastiques ou mitochondriaux ne révèle pas de séquences nucléotidiques consensus entourant les sites d'édition à l'exception de l'une ou des deux bases adjacentes au site d'édition (Giegé et Brennicke, 1999 ; Mulligan et *al.*, 1999 ; Cummings et Myers, 2004). Chaque site d'édition ou un groupe de sites d'édition présenterait donc un signal indépendamment reconnu par un facteur d'édition spécifique agissant en *trans*. Si tel est le cas, le chloroplaste et la mitochondrie d'*Arabidopsis* nécessiteraient un grand nombre de *trans*-facteurs.

Des expériences de transformation du génome chloroplastique chez le tabac ont permis l'étude des éléments *cis* (Figure 13). Des tabacs transgéniques exprimant un minigène dans le chloroplaste du tabac qui code les 98 nucléotides entourant le site d'édition *psbL* présente une diminution de l'efficacité d'édition du site *psbL* endogène. Ceci indique qu'un *trans*-facteur reconnaît la séquence en *cis* introduite et qu'il est présent en quantité limitante dans l'organite (Chaudhuri et *al.*, 1995).

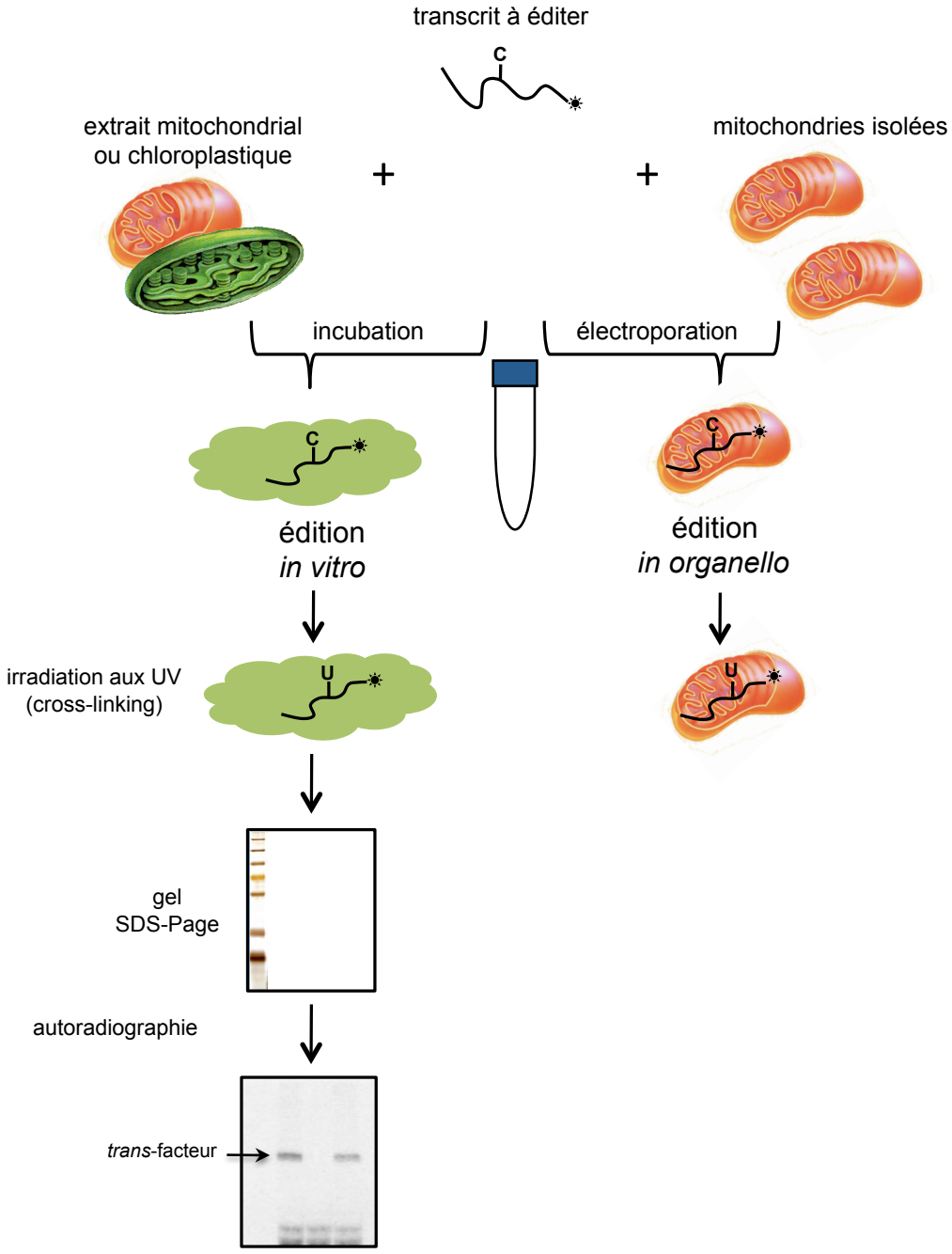


Figure 14: Principe des expériences d'édition *in vitro* et *in organello* chez les plantes.

Dans le cas de l'édition *in vitro*, le transcrit contenant une cytidine à éditer est incubé avec un extrait mitochondrial ou chloroplastique obtenu à partir d'organites purifiés (Farré et Araya, 2001; Takenaka et Brennicke, 2003; Hirose et Suguira, 2001). Durant la réaction d'édition, le *trans*-facteur qui lie le transcrit radiomarqué peut être piégé par irradiation aux UV du milieu réactionnel puis visualisé par autoradiographie après migration sur gel SDS-Page (Hirose et Sugiura, 2001).

Dans le cas de l'édition *in organello*, les mitochondries purifiées intactes sont électroporées en présence du transcrit à éditer (Farré et Araya, 2001; Choury et Araya, 2003).

Plus tard, Chateigner-Boutin et Hanson, 2002 ont réalisé le même type d'expérience en surexprimant des minigènes dans le chloroplaste du tabac. Ces minigènes entourent les sites d'édition *rpoB-2* ou *ndhF-2*. De même que les résultats mentionnés précédemment, leur expression dans le chloroplaste a un effet délétère sur l'édition des sites endogènes et plus surprenant, d'autres sites d'édition sont touchés. Ces autres sites d'édition partagent des nucléotides conservés avec les sites *rpoB-2* ou *ndhF-2*. Les auteurs proposent que ces séquences partagent une information commune qui recruterait un même *trans*-facteur. Un même *trans*-facteur pourrait donc reconnaître plusieurs sites d'édition. Ceci diminuerait le nombre de facteurs nécessaires à la reconnaissance des sites d'édition dans les organites.

Des expériences intensives d'édition *in vitro* et *in organello* en utilisant des extraits chloroplastiques et mitochondriaux ou des mitochondries isolés (Figure 14) ont permis de déterminer les séquences nucléotidiques minimales en 5' et 3' de certains sites qui sont requises à leur édition. Pour les différents sites testés, les éléments-*cis* en amont du site d'édition et nécessaires à son édition se limitent à environ 30 nucléotides (Farré et *al.*, 2001 ; Staudinger et *al.*, 2005 ; Miyamoto et *al.*, 2002 ; Hegeman et *al.*, 2005 ; Hayes et *al.*, 2006 ; Neuwirt et *al.*, 2005 ; Takenaka et *al.*, 2004, 2008 ; Verbitskiy et *al.*, 2008). Pour certains sites d'édition, une courte séquence de 5 ou 6 nucléotides en aval du site est également requise (Farré et *al.*, 2001 ; Choury et *al.*, 2004).

Des mutations ponctuelles ou plus larges dans la séquence nucléotidique de l'élément *cis* peuvent réduire ou totalement abolir l'édition d'un transcrit (Reed et *al.*, 2001 ; Neuwirt et *al.*, 2005 ; Sasaki et *al.*, 2006). Des expériences d'édition *in vitro* en présence de compétiteurs ARN couvrant les différentes parties de l'élément *cis* ont mis en évidence des groupes de nucléotides requis pour l'édition efficace de certains sites chloroplastiques du tabac (Kobayashi et *al.*, 2008).

2.2.4.3. LES FACTEURS AGISSANT EN *TRANS*

Les éléments *cis* seraient reconnus par des facteurs interagissant en *trans* (ou *trans*-facteurs) qui guideraient la machinerie d'édition dans la reconnaissance spécifique de chaque site d'édition. La nature de ces facteurs peut être de deux types : acide nucléique ou protéine. Des études biochimiques et génétiques sur le chloroplaste du tabac ont montré que ces facteurs sont des protéines codées par le noyau (Figure 15) (Bock et Koop, 1997 ; Hirose et Sugiura, 2001 ; Miyamoto et *al.*, 2002, 2004 ; Kobayashi et *al.*, 2008).

Chez le tabac, un système d'édition *in vitro* développé par Hirose et Suguira, 2001 a permis une avancée importante dans l'identification des *trans*-facteurs (Figure 14). Un substrat ARN radiomarqué et présentant un site d'édition est mis en présence d'un extrait chloroplastique. Les protéines liant le substrat ARN sont par la suite piégées par irradiation aux UV et révélées par autoradiographie après une migration sur gel SDS-PAGE. Ainsi, des protéines de poids moléculaires distincts de 25, 56, 70, 91 et 93 kDa lient spécifiquement les éléments *cis* des transcrits chloroplastiques *psbL*, *psbE*, *petB*, *rpoB* et *rpoA*, respectivement (Hirose et Sugiura, 2001 ; Miyamoto et *al.*, 2002, 2004 ; Kobayashi

Nom	Espèce	Famille	Site(s) d'édition cible(s)	Référence
Chloroplaste				
OTP80	*Arabidopsis*	PPR	*rpl23 (86056)*	Hammani et *al*. (2009)
OTP81			*rps12 intron (69553)*	
OTP82			*ndhG (118858), ndhB (95644)*	Okuda et *al*. (2010)
OTP84			*ndhF (112349), psbZ (35800), ndhB (94999)*	Hammani et *al*. (2009)
OTP85			*ndhD (116494)*	
OTP86			*rps14 (37161)*	
CRR22			*ndhB (96419), ndhD (116281), rpoB (25779)*	Okuda et *al*. (2009)
CRR28			*ndhB (96698), ndhD (116290)*	
YS1			*rpoB (25992)*	Zhou et *al*. (2009)
RARE1			*accD (57868)*	Robbins et *al*. (2009)
LPA66			*psbF (63985)*	Cai et *al*. (2009)
CRR4			*ndhD (117166)*	Kotera et *al*. (2005)
CLB19			*rpoA (78691), clpP (69942)*	Chateigner-Boutin et *al*. (2008)
CRR21			*ndhD (116785)*	Okuda et *al*. (2007)
CP31a		RRM	*Multiple*	Tillich et *al*. (2009)
CP31b				
Mitochondrie				
MEF1	*Arabidopsis*	PPR	*rps4-956, nad7-963, nad2-1160*	Zehrmann et *al*. (2009)
MEF9			*nad7-200*	Takenaka (2010)
MEF11			*cox3-422, nad4-124, ccb203-344*	Verbitskiy et *al*. (2009)
OGR1	*Oryza sativa*		*cox2-167, cox3-572, ccmC-458, nad2-1457, nad4-401, nad4-416, nad4-433*	Kim et *al*. (2009)
PpPPR_71	*Physcomitrella*		*ccmF$_c$-122*	Tasaki et *al*. (2010)

Tableau 1 : Liste des différents facteurs d'édition et de leur(s) site(s) d'édition cible(s) connus des organites de plantes.

Le(s) site(s) d'édition chloroplastiques sont indiqués selon leur position dans la séquence nucléotidique du génome chloroplastique d'*Arabidopsis* (numéro d'accession GenBank : NP_000423). Le(s) site(s) d'édition mitochondriaux sont définis selon leur position dans la séquence codante du gène (CDS). Les cases bleues (OTP80, 81, 82, 84, 85, 86) indiquent les facteurs d'édition chloroplastiques identifiés lors de mes travaux de thèse et la case verte (CRR4) indique l'unique facteur d'édition connu au début de ma thèse en 2006.

et *al.*, 2008). Une même protéine de 95 kDa lie spécifiquement les éléments *cis* des deux transcrits distincts *ndhB-9* et *ndhF-2*. Cette preuve biochimique atteste l'hypothèse qu'un unique facteur puisse reconnaître plusieurs sites d'édition dans le chloroplaste (Kobayashi et *al.*, 2008).

Toutefois, ces expériences *in vitro*, ne révèlent pas l'identité de ces protéines. L'identification de ces facteurs provient d'études de génétique directe chez *Arabidopsis*. Ces études permettent d'associer le phénotype particulier d'une plante mutante à un gène nucléaire spécifique. Les phénotypes sélectionnés au sein d'une population de mutants sont de manière générale ceux liés à un dysfonctionnement du chloroplaste (pour exemples, défauts de fluorescence et/ou de pigmentation des plantes). Kotera et *al.*, 2005 ont ainsi identifié le premier facteur CRR4 (pour chlororespiratory reduction) impliqué dans l'édition du site *ndhD-1*. Ce site d'édition permet la création du codon d'initiation du gène *ndhD*. Dans la plante mutante *crr4*, l'édition du site est complètement abolie et l'absence de la protéine NdhD non traduite déstabilise l'assemblage du complexe NDH (NADH déshydrogénase) dans le chloroplaste.

Par la suite de nombreuses études génétiques ont identifié de nouveaux facteurs d'édition spécifiques d'un ou plusieurs sites d'édition dans les mitochondries et chloroplastes d'*Arabidopsis*, du riz et des mousses (Tableau 1).

Les preuves que ces facteurs d'édition lient les transcrits à éditer ont été apportées par deux études biochimiques des protéines CRR4 et PpPPR_71 (cette dernière est impliquée dans l'édition de l'un des sites (*ccmF-2*) du transcrit mitochondrial *ccmFc* chez les mousses). Des expériences de gel retard ont montré que les protéines recombinantes bactériennes ont la capacité de lier *in vitro* les éléments *cis* (-25/+10) et (-40/+5) des sites d'édition *ndhD-1* et *ccmF-2* respectivement (Figure 16) (Okuda et *al.*, 2006 ; Tasaki et *al.*, 2010).

A l'exception des protéines CP31a et CP31b, l'ensemble des facteurs d'édition connus chez les plantes appartient à un sous groupe de la famille des protéines PPR. Ces protéines sont en grand nombre chez les plantes (450 membres chez *Arabidopsis*) et sont des candidats idéals aux rôles de facteurs d'édition dans les organites.

2.2.4.4. ROLE DE L'EDITION

Comme mentionné précédemment, l'édition des transcrits restaure très fréquemment dans les protéines un acide aminé conservé chez les autres espèces de plantes et chez leurs homologues bactériens. Cette observation a permis de suggérer que l'un des rôles premiers de l'édition dans les organites de plantes serait de corriger les mutations génomiques T en C apparues au cours de la colonisation des terres par les plantes et de permettre ainsi la traduction de protéines fonctionnelles (Shikanai et *al.*, 2006 ; Takenaka et *al.*, 2007).

On peut penser que la création du codon initiateur ou terminateur par l'édition du transcrit, soit un mode de régulation de la traduction des gènes. Dans certains tissus, certains sites ne sont pas édités et ne donnent pas lieu à la synthèse de protéines (cas du codon initiateur) ou

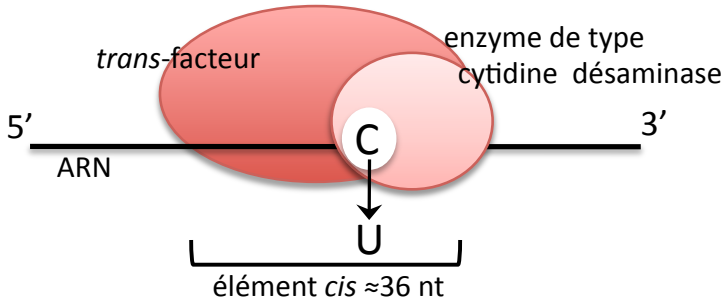


Figure 15: Modèle de reconnaissance d'un site d'édition par un *trans*-facteur spécifique dans les organites de plantes.

Un *trans*-facteur protéique reconnaitrait et lierait spécifiquement la séquence entourant le site d'édition ou élément *cis*. Le *trans*-facteur fixé à l'élément *cis* recruterait l'enzyme portant l'activité catalytique d'édition (modifié d'après Okuda et *al*., 2006).

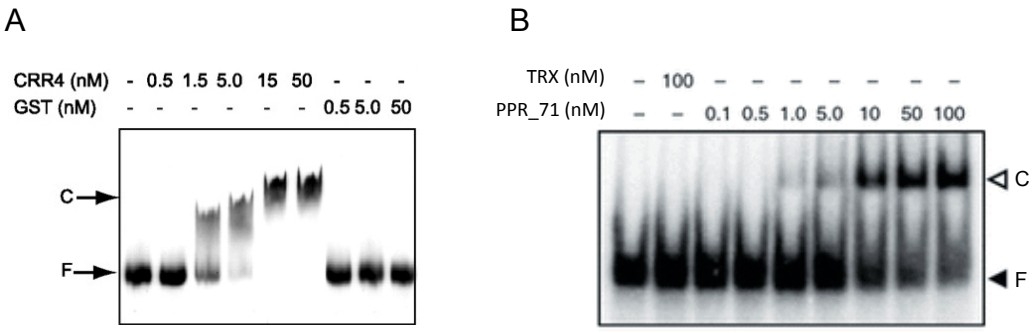


Figure 16: Expériences de gel retard des subtrats ARN en présence des *trans*-facteurs recombinants CRR4 et PPR_71.

Différentes concentrations de protéines bactériennes recombinantes CRR4 et PPR_71 fusionnées avec la glutathion-S-transférase (GST) ou thiorédoxine (TRX) ont été incubées respectivement avec les 36 nt entourant le site d'édition *ndhD-1* (CRR4) et les 46 nt entourant *ccmF-2* (PPR_71) avant leur migration sur gel. La liaison de la protéine à l'ARN forme un complexe **C** retardé sur gel par rapport à l'ARN libre **F** (d'après Okuda et *al*., 2006 ; Tasaki et *al*., 2010).

à une synthèse non terminée (cas du codon terminateur). Ainsi, le site d'édition *ndhD-1* créant le codon initiateur du transcrit n'est pas édité dans les racines et la protéine NdhD y est absente (Chateigner-Boutin et Hanson, 2003; Kotera et *al.*, 2005). Jusqu'à ce jour, aucune corrélation entre le niveau d'expression d'un *trans*-facteur et le niveau d'édition de son site respectif n'a pu être mise en évidence (Kotera et *al.*, 2005). L'édition est dépendante de facteurs protéiques codés par le noyau mais son efficacité serait limitée par d'autres éléments.

Après maturation des précurseurs d'ARN, les transcrits matures sont pris en charge par la machinerie traductionnelle.

2.3. LA TRADUCTION

La traduction dans les organites nécessite une machinerie complète composée des ribosomes, des facteurs de traduction (facteurs d'initiation, d'élongation et de terminaison), des ARNt et des aminoacyl-ARNt synthétases. De plus, comme pour les différentes étapes post-transcriptionnelles, la traduction dans les organites nécessite des facteurs codés par le noyau qui réguleraient de manière spécifique la traduction des gènes. Durant l'évolution des organites au sein de la cellule hôte, de nombreux gènes ont été transférés vers le noyau. En conséquence, le noyau code aujourd'hui pour une partie de la machinerie traductionnelle des organites.

Les génomes des organites des plantes supérieures codent, pour la totalité des ARN ribosomaux, une partie des protéines ribosomales et un set incomplet d'ARNt. Les protéines ribosomales et les ARNt manquants, ainsi que l'ensemble des aminoacyl-ARNt synthétases et facteurs de traduction sont importés du cytoplasme (Duchêne et *al.*, 2005; Salinas et *al.*, 2008).

Dans le chloroplaste, des séquences Shine et Dalgarno (SD) sont présentes en amont de l'AUG initiateur de la majorité des transcrits et permettent à la sous unité 30S de s'apparier (Sugiura et *al.*, 1998). Toutefois, l'utilisation d'un système de traduction *in vitro* de chloroplaste de tabac et des expériences de transformation de chloroplastes chez *Chlamydomonas* ont montré que la conservation de ces séquences SD ne sont pas toujours essentielles à la traduction des *trans*crits (Kim et *al.*, 1994 ; Mayfield et *al.*, 1994 ; Sakamoto et *al.*, 1994 ; Fargo et *al.*, 1998). Les séquences SD sont très rares dans les *trans*crits des mitochondries. L'indisponibilité d'un système de traduction *in vitro* pour la mitochondrie n'a pu confirmer le rôle des ces séquence dans la traduction (Pring et *al.*, 1992).

AUG est le codon initiateur typique dans les organites de plantes. La traduction peut cependant être initiée au niveau de codons alternatifs dans la mitochondrie et le chloroplaste (Bock et *al.*, 1994 ; Siculella et *al.*, 1996 ; Dong et *al.*, 1998 ; Kuroda et *al.*, 2007).

La traduction dans les chloroplastes et les mitochondries ne semble donc pas totalement identique à celle des bactéries et présente des caractéristiques propres et non totalement élucidées.

Dans le chloroplaste du maïs, une protéine CRP1 de la famille PPR joue le rôle d'activateur de la traduction des deux transcrits chloroplastiques *petA* et *psaC* en liant leur région 5'-UTR (Schmitz-Linneweber et *al.*, 2005). Le mode d'activation de la traduction par CRP1 est obscur mais il est raisonnable de penser que ce facteur définisse et stabilise les extrémités 5' de ces transcrits en les protégeant de la dégradation par des exoribonucléases selon le mécanisme développé paragraphe 2.2.2. Les extrémités 5'-UTR des transcrits permettent le recrutement de la machinerie traductionnelle. La propriété de certaines protéines PPR à définir et stabiliser les extrémités 5' et 3' des transcrits dans les organites, joue un rôle important dans le mécanisme de traduction.

Le rôle de protéines PPR dans la traduction a également été suggéré par l'association de certaines protéines PPR (PPR336) avec les polysomes des mitochondries de chou-fleur (Uyttewaal et *al.*, 2008).

2.4. COORDINATION DE L'EXPRESSION DES GENOMES DES ORGANITES ET DU NOYAU

Le nombre de gènes nucléaires codant des protéines adressées aux organites a été estimé par des approches de protéomique et de bioinformatique (Richly et Leister, 2004 ; Stern et *al.*, 2004 ; Millar et *al.*, 2006 ; Ferro et *al.*, 2010). Le noyau d'*Arabidopsis* contiendrait environ 3000 et 2000 gènes codant des protéines adressées aux chloroplastes et aux mitochondries respectivement. La biogénèse des organites est donc dépendante de l'expression de leur génome mais aussi de celui du noyau et de leur coordination réciproque. Pour exemple, l'assemblage du complexe NADH déshydrogénase (ou complexe I) de la chaine respiratoire mitochondriale est constitué d'au moins 49 sous-unités dont seules 9 ont une origine mitochondriale (Klodmann et *al.*, 2010). Jusqu'à ce jour, les facteurs jouant un rôle dans la coordination de l'expression des gènes des organites et du noyau, de leur *trans*cription à leur traduction sont encore méconnus.

Un modèle de régulation au niveau de la traduction a été proposé dans le chloroplaste. Le contrôle par épistasie de synthèse (CES) observé chez *chlamydomonas* est un mécanisme permettant de coordonner la traduction de protéines dans le chloroplaste en fonction de la présence de leur partenaires codés par le noyau (Choquet et *al.*, 1998 ; Choquet et *al.*, 2001 ; Choquet et Wollman, 2002). Pour exemple, l'absence d'une des sous-unités du complexe du cytochrome b_6f inhibe la traduction du cytochrome *f* dans le chloroplaste. Cette inhibition est médiée par l'interaction de la région 5'-UTR du *trans*crit *petA* (codant le cytochrome *f*) avec la partie C-terminale de protéine du cytochrome *f* non assemblée. Ainsi, la protéine autorégulerait sa propre traduction. L'interaction entre le cytochrome *f* et le *trans*crit pourrait être indirecte et nécessiter un tiers facteur protéique (Choquet et *al.*, 1998).

Chez les plantes supérieures, ce phénomène a également été démontré dans le tabac où la traduction de la sous-unité LSU de la rubisco, codée par le gène chloroplastique *rbcL*, est dépendante de la présence de la petite sous-unité SSU d'origine nucléaire dans le chloroplaste (Wostrikoff et Stern, 2007).

Afin de réaliser cette coordination, les mitochondries et les chloroplastes doivent envoyer des signaux aux noyaux afin de réguler l'expression de ses gènes: ce sont les signaux rétrogrades d'origine non protéique. Un modèle précis des signaux rétrogrades entre les organites et le noyau reste à établir.

Toutefois, des études génétiques et biochimiques ont permis d'identifier des voies potentielles de signaux rétrogrades entre les chloroplastes et le noyau. L'une des voies les plus étudiées implique les précurseurs des voies de biosynthèse des tétrapyrroles (Susek et *al.*, 1993; Mochizuki et *al*., 2001; Larkin et *al.*, 2003; Strand et *al.*, 2003; Kobayashi et *al*., 2009). Cette voie ferait intervenir des facteurs protéiques codées par le noyau pouvant interférer avec la synthèse des molécules (Susek et *al.*, 1993).

Une seconde voie appelée PGE pour "plastid gene expression pathway" a été découverte en utilisant des inhibiteurs de la traduction chloroplastique qui entraîne la répression de gènes nucléaires photosynthétiques (Gray et *al.*, 2003).

La dernière voie est associée à l'état d'oxydoréduction de la chaîne photosynthétique du chloroplaste. Les formes réactives d'oxygène (ROS) semblent jouer un rôle dans l'expression des gènes nucléaires dans les chloroplastes (Karpinski et *al.*, 1999).

Dans les mitochondries, la signalisation rétrograde serait également associée à la production des formes réactives d'oxygène (ROS) (Woodson et Chory, 2008).

Des facteurs protéiques agissant en *trans* ont certainement un rôle à jouer dans cette coordination de l'expression des génomes du noyau et des organites. Les protéines à double adressage noyau/organites seraient de bons candidats dans ce rôle (Krause et Krupinska, 2009). Les candidats les plus pertinents seraient, les facteurs de *trans*cription, de traduction ou les protéines ayant une fonction dans l'expression des gènes, et qui présenteraient une double localisation. Chez les plantes, la protéine de liaison à l'ADN, Whirly1 est doublement adressée au noyau et chloroplastes des cellules et suscite un grand intérêt (Grabowski et *al.*, 2008; Krause et *al*., 2009). Dans les noyaux d'*Arabidopsis*, la protéine lierait l'ADN simple brin et régulerait la longueur des télomères (Yoo et *al.*, 2007). Dans les chloroplastes d'*Arabidopsis*, la protéine participerait à la stabilité du génome en prévenant des recombinaisons illégitimes (Maréchal et *al*., 2009).

Dans les cellules humaines, le gène codant la protéine LRP130 (ou LRPPRC pour leucine-rich pentatricopeptide repeat cassette), de la famille PPR a été mis en cause dans une forme du syndrome de Leigh, une maladie neurodégénérative causée par une déficience en cytochrome *c*. Cette protéine est doublement adressée au noyau et aux mitochondries des cellules et est proposée pour jouer un rôle dans la communication entre les deux organites

Figure 17: Alignement des séquences consensus des motifs PPR P, L et S d'*Arabidopsis*.

La séquence consensus du motif P présente un décalage de deux acides aminés par rapport au motif PPR PF01535 de la base de données PFAM (Bateman et *al*., 2004). Les lettres capitales représentent les résidus les plus conservés de chaque motif et ceux en gras représentent les résidus conservés entre motifs apparentés. Les boites jaune et rouge délimitent les deux hélices α prédites (modifié d'après Lurin et *al*., 2004).

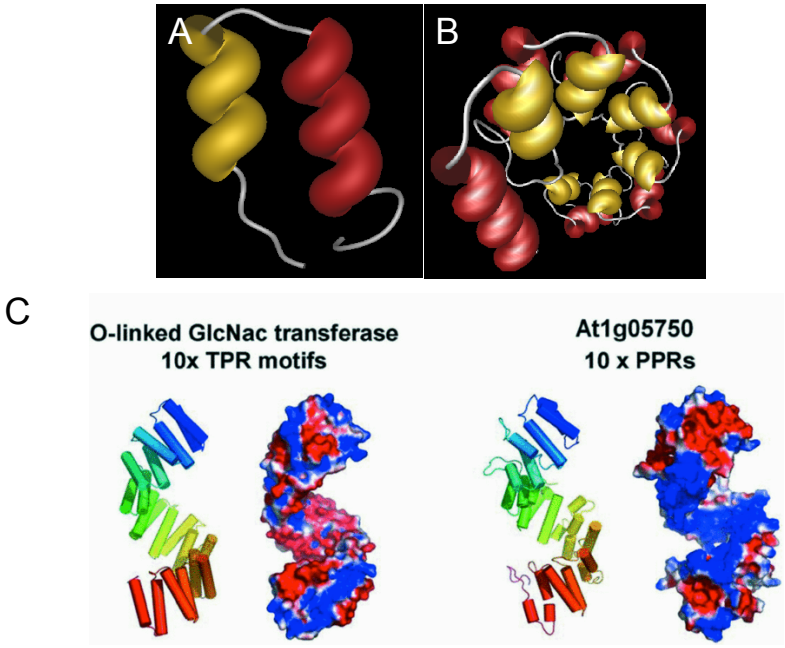


Figure 18: Structure prédite d'un motif PPR (A) et d'une succession de motifs PPR (B). Le code couleur est respecté par rapport à la Figure 17.

Le motif PPR est prédit pour former deux hélices α antiparallèles et la succession des motifs PPR formerait une superhélice (Small et Peeters, 2000).

(C) Comparaison de la structure tertiaire d'une protéine à motifs TPR GlcNAc transférase (gauche) et d'une protéine à motifs PPR At1g05750 (droite).

Contrairement à la protéine TPR ou les résidus exposés dans le sillon central sont majoritairement acides, ceux de la protéine PPR At1g0570 sont majoritairement basiques. Cette surface basique fournit une interface d'interaction adéquate pour les ARN (Delannoy et *al*., 2007).

en influençant l'expression de leur génome (Gohil et *al.*, 2010 ; Sasarman et *al.*, 2010 ; Xu et *al.*, 2004 ; Mili et Piñol-Roma, 2003).

Chez les plantes, aucune étude n'a montré une protéine PPR doublement adressée dans un organite et au noyau pouvant être impliquée dans un tel rôle. Toutefois, un gène *GRP23* est essentiel chez *Arabidopsis* et code pour une protéine PPR présentant un domaine "Leucine Zipper" et une séquence de localisation nucléaire. Cette protéine localisée dans le noyau des cellules interagit avec une sous unité de l'ARN polymérase et influencerait la transcription de gènes nucléaires (Ding et *al.*, 2006).

La première partie de l'introduction a permis de présenter les principales étapes de l'expression des génomes dans les organites chez les plantes en mettant l'accent sur le rôle important des facteurs nucléaires dans le contrôle de celles-ci. Plus particulièrement, la grande famille des protéines PPR semble jouer un rôle prépondérant dans l'expression des génomes des organites de plantes. Ce sont les protéines de cette famille qui ont fait l'objet de mes travaux de thèse.

2.5. LES PROTEINES PPR

2.5.1. DEFINITION

La famille des gènes PPR ou "pentatricopeptide repeat" code pour des protéines présentant des motifs PPR (ou motifs P). Ce motif qui donne son nom aux protéines a été récemment caractérisé par Small et Peeters, 2000, lors de recherche de gènes nucléaires codant des protéines pouvant être adressées aux organites au sein du génome d'*Arabidopsis*. Le motif PPR correspond à une séquence dégénérée de 35 acides aminés qui peut être répétée au sein de la séquence de la protéine (la moyenne du nombre de motifs par protéine est de 12) (Figure 17). Bien que la structure tridimentionnelle d'un de ces motifs n'ai pas été déterminée expérimentalement, les modélisations de structures suggèrent que le motif PPR adopte une structure en paire d'hélices α antiparallèles (Small et Peeters, 2000) (Figures 17 et 18). Cette structure s'apparente fortement au motif protéique TPR ou "tetratricopeptide repeat" composé d'une séquence dégénérée de 34 acides aminés. Ce motif est connu pour être un domaine d'interaction protéine-protéine (Blatch et Lassle, 1999). Les protéines TPR adoptent une structure hélicoïdale créée par la succession des différentes paires d'hélices α antiparallèles (Jinek et *al.,* 2004). Cette super-hélice renferme un sillon central servant de plateforme d'interaction protéine-protéine. Par analogie, les protéines PPR sont prédites d'adopter une structure similaire. La différence majeure étant que les résidus exposés dans le sillon central seraient majoritairement hydrophiles et chargés positivement (Figure 18). Cette surface hydrophile permettrait l'interaction avec des ligands hydrophiles chargés négativement tels que les ARNs (Small et Peeters, 2000). Ces observations ont conduit à proposer les motifs PPR comme étant des motifs putatifs de liaison à l'ARN. Cette propriété a par la suite été démontrée par des approches biochimiques de gel retard (Okuda et *al.*, 2006 ; Beick et *al.*, 2008 ; Pfalz et *al.*, 2009 ; Tasaki et al., 2010), de pontage des protéines PPR à leur substrat ARN par irradiation aux UV (Meierhoff et *al.*, 2003 ;

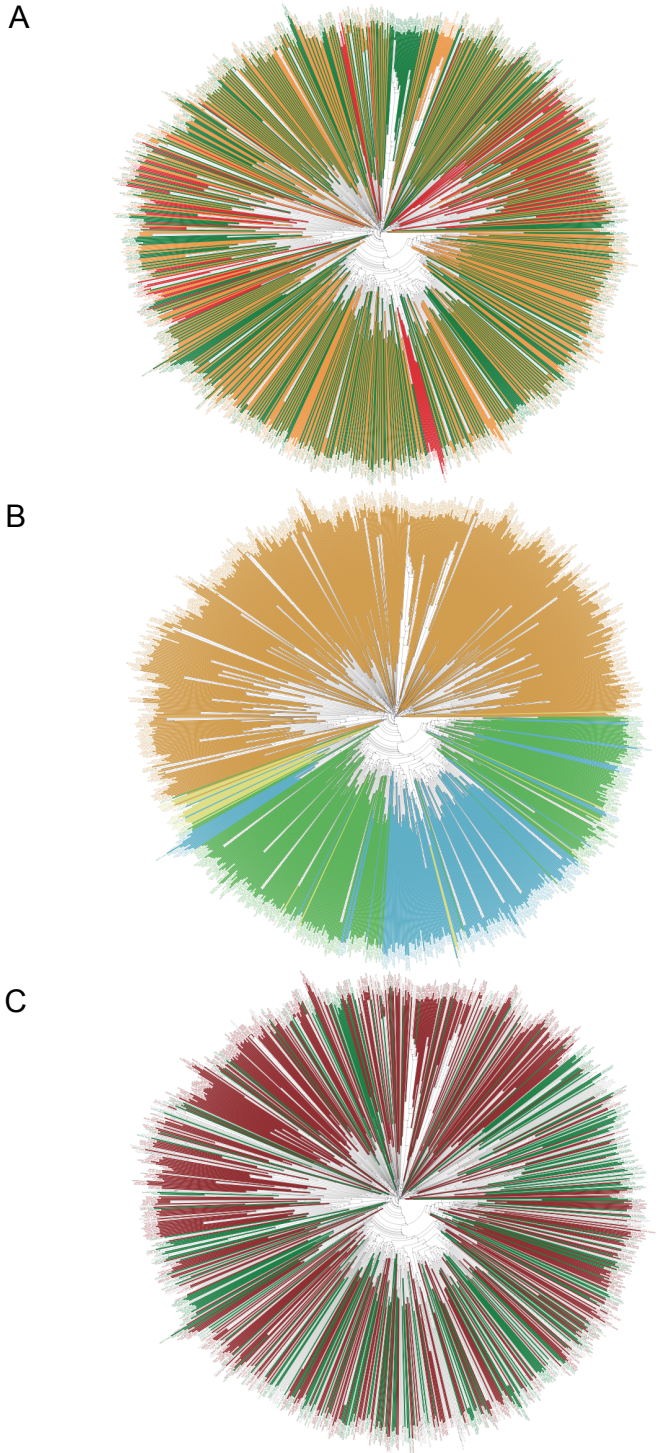


Figure 19: Arbre phylogénétique radial des protéines PPR chez *Arabidopsis*, le riz et la mousse.

Les branches terminales de l'arbre sont colorées en différentes couleurs pour indiquer: **(A)** les espèces; vert: *Arabidopsis*, orange: le riz, rouge: *Physcomitrella*. **(B)** les sous-groupes chez *Arabidopsis*; orange: P, jaune: PLS, vert: E/E+, bleu: DYW. **(C)** la localisation prédite par le programme Predotar (Small et *al*., 2004); rouge: mitochondrie, vert: chloroplaste, gris: non définie (d'après O'Toole et *al*., 2008).

Mili et Piñol-Roma, 2003) ou encore par des expériences d'affinité au substrat (Nakamura et *al.*, 2003 ; Lurin et *al*., 2004).

Chez les plantes, la famille des protéines PPR est remarquablement vaste avec respectivement 450 et 477 membres chez les espèces *Arabidopsis thaliana* et *Oryza sativa* (riz) (Figure 19). Le génome de la mousse *Physcomitrella patens* code quant à lui 103 membres (O'Toole et *al.*, 2008 ; Lurin et *al*., 2004). Ces protéines sont en nombre plus limité chez les organismes pluricellulaires et non plantes avec seulement 6 et 2 gènes PPR chez l'homme et la drosophile respectivement.

Chez les eucaryotes unicellulaires ; trypanosome, levure *Saccharomyces cerevisiae*, et algue verte *Chlamydomonas reinharditii*, on dénombre 23, 5 et 9 membres respectivement (Lurin et *al.*, 2004 ; O'Toole et *al.*, 2008).

Une analyse bioinformatique des séquences des protéines PPR a révélé que la majorité de celles-ci présente au niveau de l'extrémité N-terminale une séquence d'adressage aux organites (Figure 19) (Lurin et *al.*, 2004 ; O'Toole et *al*., 2008). Ces séquences d'adressage prédisent un rôle de ces protéines au niveau de l'ARN présent dans les mitochondries et/ou chloroplastes des cellules. Les différentes études ont effectivement confirmé cette hypothèse (voir chapitre 2.2). Certaines protéines PPR sont cependant prédites pour avoir une localisation en dehors des organites comme dans le noyau ou le cytosol (Andrès et *al*., 2007). Chez *Arabidopsis*, une protéine PPR, GRP23 présente une séquence de localisation nucléaire et est effectivement adressée au noyau des cellules où elle jouerait un rôle dans la transcription (Ding et al., 2006).

La famille des gènes PPR est donc spécifique des organismes eucaryotes et leur nombre a explosé chez les plantes terrestres ce qui suggère une importance particulière de ces protéines PPR dans les organites de plantes.

Chez *Arabidopsis*, la moitié des protéines PPR identifiées présentent des motifs PPR (ou motifs P) répétés en tandem sans espace entre ceux-ci. La seconde moitié des protéines quant à elle, présente entre les motifs PPR des insertions d'environ 65 à 70 acides aminés. Ces insertions représentent deux motifs apparentés aux protéines PPR : un motif court S (pour short) de 31 acides aminés et un motif plus long L (pour long) de 35 ou 36 acides aminés (Figure 17). Dans ce dernier groupe, les triplets de motifs PLS sont successivement répétés au sein des protéines PPR (Lurin et *al*., 2004 ; Rivals et *al.*, 2006).

La présence ou non des motifs L et S entre les motifs P permet la division de la famille PPR en deux sous-familles : la famille P (présentant uniquement des motifs PPR en tandem) et la famille PLS (Figure 20) (Lurin et *al*., 2004).

La sous-famille PLS est spécifique des plantes terrestres et a été à l'origine décrite par Aubourg et *al.*, 2000 comme étant la famille des protéines PCMP (pour Plant Combinatorial and Modular Protein) au sein desquelles le triplet PLS est le bloc de motifs le plus représenté au sein des séquences protéiques. Cependant, le motif S peut aussi être

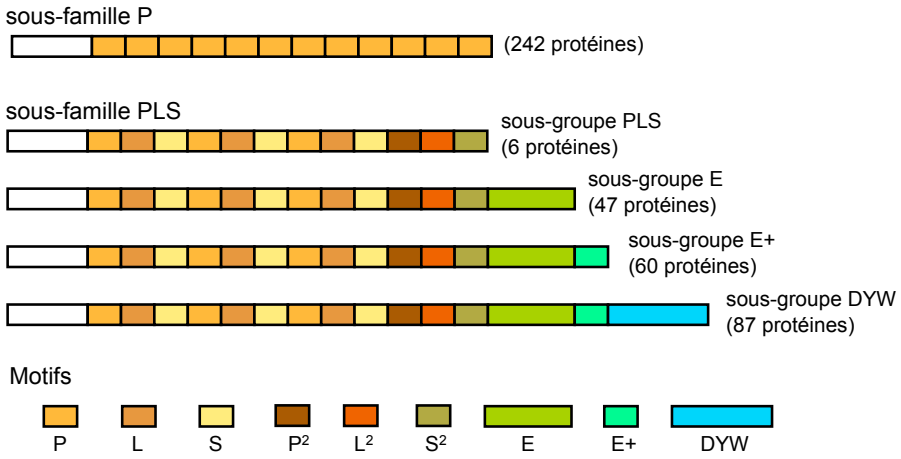


Figure 20: Organisation et architectures typiques des protéines PPR d'*Arabidopsis*.

L'architecture caractéristique des protéines PPR de chaque sous-groupe ainsi que leur nombre sont indiqués. Le nombre de motifs et répétitions varie d'une protéine à l'autre et les répétitions en tandem peuvent être interrompus (modifié d'après Lurin et *al*., 2004)

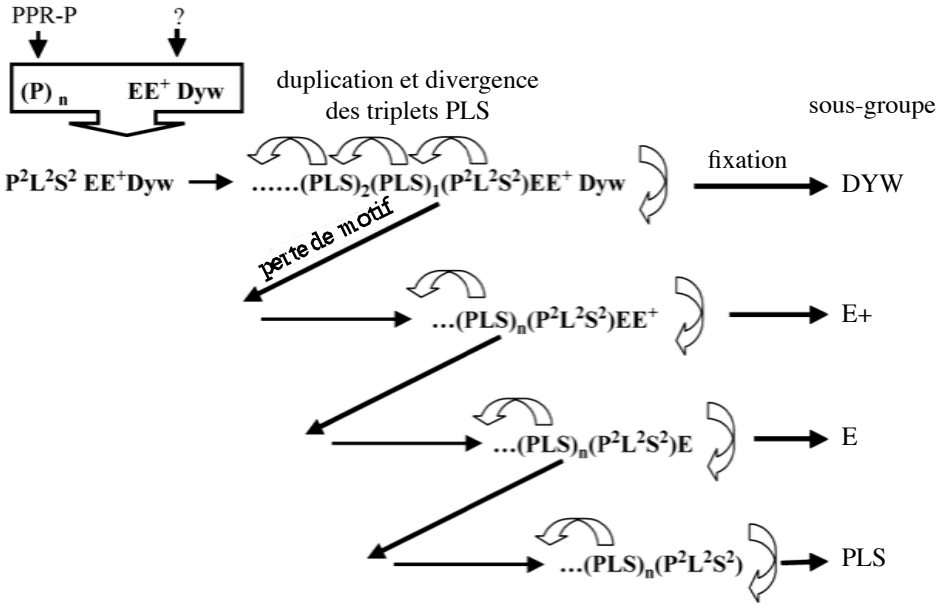


Figure 21: Modèle expliquant la formation des protéines PPR de la sous-famille PLS.

Deux gènes ancestraux codant respectivement un triplet de motifs P adjacents et les domaines E/E+/DYW auraient fusionné. La mutation du triplet P aurait généré le bloc $P^2L^2S^2$ qui après duplication en tandem et divergence aurait donné naissance aux répétitions PLS et au sous-groupe DYW. Les sous-groupes E+, E et PLS seraient apparus par pertes ultérieures de motifs C-terminaux (modifié d'après Rivals et *al*., 2006).

répété en tandem au sein d'un bloc. Un motif $P^2L^2S^2$ dont la séquence diverge du motif PLS classique est toujours retrouvé en position C-terminale après les répétitions de triplets PLS (Figure 20) (Lurin et *al.*, 2004 ; Rivals et *al.*, 2006).

La sous-famille PLS est elle même subdivisée en 4 sous-classes en fonction de la présence ou non d'une extension C-terminale de type E, E+ ou DYW. Contrairement aux motifs PLS, ces domaines ne sont jamais répétés et terminent donc la séquence de la protéine (Figure 20).

Le domaine DYW fortement conservé se termine par un tripeptide quasiment invariant aspartate, tyrosine et tryptophane (Aubourg et *al*., 2000 ; Lurin et *al*., 2004).

Rivals et *al.*, 2006 propose un modèle expliquant la formation de la sous-famille PLS chez les plantes terrestre dans lequel un gène ancestral codant une protéine PPR de type P fusionne avec un second gène ancestral codant les domaines E/E+/DYW. Après des événements de duplications et divergences des motifs P en blocs PLS ou $P^2L^2S^2$, la sous-classe des protéines PPR de la sous-famille PLS du sous-groupe E/E+/DYW serait née. Les sous-groupes E+, E ou PLS seraient apparues par pertes ultérieures et successives des motifs DYW, E+ puis E (Figure 21).

2.5.2. FONCTIONS BIOLOGIQUES DES PROTEINES PPR

Les motifs PPR liant l'ARN confèrent aux protéines PPR un rôle premier dans la régulation des différentes étapes impliquant les ARNs dans les organites et décrites ci dessus.

Bien que le nombre encore relativement limité d'études publiées à ce jour empêche de tirer des conclusions définitives, il apparaît que les protéines PPR de la sous-famille P (présentant seulement des motifs P) aient des fonctions variées dans un large éventail de mécanismes post-transcriptionnels dans les organites (Tableau 2) alors que l'édition de l'ARN semble être spécifique aux protéines PPR de la sous-famille PLS.

Des études approfondies de la distribution phylogénétique des protéines PPR de la sous-famille PLS ont par ailleurs montrées que les protéines PPR de la sous-classe DYW sont présentes chez les différentes espèces de plantes terrestres à l'exception des *Marchantiidae* (ou Marchantia) et absentes chez les algues (Salone et *al.*, 2007 ; Rudinger et *al*., 2008). Ces résultats corrèlent avec la présence ou l'absence du mécanisme d'édition dans les organites des plantes. La sous-famille des protéines PLS semble donc être spécifique des plantes réalisant l'édition. Ceci a conduit à suggérer que le domaine DYW aurait un rôle particulier dans l'édition (Salone et *al.*, 2007). Une analyse de la séquence protéique du domaine DYW a révélé un site de liaison au zinc $HxEx_nCxxC$ typique des enzymes de type cytidine désaminase (CDA). De plus, les prédictions structurales du domaine DYW s'apparentent à la structure résolues des CDAs (Salone et *al.*, 2007). Ces observations ont permis d'émettre l'hypothèse que le domaine DYW de certaines protéines PPR pourrait posséder l'activité catalytique de cytidine désaminase dans les organites de plantes (Salone

Gene	Organisme	Classe	Mécanisme moléculaire	Localisation	Référence
Pet309	*Saccharomyces*	P	RNA stabilization and translation cox1	M	Manthey et *al*., 1995
Cya5	*Neurospora*	P	Translation cox1	M	Coffin et *al*., 1997
p63	*Triticum*	P	Transcription cox2	M	Ikeda et *al*., 1999
CRP1	*Zea*	P	Translation petA, psaC	C	Fisk et *al*., 1999
BSF	*Drosophila*	P	RNA stabilization 3-UTR of bicoid	Cy	Mancebo et *al*., 2001
LRPPRC, hLRP130	*Human*	P	RNA stabilization cox1, cox3	M, N	Mili et *al*., 2003; Liu et *al*., 2002
HCF152	*Arabidopsis*	P	Splicing and maturation petB	C	Meierhoff et *al*., 2003
PPR2	*Zea*	P	Translation	C	Williams and Barkan, 2003
Rfk1	*Raphanus*	P	Translation orf125	M	Koizuka et *al*., 2003
CRR2	*Arabidopsis*	DYW	RNA cleavage ndhB	C	Hashimoto et *al*., 2003
Rfo	*Raphanus*	P	Translation orf138	M	Brown et *al*., 2003
Aep3	*Saccharomyces*	P	RNA stabilization and maturation atp6 atp8	M	Ellis et *al*., 2004
PGR3	*Arabidopsis*	P	RNA stabilization and translation petL operon	C	Yamazaki et *al*., 2004
Rf1, Rf1a	*Oryza*	P	RNA cleavage atp6-orf79	M	Akagi et *al*., 2004 Wang et *al*., 2006
CRR4	*Arabidopsis*	E	Editing ndhD (117166)	C	Kotera et *al*., 2005
GRP23	*Arabidopsis*	P	Transcription	N	Ding et *al*., 2006
Rf1b	*Oryza*	P	RNA degradation atp6-orf79	M	Wang et *al*., 2006
pTAC2	*Arabidopsis*	P	Transcription	C	Pfalz et *al*., 2006
PPR4	*Zea*	P	Splicing rps12	C	Schmitz-Linneweber et *al*., 2006
EMP4	*Zea*	P	RNA stabilization rps2A, rps2B, rps3 rpl16, orfX	M	Gutiérrez-Marcos et *al*., 2007
CRR21	*Arabidopsis*	E	Editing ndhD (116785)	C	Okuda et *al*., 2007
Pp_PPR531	*Physcomitrella*	P	RNA splicing and maturation clpP	C	Hattori et *al*., 2007
OTP43	*Arabidopsis*	P	Splicing nad1	M	de Longevialle et *al*., 2007
OTP51	*Arabidopsis*	P	Splicing ycf3	C	de Longevialle et *al*., 2008
PPR5	*Zea*	P	RNA stabilisation trnG-UCC	C	Beick et *al*., 2008
PPR336	*Oryza*	P	Translation ?	M	Uyttewaal et *al*., 2008
PTCD2	*Human*	P	Maturation nd5 cob	M	Xu et *al*., 2008
MCA1	*Chlamydomonas*	P	RNA stabilization petA	C	Loiselay et *al*., 2008
CLB19	*Arabidopsis*	E	Editing rpoA (78691), clpP (69942)	C	Chateigner-Boutin et *al*., 2008
PTCD3	*Human*	P	Translation	M	Davies et *al*., 2009
PTCD1	*Human*	P	RNA stabilization precursor Leu-tRNA	M	Rackham et *al*., 2009
OTP80	*Arabidopsis*	E	Editing rpl23 (86056)	C	Hammani et *al*., 2009
OTP81	*Arabidopsis*	DYW	Editing rps12 intron (69553)	C	Hammani et *al*., 2009
OTP84	*Arabidopsis*	DYW	Editing ndhF (112349), psbZ (35800), ndhB (94999)	C	Hammani et *al*., 2009
OTP85	*Arabidopsis*	DYW	Editing ndhD (116494)	C	Hammani et *al*., 2009
OTP86	*Arabidopsis*	DYW	Editing rps14 (37161)	C	Hammani et *al*., 2009

Tableau 2: Principales protéines PPR aux fonctions résolues.

M: mitochondrie, C: chloroplaste, Cy: cytoplasme, N: noyau.

Suite du tableau en page suivante.

CRR22	*Arabidopsis*	DYW	Editing ndhB (96419), ndhD (116281), rpoB (25779)	C	Okuda et *al*., 2009
CRR28	*Arabidopsis*	DYW	Editing ndhB (96698), ndhD (116290)	C	Okuda et *al*., 2009
YS1	*Arabidopsis*	DYW	Editing rpoB (25992)	C	Zhou et *al*., 2009
RARE1	*Arabidopsis*	DYW	Editing accD (57868)	C	Robbins et *al*., 2009
LPA66	*Arabidopsis*	DYW	Editing psbF (63985)	C	Cai et *al*., 2009
MEF1	*Arabidopsis*	DYW	Editing rps4-956, nad7-963, nad2-1160	M	Zehrmann et *al*., 2009
MEF11	*Oryza sativa*	DYW	Editing cox3-422, nad4-124, ccb203-344	M	Verbitskiy et *al*., 2009
OGR1	*Oryza sativa*	DYW	Editing cox2-167, cox3-572, ccmC-458, nad2-1457, nad4-401, nad4-416, nad4-433	M	Kim et *al*., 2009
PPR10	*Zea*	P	RNA stabilization atpH, psaJ	C	Pfalz et *al*., 2009
MEF9	*Arabidopsis*	E	Editing nad7-200	M	Takenaka, 2010
Pp_PPR71	*Physcomitrella*	DYW	Editing $ccmF_c$-122	M	Tasaki et *al*., 2010
MRL1	*Chlamydomonas*, *Arabidopsis*	P	RNA stabilization rbcL	M	Johnson et *al*., 2010
RPF2	*Arabidopsis*	P	RNA stabilization nad9, cox3	M	Jonietz et *al*., 2010
OTP82	*Arabidopsis*	DYW	Editing ndhG (118858), ndhB (95644)	C	Okuda et *al*., 2010

Tableau 2 (suite): Principales protéines PPR aux fonctions résolues.

M: mitochondrie, C: chloroplaste, Cy: cytoplasme, N: noyau.

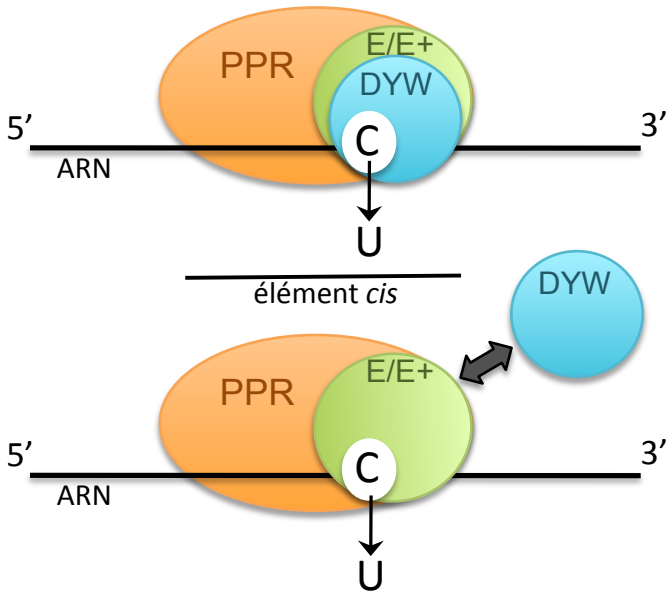


Figure 22: Modèle hypothétique d’action des *trans*-facteurs PPR-DYW et PPR-E/E+.

La protéine PPR reconnaît et lie spécifiquement la séquence entourant le site d’édition du transcrit grâce aux motifs PPR. Le domaine DYW porté par certaines PPR posséderait l’activité catalytique cytidine désaminase et agirait directement sur le site à éditer. Quand le domaine DYW est absent de la protéine (cas des protéines PPR-E/E+), un domaine DYW d’une autre protéine PPR ou une autre désaminase serait recruté et agirait en *trans* sur le site d’édition.

et *al*., 2007). Cependant, les tentatives pour démontrer l'activité cytidine désaminase de ce motif *in vitro* ont toutes échouées (Nakamura et *al.*, 2008 ; Okuda et *al.*, 2009).

Jusqu'à ce jour, les analyses génétiques ont uniquement montré l'implication directe de protéines PPR de la sous-famille PLS E/E+ et DYW dans l'édition des transcrits chloroplastiques et mitochondriaux (Tableau 2).

L'ensemble des observations à permis de proposer un modèle du rôle joué par les *trans*-facteurs PPR dans l'édition : les motifs PPR de la protéine permettraient la reconnaissance et la liaison d'éléments spécifiques du transcrit en *cis* du site d'édition. Le domaine DYW catalyserait la réaction d'édition en *cis* dans le cas des protéines PPR-DYW ou cette activité serait recruté en *trans* par les protéines PPR-E/E+ (Figure 22).

3. OBJECTIFS DES TRAVAUX DE THESE

Au début de ma thèse en Octobre 2006, les fonctions d'un nombre limité de protéines PPR étaient élucidées chez différents organismes eucaryotes (Tableau 2).

Nous avons donc décidé de réaliser des études fonctionnelles de gènes codant des protéines PPR chez la plante modèle *Arabidopsis*. Pour ce faire, nous avons tiré avantage des banques publiques de mutants d'insertion de type T-DNA (ADN de transfert) pour des gènes PPR d'*Arabidopsis* et avons mené une approche de génétique inverse afin d'élucider leur fonction.

Plus particulièrement, nous nous sommes intéressés à l'implication de protéines PPR dans deux mécanismes post-transcriptionnels dans les organites: l'édition de l'ARN dans le chloroplaste et la coordination de l'expression des génomes de la mitochondrie et du noyau.

4. RESULTATS

4.1. IMPLICATION DE PROTEINES PPR DANS L'EDITION D'ARN CHLOROPLASTIQUES

L'édition des ARNs nécessite la reconnaissance spécifique de cytidines cibles parmi la séquence nucléotidique des transcrits. Cette reconnaissance est achevée par la liaison de facteurs protéiques à la séquence nucléotidique ou élément *cis* entourant le site à éditer.

Des expériences d'édition *in vitro* de transcrits à l'aide d'extraits chloroplastiques de tabac suivies par une irradiation aux UV ont permis d'identifier des protéines de poids moléculaires distincts de 25, 56, 70, 91 et 93 kDa liant spécifiquement les éléments *cis* des transcrits chloroplastiques *psbL*, *psbE*, *petB*, *rpoB* et *rpoA*, respectivement (Hirose et Sugiura, 2001 ; Miyamoto et *al.*, 2002, 2004 ; Kobayashi et *al.*, 2008). Une même protéine de 95 kDa lie spécifiquement les éléments *cis* des deux transcrits distincts *ndhB-9* et *ndhF-2*. Les éléments *cis* liés par les *trans*-facteurs ont été délimités à 15 nucléotides en amont du site d'édition. Toutefois, ces expériences ne révèlent pas l'identité de ces protéines.

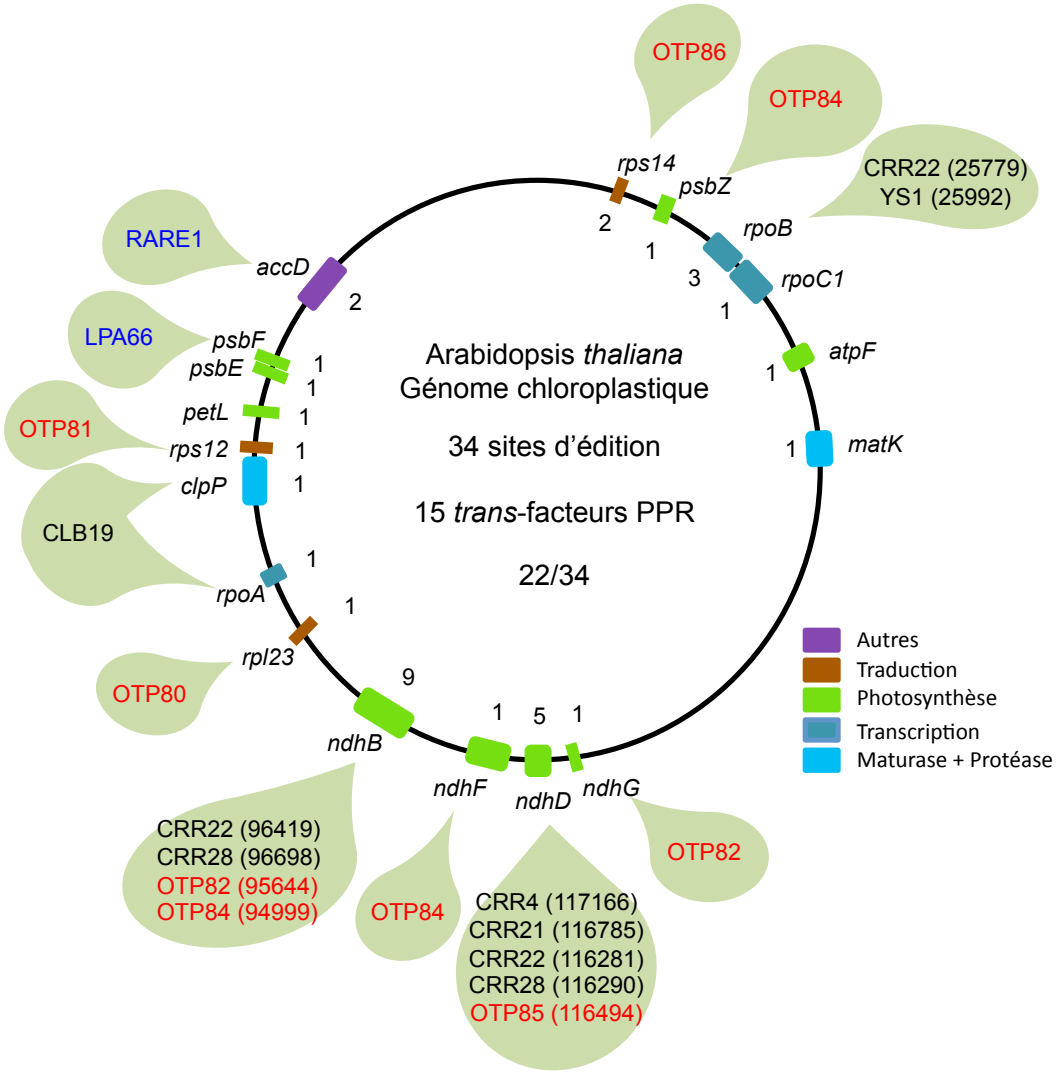


Figure 23: Représentation des gènes possédant des sites d'édition du génome chloroplastique d'*Arabidopsis thaliana* et des facteurs d'édition PPR identifiés.

Les chiffres adjacents aux gènes correspondent au nombre de sites d'édition par gène. Les *trans*-facteurs PPR sont représentés dans des bulles pointant leur(s) site(s) d'édition respectifs. Les chiffres entre parenthèses représentent la position génomique du site d'édition dans le cas de gène présentant plusieurs sites. En noir sont représentés les *trans*-facteurs connus au début de mon projet, en rouge les trans-facteurs identifiés par moi même (Hammani et *al*., 2009; Okuda et *al*., 2010) et en bleu ceux identifiés indépendamment par d'autres groupes au cours de mon projet.

Plus tard, des études génétiques ont identifié des protéines PPR comme étant des *trans*-facteurs de certains sites d'édition chloroplastiques. Au début de ce projet, 6 protéines PPR; CRR4, CRR21, CLB19, CRR22, CRR28, YS1 ont été impliquées dans l'édition de sites du génome chloroplastique (Figure 23). CLB19, CRR21 et CRR22 reconnaissent plusieurs sites d'édition sur des transcrits distincts suggérant que ces différents ARNs possèdent une information commune recrutant une même protéine PPR. Ces protéines PPR appartiennent toutes aux sous-groupes E/E+ et DYW.

La corrélation entre la présence du mécanisme d'édition et la présence des protéines PPR-DYW au sein des plantes terrestres suggère que les protéines PPR de la sous-famille PLS joueraient un rôle spécifique dans l'édition des transcrits des organites (Figure 24) (Salone et *al.*, 2007 ; Rudinger et *al*., 2008).

Afin d'identifier de nouveaux facteurs d'édition PPR chez *Arabidopsis*, nous avons décidé d'entamer une approche de génétique inverse en sélectionnant des gènes codant des protéines PPR des sous-groupes E/E+ et DYW prédites bio-informatiquement pour être adressées aux chloroplastes. Nous avons ensuite analysé des mutants *Arabidopsis* d'insertion T-DNA pour ces gènes afin d'identifier d'éventuel défaut d'édition au sein du transcriptome chloroplastique.

Deux méthodes facilitant la compréhension des résultats des publications 1 et 2 sont décrites en annexe du manuscrit.

4.1.1. PUBLICATION 1: A STUDY OF NEW ARABIDOPSIS CHLOROPLAST RNA EDITING MUTANTS REVEALS GENERAL FEATURES OF EDITING FACTORS AND THEIR TARGET SITES, THE PLANT CELL (2009)

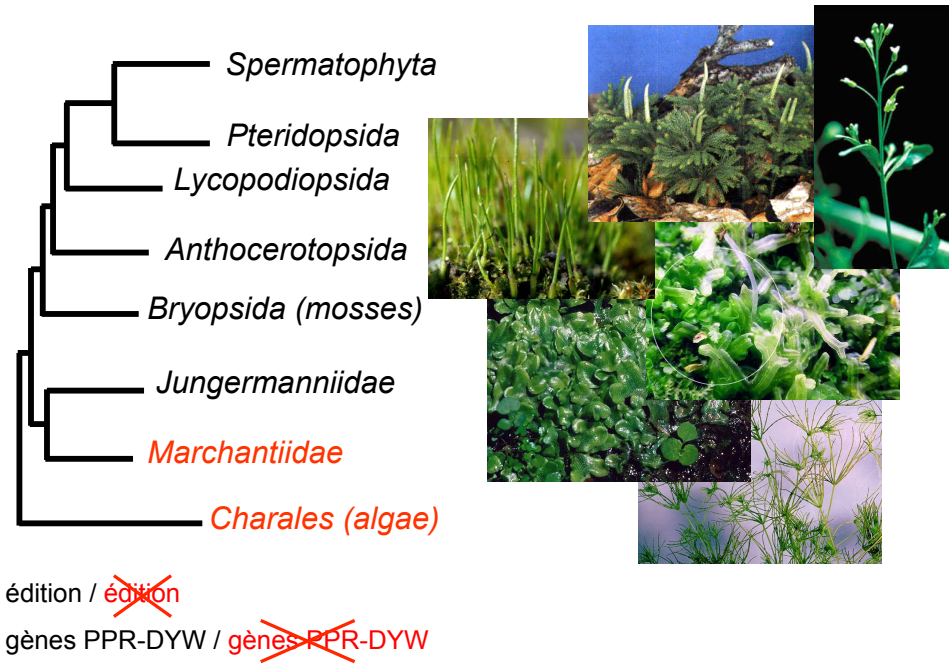


Figure 24: Distribution phylogénétique de l'édition de l'ARN au sein des plantes vertes et des protéines PPR à domaine DYW.

L'édition de l'ARN est présent chez toutes les plantes vertes terrestres à l'exception des *Marchantiidae*. Ce mécanisme est absent des algues vertes. La présence du mécanisme d'édition chez les plante vertes corrèle avec la présence de gènes codant des protéines PPR à domaine DYW au sein de leur génome (Salone et *al*., 2007; Rudinger et *al*., 2008).

The Plant Cell, Vol. 21: 3686–3699, November 2009, www.plantcell.org

A Study of New *Arabidopsis* Chloroplast RNA Editing Mutants Reveals General Features of Editing Factors and Their Target Sites [W] [OA]

Kamel Hammani,[a,b,1] Kenji Okuda,[c,1] Sandra K. Tanz,[a] Anne-Laure Chateigner-Boutin,[a] Toshiharu Shikanai,[c] and Ian Small[a,2]

[a] Australian Research Council Centre of Excellence in Plant Energy Biology, University of Western Australia, Crawley 6009 WA, Australia

[b] Institut de Biologie Moléculaire des Plantes du Centre National de la Recherche Scientifique, Université de Strasbourg, 67084 Strasbourg Cedex, France

[c] Department of Botany, Graduate School of Science, Kyoto University, Kyoto 606-8502 Japan

RNA editing in higher plant organelles results in the conversion of specific cytidine residues to uridine residues in RNA. The recognition of a specific target C site by the editing machinery involves *trans*-acting factors that bind to the RNA upstream of the C to be edited. In the last few years, analysis of mutants affected in chloroplast biogenesis has identified several pentatricopeptide repeat (PPR) proteins from the PLS subfamily that are essential for the editing of particular RNA transcripts. We selected other genes from the same subfamily and used a reverse genetics approach to identify six new chloroplast editing factors in *Arabidopsis thaliana* (OTP80, OTP81, OTP82, OTP84, OTP85, and OTP86). These six factors account for nine editing sites not previously assigned to an editing factor and, together with the nine PPR editing proteins previously described, explain more than half of the 34 editing events in *Arabidopsis* chloroplasts. OTP80, OTP81, OTP85, and OTP86 target only one editing site each, OTP82 two sites, and OTP84 three sites in different transcripts. An analysis of the target sites requiring the five editing factors involved in editing of multiple sites (CRR22, CRR28, CLB19, OTP82, and OTP84) suggests that editing factors can generally distinguish pyrimidines from purines and, at some positions, must be able to recognize specific bases.

INTRODUCTION

The term RNA editing covers a multitude of processes that lead to posttranscriptional sequence alterations in RNAs. In plants, RNA editing generally indicates highly specific cytidine-to-uridine conversions that are observed in both mitochondrial and plastid transcripts. Although similar editing events occur in other organisms (see review by Simpson and Emeson, 1996), it is clear that this form of RNA editing arose in an immediate ancestor of land plants and is not homologous to editing processes found in other phyla and may not be mechanistically related. It is likely that >500 sites are specifically altered in *Arabidopsis thaliana* mitochondrial transcripts (Giegé and Brennicke, 1999; Bentolila et al., 2008; Zehrmann et al., 2008), and at least 34 sites are known to be changed in *Arabidopsis* plastid transcripts (Chateigner-Boutin and Small, 2007). RNA editing typically affects the transcripts of protein-coding genes but has also been found to modify noncoding transcribed regions, structural RNAs, and intron sequences. RNA editing is essential for correct gene expression: proteins translated from edited transcripts are usually different from the ones deduced from the gene sequence and usually present higher similarity to the corresponding nonplant homologs (reviewed in Shikanai, 2006).

A major question concerning the editing process is the manner by which the hundreds of editing sites are specifically targeted. In vivo experiments using transgenic chloroplasts have shown that mRNA editing sites are recognized via *cis*-acting elements that are generally located within ~30 nucleotides of the editing site (Chaudhuri et al., 1995; Bock et al., 1996, 1997; Chaudhuri and Maliga, 1996; Reed et al., 2001). This has been confirmed in vitro using synthetic RNA substrates to explore editing specificity in both chloroplast (Miyamoto et al., 2002; Hegeman et al., 2005; Hayes et al., 2006) and mitochondrial extracts (Neuwirt et al., 2005; Takenaka et al., 2004, 2008; Verbitskiy et al., 2008). In other systems with such pervasive RNA editing, target sites are recognized by guide RNAs complementary to the RNA strand to be edited (Simpson and Emeson, 1996). However, there is good biochemical evidence that the *trans*-factors that bind to the *cis*-elements neighboring editing sites in plant organelles are proteins (Chaudhuri et al., 1995; Bock and Koop, 1997; Hirose and Sugiura, 2001). UV cross-linking experiments in an in vitro tobacco (*Nicotiana tabacum*) editing system reveals that proteins with distinct molecular masses of 25, 56, 70, 91, and 93 kD specifically bind to the *cis*-acting elements required for editing in

[1] These authors contributed equally to this work.

[2] Address correspondence to iansmall@cyllene.uwa.edu.au.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Ian Small (iansmall@cyllene.uwa.edu.au).

[W] Online version contains Web-only data.

[OA] Open Access articles can be viewed online without a subscription.

www.plantcell.org/cgi/doi/10.1105/tpc.109.071472

psbL, *psbE*, *petB*, *rpoB*, and *rpoA*, respectively (Hirose and Sugiura, 2001; Miyamoto et al., 2002, 2004; Kobayashi et al., 2008). The same protein of 95 kD specifically binds *cis*-acting elements of two editing sites, ndhB-9 and ndhF-2, showing that the same *trans*-factor can recognize more than one site (Kobayashi et al., 2008). However, none of these biochemically investigated factors has been identified.

A lack of editing of many of the sites identified in plant organelles would be expected to lead to striking phenotypes due to misexpression of proteins involved in photosynthesis or respiration. Investigation of *Arabidopsis* mutants with such phenotypes has identified a number of nuclear genes suspected to encode RNA-editing specificity factors (Kotera et al., 2005; Okuda et al., 2007, 2009; Chateigner-Boutin et al., 2008; Cai et al., 2009; Kim et al., 2009; Robbins et al., 2009; Yu et al., 2009; Zehrmann et al., 2009; Zhou et al., 2009). In each case, loss of one of these factors leads to a specific loss of editing of one or at most a few sites. Proof that any of these genetically identified factors directly targets its editing site is limited to a study of the protein CRR4. This factor is required for editing of the initiation codon of plastid *ndhD* transcripts and binds specifically within the region −25/+10 spanning the editing site (Okuda et al., 2006). CRR4, like of all of the editing factors described to date, is a pentatricopeptide repeat (PPR) protein (Schmitz-Linneweber and Small, 2008), characterized by tandem arrays of the 35–amino acid motif (Small and Peeters, 2000) for which this family of proteins is named. PPR proteins form a large family of >450 RNA binding proteins in *Arabidopsis*, the majority of which are thought to be targeted to mitochondria or chloroplasts (Lurin et al., 2004). More precisely, the reported organelle editing factors are all members of the plant-specific E and DYW subclasses of the PPR family (Table 1), characterized by distinctive C-terminal domains (Lurin et al., 2004; O'Toole et al., 2008). The E (extended) domain is a degenerate motif with some similarities to PPR motifs, while the DYW domain, named for its typical Asp-Tyr-Trp C-terminal tripeptide, is much more highly conserved, and its presence correlates phylogenetically with plant organelle RNA editing (Salone et al., 2007; Rudinger et al., 2008).

Most of the editing factors described so far were identified through forward genetic screens, generally via the distinctive phenotype caused by loss of editing. However, the similarity among all of them suggested to us that a reverse genetics approach concentrating on genes encoding similar proteins might identify new editing specificity factors. Using a high-resolution melting of amplicons technique for screening for editing defects (Chateigner-Boutin and Small, 2007), we identified six novel *Arabidopsis* mutants affected in editing of plastid mRNAs. An analysis of these mutants and comparison with the previously described mutants allows us to make some progress toward understanding the recognition of editing sites in plant organelles.

RESULTS

Identification of PPR Genes Required for RNA Editing in Plastids

All previously identified factors required for editing of specific sites are proteins of the E or DYW subclasses of the PPR family. As the number of editing sites in chloroplasts is much lower than that in mitochondria and therefore easier to screen, we chose to focus on E and DYW members predicted to be targeted to plastids by TargetP (Emanuelsson et al., 2000) or Predotar (Small et al., 2004). Sixteen candidate genes were selected because of the availability of T-DNA insertion mutants for each gene (Table 1). For each mutant, the location of the T-DNA insertion was verified by PCR and sequencing, and homozygous mutant lines identified (see Methods for the insertion sites). To test if a T-DNA insertion in these genes leads to an RNA editing defect, the status of the 34 editing sites in *Arabidopsis* plastids was systematically examined in each mutant using a high-resolution melting screen (Chateigner-Boutin and Small, 2007).

Among these 16 genes, the disruption of six of them leads to a defect in RNA editing in plastids. We respectively named these genes *OTP80*, *OTP81*, *OTP82, OTP84*, *OTP85*, and *OTP86* (for *Organelle Transcript Processing*). *OTP80* encodes a PPR protein from the E subclass, whereas *OTP81*, *82*, *84*, *85*, and *86* encode members of the DYW subclass. *OTP82* was independently identified in another screen and will be described elsewhere. Disruption of the other 10 genes did not lead to an observable defect in RNA editing of any of the 34 sites we tested.

In the mutants *otp80*, *otp81*, *otp85*, and *otp86*, a single editing site was affected, but in *otp84*, three different editing sites appeared to remain unedited (Table 1; see Supplemental Figure 1 online). All of the editing defects were confirmed by a more sensitive and quantitative poisoned primer extension assay (Figure 1). In *otp81*, editing in the *rps12* intron (at genome position 69,553) is reduced to 2% (Figure 1B). We consider the residual editing to indicate that this is a weak allele, since the T-DNA is inserted into the 5′-untranslated region of *OTP81* (see Methods). The mutants *otp80*, *otp85*, and *otp86* are totally impaired in the editing of *rpl23* (86,056), *ndhD* (116494), and *rps14* (37,161), respectively (Figures 1A to 1D and 1E). In *otp84*, three defects in RNA editing were observed: *ndhB* (94,999), *ndhF* (112349), and *psbZ* (35,800) (Figure 1C). All three sites remain completely unedited in the mutant.

To investigate whether the editing defects we observed in these mutants were secondarily caused by altered RNA processing or modified RNA stability, the transcripts affected were analyzed by RNA gel blots (Figure 2). The hybridization patterns reveal no striking differences in the pattern and level of the transcripts whose editing is impaired in the mutants (*rpl23*, *rps12* intron, *ndhB*, *ndhF*, *psbZ*, *ndhD*, and *rps14*). Furthermore, none of the other editing sites in the same transcripts is affected in any of these mutants (see Supplemental Figure 1 online).

Complementation experiments were performed for all the *otp* mutants. The introduction of a wild-type copy of the corresponding *OTP* gene into each mutant restores the editing at the defective sites to wild-type levels (Figure 1). These results support a direct and specific role of these OTP factors in RNA editing of these sites.

OTP80, OTP81, OTP84, OTP85, and OTP86 Are Localized in Plastids

The chloroplast editing defects in these mutants suggest that the PPR proteins encoded by these genes reside in plastids. To

Table 1. Candidate and Known Genes Affecting Cchloroplast RNA Editing

Arabidopsis Locus	Subclass	Predotar	TargetP	Gene Name	Editing Defect (Genome Position)	AA Change	T-DNA Insertion Lines	T-DNA Location
At2g02980	**DYW**	**pC**	**C**	***OTP85***	***ndhD* (116,494)**	**S>L**	**SAIL_544_B03**	**CDS**
At2g29760	**DYW**	**C**	**C**	***OTP81***	***rps12* intron (69,553)**		**SALK_092402**	**5′UTR**
At3g03580	DYW	pC	C		None		GABI_895H11	CDS
At3g08820	DYW	C	N		None		SALK_023916	CDS
At3g57430	**DYW**	**ER**	**C**	***OTP84***	***ndhF* (112,349), *psbZ* (35,800)**	**S>L, S>L**	**SAIL_568_C04**	**CDS**
					***ndhB* (94,999)**	**P>L**	**SALK_120902**	**CDS**
At3g62890	DYW	N	C		None		SAIL_1249_D04	CDS
							SALK_044324	CDS
At3g63370	**DYW**	**pC**	**C**	***OTP86***	***rps14* (37161)**	**S>L**	**SALK_102445**	**CDS**
At4g35130	DYW	C	C		None		SALK_118555	CDS
At1g74600	E	C	C		None		GABI_073C06	CDS
At1g77170	E	pM	C		None		SAIL_1291_C04	CDS
At3g22150	E	C	C		None		SALK_040629	CDS
At3g29230	E	pC	M		None		SAIL_205_G08	CDS
At4g04370	E	pC	N		None		SALK_025427	CDS
At4g25270	E	pC	C		None		SALK_090845	CDS
At5g59200	**E**	**C**	**C**	***OTP80***	***rpl23* (86,056)**	**S>L**	**SALK_060533**	**CDS**
							SALK_111721	**CDS**
At1g08070	**DYW**	**pC**	**C**	***OTP82***	***ndhG* (118,858), *ndhB* (95,644)**	**S>F, S>L**	**SAIL_851_G04**	**CDS**
							SALK_027812	**CDS**
At1g11290	DYW	N	C	*CRR22*	*ndhB* (96,419), *ndhD* (116,281) *rpoB* (25,779)	Okuda et al. (2009)		
At1g15510	DYW	C	C	*AtECB2*	*accD* (57,868)	Yu et al. (2009)		
At1g59720	DYW	N	C	*CRR28*	*ndhB* (96,698), *ndhD* (116,290)	Okuda et al. (2009)		
At3g22690	DYW	pC	C	*YS1*	*rpoB* (25,992)	Zhou et al. (2009)		
At5g13270	DYW	C	C	*RARE1*	*accD* (57,868)	Robbins et al. (2009)		
At5g48910	DYW	N	C	*LPA66*	*psbF* (63,985)	Cai et al. (2009)		
At2g45350	E	pC	N	*CRR4*	*ndhD* (117,166)	Kotera et al. (2005)		
At1g05750	E+	C	C	*CLB19*	*rpoA* (78,691), *clpP* (69,942)	Chateigner-Boutin et al. (2008)		
At5g55740	E+	C	C	*CRR21*	*ndhD* (116,785)	Okuda et al. (2007)		

pC, potentially plastid; C, plastid; pM, potentially mitochondrial; M, mitochondrial; ER, secretory pathway; N, none; AA, amino acid; UTR, untranslated region; CDS, coding sequence. The new editing factors revealed by this study are indicated in bold.

verify this, the N-terminal 100 amino acids of the proteins (72 amino acids in the case of OTP85) were fused with green fluorescent protein (GFP) driven by expression from the cauliflower mosaic virus 35S promoter. The plasmids containing the chimeric genes were transiently expressed in wild-type *Arabidopsis* cells by bombardment. Analysis of GFP fluorescence in transformed cells revealed that the fluorescence colocalized with a chloroplast marker: the fusion of red fluorescent protein (RFP) with the small subunit of *Arabidopsis* ribulose biphosphate carboxylase (Carrie et al., 2009) (see Supplemental Figure 2 online). Thus, OTP80, OTP81, OTP84, OTP85, and OTP86 are localized in plastids consistent with their role in editing plastid transcripts.

Requirement of RNA Editing at the Sites Affected by OTP Factors

The site that is not edited in *otp81* lies in the intron of *rps12*, and this altered sequence might affect splicing of *rps12* transcripts. However, RNA gel blot hybridization shows no difference in the pattern of *rps12* transcripts between *otp81* and the wild type (Figure 2B). Thus, editing of the *rps12* intron is not necessary for the stabilization of the precursor transcripts, nor correct splicing, and there is no reason to think that the amount or function of Rps12 is deleteriously affected in the *otp81* mutant.

The conversion of cytidine to uridine at the sites affected in the other *otp* mutants is not silent and changes the nature of the amino acids in the proteins encoded by these transcripts (Table 1). The alteration of editing at these sites might be expected to have an impact on the function of these proteins. Nevertheless, *otp80*, *otp84*, *otp85*, and *otp86* show normal growth under standard conditions (see Supplemental Figure 3 online).

The mutants *otp80* and *otp86* are both affected in editing of transcripts encoding ribosomal proteins and thus could be expected to show translation defects. Translation in *Arabidopsis* plastids is thought to be essential (reviewed in Schmitz-Linneweber and Small, 2008). However, methylene blue staining of total RNA from *otp80* and *otp86* did not reveal any major defects in the integrity or quantity of the chloroplast rRNAs (indicated by arrows in Figures 2A and 2G). This suggests that these editing events are not essential for normal ribosome accumulation in plastids, in agreement with their normal growth phenotype, despite the presumed effects on the amino acid sequences of Rps14 and Rpl23 (Table 1).

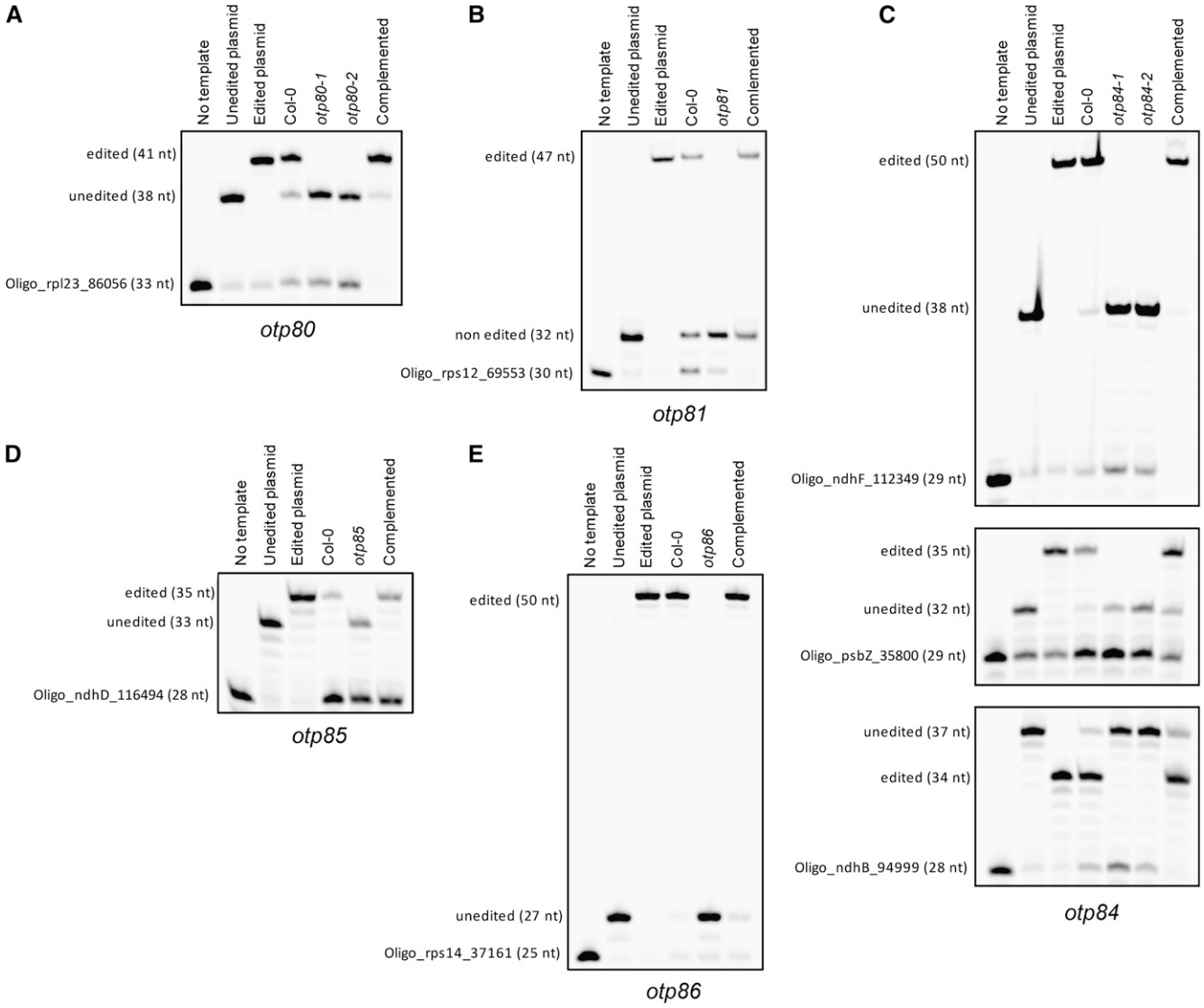


Figure 1. Editing Defects in *otp80*, *otp81*, *otp84*, *otp85*, and *otp86* Mutants.

Poisoned primer extension assays were conducted on the editing sites *rpl23* (nucleotide 86,056) for *otp80* **(A)**, *rps12* intron (69,553) for *otp81* **(B)**, *ndhF* (112,349), *psbZ* (35,800), and *ndhB* (94,999) for *otp84* **(C)**, top to bottom; *ndhD* (116494) for *otp85* **(D)**, and *rps14* (37,161) for *otp86* **(E)**. The editing sites are specified relative to the nucleotide sequence of the complete *Arabidopsis* chloroplast genome.

The mutant *otp85* lacks editing of *ndhD* (nucleotide 116494). RNA editing at this site changes a Ser codon to a Leu codon. The mutant *otp84* is impaired in the editing of the three editing sites *ndhB* (94,999), *ndhF* (112,349), and *psbZ* (35,800). RNA editing at *ndhB* (94,999) converts a Pro codon to a Leu codon, while editing at *ndhF* (112349) and *psbZ* (35,800) convert Ser codons to Leu codons (Table 1). NdhB, NdhD, and NdhF are subunits of the chloroplast NAD(P)H dehydrogenase (NDH) complex involved in photosystem I cyclic electron flow (Shikanai, 2007). The NDH complex catalyzes electron donation to plastoquinone from the stromal electron pool. PsbZ is a core subunit of photosystem II (Swiatek et al., 2001). The editing defects in *ndhB*, *ndhF*, and *psbZ* could affect the function of the encoded proteins, which should alter photosynthetic parameters.

NDH complex activity can be monitored as a transient increase in chlorophyll fluorescence reflecting plastoquinone reduction after turning off actinic light (Shikanai et al., 1998). Figure 3 shows a typical chlorophyll fluorescence trace from wild-type *Arabidopsis* and in *crr21*, a mutant lacking NDH activity (Okuda et al., 2007). In *otp84*, the postillumination increase of chlorophyll fluorescence is modified but, unlike *crr21*, is not completely abolished. This indicates that NDH activity is diminished, but not absent, in the mutant. This also suggests that Leu residues at positions 494 in NdhB and/or 97 in NdhF are required for optimal NDH stability or activity. The postillumination increase of chlorophyll fluorescence in *otp85* is not modified (Figure 4), indicating that a Leu at position 225 in NdhD is not essential for NDH activity.

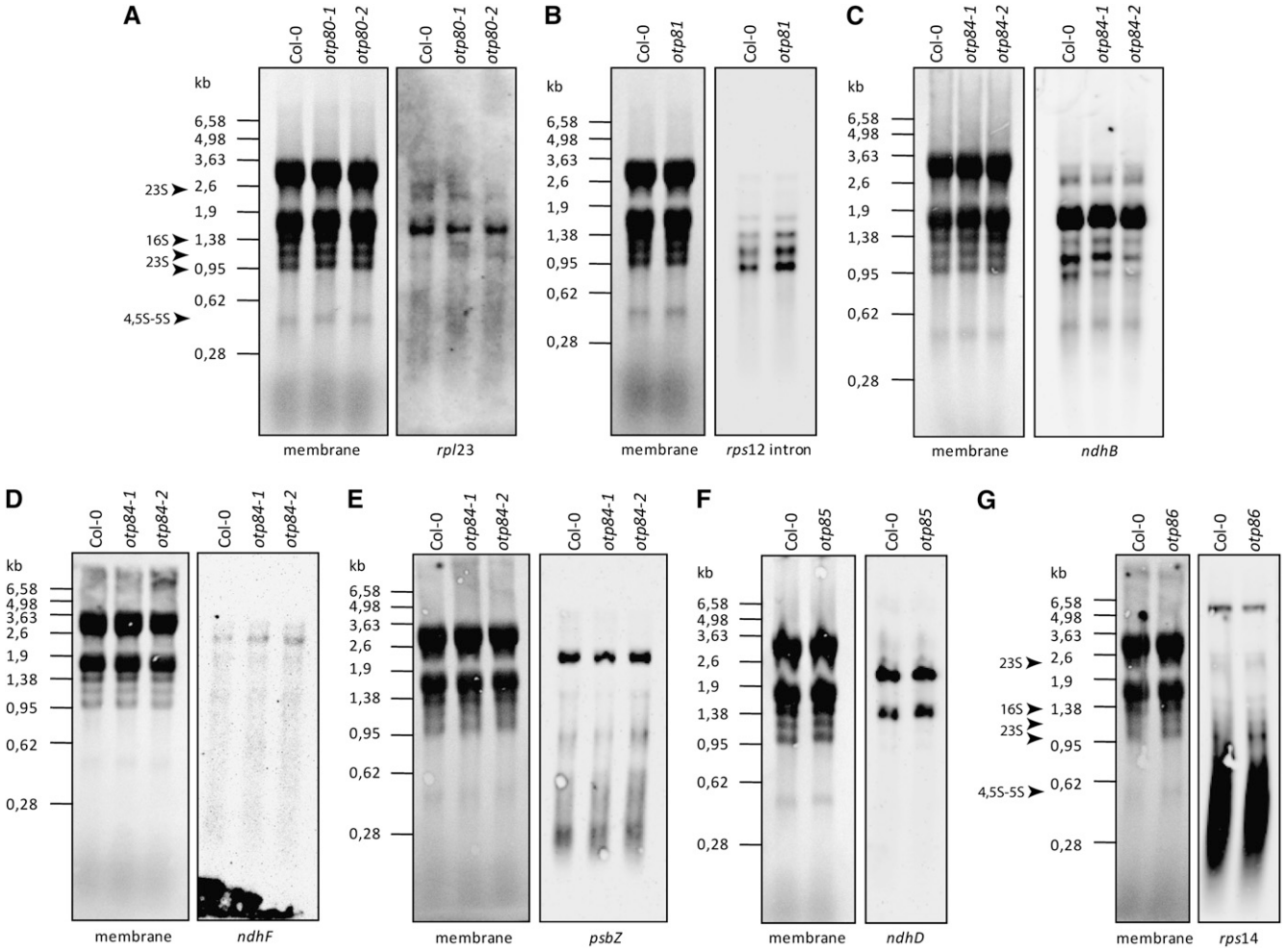


Figure 2. Transcript Profiles of Genes with the Editing Defects in *otp80*, *otp81*, *otp84*, *otp85*, and *otp86* Mutants.

otp80 **(A)**, *otp81* **(B)**, *otp84* (**[C]** to **[E]**), *otp85* **(F)**, and *otp86* **(G)** mutants. Fifteen micrograms of RNA from 15-d-old seedlings was loaded on formaldehyde agarose gels and transferred onto a membrane. Hybridizations were performed under high-stringency conditions using antisense RNA probes for the genes specified for each blot. The sizes of RNA markers are shown in kilobases. rRNA stained with methylene blue on the membranes is shown as a loading control (membrane). The arrows indicate the plastid rRNAs in *otp80* and *otp85* mutants.

RNA editing defects result in amino acid changes that may directly alter protein function or act by destabilizing the protein or by affecting its ability to form complexes with other proteins. To assess whether the subunits affected in *otp84* and *otp85* stably accumulate in vivo, protein blots were analyzed using antibodies against NdhH and PsbZ (Figure 4). The NDH complex is unstable without NdhB and NdhD (Peng et al., 2008), and an antibody against NdhH can be used to monitor accumulation of the complex. In *otp84*, the level of NdhH is reduced to 50 to 25% of the wild type. This result suggests that Leu residues at positions 494 in NdhB and/or 97 in NdhF are important for normal accumulation of the NDH complex. The decrease in the amount of the NDH complex largely explains the observed decrease in activity, although we cannot exclude that the specific activity of the complex is also decreased. In *otp85*, the accumulation of the NDH complex is not affected.

In addition to *ndhB* (94,999) and *ndhF* (112349), *otp84* is impaired in editing of *psbZ* (35,800), leading to retention of a Ser (rather than Leu) codon at position 17 in PsbZ, a core subunit of photosystem II (Swiatek et al., 2001). Immunodetection of PsbZ on blots of *otp84* extracts (Figure 4) reveals no defect in the accumulation of the protein in the mutant. Consistent with the normal growth of *otp84* (see Supplemental Figure 3 online), there were no alterations to the rate of electron transport through photosystem II nor to nonphotochemical quenching of chlorophyll fluorescence, which reflects ΔpH formation (see Supplemental Figure 4 online). These results suggest that a Leu at position 17 in PsbZ is not important for its function under standard growth conditions.

Multiple Editing Sites and Recognition Specificity

A major outstanding question is how editing factors recognize their target sites and of particular interest are those factors that

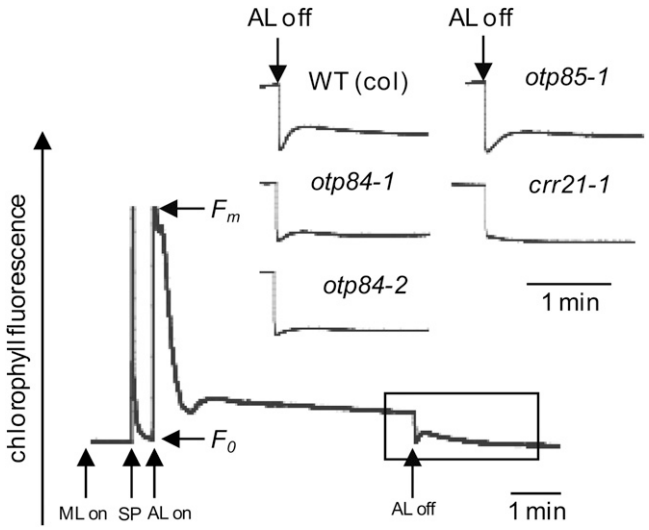


Figure 3. Monitoring of NDH Activity Using Chlorophyll Fluorescence Analysis for *otp84* and *otp85* Mutants.

The curve shows a typical trace of chlorophyll fluorescence in the wild type and a mutant totally impaired in NDH activity (*crr21*) (Okuda et al., 2007) compared with traces from *otp85* and two independent *otp84* mutants (*otp84-1* and *otp84-2*). Leaves were exposed to actinic light (AL) (50 μmol of photons m^{-2} s^{-1}) for 5 min. AL was turned off, and the subsequent change in chlorophyll fluorescence level was monitored. The transient increase in chlorophyll fluorescence is due to the plastoquinone reduction based on NDH activity. Insets are magnified traces from the boxed area. The fluorescence levels were normalized by the maximum fluorescence at closed photosystem II centers in the dark (*Fm*) levels. ML, measuring light; SP, a saturating pulse of white light.

appear to recognize more than one site. Biochemical evidence suggests that most of the sequence recognition is concentrated within the 15 nucleotides immediately upstream of the editing site (Hirose and Sugiura, 2001; Miyamoto et al., 2002, 2004; Kobayashi et al., 2008). However, as remarked in previous work (Chateigner-Boutin et al., 2008; Okuda et al., 2009), editing sites requiring the same factor do not show extensive sequence identity in this crucial region, raising doubts as to how these putative specificity factors might operate. To investigate these questions, we undertook a systematic bioinformatic analysis of all plastid editing sites in *Arabidopsis* and the factors that recognize them.

Confirmation that editing sites requiring the same factor are not particularly similar to each other is shown in Figure 5. Only *ndhB* (94,999) and *ndhF* (112349) (requiring OTP84) pair together, and even this grouping is only weakly supported. The only strongly supported grouping in this tree is between *rpoB* (25,992) and *ndhB*(96,419), but these require different editing factors (YS1 and CRR22, respectively). At first sight, these data raise doubts that the identified editing factors can work as commonly assumed (i.e., by binding to the region immediately upstream of the target C). We reasoned that we had sufficient data to test this assumption and therefore examined whether the apparently dissimilar target sites requiring the same factors nevertheless contain sufficient conserved nucleotides to unambiguously define a binding site consensus. If an unambiguous consensus could be found, then this would be consistent with the protein in question acting as a specificity factor by binding to this consensus sequence.

For each set of editing sites requiring a single factor, we calculated six possible consensus sequences using the following assumptions: (1) that PPR proteins can distinguish all four nucleotides; (2) that PPR proteins can distinguish purines (A and G) from pyrimidines (C and U); (3) that PPR proteins can distinguish the number of hydrogen bonding groups in each base (i.e., can distinguish A or U from G or C) or various combinations of these criteria (Table 2). Each consensus was then matched against both strands of the entire *Arabidopsis* chloroplast genome and the number of matches recorded. Finally, matches within duplicated or untranscribed sequences were discounted. The results indicate that with the exception of the sites recognized by CRR22, in all other cases there is sufficient conservation of sequence to unambiguously identify the edited sites from all other transcribed sequences within plastids, based on the assumptions (1) and (2) listed above.

DISCUSSION

A reverse genetic screen of T-DNA insertion mutants allowed us to identify six new PPR proteins required for RNA editing of nine sites in plastid RNAs. Unlike the majority of the previously characterized editing mutants (Okuda et al., 2006, 2007, 2009; Chateigner-Boutin et al., 2008; Kim et al., 2009; Yu et al., 2009; Zhou et al., 2009), which were identified via visible growth or fluorescence phenotypes, these mutants were identified directly via a screen for unedited RNAs. It is striking that of the eight mutants discovered in this way, including RARE1 (Robbins et al., 2009) and MEF1 (Zehrmann et al., 2009), none have growth or

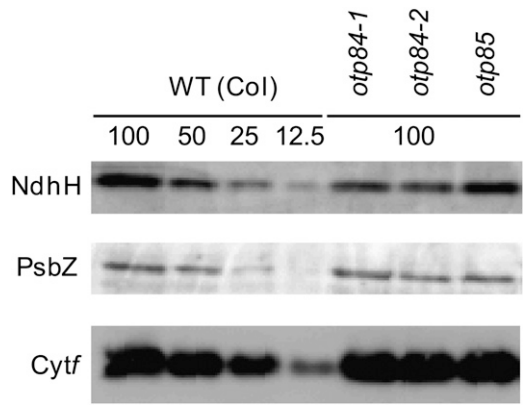


Figure 4. Protein Blot Analysis of the NDH Complex and the Major Photosynthetic Complexes.

Immunodetection of an NDH subunit, NdhH; a subunit of photosystem II, PsbZ; and a subunit of the Cytb_6f complex, cytochrome *f*. The proteins were extracted from thylakoid membrane fractions. Lanes were loaded with protein samples corresponding to 0.5 μg chlorophyll for Cyt*f*, 1 μg chlorophyll for PsbZ, and 5 μg chlorophyll for NdhH (100%) and the series of dilutions indicated as a percentage of the original sample.

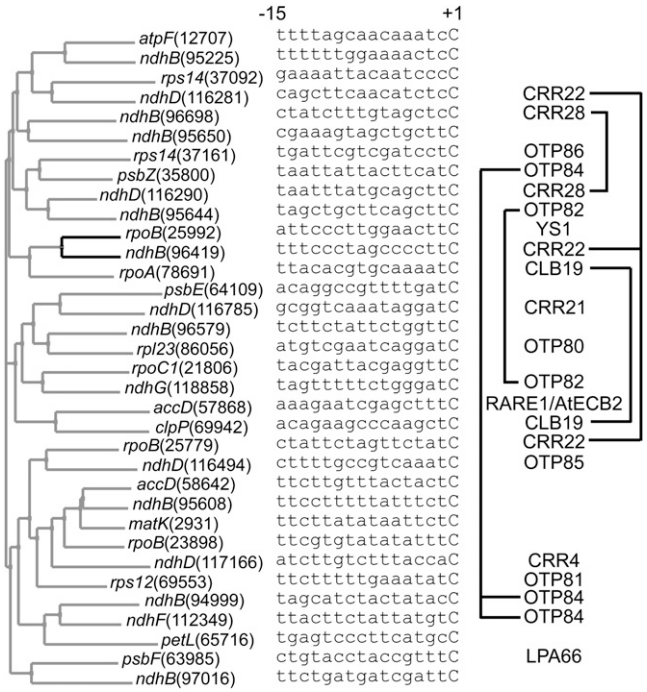


Figure 5. Assessment of Sequence Similarities among Editing Sites.

For the 34 known editing sites in the *Arabidopsis* chloroplast genome (Chateigner-Boutin and Small, 2007), the 15 nucleotides upstream of the editing site were aligned with respect to the editing site (the final C in each sequence) and a distance tree was constructed. Only the grouping of *rpoB* (25,992) with *ndhB* (96,419) has bootstrap support of >50%, and this is indicated by the black lines in the tree. The right panel indicates the PPR proteins involved in the recognition of each editing site.

physiological defects, apart from the partial decrease in NDH activity observed for *otp84*. This rather suggests that many more RNA editing events may be phenotypically silent than previously expected, at least under optimal growth conditions. This observation has implications for the evolutionary stability of editing events (Tillich et al., 2009), as it implies that selection pressure on both the editing sites and the factors that recognize them may be weaker than often thought. It should be noted, however, that as editing almost always restores codons for conserved amino acids (Figure 6), it is likely that under some conditions it would be possible to observe deleterious effects of a lack of editing of these apparently silent sites.

More than half of the candidate genes we examined could not be linked to RNA editing in plastids. This might be for a number of different reasons, including genetic redundancy, errors in targeting predictions (such that the protein functions in mitochondria, not plastids), incomplete surveying of editing sites (it is still possible that there are more than the 34 known editing sites in *Arabidopsis* plastid RNAs), or because the protein has functions unrelated to editing. As an example of the last case, the DYW protein CRR2 is required for RNA processing (accumulation of monocistronic *ndhB* mRNA), not RNA editing (Hashimoto et al., 2003).

Although prediction of editing factors is not yet infallible, nevertheless 22 of the 34 known editing sites in plastids have been assigned corresponding PPR *trans*-factors, and we expect that the remaining editing events are also likely to require a PPR protein from the plant-specific PLS subfamily. A total of 44 of these proteins are predicted to be targeted to plastids (Lurin et al., 2004), which is more than sufficient to account for all the known editing sites in plastids. In mitochondria, the number of editing sites is considerably higher than the number of PLS subfamily PPR proteins predicted to be targeted to mitochondria: 82 PPR proteins (Lurin et al., 2004) for >500 editing sites (Giegé and Brennicke, 1999; Bentolila et al., 2008; Zehrmann et al., 2008). This suggests that in mitochondria, a single PPR editing factor should, on average, cover more sites than in plastids. Indeed, in plastids, the majority of editing factors are required for a single site, with a maximum of three in the cases of CRR22 (Okuda et al., 2009) and OTP84. The first mitochondrial editing factors to be found, OGR1 and MEF-1, are required for at least seven and three editing sites in rice (*Oryza sativa*) and *Arabidopsis* mitochondria, respectively (Kim et al., 2009; Zehrmann et al., 2009).

The sequences surrounding the editing sites identified in plant organelles do not show any unequivocal consensus, apart the from the one or two bases immediately surrounding the site (reviewed in Mulligan et al., 1999 and reexamined in Cummings and Myers, 2004). This probably represents preferences of the enzyme catalyzing the editing reaction rather than a target for postulated specificity factor(s). The key question of how specific C residues are recognized for editing from all other C residues in organelle transcripts remained unanswered for many years. Experiments with transgenic plastids and in vitro RNA editing assays using organelle extracts delimited the primary region recognized by putative *trans*-factors to the 20 or so nucleotides immediately upstream of the editing site, with, in most cases, the most important recognition elements situated from −5 to −15 with respect to the edited C (reviewed in Shikanai, 2006). Binding to this region was confirmed for CRR4, the first editing factor to be genetically identified (Okuda et al., 2006), and it has been assumed that the other PPR proteins identified as editing factors function in a similar way (Okuda et al., 2007, 2009; Chateigner-Boutin et al., 2008; Cai et al., 2009; Kim et al., 2009; Robbins et al., 2009; Yu et al., 2009; Zehrmann et al., 2009; Zhou et al., 2009). However, no proof of specific binding has been shown for any of these subsequent editing factors, including the six new ones identified in the screen described here.

Some confirmation of specific binding comes from comparing the factors genetically identified in *Arabidopsis* to those biochemically identified in tobacco. Cross-linking experiments trapping putative protein *trans*-factors bound to the *cis*-acting elements of editing sites (Kobayashi et al., 2008) identified a 91-kD protein binding the rpoB-1 site which, based on size, could correspond to the tobacco ortholog of YS1, a PPR protein required for the editing of the same site in *Arabidopsis* plastids (Zhou et al., 2009) and whose predicted molecular mass (discounting the plastid targeting sequence) is 90.3 kD. An even more compelling match is seen between the 95-kD protein binding the two RNA editing sites ndhB-9 and ndhF-1 in tobacco (Kobayashi et al., 2008) and OTP84, which we show to recognize the equivalent *Arabidopsis* editing sites *ndhB* (94,999) and *ndhF* (112349). The predicted molecular mass of OTP84 (discounting

Table 2. Analyses of the Editing Sites Requiring the Same PPR Proteins

PPR		Consensus	Hits	Filtered Hits	Location
OTP82	a	UAG-U-U—G-UC	15		
	b	YRRYY-YYYYRR-YC	27		
	c	UAGYU-YUYYRG-UC	3	2	***ndhB*, *ndhG***
	d	WWS-W-W—SSWWC	260		
	e	UAG-U-U—GSWUC	4		
	f	UAGYU-YUYYRGSWUC	3	2	***ndhB*, *ndhG***
CRR22	a	——-A——-C	17,078		
	b	Y-YYYYRRY-YY-YC	59		
	c	Y-YYYYARY-YY-YC	35	12	*psbK*, *ycf2*, *ycf1*, *matK*, 3′UTR *atpH*, ***rpoB***, 3′UTR *rps4*, ***ndhB***, ***ndhD***, *ndhG*, *ndhA*, *ycf1*
	d	-W——W——W-C	15,439		
	e	-W——A——W-C	7,646		
	f	YW-YYYYARY-YYWYC	23	8	*psbK*, *ycf1*, 3′UTR *atpH*, ***rpoB***, ***ndhB***, ***ndhD***, *ndhG*, *ndhA*
OTP84	a	U——UA-U—-C	444		
	b	Y-RY-YYRYY-YRYC	28		
	c	U-RY-YUAYU-YRYC	7	3	***psbZ*, *ndhB*, *ndhF***
	d	WW-WW-WW-WW—C	1,732		
	e	UW-WW-UA-UW—C	88		
	f	UWRYWWYUAYUWYRYC	5	3	***psbZ*, *ndhB*, *ndhF***
CLB19	a	-A-A-G-CAA-UC	16		
	b	-YR-R-RY-YRRR-YC	74		
	c	-YA-A-GY-CAAR-UC	4	2	***clpP*, *rpoA***
	d	W-WSW-S-SSWW-WC	39		
	e	W-ASA-G-SCAA-UC	2		
	f	WYASA-GYSCAAR-UC	2	2	***clpP*, *rpoA***
CRR28	a	-AU-U-UG-AGCU-C	16		
	b	Y-RYYY-YRYRRYYYC	6		
	c	Y-AUYU-UG-AGCU-C	3	2	***ndhB*, *ndhD***
	d	-WWW-WWWS-WSSW-C	79		
	e	-WAU-UWUG-AGCU-C	3		
	f	YWAUYUWUG-AGCU-C	3	2	***ndhB*, *ndhD***

Consensus sequences were derived as follows from the 15 nucleotides immediately upstream of the edited C (and including the editing site itself): (a) full conservation of nucleotides (A, U, G, and C); (b) conservation of purines (A or G = R) or pyrimidines (U or C = Y); (c) combination of a and b; (d) conservation of number of hydrogen bonding groups (A or U = W, G or C = S); (e) combination of a and d; (f) combination of a, b, and d. The "Hits" column indicates the number of times each consensus is found within both strands of the *Arabidopsis* plastid genome. "Filtered Hits" removes duplicate matches within the inverted repeats, matches to noncoding strands, and matches to intergenic regions that do not form stably accumulated transcripts. The editing sites known to be recognized by each factor are highlighted in bold. UTR, untranslated region.

the plastid targeting sequence) is 94.8 kD. All the observations so far are entirely consistent with editing site specificity being determined by binding of the relevant PPR protein to the sequence upstream of the target C, in accordance with the general view of how PPR proteins function in RNA metabolism (Delannoy et al., 2007). It is worth noting that there is no significant difference in length, number of motifs, or any other aspect of the protein sequence or structure that we can find to distinguish PPR proteins recognizing multiple target sites from those that recognize single sites (Figure 7).

Only a few PPR proteins have been shown unambiguously to bind to multiple target sites, of which the best-studied are the maize (*Zea mays*) proteins CRP1 (Schmitz-Linneweber et al., 2005) and PPR10 (Pfalz et al., 2009). In these two cases the multiple target sites have almost identical sequences. This is not true of the multiple editing sites targeted by single factors. In our study, a distance tree relating the *Arabidopsis* plastid editing sites (Figure 5) does not group the editing sites known to be recognized by a single PPR factor except for the *ndhB* (94,999) and *ndhF* (112349) sites mentioned above. Previously proposed clusters based on apparent sequence similarity (Chateigner-Boutin and Hanson, 2002; Chateigner-Boutin and Hanson, 2003) do not exactly correspond to groups of sites recognized by single factors. This prompted us to test whether the commonly held assumption that these editing factors bind the biochemically defined *cis*-elements is consistent with the data.

We reasoned that an unambiguous consensus between sites requiring the same editing factor would be consistent with the

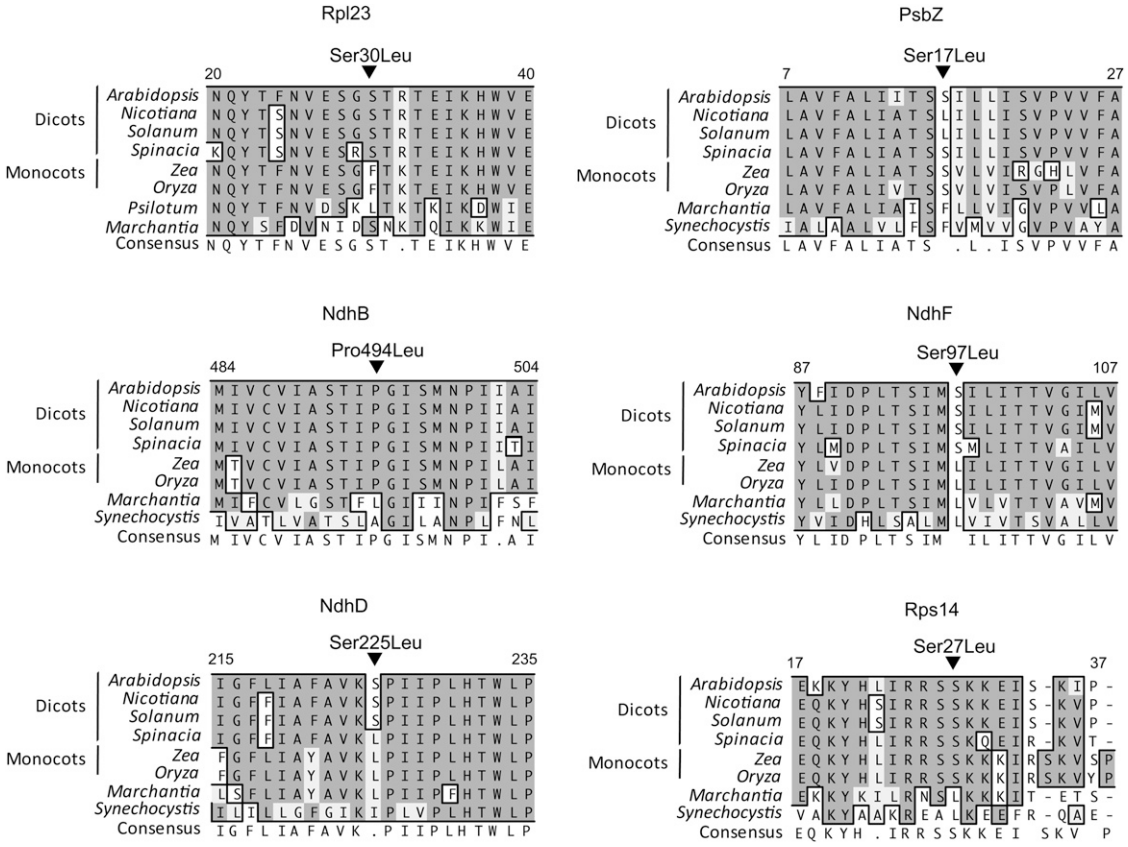


Figure 6. Partial Sequence Alignments of Rpl23, PsbZ, NdhB, NdhF, Rps14, and NdhD around the Amino Acids Affected by RNA Editing.

Arabidopsis Rpl23, PsbZ, NdhB, NdhF, Rps14, and NdhD proteins were aligned with their homologs from other species. The alignment was performed using ClustalW (Thompson et al., 1994). Amino acids that are fully or semiconserved are shaded black or gray, respectively. Numbers indicate amino acid positions in the protein. The arrows above the sequences indicate the positions of edited codons.

hypothesis that the factor is involved in target site recognition. We addressed this by generating various consensus sequences (Table 2) from the multiple target sites, using the results from published in vitro experiments (Hirose and Sugiura, 2001; Miyamoto et al., 2002, 2004; Kobayashi et al., 2008) to delimit the region of interest from −15 to the editing site. We found that it was possible to define an unambiguous 15-nucleotide consensus for each of the editing factors required for editing of multiple sites, except in the case of CRR22. We conclude that the data are consistent with the hypothesis that these PPR editing factors bind to multiple target sites and consistent with the biochemical evidence that recognition involves the nucleotides upstream of the editing site. The one exception is CRR22, which will be discussed in more detail later.

In the cases where a consensus can be found, the pattern of nucleotide similarity is informative regarding what sequence features might be being recognized. RNA nucleotides tend to interact base-specifically with proteins in one of two ways: via stacking interactions with the aromatic rings or via hydrogen bonding. Sequence specificity can be achieved by positive base-specific bonding interactions or by negative steric hindrance (Auweter et al., 2006). Based on these considerations, we considered three (mutually compatible) hypotheses: that editing factors can distinguish all four RNA nucleotides, that they can distinguish purines (A and G) from pyrimidines (C and U), or that they can distinguish bases by Watson-Crick hydrogen bonding patterns (G and C versus A and U). None of these consensuses alone provided sufficient specificity to unambiguously define the edited sites within the plastid transcriptome, but combining conserved nucleotide positions with conserved purines or pyrimidines is sufficient (in the cases where a consensus could be achieved). We conclude that to achieve the required specificity, editing factors must be able to distinguish purine nucleotides from pyrimidine nucleotides, and at least at some positions, be able to uniquely distinguish one of the four nucleotides. The relatively poor performance of the consensus based on A/U versus C/G suggests that this distinction is not important in determining binding specificity.

This analysis is consistent with previous biochemical data. Kobayashi et al. (2008) conducted experiments using RNAs competing with the factor required for editing ndhB-9 and ndhF-1 (as explained above, this is almost certainly the tobacco ortholog of OTP84). Competitor RNAs were constructed by scanning mutagenesis in which each successive five-nucleotide

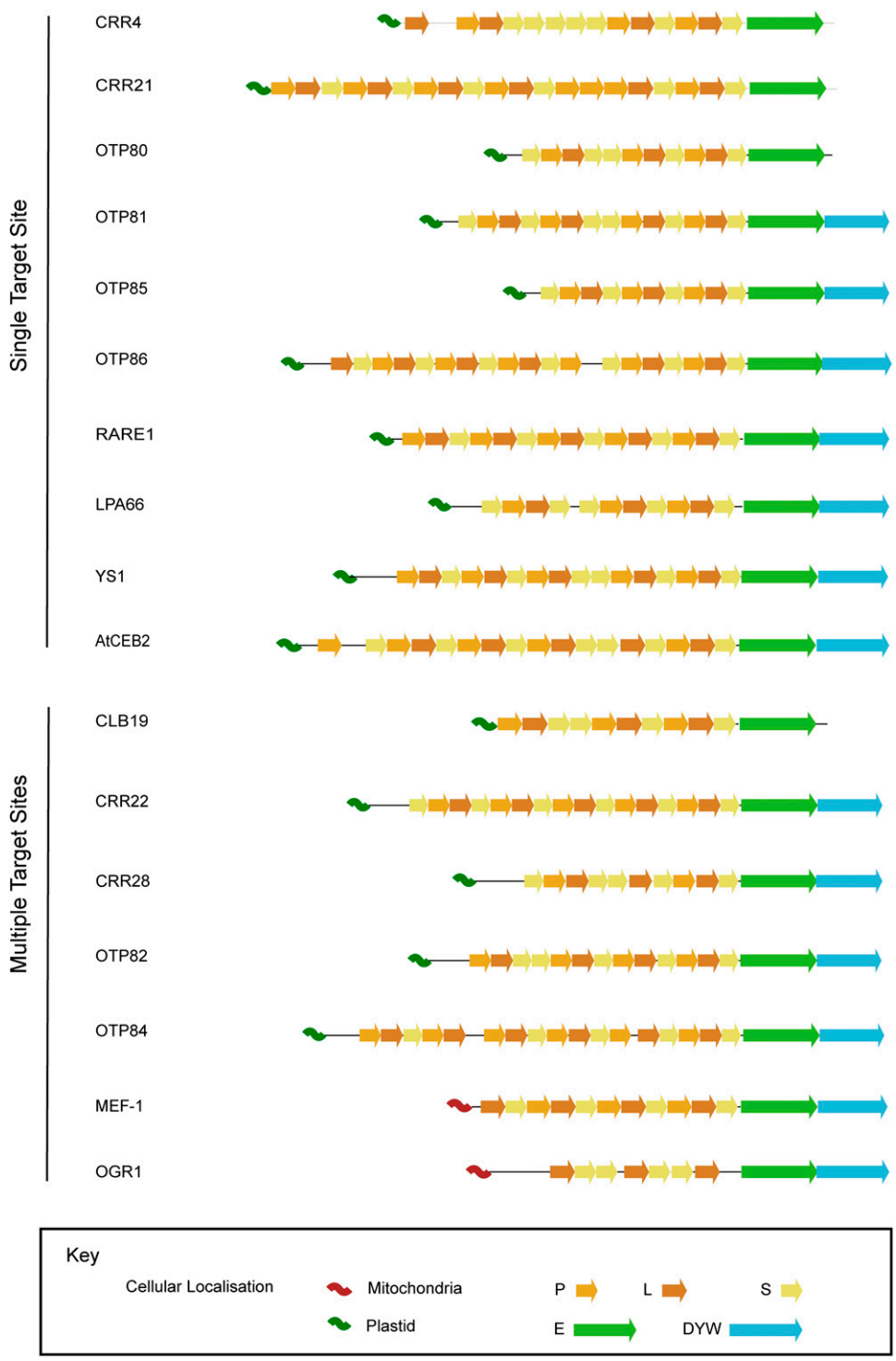


Figure 7. Schematic Structure of Plant PPR Proteins Involved in Editing of Plastid or Mitochondrial Transcripts.

PPR proteins are grouped according to their number of target sites (single or multiple). All the identified PPR motifs are indicated (as defined in Lurin et al., 2004), as well as the C-terminal E and DYW domains and the N-terminal organelle targeting sequences. P represents canonical 35–amino acid PPR motifs; L (long) and S (short) represent variant PPR motifs with generally 36 and 31 amino acids, respectively.

block in the region from −15 to −1 was substituted with its complementary nucleotide sequence. These mutations therefore switched the purine/pyrimidine nature of the residues in the upstream *cis*-elements. For both sites, mutations introduced into the −15 to −6 region abolished competition for the binding of the editing factor, whereas mutations in the −5 to −1 region had a weaker effect. Consistent with this, mutagenesis of the *ndhF* RNA substrate showed that purine/pyrimidine exchanges in the −15 to −1 upstream region abolish in vitro editing of the ndhF-1 site (Sasaki et al., 2006).

Throughout the analysis discussed above, the exception is CRR22. This PPR editing factor is required for editing the three sites *ndhB* (96,419), *ndhD* (116,281), and *rpoB* (25,779) (Okuda et al., 2009), but the sequence conservation between these three sites is not sufficient to explain how these three sites can be distinguished from other equally similar sequences scattered throughout the plastid transcriptome (Table 2). One possibility might be that CRR22 recognizes sequences outside the region studied here, but the sequence similarity around these sites is no greater over a more extended region. Another possibility is that different arrays of PPR motifs within CRR22 recognize different target sites, thus allowing a single protein to bind to several unrelated targets. However, CRR22 is not longer than other editing factors (Figure 7), which might be expected if it contained two or more independent binding domains. We therefore need to consider whether CRR22 truly acts as a specificity factor at all three sites it is required for or whether there might be an alternative explanation for its involvement in editing at one or more of these sites. Recently, two different PPR editing factors, RARE1 and ECB2, were demonstrated to be required for editing of a single site, *accD* (57,868), in *Arabidopsis* chloroplasts (Robbins et al., 2009; Yu et al., 2009). As all the other evidence suggests that a single PPR protein is sufficient to define editing specificity, the implication of two proteins in the editing of *accD* (57,868) implies that one of them may have a different role, as might be the case for CRR22. The most attractive hypothesis would be that this different role is to recruit the editing enzyme or even to catalyze the editing reaction itself. Like CRR22, both RARE1 and AtECB2 belong to the DYW subfamily, and one model for RNA editing proposes that the C-terminal DYW domain carries the cytidine deaminase editing activity (Salone et al., 2007).

METHODS

All the primers used in this study are listed in Supplemental Table 1 online.

Plant Material

Arabidopsis thaliana ecotype Columbia (Col-0) was used in this study. The T-DNA insertion mutant lines were obtained from the ABRC Stock Center. Accession numbers are provided in Table 1.

Genetic Analysis

Total cellular DNA was isolated as described by Edwards et al. (1991). Plants were genotyped for homozygous lines by PCR, and the insertion position was confirmed by sequencing with a T-DNA left border primer. *otp80-1* (SALK_111721, Col-0, insertion site +141), *otp80-2* (SALK_060533, Col-0, insertion site +221), *otp81* (SALK_092402, Col-0, insertion site −15), *otp84-1* (SAIL_568_C04, Col-0, insertion site +932), *otp84-2* (SALK_120902, Col-0, insertion site +642), *otp85* (SAIL_544_B03, Col-0, insertion site +1328), and *otp86* (SALK_102445, Col-0, insertion site +698).

Analysis of Targeting via GFP Fusions

The first 300 bp of the coding sequences of the PPR genes were amplified using Phusion DNA polymerase (Finnzymes) with primers listed in Supplemental Table 1 containing the *attB* sites for Gateway cloning according to the manufacturer's instructions (Invitrogen). The GFP vector used and the chloroplast targeting marker encoding RFP fused to the small subunit of *Arabidopsis* ribulose biphosphate carboxylase were kindly provided by James Whelan (University of Western Australia) (Carrie et al., 2009). Biolistic transformations of GFP and RFP constructs were performed on *Arabidopsis* cell culture (Carrie et al., 2007). The GFP construct and the chloroplast RFP marker (5 μg each) were coprecipitated onto gold particles and transformed using the biolistic PDS-1000/He system (Bio-Rad). Particles were bombarded onto 2 mL of *Arabidopsis* cell suspension resting on filter paper on osmoticum plates (2.17 g/L Murashige and Skoog modified basal salt mixture, 30 g/L sucrose, 0.5 mg/L naphthalene acetic acid, 0.05 mg/L kinetin, and 36.44 g/L mannitol). After bombardment, the cells were placed in the dark at 22°C for 24 h. Observation of transient GFP and RFP expression was performed using an Olympus BX61 fluorescence microscope with excitation wavelengths of 460/480 nm (GFP) and 535/555 nm (RFP) and emission wavelengths of 495 to 540 nm (GFP) and 570 to 625 nm (RFP). Subsequent images were captured using Cell imaging software as previously described (Carrie et al., 2007; Murcha et al., 2007).

Genetic Complementation

The 1870-, 3138-, 2673-, 2278-, and 2883-bp fragments containing the respective coding sequence of *OTP80*, *OTP81*, *OTP84*, *OTP85*, and *OTP86* were amplified by PCR on total cellular DNA. These constructs were cloned into pGWB1 (*OTP80*, *OTP81*, and *OTP85*) or pGWB2 (*OTP84* and *OTP86*) binary vectors and introduced into *otp* mutants via *Agrobacterium tumefaciens* GV3101. Transformants were obtained by selection on Murashige and Skoog agar plates containing 25μg/mL hygromycin and confirmed by PCR.

Analysis of RNA Editing

High-resolution melting analysis of amplicons was performed as previously described (Chateigner-Boutin and Small, 2007) using the primers listed by Okuda et al. (2009). Poisoned primer extension of RT-PCR products was performed as described by Chateigner-Boutin and Small (2007). RT-PCR products were obtained with primers surrounding the editing sites and serve as templates for the extension reaction from a labeled 6-carboxyfluorescein primer that anneals next to the target editing site. The extension is stopped by the incorporation of ddGTP or ddCTP at the location of the editing site for unedited molecules producing a short unedited product. The extension is stopped at the next G/C for the edited molecules producing a longer edited product. For the *ndhB* (94,999) site, the extension is stopped by the incorporation of ddATP leading to a short edited product and a long unedited product.

RNA Preparation and Analysis

Total RNA from leaves of 15-d-old plantlets was isolated using TRIzol reagent (Invitrogen) as recommended by the manufacturer. Fifteen micrograms of RNA was fractionated on 1.2% (w/v) formaldehyde agarose gels and transferred onto Hybond N^+ nylon membranes

(GE Healthcare). RNA integrity, loading, and transfer were checked by staining the membrane with methylene blue.

RNA probes were internally labeled with biotinylated cytidine by transcription of PCR products cloned in pGEM-T Easy vector (Promega). The primers used for the PCR are listed in Supplemental Table 1 online. Clones with inserts in antisense orientation were amplified by PCR using the forward primer and M13/pUC reverse. The PCR products served as a template for in vitro transcription with SP6 polymerase following the manufacturer's instructions (Maxiscript Ambion).

To prepare the *ndhF* RNA probe, the *ndhF* sequence was amplified using an antisense primer linked to the T7 promoter. The PCR product served as template for in vitro transcription with T7 polymerase (Maxiscript Ambion).

Prehybridization of the membrane was performed for 1 h in hybridization buffer (5× SSC, 50% [v/v] formamide, 0.5% SDS, and 100 μg/mL heparin) at 68°C. Hybridization with RNA probes was performed in the same buffer overnight at 68°C, followed by three 15-min washes at room temperature in 1× SSC/0.5% SDS and two washes at 60°C in 0.1× SSC/0.1% SDS for 20 min and 1 h, respectively. Signal detection was performed using the Chemiluminescent Nucleic Acid Detection Module (Pierce) and read in an ImageQuant-RT ECL (Amersham).

Chlorophyll Fluorescence Analysis

Chlorophyll fluorescence was measured using a MINI-PAM portable chlorophyll fluorometer (Waltz). The transient increase in chlorophyll fluorescence after turning off actinic light was monitored as previously described (Shikanai et al., 1998).

Immunoblot Analysis

Chloroplasts were isolated from the leaves of 4-week-old plants as previously described (Okuda et al., 2007). Samples were normalized by measuring chlorophyll concentration. The protein samples were separated by 12.5% SDS-PAGE. After electrophoresis, the proteins were transferred onto a Hybond-P membrane (GE Healthcare) and incubated with specific antibodies. The signals were detected using an ECL Advance Western Blotting Detection Kit (for NdhH; GE Healthcare) or an ECL Plus Western Blotting Detection Kit (for the others; GE Healthcare). The signals were visualized by a LAS1000 chemiluminescence analyzer (Fuji Film).

Bioinformatic Analysis

The 15 nucleotides sequences upstream of editing sites were aligned and clustered with ClustalW 1.83 (Thompson et al., 1994) using the default parameters. Consensuses for editing sites recognized by the same factor were calculated by hand and searched against the *Arabidopsis* plastid genome sequence using fuzznuc from the EMBOSS package (Rice et al., 2000).

Accession Numbers

The genes described in this article correspond to the following Arabidopsis Genome Initiative codes: At5g59200 (*OTP80*), At2g29760 (*OTP81*), At3g57430 (*OTP84*), At2g02980 (*OTP85*), and At3g63370 (*OTP86*). Accession information for T-DNA insertion lines is provided in Table 1. Editing sites are specified relative to the nucleotide sequence of the complete *Arabidopsis* chloroplast genome (GenBank accession number AP000423).

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure 1. High-Resolution Melting Screen of *otp80*, *otp81*, *otp84*, *otp85*, and *otp86* Mutants.

Supplemental Figure 2. Analysis of Subcellular Localization of the OTP80, OTP81, OTP84, OTP85, and OTP86 Proteins.

Supplemental Figure 3. Phenotype of *otp80*, *otp81*, *otp84*, *otp85*, *otp86*, and Wild-Type *Arabidopsis* Plants.

Supplemental Figure 4. In Vivo Analysis of Electron Transport Activity Using Light Intensity Dependence of Electron Transport Rate.

Supplemental Table 1. Oligonucleotides Used in This Study.

Supplemental Data Set 1. Text File of the Alignment Used for the Analysis Shown in Figure 5.

ACKNOWLEDGMENTS

This work is funded through the Australian Research Council Centres of Excellence scheme (CE0561495), an International Science Linkages grant (CG120098) from the Australian Government Department of Innovation, Industry, Science, and Research, and scholarships from the Region of Alsace and the University of Western Australia. K.H. is the holder of a Lavoisier Fellowship, and I.S. is a West Australian State Premier's Fellow. T.S. was supported by Grants-in-Aid (17GS0316 and 16085206) and by Genomics for Agricultural Innovation (GPN0008).

Received September 17, 2009; revised October 9, 2009; accepted October 30, 2009; published November 29, 2009.

REFERENCES

Auweter, S.D., Oberstrass, F.C., and Allain, F.H. (2006). Sequence-specific binding of single-stranded RNA: Is there a code for recognition? Nucleic Acids Res. **34:** 4943–4959.

Bentolila, S., Elliott, L.E., and Hanson, M.R. (2008). Genetic architecture of mitochondrial editing in *Arabidopsis thaliana.* Genetics **178:** 1693–1708.

Bock, R., Hermann, M., and Fuchs, M. (1997). Identification of critical nucleotide positions for plastid RNA editing site recognition. RNA **3:** 1194–1200.

Bock, R., Hermann, M., and Kössel, H. (1996). *In vivo* dissection of *cis*-acting determinants for plastid RNA editing. EMBO J. **15:** 5052–5059.

Bock, R., and Koop, H.U. (1997). Extraplastidic site-specific factors mediate RNA editing in chloroplasts. EMBO J. **16:** 3282–3288.

Cai, W., Ji, D., Peng, L., Guo, J., Ma, J., Zou, M., Lu, C., and Zhang, L. (2009). LPA66 is required for editing *psbF* chloroplast transcripts in *Arabidopsis.* Plant Physiol. **150:** 1260–1271.

Carrie, C., Kühn, K., Murcha, M.W., Duncan, O., Small, I.D., O'Toole, N., and Whelan, J. (2009). Approaches to defining dual-targeted proteins in *Arabidopsis.* Plant J. **57:** 1128–1139.

Carrie, C., Murcha, M.W., Millar, A.H., Smith, S.M., and Whelan, J. (2007). Nine 3-ketoacyl-CoA thiolases (KATs) and acetoacetyl-CoA thiolases (ACATs) encoded by five genes in *Arabidopsis thaliana* are targeted either to peroxisomes or cytosol but not to mitochondria. Plant Mol. Biol. **63:** 97–108.

Chateigner-Boutin, A.L., and Hanson, M.R. (2002). Cross-competition in transgenic chloroplasts expressing single editing sites reveals shared *cis* elements. Mol. Cell. Biol. **22:** 8448–8456.

Chateigner-Boutin, A.L., and Hanson, M.R. (2003). Developmental co-variation of RNA editing extent of plastid editing sites exhibiting similar *cis*-elements. Nucleic Acids Res. **31:** 2586–2594.

Chateigner-Boutin, A.L., Ramos-Vega, M., Guevara-García, A., Andrés, C., de la Luz Gutiérrez-Nava, M., Cantero, A., Delannoy,

E., Jiménez, L.F., Lurin, C., Small, I., and León, P. (2008). CLB19, a pentatricopeptide repeat protein required for editing of rpoA and clpP chloroplast transcripts. Plant J. **56:** 590–602.

Chateigner-Boutin, A.L., and Small, I. (2007). A rapid high-throughput method for the detection and quantification of RNA editing based on high-resolution melting of amplicons. Nucleic Acids Res. **35:** e114.

Chaudhuri, S., Carrer, H., and Maliga, P. (1995). Site-specific factor involved in the editing of the *psbL* mRNA in tobacco plastids. EMBO J. **14:** 2951–2957.

Chaudhuri, S., and Maliga, P. (1996). Sequences directing C to U editing of the plastid *psbL* mRNA are located within a 22 nucleotide segment spanning the editing site. EMBO J. **15:** 5958–5964.

Cummings, M.P., and Myers, D.S. (2004). Simple statistical models predict C-to-U edited sites in plant mitochondrial RNA. BMC Bioinformatics **5:** 132.

Delannoy, E., Stanley, W.A., Bond, C.S., and Small, I.D. (2007). Pentatricopeptide repeat (PPR) proteins as sequence-specificity factors in post-transcriptional processes in organelles. Biochem. Soc. Trans. **35:** 1643–1647.

Edwards, K., Johnstone, C., and Thompson, C. (1991). A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. Nucleic Acids Res. **19:** 1349.

Emanuelsson, O., Nielsen, H., Brunak, S., and von Heijne, G. (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J. Mol. Biol. **300:** 1005–1016.

Giegé, P., and Brennicke, A. (1999). RNA editing in *Arabidopsis* mitochondria effects 441 C to U changes in ORFs. Proc. Natl. Acad. Sci. USA **96:** 15324–15329.

Hashimoto, M., Endo, T., Peltier, G., Tasaka, M., and Shikanai, T. (2003). A nucleus-encoded factor, CRR2, is essential for the expression of chloroplast ndhB in Arabidopsis. Plant J. **36:** 541–549.

Hayes, M.L., Reed, M.L., Hegeman, C.E., and Hanson, M.R. (2006). Sequence elements critical for efficient RNA editing of a tobacco chloroplast transcript *in vivo* and *in vitro.* Nucleic Acids Res. **34:** 3742–3754.

Hegeman, C.E., Hayes, M.L., and Hanson, M.R. (2005). Substrate and cofactor requirements for RNA editing of chloroplast transcripts in *Arabidopsis in vitro.* Plant J. **42:** 124–132.

Hirose, T., and Sugiura, M. (2001). Involvement of a site-specific trans-acting factor and a common RNA-binding protein in the editing of chloroplast mRNAs: Development of a chloroplast *in vitro* RNA editing system. EMBO J. **20:** 1144–1152.

Kim, S.R., Yang, J.I., Moon, S., Ryu, C.H., An, K., Kim, K.M., Yim, J., and An, G. (2009). Rice OGR1 encodes a pentatricopeptide repeat-DYW protein and is essential for RNA editing in mitochondria. Plant J. **59:** 738–749.

Kobayashi, Y., Matsuo, M., Sakamoto, K., Wakasugi, T., Yamada, K., and Obokata, J. (2008). Two RNA editing sites with *cis*-acting elements of moderate sequence identity are recognized by an identical site-recognition protein in tobacco chloroplasts. Nucleic Acids Res. **36:** 311–318.

Kotera, E., Tasaka, M., and Shikanai, T. (2005). A pentatricopeptide repeat protein is essential for RNA editing in chloroplasts. Nature **433:** 326–330.

Lurin, C., et al. (2004). Genome-wide analysis of *Arabidopsis* pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. Plant Cell **16:** 2089–2103.

Miyamoto, T., Obokata, J., and Sugiura, M. (2002). Recognition of RNA editing sites is directed by unique proteins in chloroplasts: Biochemical identification of cis-acting elements and trans-acting factors involved in RNA editing in tobacco and pea chloroplasts. Mol. Cell. Biol. **22:** 6726–6734.

Miyamoto, T., Obokata, J., and Sugiura, M. (2004). A site-specific factor interacts directly with its cognate RNA editing site in chloroplast transcripts. Proc. Natl. Acad. Sci. USA **101:** 48–52.

Mulligan, R.M., Williams, M.A., and Shanahan, M.T. (1999). RNA editing site recognition in higher plant mitochondria. J. Hered. **90:** 338–344.

Murcha, M.W., Elhafez, D., Lister, R., Tonti-Filippini, J., Baumgartner, M., Philippar, K., Carrie, C., Mokranjac, D., Soll, J., and Whelan, J. (2007). Characterization of the preprotein and amino acid transporter gene family in *Arabidopsis.* Plant Physiol. **143:** 199–212.

Neuwirt, J., Takenaka, M., van der Merwe, J.A., and Brennicke, A. (2005). An *in vitro* RNA editing system from cauliflower mitochondria: editing site recognition parameters can vary in different plant species. RNA **11:** 1563–1570.

Okuda, K., Chateigner-Boutin, A.L., Nakamura, T., Delannoy, E., Sugita, M., Myouga, F., Motohashi, R., Shinozaki, K., Small, I., and Shikanai, T. (2009). Pentatricopeptide repeat proteins with the DYW motif have distinct molecular functions in RNA editing and RNA cleavage in *Arabidopsis* chloroplasts. Plant Cell **21:** 146–156.

Okuda, K., Myouga, F., Motohashi, R., Shinozaki, K., and Shikanai, T. (2007). Conserved domain structure of pentatricopeptide repeat proteins involved in chloroplast RNA editing. Proc. Natl. Acad. Sci. USA **104:** 8178–8183.

Okuda, K., Nakamura, T., Sugita, M., Shimizu, T., and Shikanai, T. (2006). A pentatricopeptide repeat protein is a site recognition factor in chloroplast RNA editing. J. Biol. Chem. **281:** 37661–37667.

O'Toole, N., Hattori, M., Andres, C., Iida, K., Lurin, C., Schmitz-Linneweber, C., Sugita, M., and Small, I. (2008). On the expansion of the pentatricopeptide repeat gene family in plants. Mol. Biol. Evol. **25:** 1120–1128.

Peng, L., Shimizu, H., and Shikanai, T. (2008). The chloroplast NAD(P)H dehydrogenase complex interacts with photosystem I in *Arabidopsis.* J. Biol. Chem. **283:** 34873–34879.

Pfalz, J., Bayraktar, O.A., Prikryl, J., and Barkan, A. (2009). Site-specific binding of a PPR protein defines and stabilizes 5′ and 3′ mRNA termini in chloroplasts. EMBO J. **28:** 2042–2052.

Reed, M.L., Peeters, N.M., and Hanson, M.R. (2001). A single alteration 20 nt 5′ to an editing target inhibits chloroplast RNA editing *in vivo.* Nucleic Acids Res. **29:** 1507–1513.

Rice, P., Longden, I., and Bleasby, A. (2000). EMBOSS: The European Molecular Biology Open Software Suite. Trends Genet. **16:** 276–277.

Robbins, J.C., Heller, W.P., and Hanson, M.R. (2009). A comparative genomics approach identifies a PPR-DYW protein that is essential for C-to-U editing of the *Arabidopsis* chloroplast *accD* transcript. RNA **15:** 1142–1153.

Rudinger, M., Polsakiewicz, M., and Knoop, V. (2008). Organellar RNA editing and plant-specific extensions of pentatricopeptide repeat proteins in jungermanniid but not in marchantiid liverworts. Mol. Biol. Evol. **25:** 1405–1414.

Salone, V., Rudinger, M., Polsakiewicz, M., Hoffmann, B., Groth-Malonek, M., Szurek, B., Small, I., Knoop, V., and Lurin, C. (2007). A hypothesis on the identification of the editing enzyme in plant organelles. FEBS Lett. **581:** 4132–4138.

Sasaki, T., Yukawa, Y., Wakasugi, T., Yamada, K., and Sugiura, M. (2006). A simple *in vitro* RNA editing assay for chloroplast transcripts using fluorescent dideoxynucleotides: Distinct types of sequence elements required for editing of *ndh* transcripts. Plant J. **47:** 802–810.

Schmitz-Linneweber, C., and Small, I. (2008). Pentatricopeptide repeat proteins: A socket set for organelle gene expression. Trends Plant Sci. **13:** 663–670.

Schmitz-Linneweber, C., Williams-Carrier, R., and Barkan, A. (2005). RNA immunoprecipitation and microarray analysis show a chloroplast pentatricopeptide repeat protein to be associated with the 5′ region of mRNAs whose translation it activates. Plant Cell **17:** 2791–2804.

Shikanai, T. (2006). RNA editing in plant organelles: Machinery, physiological function and evolution. Cell. Mol. Life Sci. **63:** 698–708.

Shikanai, T. (2007). Cyclic electron transport around photosystem I: Genetic approaches. Annu. Rev. Plant Biol. **58:** 199–217.

Shikanai, T., Endo, T., Hashimoto, T., Yamada, Y., Asada, K., and Yokota, A. (1998). Directed disruption of the tobacco *ndhB* gene impairs cyclic electron flow around photosystem I. Proc. Natl. Acad. Sci. USA **95:** 9705–9709.

Simpson, L., and Emeson, R.B. (1996). RNA editing. Annu. Rev. Neurosci. **19:** 27–52.

Small, I., Peeters, N., Legeai, F., and Lurin, C. (2004). Predotar: A tool for rapidly screening proteomes for N-terminal targeting sequences. Proteomics **4:** 1581–1590.

Small, I.D., and Peeters, N. (2000). The PPR motif - A TPR-related motif prevalent in plant organellar proteins. Trends Biochem. Sci. **25:** 46–47.

Swiatek, M., Kuras, R., Sokolenko, A., Higgs, D., Olive, J., Cinque, G., Müller, B., Eichacker, L.A., Stern, D.B., Bassi, R., Herrmann, R. G., and Wollman, F.A. (2001). The chloroplast gene *ycf9* encodes a photosystem II (PSII) core subunit, PsbZ, that participates in PSII supramolecular architecture. Plant Cell **13:** 1347–1367.

Takenaka, M., Neuwirt, J., and Brennicke, A. (2004). Complex *cis*-elements determine an RNA editing site in pea mitochondria. Nucleic Acids Res. **32:** 4137–4144.

Takenaka, M., Verbitskiy, D., van der Merwe, J.A., Zehrmann, A., and Brennicke, A. (2008). The process of RNA editing in plant mitochondria. Mitochondrion **8:** 35–46.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:** 4673–4680.

Tillich, M., Le Sy, V., Schulerowitz, K., von Haeseler, A., Maier, U.G., and Schmitz-Linneweber, C. (2009). Loss of matK RNA editing in seed plant chloroplasts. BMC Evol. Biol. **9:** 201.

Verbitskiy, D., van der Merwe, J.A., Zehrmann, A., Brennicke, A., and Takenaka, M. (2008). Multiple specificity recognition motifs enhance plant mitochondrial RNA editing in vitro. J. Biol. Chem. **283:** 24374–24381.

Yu, Q.B., Jiang, Y., Chong, K., and Yang, Z.N. (2009). AtECB2, a pentatricopeptide repeat protein, is required for chloroplast transcript *accD* RNA editing and early chloroplast biogenesis in *Arabidopsis thaliana.* Plant J. **59:** 1011–1023.

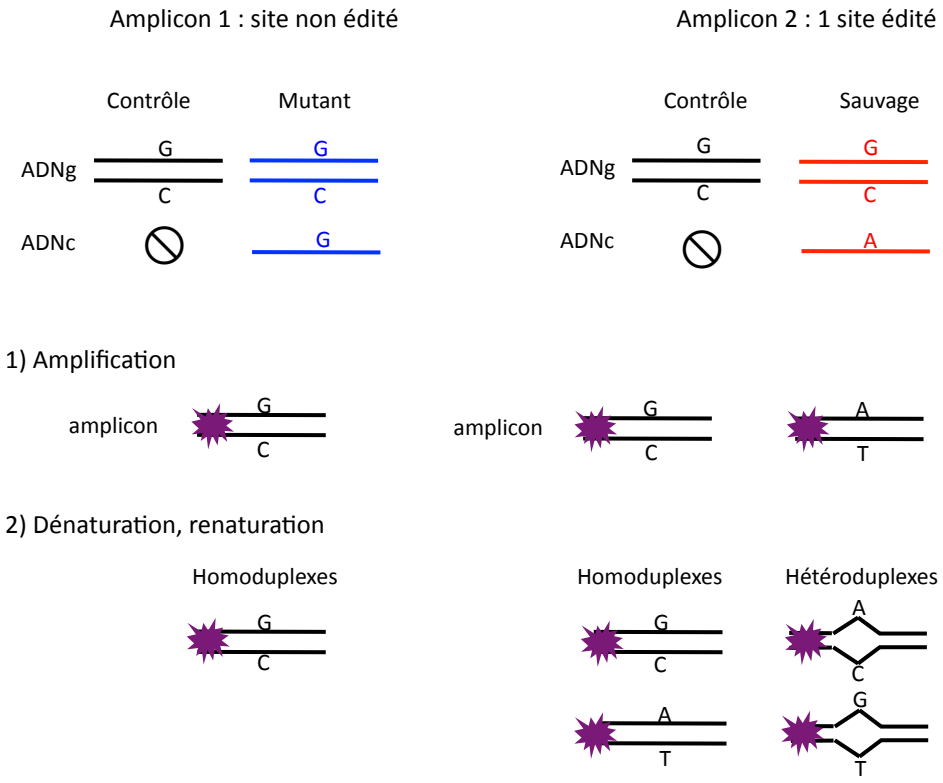
Zehrmann, A., van der Merwe, J.A., Verbitskiy, D., Brennicke, A., and Takenaka, M. (2008). Seven large variations in the extent of RNA editing in plant mitochondria between three ecotypes of *Arabidopsis thaliana*. Mitochondrion **8:** 319–327.

Zehrmann, A., Verbitskiy, D., van der Merwe, J.A., Brennicke, A., and Takenaka, M. (2009). A DYW domain-containing pentatricopeptide repeat protein is required for RNA editing at multiple sites in mitochondria of *Arabidopsis thaliana.* Plant Cell **21:** 558–567.

Zhou, W., Cheng, Y., Yap, A., Chateigner-Boutin, A.L., Delannoy, E., Hammani, K., Small, I., and Huang, J. (2009). The *Arabidopsis* gene *YS1* encoding a DYW protein is required for editing of *rpoB* transcripts and the rapid development of chloroplasts during early growth. Plant J. **58:** 82–96.

Annexe 1

Principe de la méthode de HRM
(modifié d'après Chateigner-Boutin et Small, NAR, 2007)

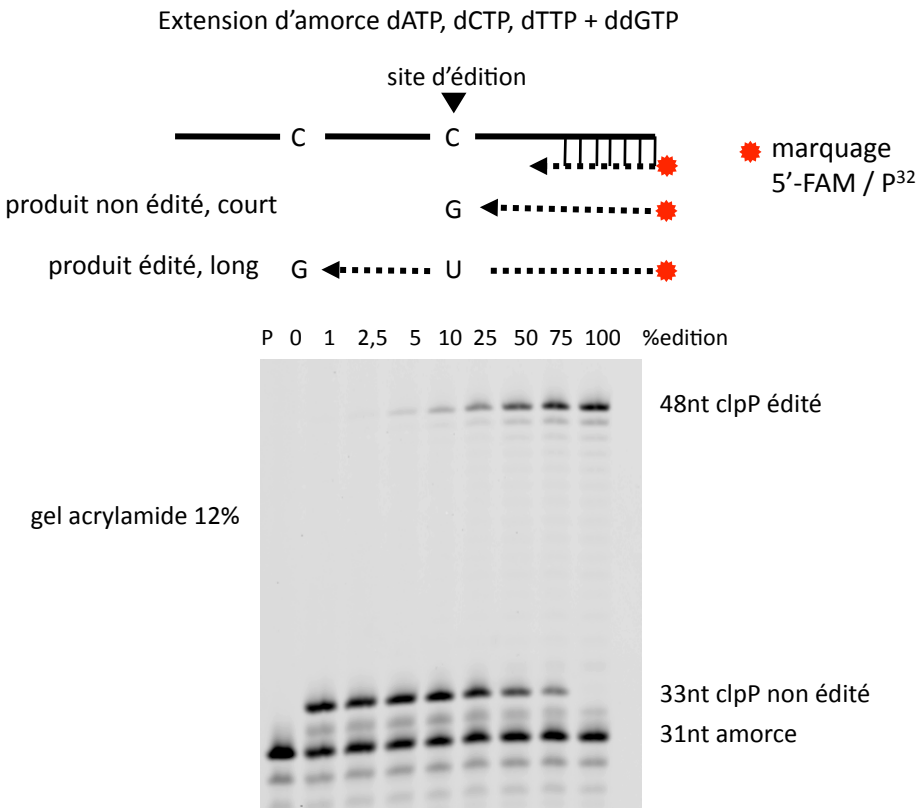


Principe de la méthode de HRM (pour High Resolution Melting) afin de cribler un défaut d'édition de l'ARN.

Les ADNc d'une plante mutante et sauvage sont synthétisés à partir d'ARN total. Ces ADNc sont mélangés indépendamment avec de l'ADN génomique sauvage (ADNg) et seront comparés à une expérience contrôle (ADNg seul). Un court produit de PCR (amplicon) est amplifié en utilisant des amorces entourant le site d'édition. Les produits de PCR sont par la suite dénaturés et renaturés. Si le site d'édition n'est pas édité (cas d'un mutant), un seul type d'amplicon (G/C) est formé. Ce sont des homoduplexes. Dans le cas où le site d'édition est édité le résidu C est converti en U au niveau de l'ARN traduit par la présence de A au niveau de l'ADNc. Après les étapes d'amplification, dénaturation, renaturation, deux populations de molécules se forment: les homoduplexes correspondant à l'appariement des brins complémentaires non édités ou édités et des hétéroduplexes correspondant à l'appariement d'un brin édité à un brin non édité. Les hétéroduplexes qui possèdent un mésappariement ont une température de fusion (Tm) plus faible que les homoduplexes. Un marqueur fluorescent fixant l'ADN double brun (étoile violette) permet de mesurer ces différences de Tm à l'aide d'une machine de PCR en temps réel développée par Roche®. L'édition du site d'édition entraine la détection de deux pics de fluorescence correspondant aux deux Tm dont l'un disparaît chez la plante mutante. Lors de la fusion des amplicons une partie du marqueur fluorescent est libérée entrainant une diminution de la fluorescence. La présence ou non d'hétéroduplexes entraine un changement de la forme de la courbe.

Annexe 2

Principe de la méthode de Poisoned Primer Extension (PPE)

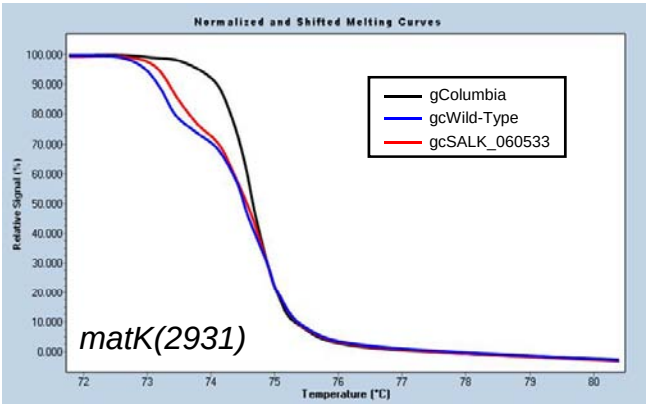


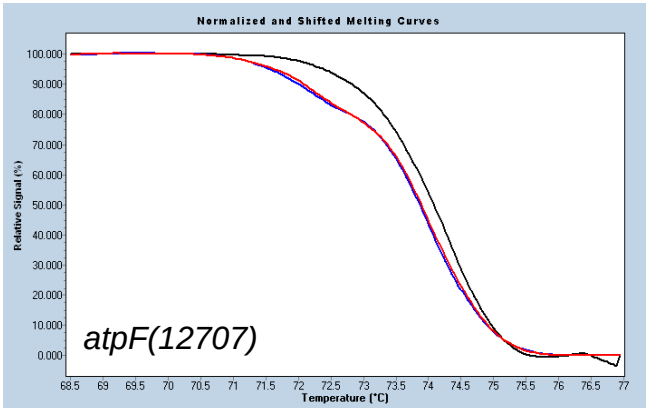
Principe de la méthode de PPE afin de confirmer et de quantifier un défaut d'édition de l'ARN.

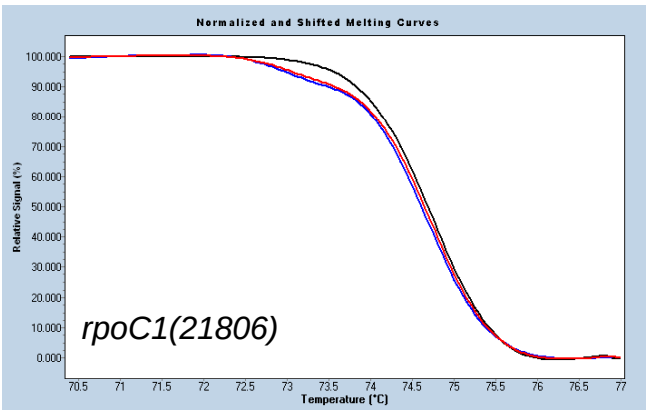
Une extension 5'-3' (PCR asymétrique) est réalisée sur l'ARN total d'une plante mutante ou sauvage en présence, d'une amorce marquée (fluorescente ou radioactive) qui s'apparie en amont du site d'édition, de trois dNTPs (ici dATP, dCTP, dTTP) et d'un ddNTP (ddGTP). Le didésoxyribonucléotide stoppera (empoisonnera) la réaction d'amplification lors de son incorporation dans le nouveau brin synthétisé. Dans notre exemple, si le site d'édition n'est pas édité, celui-ci conduira à l'incorporation d'un ddGTP au niveau du site d'édition lors de l'amplification et permettra la synthèse d'un court fragment. Si le site est édité, la cytidine du site d'édition sera convertie en uridine au niveau de l'ARN permettant une amplification du brin jusqu'au prochain résidu cytidine en aval du site d'édition générant un fragment plus long. Après l'étape d'amplification, les produits nouvellement synthétisés et marqués sont séparés sur gel d'acrylamide concentré permettant après révélation de visualiser et quantifier l'intensité des bandes correspondant aux fragments édité ou non. Le rapport des intensités des fragments édité et non édité détermine le pourcentage d'édition. Dans notre exemple, le gel représente les résultats obtenus sur le site d'édition du transcrit chloroplastique *clpP* à partir de différents échantillons de pourcentage d'édition croissant pour le site. Cette méthode très sensible permet de détecter une variation de l'ordre du pourcent. La PCR asymétrique peut également être réalisée sur un produit de RT-PCR du site d'édition.

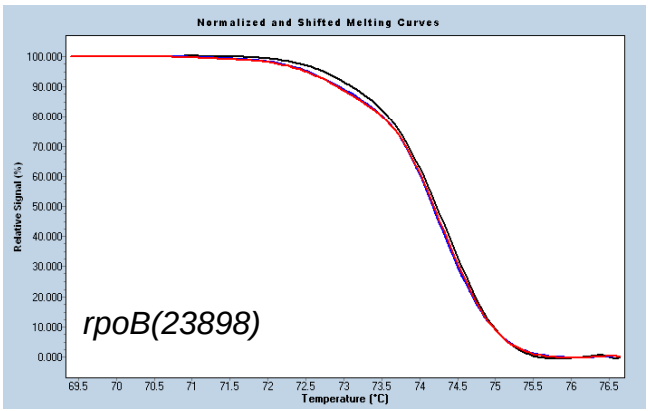
Supplemental Data, Hammani et al., (2009). **A study of new *Arabidopsis* chloroplast RNA editing mutants reveals general features of editing factors and their target sites.**

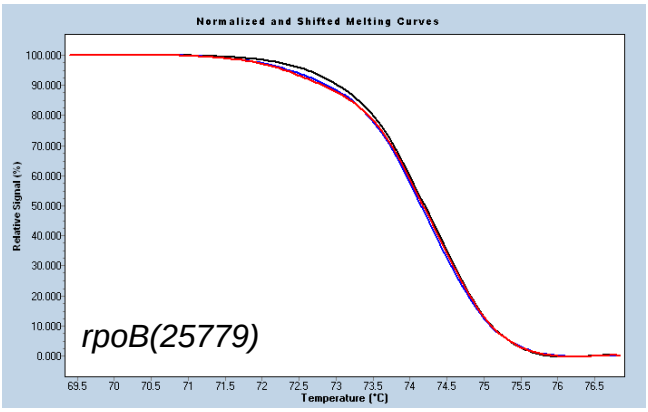
otp80

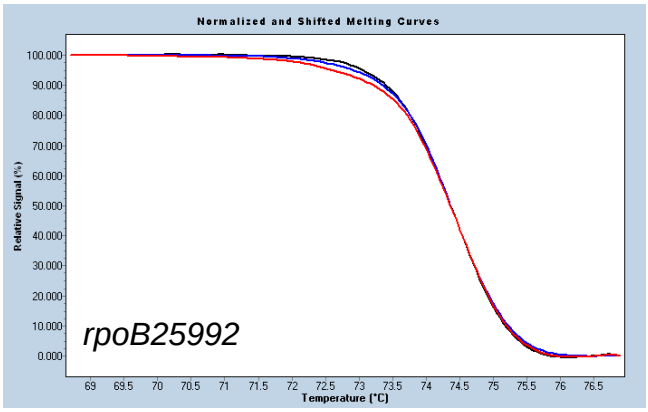


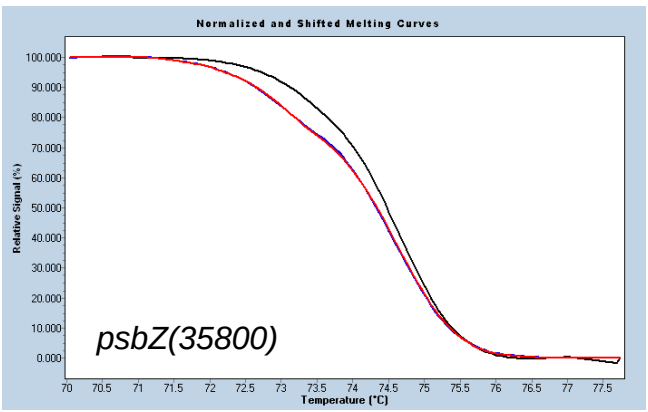


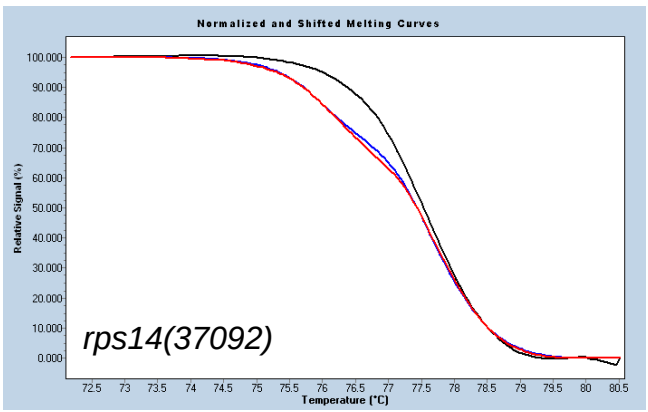




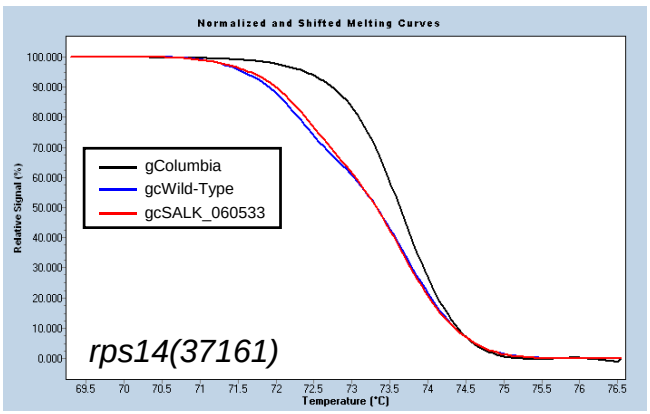


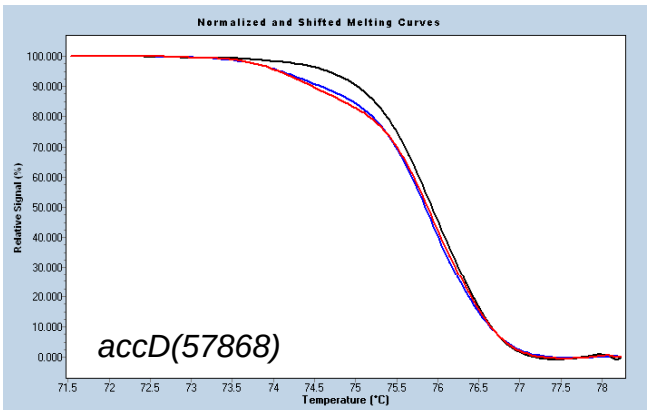


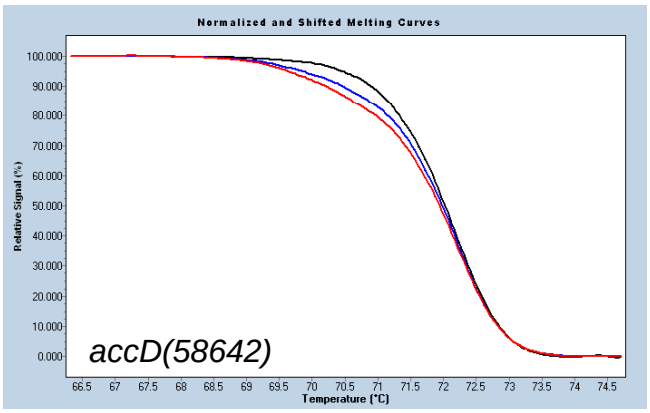


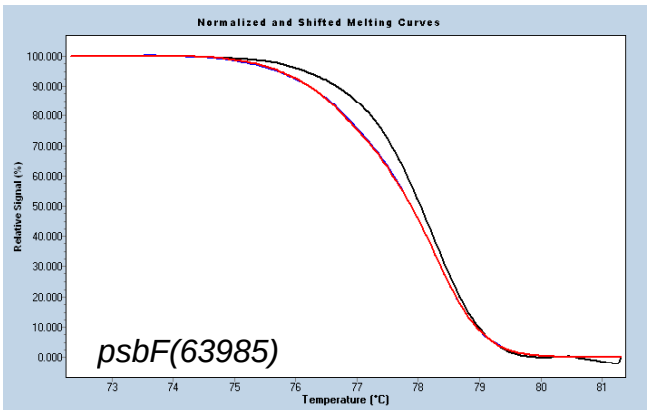


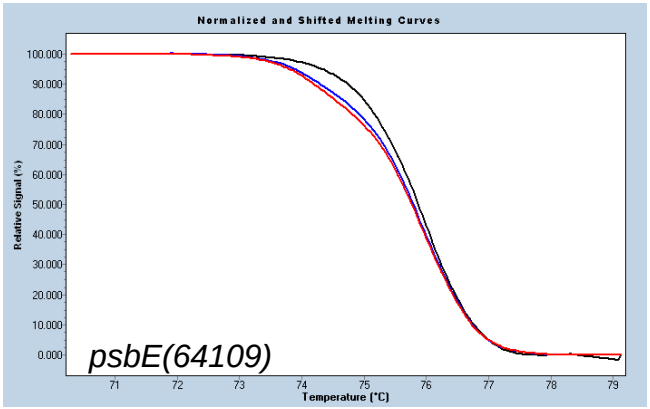
otp80

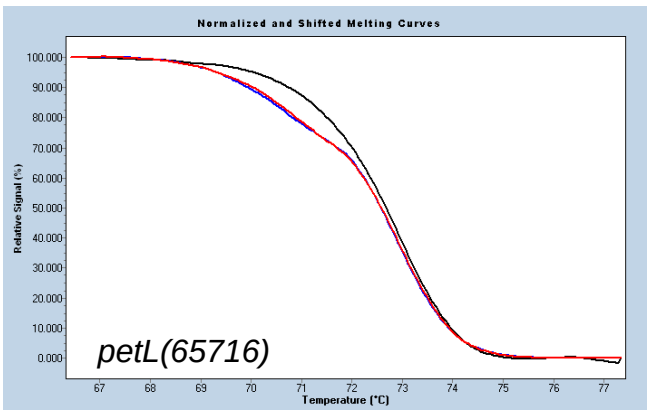

Normalized and Shifted Melting Curves
gColumbia
gcWild-Type
gcSALK_060533
Relative Signal (%)
Temperature (°C)
rps14(37161)

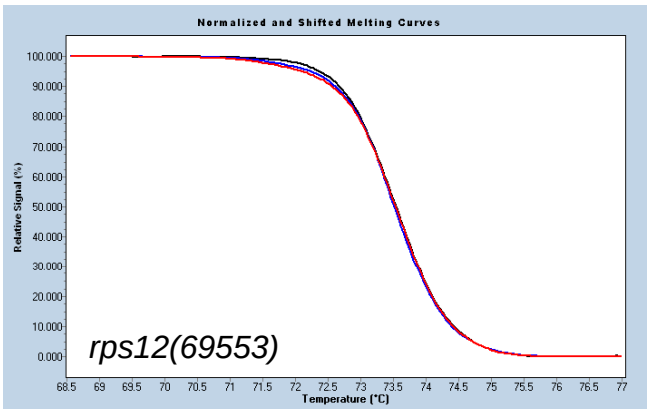

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
accD(57868)

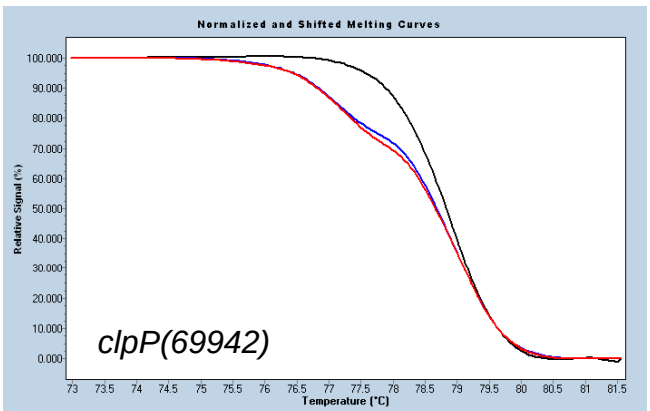

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
accD(58642)


Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
psbF(63985)

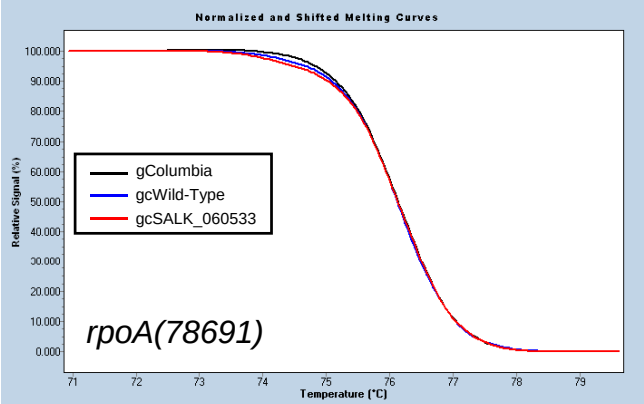

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
psbE(64109)

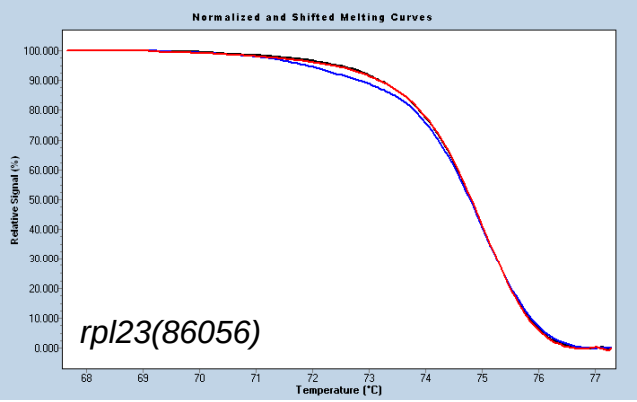

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
petL(65716)

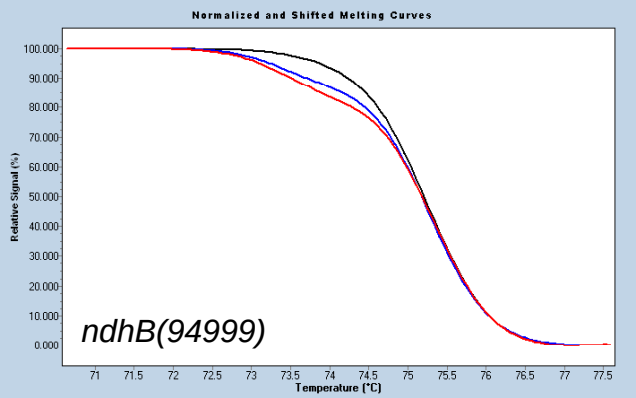

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
rps12(69553)

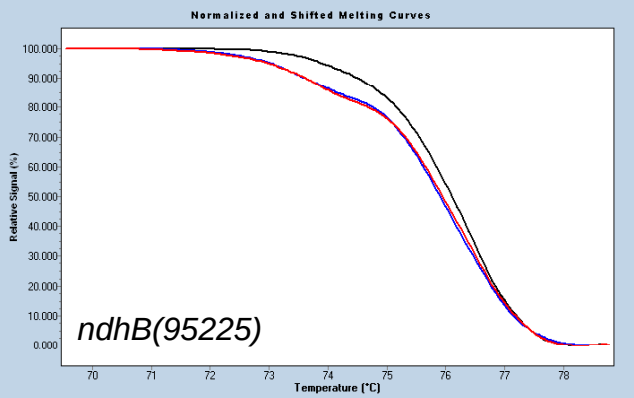

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
clpP(69942)

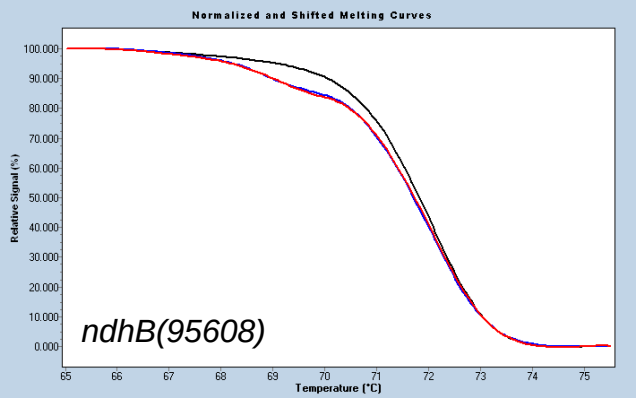
otp80

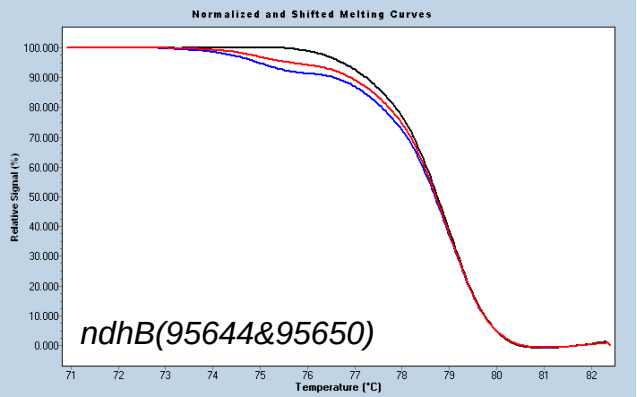

Normalized and Shifted Melting Curves
gColumbia
gcWild-Type
gcSALK_060533
rpoA(78691)
Relative Signal (%)
Temperature (°C)

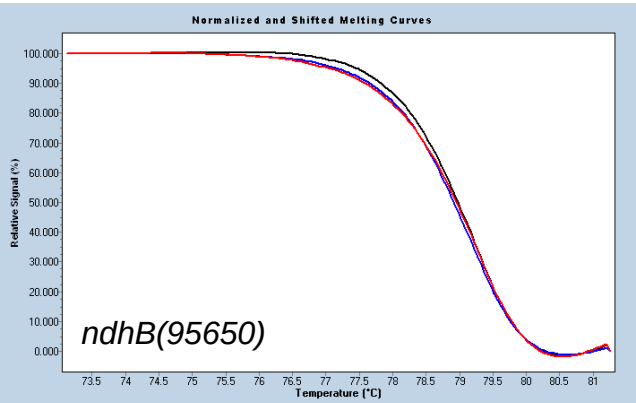

Normalized and Shifted Melting Curves
rpl23(86056)
Relative Signal (%)
Temperature (°C)

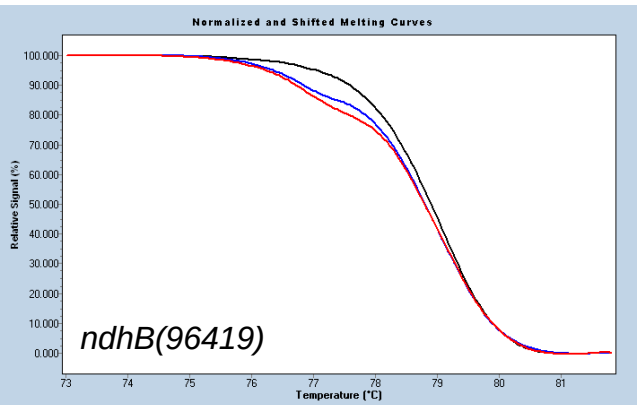

Normalized and Shifted Melting Curves
ndhB(94999)
Relative Signal (%)
Temperature (°C)


Normalized and Shifted Melting Curves
ndhB(95225)
Relative Signal (%)
Temperature (°C)

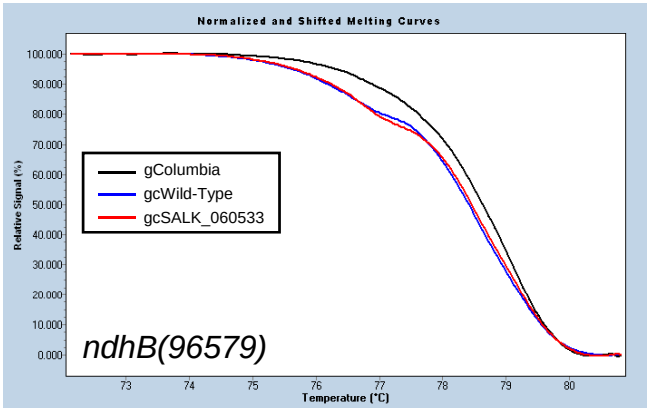

Normalized and Shifted Melting Curves
ndhB(95608)
Relative Signal (%)
Temperature (°C)

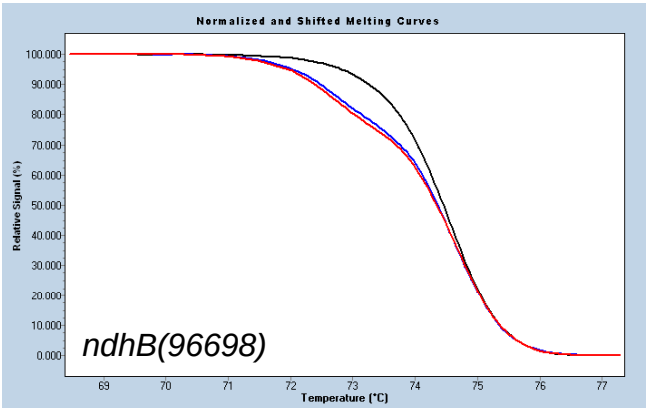

Normalized and Shifted Melting Curves
ndhB(95644&95650)
Relative Signal (%)
Temperature (°C)

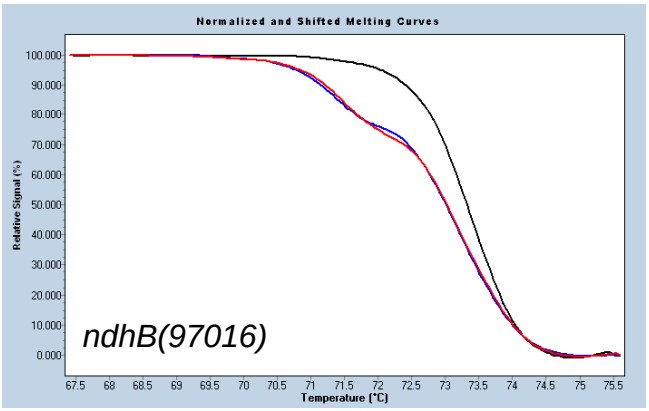

Normalized and Shifted Melting Curves
ndhB(95650)
Relative Signal (%)
Temperature (°C)

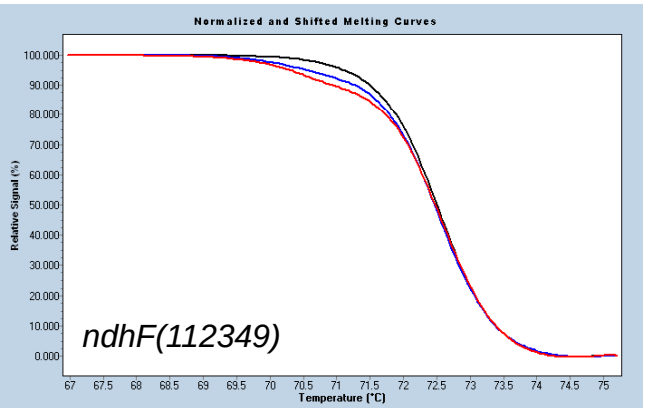

Normalized and Shifted Melting Curves
ndhB(96419)
Relative Signal (%)
Temperature (°C)

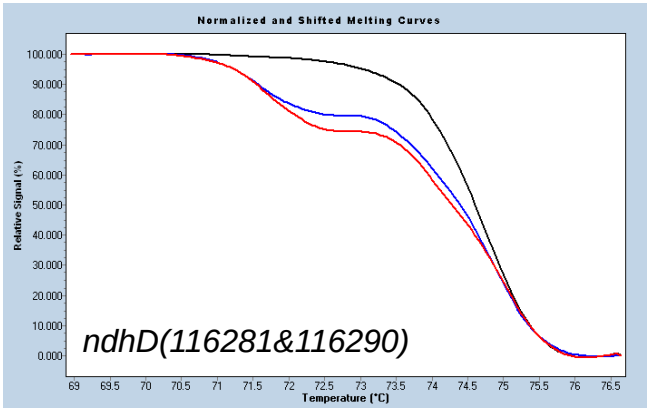
otp80

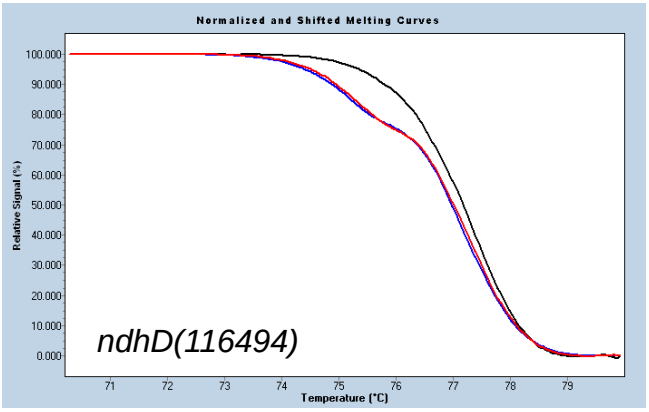

Normalized and Shifted Melting Curves
gColumbia
gcWild-Type
gcSALK_060533
Relative Signal (%)
Temperature (°C)
ndhB(96579)

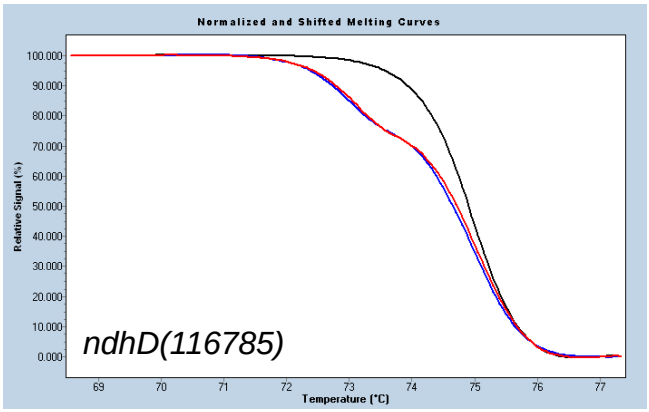

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhB(96698)

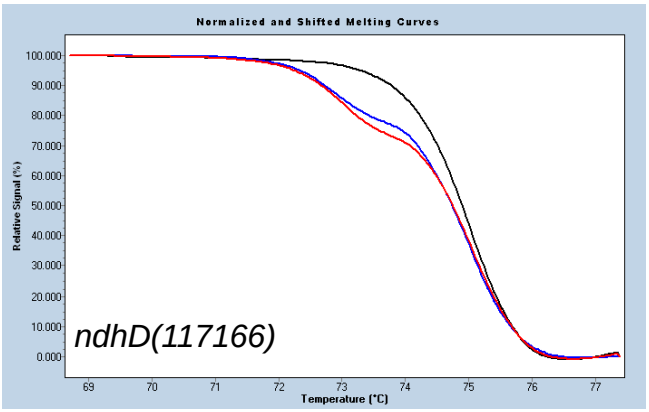

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhB(97016)


Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhF(112349)

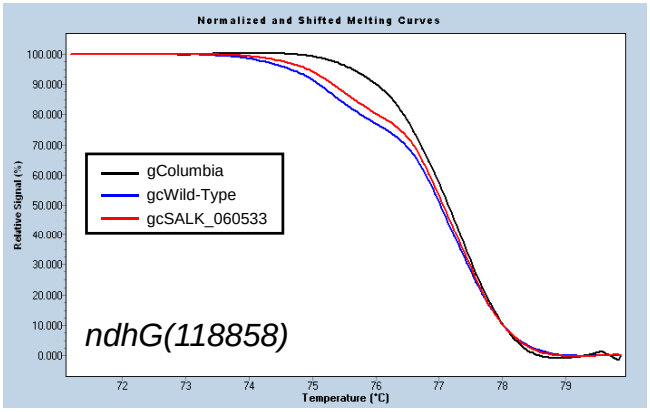

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhD(116281&116290)


Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhD(116494)

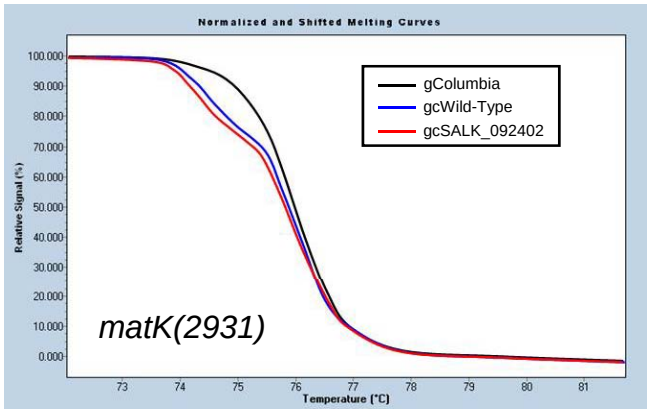

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhD(116785)

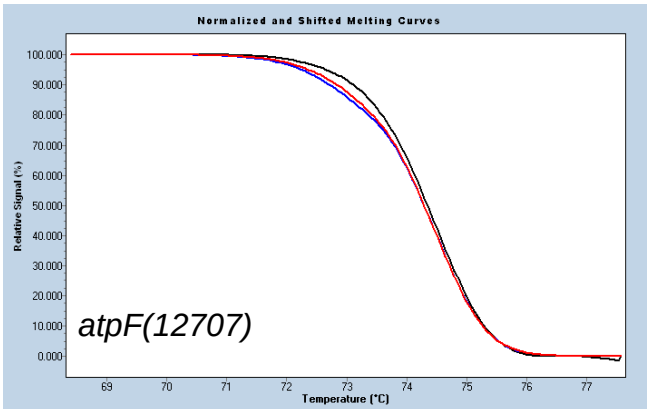

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhD(117166)

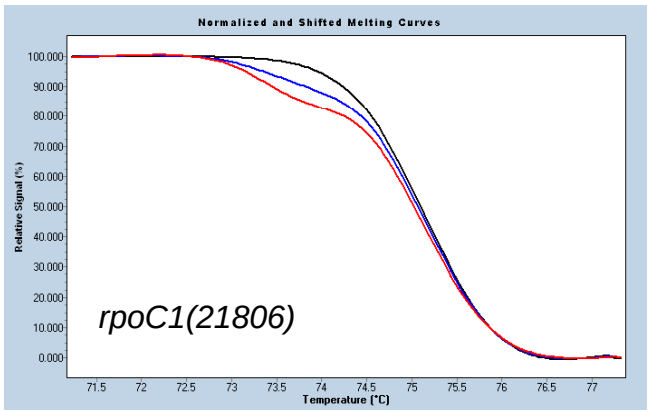
otp80

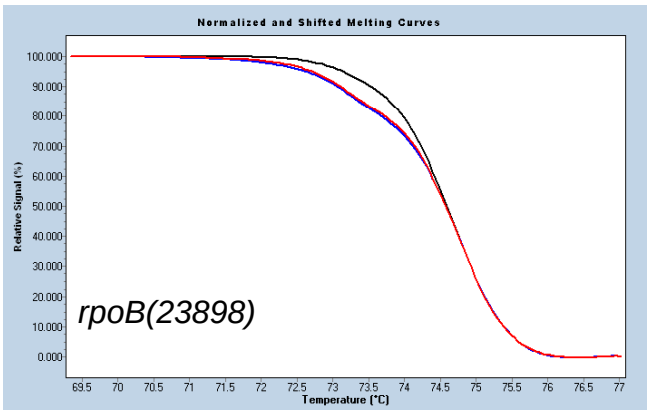

Normalized and Shifted Melting Curves
Relative Signal (%)
100.000
90.000
80.000
70.000
60.000
50.000
40.000
30.000
20.000
10.000
0.000
72
73
74
75
76
77
78
79
Temperature (°C)
gColumbia
gcWild-Type
gcSALK_060533
ndhG(118858)

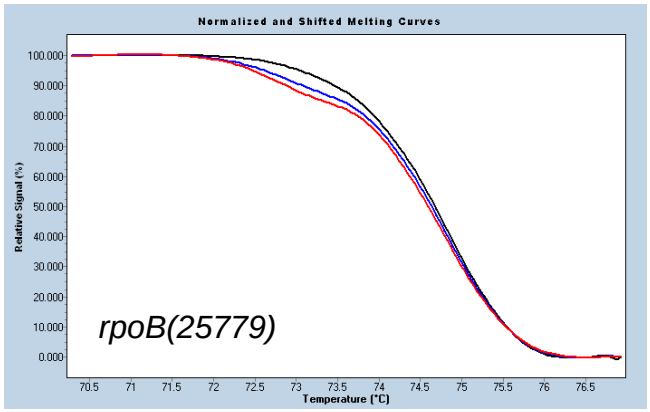
otp81

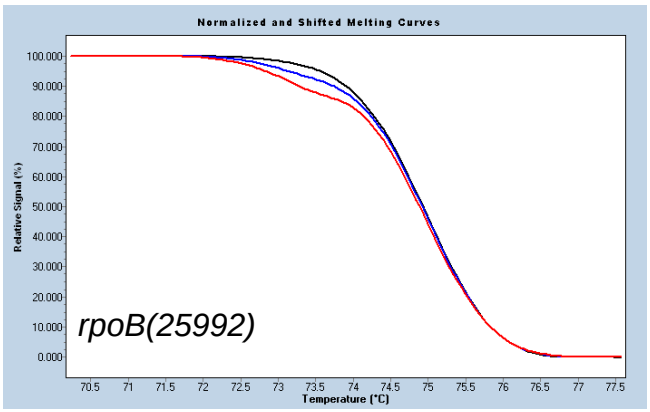

Normalized and Shifted Melting Curves
gColumbia
gcWild-Type
gcSALK_092402
Relative Signal (%)
Temperature (°C)
matK(2931)

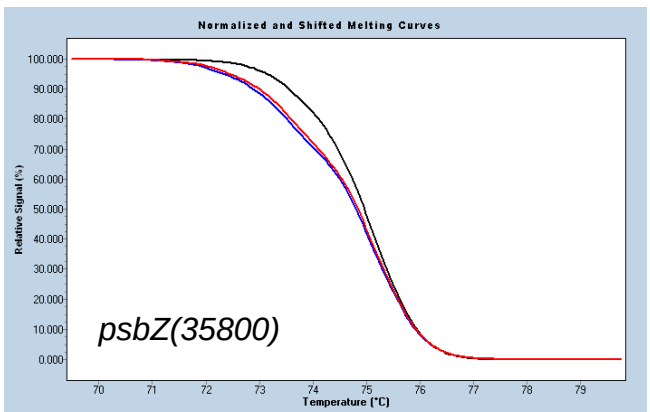

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
atpF(12707)

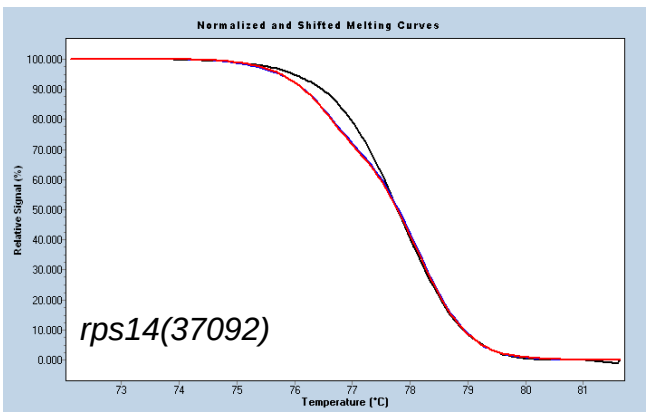

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
rpoC1(21806)


Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
rpoB(23898)

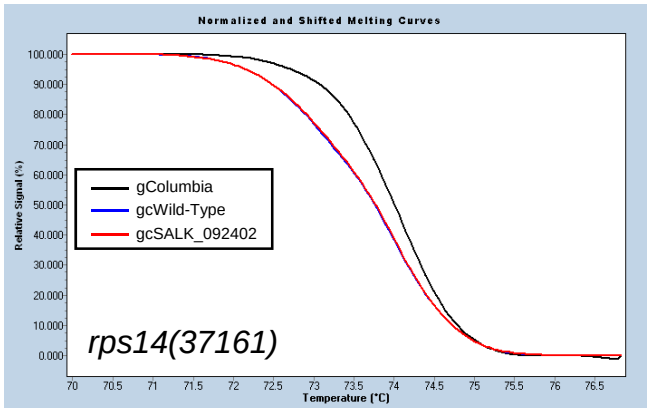

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
rpoB(25779)

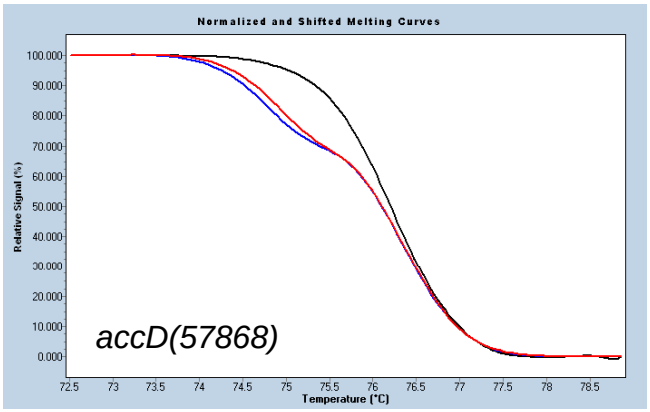

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
rpoB(25992)

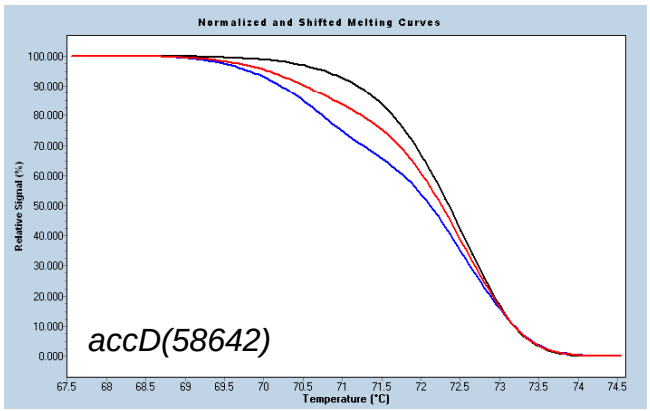

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
psbZ(35800)

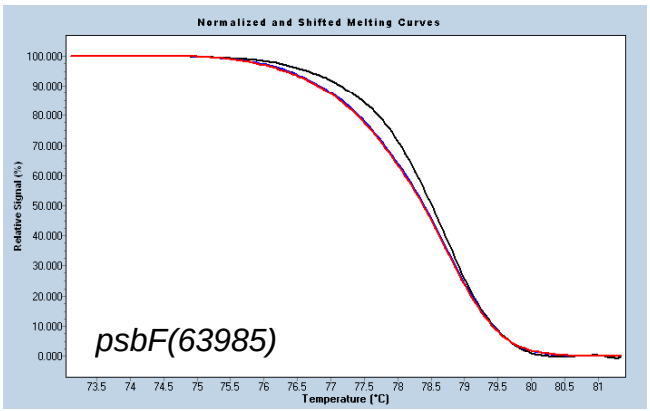

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
rps14(37092)

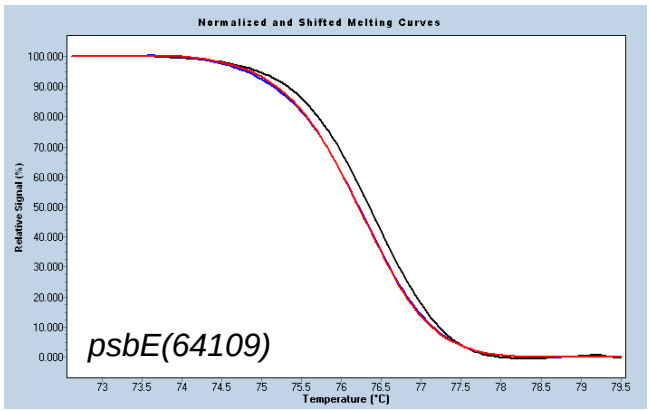
otp81

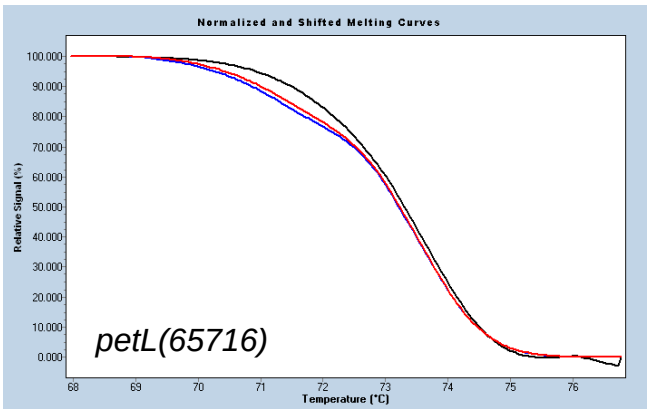

Normalized and Shifted Melting Curves
gColumbia
gcWild-Type
gcSALK_092402
rps14(37161)
Relative Signal (%)
Temperature (°C)

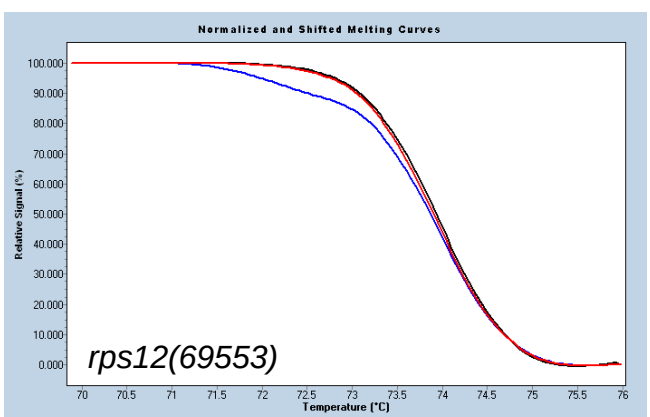

Normalized and Shifted Melting Curves
accD(57868)
Relative Signal (%)
Temperature (°C)

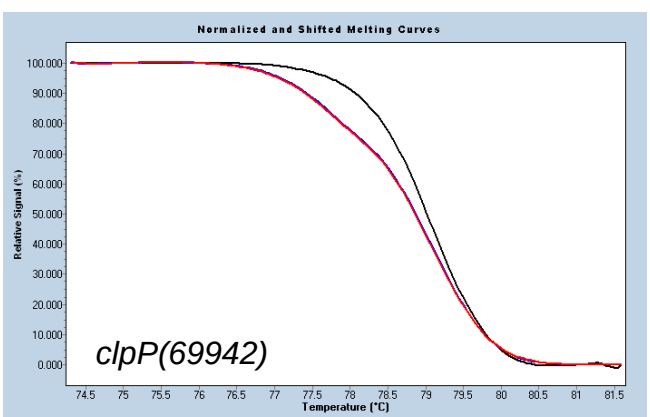

Normalized and Shifted Melting Curves
accD(58642)
Relative Signal (%)
Temperature (°C)


Normalized and Shifted Melting Curves
psbF(63985)
Relative Signal (%)
Temperature (°C)

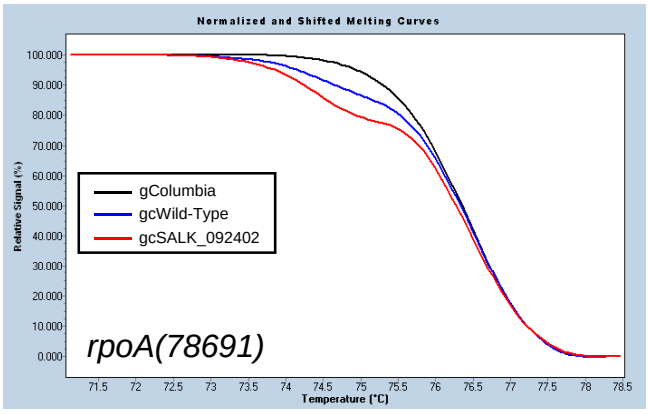

Normalized and Shifted Melting Curves
psbE(64109)
Relative Signal (%)
Temperature (°C)

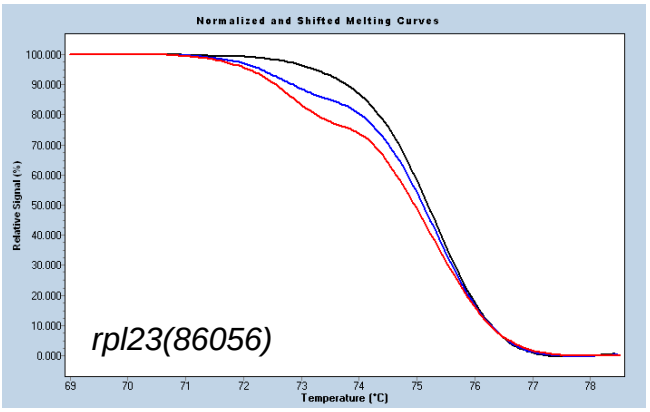

Normalized and Shifted Melting Curves
petL(65716)
Relative Signal (%)
Temperature (°C)

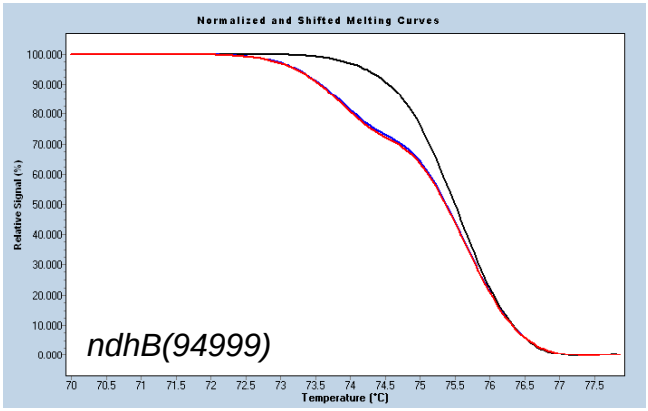

Normalized and Shifted Melting Curves
rps12(69553)
Relative Signal (%)
Temperature (°C)

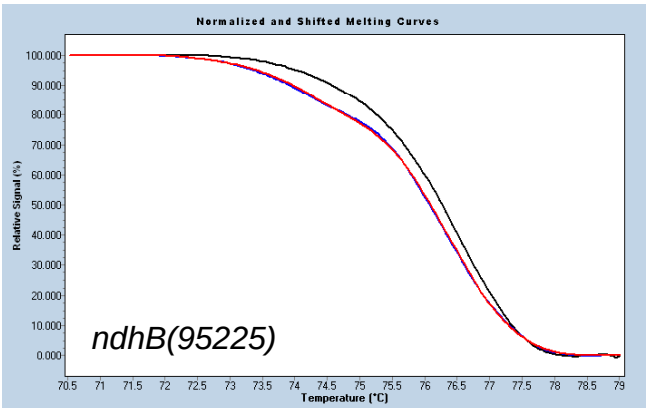

Normalized and Shifted Melting Curves
clpP(69942)
Relative Signal (%)
Temperature (°C)

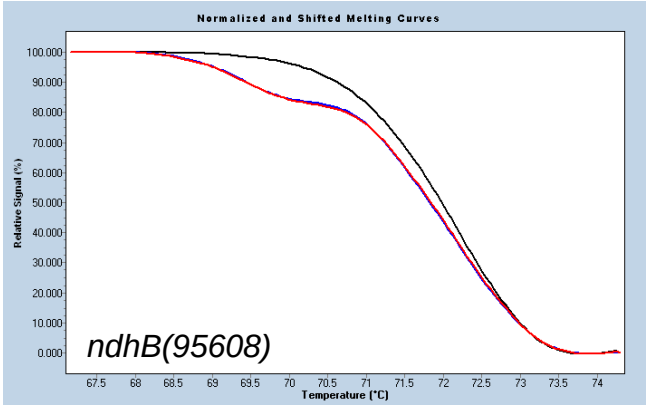
otp81

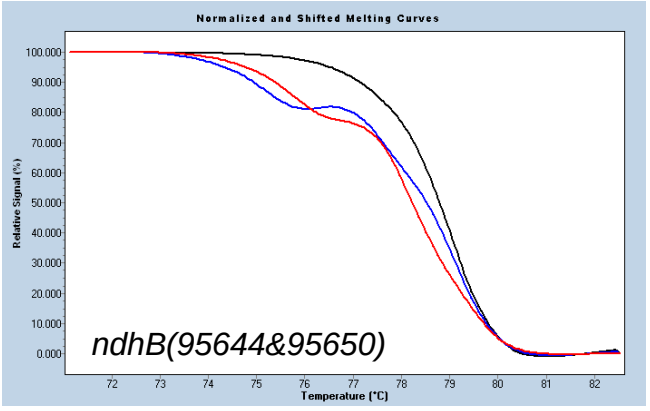

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
gColumbia
gcWild-Type
gcSALK_092402
rpoA(78691)

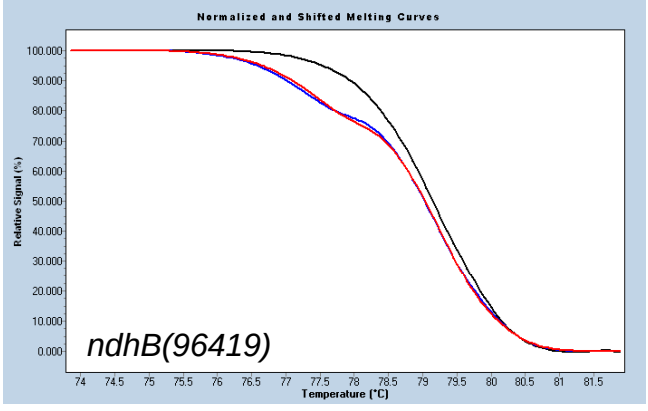

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
rpl23(86056)

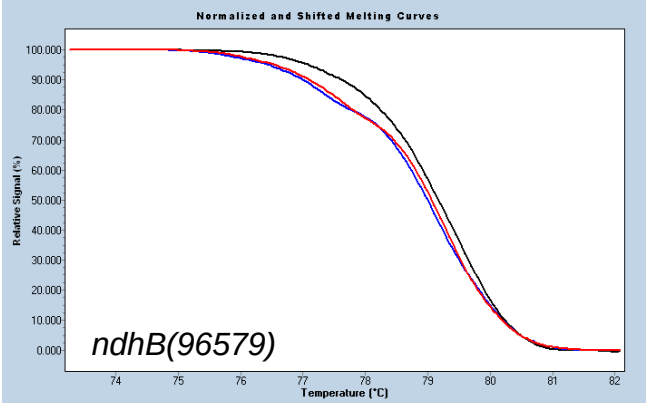

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhB(94999)


Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhB(95225)

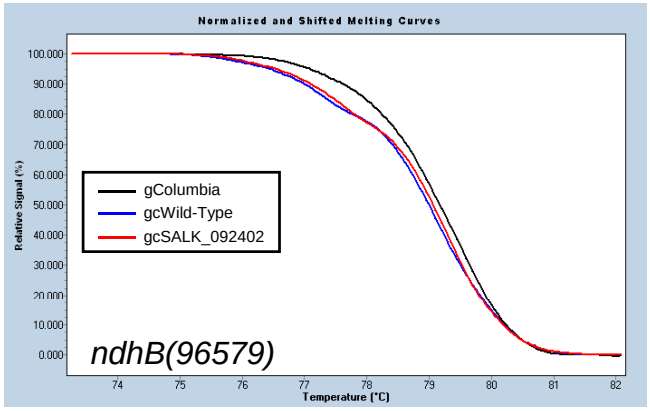

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhB(95608)

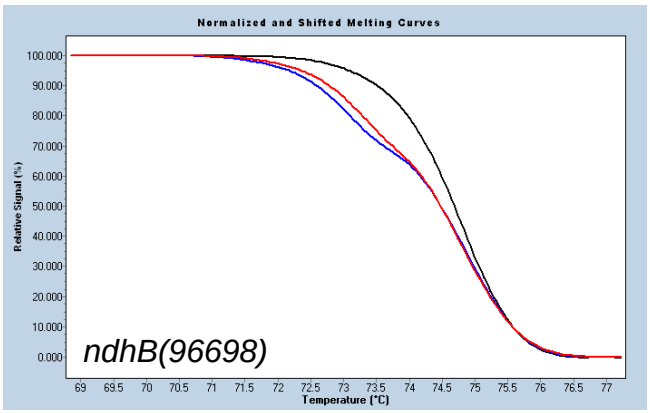

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhB(95644&95650)

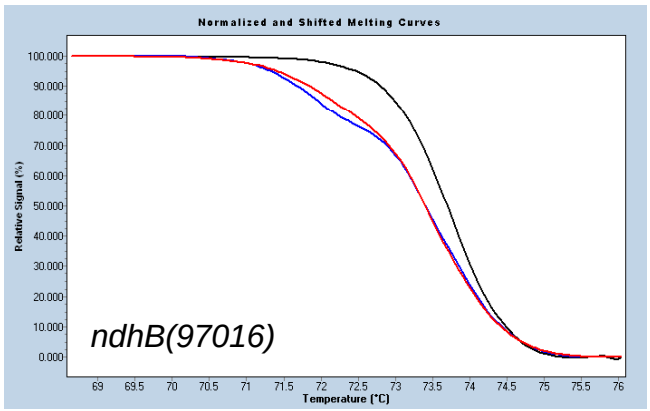

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhB(96419)

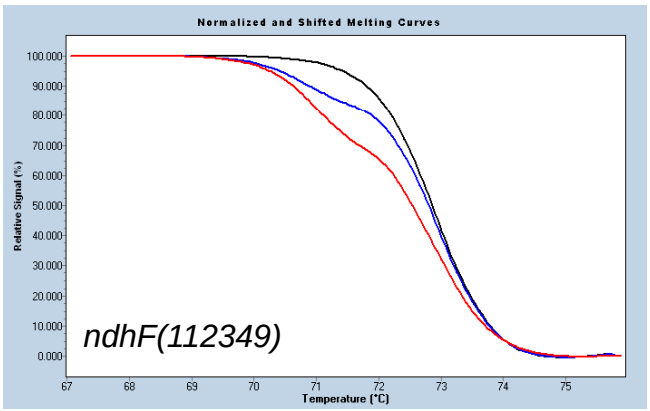

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhB(96579)

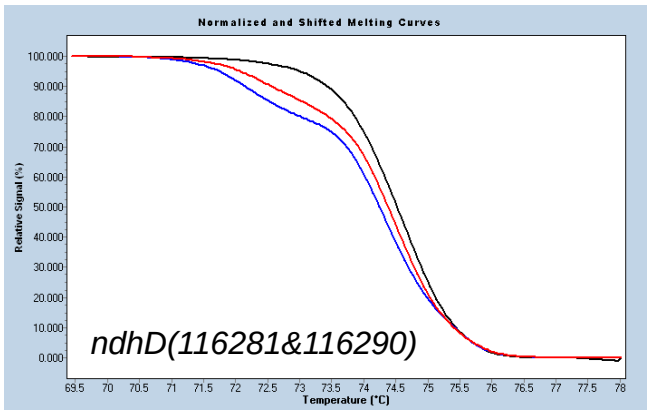
otp81

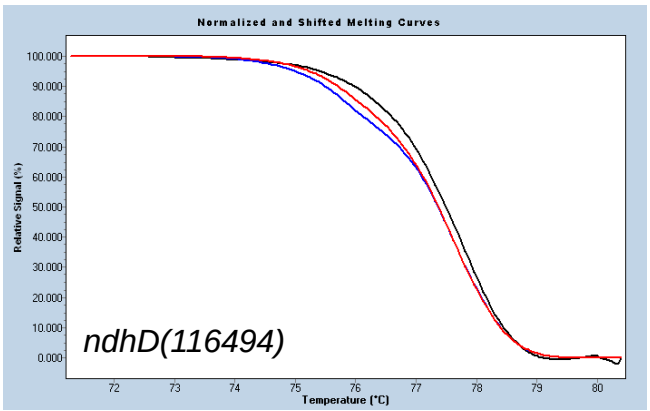

Normalized and Shifted Melting Curves
gColumbia
gcWild-Type
gcSALK_092402
ndhB(96579)
Relative Signal (%)
Temperature (°C)

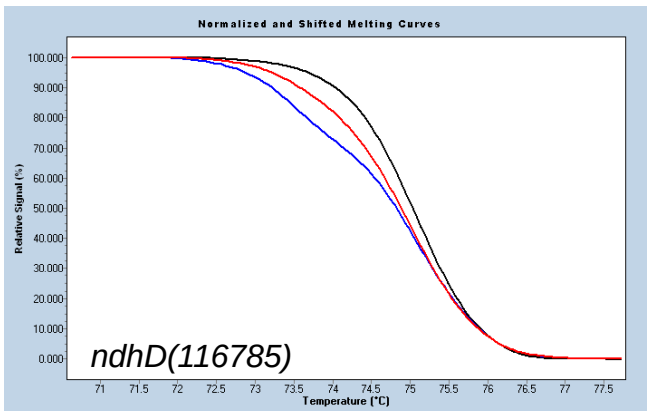

Normalized and Shifted Melting Curves
ndhB(96698)
Relative Signal (%)
Temperature (°C)

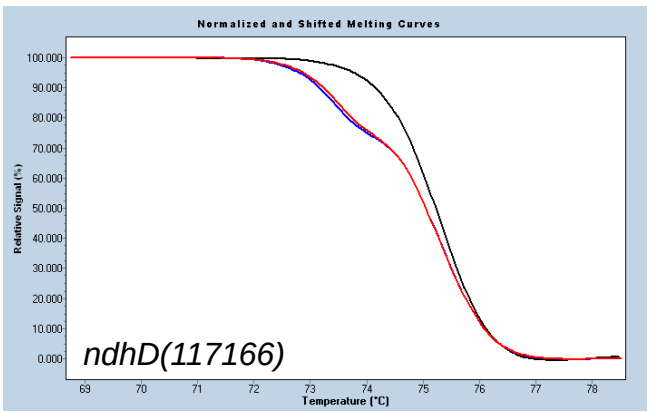

Normalized and Shifted Melting Curves
ndhB(97016)
Relative Signal (%)
Temperature (°C)


Normalized and Shifted Melting Curves
ndhF(112349)
Relative Signal (%)
Temperature (°C)

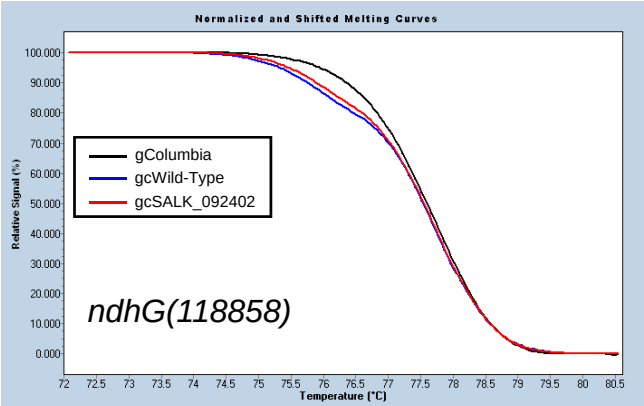

Normalized and Shifted Melting Curves
ndhD(116281&116290)
Relative Signal (%)
Temperature (°C)


Normalized and Shifted Melting Curves
ndhD(116494)
Relative Signal (%)
Temperature (°C)

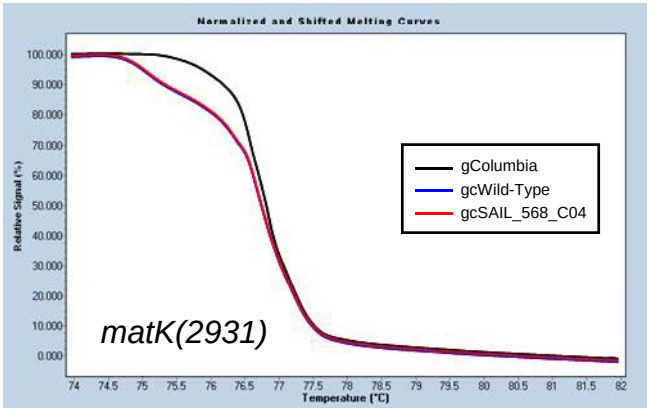

Normalized and Shifted Melting Curves
ndhD(116785)
Relative Signal (%)
Temperature (°C)

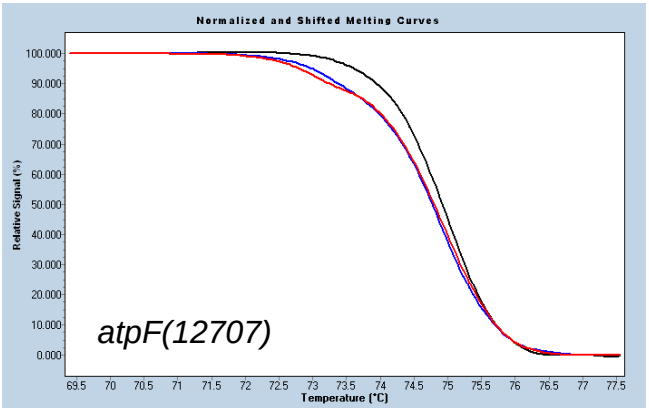

Normalized and Shifted Melting Curves
ndhD(117166)
Relative Signal (%)
Temperature (°C)

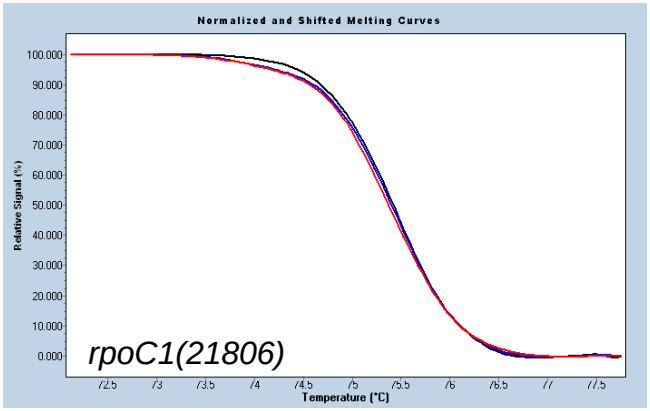
otp81

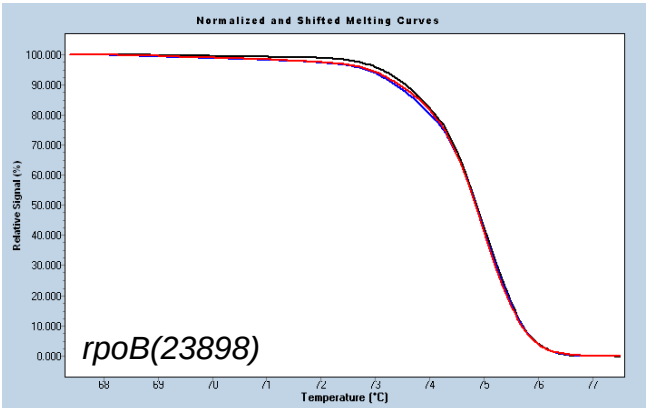

Normalized and Shifted Melting Curves
Relative Signal (%)
100.000
90.000
80.000
70.000
60.000
50.000
40.000
30.000
20.000
10.000
0.000
gColumbia
gcWild-Type
gcSALK_092402
ndhG(118858)
72
72.5
73
73.5
74
74.5
75
75.5
76
76.5
77
77.5
78
78.5
79
79.5
80
80.5
Temperature (°C)

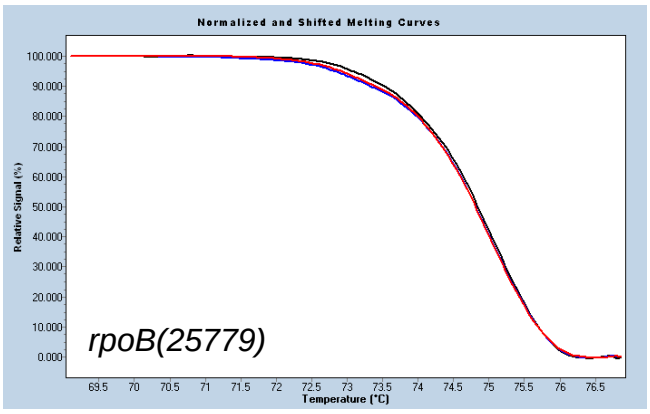
otp84

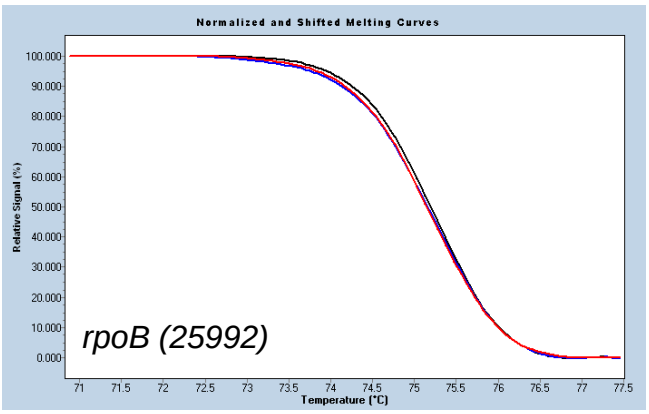

Normalized and Shifted Melting Curves
gColumbia
gcWild-Type
gcSAIL_568_C04
matK(2931)
Relative Signal (%)
Temperature (°C)

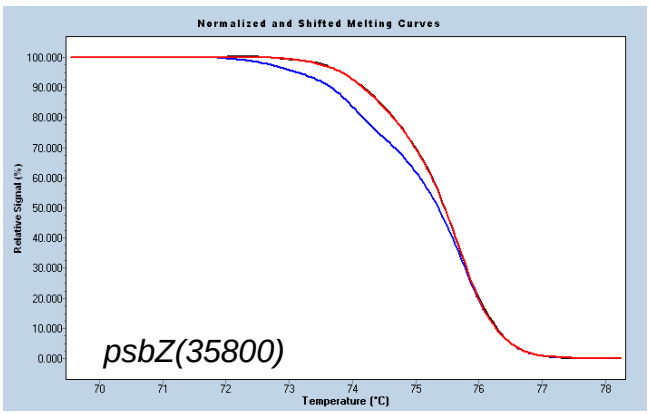

Normalized and Shifted Melting Curves
atpF(12707)
Relative Signal (%)
Temperature (°C)

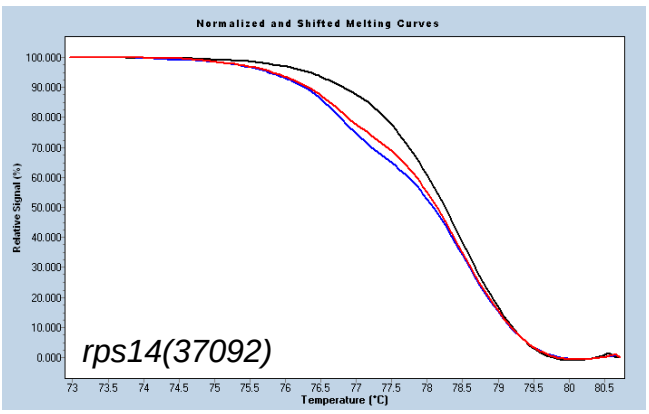

Normalized and Shifted Melting Curves
rpoC1(21806)
Relative Signal (%)
Temperature (°C)


Normalized and Shifted Melting Curves
rpoB(23898)
Relative Signal (%)
Temperature (°C)

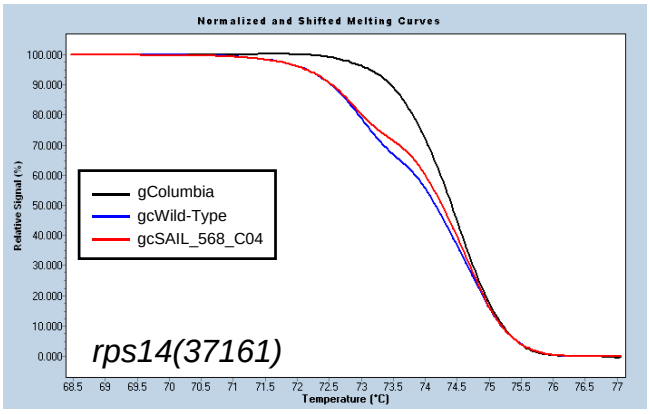

Normalized and Shifted Melting Curves
rpoB(25779)
Relative Signal (%)
Temperature (°C)

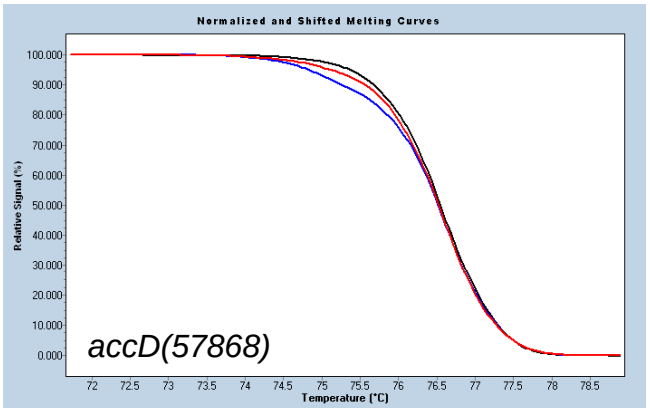

Normalized and Shifted Melting Curves
rpoB (25992)
Relative Signal (%)
Temperature (°C)

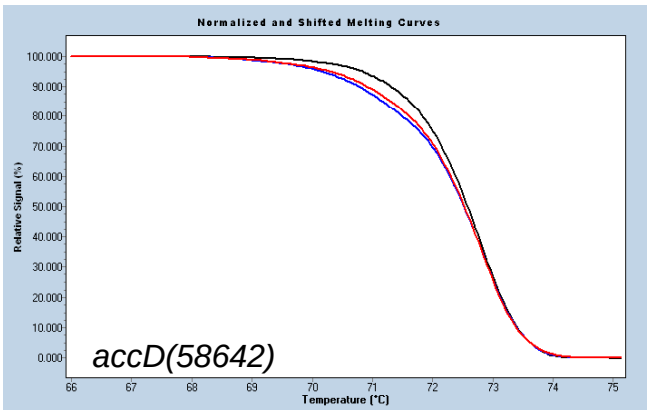

Normalized and Shifted Melting Curves
psbZ(35800)
Relative Signal (%)
Temperature (°C)

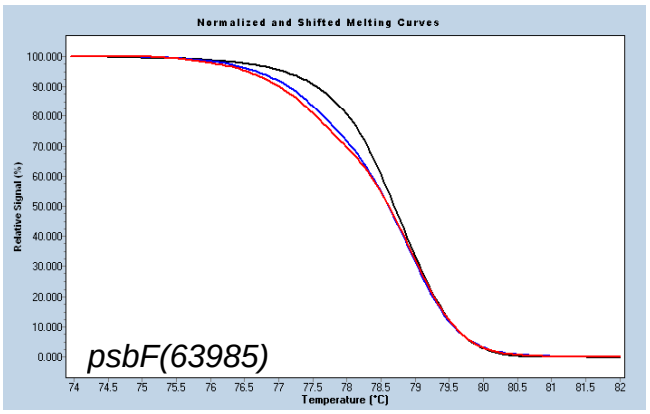

Normalized and Shifted Melting Curves
rps14(37092)
Relative Signal (%)
Temperature (°C)

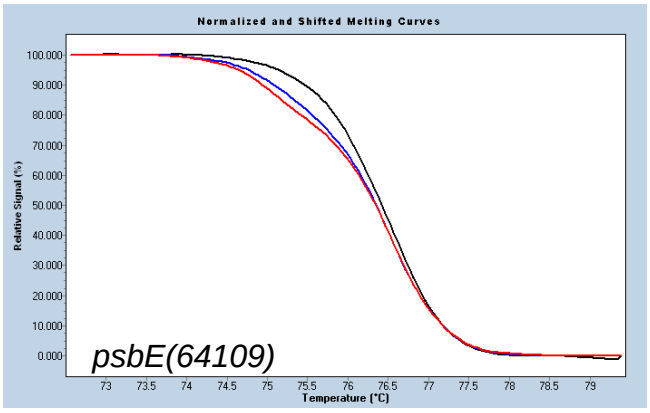
otp84

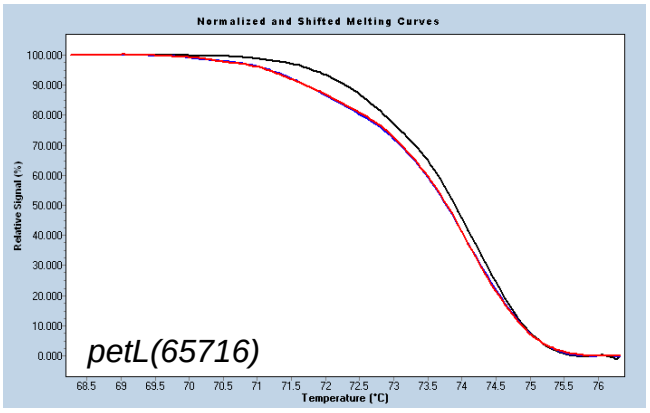

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
gColumbia
gcWild-Type
gcSAIL_568_C04
rps14(37161)

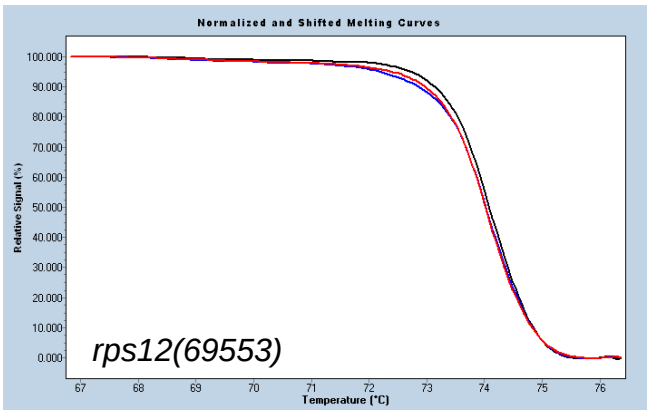

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
accD(57868)

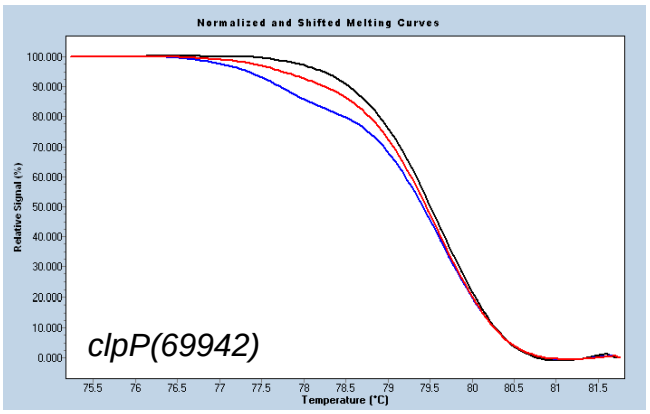

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
accD(58642)


Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
psbF(63985)

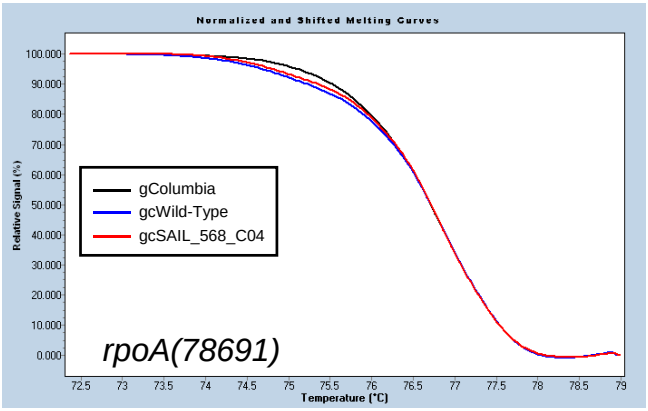

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
psbE(64109)

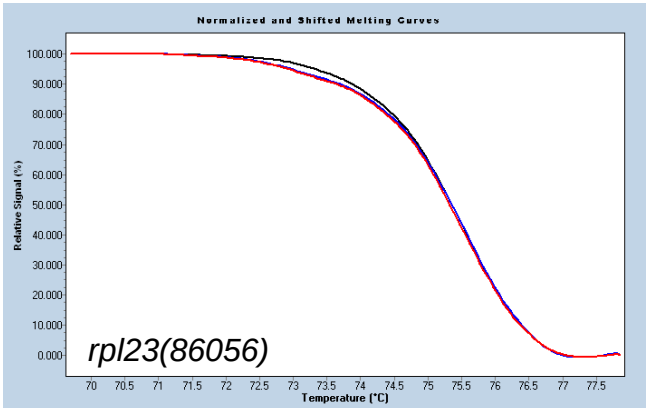

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
petL(65716)

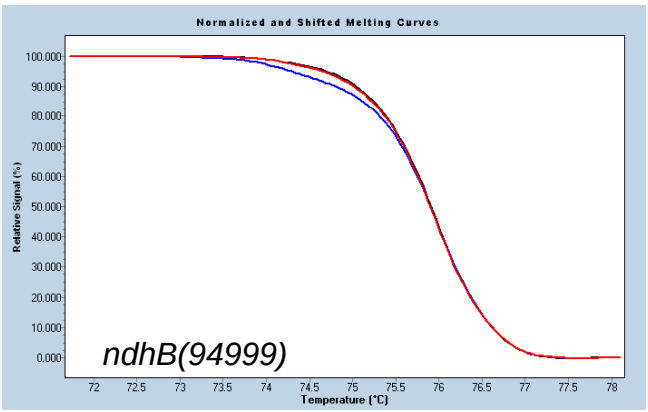

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
rps12(69553)

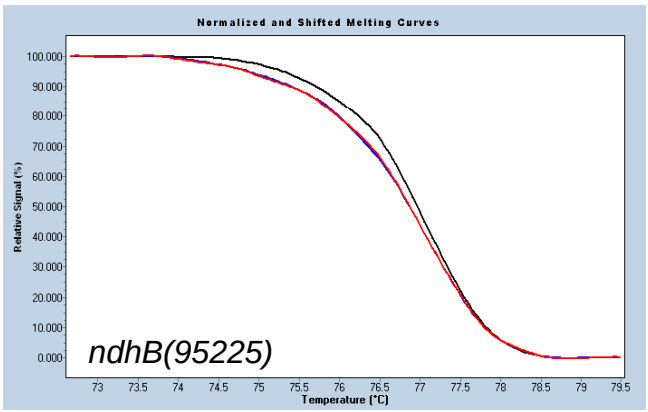

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
clpP(69942)

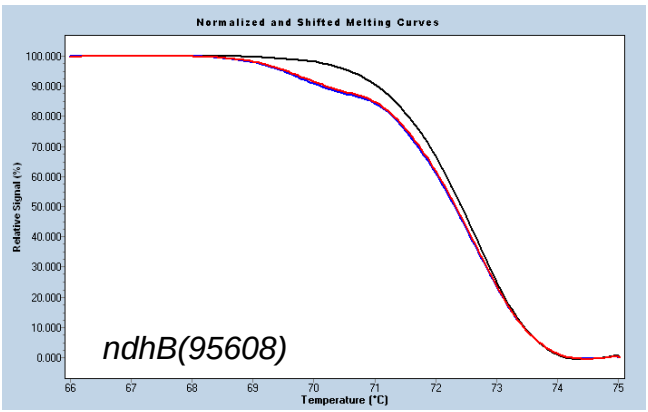
otp84

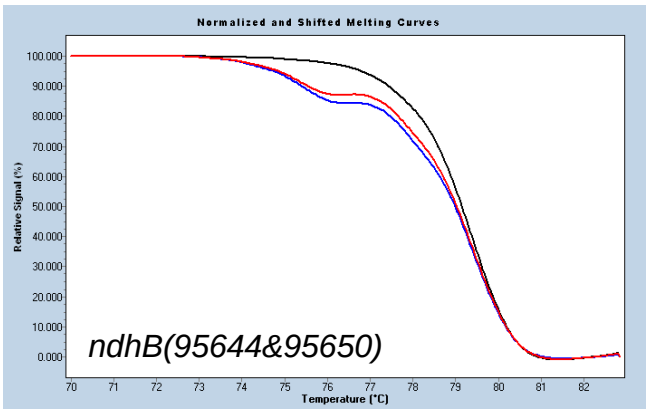

Normalized and Shifted Melting Curves
gColumbia
gcWild-Type
gcSAIL_568_C04
rpoA(78691)
Relative Signal (%)
Temperature (°C)

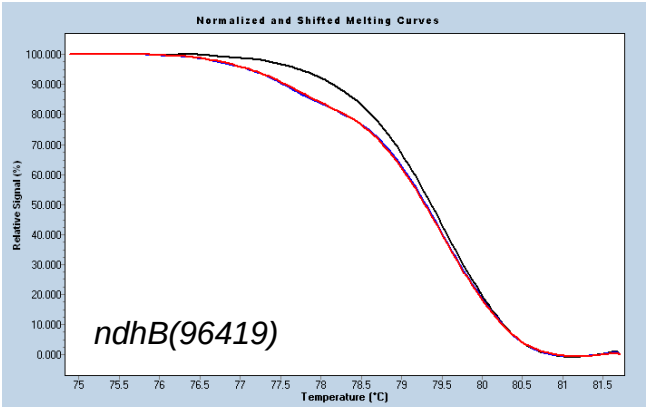

Normalized and Shifted Melting Curves
rpl23(86056)
Relative Signal (%)
Temperature (°C)

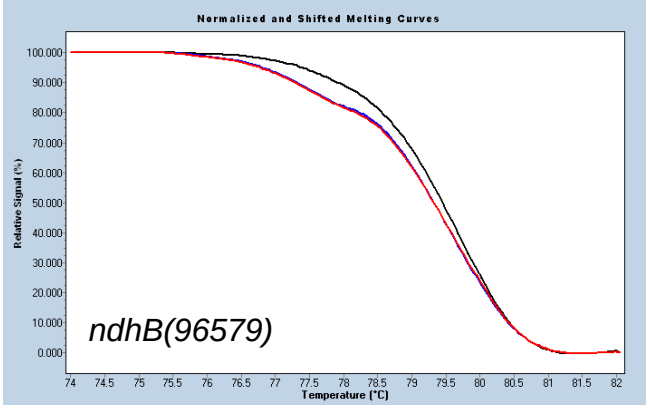

Normalized and Shifted Melting Curves
ndhB(94999)
Relative Signal (%)
Temperature (°C)


Normalized and Shifted Melting Curves
ndhB(95225)
Relative Signal (%)
Temperature (°C)

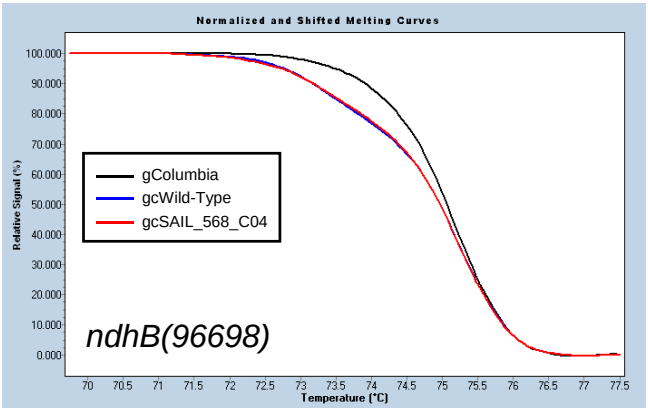

Normalized and Shifted Melting Curves
ndhB(95608)
Relative Signal (%)
Temperature (°C)

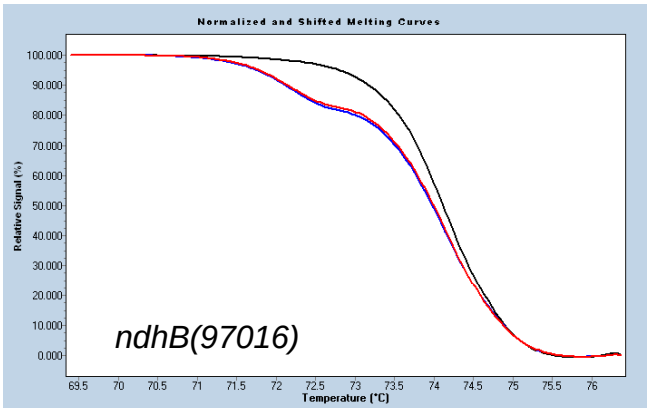

Normalized and Shifted Melting Curves
ndhB(95644&95650)
Relative Signal (%)
Temperature (°C)

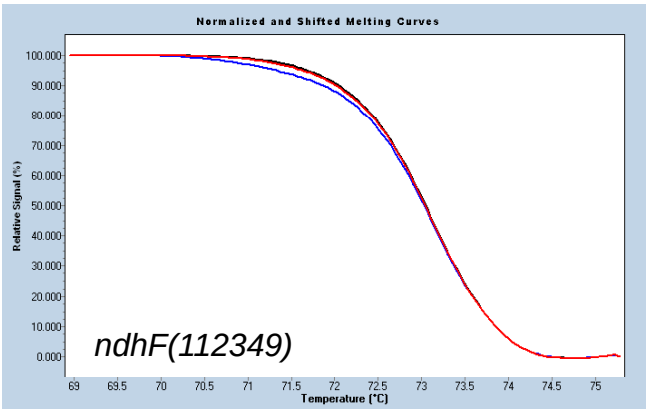

Normalized and Shifted Melting Curves
ndhB(96419)
Relative Signal (%)
Temperature (°C)

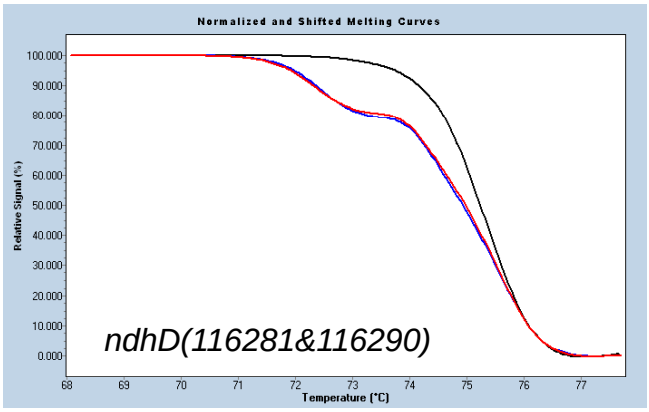

Normalized and Shifted Melting Curves
ndhB(96579)
Relative Signal (%)
Temperature (°C)

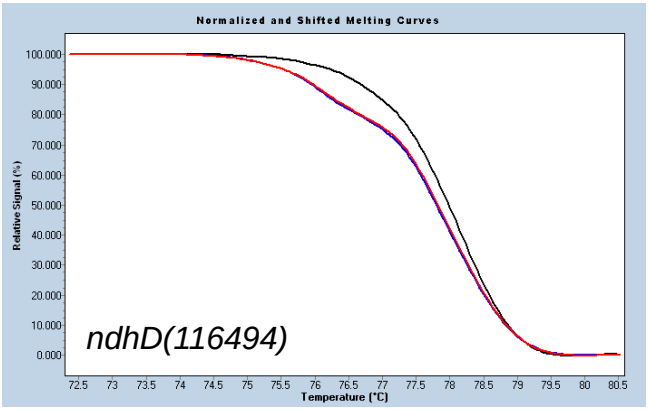
otp84

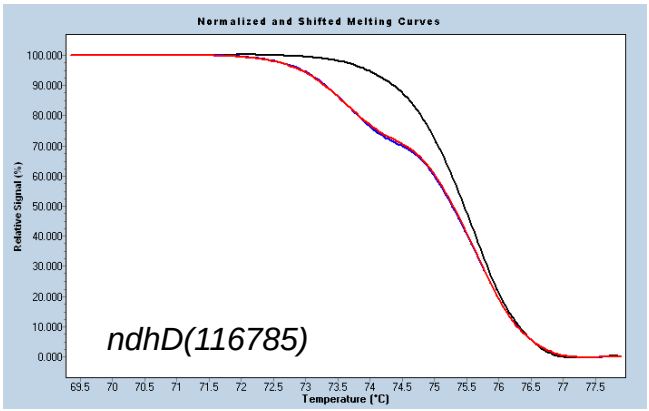

Normalized and Shifted Melting Curves
gColumbia
gcWild-Type
gcSAIL_568_C04
ndhB(96698)
Relative Signal (%)
Temperature (°C)

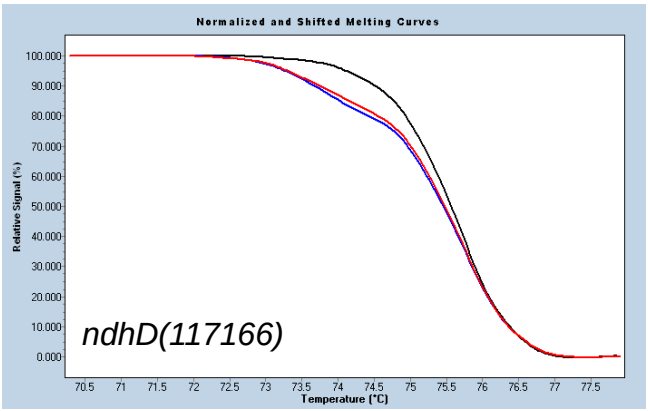

Normalized and Shifted Melting Curves
ndhB(97016)
Relative Signal (%)
Temperature (°C)

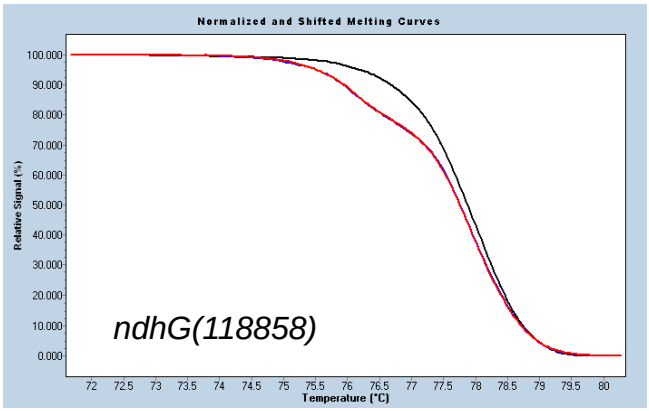

Normalized and Shifted Melting Curves
ndhF(112349)
Relative Signal (%)
Temperature (°C)


Normalized and Shifted Melting Curves
ndhD(116281&116290)
Relative Signal (%)
Temperature (°C)

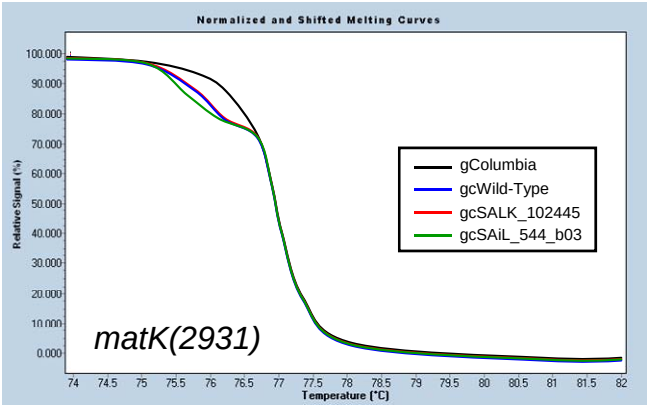

Normalized and Shifted Melting Curves
ndhD(116494)
Relative Signal (%)
Temperature (°C)

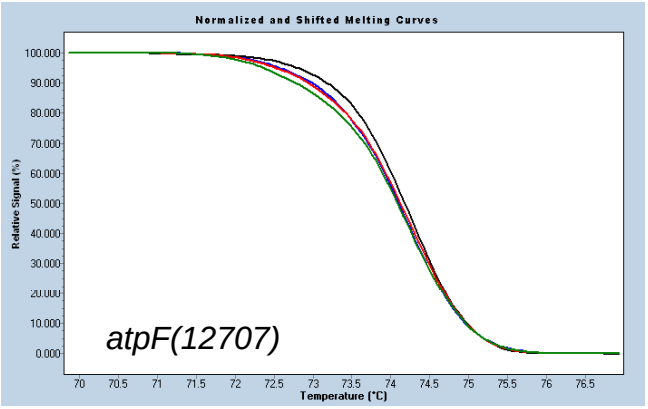

Normalized and Shifted Melting Curves
ndhD(116785)
Relative Signal (%)
Temperature (°C)

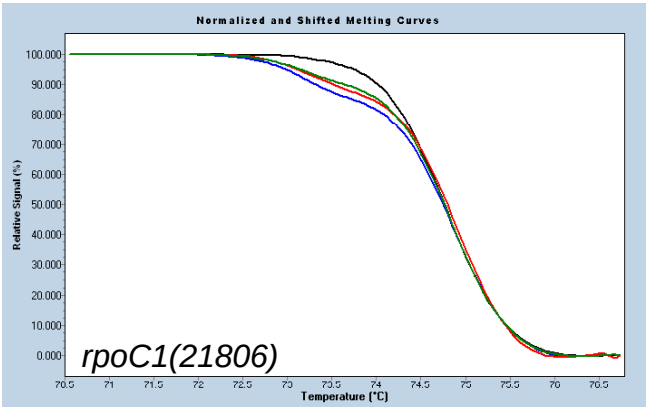

Normalized and Shifted Melting Curves
ndhD(117166)
Relative Signal (%)
Temperature (°C)

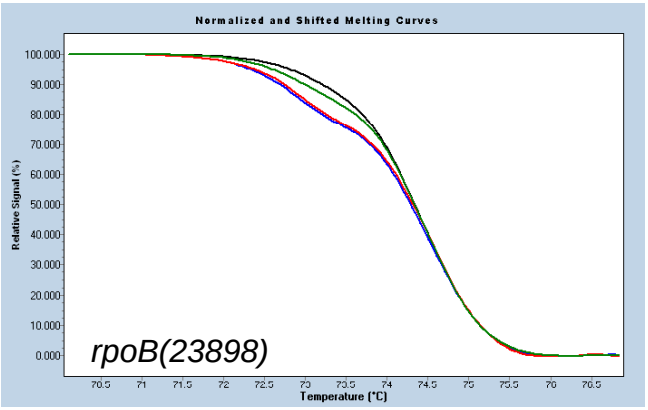

Normalized and Shifted Melting Curves
ndhG(118858)
Relative Signal (%)
Temperature (°C)

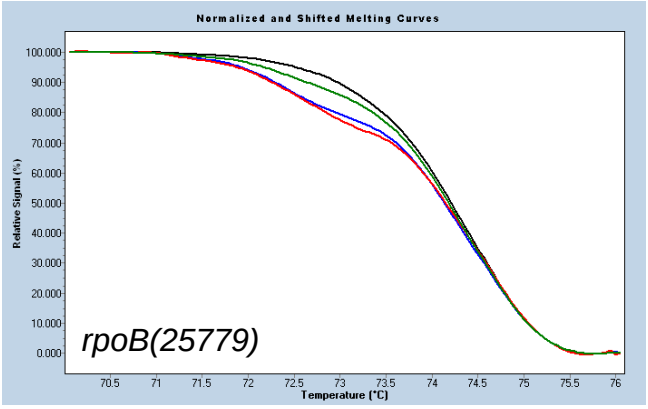
otp85, otp86

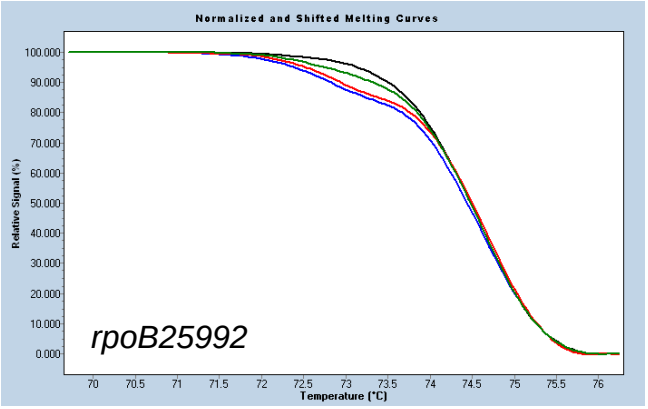

Normalized and Shifted Melting Curves
gColumbia
gcWild-Type
gcSALK_102445
gcSAiL_544_b03
matK(2931)
Relative Signal (%)
Temperature (°C)

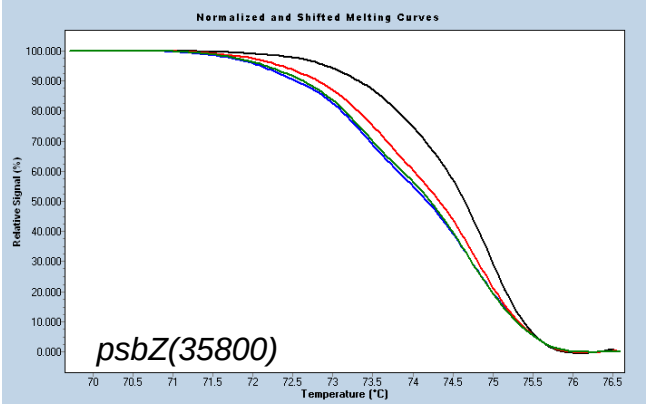

Normalized and Shifted Melting Curves
atpF(12707)
Relative Signal (%)
Temperature (°C)

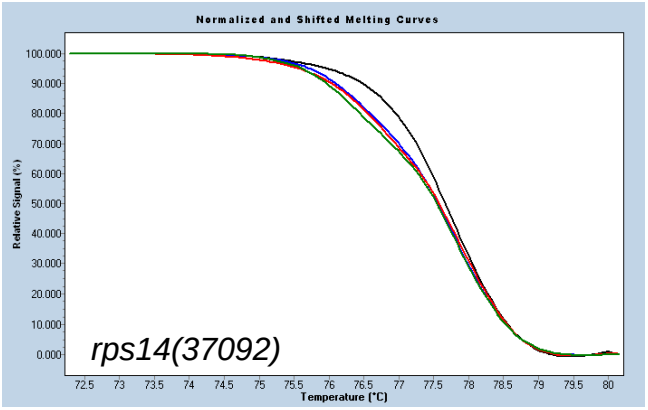

Normalized and Shifted Melting Curves
rpoC1(21806)
Relative Signal (%)
Temperature (°C)


Normalized and Shifted Melting Curves
rpoB(23898)
Relative Signal (%)
Temperature (°C)

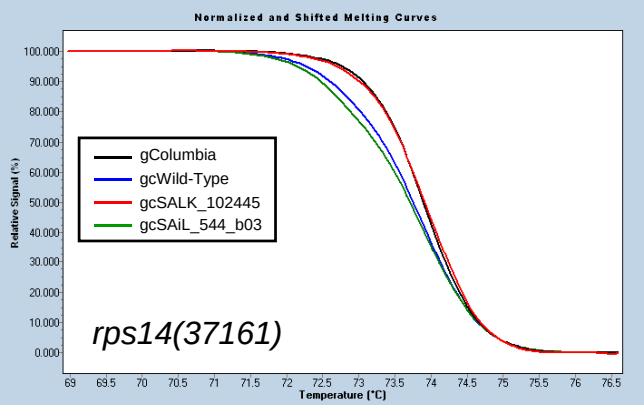

Normalized and Shifted Melting Curves
rpoB(25779)
Relative Signal (%)
Temperature (°C)

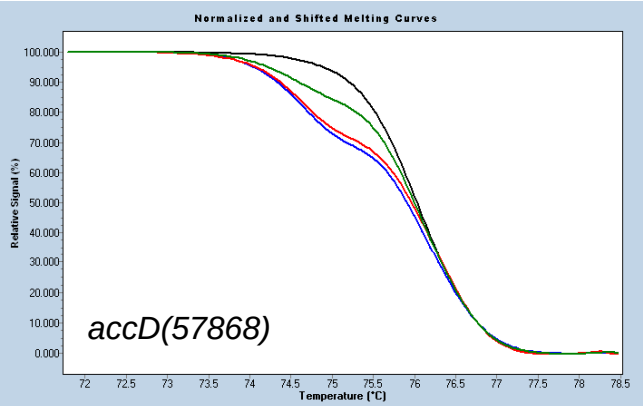

Normalized and Shifted Melting Curves
rpoB25992
Relative Signal (%)
Temperature (°C)

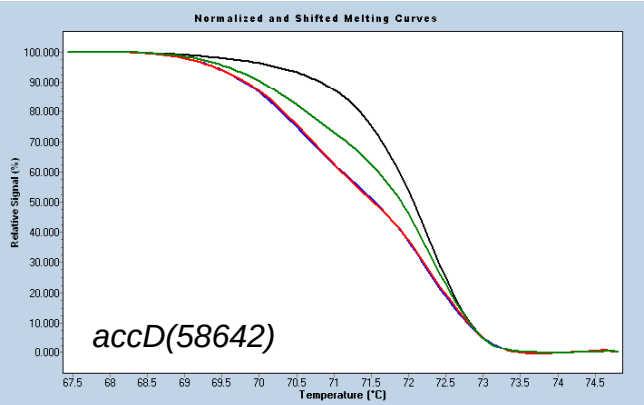

Normalized and Shifted Melting Curves
psbZ(35800)
Relative Signal (%)
Temperature (°C)

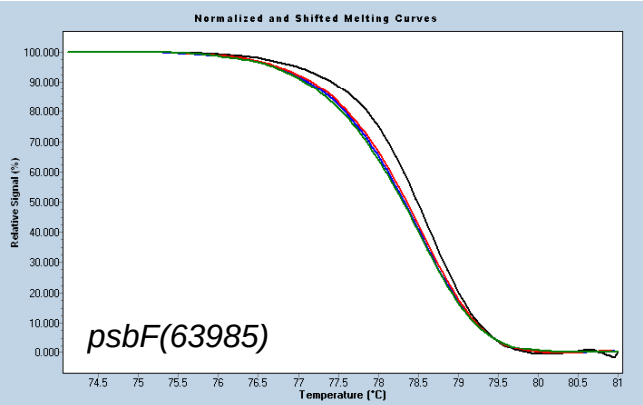

Normalized and Shifted Melting Curves
rps14(37092)
Relative Signal (%)
Temperature (°C)

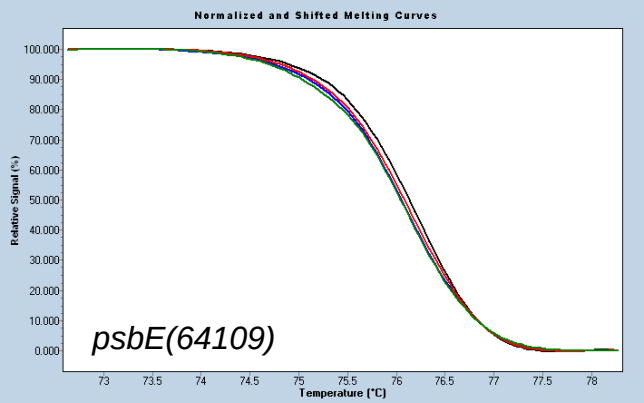
otp85, otp86

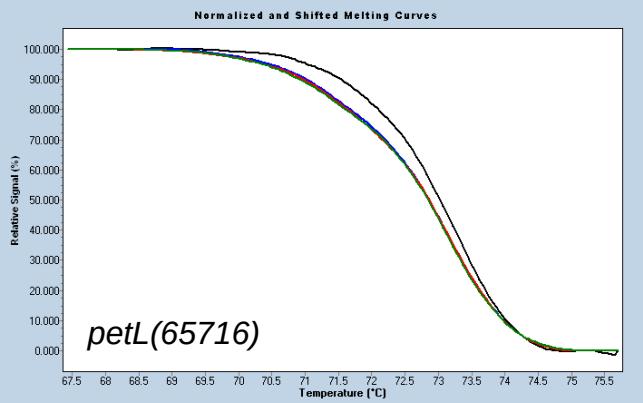

Normalized and Shifted Melting Curves
gColumbia
gcWild-Type
gcSALK_102445
gcSAiL_544_b03
rps14(37161)
Relative Signal (%)
Temperature (°C)

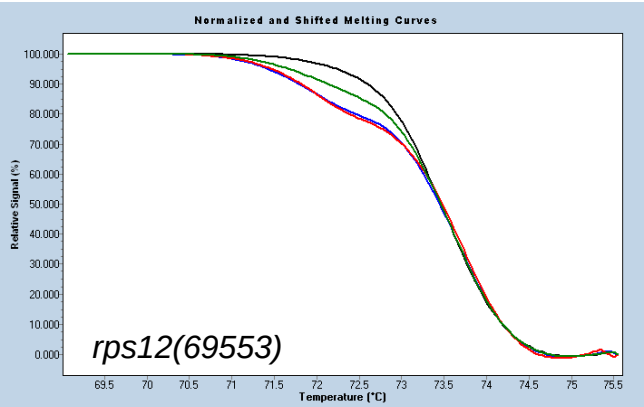

Normalized and Shifted Melting Curves
accD(57868)
Relative Signal (%)
Temperature (°C)

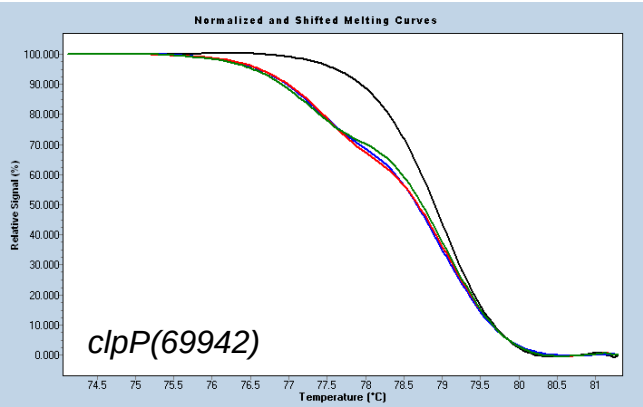

Normalized and Shifted Melting Curves
accD(58642)
Relative Signal (%)
Temperature (°C)


Normalized and Shifted Melting Curves
psbF(63985)
Relative Signal (%)
Temperature (°C)

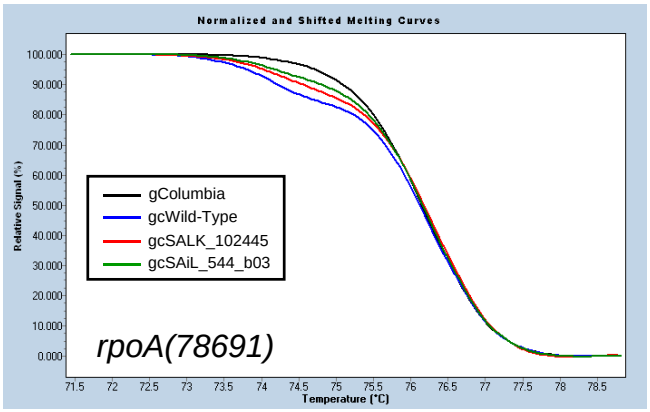

Normalized and Shifted Melting Curves
psbE(64109)
Relative Signal (%)
Temperature (°C)

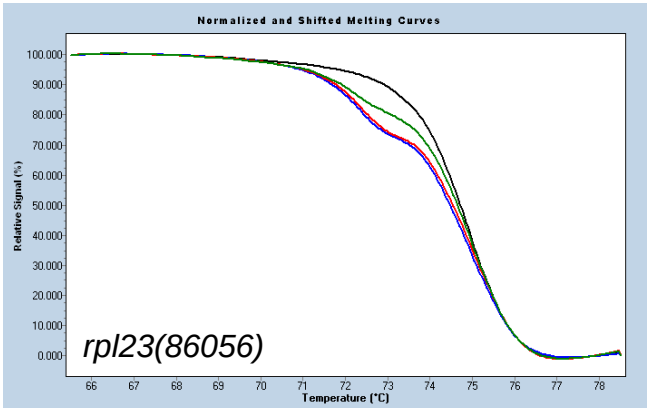

Normalized and Shifted Melting Curves
petL(65716)
Relative Signal (%)
Temperature (°C)

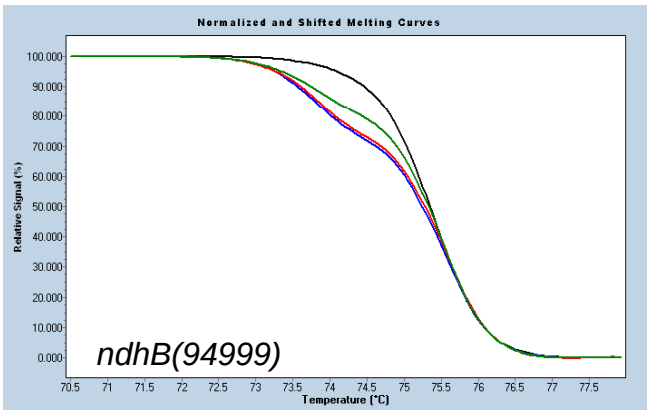

Normalized and Shifted Melting Curves
rps12(69553)
Relative Signal (%)
Temperature (°C)

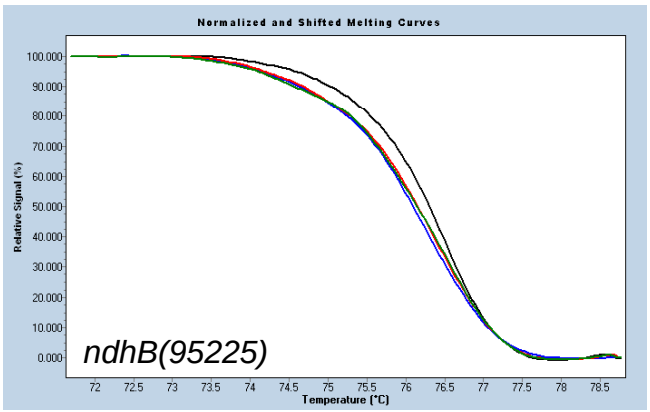

Normalized and Shifted Melting Curves
clpP(69942)
Relative Signal (%)
Temperature (°C)

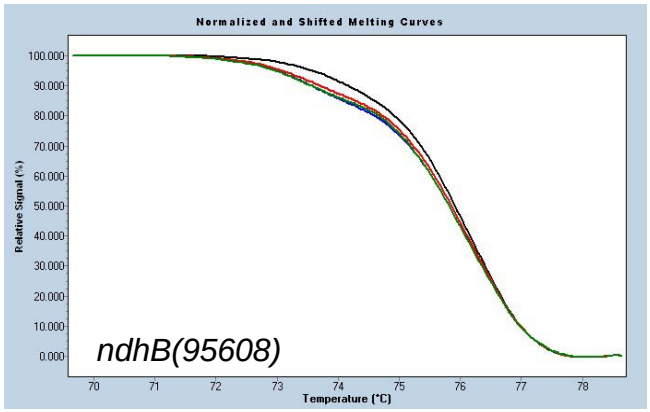
otp85, otp86

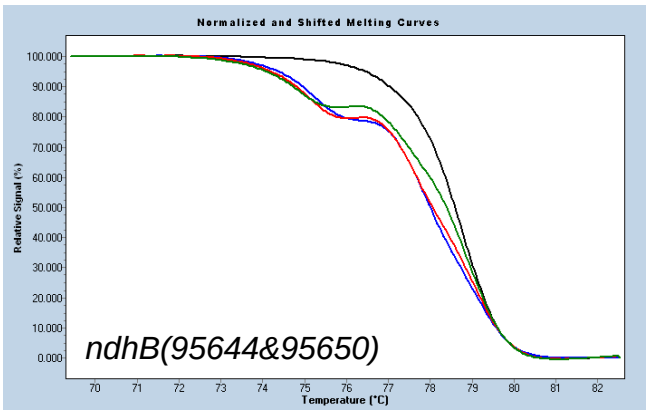

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
gColumbia
gcWild-Type
gcSALK_102445
gcSAiL_544_b03
rpoA(78691)

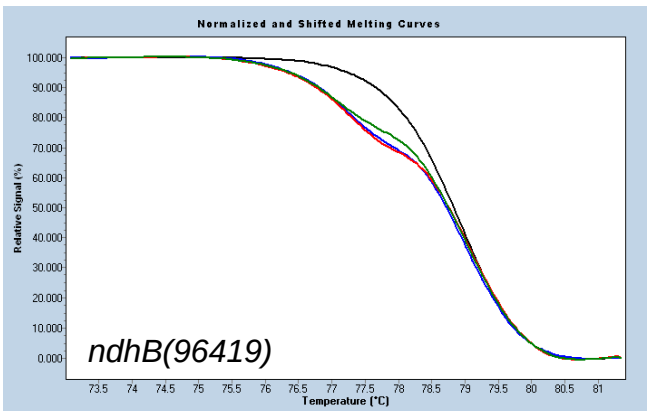

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
rpl23(86056)

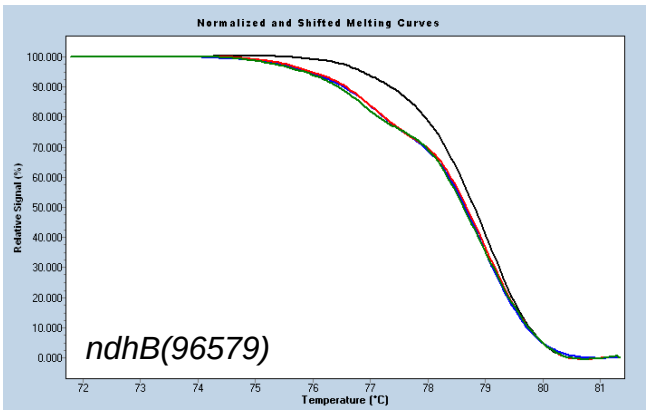

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhB(94999)


Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhB(95225)

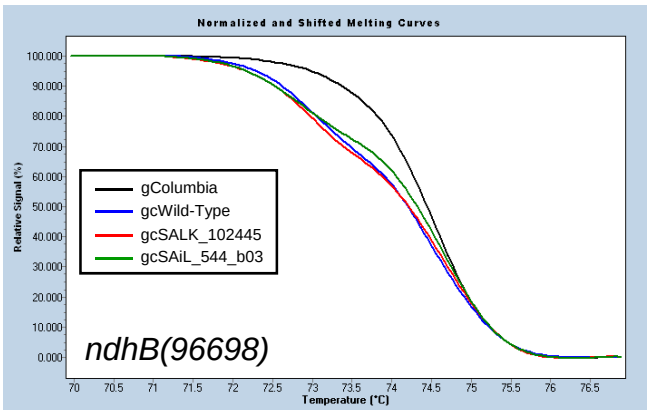

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhB(95608)

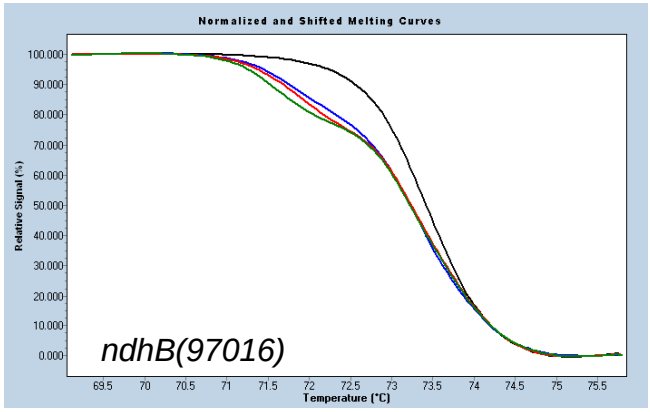

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhB(95644&95650)

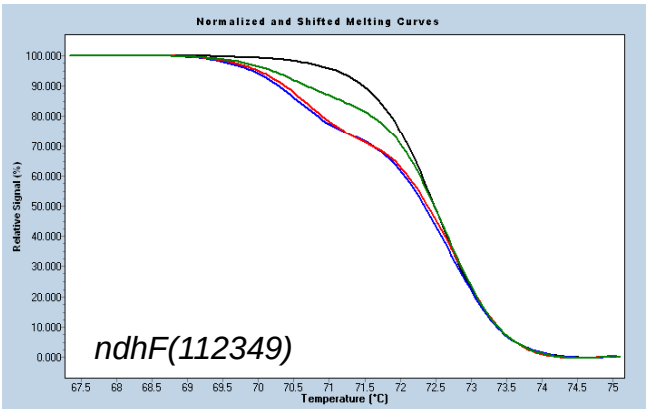

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhB(96419)

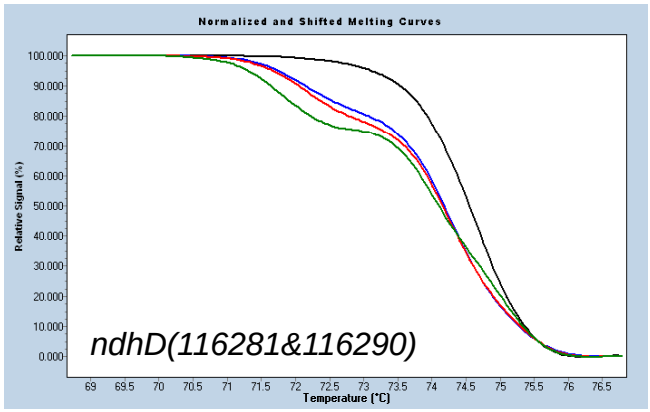

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhB(96579)

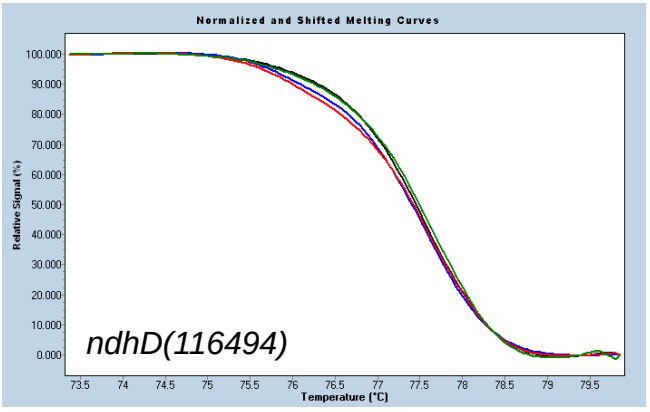
otp85, otp86

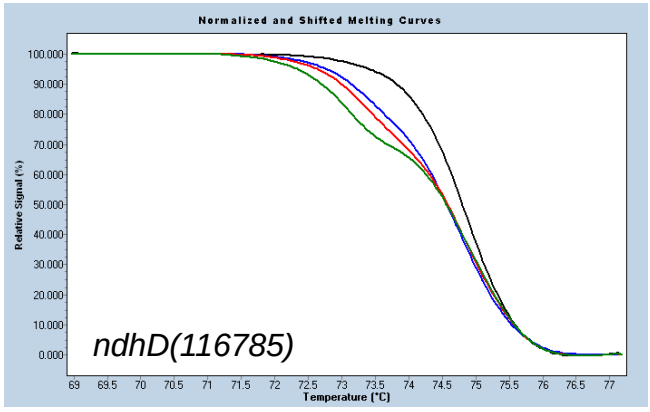

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
gColumbia
gcWild-Type
gcSALK_102445
gcSAiL_544_b03
ndhB(96698)

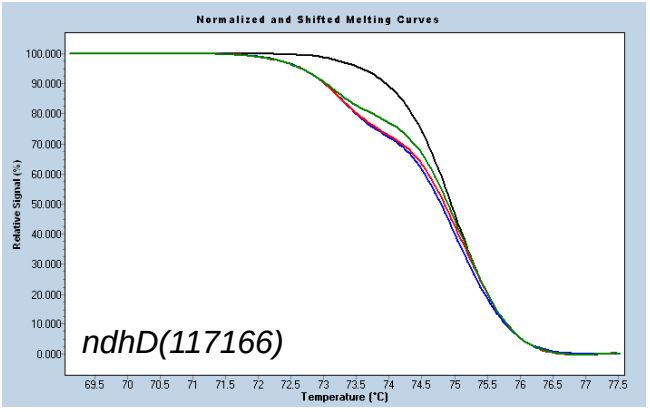

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhB(97016)


Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhF(112349)


Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhD(116281&116290)


Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhD(116494)


Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhD(116785)


Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhD(117166)

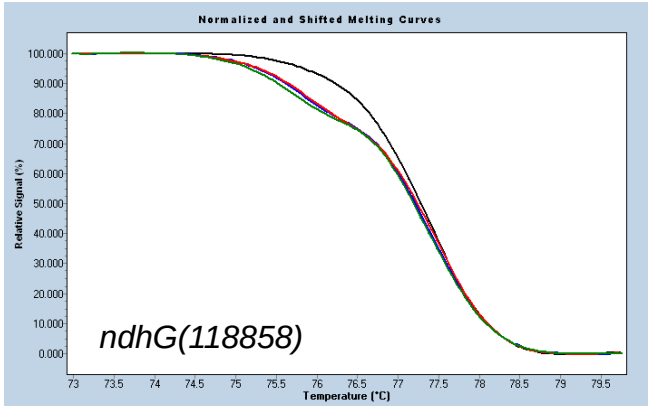

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhG(118858)

Figure 1. High resolution melting screen of *otp80*, *otp81*, *otp84*, *otp85* and *otp86* mutants.
Real-time PCR was done using primers surrounding the 34 known editing sites of Arabidopsis in the presence of a fluorescent double-strand-specific dye. At the end of the amplification, amplicons were denatured, renatured and then melted using precise incremental increases in temperature. Melting of DNA duplexes releases the dye, causing a decrease in fluorescence. The presence of less thermostable heteroduplexes in a sample alters the shape of the melting curves. For each editing site, the melting curve of a control consisting of genomic DNA (g) from the wt sample was compared to the melting curves of a mix of genomic DNA and cDNA (gc) from the mutant lines. Editing produces cDNA with a nucleotide change and thus the mix of genomic DNA and cDNA exhibits a different melting curve from that of genomic DNA alone. If an editing site is unedited in the mutant, then the melting curves from the genomic DNA sample and the mixed genomic/cDNA sample are identical.

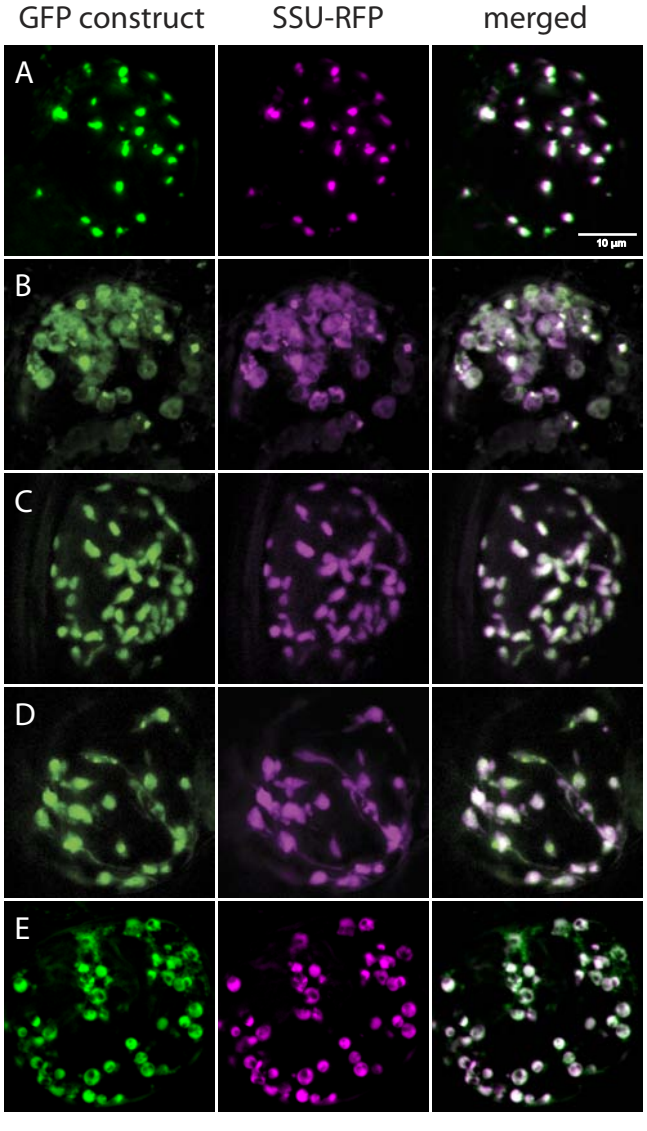
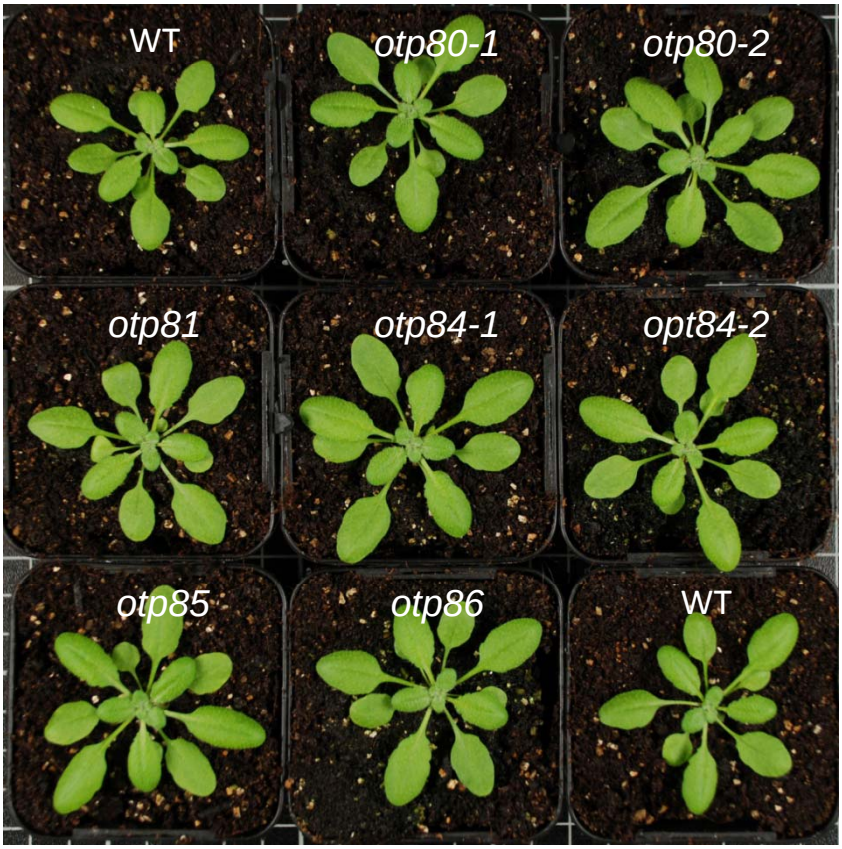


Figure 2. Analysis of subcellular localization of the OTP80 (A), OTP81 (B), OTP84 (C), OTP85 (D) and OTP86 (E) proteins.
Chimeric proteins consisting of the first 100 amino acids of OTP80, OTP81, OTP84, OTP86 or the first 72 amino acids of OTP85 fused to GFP were targeted to chloroplasts in transformed Arabidopsis cells. Fluorescence of the green fluorescent protein (GFP construct, shown in green), fluorescence of the red fluorescent protein in fusion with the small subunit of the ribulose biphosphate carboxylase (SSU-RFP, shown in magenta), and the overlay of the two fluorescence images are shown (merged).

A



B

Figure 3. Phenotype of *otp80*, *otp81*, *otp84*, *otp85*, *otp86* and wild-type (WT) plants. Three- **(A)** and seven- **(B)** week-old plants were grown under medium day conditions.

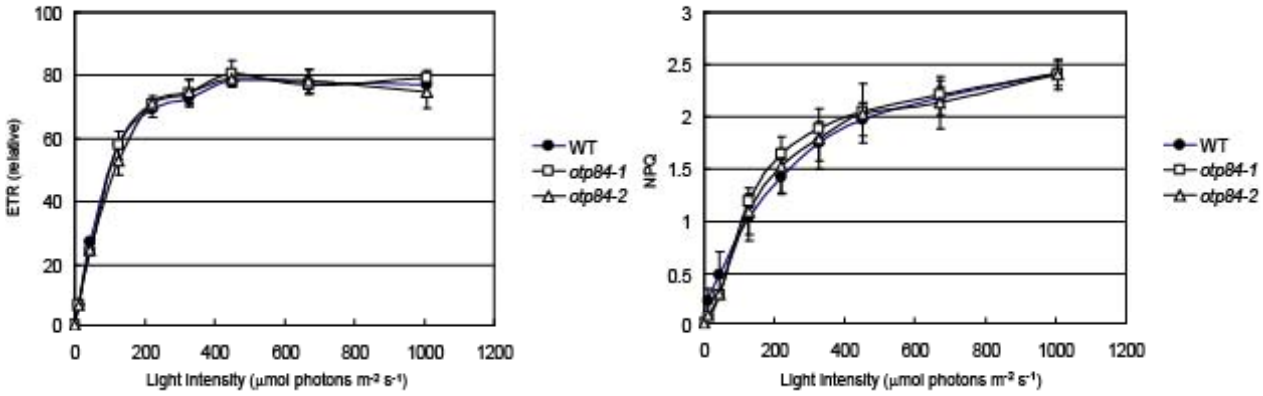


Figure 4. *In vivo* analysis of electron transport activity, light intensity dependence of ETR (electron transport rate). ETR is plotted relative to ϕ_{PSII} × light intensity (μmol photons $m^{-2}.s^{-1}$), where ϕ_{PSII} is quantum yield of photosystem II. Light intensity dependence of NPQ (non-photochemical quenching) of chlorophyll fluorescence. All values represent the mean ± SD (n=3).

Supplemental Method

Chlorophyll Fluorescence Analysis

Chlorophyll fluorescence was measured with a MINI-PAM portable chlorophyll fluorometer (Walz, Effeltrich, Germany) in ambient air at room temperature (25oC). Minimum fluorescence at open PSII centers in the dark-adapted state (Fo) was excited by a weak measuring light (wavelength 650 nm) at a PFD of 0.05-0.1 μmol m-2 s-1. A saturating pulse of white light (800 ms, 3,000 μmol photons m-2 s-1) was applied to determine the maximum fluorescence at closed PSII centers in the dark-adapted state (Fm) and during AL illumination (Fm'). The steady state fluorescence level (Fs) was recorded during AL illumination (15-1,000 μmol photons m-2 s-1). These photosynthetic parameters were determined 2 min after the change of AL (Actinic Light) intensity. NPQ was calculated as (Fm-Fm')/Fm'. The quantum yield of PSII (PSII) was calculated as (Fm'-Fs)/Fm' (Genty et al., 1989). ETR was calculated as PSIIxPFD. The transient increase in chlorophyll fluorescence after AL had been turned off was monitored as described (Shikanai et al., 1998).

Supplemental Table 1. Oligonucleotides used in this study.

For genotyping

LB SALK	GCGTGGACCGCTTGCTGCAAC
LB SAIL	AGCATCTGAATTTCATAACCAATCTCGATACAC
LB GABI	ATATTGACCATCATACTCATTGC
SALK_060533RP	TTATCAGGCAAAACCGAATTG
SALK_060533LP	TGTTCAAATTGGTTTCCTTGG
SALK_111721RP	TCCATCGATTCACGCCAAGAT
SALK_111721LP	CCTCCATGGCTACATGCGTTT
SALK_092402RP	TGCTGACTTAACAGCCATCCCA
SALK_092402LP	GATGGATCCCCCAAGTTGACC
SALK_120902LP	GTCATGTCCTCTTGCCTTCAC
SALK_120902RP	GGAATTCTCATCCAAAGAGCC
SAIL_568_C04LP	GTAAGCAGGTTCACGCCTATG
SAIL_568_C04RP	TTAATCCTTCATCCACCATGC
SAIL_544_B03LP	GTAGCACTTCGGTCGAGAATG
SAIL_544_B03RP	TACGCGAGAAACAAGAAATGG
SALK_102445LP	CTACGTTCTCGAGCTTTGTGG
SALK_102445RP	TCCCAGACGTGGAATATGAAG
SALK_025427LP	CCTTTCAAGACTTCACGCATC
SALK_025427RP	TTCGTGGAATGCGATTATCTC
SAIL_1291_C04LP	GATAATCCCGGTTCTAGCTCG
SAIL_1291_C04RP	CGATCTTTCACCGTTTGAATC
SAIL_205_G08LP	GCTTGTTCTGGTCAATCATGG
SAIL_205_G08RP	CGCCCTACTCGACCTAATAGG
SALK_118555RP	TCAAGGAATCTCTTGGCTGCG
SALK_118555LP	CGGTATGATCCAGAGGAACATCG
SALK_23916LP	GAATCACTTGGCCATGATCCG
SALK_23916RP	TGCAAGCTTAGGAGCCCTTGA
SALK_090845LP	TCTGGTTTAGCATTAGCACGG
SALK_090845RP	TCGAAAAACAATCCATTCACC
SALK_040629LP	TTTAGCCAAACCGTCGATTAC

SALK_040629RP	TCAGTTCGAGGGAATGAACAG
GABI_073C06LP	ATCTTCCGTTACCAATTTCCG
GABI_073C06RP	GGCTTGCTTCACAGACCATAG
SAIL_1249_D04RP	TTCCGGGATGTAGCTATGTTG
SAIL_1249_D04LP	TTTCAAAACCTCCAAAATTCAG
SALK_044324RP	GGCCTTGTCTTCATCAAACAG
SALK_044324LP	TAAGCGGATCAAGAATGCTTG
GABI_895H11RP	GCCTGTGAGAGATCTCGTGTC
GABI_895H11LP	TGCATTCCCTATTTGCAACTC

For GFP and complementation construct cloning

z0779 GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC atcaggaaccaagggcacact at5g59200:OTP80 oligo forward complementation

z0781 GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC atgatttcttccttggcagcg at5g59200:OTP80 oligo forward GFP fusion

z0782 GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC ctagaaacccattataatatc at5g59200:OTP80 oligo reverse complementation

z0783 GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC atcgatcatagcagtgtatag at5g59200:OTP80 oligo reverse GFP fusion

z0789 GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC atggctatcttctccacagc at2g29760:OTP81 oligo forward GFP fusion

z0790 GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC tcaccagaaatcgttacagg at2g29760:OTP81 oligo reverse complementation

z0792 GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC cttgattcctgcctctagg at2g29760:OTP81 oligo forward complementation

z0816 GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC atgtcatgtcctcttgcc at3g57430:OTP84 oligo forward complementation and GFP fusion

z0817 GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC tcaccaatagtctccacagg at3g57430:OTP84 oligo reverse complementation

z0819 GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC gaacgcataattatcggg at3g57430:OTP84 oligo reverse GFP fusion

z0826 GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC cgcatcaacaattcaaaaaccg At2g02980:*OTP85*

oligo forward complementation

z0827 GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC atggcgatatcttcagcttc At2g02980:*OTP85*

oligo forward GFP fusion

z0828 GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC TCACCAAAAATCCCCACAAG At2g02980:*OTP85*

oligo reverse complementation

z0830 GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC CTCAGTGCAGAAATTTATCAGC At2g02980:*OTP85*

oligo reverse GFP fusion

z0836 GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC atggaatacgcggtaactaac At3g63370:*OTP86*

oligo forward complementation and GFP fusion

z0837 GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC TCACCAAGAATCTCCGCAAG At3g63370:*OTP86*

Oligo reverse complemenation

z0839 GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC CCGACCTTGTGAGACAGCTCT At3g63370:*OTP86*

oligo reverse GFP fusion

For generating PCR templates for PPE assay

chloro140R	CAAGACAGCCAATCCGAAAC	*rps12*
chloro140F	CTTGTACAATTCACATTCTTTGGC	
ndhB.AT.rev2	CTAAAAAAGGGTATCCTGAGC	*ndhB*
AndhB	GTCGTTGCTTTTCTTTCTG	
chloro68F	AGAACATAGCCCTATGAGTTAATACGA	*psbZ*
chloro68R	GATAAGAGAATTAAGGATACCCACCA	
rps14.AT.for	TTATAGGGAGAAGAAGAGGC	*rps14*
rps14.AT.rev	TACCAGCTTGATCTTGTTGC	
QC840F	CGGTTATTGGGGAAAAATCA	*rpl23*
QC840R	TTTTAACCTTTCCGGGGAGT	
ndhD.AT.for	CAAGCCTAATTCTATCATAACTCG	*ndhD*
ndhD.AT.rev	AAGTTTATATGGTTCGAACG	

For PPe analysis

rps14(37161)_PPE_C	AAAATATCATTTGATTCGTCGATCC
ndhD(116494)_PPE_C	TATTGGATTTCTTATTGCTTTTGCCGTC

ndhB(94999)_PPE_A	GCAATCGCAATAATCGGGTTCATTGATA	
rpl23_PPE_C	GGGGAAAAATCAATATACTTTTAATGTCGAATC	
psbZ_PPE_C	TTGGCAGTTTTTGCATTAATTATTACTTC	
Rps12_PPE_G	CATTCTATTTTATTCATAGGTACTGATCCG	
ndhF_PPE_C	TTCGGTTACTTTATCGATCCACTTACTTC	

For RNA gel blot analysis

T7_ndhF_R	TAATACGACTCACTATAGGGGCATGTATAAGAGCCGAAATGGGAGTAGGGCCC	ndhF
ndhF_F	GGGAATTAGTTGGAATGTGTTCGTATTTATTAA	
ndhB.AT.rev2	CTAAAAAAGGGTATCCTGAGC	ndhB
AndhB	GTCGTTGCTTTTCTTTCTG	
chloro213F	AATAGTATTTCAAGTGTAACAGGATAGGA	ndhD
chloro213R	TGGACGATCCATTAATTCAACT	
QC840F	CGGTTATTGGGGAAAAATCA	rpl23
QC840R	TTTTAACCTTTCCGGGGAGT	
chloro68F	AGAACATAGCCCTATGAGTTAATACGA	psbZ
chloro68R	GATAAGAGAATTAAGGATACCCACCA	
chloro140R	CAAGACAGCCAATCCGAAAC	rps12
chloro140F	CTTGTACAATTCACATTCTTTGGC	
rps14.AT.for	TTATAGGGAGAAGAAGAGGC	rps14
rps14.AT.rev	TACCAGCTTGATCTTGTTGC	
M13/pUC reverse	GAGCGGATAACAATTTCACACAGG	

For HRM analysis

matK_HRM_for:	TGAATCCAAGATTTTTCTTGTTCTT
matK_HRM_rev:	AAAAAGATGGATTCGTATTCACA
atpF_HRM_for:	GCCGGGAGTTTCGGATTTA
atpF_HRM_rev:	GATCAATACACCGAAAACTACACTT
rpoC1_HRM_F:	TTCTTTTGCTAGGCCCATAA
rpoC1_HRM_R:	TGCTGTATTTCCAGGATTGAA
rpoB23898_HRM_for:	TGGTTCAAGTTATAACCCAGAAAT
rpoB23898_HRM_rev:	CCATGTCTTCCAGCTACTTTATCA

rpoB25779_HRM_for:	GCTCGGGTGAGTAGGAAACA
rpoB25779_HRM_rev:	TCTAGAATTTCTCGTAGATTCAAACC
rpoB25992_HRM_for:	TCCTTTAATGAATTCCCTTGGA
rpoB25992_HRM_rev:	CGGGACTTTGCAATATTTGAT
psbZ_HRM_for:	TGCTTTCCAATTGGCAGTTT
psbZ_HRM_rev:	GAGACGCAAATACAACGGGTA
RPS14(37092)HRM_for:	TCGCTAAGTGAGAAATGGAAAA
RPS14(37092)HRM_rev:	CGTCGATGAAGACGTGTAGG
RPS14(37161)HRM_for:	GAAGAAGAGGCAAAAATTGGAA
RPS14(37161)HRM_rev:	TTCTCACTTAGCGACGGAATC
accD_HRMfor:	TGAATGTTTGTGAACAATGTGG
accD_HRMrev:	GATTCCAAGTACCCGGATCA
accD(58642)HRM_rev:	ACTTTTTAGCTTGTTGATAGAGGTTT
accD(58642)HRM_for:	GCATTTTTCTTTCAAATCATTTTT
psbF_HRM_rev:	TCGTTGGATGAACTGCATTG
psbF_HRM_for:	GGCTGTTCATGGACTAGCTGTA
psbE_HRM_rev:	TGGGTCCTCCTAAAAAGATCTAC
psbE_HRM_for:	AAGGCATTCCATTAATAACAGG
petL_HRM_for:	ATTTTATTGAGTCCCTTCATGC
petL_HRM_rev:	AAAGCTGCTAGTAGAAAACCGAAA
Rps12_HRM_rev:	TCGAGTTCGTCCATTCTATTTT
Rps12_HRM_for:	AGACTGAGACATAAAAAGGAAATTCT
clpPAt2rev:	TGAACCGCTACAAGATCAAC
clpP_HRM70for:	GAAGACATGGAACGGGATGT
rpoA_HRMfor:	TTGGCGAAATAGAAGGAACA
rpoA_HRMrev:	CCTGCTATGTTAGAATAGTCATGTGG
Rpl23_HRM_r:	GAAGAGTTCGACCCAATGCT
Rpl23_HRM_f:	TCAATATACTTTTAATGTCGAATCAGG
ndhB(94999)HRM_for:	TTCCATCGAATTGAGTATGATTG
ndhB(94999)HRM_rev:	AATCGCAATAATCGGGTTCA
ndhB95225_HRM_for:	GGTCTTCCTCCACTAGCAGGT
ndhB95225_HRM_rev:	CTGCCATCCACACCAGAATA

95640R:	CACTCGAATTTTCGATATTCCTTT
ndhB(95608)HRM_rev:	TCCAGAAGAAGATGCCATTCA
95618F:	AAGGAATATCGAAAATTCGAGTGG
95693R:	CCAGTCGTTGCTTTTCTTTCTG
95627F:	CGAAAATTCGAGTGGCTGA
95698R:	CCACTCCAGTCGTTGCTTTT
96461R:	TCATCACTGTAGGAATTGGGTTC
ndhB(96419)HRM_rev:	CTTCGTATACGTCAGGAGTCCA
96606R:	GGTGGGGCAAGCTCTTCTAT
96537F:	CAATCTCTCCCCCGGATG
ndhB(96698)HRM_for:	TTTTTATGTGGTGCTAACGATTT
ndhB(96698)HRM_rev:	CCAGATAATAGGTAGGAGCATAAACTG
ndhB(97016)HRM_for:	TTGGCCTAATTCTTCTTCTGATG
ndhB(97016)HRM_rev:	TTGAAGAGATGAAATATAACCAAGG
ndhF(112349)HRM_for:	TTCGGTTACTTTATCGATCCAC
ndhF(112349)HRM_rev:	TCAGAACCAAAATCCCAACAG
116318R:	GGTACAATCCAAATAATTTATGCAG
116228F:	CCATATGAGATACAGAAGAATAGGC
ndhD(116494)HRM_for:	TTGGATTTCTTATTGCTTTTGC
ndhD(116494)HRM_rev:	CACCGTGGGTGTCAGGTAAC
ndhD(116785)HRM_for:	AGCAATGTACAGCGGTCAAA
ndhD(116785)HRM_rev:	TTCTAATTCCCACATGATGAAAAA
ndhD(117166)HRM_for:	TTGGACCTGGTGTATCTTGTC
ndhD(117166)HRM_rev:	CCAGCAGATATTGGAAAAACAA
ndhG_HRMfor:	TTGCCTGGACCAATACATGA
ndhG_HRMrev:	CACTCCCAGACCCCCTACTA

Supplemental Dataset 1. Text file of the alignment used for the analysis shown in Figure 5

```
ndhB(95225)       TTTTTTGGAAAACTCC
atpF(12707)       TTTTAGCAACAAATCC
ndhD(116281)      CAGCTTCAACATCTCC
rpoA(78691)       TTACACGTGCAAAATC
ndhD(116785)      GCGGTCAAATAGGATC
rpl23(86056)      ATGTCGAATCAGGATC
ndhD(116494)      CTTTTGCCGTCAAATC
ndhB(96579)       TCTTCTATTCTGGTTC
rpoB(25779)       CTATTCTAGTTCTATC
ndhF(112349)      TTACTTCTATTATGTC
ndhB(94999)       TAGCATCTACTATACC
ndhB(95608)       TTCCTTTTTATTTCTC
matK(2931)        TTCTTATATAATTCTC
rps12(69553)      TTCTTTTTGAAATATC
rpoB(23898)       TTCGTGTATATATTTC
ndhD(117166)      ATCTTGTCTTTACCAC
accD(58642)       TTCTTGTTTACTACTC
ndhB(97016)       TTCTGATGATCGATTC
psbZ(35800)       TAATTATTACTTCATC
ndhG(118858)      TAGTTTTTCTGGGATC
rps14(37161)      TGATTCGTCGATCCTC
ndhB(95644)       TAGCTGCTTCAGCTTC
ndhB(95650)       CGAAAGTAGCTGCTTC
ndhB(96698)       CTATCTTTGTAGCTCC
ndhD(116290)      TAATTTATGCAGCTTC
ndhB(96419)       TTTCCCTAGCCCCTTC
rpoB(25992)       ATTCCCTTGGAACTTC
petL(65716)       TGAGTCCCTTCATGCC
psbF(63985)       CTGTACCTACCGTTTC
psbE(64109)       ACAGGCCGTTTTGATC
rps14(37092)      GAAAATTACAATCCCC
clpP(69942)       ACAGAAGCCCAAGCTC
rpoC1(21806)      TACGATTACGAGGTTC
accD(57868)       AAAGAATCGAGCTTTC
```

La stratégie de génétique inverse utilisée a permis l'identification de 6 nouveaux *trans*-facteurs PPR impliqués dans l'édition de 9 sites non précédemment assignées à une protéine PPR. En comptant les facteurs d'édition PPR préalablement identifiés à cette étude et ceux identifiés durant ce projet, on dénombre désormais 15 facteurs d'édition PPR (Figure 23). Le facteur d'édition AtECB2du transcrit *accD* a été exclu de ceux-ci. En effet, après parution de mes résultats, une publication par Tseng et *al.*, 2010 a démontré que les résultats concernant le mutant *AtECB2* (Yu et *al.*, 2009) étaient erronnés.

Chaque site d'édition n'est assigné qu'à une seule protéine PPR suggérant une reconnaissance spécifique des sites d'édition. Cinq de ces facteurs sont impliqués dans l'édition de plusieurs sites de transcrits distincts. L'étude bio-informatique conduite sur les éléments *cis* des différents transcrits reconnus par une même protéine PPR et l'utilisation de données biochimiques précédemment publiées sur la nature des éléments *cis* liés par des *trans*-facteurs nous a permis d'émettre une hypothèse sur le mode de reconnaissance spécifique des sites d'édition chloroplastiques par ces protéines PPR. Toutefois, cette hypothèse devra être testée par des expériences biochimiques.

Une autre information très importante tirée de ces travaux est le nombre de plantes d'*Arabidopsis* mutantes pour l'édition d'un ou plusieurs sites chloroplastiques ne présentant aucun phénotype dans des conditions de culture du laboratoire. Ces plantes expriment des protéines chloroplastiques mutées au niveau d'un acide-aminé très conservé au cours de l'évolution. A l'exception du mutant *otp84*, la fonctionnalité de ces protéines ne semble pas affectée par ces mutations ponctuelles. Contrairement à ce qu'il était pensé, l'édition de certains sites chloroplastiques ne serait pas essentielle à la fonctionnalité des protéines codées. En d'autres termes, les protéines codées par le transcrit édité et non édité pourraient être fonctionnellement équivalentes. Afin d'examiner cette possibilité, nous avons décidé d'étudier de manière plus extensive le mutant d'édition *otp82* en collaboration avec l'équipe du Professeur Toshiharu Shikanai (Université de Kyoto) spécialiste du complexe NAD(P)H déshydrogénase chloroplastique. Ce mutant d'*Arabidopsis* est affecté dans l'édition de deux sites de transcrits chloroplastiques distincts codant des sous-unités du complexe NAD(P)H déshydrogénase. L'impact de ces défauts d'édition sur la fonctionnalité physiologique et biochimique des protéines a été évalué.

4.1.2. PUBLICATION 2: THE PENTATRICOPEPTIDE REPEAT PROTEIN OTP82 IS REQUIRED FOR RNA EDITING OF PLASTID NDHB AND NDHG TRANSCRIPTS, THE PLANT JOURNAL (2010)

the plant journal

SEB SOCIETY FOR EXPERIMENTAL BIOLOGY

The Plant Journal (2010) **61**, 339–349 doi: 10.1111/j.1365-313X.2009.04059.x

The pentatricopeptide repeat protein OTP82 is required for RNA editing of plastid ndhB and ndhG transcripts

Kenji Okuda[1,†], Kamel Hammani[2,†], Sandra K. Tanz[2], Lianwei Peng[1], Yoichiro Fukao[3], Fumiyoshi Myouga[4], Reiko Motohashi[5], Kazuo Shinozaki[4], Ian Small[2] and Toshiharu Shikanai[1,*]

[1]*Department of Botany, Graduate School of Science, Kyoto University, Kyoto 606-8502, Japan,*

[2]*ARC Centre of Excellence in Plant Energy Biology, Molecular and Chemical Sciences Building (M316), University of Western Australia, 35 Stirling Highway, Crawley, Perth 6009 WA, Australia,*

[3]*Plant Science Education Unit, Graduate School of Biological Sciences, Nara Institute of Science and Technology, Takayama, Ikoma, Nara 630-0101, Japan,*

[4]*Plant Genome Network Research Team, RIKEN Plant Science Center, 1-7-22 Yokohama 203-0045, Japan, and*

[5]*Faculty of Agriculture, University of Shizuoka, Shizuoka 422-8529, Japan*

Received 29 August 2009; revised 6 October 2009; accepted 13 October 2009; published online 18 November 2009.
*For correspondence (fax +81 75 753 4257; e-mail shikanai@pmg.bot.kyoto-u.ac.jp).
†These authors contributed equally.

SUMMARY

Several hundred nucleus-encoded factors are required for regulating gene expression in plant organelles. Among them, the most numerous are the members of the pentatricopeptide repeat (PPR) protein family. We found that PPR protein OTP82 is essential for RNA editing of the ndhB-9 and ndhG-1 sites within transcripts encoding subunits of chloroplast NAD(P)H dehydrogenase. Despite the defects in RNA editing, *otp82* did not show any phenotypes in NDH activity, stability or interaction with photosystem I, suggesting that the RNA editing events mediated by OTP82 are functionally silent even though they induce amino acid alterations. In agreement with this result, both sites are partially edited even in the wild type, implying the possibility that a single gene produces heterogeneous proteins that are functionally equivalent. Although only five nucleotides separate the ndhB-8 and ndhB-9 sites, the ndhB-8 site is normally edited in *otp82* mutants, suggesting that both sites are recognized by different PPR proteins. OTP82 falls into the DYW subclass containing conserved C-terminal E and DYW motifs. As in CRR22 and CRR28, the DYW motif present in OTP82 is not essential for RNA editing *in vivo*.

Keywords: RNA editing, pentatricopeptide repeat (PPR) protein, NAD(P)H dehydrogenase, plastid, DYW motif, Arabidopsis.

INTRODUCTION

Plastids were derived from cyanobacteria-like ancestors via endosymbiosis. Current plastid genomes, containing approximately 120 genes required for photosynthesis and housekeeping functions, were shaped after large-scale transfer of genes to the host genomes (Timmis *et al.*, 2004). Accordingly, plant nuclear genomes encode many proteins with prokaryotic origin that retain the ancestral function in plastids (Barkan and Goldschmidt-Clermont, 2000). Eukaryotic phototrophs evolved mechanisms to coordinately regulate the expression of nuclear and plastid genomes. Several hundred nucleus-encoded proteins of eukaryotic origin are estimated to be involved in the regulation of plastid gene expression, based on the current data (Barkan *et al.*, 2007; Schmitz-Linneweber and Small, 2008; Kroeger *et al.*, 2009). However, the functions of many of them remain to be elucidated.

Among these proteins, much of the focus has concentrated recently on the pentatricopeptide repeat (PPR) protein family. PPR proteins form a huge protein family that is particularly prevalent in land plants: 450 members are encoded in the *Arabidopsis thaliana* genome (Lurin *et al.*, 2004; O'Toole *et al.*, 2008). The family members are defined by a tandem array of PPR motifs: each is a highly degenerate unit consisting of 35 amino acids, and is expected to be folded into a pair of antiparallel helices (Small and Peeters, 2000). Most PPR proteins are predicted to be localized in

plastids or mitochondria (Lurin *et al.*, 2004). A growing mass of information indicates that PPR proteins are involved in almost all stages of gene expression in both plastids and mitochondria (Schmitz-Linneweber and Small, 2008). The most probable explanation for their divergent roles is that they are sequence-specific RNA binding adaptors recruiting effector enzymes to the target RNA (Delannoy *et al.*, 2007). The PPR protein family is divided into the P and PLS subfamilies (Lurin *et al.*, 2004): the latter accounts for roughly half of the members in Arabidopsis, and is specific to land plants (O'Toole *et al.*, 2008). Based on the differences in C-terminal motifs, the PLS subfamily is further classified into the PLS, E and DYW subclasses (O'Toole *et al.*, 2008). The expansion of the PLS subfamily in plants is correlated with the specific characteristics of plant organelles, particularly RNA editing (Salone *et al.*, 2007; O'Toole *et al.*, 2008; Rüdinger *et al.*, 2008).

RNA editing is a post-transcriptional process that alters specific cytidine residues to uridine in the mitochondrial and plastid RNA of higher plants (Shikanai, 2006). Thirty-four sites are edited in Arabidopsis plastids (Chateigner-Boutin and Small, 2007), whereas more than 450 editing sites are edited in Arabidopsis mitochondria (Giege and Brennicke, 1999; Bentolila *et al.*, 2008; Zehrmann *et al.*, 2008). *In vivo* approaches using plastid transformation and *in vitro* RNA editing assays clarified that a *cis*-element consisting of fewer than 30 nucleotides surrounding the editing site is essential for site recognition (Chaudhuri and Maliga, 1996; Hirose and Sugiura, 2001). The case is also similar in mitochondria (Takenaka *et al.*, 2004). Our genetic and biochemical studies revealed that the PPR protein CHLORORESPIRATORY REDUCTION 4 (CRR4) acts as the site-specific factor for recognizing the site 1 of RNA editing (ndhD-1) in the plastid *ndhD* transcript encoding a subunit of NAD(P)H dehydrogenase (NDH) (Kotera *et al.*, 2005; Okuda *et al.*, 2006). Subsequently, two PPR proteins, CRR21 and CLB19, were identified as site-specific factors required for the RNA editing of ndhD-2 or RNA editing of *rpoA* and *clpP*, respectively (Okuda *et al.*, 2007; Chateigner-Boutin *et al.*, 2008). CRR4, CRR21 and CLB19 belong to the E subclass (Kotera *et al.*, 2005; Okuda *et al.*, 2007; Chateigner-Boutin *et al.*, 2008). Subsequently, multiple reports indicated that DYW subclass members are also required for RNA editing. The DYW subclass member YS1 is involved in RNA editing of rpoB-1 (Zhou *et al.*, 2008), whereas both CRR22 and CRR28 are involved in the RNA editing of multiple plastid transcripts (Okuda *et al.*, 2009). Likewise, the DYW subclass members RARE1, LPA66 and ECB2 were found to be involved in RNA editing of plastid transcripts (Cai *et al.*, 2009; Robbins *et al.*, 2009; Yu *et al.*, 2009). Similar DYW subclass proteins, MEF1 and OGR1, were also shown to be required for RNA editing in mitochondria of Arabidopsis and rice, respectively (Kim *et al.*, 2009; Zehrmann *et al.*, 2009). Although the involvement of PPR proteins in RNA editing of plant organelles is already beyond any doubt, questions still remain about the molecular mechanisms by which these proteins recognize editing sites, not to mention the unresolved mystery concerning the identity of the editing enzyme itself. Further examples of site-specific factors will be helpful for progressing our understanding of this process. In this context, the best-studied model system is the Arabidopsis plastid, where 34 sites are edited.

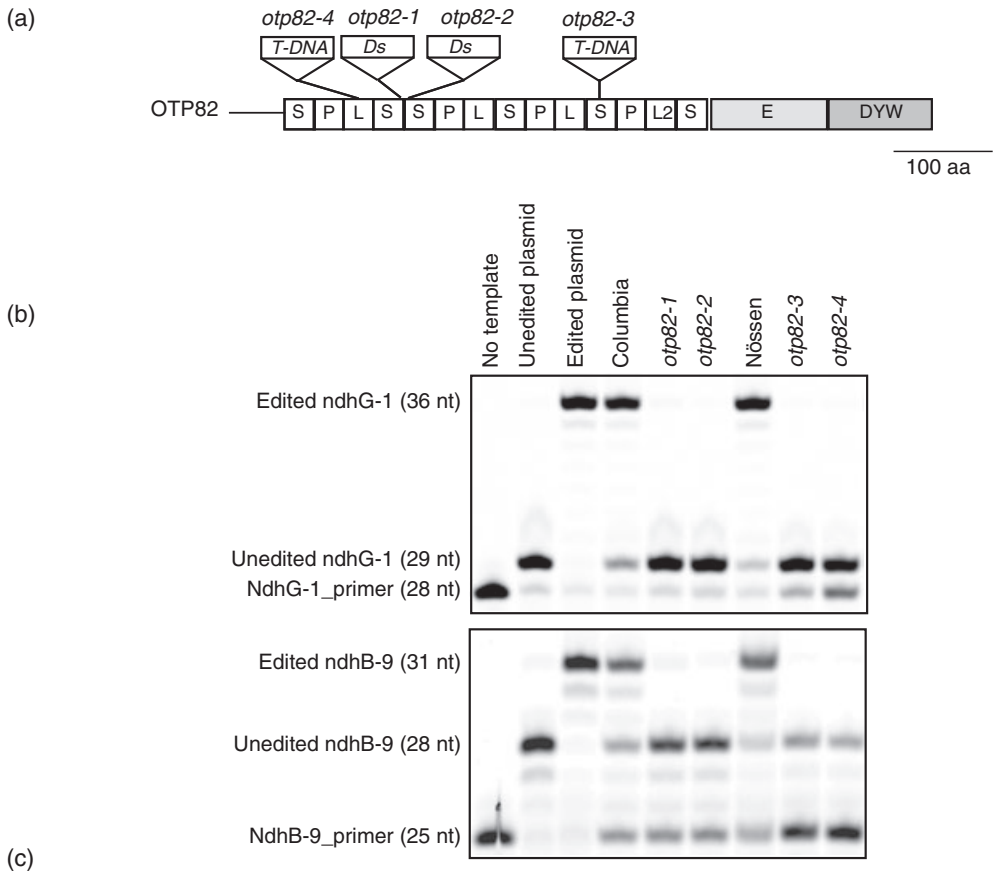
To identify further factors involved in plastid RNA editing, we are performing a comprehensive survey of E and DYW subclass members predicted to target to plastids. Here, we report the identification of a new member of the DYW subclass required for RNA editing of plastid *ndh* transcripts.

RESULTS

At1g08070 (*OTP82*) encodes a DYW member required for multiple RNA editing events in plastids

To clarify the genes involved in plastid RNA editing, we focused on the chlorophyll fluorescence changes after actinic light (AL) illumination, reflecting NDH activity *in vivo* (Shikanai *et al.*, 1998), and identified the mutants defective in RNA editing of plastid *ndh* genes (Kotera *et al.*, 2005; Okuda *et al.*, 2007, 2009). However, the strategy cannot identify mutants impaired in RNA editing in which the defect does not severely influence NDH activity. To identify mutants defective in plastid RNA editing more systematically, we focused on E and DYW members that are predicted to target to plastids. Among them we selected a DYW member, *At1g08070*, that we named *ORGANELLE TRANSCRIPT PROCESSING 82* (*OTP82*) after the molecular phenotype of its mutants (see below). *OTP82* is disrupted by two independent *Ds* insertions in the lines *otp82-1* and *otp82-2* in the Nössen background (Kuromori *et al.*, 2004; Ito *et al.*, 2005) (Figure 1a). In addition, two T-DNA insertion lines, *otp82-3* and *otp82-4*, in the Col-0 background were independently characterized (Figure 1a). The homozygous plants for each mutant allele did not show any obvious visible phenotypes under the standard culture conditions. The *OTP82* gene does not contain introns, and encodes a putative DYW member of the PPR family consisting of 741 amino acids (Figure 1a). Putative orthologs of Arabidopsis *OTP82* were found in several plants, including rice (*Os01g08120*), by the program of POGs/PlantRBP (plant RNA-binding Protein Database) (http://plantrbp.uoregon.edu). OTP82 contains 14 PPR or PPR-like (P, L, L2 and S) motifs, as well as the E and DYW motifs that are characteristic of the DYW subclass (Figure 1a). Figure S1 shows the alignment of E and DYW motifs of OTP82 with those of PPR proteins involved in RNA editing, and with that of CRR2 involved in intergenic RNA cleavage (Hashimoto *et al.*, 2003), indicating the high level of sequence similarity among the family members.

To test the possibility that the *otp82* disruptants are defective in RNA editing, we systematically examined the



ndhG	pU	pE	Nössen	*otp82-1*	*otp82-2*	Columbia	*otp82-3*	*otp82-4*
% editing	0	100	67.2	2.4	–0.3	78.9	0.0	0.1
ndhB	pU	pE	Nössen	*otp82-1*	*otp82-2*	Columbia	*otp82-3*	*otp82-4*
% editing	0	100	71.1	0.8	0.1	86.3	–1.1	–1.8

Figure 1. Editing defects in *otp82* mutants.
(a) Predicted motif structure of OTP82. pentatricopeptide repeat (PPR) (or PPR-like), E and DYW motifs are depicted as boxes with letters. The designation of the P, L and S corresponds to the PPR motif, PPR-like S (for short) motif, and PPR-like L (for long) motif, respectively, proposed by Lurin *et al.* (2004). Sites of *Ds* and T-DNA insertions in mutant alleles are indicated.
(b) Poisoned primer extension assays were conducted on the editing sites ndhG-1 (118 858) and ndhB-9 (95 644). The editing sites are specified relative to the nucleotide sequence of the complete Arabidopsis chloroplast genome (Genbank accession number AP000423). RT-PCR products were obtained with primers surrounding the editing sites, and serve as templates for the extension reaction from a labeled 6-carboxyfluorescein primer that anneals next to the target editing site (a forward poisoned primer extension primer was used for all sites). The extension is stopped by the incorporation of ddGTP at the location of the editing site for unedited molecules, producing a short unedited product. The extension is stopped at the next G/C for the edited molecules, producing a longer edited product.
(c) RNA editing efficiency; pU, plasmid not edited; pE, plasmid edited.

editing status of chloroplast transcripts using a new high-resolution melting screen (Chateigner-Boutin and Small, 2007). Among the 34 RNA editing sites present in Arabidopsis plastids (Chateigner-Boutin and Small, 2007), we identified defects in the RNA editing of ndhB-9 and ndhG-1 in *otp82* (Figure S2). These defects were confirmed by more sensitive poisoned primer extension assays (Figure 1b). The sites at ndhB-9 and ndhG-1 were partially edited in the wild type at efficiencies estimated to be 71.1 and 67.2% in the Nössen background, respectively, and 78.9 and 86.3% in the Col-0 background, respectively (Figure 1c). RNA editing of these two sites was completely impaired in *otp82* mutants, except that *otp82-1* showed residual editing at the ndhB-9 and ndhG-1 sites (efficiency of 0.8 and 2.4%, respectively) (Figure 1c). We also confirmed that the introduction of the wild-type genomic sequence of OTP82 fully restored the editing of ndhB-9 and ndhG-1 (Figure 2), confirming that the defect in RNA editing results from the disruption of *OTP82*. Although only five nucleotides separate the ndhB-8 and ndhB-9 sites, the ndhB-8 site is normally edited in *otp82*

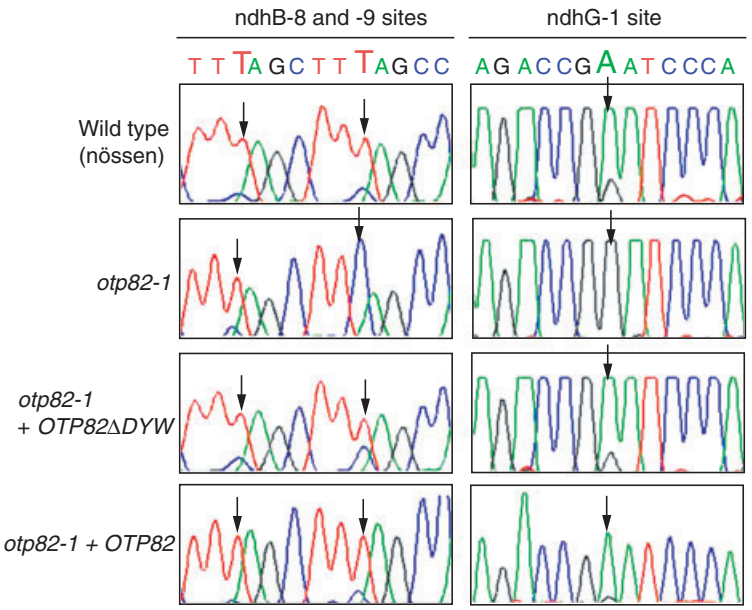


Figure 2. Effect of the deletion of the DYW motif in OTP82.
Nucleotide sequences including the RNA editing sites of ndhB-9 and ndhG-1 are shown as sequence chromatograms. Editing sites of ndhB-8, ndhB-9 and ndhG-1 are indicated by arrows pointing to the corresponding peaks; *otp82-1 + OTP82ΔDYW*, *otp82-1* transformed with *OTP82* lacking the DYW motif; *otp82-1 + OTP82*, *otp82-1* complemented by the introduction of the wild-type genomic *OTP82*.

mutants (Figures 2 and S2), suggesting that these two adjacent RNA editing sites are recognized by different site-specificity factors. All other known sites were also edited correctly in *otp82* mutants (Figure S2).

Defects in RNA editing may be secondarily caused by aberrant RNA processing. To test this possibility, the levels and patterns of transcripts were analyzed by RNA gel blot. Figure 3 shows that there are no obvious alterations in *ndhB* and *ndhG* transcripts in *otp82*. We conclude that *otp82* is primarily defective in multiple RNA editing events in *ndhB* and *ndhG* transcripts.

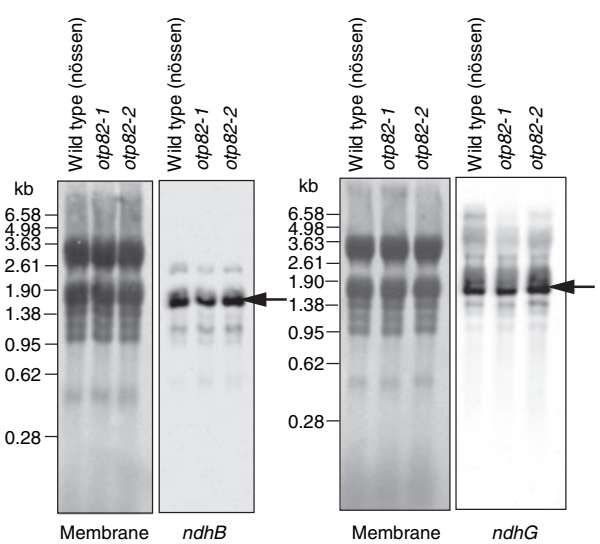


Figure 3. Transcript profiles of genes with the editing defects in *otp82*.
A 10 μg portion of RNA from 15-day-old seedlings was loaded onto formaldehyde agarose gels and transferred onto a membrane. Hybridizations were performed under high-stringency conditions using antisense RNA probes for the genes specified for each blot. The arrows point at the mature transcript. The sizes of RNA markers are shown in kilobases. Ribosomal RNA stained with methylene blue on the membranes is shown as a loading control (membrane).

OTP82 is localized to plastids

OTP82 is predicted to be targeted to chloroplasts by TargetP (Emanuelsson *et al.*, 2000) and Predotar (Small *et al.*, 2004). To confirm the localization experimentally, we constructed a chimeric gene encoding a fusion protein consisting of the N-terminal 100 amino acids of OTP82 and green-fluorescent protein (GFP) under the control of the Cauliflower mosaic virus *35S* promoter. As the cleavage site of the transit peptide was unknown, we used the N-terminal 100 amino acids including the first PPR motif to ensure that the complete targeting information was included. The plasmid containing the chimeric gene was introduced into wild-type Arabidopsis cells by bombardment. Analysis of GFP fluorescence in transformed cells revealed that the fluorescence co-localized with a chloroplast marker: the fusion of red-fluorescent protein (RFP) with the small subunit of Arabidopsis ribulose biphosphate carboxylase (SSU) (Carrie *et al.*, 2007) (Figure 4). We conclude that OTP82 is localized to plastids, consistent with the *otp82* phenotype specifically defective in plastid RNA editing.

RNA editing mediated by OTP82 is not required for NDH activity, stability and the supercomplex formation with PSI

The RNA editing of ndhB-9 and ndhG converts Ser279 of NdhB (uCa) to leucine (uUa) and Ser17 of NdhG (uCc) to phenylalanine (uUc), respectively (Figure 5d). The alterations between amino acids with different characters might be expected to have a strong impact on the encoded protein. However, both the ndhB-9 and ndhG-1 sites are only partially edited, even in the wild type (Figure 1c), implying that RNA editing might be dispensable for the protein function.

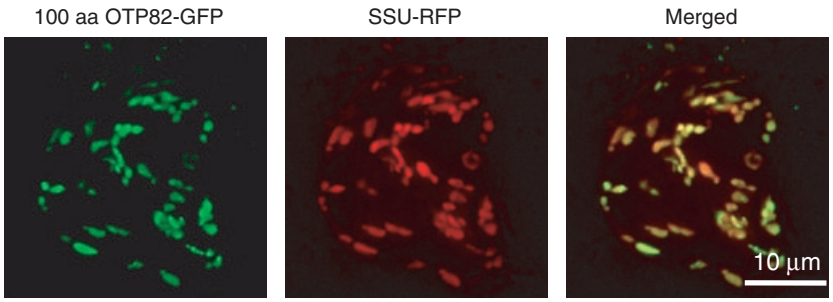


Figure 4. Analysis of subcellular localization of the OTP82 protein.
A chimeric protein consisting of the first 100 amino acids of *At1g08070* fused to GFP was targeted to chloroplasts in transformed Arabidopsis cells. Fluorescence of the green fluorescent protein (100 aa OTP82-GFP), red fluorescence of the red fluorescent protein in fusion with the targeting sequence of the small subunit of the ribulose biphosphate carboxylase (SSU-RFP), and the overlay of the two fluorescence images are shown.

First, we analyzed the level of NDH complex in *otp82*. As the NDH complex is unstable without the membrane subunits, NdhA, NdhB, NdhC, NdhD, NdhE, NdhF and NdhG, as reported for mutants defective in NdhB, NdhD or NdhF (Peng *et al.*, 2008), antibodies against NdhH and NdhL can be used to monitor the accumulation of the NDH complex, and consequently that of NdhB and also probably that of NdhG. In *otp82*, neither NdhH nor NdhL levels were affected (Figure 5a). As editing at the ndhB-9 and ndhG-1 site was below the detection limit in *otp82-2* (Figure 1c), all the NdhB and NdhG accumulating in the mutants was probably translated from unedited RNA. These results suggest that Leu279 in NdhB and Phe17 in NdhG are unlikely to be essential for stabilizing the NDH complex.

Chloroplast NDH interacts with the photosystem I (PSI) complex to form a supercomplex (NDH-PSI) in Arabidopsis (Peng *et al.*, 2008). It is possible that Leu279 of NdhB and Phe17 of NdhG are required for the supercomplex formation, although they are not essential for stabilizing NDH. To test this possibility, we analyzed the level of the NDH-PSI supercomplex in *otp82* by Blue Native (BN)-PAGE. The high molecular weight green band (band I) corresponds to the NDH-PSI supercomplex in the wild type, and is missing in the *crr4-3* mutant defective in *ndhD* expression, and is greatly reduced in the *crr2-2* mutant defective in *ndhB* expression (Figure 5b), consistent with previous results (Peng *et al.*, 2008). The BN-PAGE detected the same level of band I in *otp82* as in the wild type (Figure 5b). The result indicates that Leu279 of NdhB and Phe17 of NdhG are not essential for the interaction between NDH and the PSI complex.

The mutants *crr22* and *crr28*, both defective in the RNA editing of *ndhB* transcripts, showed altered NDH activity, despite the stable accumulation of the NDH complex (Okuda *et al.*, 2009). It is possible that Leu279 of NdhB and Phe17 of NdhG are essential for NDH activity, rather than for its stability. To examine this possibility, we analyzed NDH activity in *otp82*. The chloroplast NDH complex catalyzes electron donation to plastoquinone from the stromal electron pool, and its activity can be monitored as a transient increase in chlorophyll fluorescence, reflecting plastoquinone reduction after turning off AL (Shikanai *et al.*, 1998). Figure 5c shows a typical trace of the chlorophyll fluorescence level in the wild type. In *otp82*, the post-illumination increase of chlorophyll fluorescence was not suppressed (Figure 5c), indicating that NDH activity was not affected in *otp82*. These results suggest that Leu279 of NdhB and Phe17 of NdhG are not required for NDH activity. This result is consistent with the fact that the mutants were not identified in our chlorophyll fluorescence screen looking for mutants affected in NDH activity (Okuda *et al.*, 2007).

Taken together, the Ser279 → Leu279 conversion of NdhB and Ser17 → Phe17 conversion of NdhG are unlikely to be essential for the activity or stability of the NDH complex, or even for NDH-PSI supercomplex formation, although the editing of ndhB-9 and ndhG-1 results in the restoration of codons for amino acids conserved in other land plants (Figure 5d) (Lutz and Maliga, 2001; Tsudzuki *et al.*, 2001), except that ndhG-1 editing in pea converts serine to leucine (Inada *et al.*, 2004). The results are consistent with the fact that the sites are only partially edited even in the wild type (Figure 1c).

Mass spectrometric analysis of the PSI-NDH supercomplex in the wild type and in *otp82*

The observations above suggest that the proteins translated from edited and unedited mRNA are functionally equivalent, and that both proteins are incorporated into the NDH complex. To examine this possibility, we analyzed the NDH-PSI supercomplex by mass spectrometry (MS). Consistent with the previous MS analysis (Peng *et al.*, 2009), we failed to detect NdhG in both the wild type and the *otp82* mutant, probably because of the hydrophobic nature of NdhG. In the wild type, the MS detected only NdhB derived from the *ndhB* transcript in which the ndhB-9 site was edited (Table 1). The ndhB-8 site is separated by only five nucleotides and the site was also edited, resulting in a Ser → Leu substitution in the detected peptide (Table 1). We also analyzed the NDH-PSI

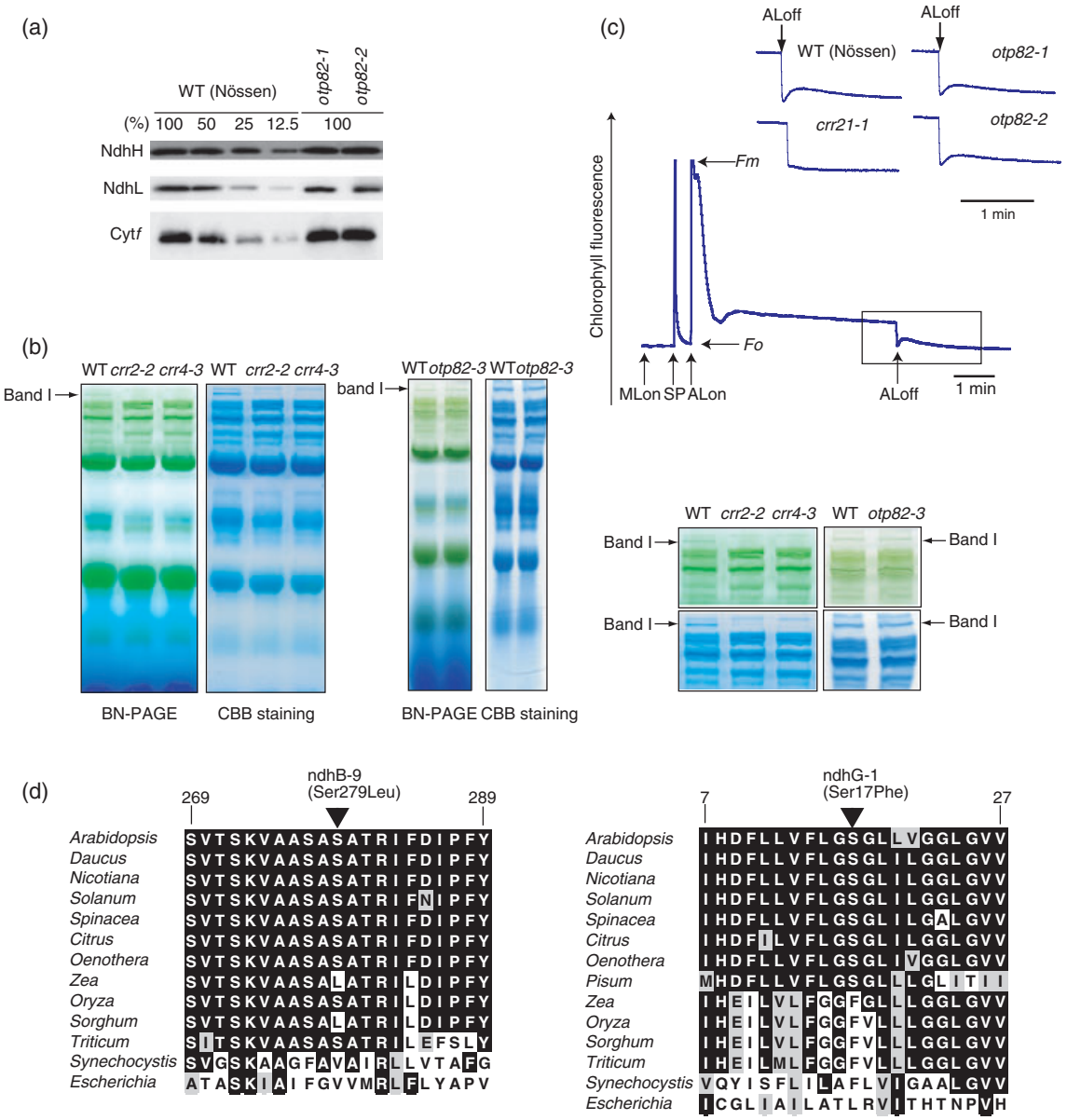


Figure 5. Analyses of the NAD(P)H dehydrogenase (NDH) complex in *otp82*.
(a) Immunodetection of NDH subunits NdhH and NdhL, and a subunit of the cytochrome (Cyt) b_6f complex, Cyt*f*. The proteins were extracted from thylakoid membrane fractions. Lanes were loaded with protein samples corresponding to 0.5 μg chlorophyll for Cyt*f*, 1 μg chlorophyll for NdhL and 5.0 μg chlorophyll for NdhH (100%), and the series of dilutions indicated.
(b) BN-PAGE analysis of thylakoid protein complex isolated from the wild type and from *otp82*. Thylakoid membrane complexes isolated from the wild type, *crr2-2*, *crr4-3* and *otp82-3* were solubilized and separated by BN-PAGE, and then stained with Coomassie Briliant Blue, and are shown on the left panel. The top part of the gels is closed up on the right panel. Positions of band I are indicated by arrows.
(c) Monitoring of NDH activity by using chlorophyll fluorescence analysis. The curve shows a typical trace of chlorophyll fluorescence in the wild type (WT). Leaves were exposed to actinic light (AL) (50 μmol of photons m^{-2} s^{-1}) for 5 min. AL was turned off and the subsequent change in chlorophyll fluorescence level was monitored. The transient increase in chlorophyll fluorescence is caused by the plastquinone reduction based on NDH activity. Insets are magnified traces from the boxed area. The fluorescence levels were normalized by the maximum fluorescence at closed PSII centers in the dark (*Fm*) levels; ML, measuring light; SP, a saturating pulse of white light.
(d) Partial sequence alignments of NdhB and NdhG around the amino acids affected by RNA editing. Arabidopsis NdhB and NdhG proteins were aligned with their homologs from other species. The alignment was performed using the CLUSTALW program (version 1.87, DNA Data Bank of Japan). Amino acids that are fully or semiconserved are shaded black and gray, respectively. Numbers indicate amino acid positions in the protein. The arrows above the sequences indicate the positions of edited codons.

supercomplex in *otp82-3*, where the ndhB-9 site was not edited at all. The MS identified oligopeptides common to NdhB translated from both edited and unedited RNAs, but did not identify the peptide reflecting the RNA editing status of ndhB-9 (Table 1). The oligopeptide derived from unedited RNA (VAALA**S**ATR) may not be detected by MS for technical

Table 1 Identified polypeptides from linear ion-trap triple quadrupole (LTQ)-Orbitrap mass analysis of the NDH-PSI supercomplex against NdhB amino acid sequences

Samples	Polypeptides
Wild type	VAA**LA**L**A**TR[a]
	DYAGLYTK
	IIKLLMTGR
	NQEITPHMR
	TGTDNIRDYAGLYTK
otp82	ISPLR
	DYAGLYTK
	IIKLLMTGR
	NQEITPHMR

[a]Specific peptide derived from NdhB that is translated from the *ndhB* transcript edited at sites 8 and 9.

reasons, although the corresponding peptide translated from edited RNA (VAALA**L**ATR) was detected in the wild type. Thus, it is still possible that two versions of NdhB translated from edited and unedited mRNA are incorporated into the NDH complex.

The DYW motif of OTP82 is not essential for RNA editing *in vivo*

The DYW motif was proposed as the catalytic site of organelle RNA editing, based on the apparent similarity to the active site of cytidine deaminases, including the human RNA editing enzyme APOBEC1, and the phylogenetic correlation between the occurrence of RNA editing and the presence of DYW motifs (Salone *et al.*, 2007; Rüdinger *et al.*, 2008). However, the DYW motifs of CRR22 and CRR28 were shown to be dispensable for RNA editing *in vivo*, even though that of CRR2 was essential for RNA cleavage *in vivo* (Okuda *et al.*, 2009). To examine whether this finding can be extended to other DYW motifs present in the PPR proteins required for RNA editing, a truncated version of OTP82 lacking its DYW motif was introduced into *otp82*. Consistent with our previous result, the mutant version of OTP82 restored the RNA editing activity of ndhB-9 and ndhG-1 to the wild-type level (Figure 2, *otp82-1* + *OTP82ΔDYW*). This result indicates that the DYW motif present in OTP82 is not essential for RNA editing *in vivo*.

DISCUSSION

The present study shows that OTP82 is a site-specific factor required for two RNA editing events in plastids. Like all other editing factors identified so far, OTP82 belongs to the PLS family of PPR proteins that is specific to land plants. Of 34 editing sites known in Arabidopsis plastids, three E and seven DYW subclass members required for 14 sites have been determined (Table S1). In total in Arabidopsis, 20 E and 24 DYW proteins are predicted to be targeted to plastids (Lurin *et al.*, 2004). Even though some of these proteins (such as CRR2) may have other roles (Hashimoto *et al.*, 2003; Nakamura and Sugita, 2008; Okuda *et al.*, 2009), the number of E and DYW proteins appears more than sufficient to deal with all editing sites present in plastids, especially considering that a single PPR protein can recognize multiple sites (Chateigner-Boutin *et al.*, 2008; Okuda *et al.*, 2009; this work). This case appears to be similar in mitochondria, where the DYW members MEF1 and OGR1 were identified as editing factors (Zehrmann *et al.*, 2009; Kim *et al.*, 2009). The number of the PLS subfamily members predicted to localize to mitochondria correlates well with the number of editing sites in mitochondria (Zehrmann *et al.*, 2009). The E and DYW domains are highly conserved in the PPR proteins involved in RNA editing, even between plastid and mitochondrial proteins (Figure S1). If these C-terminal motifs, especially the E motif, are the binding sites for the still-unidentified editing enzyme (Okuda *et al.*, 2007), similar or identical editing enzymes are likely to be recruited in both organelles.

In vitro RNA editing assays have suggested that the *cis*-elements recognized by the same *trans*-factor show high sequence identity (a 60% identity in nucleotide sequence) (Kobayashi *et al.*, 2007). Consistent with this idea, thirteen nucleotides are conserved within the 35 nucleotides surrounding the ndhB-9 and ndhG-1 sites that are recognized by OTP82 (Figure 6). The ndhB-8 editing site that is 6-bp upstream from the ndhB-9 site increases the number of identical nucleotides from 13 to 14 (Figure 6). Similarly, the nucleotides surrounding the two editing sites recognized by CRR28 are highly conserved (Okuda *et al.*, 2009). However, it is not simple to interpret the recognition of multiple *cis*-elements by a single PPR protein. The sets of putative *cis*-elements recognized by CLB19, CRR22, MEF1 and OGR1 show no obvious similarity (Chateigner-Boutin *et al.*, 2008; Kim *et al.*, 2009; Okuda *et al.*, 2009; Zehrmann *et al.*, 2009). Different sets of PPR motifs within a single protein may independently recognize unrelated *cis*-sequences.

The ndhB-8 and ndhB-9 sites are separated by only five nucleotides in Arabidopsis (Figure 6). Are the closely adjacent sites recognized by distinct proteins, or by a common factor? In the tobacco *ndhB* gene a pair of editing sites, IV and V, are separated by only eight nucleotides. Both sites were independently edited in

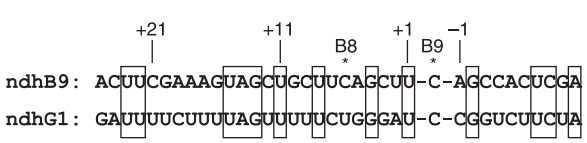


Figure 6. Comparison of the nucleotide sequences in the region surrounding the editing sites affected in *otp82*.
Sequence alignment of the region surrounding the ndhB-9 and ndhG-1 editing sites. The alignment includes the sequence from −30 to +10 around the edited C (asterisk), with identical nucleotides shown in boxes. The ndhB-8 site is indicated by the asterisk.

partially edited transcripts, implying the existence of two independent site-specificity factors (Bock *et al.*, 1997). Both ndhB-8 and ndhB-9 sites are edited in *Nicotiana tabaccum* and *Nicotiana sylvestris*, whereas the ndhB-8 site is not edited in *Nicotiana tomentosiformis*, implying that distinct site-specificity factors are involved in the two editing events (Sasaki *et al.*, 2003). Consistent with this idea, *otp82* lacks only the editing activity of ndhB-9, and the ndhB-8 site is edited as in the wild type (Figure 2). Our result provides evidence that the ndhB-8 and ndhB-9 sites are recognized by distinct proteins, despite their close proximity.

otp82 is defective in RNA editing of two sites, ndhB-9 and ndhG-1, leading to amino acid alterations. However, the amino acid alterations induced by these two RNA editing events are functionally silent, at least under the culture conditions used in this study. Even leaky defects in NDH activity cause a severe growth phenotype in the *pgr5* mutant background impaired in the main pathway of PSI cyclic electron transport (Munekage *et al.*, 2004; Peng *et al.*, 2009). To test the possibility that NDH activity is subtly affected in *otp82*, we generated the double mutant *otp82-2 pgr5*. The phenotype of *otp82-2 pgr5* was identical to that of *pgr5*, supporting our idea that NDH activity was not affected in *otp82* (data not shown).

The ndhB-9 and ndhG-1 sites are not edited completely, even in wild-type plants (Figure 1c), suggesting the possibility that proteins translated from edited and unedited mRNA are incorporated into the NDH complex. However, our MS analysis preferentially detected NdhB that originated from edited RNA in the wild type (Table 1). We cannot conclude that the NDH-PSI supercomplex solely consists of NdhB originating from edited mRNA, despite the high mass accuracy, high resolution and high sensitivity of linear ion-trap triple quadrupole (LTQ)-Orbitrap MS.

RNA editing at ndhB-9 and ndhG-1 may be physiologically significant under certain conditions, because the amino acids restored by editing are highly conserved in plants, presumably by natural selection. We reported that CRR22 is involved in RNA editing of rpoB-3, and that the resulting amino acid alteration was also functionally silent (Okuda *et al.*, 2009). At this site a trace level of unedited transcript was also detected even in the wild type (Okuda *et al.*, 2009). We also observed that editing of the ndhD-1 site that creates the translational initiation site of *ndhD* is only partial (Kotera *et al.*, 2005), and that the efficiency of this RNA editing event can be determined by CRR4 activity (Okuda *et al.*, 2008). Partial editing seems to be restricted to cases where the editing event is not essential for protein expression or activity. As yet we cannot conclude whether this is simply a result of lower selection pressure for efficient editing, or whether there is active selection for partial editing to generate genetic diversity.

EXPERIMENTAL PROCEDURES

Plant materials

Arabidopsis thaliana ecotypes Nössen and Col-0 were used in this study. *otp82-1* (11-6902-1, insertion site +465 bp, Nössen) and *otp82-2* (54-2824-1, insertion site +793 bp, Nössen) carry *Ds* transposon insertions (Kuromori *et al.*, 2004; Ito *et al.*, 2005). *otp82-3* (SALK_027812, Col-0) and *otp82-4* (SAIL_851_G04, Col-0) are T-DNA insertion mutants obtained from the ABRC Stock Center (http://www.biosci.ohio-state.edu/pcmb/Facilities/abrc/abrchome.htm).

Chlorophyll fluorescence analysis

Chlorophyll fluorescence was measured using a MINI-PAM portable chlorophyll fluorometer (Waltz, http://www.walz.com). The transient increase in chlorophyll fluorescence after turning off AL was monitored as previously described (Shikanai *et al.*, 1998).

Construction of GFP fusion proteins to analyze targeting

The first 300 bp of the coding sequence of *At1g08070* was amplified using Phusion DNA polymerase (Finnzymes, http://www.finnzymes.fi) with primers z0797 and z0798 containing the *attB* sites for Gateway® cloning according to the manufacturer's instructions (Invitrogen, http://www.invitrogen.com). The GFP vector used was the same as that previously described (Carrie *et al.*, 2009). The chloroplast targeting marker consists of RFP fused to SSU (Carrie *et al.*, 2009). The primers are listed in Table S2.

Subcellular localization

Biolistic transformations of GFP and RFP constructs were performed on Arabidopsis cell culture, as previously reported (Carrie *et al.*, 2007). The GFP construct and the chloroplast RFP marker (5 μg each) were co-precipitated onto gold particles and transformed using the biolistic PDS-1000/He system (Bio-Rad, http://www.bio-rad.com). Particles were bombarded onto 2 ml of Arabidopsis cell suspension resting on filter paper on osmoticum plates. After bombardment, the cells were placed in the dark at 22°C for 24 h. Observation of transient GFP and RFP expression was performed using an Olympus BX61 fluorescence microscope (Olympus, http://www.olympus.com) with excitation wavelengths of 460/480 nm (GFP) and 535/555 nm (RFP), and with emission wavelengths of 495–540 nm (GFP) and 570–625 nm (RFP). Subsequent images were captured using CELL® imaging software, as previously described (Carrie *et al.*, 2007; Murcha *et al.*, 2007).

Analysis of RNA editing

A high-resolution melting analysis of amplicons was performed as previously described (Chateigner-Boutin and Small, 2007), except that the primers used for the PCR were designed to give shorter amplicons than in the previous study. Poisoned primer extension of reverse transcription (RT)-PCR products was performed as previously described (Chateigner-Boutin and Small, 2007). RNA editing was also analyzed by directly sequencing the RT-PCR products, including editing sites, as previously described (Okuda *et al.*, 2009). The primers are listed in Table S2.

Plant transformation

For complementation of the *otp82* mutation, the wild-type genomic sequence surrounded by 5′-GTACAAGATCGGAAGAGC-3′ and 5′-CTACCAGTAGTCATTGCAG-3′ was cloned in pGEW-NB1 vector. For the expression of OTP82 truncated in the DYW motif, the wild-type genomic sequence surrounded by 5′-GTACAAGATCGGAAGAGC-3′

and 5′-CTACTCTAGTAACACCTCCATT-3′ was cloned in the pGEW-NB1 vector. The resultant plasmids were introduced into *otp82-1* via *Agrobacterim tumefaciens* strain ASE.

RNA preparation and RNA gel blot analysis

Total RNA from leaves of 15-day-old plantlets was isolated using TRIzol reagent (Invitrogen), as recommended in the manufacturer's instructions. Fifteen micrograms of RNA was fractionated on 1.2% (w/v) formaldehyde agarose gels, and then transferred onto Hybond N+ nylon membranes (GE Healthcare, http://www.gehealthcare.com). RNA integrity, loading and transfer were checked by staining the membrane with methylene blue. RNA probes were internally labeled with biotinylated cytidine by transcription of PCR products cloned in pGEM-T Easy vector (Promega, http://www.promega.com). The primers used for the PCR were ndhB.AT.rev2 and AndhB, corresponding to the second exon of *ndhB*, and ndhG.AT.for and ndhG.AT.rev for *ndhG*. Clones with inserts in antisense orientation were amplified by PCR using the forward primer and M13/pUC reverse primer. The PCR products served as a template for *in vitro* transcription with SP6 polymerase, following the manufacturer's instructions (Maxiscript; Ambion, http://www.ambion.com). Prehybridization of the membrane was carried out for 1 h in hybridization buffer containing 5 x SSC, 50% (v/v) formamide, 0.5% SDS and 100 μg ml^{-1} heparin, at 68°C. Hybridization with RNA probes was carried out in the same buffer overnight at 68°C, followed by three 15-min washes at 25°C in 1 x SSC/0.5% SDS, and two washes at 60°C in 0.1 x SSC/0.1% SDS for, respectively, 20 min and 1 h. Signal detection was performed using the Chemiluminescent Nucleic Acid Detection Module (Pierce, http://www.piercenet.com), read in an ImageQuant-RT ECL (GE Healthcare). The primers are listed in Table S2.

Immunoblot analysis

Chloroplasts were isolated from the leaves of 4-week-old plants as previously described (Okuda *et al.*, 2007). Samples were normalized by measuring chlorophyll concentration. The protein samples were separated by 12.5% SDS–PAGE. After electrophoresis, the proteins were transferred onto a Hybond-P membrane (GE Healthcare) and incubated with specific antibodies. The signals were detected using an ECL Advance Western Blotting Detection Kit (for NdhH) (GE Healthecare) or an ECL Plus Western Blotting Detection Kit (for the others) (GE Healthcare), and were visualized by an LAS1000 chemiluminescence analyzer (Fuji Film, http://www.fujifilm.com).

Thylakoid membrane preparation and BN-PAGE

Chloroplasts were isolated as previously described (Okuda *et al.*, 2007) and osmotically ruptured in buffer containing 20 mM HEPES/KOH (pH 7.6), 5 mM $MgCl_2$ and 2.5 mM EDTA. Thylakoid membranes were pelleted by centrifugation (7700 ***g*** for 3 min) and resuspended in the same buffer.

BN-PAGE was performed as previously described (Peng *et al.*, 2006) with some minor modifications. The freshly isolated thylakoid membranes were gently washed twice with buffer containing 25 mM BisTris–HCl (pH 7.0), 20% glycerol, and solubilized in 25 mM BisTris–HCl (pH 7.0), 20% glycerol, 1% *N*-dodecyl-β-D-maltoside, at final chlorophyll concentration of 1 mg ml^{-1}. After incubation on ice for 10 min, the supernatants were supplemented with a 1/10 volume of BN sample buffer [100 mM BisTris–HCl, pH 7.0, 5% Serva blue G, 0.5 M 6-amino-*N*-caproic acid, 30% sucrose (w/v)]. Thylakoid protein complexes were separated by 5–12% gradient BN-PAGE in 0.75-mm-thick gels connected to a circulating cooler.

Peptide preparation for MS/MS analysis

Thylakoid membrane complexes isolated from wild-type and *otp82* mutant plants were solubilized and separated by BN-PAGE. Band I (described in Peng *et al.*, 2008) was excised from the gel. The excised band was treated twice with 25 mM ammonium bicarbonate in 30% (v/v) acetonitrile for 10 min and 100% (v/v) acetonitrile for 15 min, and then dried in a vacuum concentrator. The dried gel pieces were treated with 0.01 mg ml^{-1} trypsin (sequence grade; Promega)/50 mM ammonium bicarbonate, and then incubated at 37°C for 16 h. The digested peptides in the gel pieces were recovered twice with 20 μl 5% (v/v) formic acid/50% (v/v) acetonitrile. The extracted peptides were combined and then dried in a vacuum concentrator.

Mass spectrometric analysis and database searching

LC-MS/MS analyses were performed on an LTQ-Orbitrap XL-HTC-PAL system. Trypsin-digested peptides were loaded on the column (ø75 μm, 15 cm; L-Column; CERI, http://www.cerij.or.jp/ceri_en/otoiawase/tokyo.html) by using a Paradigm MS4 HPLC pump (Michrom BioResources, http://www.michrom.com) and an HTC-PAL autosampler (CTC Analytics), and were eluted by a gradient of 5–45% (v/v) acetonitrile in 0.1% (v/v) formic acid over 70 min. The eluted peptides were introduced directly into the LTQ-Orbitrap XL MS at a flow rate of 300 nl min^{-1}, and with a spray voltage of 2.0 kV. The range of MS scan was *m/z* 450–1500, and the top three peaks were analyzed by MS/MS analysis. MS/MS spectra were compared by MASCOT 2.2 against NdhB amino acid sequences (Figure S3) with the following search parameters: set-off threshold at 0.05 in the ion score cut-off; peptide tolerance, 10 ppm; MS/MS tolerance, ±0.8 Da; peptide charge, 2+ or 3+; trypsin as enzyme allowing up to one missed cleavage; carboxymethylation on cysteines as a fixed modification and oxidation on methionine as a variable modification.

ACKNOWLEDGEMENTS

We thank Asako Tahara (Kyoto University, Kyoto, Japan) for skilled technical support. We also thank Tsuyoshi Endo (Kyoto University, Kyoto, Japan) for giving us antibodies. We are also grateful to Tsuyoshi Nakagawa (Shimane University, Matsue, Japan) for giving us the binary vector, pGWB-NB1. This work was supported by: Grant-in-Aid 16085206 for Scientific Research on Priority Areas; Grant 17GS0316 for Creative Science Research from the Ministry of Education, Culture, Sports, Science, and Technology of Japan; a grant from the Ministry of Agriculture, Forestry and Fisheries of Japan (Genomics for Agricultural Innovation, GPN0008); a grant from the Australian Research Council (CE0561495); an International Science Linkages grant (CG120098) from the Australian Government Department of Innovation, Industry, Science and Research; and scholarships from the Region of Alsace and the University of Western Australia. KH is the holder of a Lavoisier Fellowship and IS is a West Australian State Premier's Fellow.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Figure S1. Comparison of the E and DYW motifs among YS1, CRR22, CRR28, RARE1, LPA66, ECB2, OTP82, OGR1, MEF1 and CRR2.

Figure S2. High-resolution melting screen of *otp82* mutants.

Figure S3. Sequence data for mass spectrometric analysis.

Table S1. A table summarizing the pentatricopeptide repeat (PPR) proteins required for RNA editing in Arabidopsis plastids.

Table S2. Oligonucleotide primers used in the study.

Please note: As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

REFERENCES

Barkan, A. and Goldschmidt-Clermont, M. (2000) Participation of nuclear genes in chloroplast gene expression. *Biochimie*, **82**, 559–572.

Barkan, A., Klipcan, L., Ostersetzer, O., Kawamura, T., Asakura, Y. and Watkins, K.P. (2007) The CRM domain: an RNA binding module derived from an ancient ribosome-associated protein. *RNA*, **13**, 55–64.

Bentolila, S., Elliott, L.E. and Hanson, M.R. (2008) Genetic architecture of mitochondrial editing in *Arabidopsis thaliana*. *Genetics*, **178**, 1693–1708.

Bock, R., Hermann, M. and Fuchs, M. (1997) Identification of critical nucleotide positions for plastid RNA editing site recognition. *RNA*, **3**, 1194–1200.

Cai, W., Ji, D., Peng, L., Guo, J., Ma, J., Zou, M., Lu, C. and Zhang, L. (2009) LPA66 is required for editing *psbF* chloroplast transcripts in *Arabidopsis*. *Plant Physiol.* **150**, 1260–1271.

Carrie, C., Murcha, M.W., Millar, A.H., Smith, S.M. and Whelan, J. (2007) Nine 3-ketoacyl-CoA thiolases (KATs) and acetoacetyl-CoA thiolases (ACATs) encoded by five genes in *Arabidopsis thaliana* are targeted either to peroxisomes or cytosol but not mitochondria. *Plant Mol. Biol.* **63**, 97–108.

Carrie, C., Kühn, K., Murcha, M.W., Duncan, O., Small, I.D., O'Toole, N. and Whelan, J. (2009) Approaches to defining dual-targeted proteins in *Arabidopsis*. *Plant J.* **57**, 1128–1139.

Chateigner-Boutin, A.L. and Small, I. (2007) A rapid high-throughput method for the detection and quantification of RNA editing based on high-resolution melting of amplicons. *Nucleic Acids Res.* **35**, e114.

Chateigner-Boutin, A.L., Ramos-Vega, M., Guevara-Garcia, A. *et al.* (2008) CLB19, a pentatricopeptide repeat protein required for editing of *rpoA* and *clpP* chloroplast transcripts. *Plant J.* **56**, 590–602.

Chaudhuri, S. and Maliga, P. (1996) Sequences directing C to U editing of the plastid *psbL* mRNA are located within a 22 nucleotide segment spanning the editing site. *EMBO J.* **15**, 5958–5964.

Delannoy, E., Stanley, W.A., Bond, C.S. and Small, I.D. (2007) Pentatricopeptide repeat (PPR) proteins as sequence-specificity factors in post-transcriptional processes in organelles. *Biochem. Soc. Trans.* **35**, 1643–1647.

Emanuelsson, O., Nielsen, H., Brunak, S. and von Heijne, G. (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol.* **300**, 1005–1016.

Giege, P. and Brennicke, A. (1999) RNA editing in *Arabidopsis* mitochondria effects 441 C to U changes in ORFs. *Proc. Natl Acad. Sci. USA*, **96**, 15324–15329.

Hashimoto, M., Endo, T., Peltier, G., Tasaka, M. and Shikanai, T. (2003) A nucleus-encoded factor, CRR2, is essential for the expression of chloroplast *ndhB* in *Arabidopsis*. *Plant J.* **36**, 541–549.

Hirose, T. and Sugiura, M. (2001) Involvement of site-specific *trans*-acting factor and a common RNA-binding protein in the editing of chloroplast mRNAs: development of a chloroplast *in vitro* RNA editing system. *EMBO J.* **20**, 1144–1152.

Inada, M., Sasaki, T., Yukawa, M., Tsudzuki, T. and Sugiura, M. (2004) A systematic search for RNA editing sites in pea chloroplasts: an editing event causes diversification from the evolutionarily conserved amino acid sequence. *Plant Cell Physiol.* **45**, 1615–1622.

Ito, T., Motohashi, R., Kuromori, T., Noutoshi, Y., Seki, M., Kamiya, A., Mizukado, S., Sakurai, T. and Shinozaki, K. (2005) A resource of 5,814 dissociation transposon-tagged and sequence-indexed lines of *Arabidopsis* transposed from start loci on chromosome 5. *Plant Cell Physiol.* **46**, 1149–1153.

Kim, S.R., Yang, J.I., Moon, S., Ryu, C.H., An, K., Kim, K.M., Yim, J. and An, G. (2009) Rice OGR1 encodes a pentatricopeptide repeat-DYW protein and is essential for RNA editing in mitochondria. *Plant J.* **59**, 738–749.

Kobayashi, Y., Matsuo, M., Sakamoto, A., Wakasugi, T., Yamada, K. and Obokata, J. (2007) Two RNA editing sites with *cis*-elements of moderate sequence identity are recognized by an identical site-recognition protein in tobacco chloroplasts. *Nucleic Acids Res.* **36**, 311–318.

Kotera, E., Tasaka, M. and Shikanai, T. (2005) A pentatricopeptide repeat protein is essential for RNA editing in chloroplasts. *Nature*, **433**, 326–330.

Kroeger, T.S., Watkins, K.P., Friso, G., van Wijk, K.J. and Barkan, A. (2009) A plant-specific RNA binding domain revealed through analysis of chloroplast group II intron splicing. *Proc. Natl Acad. Sci. USA*, **106**, 4537–4542.

Kuromori, T., Hirayama, T., Kiyosue, Y., Takabe, H., Mizukado, S., Sakurai, T., Akiyama, K., Kamiya, A., Ito, T. and Shinozaki, K. (2004) A collection of 11,800 single-copy *Ds* transposon insertion lines in *Arabidopsis*. *Plant J.* **37**, 897–905.

Lurin, C., Andres, C., Aubourg, S. *et al.* (2004) Genome-wide analysis of *Arabidopsis* pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell*, **16**, 2089–2103.

Lutz, K.A. and Maliga, P. (2001) Lack of conservation of editing sites in mRNAs that encode subunits of the NAD(P)H dehydrogenase complex in plastids and mitochondria of *Arabidopsis thaliana*. *Curr. Genet.* **40**, 214–219.

Munekage, Y., Hashimoto, M., Miyake, C., Tomizawa, K., Endo, T., Tasaka, M. and Shikanai, T. (2004) Cyclic electron flow around photosystem I is essential for photosynthesis. *Nature*, **429**, 579–582.

Murcha, M.W., Elhafez, D., Lister, R., Tonti-Filippini, J., Baumgartner, M., Philippar, K., Carrie, C., Mokranjac, D., Soll, J. and Whelan, J. (2007) Characterization of the preprotein and amino acid transporter gene family in *Arabidopsis*. *Plant Physiol.* **143**, 199–212.

Nakamura, T. and Sugita, M. (2008) A conserved DYW domain of the pentatricopeptide repeat protein possesses a novel endoribonuclease activity. *FEBS Lett.* **582**, 4163–4168.

O'Toole, N., Hattori, M., Andres, C., Iida, K., Lurin, C., Schmitz-Linneweber, C., Sugita, M. and Small, I. (2008) On the expansion of the pentatricopeptide repeat gene family in plants. *Mol. Biol. Evol.* **25**, 1120–1128.

Okuda, K., Nakamura, T., Sugita, M., Shimizu, T. and Shikanai, T. (2006) A pentatricopeptide repeat protein is a site recognition factor in chloroplast RNA editing. *J. Biol. Chem.* **281**, 37661–37667.

Okuda, K., Myouga, F., Motohashi, R., Shinozaki, K. and Shikanai, T. (2007) Conserved domain structure of pentatricopeptide repeat proteins involved in chloroplast RNA editing. *Proc. Natl Acad. Sci. USA*, **104**, 8178–8183.

Okuda, K., Habata, Y., Kobayashi, Y. and Shikanai, T. (2008) Amino acid sequence variations in *Nicotiana* CRR4 orthologs determine the species-specific efficiency of RNA editing in plastids. *Nucleic Acids Res.* **36**, 6155–6164.

Okuda, K., Chateigner-Boutin, A., Nakamura, T., Delannoy, E., Sugita, M., Myouga, F., Motohashi, R., Shinozaki, K., Small, I. and Shikanai, T. (2009) Pentatricopeptide repeat proteins with the DYW motif have distinct molecular functions in RNA editing and RNA cleavage in *Arabidopsis* chloroplasts. *Plant Cell*, **21**, 146–156.

Peng, L., Ma, J., Chl, W., Guo, J., Zhu, S., Lu, Q., Lu, C. and Zhang, L. (2006) Low PSII accumulation1 is involved in efficient assembly of photosystem II in *Arabidopsis thaliana*. *Plant Cell*, **18**, 955–969.

Peng, L., Shimizu, H. and Shikanai, T. (2008) The chloroplast NAD(P)H dehydrogenase complex interacts with photosystem I in *Arabidopsis*. *J. Biol. Chem.* **83**, 34873–34879.

Peng, L., Fukao, Y., Fujiwara, M., Takami, T. and Shikanai, T. (2009) Efficient operation of NAD(P)H dehydrogenase requires the supercomplex formation with photosystem I via minor LHCI in *Arabidopsis*. *Plant Cell*, doi: 10.1105/tpc.109.068791.

Robbins, J.C., Heller, W.P. and Hanson, M.R. (2009) A comparative genomics approach identifies a PPR-DYW protein that is essential for C-to-U editing of *Arabidopsis* chloroplast *accD* transcript. *RNA*, **15**, 1142–1153.

Rüdinger, M., Polsakiewicz, M. and Knoop, V. (2008) Organellar RNA editing and plant-specific extensions of pentatricopeptide repeat proteins in *jungermanniid* but not in *marchantiid* liverworts. *Mol. Biol. Evol.* **25**, 1405–1414.

Salone, V., Rüdinger, M., Polsakiewicz, M., Hoffmann, B., Groth-Malonek, M., Szurek, B., Small, I., Knoop, V. and Lurin, C. (2007) A hypothesis on the identification of the editing enzyme in plant organelles. *FEBS Lett.* **581**, 4132–4138.

Sasaki, T., Yukawa, Y., Miyamoto, T., Obokata, J. and Sugiura, M. (2003) Identification of RNA editing siets in chloroplast transcripts from the maternal and paternal progenitors of tobacco (*Nicotiana tabacum*): comparative analysis shows the involvement of distinct *trans*-factors for *ndhB* editing. *Mol. Biol. Evol.* **20**, 1028–1035.

Schmitz-Linneweber, C. and Small, I. (2008) Pentatricopeptide repeat proteins: a socket set for organelle gene expression. *Trends Plant Sci.* **13**, 663–670.

Shikanai, T. (2006) RNA editing in plant organelles: machinery, physiological function and evolution. *Cell. Mol. Life Sci.* **63**, 698–708.

Shikanai, T., Endo, T., Hashimoto, T., Yamada, Y., Asada, K. and Yokota, A. (1998) Directed disruption of the tobacco *ndhB* gene impairs cyclic electron flow around photosystem I. *Proc. Natl Acad. Sci. USA*, **95**, 9705–9709.

Small, I.D. and Peeters, N. (2000) The PPR motif – a TPR-related motif prevalent in plant organellar proteins. *Trends Biochem. Sci.* **25**, 46–47.

Small, I., Peeters, N., Legeai, F. and Lurin, C. (2004) Predotar: a tool for rapidly screening proteomes for *N*-terminal targeting sequences. *Proteomics*, **4**, 1581–1590.

Takenaka, M., Neuwirt, J. and Brennicke, A. (2004) Complex *cis*-elements determine an RNA editing site in pea mitchoncdria. *Nucleic Acids Res.* **32**, 4137–4144.

Timmis, J.N., Ayliffe, M.A., Huang, C.Y. and Martin, W. (2004) Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nat. Rev. Genet.* **5**, 123–135.

Tsudzuki, T., Wakasugi, T. and Sugiura, M. (2001) Comparative analysis of RNA editing sites in higher plant chloroplasts. *J. Mol. Evol.* **53**, 327–332.

Yu, Q.B., Jiang, Y., Chong, K. and Yang, Z.N. (2009) AtECB2, a pentatricopeptide repeat protein, is required for chloroplast transcript *accD* RNA editing and early chloroplast biogenesis in *Arabidopsis thaliana*. *Plant J.* **59**, 1011–1023.

Zehrmann, A., van der Merwe, J.A., Verbitskiy, D., Brennicke, A. and Takenaka, M. (2008) Seven large variations in the extent of RNA editing in plant mitochondria between three ecotypes of *Arabidopsis thaliana*. *Mitochondrion*, **8**, 319–327.

Zehrmann, A., Verbitsky, D., van der Merwe, J.A., Brennicke, A. and Takenaka, M. (2009) A DYW domain-containing pentatricopeptide repeat proteins is required for RNA editing at multiple sites in mitochondria of *Arabidopsis thaliana*. *Plant Cell*, **21**, 558–567.

Zhou, W., Cheng, Y., Yap, A., Chateigner-Boutin, A.L., Delannoy, E., Hammani, K., Small, I. and Huang, J. (2008) The *Arabidopsis* gene *YS1* encoding a DYW protein is required for editing of *rpoB* transcripts and the rapid development of chloroplasts during growth. *Plant J.* **58**, 82–96.

Table S1. A table summarizing the PPR proteins required for RNA editing in *Arabidopsis* plastids.

Site-specificity factors	Subclass	Target editing sites	Physiological process	References
CRR4	E	ndhD-1	Photosynthesis	Kotera et al., 2005
CRR21	E	ndhD-2	Photosynthesis	Okuda et al., 2007
CLB19	E	clpP and rpoA	Plastid biogenesis and transcription	Chateigner-Boutin et al., 2008
YS1	DYW	rpoB-1	Plastid transcription	Zhou et al., 2008
CRR22	DYW	ndhB-7, ndhD-5, and rpoB-3	Photosynthesis and plastid transcription	Okuda et al., 2009
CRR28	DYW	ndhB-2 and ndhD-3	Photosynthesis	Okuda et al., 2009
RARE1	DYW	accD	Plastid biogenesis	Robins et al., 2009
LPA66	DYW	psbF	Photosynthesis	Cai et al., 2009
ECB2	DYW	accD	Plastid biogenesis	Yu et al., 2009
OTP82	DYW	ndhB-9 and ndhG-1	Photosynthesis	This study

Table S2. Oligonucleotide primers used.

name	sequence
For genotyping	
LB SALK	GCGTGGACCGCTTGCTGCAAC
LB SAIL	AGCATCTGAATTTCATAACCAATCTCGATACAC
SALK_027812RP	TACGGAGAGAAGAAGCATTCG
SALK_027812LP	TTACCCATTCCATTTCCTTCC
SAIL_851_G04RP	TACCGGTTTCAGCATATCCTG
SAIL_851_G04LP	TTTATCGTCCACGAAAATAATGG
For GFP construct cloning	
z0797	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCAT GATGCTCTCGTGTTCTCC
z0798	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTA CCAGTAGTCATTGCAGG
For generating PCR templates for PPE assay	
ndhG.AT.for	AGAATTATTGAAATGAGTTC
ndhG.AT.rev	TATAATGGATTTGCCTGGAC
ndhB.AT.rev2	CTAAAAAAGGGTATCCTGAGC
AndhB	GTCGTTGCTTTTCTTTCTG
For PPe analysis	
ndhB(95644)_PPe_G	AAAGGAATATCGAAAATTCGAGTGG
ndhG_PPe_G	CACTCCCAGACCCCCTACTAGAAGACCG
For RNA gel blot analysis	
ndhB.AT.rev2	CTAAAAAAGGGTATCCTGAGC
AndhB	GTCGTTGCTTTTCTTTCTG
ndhG.AT.for	AGAATTATTGAAATGAGTTC
ndhG.AT.rev	TATAATGGATTTGCCTGGAC
M13/pUC reverse	GAGCGGATAACAATTTCACACAGG
For HRM analysis	
matK_HRM_for:	TGAATCCAAGATTTTTCTTGTTCTT

matK_HRM_rev:	AAAAAGATGGATTCGTATTCACA
atpF_HRM_for:	GCCGGGAGTTTCGGATTTA
atpF_HRM_rev:	GATCAATACACCGAAAACTACACTT
rpoC1_HRM_F:	TTCTTTTGCTAGGCCCATAA
rpoC1_HRM_R:	TGCTGTATTTCCAGGATTGAA
rpoB23898_HRM_for:	TGGTTCAAGTTATAACCCAGAAAT
rpoB23898_HRM_rev:	CCATGTCTTCCAGCTACTTTATCA
rpoB25779_HRM_for:	GCTCGGGTGAGTAGGAAACA
rpoB25779_HRM_rev:	TCTAGAATTTCTCGTAGATTCAAACC
rpoB25992_HRM_for:	TCCTTTAATGAATTCCCTTGGA
rpoB25992_HRM_rev:	CGGGACTTTGCAATATTTGAT
psbZ_HRM_for:	TGCTTTCCAATTGGCAGTTT
psbZ_HRM_rev:	GAGACGCAAATACAACGGGTA
RPS14(37092)HRM_for:	TCGCTAAGTGAGAAATGGAAAA
RPS14(37092)HRM_rev:	CGTCGATGAAGACGTGTAGG
RPS14(37161)HRM_for:	GAAGAAGAGGCAAAAATTGGAA
RPS14(37161)HRM_rev:	TTCTCACTTAGCGACGGAATC
accD_HRMfor:	TGAATGTTTGTGAACAATGTGG
accD_HRMrev:	GATTCCAAGTACCCGGATCA
accD(58642)HRM_rev:	ACTTTTTAGCTTGTTGATAGAGGTTT
accD(58642)HRM_for:	GCATTTTTCTTTCAAATCATTTTT
psbF_HRM_rev:	TCGTTGGATGAACTGCATTG
psbF_HRM_for:	GGCTGTTCATGGACTAGCTGTA
psbE_HRM_rev:	TGGGTCCTCCTAAAAAGATCTAC
psbE_HRM_for:	AAGGCATTCCATTAATAACAGG
petL_HRM_for:	ATTTTATTGAGTCCCTTCATGC
petL_HRM_rev:	AAAGCTGCTAGTAGAAAACCGAAA
Rps12_HRM_rev:	TCGAGTTCGTCCATTCTATTTT
Rps12_HRM_for:	AGACTGAGACATAAAAAGGAAATTCT
clpPAt2rev:	TGAACCGCTACAAGATCAAC
clpP_HRM70for:	GAAGACATGGAACGGGATGT
rpoA_HRMfor:	TTGGCGAAATAGAAGGAACA
rpoA_HRMrev:	CCTGCTATGTTAGAATAGTCATGTGG
Rpl23_HRM_r:	GAAGAGTTCGACCCAATGCT
Rpl23_HRM_f:	TCAATATACTTTTAATGTCGAATCAGG
ndhB(94999)HRM_for:	TTCCATCGAATTGAGTATGATTG

ndhB(94999)HRM_rev:	AATCGCAATAATCGGGTTCA
ndhB95225_HRM_for:	GGTCTTCCTCCACTAGCAGGT
ndhB95225_HRM_rev:	CTGCCATCCACACCAGAATA
95640R:	CACTCGAATTTTCGATATTCCTTT
ndhB(95608)HRM_rev:	TCCAGAAGAAGATGCCATTCA
95618F:	AAGGAATATCGAAAATTCGAGTGG
95693R:	CCAGTCGTTGCTTTTCTTTCTG
95627F:	CGAAAATTCGAGTGGCTGA
95698R:	CCACTCCAGTCGTTGCTTTT
96461R:	TCATCACTGTAGGAATTGGGTTC
ndhB(96419)HRM_rev:	CTTCGTATACGTCAGGAGTCCA
96606R:	GGTGGGGCAAGCTCTTCTAT
96537F:	CAATCTCTCCCCCGGATG
ndhB(96698)HRM_for:	TTTTTATGTGGTGCTAACGATTT
ndhB(96698)HRM_rev:	CCAGATAATAGGTAGGAGCATAAACTG
ndhB(97016)HRM_for:	TTGGCCTAATTCTTCTTCTGATG
ndhB(97016)HRM_rev:	TTGAAGAGATGAAATATAACCAAGG
ndhF(112349)HRM_for:	TTCGGTTACTTTATCGATCCAC
ndhF(112349)HRM_rev:	TCAGAACCAAAATCCCAACAG
116318R:	GGTACAATCCAAATAATTTATGCAG
116228F:	CCATATGAGATACAGAAGAATAGGC
ndhD(116494)HRM_for:	TTGGATTTCTTATTGCTTTTGC
ndhD(116494)HRM_rev:	CACCGTGGGTGTCAGGTAAC
ndhD(116785)HRM_for:	AGCAATGTACAGCGGTCAAA
ndhD(116785)HRM_rev:	TTCTAATTCCCACATGATGAAAAA
ndhD(117166)HRM_for:	TTGGACCTGGTGTATCTTGTC
ndhD(117166)HRM_rev:	CCAGCAGATATTGGAAAAACAA
ndhG_HRMfor:	TTGCCTGGACCAATACATGA
ndhG_HRMrev:	CACTCCCAGACCCCCTACTA

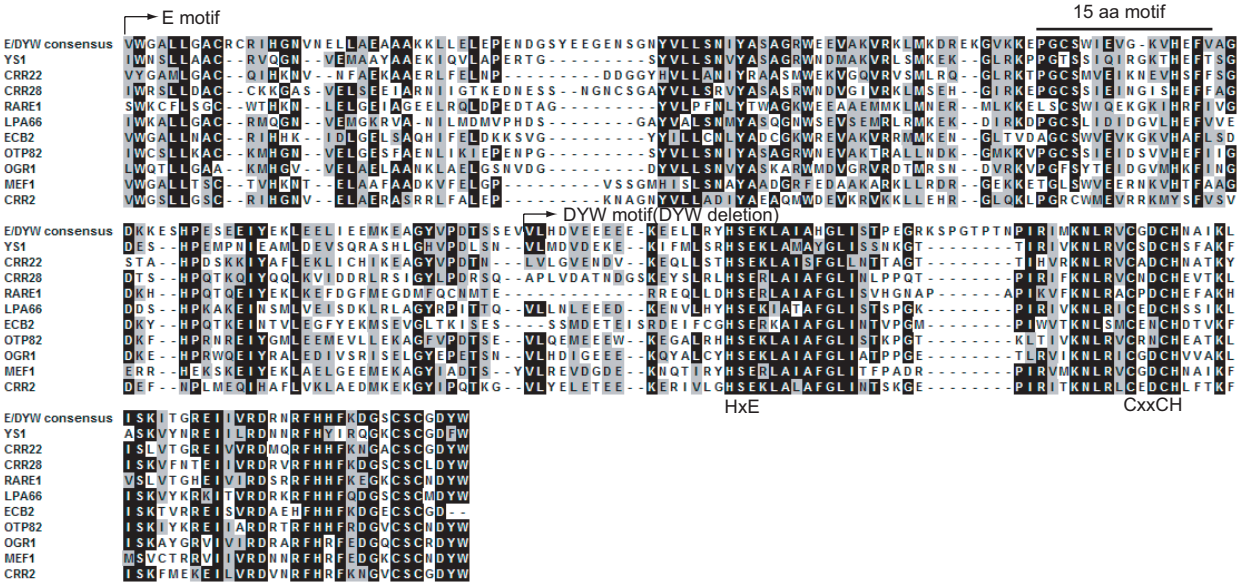
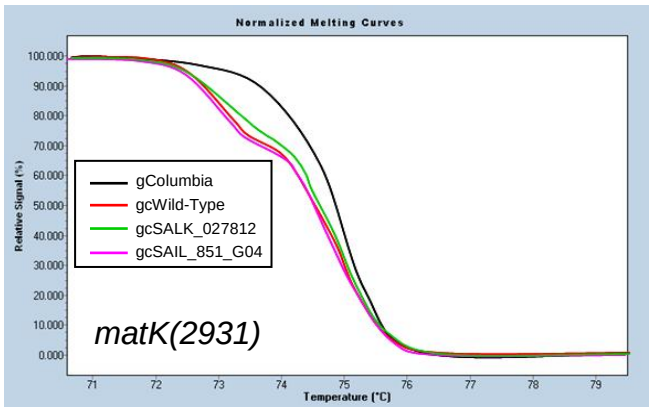
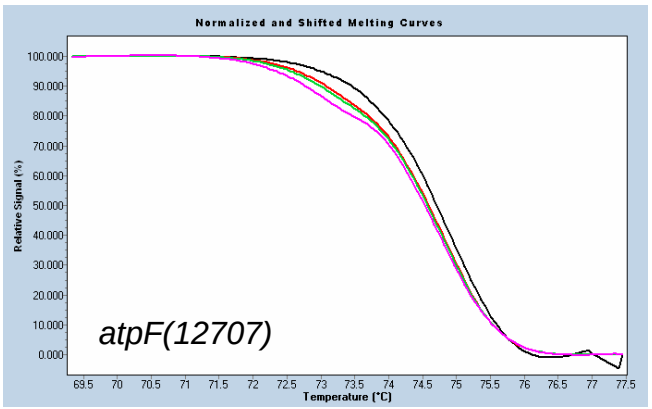
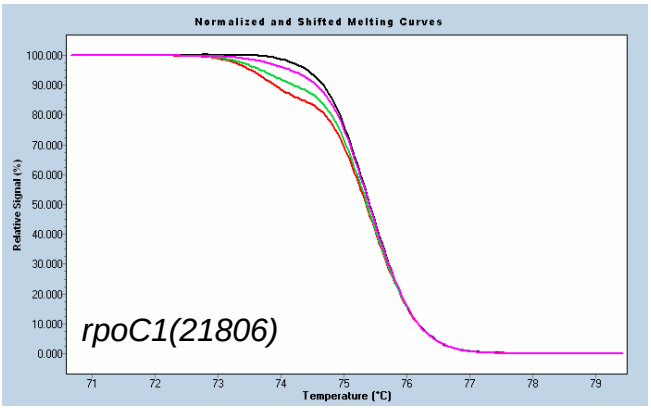


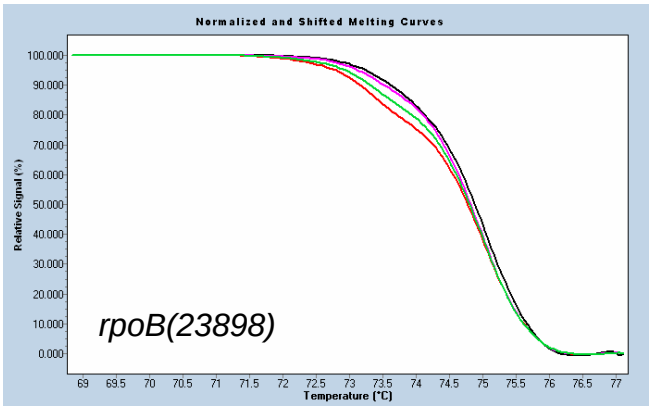
Figure S1. Comparison of the E and DYW motifs among YS1, CRR22, CRR28, RARE1, LPA66, ECB2, OTP82, OGR1, MEF1, and CRR2.

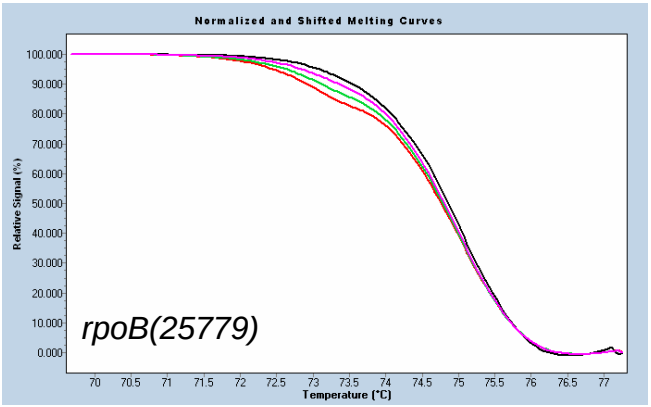
Alignment was performed using ClustalW program. The consensus sequence of the E and DYW motifs according to Lurin *et al*., (2004) is shown at the top. Amino acids that are fully or semiconserved are shaded black and gray, respectively. The invariant Cys and His residues in the DYW motif (Salone *et al*., 2007) are indicated above the sequences. The 15 amino acids motif (Kotera *et al*., 2005) is indicated above the sequences. The point at which the sequence was truncated in the *OTP82ΔDYW* construct is specified.

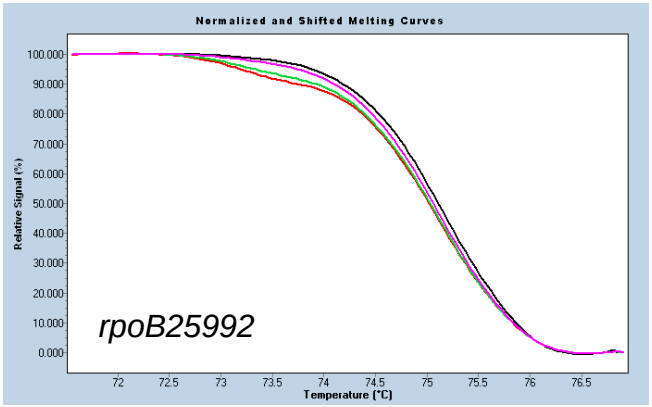

Normalized Melting Curves
gColumbia
gcWild-Type
gcSALK_027812
gcSAIL_851_G04
matK(2931)
Relative Signal (%)
Temperature (°C)

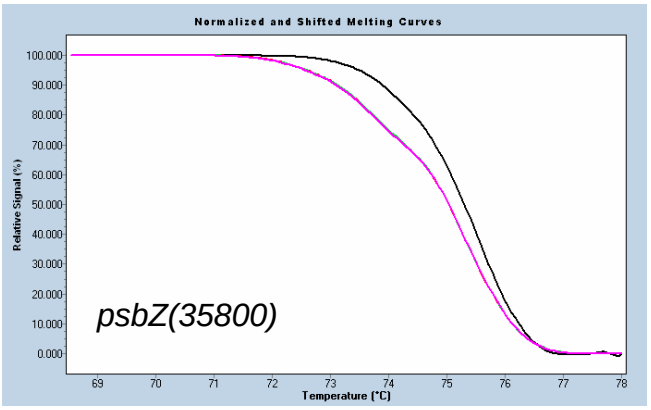

Normalized and Shifted Melting Curves
atpF(12707)
Relative Signal (%)
Temperature (°C)

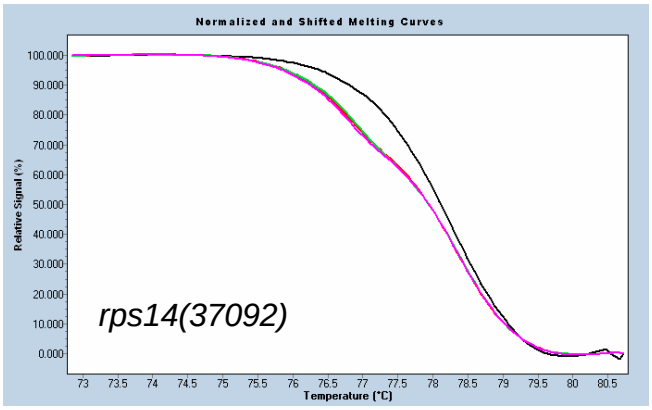

Normalized and Shifted Melting Curves
rpoC1(21806)
Relative Signal (%)
Temperature (°C)

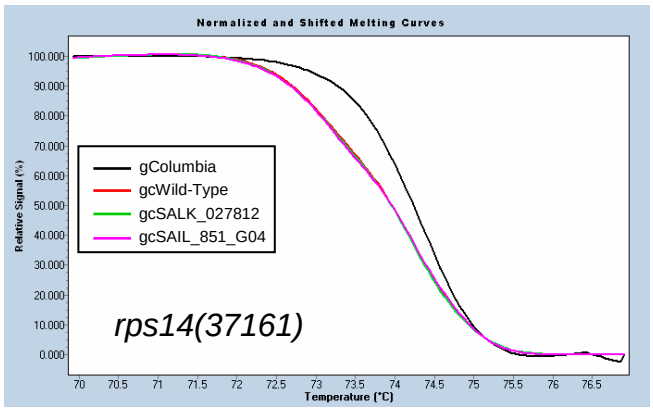

Normalized and Shifted Melting Curves
rpoB(23898)
Relative Signal (%)
Temperature (°C)

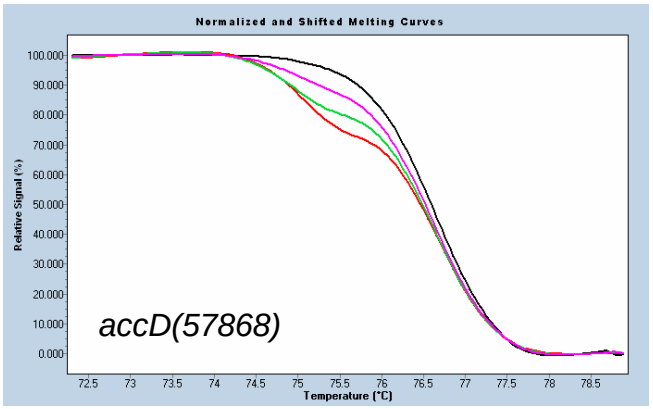

Normalized and Shifted Melting Curves
rpoB(25779)
Relative Signal (%)
Temperature (°C)

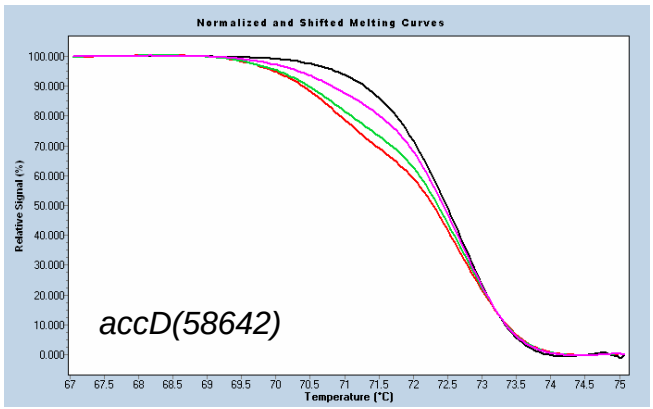

Normalized and Shifted Melting Curves
rpoB25992
Relative Signal (%)
Temperature (°C)

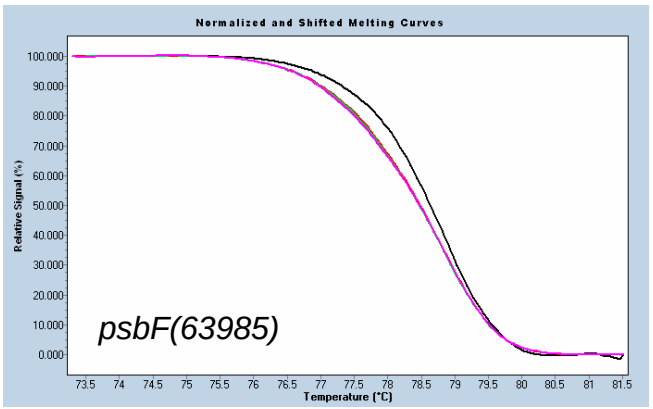

Normalized and Shifted Melting Curves
psbZ(35800)
Relative Signal (%)
Temperature (°C)

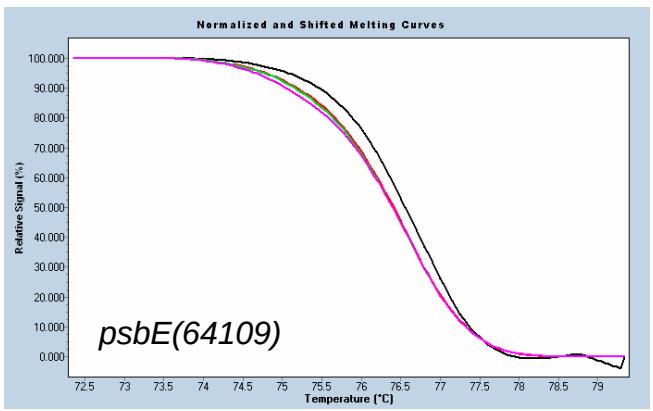

Normalized and Shifted Melting Curves
rps14(37092)
Relative Signal (%)
Temperature (°C)

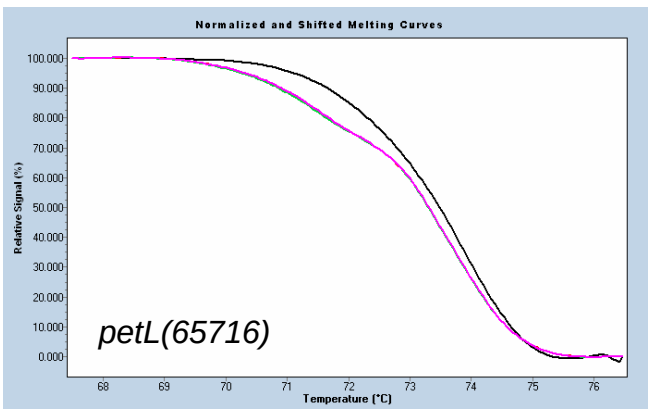

Normalized and Shifted Melting Curves
gColumbia
gcWild-Type
gcSALK_027812
gcSAIL_851_G04
Relative Signal (%)
Temperature (°C)
rps14(37161)

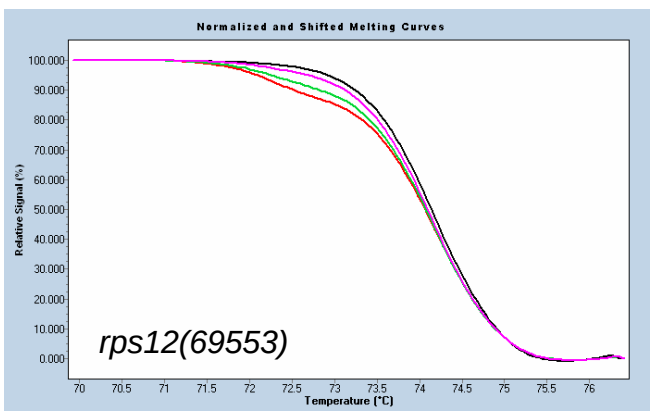

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
accD(57868)

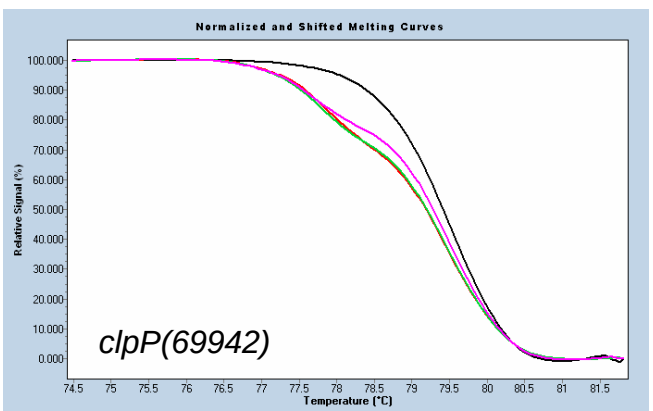

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
accD(58642)

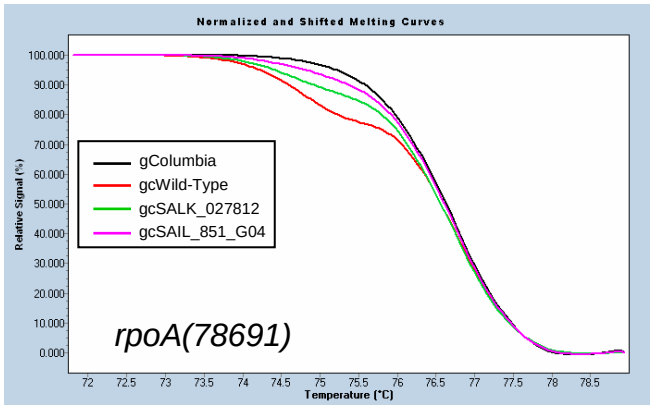

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
psbF(63985)

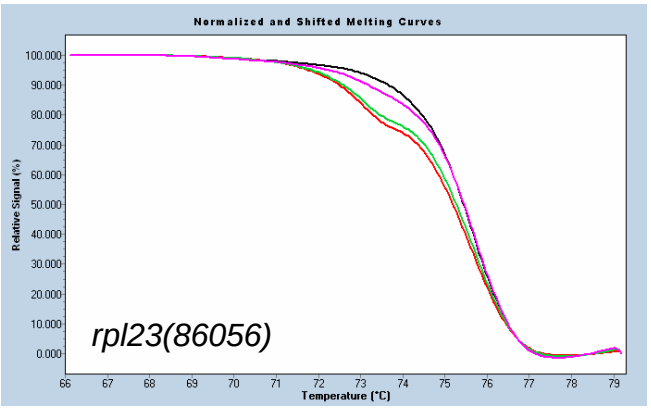

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
psbE(64109)

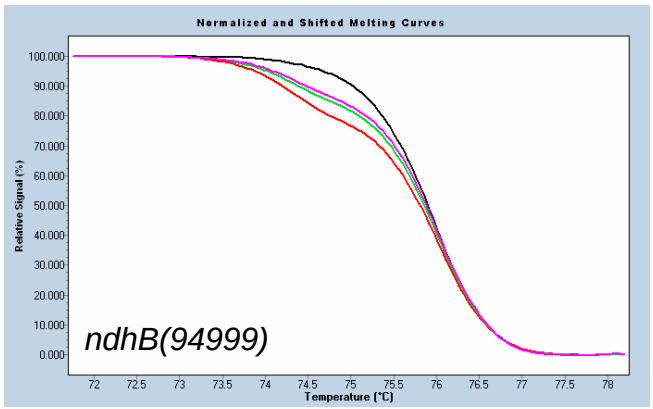

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
petL(65716)

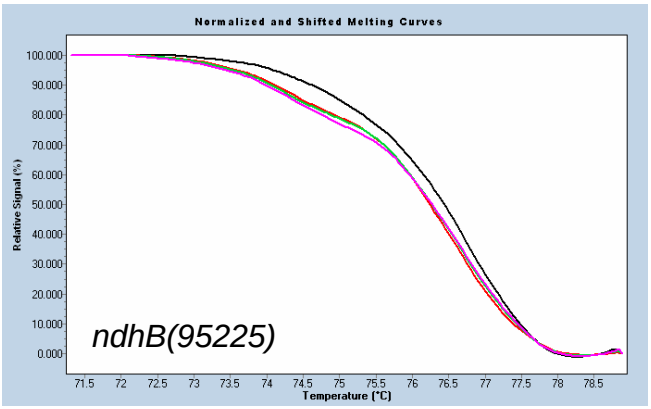

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
rps12(69553)

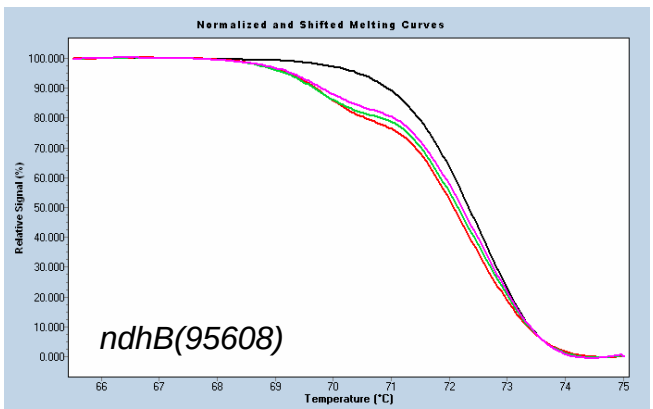

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
clpP(69942)

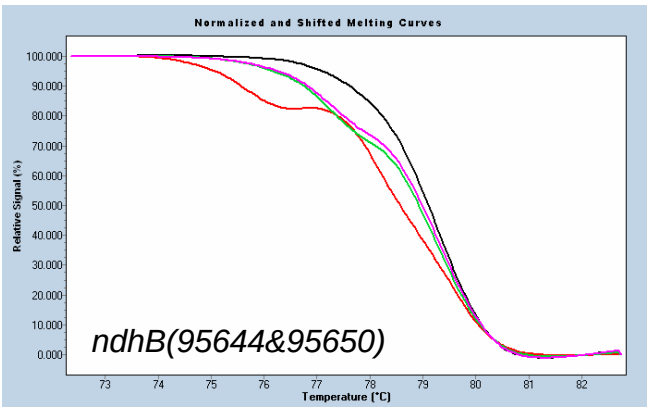

Normalized and Shifted Melting Curves
gColumbia
gcWild-Type
gcSALK_027812
gcSAIL_851_G04
rpoA(78691)
Relative Signal (%)
Temperature (°C)

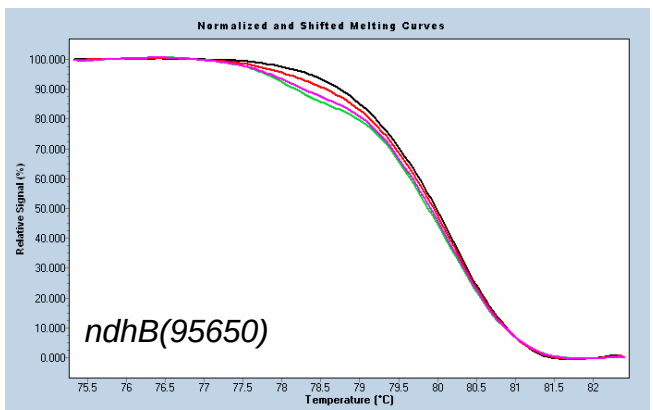

Normalized and Shifted Melting Curves
rpl23(86056)
Relative Signal (%)
Temperature (°C)

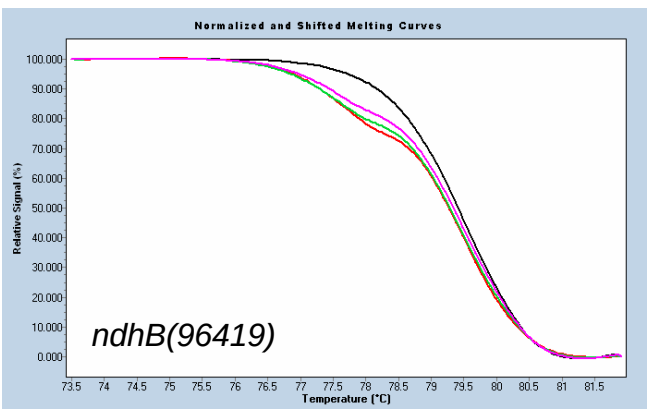

Normalized and Shifted Melting Curves
ndhB(94999)
Relative Signal (%)
Temperature (°C)

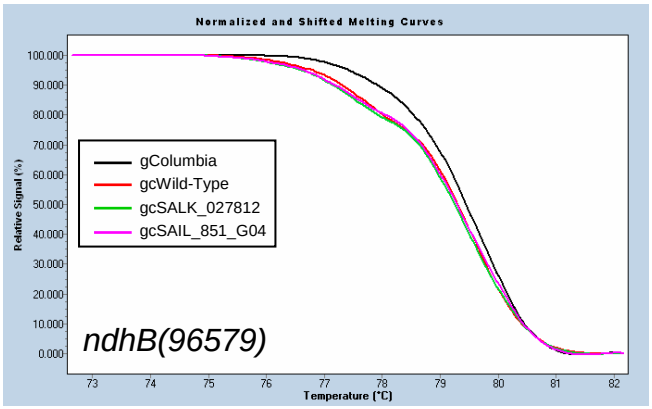

Normalized and Shifted Melting Curves
ndhB(95225)
Relative Signal (%)
Temperature (°C)

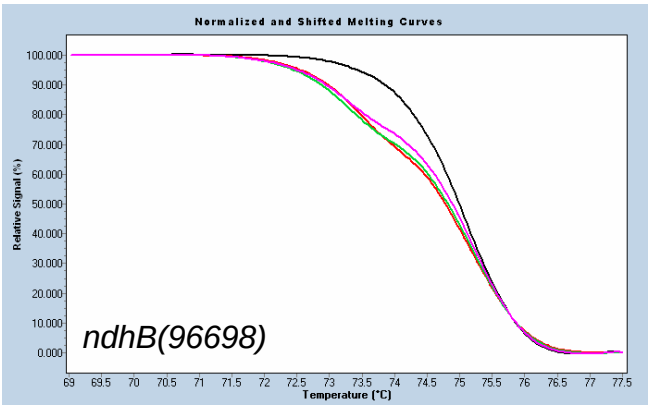

Normalized and Shifted Melting Curves
ndhB(95608)
Relative Signal (%)
Temperature (°C)

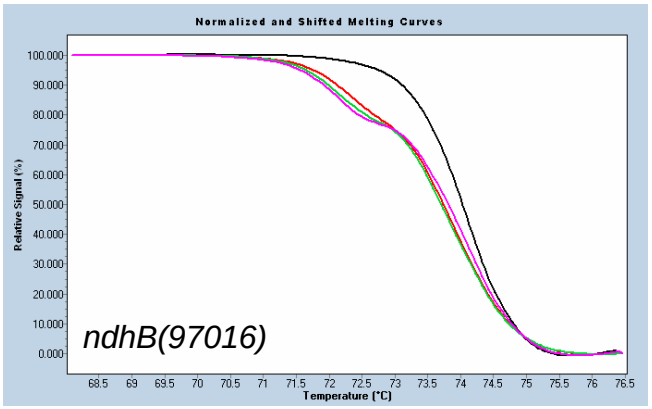

Normalized and Shifted Melting Curves
ndhB(95644&95650)
Relative Signal (%)
Temperature (°C)

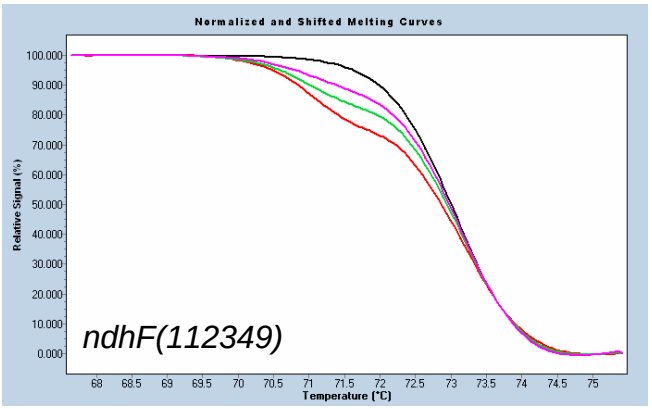

Normalized and Shifted Melting Curves
ndhB(95650)
Relative Signal (%)
Temperature (°C)

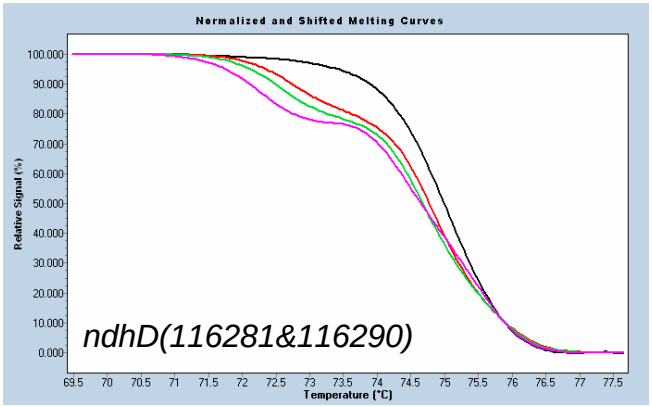

Normalized and Shifted Melting Curves
ndhB(96419)
Relative Signal (%)
Temperature (°C)

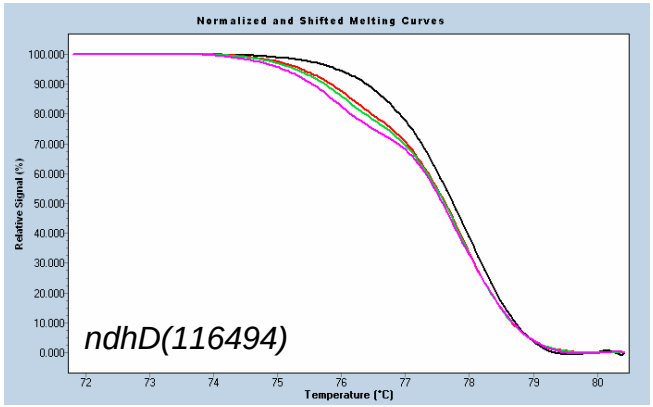

Normalized and Shifted Melting Curves
gColumbia
gcWild-Type
gcSALK_027812
gcSAIL_851_G04
Relative Signal (%)
Temperature (°C)
ndhB(96579)

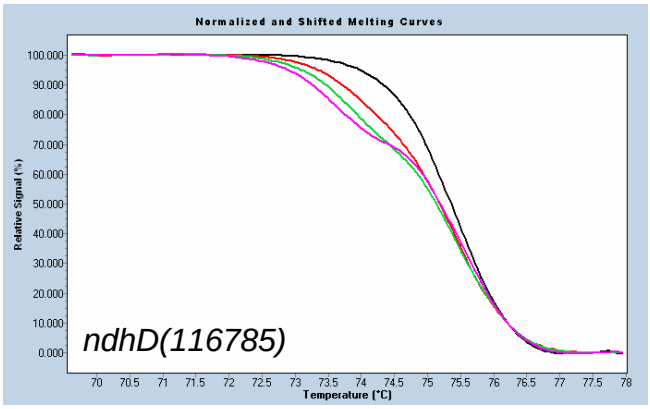

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhB(96698)

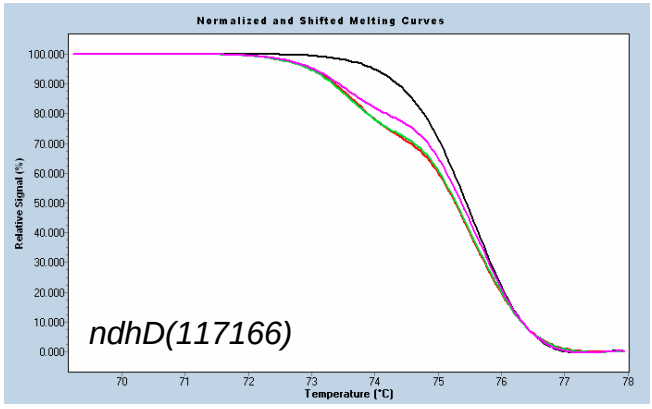

Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhB(97016)


Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhF(112349)


Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhD(116281&116290)


Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhD(116494)


Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhD(116785)


Normalized and Shifted Melting Curves
Relative Signal (%)
Temperature (°C)
ndhD(117166)

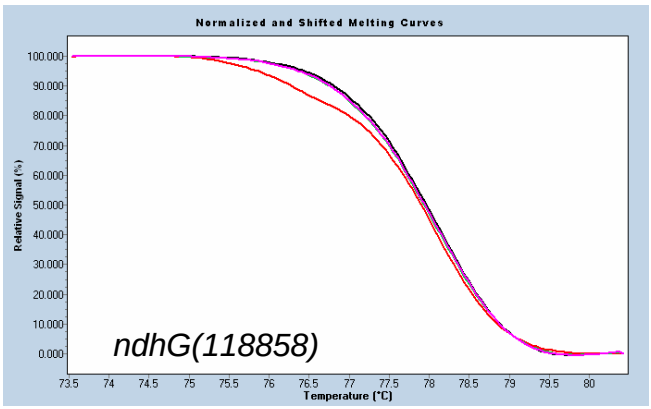


Figure S2. High resolution melting screen of *otp82* mutants.
Real-time PCR was done using primers surrounding the 34 known editing sites of *Arabidopsis* in the presence of a fluorescent double-strand-specific dye. At the end of the amplification, amplicons were denatured, renatured and then melted using precise incremental increases in temperature. Melting of DNA duplexes releases the dye, causing a decrease in fluorescence. The presence of less thermostable heteroduplexes in a sample alters the shape of the melting curves. For each editing site, the melting curve of a control consisting of genomic DNA (g) from the wild type sample was compared to the melting curves of a mix of genomic DNA and cDNA (gc) from the mutant lines. Editing produces cDNA with a nucleotide change and thus the mix of genomic DNA and cDNA exhibits a different melting curve from that of genomic DNA alone. If an editing site is unedited in the mutant, then the melting curves from the genomic DNA sample and the mixed genomic/cDNA sample are identical.

Figure S3. Sequence data for mass spectrometric analysis

>NdhB (all sites are edited)
MIWHVQNENFILDSTRIFMKAFHLLLFDGSFIFPECILIFGLILLLMID**L**TSDQKDIPW
LYFISSTSFVMSITALLFRWREEPMISFSGNFQTNNFNEIFQFLILLCSTLCIPLSVEY
IECTEMAITEFLLFILTATLGGMFLCGANDLITIFVA**L**ECFSLCSYLLSGYTKKDIRSN
EATMKYLLMGGASSSILV**Y**GFSWLYG**L**SGGEIELQEIVNGLINTQMYNSPGISIALIFI
TVGIGFKLSLAP**F**HQWTPDVYEGSPTPVVAFLSVTSKVAA**L**A**L**ATRIFDIPFYF**L**SNEW
HLLLEILAILSMIFGNLIAITQTSMKRMLAYSSIGQIGYVIIGIIVGDSNGGYASMITY
MLFYIAMNLGTFACIILFGLRTGTDNIRDYAGLYTKDPFLALSLALCLLSLGGLPPLAG
FFGKL**Y**LFWCGWQAGLYFLVSIGLLTSVLSIYYYLKIIKLLMTGRNQEITPHMRNYRIS
PLRSNNSIELSMIVCVIASTI**L**GISMNPIIAIAQDTLFSF
>NdhB (ndhB-9 site is unedited)
MIWHVQNENFILDSTRIFMKAFHLLLFDGSFIFPECILIFGLILLLMID**L**TSDQKDIPW
LYFISSTSFVMSITALLFRWREEPMISFSGNFQTNNFNEIFQFLILLCSTLCIPLSVEY
IECTEMAITEFLLFILTATLGGMFLCGANDLITIFVA**L**ECFSLCSYLLSGYTKKDIRSN
EATMKYLLMGGASSSILV**Y**GFSWLYG**L**SGGEIELQEIVNGLINTQMYNSPGISIALIFI
TVGIGFKLSLAP**F**HQWTPDVYEGSPTPVVAFLSVTSKVAA**L**A**S**ATRIFDIPFYF**L**SNEW
HLLLEILAILSMIFGNLIAITQTSMKRMLAYSSIGQIGYVIIGIIVGDSNGGYASMITY
MLFYIAMNLGTFACIILFGLRTGTDNIRDYAGLYTKDPFLALSLALCLLSLGGLPPLAG
FFGKL**Y**LFWCGWQAGLYFLVSIGLLTSVLSIYYYLKIIKLLMTGRNQEITPHMRNYRIS
PLRSNNSIELSMIVCVIASTI**L**GISMNPIIAIAQDTLFSF
>NdhB (ndhB-8 site is unedited)
MIWHVQNENFILDSTRIFMKAFHLLLFDGSFIFPECILIFGLILLLMID**L**TSDQKDIPW
LYFISSTSFVMSITALLFRWREEPMISFSGNFQTNNFNEIFQFLILLCSTLCIPLSVEY
IECTEMAITEFLLFILTATLGGMFLCGANDLITIFVA**L**ECFSLCSYLLSGYTKKDIRSN
EATMKYLLMGGASSSILV**Y**GFSWLYG**L**SGGEIELQEIVNGLINTQMYNSPGISIALIFI
TVGIGFKLSLAP**F**HQWTPDVYEGSPTPVVAFLSVTSKVAA**S**A**L**ATRIFDIPFYF**L**SNEW
HLLLEILAILSMIFGNLIAITQTSMKRMLAYSSIGQIGYVIIGIIVGDSNGGYASMITY
MLFYIAMNLGTFACIILFGLRTGTDNIRDYAGLYTKDPFLALSLALCLLSLGGLPPLAG
FFGKL**Y**LFWCGWQAGLYFLVSIGLLTSVLSIYYYLKIIKLLMTGRNQEITPHMRNYRIS
PLRSNNSIELSMIVCVIASTI**L**GISMNPIIAIAQDTLFSF
>NdhB (ndhB-8 and -9 sites are unedited)
MIWHVQNENFILDSTRIFMKAFHLLLFDGSFIFPECILIFGLILLLMID**L**TSDQKDIPW
LYFISSTSFVMSITALLFRWREEPMISFSGNFQTNNFNEIFQFLILLCSTLCIPLSVEY
IECTEMAITEFLLFILTATLGGMFLCGANDLITIFVA**L**ECFSLCSYLLSGYTKKDIRSN

EATMKYLLMGGASSSILV**Y**GFSWLYG**L**SGGEIELQEIVNGLINTQMYNSPGISIALIFI
TVGIGFKLSLAP**F**HQWTPDVYEGSPTPVVAFLSVTSKVAA**S**A**S**ATRIFDIPFYF**L**SNEW
HLLLEILAILSMIFGNLIAITQTSMKRMLAYSSIGQIGYVIIGIIVGDSNGGYASMITY
MLFYIAMNLGTFACIILFGLRTGTDNIRDYAGLYTKDPFLALSLALCLLSLGGLPPLAG
FFGKL**Y**LFWCGWQAGLYFLVSIGLLTSVLSIYYYLKIIKLLMTGRNQEITPHMRNYRIS
PLRSNNSIELSMIVCVIASTI**L**GISMNPIIAIAQDTLFSF

L'investigation plus avancée du mutant *otp82* a montré que l'édition des sites *ndhG-11* et *ndhB-9* n'est pas essentielle pour la fonction biologique des protéines NdhG et NdhB dans les conditions de culture du laboratoire. Ces résultats démontrent pour la première fois qu'une protéine traduite d'un transcrit non éditée peut être fonctionnellement équivalente à la protéine résultante de la traduction du transcrit édité. Malheureusement, l'incorporation du peptide de forme "non édité" au sein du complexe NAD(P)H déshydrogénase des thylakoïdes du mutant *otp82* n'a pu être prouvé par spectrométrie de masse.

4.2. IMPLICATION D'UNE PROTEINE PPR DANS LA REGULATION DE L'EXPRESSION DES GENOMES DU NOYAU ET DES MITOCHONDRIES

L'analyse des séquences protéiques des protéines PPR a révélé que trois quarts des protéines PPR présentent des séquences putatives d'adressage aux mitochondries ou chloroplastes (Lurin et *al*., 2004). Ceci laisse la possibilité pour certaines protéines PPR de fonctionner en dehors des organites probablement dans d'autres compartiments possédant un génome tel que le noyau. Il a été démontré pour la première fois chez *Arabidopsis* qu'une protéine PPR, GRP23 (pour Glutamine Rich Protein23) était adressée au noyau des cellules où elle interagirait avec la sous-unité III de l'ARN polymérase nucléaire (Ding et *al*., 2006). GRP23 présente également un domaine basique de liaison à l'ADN caractéristique des facteurs de transcription. La protéine GRP23 est donc proposée comme régulateur de l'expression du génome nucléaire. Il est probable que d'autres protéines PPR puissent avoir le même type de fonction dans le noyau.

Chez l'humain, de manière surprenante, l'un des 6 gènes *PPR* code pour une protéine LRP130 (ou LRPPRC) doublement adressée au noyau et aux mitochondries des cellules et pourrait jouer un rôle dans la communication entre les deux organites en influençant l'expression de leur génome (Gohil et *al.*, 2010 ; Sasarman et *al.*, 2010 ; Xu et *al.*, 2004 ; Mili et Piñol-Roma, 2003).

L'identification de ces deux protéines laisse la porte ouverte à la découverte de nouvelles protéines PPR localisés en dehors des organites et aux fonctions nouvelles.

Lors d'un criblage de mutants d'insertion de gènes *PPR* chez *Arabidopsis*, je me suis intéressé au gène particulier *PPR9*, dont l'extinction chez *Arabidopsis* entraine la létalité de l'embryon. De manière très intéressante, la séquence protéique PPR9 présente deux signaux de localisation cellulaire: une séquence d'adressage aux mitochondries et un signal de localisation nucléaire. Nous avons donc émis l'hypothèse que PPR9 serait une nouvelle protéine PPR doublement adressée au noyau et aux mitochondries des cellules d'*Arabidopsis* qui pourrait avoir de nouvelles fonctions non assignées précédemment. Nous avons donc conduit une analyse fonctionnelle génétique et biochimique du gène *PPR9* et de la protéine PPR9.

4.2.1. PUBLICATION 3: ARABIDOPSIS PPR9 IS DUAL LOCALIZED IN MITOCHONDRIA AND THE NUCLEUS WHERE IT INTERACTS WITH TRANSCRIPTION FACTORS. (EN PREPARATION)

Arabidopsis PPR9 is dual localized in mitochondria and the nucleus where it interacts with transcription factors.

Kamel Hammani[a,b], Anthony Gobert[a] Kamal Hleibieh[a], Laurence Choulier[c], Ian Small[b], Philippe Giegé[a,1]

[a] Institut de Biologie Moléculaire des Plantes du CNRS, Université de Strasbourg, 12 rue du Général Zimmer, 67084 Strasbourg Cedex, France

[b] Australian Research Council Centre of Excellence in Plant Energy Biology, University of Western Australia, Crawley 6009 WA, Australia

[c]UMR7100-CNRS, ESBS, Boulevard Sébastien Brandt, 67400 Illkirch Cedex, France

[1]Author for correspondence

Philippe Giegé
Institut de Biologie Moléculaire des Plantes du CNRS
Université de Strasbourg,
12 rue du Général Zimmer,
67084 Strasbourg, FRANCE
Fax: 33-(0)-3 8861 4442
Email: philippe.giege@ibmp-cnrs.unistra.fr

Running title: PPR9 is in plant mitochondria and nuclei.

Abstract

Following the acquisition of mitochondria by eukaryotic cells during endosymbiotic evolution, most of the genes in this organelle were transferred to the nucleus. To maintain mitochondrial biogenesis and function, the nuclear and mitochondrial genomes requires regulated and coordinated expression. In plant organelles, nuclear encoded proteins targeted to the organelles control post-transcriptional and post-translational mechanisms. Pentatricopeptide repeat proteins are good candidates to play such regulating roles. We report that loss-of-function mutations of *PPR9* are lethal for the embryo in *Arabidopsis thaliana*. *PPR9* encodes a novel pentatricopeptide repeat (PPR) protein dual-localized to mitochondria and nuclei as observed by GFP fusions, immunoblots using anti-PPR9 antibodies on sub-cellular fractions and immunohistochemistry experiments. Genetic complementation showed that loss of *PPR9* function in mitochondria (but not in nuclei) is lethal for the embryo. PPR9 is able to bind RNA. In mitochondria, it is associated with polysomes and may play a role in translation, whereas, in the nucleus, PPR9 interacts with the nucleosome assembly protein NAP1 and the transcription factor TCP8. Altogether, our findings suggest that PPR9 might be involved in coordinating gene expression between mitochondria and the nucleus.

Introduction

In eukaryotic cells, mitochondria originated from free-living bacterial ancestors (Gray et al., 2001). In plants, mitochondria have retained a genome. However, during evolution, the vast majority of the endosymbiont's genes have been transferred to the nucleus, allowing the nucleus to consolidate its genetic control over the organelle. As a result, the mitochondrial genomes encode a limited number of proteins which are mainly components of the respiratory chain, plus the transcription and translational machineries to express their genes (Burger et al., 2003). On the other hand, proteomic and genomic analyses of protein localization sequences estimate that organelles might contain up to several thousand different proteins (Millar et al., 2005). As these numbers imply, most proteins found in organelles are encoded in the nucleus, synthezised in the cytoplasm and then imported into the organelles. Hence, components of the respiratory chain as well as organellar ribosomes are multiprotein complexes made up of nuclearly encoded and mitochondrially encoded subunits. This implies the existence of mechanisms that coordinate the expression of organellar and nuclear genomes to ensure correct assembly of complexes and to maintain organelle function. However, we are still

largely ignorant of the mechanisms that lead to the coordination of expression of nuclear and organellar genomes in individual cells and tissues.

Among the estimated 1000 or more proteins that constitute the functional plant mitochondrion (Millar et al., 2004), only 33 proteins, three rRNAs, and 20 tRNAs are encoded in the *Arabidopsis thaliana* mitochondrial genome (Unseld et al., 1997). The remaining proteins and tRNAs are encoded in the nuclear genome and imported into the mitochondrion from the cytosol (Braun and Schmitz, 1999; Duchene et al., 2005; Salinas et al., 2008). Although our understanding of the signaling pathways between the nucleus and chloroplast has improved significantly in recent years (Pogson et al., 2008; Woodson and Chory, 2008), comparable studies on plant mitochondria have mainly focused on understanding the nuclear control of mitochondrial biogenesis e.g., (Zabaleta et al., 1998; Mackenzie and McIntosh, 1999; Edqvist and Bergman, 2002; Giraud et al., 2009). Studies in *Arabidopsis* by Brennicke and coworkers have shown that regulation of mitochondrial gene expression involves complex transcriptional and post-transcriptional processes, including RNA processing, intron splicing, RNA editing, and RNA stability (Binder and Brennicke, 2003). A comparative study of the transcriptional activities and steady state levels of the various mRNAs encoded in the *Arabidopsis* mitochondrial genome revealed little correlation between relative promoter activities and transcript abundance, suggesting extensive post-transcriptional regulation (Giegé et al., 2000).

Genetic analysis of mutants defective in organellar functions in yeast, *Chlamydomonas* and land plants has revealed many nuclear-encoded post-transcriptional regulators of organelle gene expression. It is likely that these proteins act as adaptors; they bind specific RNA transcripts and recruit translation machinery or RNA metabolism enzymes. Many of these regulators belong to the pentatricopeptide repeat (PPR) protein family. The PPR family is a eukaryote-specific protein family and is particularly large in land plants with more than 400 members in Arabidopsis. PPR proteins are characterized by the signature motif of a degenerate 35 amino-acids repeat often arranged in tandem arrays. Many PPR proteins are predicted to be targeted to either mitochondria or chloroplasts (Lurin et al., 2004). A rapidly growing number of publications relate the function of PPR proteins with post transcriptional processes in organelles (Schmitz-Linneweber and Small, 2008). Among them, the function of PPR proteins is connected to translation as suggested by the involvement of CRP1 as a chloroplast translation regulator (Schmitz-Linneweber et al., 2005), by the requirement of Pet309 for translation in yeast mitochondrial (Tavares-Carreon et al., 2008) and by the association of PPR336 with polysomes in plant mitochondria (Uyttewaal et al., 2008).

Many post-transcriptional steps are shared between organelles and the nucleus leading to the expectation to find proteins that are targeted to all three compartments in plant cells and that are involved in DNA or RNA metabolism (Small et al., 1998). Such factors would be good candidates in coordinating the expression of nuclear and organellar genomes (Krause and Krupinska, 2009). In most cases, the dual targeted proteins (nuclear-organellar proteins) that we know today are RNA and/or DNA binding proteins e.g. transcription factors, telomere-binding proteins such as Whirly proteins (Desveaux et al., 2005; Prikryl et al., 2008).

Here, we identified an *Arabidopsis* embryo-lethal mutant designated *ppr9*. *PPR9* encodes a novel pentatricopeptide repeat protein that is dual-localized to mitochondria and nuclei. PPR9 has RNA binding capacity and has an essential function in mitochondria that may be related to translation. We showed that PPR9 interacts with the nucleosome assembly protein NAP1 and the nuclear transcription factor TCP8 in yeast and in the nucleus of plant cells. Altogether, our findings suggest that PPR9, by interaction with both nuclear transcriptional factors and mitochondrial ribosomes is a potential coordinator of the expression of the two genomes.

Results

PPR9 is essential for embryo development

In order to gain insights into the involvement of PPR proteins in mitochondrial translation as suggested by the interaction of PPR336 with mitochondrial polysomes (Uyttewaal et al., 2008) we initially performed a reverse genetic screen of *Arabidopsis* mutants representing PPR genes of unknown function. The collections from the Syngenta *Arabidopsis* Insertion Library and Salk Institute were screened to look for T-DNA insertions in *Arabidopsis* PPR genes. We were unable to find any homozygous plants for two independent T-DNA insertion alleles for At5g60960 when we looked at progeny from a large number of heterozygous plants. We named this gene *PPR9* (Figure 1A). The dissection of siliques from heterozygous mutant plants showed that about one quarter of the embryos are aborted (Figure 1B), indicating that homozygous mutation of *PPR9* is lethal at an early stage of embryo development and thus that PPR9 has an essential function.

PPR9 is expressed in young tissues and pollen grains

To investigate the expression pattern of *PPR9* in different tissues, we used a promoter–β-glucuronidase (GUS) reporter system. The sequence between the start codon of *PPR9* and the

stop codon of the upstream gene was used to drive the *GUS* reporter gene (P_{PPR9}-*GUS* construct) and stably introduced in *Arabidopsis* by agroinfection. In five independent P_{PPR9}-*GUS* plant lines, GUS activity was visible in the apical root meristem, lateral root primordia, shoot apex and leaf primordia (Figure 1C). In addition, a high level of GUS activity was detected in the mature pollen grains. These data indicate that *PPR9* is mostly expressed in pollen grains and in actively dividing cells during vegetative growth.

PPR9 is dual localized in mitochondria and the nucleus

At5g60960 encodes a putative protein of 521 amino acids with an estimated molecular mass of 59 kDa. Sequence analysis revealed that the protein harbors 7 PPR motifs and two predicted localization signals (Figure 1), a mitochondrial target signal at its N terminus (MTS) and a putative nuclear localization signal (NLS) rich in basic residues at its C terminus. Alignment of PPR9 with its orthologs from *Oryza sativa* and *Vitis vinifera* shows that this basic region is highly conserved and thus might be functional (Figure 2). The presence of the two predicted localization signals suggested that PPR9 could be dual targeted to mitochondria and nuclei and led us to investigate the precise distribution of PPR9 in *Arabidopsis* cells. To analyze the sub-cellular localization of PPR9 in plant cells, we first transiently expressed in tobacco cells different GFP tagged proteins under the control of a 35S promoter (Figure 3A). The C-terminal translational fusion of the full length *PPR9* with GFP (PPR9_GFP) localizes to mitochondria of tobacco cells. The N-terminal fusion of GFP to PPR9 prevents its import in mitochondria and the GFP_PPR9 protein shows a nuclear-cytosolic localization. Deletion of the mitochondrial target signal (PPR9-MTS_GFP) restricts the localization of the protein to the nucleus. These results suggest that PPR9 harbors a signal targeting that can target the protein to the nucleus. To verify that this nuclear targeting property is indeed encoded by the C-terminal putative NLS signal, the last 174 bp of *PPR9* were fused in frame with *GUS* gene (GUS_NLS_GFP) encoding the cytosolic β-glucuronidase protein (GUS). The GUS_NLS-GFP protein was localized exclusively to the nucleus. This shows that the putative NLS signal is active in plant cells. To confirm the mitochondrial and nuclear localizations of PPR9 *in vivo*, immunodetection was performed on *Arabidopsis* sub-cellular fractions using polyclonal antibodies generated against the full-length PPR9. The antibodies detected a unique 55 kDa protein in the mitochondrial and nuclear fractions. This size corresponds to the predicted processed form of PPR9. The purity of the respective sub-cellular fractions was demonstrated with antibodies specific to each sub cellular fraction (Figure 3B). Localization was further investigated by immunohistochemistry. PPR9 antibodies were used to detect the protein in

histochemical sections from *Arabidopsis* seedlings. Signal was observed in structures that correspond to nuclei as indicated by the co-localisation with DAPI staining (Figure 3C). Signal specificity was controlled by substituting the PPR9 primary antibody with its preimmune serum. Altogether, the results show that PPR9 is a novel dual localized PPR protein in *Arabidopsis* cells.

The function of PPR9 is essential in mitochondria

The dual localization of PPR9 in *Arabidopsis* cells reveals that the protein plays an essential role in either mitochondria or the nucleus or in both organelles. To determine which localization is essential for embryogenesis, heterozygous 9B mutant plants were respectively complemented with constructs driven by the endogenous *PPR9* promoter and expressing either the wild type *PPR9* gene or truncated versions deleted of either the mitochondrial or the nuclear localization signal. The self-progeny of three independent transgenic and heterozygous 9B plants (T1) for each construct were analyzed by genotyping. The wild type transgene successfully complemented the *ppr9* mutation indicating that the embyro-lethal phenotype was indeed caused by the loss of function of *At5g60960*. In contrast, the truncated gene lacking a mitochondrial target signal did not complement the *ppr9* mutation. Finally, the version deleted of the C-terminal part of *PPR9* gene complemented the *ppr9* mutation and resulted in viable plants. These results indicate that PPR9 has an essential function in mitochondria and a distinct, although non-essential, role in the nuclei of plant cells.

PPR9 is an RNA binding protein

PPR proteins were predicted to be RNA binding proteins (Small and Peeters, 2000) and studies have already shown that some PPR proteins can indeed directly bind RNA (Tasaki et al., ; Okuda et al., 2006). Here, we used Biacore technology based on surface plasmon resonance (Jason-Moller et al., 2006) to establish whether PPR9 was also able to bind RNA. The full-length and mature recombinant PPR9 were immobilized on sensor chips and RNA solutions were used as potential ligands. High affinity binding was observed when *Arabidopsis* total mitochondrial RNA was used with both versions of PPR9. The release of RNA from PPR9 was slow which suggests that the interaction with RNA is stable rather than transient (Figure 4A). When oligoribonucleotides representing poly A, C, G and U were used, PPR9 was able to bind poly G with the highest affinity (Supplemental Figure 1) as already observed for other PPR proteins (Lurin et al., 2004). Thioredoxin was used as a negative control to show that the affinity response observed was indeed due to RNA / protein

interaction and not to artefactual binding of RNA to sensor chips. Altogether, the results indicate that PPR9 is able to bind RNA.

PPR9 is associated with polysomes in plant mitochondria

PPR9 plays an essential function in *Arabidopsis* mitochondria. Its capacity to bind RNA suggests that its function could be related to an essential post-transcriptional process in mitochondria, as observed for a rapidly growing number of PPR proteins in plant organelles (Schmitz-Linneweber and Small, 2008). We first investigated the submitochondrial localization of PPR9. Mitochondria were fractionated into mitoplasts (i.e., matrix and inner membrane), a soluble fraction and a membrane fraction. Mitochondrial membrane proteins were further fractionated into extrinsic and intrinsic proteins by carbonate treatment. PPR9 was detected in the mitoplast, soluble and extrinsic membrane protein fractions (Figure 4B). These results indicate that PPR9 is a soluble protein in the mitochondrial matrix and a peripheral protein of the inner membrane as well. The quality of the fractions was controlled with specific antibodies. Subsequently, the association of PPR9 with complexes containing RNA in mitochondria was investigated through the immunodetection of PPR9 in complexes separated by sucrose density gradients. For this analysis, mitochondria were extracted from cauliflower and not from *Arabidopsis* due to the necessity of large amounts of pure mitochondria necessary for this biochemical approach (i.e., the equivalent of 100 mg of mitochondrial proteins). Prior to this, we controlled that PPR9 was also detected in the cauliflower mitochondrial extracts with our antibody (data not shown). Mitochondrial extracts were separated on sucrose gradients, and fractions representing the entire gradients were collected. Equivalent amounts of proteins from each fraction were separated on SDS-PAGE, blotted and reacted with PPR9 antibodies. The 55-kDa signal was detected for fractions at the bottom of the gradients (Figure 4C). The same protein blots were reacted with antibodies specific for the ribosomal protein RPS1 and NAD9 from respiratory complex I. NAD9 was only detected in the top fractions of the gradients, suggesting that they contain complexes of sizes up to 2000 kDa. On the other hand, RPS1 was detected in the same bottom fractions of the gradient as PPR9. These top fractions were proposed to contain free ribosomes, whereas the bottom ones were proposed to contain polysomes (Delage et al., 2007). This suggested that PPR9 could be associated with polysomes. Mitochondrial samples were subjected to treatments leading to the destabilization of polysomes, in order to demonstrate this assumption. Samples were treated with puromycin and RNases. Puromycin specifically destabilises ribosomes (Lu and Draper, 1994) whereas RNase results in the degradation of

RNA necessary to maintain polysome integrity. In all cases, after treatment and separation on gradients, PPR9, similar to RPS1, was no longer detected in the fractions at the bottom of the gradients (Figure 4C). This confirms that PPR9 is indeed associated with polysomes in an RNA dependent manner in *Arabidopsis* mitochondria, which is consistent with the capacity of PPR9 to bind RNA, although the precise nature of this association with polysomes is unknown.

In order to identify the precise RNA targets of PPR9 in mitochondria, extensive efforts were deployed to coimmunoprecipitate PPR9 bound to RNA partners from a solubilized *Arabidopsis* mitochondrial extract with either anti-PPR9 serum or purified anti-PPR9 polyclonal antibodies as described previously (Schmitz-Linneweber et al., 2005). Surprisingly, we were unable to immunoprecipitate PPR9 presumably because the protein is in tight complex with polysome components that hide epitopes from antibodies (data not shown).

PPR9 interacts with nuclear transcription factors

The identification of protein partners should give clues to understanding the function(s) of PPR9. For this a genetic screen using a system derived from yeast two-hybrid was performed. The DUAL hunter system, (DualsystemsBiotech®) enables the use of a nuclear protein as a bait. Briefly, a bait protein of interest is inserted in frame with the membrane protein Ost4p, the C-terminal half of ubiquitin and the transcription factor LexA-VP16 (X-Cub constructs). A second protein of interest (the prey) is fused at the C-terminal of a mutated version of the N-terminal domain of ubiquitin (Nub-X constructs). If bait and prey interact, Cub and NubG complement to form split-ubiquitin. Then, ubiquitin-specific proteases release LexA-VP16 that migrates to the nucleus and activates the transcription of reporter genes (Mockli et al., 2007). We screened an *Arabidopsis* seedling cDNA library for proteins that interact with PPR9. Of 10^7 transformants screened, eleven positive clones were identified. Eight different cDNA clones contained sequences of *At4g26110* encoding NAP1, the Nucleosome Assembly Protein 1. The cDNAs rescued from the three other clones represented respectively, a nuclear transcriptional factor TCP8 (AT1G58100) and two proteins of unknown functions (AT5G24680 and AT1G44920). AT5G24680 contains a ubiquitin fold modifier-specific peptidase domain (IPR012462) suggesting that the protein could interact with the ubiquitin protein fused to PPR9 rather than with PPR9 itself. AT5G24680 and AT1G44920 do not contain any predictable localization or function in mitochondria and / or the nucleus and were thus not further investigated at this stage. The full-length mature PPR9, NAP1 and TCP8

cDNA sequences were subsequently cloned in the respective DUAL hunter vectors to confirm protein interactions (Figure 5A). The results showed that PPR9 is indeed able to interact directly with both NAP1 and TCP8, two proteins involved in gene expression and localized in plant nuclei.

The protein interactions observed in yeast were investigated *in planta*. For this, we used the bimolecular fluorescence complementation (BiFC) approach. This method relies on expression of the two proteins of interest as translational fusions either to the non-fluorescing N-terminal (YN) or C-terminal (YC) halves of yellow fluorescent protein (YFP). Only when the YFP halves are brought together by interaction of the fused proteins can they form a functional YFP fluorophore, which can be detected with standard epifluorescence microscopy equipment (Hu et al., 2002). A set of vectors was generated to express PPR9, NAP1 and TCP8 fused with the N-terminal (YN) or the C-terminal (YC) part of the YFP after transient transformation of onion epidermal cells using particle bombardment. YFP fluorescence was observed in the nucleus of plant cells when the vectors encoding PPR9 were co-bombarded in combination with the vectors expressing NAP1 and TCP8 respectively (Figure 5B). These results show that PPR9 interacts physically with NAP1 and TCP8 in the nuclei of living plant cells and confirms interaction with the protein partners obtained by the yeast split-ubiquitin screen.

Discussion

Through genetic studies, we have shown that the loss of *PPR9* is lethal for the embryo in *Arabidopsis thaliana*. PPR9 encodes a pentatricopeptide repeat protein. PPR proteins play crucial functions in plant organellar gene expression associated with RNA cleavage (Hashimoto et al., 2003) (Kazama et al., 2008), RNA processing (Meierhoff et al., 2003; Hattori et al., 2004), RNA splicing (Schmitz-Linneweber et al., 2006; de Longevialle et al., 2007), RNA editing e.g. (Kotera et al., 2005; Chateigner-Boutin et al., 2008; Hammani et al., 2009), and translational activation (Schmitz-Linneweber et al., 2005). Several reports have shown that loss of function of PPR genes in *Arabidopsis* often causes an embryo-lethal phenotype (Lurin et al., 2004; Schmitz-Linneweber et al., 2006; Gobert et al., 2010).

TargetP (Emanuelsson et al., 2000) and Predotar (Small et al., 2004) programs predict three-quarters of PPR proteins to be targeted to mitochondria or plastids leading to the possiblity that some PPR proteins could function outside of organelles (Lurin et al., 2004; O'Toole et al., 2008). This has been demonstrated with GRP23 (Glutamine Rich Protein23), the only known

PPR protein located in the nucleus in plants (Ding et al., 2006). The GRP23 protein interacts physically with subunit III of RNA polymerase II through its C-terminal Gln-rich WQQ domain. Although the exact role of GRP23 has not been elucidated, GRP23 is speculated to play a role in the nucleus as a transcriptional regulator by interacting with the RNA polymerase II subunit.

In this report, we showed that PPR9 may play a role outside of organelles. PPR9 protein harbors two sub-cellular localization signals: a C-terminal bipartite Nuclear Localization Signal and a N-terminal Mitochondrial Target Sequence. GFP fusions and immunodetection on sub-cellular fractions have shown that PPR9 is indeed dual localized to nuclei and mitochondria. PPR9 is expected to play a dual function in the nucleus and mitochondria in *Arabidopsis* cells. Functional analysis of *PPR9* knock-out *Arabidopsis* plants with site-specific accumulation of PPR9 in either the nucleus or mitochondria showed that only the mitochondrial localization of PPR9 is required for the embryo development. This result suggests that PPR9 represents a novel dual targeted PPR protein playing an essential function for embryogenesis in mitochondria and a distinct role in the nucleus.

In plant mitochondria, PPR9 is associated with ribosomes in an RNA-dependent manner. It is possible that the association of PPR9 with polysomes is indirect - that PPR9 is part of an RNA maturation complex that could be attached to polysomes. Alternatively, it is possible that PPR9 interacts directly with ribosomes or other components of the translation apparatus, thus having a direct role in mitochondrial translation or its control.

In the nucleus, PPR9 can interact with both TCP8 and NAP1, two nuclear factors involved in the control of gene expression. NAP1 is the Nucleosome Assembly Protein 1. Nucleosome assembly proteins are conserved from yeast to human and facilitate the *in vitro* assembly of nucleosomes as histone chaperones. They shuttle histones from the cytosol into the nucleus, assemble nucleosomes, and promote chromatin fluidity, thereby affecting the transcription of many genes (Ishimi and Kikuchi, 1991; Simon et al., 1994; Yoon et al., 1995; Ito et al., 1996; Rodriguez et al., 1997; Shen et al., 2001; Dong et al., 2003; Dong et al., 2005; Liu et al., 2009). In higher plants, several NAP1 homologs have been identified in a given species including tobacco, rice, maize, and *Arabidopsis* (Dong et al., 2003; Liu et al., 2009). In *Arabidopsis* and tobacco, NAP1 proteins bind histones thus suggesting a conserved function of these proteins among eukaryotes (Dong et al., 2005; Liu et al., 2009). The functional reason for the association of PPR9 with NAP1 is unknown. However, PPR9 might be part of a chromatin remodeling complex involved in gene expression control in the nucleus. Alternatively, PPR9 might associate with NAP1 for its transit from the cytosol to the nucleus.

In contrast, the TCP class of genes is found only in plants and is represented by the first three identified members *TEOSINTE BRANCHED1* in maize, *CYCLOIDEA* in *Antirrhinum majus*, and PCF coding gene in rice (Cubas et al., 1999). Members belonging to this class are important regulators of plant growth and development and control multiple traits in diverse plant species, including flower and petal asymmetry, plant architecture (Aguilar-Martinez et al., 2007), leaf morphogenesis (Ori et al., 2007) and senescence (Schommer et al., 2008), embryo growth (Tatematsu et al., 2008), and circadian rhythm (Pruneda-Paz et al., 2009). The TCP genes code for transcription factors that share a ≈60-residue homologous region called the TCP domain (Cubas et al., 1999), common to all the members. The TCP domain mediates binding of the TCP proteins to GC-rich DNA sequence motifs *in vitro* (Kosugi and Ohashi, 1997). These motifs, called site II elements, have been identified as *cis*-elements in many plant genes coding for proteins such as *PCNA* (Kosugi and Ohashi, 1997), *RADIALIS* (Costa et al., 2005), *LIPOXIGENASE2* (Schommer et al., 2008), *CIRCADIAN ASSOCIATED1* (Pruneda-Paz et al., 2009). Welchen et Gonzalez have observed the prevalence of site II-elements in the promoters of nuclear genes encoding components of the oxidative phosphorylation (OxPhos) machinery from both *Arabidopsis* and rice and have thus proposed that TCP transcription factors might be involved in the coordinated expression of the nuclear encoded subunits of the mitochondrial OxPhos machinery (Welchen and Gonzalez, 2005, 2006).

Biogenesis of the OxPhos complexes requires the expression of two separate genomes within one cell. This is because some of the proteins that compose these complexes are encoded within the organelle genome, while the rest are encoded in the nucleus (Mackenzie and McIntosh, 1999). It is generally assumed that the expression of genes encoded in these genomes must be somehow coordinated to ensure correct complex assembly. A global study of the effect of sucrose starvation on mitochondrial biogenesis suggests that coordination of the expression of genes encoded in the nuclear and the mitochondrial compartments occurs at the post-translational level (Giege et al., 2005). Regarding the coordination of nuclear genes themselves, it is more likely that regulation takes place at the transcriptional level, perhaps through the interaction of common sets of transcription factors like TCP with cognate binding sites present in the respective promoters (Welchen and Gonzalez, 2006). If this is true, the activity of the nuclear transcription factors should be modulated by a retrograde signal coming from mitochondria.

It has been hypothesized that proteins that are dual-targeted to both the nucleus and an organelle play a role in this coordinating mechanism (Krause and Krupinska, 2009). Thus, we

proposed that PPR9 represents a novel PPR protein and a potential regulator of the expression of two distinct genomes by its association with the ribosomes in mitochondria and transcriptional factors in nuclei of plant cells. Despite the fact that an *Arabidops*is *PPR9* knock-out mutant only accumulating PPR9 in mitochondria does not show obvious macroscopic phenotypes under standard growth conditions, it would be interesting to compare gene expression profiles between the mutant and the wild-type plants. Finally, it is possible that the precise role of PPR9 in the nucleus might only be revealed during particular developmental stages or in response to specific environmental stimuli.

Methods

All the primers used in this study are listed in Supplemental Table 1.

Plant Material

Arabidopsis thaliana ecotype Columbia (Col-0) was used in this study. The T-DNA insertion mutant lines were obtained from the ABRC Stock Center.

Arabidopsis thaliana var. *Landsberg erecta* suspension cultures were used for cell fractionation.

Genetic Analysis

Total cellular DNA was isolated as described by Edwards et al. (1991). Plants were genotyped for homozygous lines by PCR and the insertion position was confirmed by sequencing with a T-DNA left border primer. *ppr9B* (SAIL_868_D10, Col-0, insertion site +958), *ppr9C* (SAIL_224_H01, Col-0, insertion site +971).

Genetic complementation

The 2218 bp containing the coding sequence of *PPR9* and the 5'-intergenic region or the 2040 pb containing the 5'-intergenic region and the coding sequence of *PPR9* truncated in the last 171 bp encoding the NLS signal were amplified by PCR using Phusion DNA polymerase (Finnzymes) with primers listed in Supplemental Table 1 containing the *attB* sites for Gateway® cloning according to the manufacturer's instructions (Invitrogen). The chimeric gene *PPR9-MTS* encoding PPR9 without its mitochondrial target sequence and driven by the

same 5'-intergenic region was amplified by mixing two 15bp-overlapping PCR fragments encoding respectively the 5'-intergenic and the 60-521 amino acids of PPR9.
These products were cloned into pGWB1 binary vector and introduced into *9B* heterozygote mutants via *Agrobacterium tumefaciens* GV3101. Transformants were obtained by selection on MS agar plates containing 50µg/mL kanamycin and confirmed by PCR.

Promoter-GUS Fusion Analysis

The 5′ intergenic region of *PPR9* (652 nucleotides between the initiation codon of At5g60960 and the stop codon of the upstream gene) was cloned into the binary vector pGWB3 upstream of the *GUS* gene and introduced into *Arabidopsis* plants (Columbia) via *Agrobacterium tumefaciens* GV3101. Transformants were obtained by selection on MS agar plates containing 50µg/mL kanamycin and confirmed by PCR. Tissues from promoter-GUS fusion plants issued from five independent lines were stained with 5-bromo-4-chloro-3-indolyl-β-glucuronic acid (Biosynth) as described by Jefferson et *al.,* 1987.

Protein Expression and antibodies production

AT5G60960 (*PPR9*) full-length coding region without the termination codon (1563 bp) was amplified by PCR and cloned into pBAD/Thio TOPO® (Invitrogen) in fusion with an N-terminal thioredoxin and C-terminal polyhistidine tags. Proteins were expressed for 2 h at 37°C in LMG194 *E. coli* induced with 0,2% arabinose (p/v). Bacteria were lysed in lysis buffer (100 mM NaH_2PO_4, 10 mM Tris base, 6 M guanidine hydrochloride, 10 mM imidazole, 1X complete protease inhibitor EDTA free cocktail (Roche), pH 8) and centrifuged at 20000 g for 30 min 4°C. The cleared lysate was incubated with Ni-NTA resin (Qiagen) over-night. Bound proteins were washed twice with wash buffer 1 (100 mM NaH_2PO_4, 150 mM NaCl, 8 M urea, 20 mM imidazole, pH 8) followed by extensive washing with renaturation buffer (50 mM NaH_2PO_4, 500 mM NaCl, 20 mM imidazole, pH 8) to eliminate urea. Proteins were eluted using renaturation buffer supplemented with 250 mM imidazole. Purified PPR9 protein was used to immunize rabbit to produce polyclonal antibodies. For RNA binding assay, the fragment encoding PPR9 without the predicted mitochondrial transit peptide (amino acids 60-521) was cloned into the same vector expressed and purified with the same procedure.

Analysis of targeting via GFP fusions

PPR9, *PPR9-MTS, GUS*, *GUS-NLS* coding sequences were amplified using Phusion DNA polymerase (Finnzymes) with primers listed in Supplemental Table 1 containing the *attB* sites for Gateway® cloning according to the manufacturer's instructions (Invitrogen). The chimeric gene *GUS-NLS* was amplified by mixing two 15bp-overlapping PCR fragments encoding respectively GUS and the last 57 amino acids of PPR9. The vectors used for N-terminal or C-terminal GFP fusion were kindly provided by Prof. James Whelan (The University of Western Australia) (Carrie et al., 2009). Biolistic transformations of GFP constructs were performed on tobacco BY-2 cell culture as described by (Haas et al., 2005). The GFP construct was co-precipitated onto gold particles and transformed using the biolistic PDS-1000/He system (Bio-Rad, http://www.bio-rad-com/). Particles were bombarded onto 3 mL of tobacco BY-2 cell suspension resting on filter paper on osmoticum plates. After bombardment, the cells were placed in the dark at 22°C for 6 h. Observation of transient GFP expression was performed using a Zeiss LSM510 confocal laser scanning microscope (Zeiss).

Immunodetection

Sera were used at dilutions of 1:5000 for PPR9 antibodies, of 1:100,000 for wheat NAD9 (Lamattina et al., 1993), of 1:10,000 for wheat RPS1, of 1:10,000 for tobacco MnSOD (Bowler et al., 1989), of 1:50,000 for yeast $Cytc_1$ (G. Schatz, Basel University, Basel, Switzerland), of 1:10,000 for RuBisCo (B. Camara, Strasbourg, France) and of 1:2000 for histone H2B antibodies (W.H. Shen, Strasbourg, France).

Immunohistochemistry

Arabidopsis seedlings were fixed in FAA (10 % (v/v) formol, 5 % (v/v) acetic acid, and 50 % (v/v) ethanol), embedded in paraffin and sectioned as previously described (Mundel et al., 2000). Pretreatment of the samples included deparaffination and rehydration. The samples were washed twice in 1 x PBS, Triton X-100 (0.05 %) and BSA (1 %), and then blocked with 1 x PBS, Triton X-100 (0.05 %), BSA (1 %) and goat serum (5 %) for 15 min. Incubation with primary PPR9 antibodies (dilution 1: 500) was performed overnight at 4°C in 1 x PBS, 0.01% Triton X-100, and 1 % BSA. After two washes, incubation with anti-rabbit alkaline phosphatase-conjugated antibodies (dilution 1:1000) in 1 x PBS was performed for 1 h at room temperature. Two washes in 1 x PBS were followed by one wash in Fast RED buffer (Tris HCl 0.1 M, pH 8.2) for 5 min at room temperature. Detection of the signal was performed with Fast RED/Naphthol AS-Mix as substrate according to manufacturers

instructions (Sigma, St Louis, USA). The reaction was stopped by rinsing the sections in stop-buffer (10 mM Tris HCl, pH 8.0, 1 mM EDTA) for 5 min and in water (5 min). Immunolabelled samples were observed by light microscopy. The specificity of labelling was checked by substituting the PPR9 primary antibody with its preimmune serum. Nuclei were visualised by DAPI staining with a Nikon E-800 flurescent microscope.

***Arabidopsis* cell fractionation**

Mitochondria and chloroplasts were isolated as described previously (Spielewoy et al., 2001; Giegé et al., 2003). Nuclei were extracted from protoplasts. 10^6 protoplasts were resuspended in 1 ml of NIB buffer (10 mM MES KOH pH 5.5, 0.2 M sucrose, 2.5 mM EDTA and 2.5 mM DTT). After protoplasts lysis, the homogenate was loaded on, 2.3 M sucrose, 1X NIB, 40 % percoll (v/v) gradients. Gradients were centrifuged for 30 min at 3200 g and nuclei were collected at the percoll/sucrose interphase, washed and resuspended with 0.35 M sucrose and 0.5 mM $MgCl_2$. For mitochondria sub-fractionation, mitochondria were fractionated into mitoplast, membrane and soluble fractions as described previously (Spielewoy et al., 2001).

RNA binding assays

Experiments were performed at 25°C on a Biacore 2000 optical biosensor instrument (GE Healthcare, Biacore). The running buffer was composed of 50 mM phosphate buffer pH 8, 250 mM NaCl, 0.1 % (v/v) triton X-100 and 10 mM $MgCl_2$. Recombinant PPR9 proteins that possess C-terminal histidine tags were bound to anti-His antibodies immobilised on a CM5 *sensor* surface using standard amine-coupling chemistry. Briefly, The carboxymethyl dextran surface was activated with an injection of a 1:1 ratio of 0.4 M EDC (1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride) and 0.1 M NHS (N-hydroxysuccinimide). The antibodies at 10 µg/ml diluted in 10 mM sodium acetate pH 6 were injected on the surface. Remaining activated groups were blocked with 1 M ethanolamine pH 8.5. Histidine-tagged proteins diluted in the running buffer at 300 ng / µl were injected and captured on the flow cells. For molecular binding assays, RNA molecules at 10 ng / µl were injected on the flow cells at 10 µl / min for 120 s. Dissociation was monitored for 120 s and was followed by a washing step of 30 s with 2 M NaCl when necessary. Responses were automatically zeroed before injection start. The sensorgrams were corrected for signals in the reference flow cell that contains the control thioredoxin. Data were interpreted using the BiaEvaluation 3.2 software (GE Healthcare).

Preparation and analysis of mitochondrial polysome-enriched fractions

Polysome-enriched fractions were prepared as described previously (Uyttewaal et al., 2008) at the exception that polysomes were isolated from cauliflower mitochondria. Succinctly, a lysate obtained from 100 mg of cauliflower (*Brassica oleracea*) mitochondria was centrifuged through a 61 % sucrose cushion at 235,000 ***g*** for 2 h. The polysome-containing pellet was further separated on 10–35 % sucrose gradients centrifuged at 155,000 ***g*** for 2 h. Fractions 1–10 (from top to bottom) were collected on the gradients and analysed on 10–17 % SDS-PAGE. For polysome destabilisation experiments, lysates were treated with 10 mM puromycin, or 600 U/ml of RNase A.

Yeast DUALhunter Screen

The DUAL hunter system (DualsystemsBiotech®) was used to identify PPR9 protein partners. A PPR9 cDNA sequence corresponding to codons 60 to 521 was cloned with *Sfi*I technology in pDHB1 in order to create the bait vector that was transformed in the yeast strain NMY51. The expression of the PPR9-Cub fusion protein was tested by western blot with PPR9 and VP16 specific antibodies. The bait strain was transformed according to manufacturers instructions with a commercially available library expressing Nub fused to cDNAs. The library was constructed with total RNA from a mixture of tissues from six days old *Arabidopsis* seedlings (DualsystemsBiotech®). Positive clones were selected for the activation of the three reporter genes *ADE2*, *HIS3* and *lacZ*. Full length cDNAs representing the positive clones were cloned in both pDHB1 and pPR3-N (the library vector) to confirm the identifications in pairwise interaction experiments.

Bimolecular fluorescence complementation assay

PPR9, NAP1 and TCP8 cDNAs were cloned with the gateway technology in pMAV-YN-GW and pMAV-YC-GW vectors to obtain constructs expressing C-terminal fusions of the proteins of interest with the N terminal part of YFP (YN) and the C terminal part of YFP (YC) respectively. Pairs of YC and YN constructs were cotransformed in onion epidermal cells by biolistic bombardment as described previously (Marrocco et al., 2006). In order to identify transformed cells and to localize nuclei, a plasmid expressing CFP fused to the parsley common plant regulatory protein 2 (CPRF2), which localizes to the nucleus was included during bombardment. Images were recorded 16 h after tranformation with a Nikon E800 fluorescent microscope equiped with YFP and CFP specific filters.

Accession number

The genes described in this article correspond to the following *Arabidopsis* Genome Initiative codes: At5g60960 (*PPR9*), At1g58100 (TCP8) and At4g26110 (NAP1).

Acknowledgements

This work is funded by the French Centre National de la Recherche Scientifique, by research grant 07-JCJC-0123 from the Agence Nationale de la Recherche (ANR) to PG, by the Australian Research Council Centres of Excellence scheme (CE0561495) and scholarships from the Region of Alsace and the University of Western Australia. KH is the holder of a Lavoisier Fellowship and IS is a West Australian State Premier's Fellow.

References

Aguilar-Martinez, J.A., Poza-Carrion, C., and Cubas, P. (2007). Arabidopsis BRANCHED1 acts as an integrator of branching signals within axillary buds. Plant Cell **19,** 458-472.

Binder, S., and Brennicke, A. (2003). Gene expression in plant mitochondria: transcriptional and post-transcriptional control. Philos Trans R Soc Lond B Biol Sci **358,** 181-188; discussion 188-189.

Bowler, C., Alliotte, T., De Loose, M., Van Montagu, M., and Inze, D. (1989). The induction of manganese superoxide dismutase in response to stress in Nicotiana plumbaginifolia. EMBO J **8,** 31-38.

Braun, H.P., and Schmitz, U.K. (1999). The protein-import apparatus of plant mitochondria. Planta **209,** 267-274.

Burger, G., Gray, M.W., and Lang, B.F. (2003). Mitochondrial genomes: anything goes. Trends Genet **19,** 709-716.

Carrie, C., Kuhn, K., Murcha, M.W., Duncan, O., Small, I.D., O'Toole, N., and Whelan, J. (2009). Approaches to defining dual-targeted proteins in Arabidopsis. Plant J **57,** 1128-1139.

Chateigner-Boutin, A.L., Ramos-Vega, M., Guevara-Garcia, A., Andres, C., de la Luz Gutierrez-Nava, M., Cantero, A., Delannoy, E., Jimenez, L.F., Lurin, C., Small, I., and Leon, P. (2008). CLB19, a pentatricopeptide repeat protein required for editing of rpoA and clpP chloroplast transcripts. Plant J **56,** 590-602.

Costa, M.M., Fox, S., Hanna, A.I., Baxter, C., and Coen, E. (2005). Evolution of regulatory interactions controlling floral asymmetry. Development **132,** 5093-5101.

Cubas, P., Lauter, N., Doebley, J., and Coen, E. (1999). The TCP domain: a motif found in proteins regulating plant growth and development. Plant J **18,** 215-222.

de Longevialle, A.F., Meyer, E.H., Andres, C., Taylor, N.L., Lurin, C., Millar, A.H., and Small, I.D. (2007). The pentatricopeptide repeat gene OTP43 is required for trans-splicing of the mitochondrial nad1 Intron 1 in Arabidopsis thaliana. Plant Cell **19,** 3256-3265.

Delage, L., Giegé, P., Sakamoto, M., and Marechal-Drouard, L. (2007). Four paralogues of RPL12 are differentially associated to ribosome in plant mitochondria. Biochimie.

Desveaux, D., Marechal, A., and Brisson, N. (2005). Whirly transcription factors: defense gene regulation and beyond. Trends Plant Sci **10,** 95-102.

Ding, Y.H., Liu, N.Y., Tang, Z.S., Liu, J., and Yang, W.C. (2006). Arabidopsis GLUTAMINE-RICH PROTEIN23 is essential for early embryogenesis and encodes a novel nuclear PPR motif protein that interacts with RNA polymerase II subunit III. Plant Cell **18,** 815-830.

Dong, A., Zhu, Y., Yu, Y., Cao, K., Sun, C., and Shen, W.H. (2003). Regulation of biosynthesis and intracellular localization of rice and tobacco homologues of nucleosome assembly protein 1. Planta **216,** 561-570.

Dong, A., Liu, Z., Zhu, Y., Yu, F., Li, Z., Cao, K., and Shen, W.H. (2005). Interacting proteins and differences in nuclear transport reveal specific functions for the NAP1 family proteins in plants. Plant Physiol **138,** 1446-1456.

Duchene, A.M., Giritch, A., Hoffmann, B., Cognat, V., Lancelin, D., Peeters, N.M., Zaepfel, M., Marechal-Drouard, L., and Small, I.D. (2005). Dual targeting is the rule for organellar aminoacyl-tRNA synthetases in Arabidopsis thaliana. Proc Natl Acad Sci U S A **102,** 16484-16489.

Edqvist, J., and Bergman, P. (2002). Nuclear identity specifies transcriptional initiation in plant mitochondria. Plant Mol Biol **49,** 59-68.

Emanuelsson, O., Nielsen, H., Brunak, S., and von Heijne, G. (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J Mol Biol **300,** 1005-1016.

Giege, P., Sweetlove, L.J., Cognat, V., and Leaver, C.J. (2005). Coordination of nuclear and mitochondrial genome expression during mitochondrial biogenesis in Arabidopsis. Plant Cell **17,** 1497-1512.

Giegé, P., Sweetlove, L., and Leaver, C. (2003). Identification of mitochondrial protein complexes in Arabidopsis using two-dimensional Blue-Native polyacrylamide gel electrophoresis. Plant Mol. Biol. reporter **21,** 133-144.

Giegé, P., Hoffmann, M., Binder, S., and Brennicke, A. (2000). RNA degradation buffers asymmetries of transcription in Arabidopsis mitochondria. EMBO Rep **1,** 164-170.

Giraud, E., Van Aken, O., Ho, L.H., and Whelan, J. (2009). The transcription factor ABI4 is a regulator of mitochondrial retrograde expression of ALTERNATIVE OXIDASE1a. Plant Physiol **150,** 1286-1296.

Gobert, A., Gutmann, B., Taschner, A., Gößringer, M., Holzmann, J., Hartmann, R.K., Rossmanith, W., and Giegé, P. (2010). A single Arabidopsis organellar protein has RNase P activity. Nature structural & molecular biology **in press**.

Gray, M.W., Burger, G., and Lang, B.F. (2001). The origin and early evolution of mitochondria. Genome Biol **2,** REVIEWS1018.

Hammani, K., Okuda, K., Tanz, S.K., Chateigner-Boutin, A.L., Shikanai, T., and Small, I. (2009). A study of new Arabidopsis chloroplast RNA editing mutants reveals general features of editing factors and their target sites. Plant Cell **21,** 3686-3699.

Hashimoto, M., Endo, T., Peltier, G., Tasaka, M., and Shikanai, T. (2003). A nucleus-encoded factor, CRR2, is essential for the expression of chloroplast ndhB in Arabidopsis. Plant J **36,** 541-549.

Hattori, M., Hasebe, M., and Sugita, M. (2004). Identification and characterization of cDNAs encoding pentatricopeptide repeat proteins in the basal land plant, the moss Physcomitrella patens. Gene **343,** 305-311.

Hu, C.D., Chinenov, Y., and Kerppola, T.K. (2002). Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation. Mol Cell **9,** 789-798.

Ishimi, Y., and Kikuchi, A. (1991). Identification and molecular cloning of yeast homolog of nucleosome assembly protein I which facilitates nucleosome assembly in vitro. J Biol Chem **266,** 7025-7029.

Ito, T., Bulger, M., Kobayashi, R., and Kadonaga, J.T. (1996). Drosophila NAP-1 is a core histone chaperone that functions in ATP-facilitated assembly of regularly spaced nucleosomal arrays. Mol Cell Biol **16,** 3112-3124.

Jason-Moller, L., Murphy, M., and Bruno, J. (2006). Overview of Biacore systems and their applications. Curr Protoc Protein Sci **Chapter 19,** Unit 19 13.

Kazama, T., Nakamura, T., Watanabe, M., Sugita, M., and Toriyama, K. (2008). Suppression mechanism of mitochondrial ORF79 accumulation by Rf1 protein in BT-type cytoplasmic male sterile rice. Plant J **55,** 619-628.

Kosugi, S., and Ohashi, Y. (1997). PCF1 and PCF2 specifically bind to cis elements in the rice proliferating cell nuclear antigen gene. Plant Cell **9,** 1607-1619.

Kotera, E., Tasaka, M., and Shikanai, T. (2005). A pentatricopeptide repeat protein is essential for RNA editing in chloroplasts. Nature **433,** 326-330.

Krause, K., and Krupinska, K. (2009). Nuclear regulators with a second home in organelles. Trends Plant Sci **14,** 194-199.

Lamattina, L., Gonzalez, D., Gualberto, J., and Grienenberger, J.M. (1993). Higher plant mitochondria encode an homologue of the nuclear-encoded 30-kDa subunit of bovine mitochondrial complex I. Eur J Biochem **217,** 831-838.

Liu, Z., Zhu, Y., Gao, J., Yu, F., Dong, A., and Shen, W.H. (2009). Molecular and reverse genetic characterization of NUCLEOSOME ASSEMBLY PROTEIN1 (NAP1) genes unravels their function in transcription and nucleotide excision repair in Arabidopsis thaliana. Plant J **59,** 27-38.

Lu, M., and Draper, D.E. (1994). Bases defining an ammonium and magnesium ion-dependent tertiary structure within the large subunit ribosomal RNA. J Mol Biol **244,** 572-585.

Lurin, C., Andres, C., Aubourg, S., Bellaoui, M., Bitton, F., Bruyere, C., Caboche, M., Debast, C., Gualberto, J., Hoffmann, B., Lecharny, A., Le Ret, M., Martin-Magniette, M.L., Mireau, H., Peeters, N., Renou, J.P., Szurek, B., Taconnat, L., and Small, I. (2004). Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. Plant Cell **16,** 2089-2103.

Mackenzie, S., and McIntosh, L. (1999). Higher plant mitochondria. Plant Cell **11,** 571-586.

Marrocco, K., Zhou, Y., Bury, E., Dieterle, M., Funk, M., Genschik, P., Krenz, M., Stolpe, T., and Kretsch, T. (2006). Functional analysis of EID1, an F-box protein involved in phytochrome A-dependent light signal transduction. Plant J **45,** 423-438.

Meierhoff, K., Felder, S., Nakamura, T., Bechtold, N., and Schuster, G. (2003). HCF152, an Arabidopsis RNA binding pentatricopeptide repeat protein involved in the processing of chloroplast psbB-psbT-psbH-petB-petD RNAs. Plant Cell **15,** 1480-1495.

Millar, A., Day, A., and Whelan, J. (2004). Mitochondrial biogenesis and function in Arabidopsis. (Rockville, MD).

Millar, A.H., Heazlewood, J.L., Kristensen, B.K., Braun, H.P., and Moller, I.M. (2005). The plant mitochondrial proteome. Trends Plant Sci **10,** 36-43.

Mockli, N., Deplazes, A., Hassa, P.O., Zhang, Z., Peter, M., Hottiger, M.O., Stagljar, I., and Auerbach, D. (2007). Yeast split-ubiquitin-based cytosolic screening system to

detect interactions between transcriptionally active proteins. Biotechniques **42,** 725-730.

Mundel, C., Baltz, R., Eliasson, A., Bronner, R., Grass, N., Krauter, R., Evrard, J.L., and Steinmetz, A. (2000). A LIM-domain protein from sunflower is localized to the cytoplasm and/or nucleus in a wide variety of tissues and is associated with the phragmoplast in dividing cells. Plant Mol Biol **42,** 291-302.

O'Toole, N., Hattori, M., Andres, C., Iida, K., Lurin, C., Schmitz-Linneweber, C., Sugita, M., and Small, I. (2008). On the expansion of the pentatricopeptide repeat gene family in plants. Mol Biol Evol **25,** 1120-1128.

Okuda, K., Nakamura, T., Sugita, M., Shimizu, T., and Shikanai, T. (2006). A pentatricopeptide repeat protein is a site recognition factor in chloroplast RNA editing. J Biol Chem **281,** 37661-37667.

Ori, N., Cohen, A.R., Etzioni, A., Brand, A., Yanai, O., Shleizer, S., Menda, N., Amsellem, Z., Efroni, I., Pekker, I., Alvarez, J.P., Blum, E., Zamir, D., and Eshed, Y. (2007). Regulation of LANCEOLATE by miR319 is required for compound-leaf development in tomato. Nat Genet **39,** 787-791.

Pogson, B.J., Woo, N.S., Forster, B., and Small, I.D. (2008). Plastid signalling to the nucleus and beyond. Trends Plant Sci **13,** 602-609.

Prikryl, J., Watkins, K.P., Friso, G., van Wijk, K.J., and Barkan, A. (2008). A member of the Whirly family is a multifunctional RNA- and DNA-binding protein that is essential for chloroplast biogenesis. Nucleic Acids Res **36,** 5152-5165.

Pruneda-Paz, J.L., Breton, G., Para, A., and Kay, S.A. (2009). A functional genomics approach reveals CHE as a component of the Arabidopsis circadian clock. Science **323,** 1481-1485.

Rodriguez, P., Munroe, D., Prawitt, D., Chu, L.L., Bric, E., Kim, J., Reid, L.H., Davies, C., Nakagama, H., Loebbert, R., Winterpacht, A., Petruzzi, M.J., Higgins, M.J., Nowak, N., Evans, G., Shows, T., Weissman, B.E., Zabel, B., Housman, D.E., and Pelletier, J. (1997). Functional characterization of human nucleosome assembly protein-2 (NAP1L4) suggests a role as a histone chaperone. Genomics **44,** 253-265.

Salinas, T., Duchene, A.M., and Marechal-Drouard, L. (2008). Recent advances in tRNA mitochondrial import. Trends Biochem Sci **33,** 320-329.

Schmitz-Linneweber, C., and Small, I. (2008). Pentatricopeptide repeat proteins: a socket set for organelle gene expression. Trends Plant Sci **13,** 663-670.

Schmitz-Linneweber, C., Williams-Carrier, R., and Barkan, A. (2005). RNA Immunoprecipitation and Microarray Analysis Show a Chloroplast Pentatricopeptide Repeat Protein to Be Associated with the 5' Region of mRNAs Whose Translation It Activates. Plant Cell **17,** 2791-2804.

Schmitz-Linneweber, C., Williams-Carrier, R.E., Williams-Voelker, P.M., Kroeger, T.S., Vichas, A., and Barkan, A. (2006). A pentatricopeptide repeat protein facilitates the trans-splicing of the maize chloroplast rps12 pre-mRNA. Plant Cell **18,** 2650-2663.

Schommer, C., Palatnik, J.F., Aggarwal, P., Chetelat, A., Cubas, P., Farmer, E.E., Nath, U., and Weigel, D. (2008). Control of jasmonate biosynthesis and senescence by miR319 targets. PLoS Biol **6,** e230.

Shen, H.H., Huang, A.M., Hoheisel, J., and Tsai, S.F. (2001). Identification and characterization of a SET/NAP protein encoded by a brain-specific gene, MB20. Genomics **71,** 21-33.

Simon, H.U., Mills, G.B., Kozlowski, M., Hogg, D., Branch, D., Ishimi, Y., and Siminovitch, K.A. (1994). Molecular characterization of hNRP, a cDNA encoding a

human nucleosome-assembly-protein-I-related gene product involved in the induction of cell proliferation. Biochem J **297 (Pt 2),** 389-397.

Small, I., Wintz, H., Akashi, K., and Mireau, H. (1998). Two birds with one stone: genes that encode products targeted to two or more compartments. Plant Mol. Biol. **38,** 265-277.

Small, I., Peeters, N., Legeai, F., and Lurin, C. (2004). Predotar: A tool for rapidly screening proteomes for N-terminal targeting sequences. Proteomics **4,** 1581-1590.

Small, I.D., and Peeters, N. (2000). The PPR motif - a TPR-related motif prevalent in plant organellar proteins. Trends Biochem Sci **25,** 46-47.

Spielewoy, N., Schulz, H., Grienenberger, J.M., Thony-Meyer, L., and Bonnard, G. (2001). CCME, a nuclear-encoded heme-binding protein involved in cytochrome c maturation in plant mitochondria. J Biol Chem **276,** 5491-5497.

Tasaki, E., Hattori, M., and Sugita, M. The moss pentatricopeptide repeat protein with a DYW domain is responsible for RNA editing of mitochondrial ccmFc transcript. Plant J.

Tatematsu, K., Nakabayashi, K., Kamiya, Y., and Nambara, E. (2008). Transcription factor AtTCP14 regulates embryonic growth potential during seed germination in Arabidopsis thaliana. Plant J **53,** 42-52.

Tavares-Carreon, F., Camacho-Villasana, Y., Zamudio-Ochoa, A., Shingu-Vazquez, M., Torres-Larios, A., and Perez-Martinez, X. (2008). The pentatricopeptide repeats present in Pet309 are necessary for translation but not for stability of the mitochondrial COX1 mRNA in yeast. J Biol Chem **283,** 1472-1479.

Unseld, M., Marienfeld, J.R., Brandt, P., and Brennicke, A. (1997). The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. Nature Genet. **15,** 57-61.

Uyttewaal, M., Mireau, H., Rurek, M., Hammani, K., Arnal, N., Quadrado, M., and Giege, P. (2008). PPR336 is associated with polysomes in plant mitochondria. J Mol Biol **375,** 626-636.

Welchen, E., and Gonzalez, D.H. (2005). Differential expression of the Arabidopsis cytochrome c genes Cytc-1 and Cytc-2. Evidence for the involvement of TCP-domain protein-binding elements in anther- and meristem-specific expression of the Cytc-1 gene. Plant Physiol **139,** 88-100.

Welchen, E., and Gonzalez, D.H. (2006). Overrepresentation of elements recognized by TCP-domain transcription factors in the upstream regions of nuclear genes encoding components of the mitochondrial oxidative phosphorylation Machinery. Plant Physiol **141,** 540-545.

Woodson, J.D., and Chory, J. (2008). Coordination of gene expression between organellar and nuclear genomes. Nat Rev Genet **9,** 383-395.

Yoon, H.W., Kim, M.C., Lee, S.Y., Hwang, I., Bahk, J.D., Hong, J.C., Ishimi, Y., and Cho, M.J. (1995). Molecular cloning and functional characterization of a cDNA encoding nucleosome assembly protein 1 (NAP-1) from soybean. Mol Gen Genet **249,** 465-473.

Zabaleta, E., Heiser, V., Grohmann, L., and Brennicke, A. (1998). Promoters of nuclear-encoded respiratory chain complex I genes from Arabidopsis thaliana contain a region essential for anther/pollen-specific expression. Plant J **15,** 49-59.

Figure Legends

Figure 1. Molecular characterization of the *PPR9* gene.

(A) Diagram of the insertion positions of T-DNAs in *PPR9*. The different domains encoded by the gene are represented. MTS: Mitochondrial Targeting Signal, PPR: PentatricoPeptide Repeats, NLS: Nuclear Localization Signal. Rectangles in the bottom panel represent the mutant *PPR9* genes cloned in binary vectors for functional complementation assays. *PPR9-MTS*: Mitochondrial Targeting Signal truncated *PPR9*, *PPR9-NLS*: C-terminal domain truncated. **(B)** *PPR9* is essential for embryo development. Phenotype of siliques from self-fertilized *PPR9/ppr9 Arabidopsis* plants for two independant T-DNA insertion alleles. Stars indicate aborted seeds with *ppr9/ppr9* embryos. **(C)** Histochemical localization of GUS activity in *Arabidopsis* plants transformed with the intergenic region upstream *PPR9* gene fused to the *GUS* reporter gene. Left: *GUS* staining of 8 old days seedlings showing *PPR9* expression in the shoot apex (SA), root meristem (RM), lateral root primordia (LRP), leaf primordium (LP). Right: *GUS* staining of flower showing *PPR9* expression in pollen grains (P) within the anthers (A).

Figure 2. Alignment of the PPR9 protein sequence with its homologs from *Vitis vinifera* and *Oryza sativa*.

Identical amino acids are shown with white letters in black boxes, and similar amino acids are shown with shaded boxes. The seven PPR motifs and the putative bipartite localization signal are underlined. The PPR9 mitochondrial targeting sequence (MTS) and the C-terminal domain, which have been deleted for functional complementation assays are delimited with arrows.

Figure 3. PPR9 is dual localized to the nuclei and mitochondria of plant cells.

(A) BY-2 cells transiently expressing GFP fusion proteins were observed by confocal microscopy 6 h after bombardment. M: mitochondria, Nu: nucleus, Cy: cytosol **(B)** Protein fractions from *Arabidopsis thaliana* were analyzed by Western Blots probed with antibodies directed against PPR9. (Nu) Protein extracts of nuclei, (M) mitochondria and (Chl) chloroplasts. The purity of the respective fractions was controlled with antibodies directed

against the chloroplast large subunit of RuBisCo, the subunit 9 of the mitochondrial complex I (Nad9) and the nuclear histone protein H2B. **(C)** Immunocytology experiments with PPR9 antibodies used on *Arabidopsis* seedling sections show a colocalization of PPR9 (Red staining) with nuclei (DAPI staining).

Figure 4. PPR9 is an RNA binding protein associated with polysomes in plant mitochondria.

(A) Sensorgram showing the interaction between PPR9 and total mitochondrial RNA obtained with a Biacore biosensor (GE Healthcare). Molecular interaction is expressed in "response units" (ru) and monitored through time (s). The black asterisk indicates when RNA injection started and the grey asterisk when it was stopped. Blue: PPR9 mature, Green: PPR9 full length, Pink: Thioredoxin **(B)** Immunodetection of PPR9 on sub-mitochondrial protein fractions. Mtp shows the mitoplast fraction; Sol, the soluble fraction; Mb, the membrane fraction; Mbi, the intrinsic membrane protein fraction; and Mbe, the extrinsic membrane protein fraction. Fractionations are controlled with antibodies specific for complex I subunit 9 (NAD9), the soluble MnSOD and the intrinsic membrane protein cytochrome c1 (Cyt c1). Molecular masses are given in kilodaltons. **(C)** Mitochondrial protein fractions were collected from sucrose gradients, separated by SDS PAGE, blotted and reacted with antibodies specific for PPR9, NAD9 and the ribosomal protein RPS1. Control (C) sample was untreated, whereas other samples were treated to destabilise polysomes with puromycin (P) or RNases (R).

Figure 5. PPR9 interacts with nuclear transcriptional factors

(A) Yeast two hybrid-like assays show that PPR9 interacts with NAP1 and TCP8. Yeast cells were cotransformed with X-Cub and Nub-X constructs. X-Cub constructs are fusions of a protein of interest with the membrane anchor protein Ost4p, the C-terminal half of ubiquitin and a transcription factor. Nub-X constructs are C-terminal fusions of a protein of interest to a mutated version of the N-terminal domain of ubiquitin. Liquid cultures of double transformants are plated at DO_{600} = 0.1 and at increasing ×10 and ×100 dilutions of the cultures (black triangles) on synthetic drop out selective medium that lacked Leucine, Histine, Adenine and Tryptophan supplemented with 10 mM 3-AT. **(B)** BiFC visualization of PPR9 interaction with NAP1 and TCP8 in nuclei (Nu) of onion epidermal cells. Left, reconstituted YFP fluorescence image of cells cotransfected with the split-YFP constructs and the control

CPRF2-CFP mainly localized in the nucleus, CFP signal in the center panel. The bright-field images of the onion cells are shown on the right.

Supplemental data

Supplemental Figure 1. PPR9 binds preferentially polyG oligoribonucleotides.

Sensorgram showing the interaction between PPR9 and polyG oligoribonucleotides obtained with a Biacore biosensor (GE Healthcare). Molecular interaction is expressed in "response units" (ru) and monitored through time (s). The black asterisk indicates when RNA injection started and the grey asterisk when it was stopped. Blue: PPR9 mature, Green: PPR9 full length, Pink: Thioredoxin

Supplemental Table 1. Oligonucleotides used in this study.

For GUS and complementation constructs cloning

z0753	ggggacaagtttgtacaaaaaagcaggcttccggagactataatgcata	oligo forward PPR9 promotor amplification
z0667	ggggaccactttgtacaagaaagctgggtcttctgatttagtgagaga	oligo reverse PPR9 promotor amplification
z0718	GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC TCA CAT TCT CCG TCT CAT CTT	oligo reverse cloning PPR9 Full length with termination codon
z0761	ttctgatttagtgagagaaaattggtaact	oligo reverse promotor PPR9 with 15 nt matching z0760
z0760	ctcactaaatcagaaatgtctgaaacgaatgcg	oligo forward PPR9 -MTS
z0759	GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC TCA ttccttagggctaacagc	oligo reverse amplification PPR9 without C-ter domain

For GFP contructs cloning

z0764	GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC ttccttagggctaacagco	oligo reverse amplification PPR9 without C-ter domain, fusion C-ter
z0765	cag cag gga ggc aaa caa aag gaa tac aga gta gat cca	oligo forward chimere end gus et Ct domain PPR9
z0772	tgg atc tac tct gta ttc ctt ttg ttt gcc tcc ctg ctg	oligo reverse amplification gus without stop codon matching 15nt Cter PPR9
z0774	gggg aca agt ttg tac aaaaaa gca ggc ttc atgtctgaaacgaatgcggaatcg	oligo forward cloning PPR9-MTS predicted targetP

z0654 GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC CAT TCT CCG TCT CAT CTT ATT
oligo reverse cloning PPR9 full lenght fusion Cter

z0721 GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC CCG CCG TCG TTA CCG TCT CTG
oligo forward cloning PPR9 fusion Nter

z0652 GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC ATG CCG CCG TCG TTA CCG TCT
oligo forward cloning PPR9 full length fusion Cter

z0729 gggg aca agt ttg tac aaaaaa gca ggc ttc atgttacgtcctgtagaaacc
oligo forward cloning gateway gusA fusion Cter

z0730 ggg gac cac ttt gta caa gaa agc tgg gtc ttgtttgcctccctgctgcgg
oligo reverse cloning gateway gusA fusion Cter

For genotyping T-DNA lines

p1165	AGCATCTGAATTTCATAACCAATCTCGATACAC	oligo LB SAIL T-DNA
z0621_RP	TCAGCATCTGGTTGACAACC	oligo RP sail_868_D10
z0622_LP	GTAAAACACTAGAATCGGCG	oligo LP sail_868_D10
z0713_LP	ggtgtacctagaaacacagagactttcaatgtg	oligo LP sail_224_H01
z0714_RP	tcagcatcacattctccgtctcatcttattgtgc	oligo RP sail_224_H01

For genotyping complemented plants

p1165	AGCATCTGAATTTCATAACCAATCTCGATACAC	
z0621_RP	TCAGCATCTGGTTGACAACC	
z0622_LP	GTAAAACACTAGAATCGGCG	
z0767	cgg att gag tag ttt cta taa	
	oligo forward upstream of PPR9 sequence cloned in pGWB1 for complementation	
z0739	gcgattgataggcttgttagag	oligo forward +541 PPR9 gene
z0770	cac cac ttt gta caa gaa agc	pGWB1 reverse oligo
k0896	GCAAAATGATGAGTACTCTTGAC	
	oligo reverse downstream of PPR9 sequence cloned in pGWB1 for complementation	

For Dual hunter vectors cloning

<u>Bait construction (pDHB1 vector)</u>

306DHPPR9F ATTAACAA GGCCATTACGGCC GAATCTTTGGATTCGAATG
306DHPPR9R AACTGATT GGCCGAGGCGGCC CC CATTCTCCGTCTCATCTTATT

<u>pPR3N vector</u>

306PR3N 9R	AACTGATT GGCCGAGGCGGCC C TCATCTCCGTCTCATCTTATT
306DH NAP1F	ATTAACAA GGCCATTACGGCC AGCAACGACAAGGATAGCTTC
306DHB1 NAP1R	AACTGATT GGCCGAGGCGGCC CC CTGTTGCTTGCATTCGGGTGG
306PR3N NAP1R	AACTGATT GGCCGAGGCGGCC C TTACTGTTGCTTGCATTCGGG
306DH TCPF	ATTAACAA GGCCATTACGGCC GATCTCTCCGACATCCGAAAC
306DHB1 TCPR	AACTGATT GGCCGAGGCGGCC CC CTCAGAGCTATTTGAGTTCTC

306PR3N TCPR AACTGATT GGCCGAGGCGGCC C TCACTCAGAGCTATTTGAGTTC

For BiFC vector cloning

306Gate PPR9F GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGAATCTTTGGATTCGAATG
306Gate PPR9R GGGGACCACTTTGTACAAGAAAGCTGGGTCTCATCTCCGTCTCATCTTATT
306Gate NAP1F GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGAGCAACGACAAGGATAGCTTC
306Gate NAP1R GGGGACCACTTTGTACAAGAAAGCTGGGTCTTACTGTTGCTTGCATTCGGG
306Gate TCPF GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGATCTCTCCGACATCCGAAAC
306Gate TCPR GGGGACCACTTTGTACAAGAAAGCTGGGTCTCACTCAGAGCTATTTGAGTTC

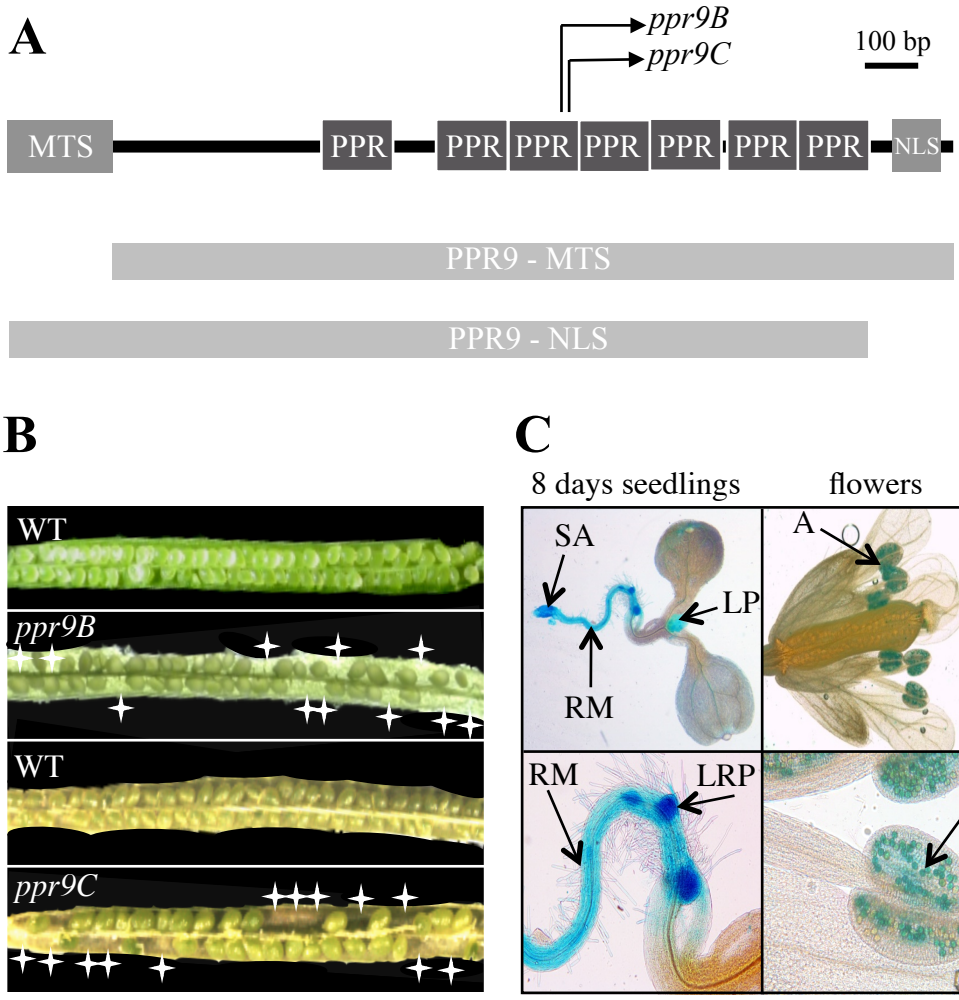

A
ppr9B
ppr9C
100 bp
MTS
PPR
PPR
PPR
PPR
PPR
PPR
PPR
NLS
PPR9 - MTS
PPR9 - NLS
B
WT
ppr9B
WT
ppr9C
C
8 days seedlings
flowers
SA
LP
RM
A
RM
LRP
P

Figure 1

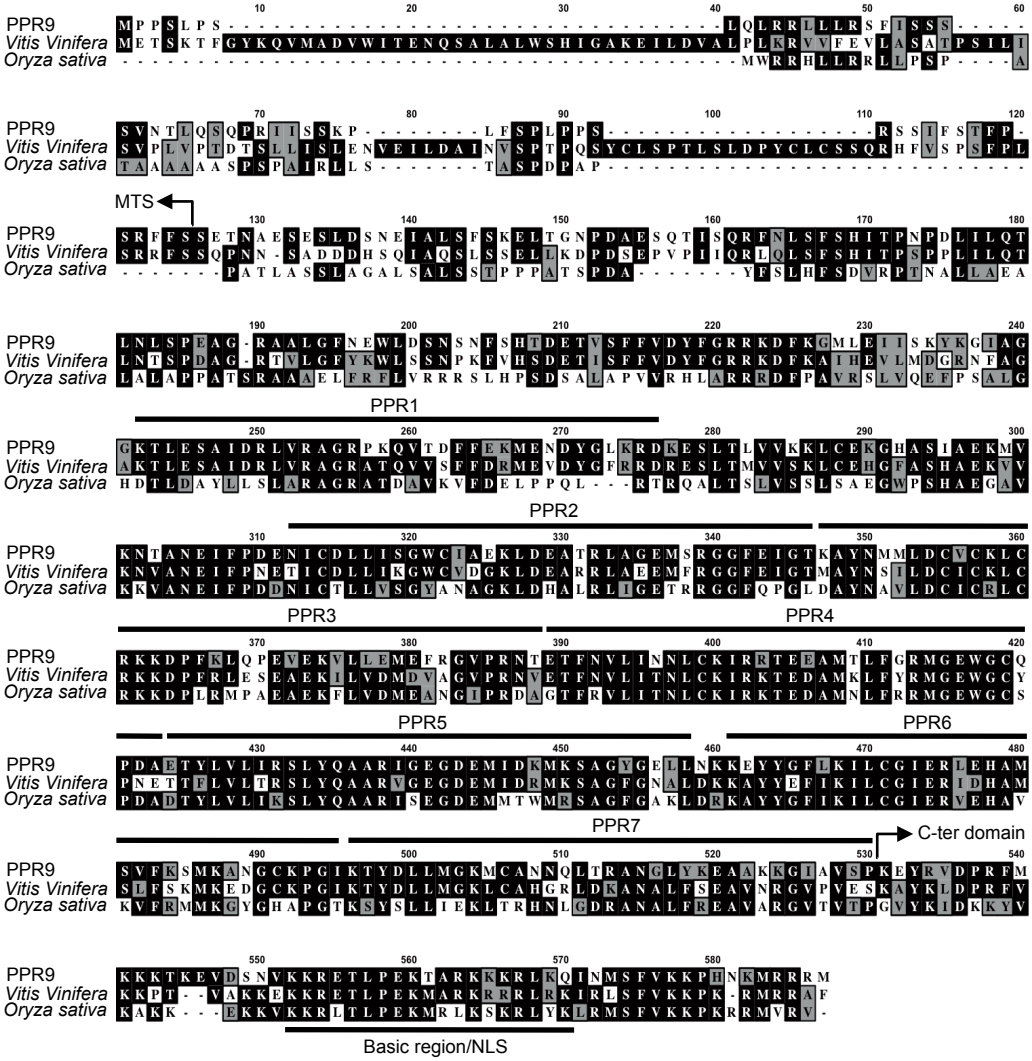

PPR9
Vitis Vinifera
Oryza sativa
MTS
PPR1
PPR2
PPR3
PPR4
PPR5
PPR6
PPR7
C-ter domain
Basic region/NLS

Figure 2

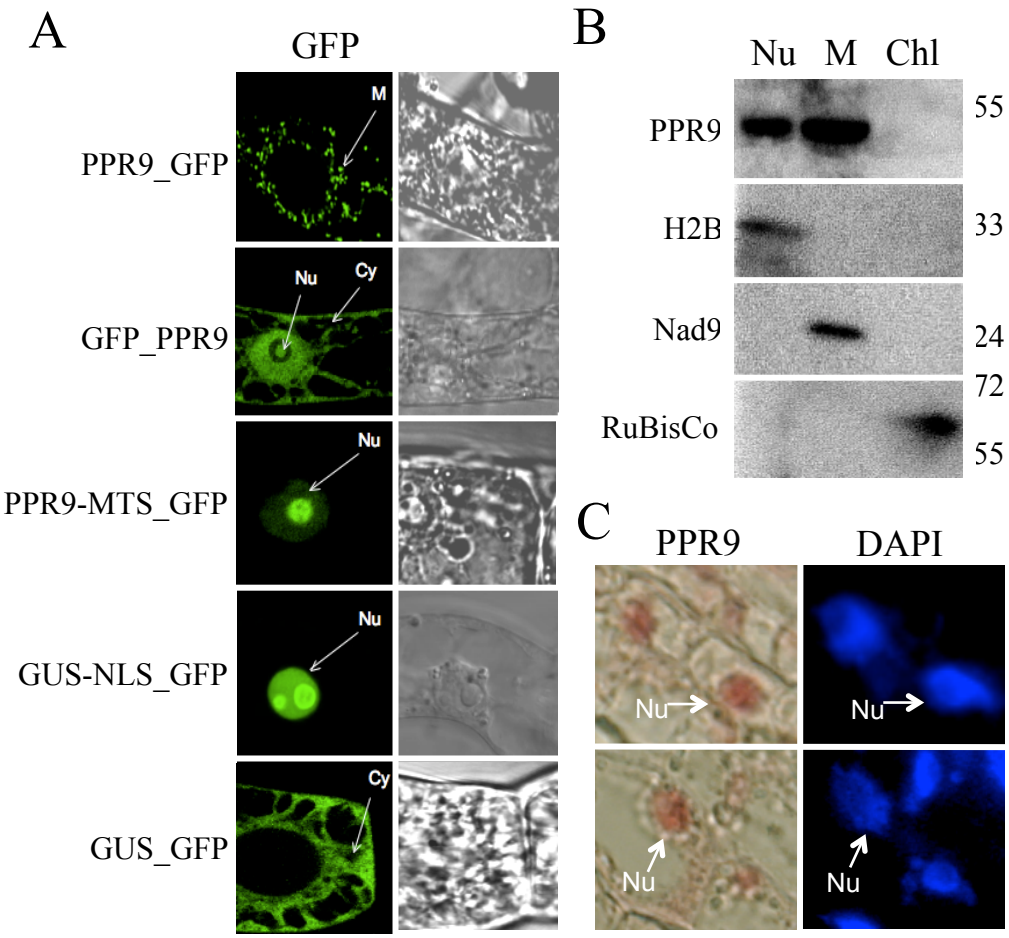

A
GFP
PPR9_GFP
GFP_PPR9
PPR9-MTS_GFP
GUS-NLS_GFP
GUS_GFP
M
Nu
Cy
B
Nu M Chl
PPR9
H2B
Nad9
RuBisCo
55
33
24
72
55
C
PPR9
DAPI

Figure 3

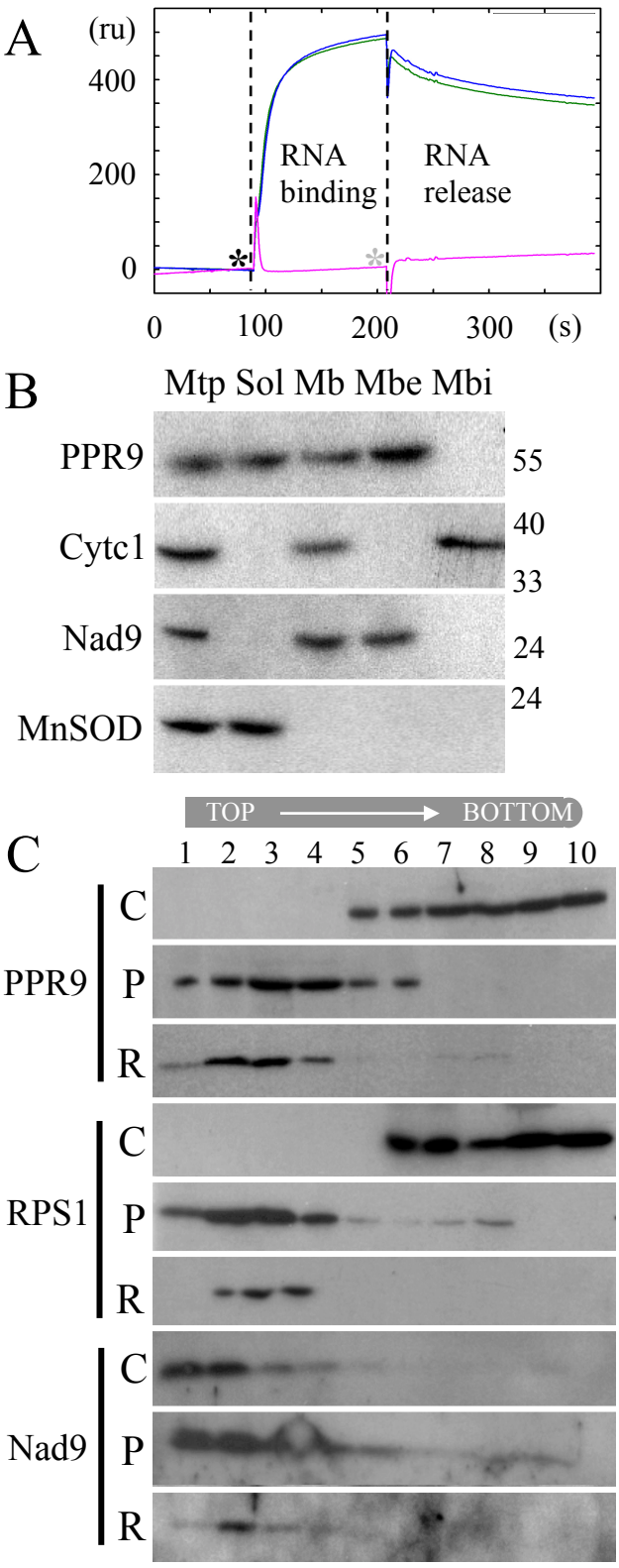

A
(ru)
400
200
0
RNA
binding
RNA
release
0
100
200
300
(s)
B
Mtp Sol Mb Mbe Mbi
PPR9
Cytc1
Nad9
MnSOD
55
40
33
24
24
C
TOP
BOTTOM
1 2 3 4 5 6 7 8 9 10
PPR9
RPS1
Nad9
C
P
R

Figure 4

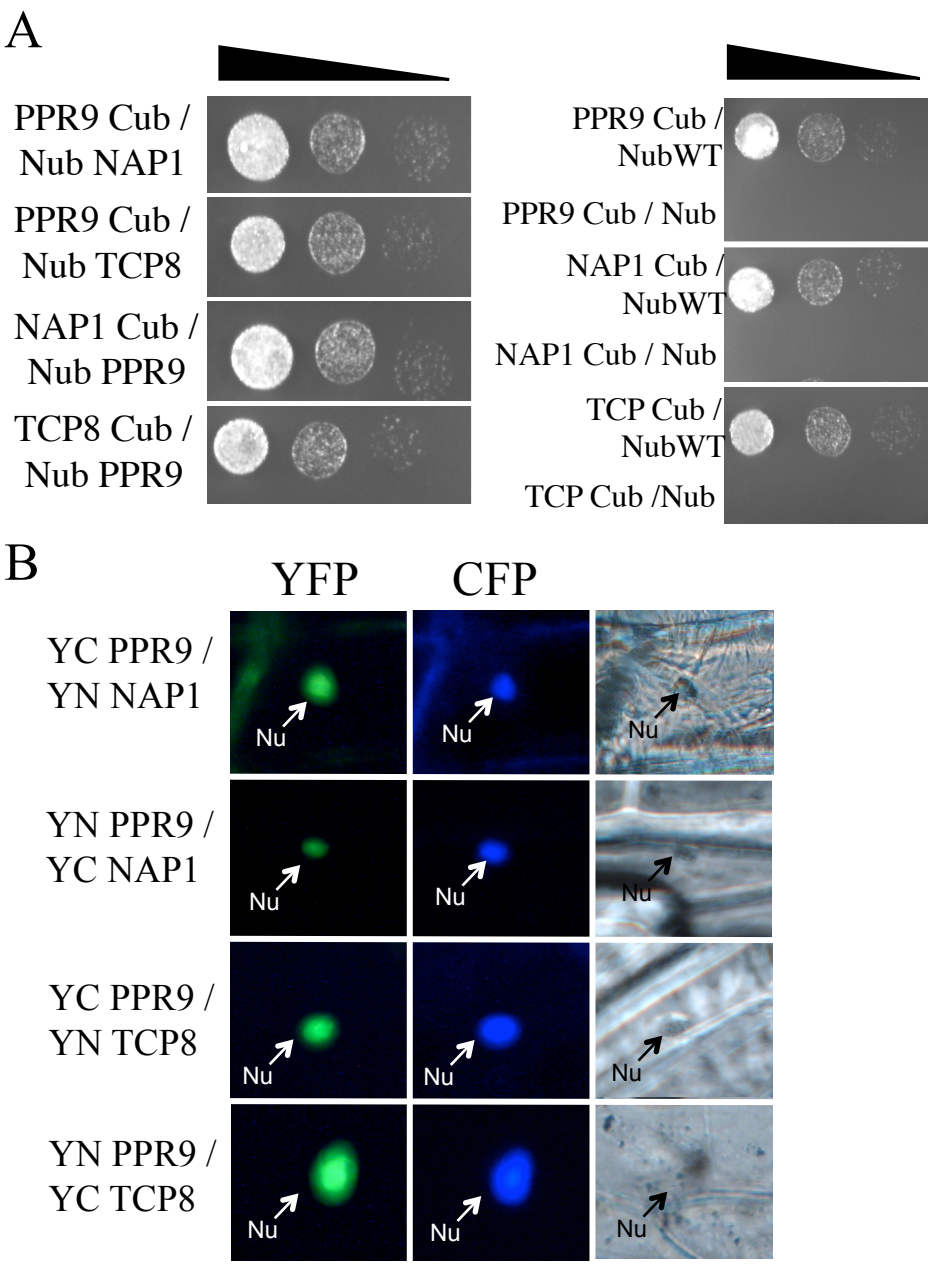
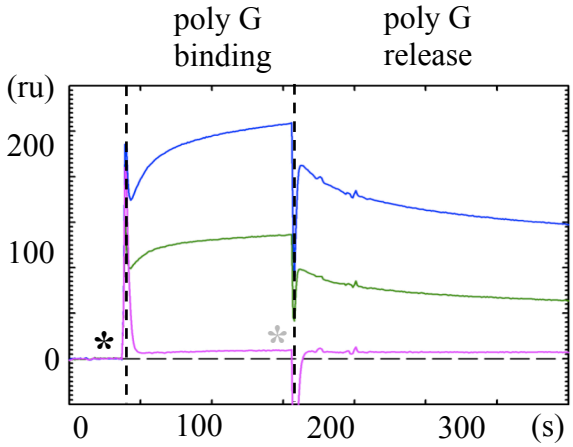

A
PPR9 Cub / Nub NAP1
PPR9 Cub / Nub TCP8
NAP1 Cub / Nub PPR9
TCP8 Cub / Nub PPR9
PPR9 Cub / NubWT
PPR9 Cub / Nub
NAP1 Cub / NubWT
NAP1 Cub / Nub
TCP Cub / NubWT
TCP Cub /Nub
B
YFP
CFP
YC PPR9 / YN NAP1
YN PPR9 / YC NAP1
YC PPR9 / YN TCP8
YN PPR9 / YC TCP8
Nu

Figure 5



Supplemental Figure 1

4.3. CONTRIBUTIONS MINEURES A PUBLICATION

4.3.1. PUBLICATION 4: THE ARABIDOPSIS GENE YS1 ENCODING A DYW PROTEIN IS REQUIRED FOR EDITING OF rpoB TRANSCRIPTS AND THE RAPID DEVELOPMENT OF CHLOROPLASTS DURING EARLY GROWTH. THE PLANT JOURNAL (2009)

the plant journal

SEB SOCIETY FOR EXPERIMENTAL BIOLOGY

The Plant Journal (2009) **58**, 82–96 doi: 10.1111/j.1365-313X.2008.03766.x

The Arabidopsis gene *YS1* encoding a DYW protein is required for editing of *rpoB* transcripts and the rapid development of chloroplasts during early growth

Wenbin Zhou[1], Yuxiang Cheng[1], Aaron Yap[2], Anne-Laure Chateigner-Boutin[2], Etienne Delannoy[2], Kamel Hammani[2], Ian Small[2] and Jirong Huang[1,*]

[1]National Key Laboratory of Plant Molecular Genetics, Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 200032 Shanghai, China, and

[2]ARC Centre of Excellence in Plant Energy Biology, Molecular and Chemical Sciences Building (M316), University of Western Australia, 35 Stirling Highway, Crawley, Perth 6009 WA, Australia

Received 18 October 2008; revised 17 November 2008; accepted 18 November 2008; published online 29 December 2008.
*For correspondence (fax +86 21 5492 4015; e-mail huangjr@sibs.ac.cn).

Summary

Virescence, a phenotype in which leaves green more slowly than usual, is recognized to play a role in protection from photo-oxidative damage before healthy chloroplasts are developed. The elucidation of the molecular mechanisms underlying virescence will provide insights into how the development of chloroplasts is controlled. In this study, we find that knockout alleles of *Yellow Seedlings 1* (*YS1*) in Arabidopsis lead to a virescent phenotype, which disappears by 3 weeks after germination. The *ys1* mutation resulted in marked decreases in photosynthetic capacity and photosynthetic pigment complexes, and disturbed ultrastructure of thylakoid membranes in 8-day-old seedlings. However, cotyledons of *ys1* seedlings pre-treated in the dark for 5 days turn green almost as fast as the wild type in light, revealing that the developmental defects in *ys1* are limited to the first few days after germination. Inspection of all known plastid RNA editing and splicing events revealed that YS1 is absolutely required for editing of site 25992 in *rpoB* transcripts encoding the beta subunit of the plastid-encoded RNA polymerase (PEP). YS1 is a nuclear-encoded chloroplast-localized pentatricopeptide repeat protein differing from previously described editing factors in that it has a C-terminal DYW motif. A defect in PEP activity is consistent with the changes in plastid transcript patterns observed in *ys1* seedlings. We conclude that the activity of PEP containing RpoB translated from unedited transcripts is insufficient to support rapid chloroplast differentiation.

Keywords: chloroplast development, pentatricopeptide repeat protein, plastid gene expression, RNA polymerase, Arabidopsis.

Introduction

The virescent phenotype is defined as presenting young leaves with reduced chlorophyll levels but can accumulate almost normal amounts of chlorophyll as they mature (Archer and Bonnett, 1987). Due to the unique features of their chloroplast biogenesis and potential value in agriculture, virescent mutants have been studied for more than half a century in various plant species, such as maize (*Zea mays*; da Costa e Silva *et al.*, 2004), rice (*Oryza sativa*; Kusumi *et al.*, 1997; Sugimoto *et al.*, 2007), cotton (*Gossypium*; Rhyne, 1955), tobacco (*Nicotiana tabacum*; Archer and Bonnett, 1987), peanut (*Arachis hypogaea*; Benedict and Ketring, 1972) and Arabidopsis (Loschelder *et al.*, 2006; Koussevitzky *et al.*, 2007). To date, however, only a few of the genes responsible for virescent phenotypes have been isolated and their molecular action remains largely unknown.

Plastids of higher plants arose from cyanobacteria through endosymbiosis and have their own genome (Dyall *et al.*, 2004) containing around 40 photosynthesis-related genes and 60 housekeeping genes that are essential for protein synthesis (Sato *et al.*, 1999). It is well documented that coordination of nuclear and plastid gene expression is crucial for chloroplast biogenesis (Mullet, 1988). Transcrip-

tion of the plastome relies on two types of plastid RNA polymerases: nucleus-encoded RNA polymerase (NEP) that resembles those of the T7 bacteriophage (Hedtke *et al.*, 1997), and plastid-encoded RNA polymerase (PEP), which is a multi-subunit eubacterial-type enzyme (Hess and Börner, 1999). Nucleus-encoded polymerase predominantly transcribes housekeeping genes (e.g. *rpoB, accD*) (Hajdukiewicz *et al.*, 1997), whereas PEP is responsible for transcription of photosynthesis-related genes (e.g. *rbcL*, *psbA*) (Liere and Maliga, 1999). Promoter specificity and transcription initiation by PEP requires nucleus-encoded proteins, such as σ-like factors whose expression is spatially and temporally regulated by environmental cues (Allison, 2000; Kasai *et al.*, 2004; Ishizaki *et al.*, 2005). At the beginning of light-induced chloroplast development, NEP is imported into plastids and initiates transcription of housekeeping genes including PEP-encoding genes, followed by an increase in PEP-dependent transcription (Hajdukiewicz *et al.*, 1997). Control of chloroplast gene expression has been demonstrated at various levels, such as transcription, RNA processing and stability, and translation (Barkan and Goldschmidt-Clermont, 2000).

Several lines of evidence have demonstrated that the virescent phenotype is closely associated with reduced activity of PEP or NEP at an early stage of chloroplast development. In Arabidopsis, mutations in *SIG6* cause a weakly virescent phenotype, with only the first pair of leaves lacking normal pigment levels (Loschelder *et al.*, 2006). Transcripts of several PEP-dependent plastid genes are markedly and specifically reduced in the *sig6* mutant (Loschelder *et al.*, 2006). Partial silencing of one of the NEP genes, *RPOTP*, leads to a typical virescent phenotype of Arabidopsis seedlings, which can become as green as the wild type after 2 weeks of growth (Swiatecka-Hagenbruch *et al.*, 2008). A knockout of *RPOTP*, however, exhibits severe deficiency in chloroplast biogenesis as well as in cell proliferation (Hricová *et al.*, 2006). Very severe delays to normal greening are also seen in *clb19* mutants, in which RNA editing of *rpoA* transcripts is lacking leading to a loss in activity of PEP (Chateigner-Boutin *et al.*, 2008). *CLB19* is one of many nuclear pentatricopeptide repeat (PPR) genes required for proper organellar gene expression. At least 450 PPR genes exist in the Arabidopsis genome (Aubourg *et al.*, 2000; Lurin *et al.*, 2004; Rivals *et al.*, 2006). They are characterized by tandem 35-amino-acid motifs, which share some features with the better-known tetratricopeptide repeat (TPR) motif (Small and Peeters, 2000). Pentatricopeptide repeat proteins are thought to bind DNA or RNA molecules and/or to mediate protein–protein interactions (Delannoy *et al.*, 2007). In general, PPR proteins have been implicated in RNA transcription (Ikeda and Gray, 1999; Pfalz *et al.*, 2006), RNA editing (Kotera *et al.*, 2005; Okuda *et al.*, 2007; Chateigner-Boutin *et al.*, 2008), RNA processing (Meierhoff *et al.*, 2003; Hattori *et al.*, 2007), RNA splicing (Schmitz-Linneweber *et al.*, 2006; Falcon de Longevialle *et al.*, 2007, 2008), RNA stability (Beick *et al.*, 2008) and translational activation (Schmitz-Linneweber *et al.*, 2005; Tavares-Carreón *et al.*, 2008). However, most PPR genes remain to be investigated with regard to their physiological functions, biochemical properties and mechanism of action. Here, we show that a chloroplast-localized PPR protein, YS1, is required for the editing of *rpoB* transcripts in Arabidopsis. YS1 is particularly interesting because it is a member of the DYW subgroup of the PPR family, suggested on phylogenetic grounds to be likely editing factors (Salone *et al.*, 2007; O'Toole *et al.*, 2008) but for which experimental proof was lacking. Lack of editing of *rpoB* transcripts leads to a significant decrease in PEP activity and, in turn, a delay in the rapid development of chloroplasts during early growth.

Results

Identification and characterization of the ys1 *mutant*

In order to gain more insight into the role of PPR proteins in early chloroplast development, we screened for virescent mutants with slow greening of cotyledons and/or leaves from a batch of Arabidopsis PPR mutants ordered from the SALK T-DNA collection (Alonso *et al.*, 2003). In this study, we investigated a virescent mutant (Salk_123515), designated *yellow seedling 1* (*ys1-1*), because the chlorophyll content of the mutant was ultimately able to completely restore to the wild-type level. The virescent phenotype of *ys1-1* was imaged at day 4 (Figure 1a), day 8 (Figure 1b), day 12 (Figure 1c) and day 25 (Figure 1d) after germination. This observation was consistent with the time-course changes in chlorophyll and carotenoids in mutant and wild-type plants (Figure 1e). In contrast, the leaf area of the mutant reached only about 55% of the wild type even at day 30 after germination (Figure 1f). Leaves emerging from more than 4-week-old mutant plants could reach the same chlorophyll level as those of the wild type within 2 days, and mature plants were almost indistinguishable between the mutant and the wild type in terms of rosette diameter, bolting time and seed fertility (data not shown).

YS1 *encodes a chloroplast-localized PPR protein*

Our genetic analysis showed that the virescent phenotype of *ys1-1* was controlled by a recessive nuclear gene and co-segregated with the kanamycin resistance marker (data not shown). Plasmid rescue and sequencing of the PCR products showed that the T-DNA was inserted in the first exon, 378 bp downstream of the start codon, of the Arabidopsis gene *At3g22690* (henceforth named *YS1*). Analyses of the deduced YS1 amino acid sequence revealed a predicted chloroplast transit peptide at its N terminus, followed by 16 PPR motifs and E and DYW motifs (Figure 1g). The predicted gene model for *YS1* in the TAIR

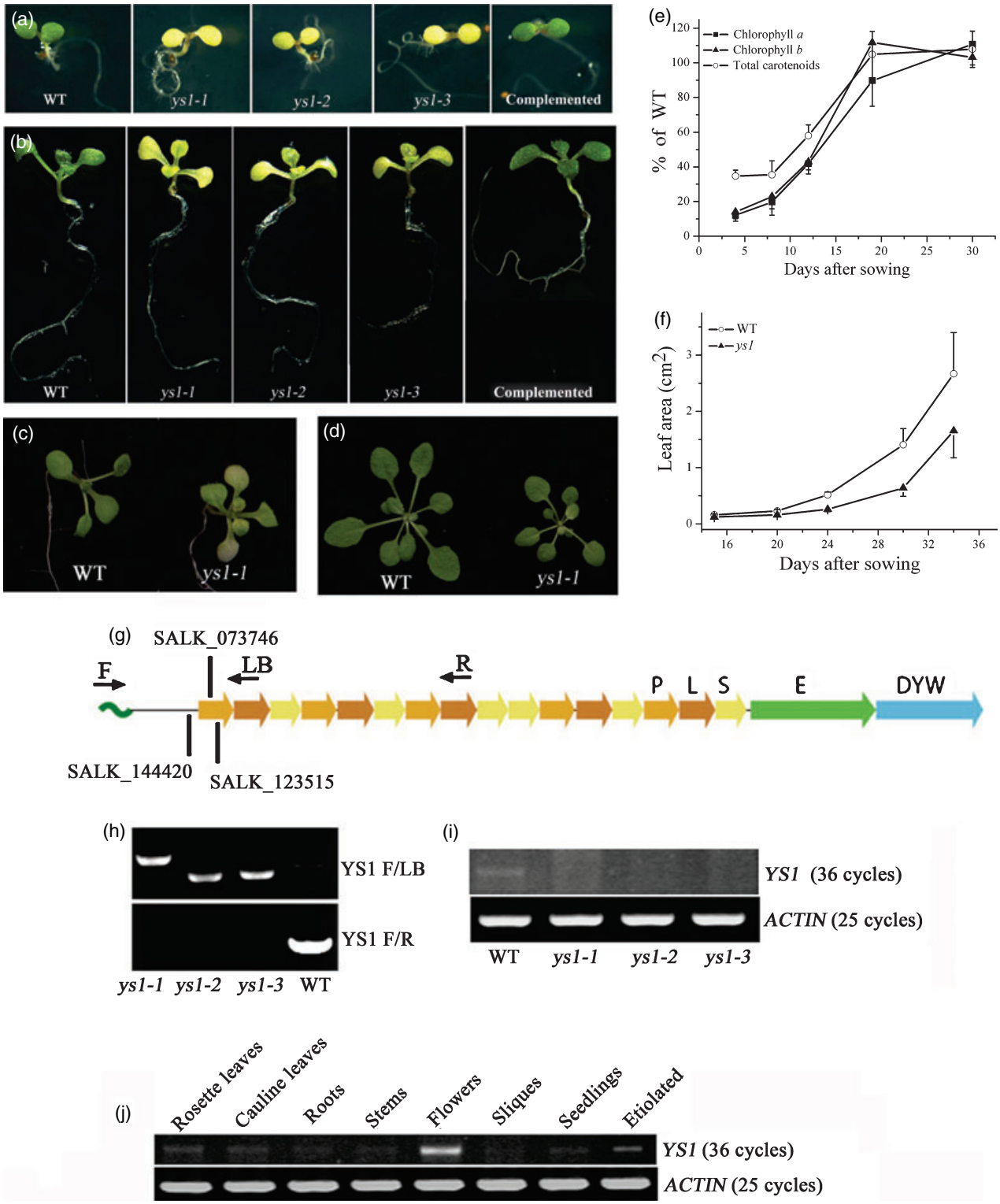


Figure 1. Identification and characterization of the *ys1* mutant.
(a) Four-day-old plants; (b) 7-day-old plants; (c) 12-day-old plants; (d) 25-day-old plants. (e) Time course of pigment accumulation. (f) Change in leaf area. (g) Localization of the T-DNA insertions in the three mutant alleles and arrangement of the various motifs within the *YS1* sequence. Arrows represent the position and orientation of the gene-specific primers used in (h) and (i). (h) Identification of the T-DNA insertions in the *ys1* mutants by PCR with genomic DNA as template. (i) Messenger RNA of *YS1* detected by RT-PCR. The actin transcript from the same cDNA preparations was used as an internal control. (j) Detection of *YS1* transcripts in various tissues. Total RNA was extracted from the indicated tissues of 8- or 30-day-old light-grown, or 5-day-old etiolated plants.

database (http://www.arabidopsis.org/) includes an intron that would disrupt the DYW motif and adds additional exons downstream, extending the reading frame. To verify this model, we cloned *YS1* cDNAs by 3′-rapid amplification of cDNA ends (RACE) and sequenced four clones of differing lengths (Figure S1 in Supporting Information). None of the cDNA clones supported the TAIR model, and all indicated that the YS1 open reading frame ends with the terminal Asp–Phe–Trp codons similar to the Asp–Tyr–Trp characteristic of DYW genes. The Arabidopsis gene *At2g29760* encodes a protein that is 60% similar to YS1, and *YS1* homologues are conserved in rice and grape; all of these homologues have a predicted gene structure similar to the one we propose for *YS1* (Figure S2).

To perform genetic complementation, we introduced, under the control of 35S-cauliflower mosaic virus (CaMV) promoter, a 3778-bp genomic sequence spanning the entire open reading frame of *YS1* into the *ys1-1* mutant by *Agrobacterium tumefaciens*-mediated transformation. More than 10 transgenic lines expressing *YS1* were identified from the transformants and displayed the same phenotype as the wild type (Figure 1a,b), confirming that the virescent phenotype indeed resulted from disruption of *At3g22690*. We identified two other mutant alleles (SALK_144420 and SALK_073746, named *ys1-2* and *ys1-3*, respectively) from the SALK mutant collection, and both mutants had the same phenotype as *ys1-1* (Figure 1a,b). The T-DNA insertions in *YS1* were also confirmed by amplification of a specific band in the homozygous mutant plants (Figure 1h). A RT-PCR analysis showed that the mRNA of *YS1* was not detectable in the mutants, indicating that *ys1-1*, *ys1-2* and *ys1-3* were null alleles (Figure 1i). In the wild type, *YS1* mRNA was detectable in roots, rosette and cauline leaves, stems, flowers and siliques of mature plants as well as green and etiolated seedlings (Figure 1j).

The chloroplast transit peptide in YS1 predicted by ChloroP (http://www.cbs.dtu.dk/services/ChloroP/; Emanuelsson *et al.*, 1999) comprises the N-terminal 30 amino acids. To confirm the intracellular localization of YS1, we made a construct in which the first 100 N-terminal amino acid residues (from Met-1 to Phe-100) were fused to green fluorescent protein (GFP) driven by the 35S CaMV promoter (*35S:YS1*$_{1–100}$*-GFP*), and transformed it into wild-type plants. The fluorescent images showed that GFP fluorescence completely overlapped with chloroplast autofluorescence, whereas GFP fluorescence of the control was detected in the cytosol (Figure S3). This result confirmed that YS1 is localized in chloroplasts.

Loss of YS1 leads to low photosynthetic capacity

To understand the physiological role of YS1 in chloroplast development, we characterized the photosynthetic capacity of the mutant. Chlorophyll fluorescence induction experiments showed that the ratio of variable fluorescence to maximum fluorescence (F_v/F_m), which reflects maximum photochemical efficiency of photosystem II (PSII), was significantly lower in *ys1* (0.20 $\pm$ 0.02) than in the wild type (0.78 $\pm$ 0.01) (Figure 2a and Table 1). This was supported by the quantum yield of PSII (ΦPSII), with 0.12 $\pm$ 0.01 in *ys1* versus 0.56 $\pm$ 0.02 in the wild type (Table 1). Loss of YS1 also led to a significant decrease in the electron transport rate (ETR), the relative flow of electrons through PSII (Table 1). The reduced ETR was coincident with the enhanced 1 – *qP* value, a parameter of the oxidation state of the plastoquinone Q_A of PSII, indicating that Q_A was in a reduced state and electron flow was impaired downstream of Q_A in *ys1*. To find out whether ETR was also blocked at PSI, we determined P700 redox kinetics under far-red light. Our data showed that the mutant had significantly lower PSI activity relative to the wild type (Figure 2b). Additionally, $\Delta A/\Delta A_{max}$ was significantly increased (Table 1), indicating that electron transport was also restricted in the intersystem chain.

We further examined the status of PSI and PSII using 77 K chlorophyll fluorescence emission spectra from seedlings. Compared with the highly reduced level of fluorescence emission from PSI (with a maximum at 735 nm), the magnitude of the fluorescence emission from PSII was not obviously altered by *YS1* mutation (Figure 2c). The signal of the high-wavelength band indicated a nearly normal accumulation of the outer antenna of PSII, suggesting that the accumulation of the PSI core complex was more decreased than PSII in *ys1*. However, a blue shift of the characteristic PSII fluorescence peak at 687 nm and 695 nm was observed in the mutant, indicating that excitons were unable to be transferred efficiently to the PSII reaction centre. A similar peak shift of the PSI fluorescence was also detected in *ys1* (Figure 2c). These shifts were probably caused by partial disruption of energy transfer from the antennae to photosystems. Additionally, the ratio of F_{PSI} to F_{PSII} was significantly lower in *ys1* relative to the wild type (Table 1). Taken together, these results clearly indicate that photosynthetic capacity is severely lowered in the *ys1* mutants.

Thylakoid membranes are disturbed in ys1

To unveil the ultrastructural basis of delayed greening of *ys1* seedlings, we examined dynamic changes in thylakoid membranes using transmission electron microscopy. In leaves of wild-type seedlings grown in the light, chloroplasts were crescent-shaped and contained a well-formed thylakoid structure including stroma thylakoids and grana thylakoids (Figure 3a). In contrast, at the same stage, the chloroplasts of mutant seedlings were small and thylakoid membrane structure was distorted in various ways. As shown in Figure 3(b), some plastids formed rudimentary thylakoids consisting only of grana lamellae and failed to accumulate stroma lamellae; some were filled with numerous vesicles

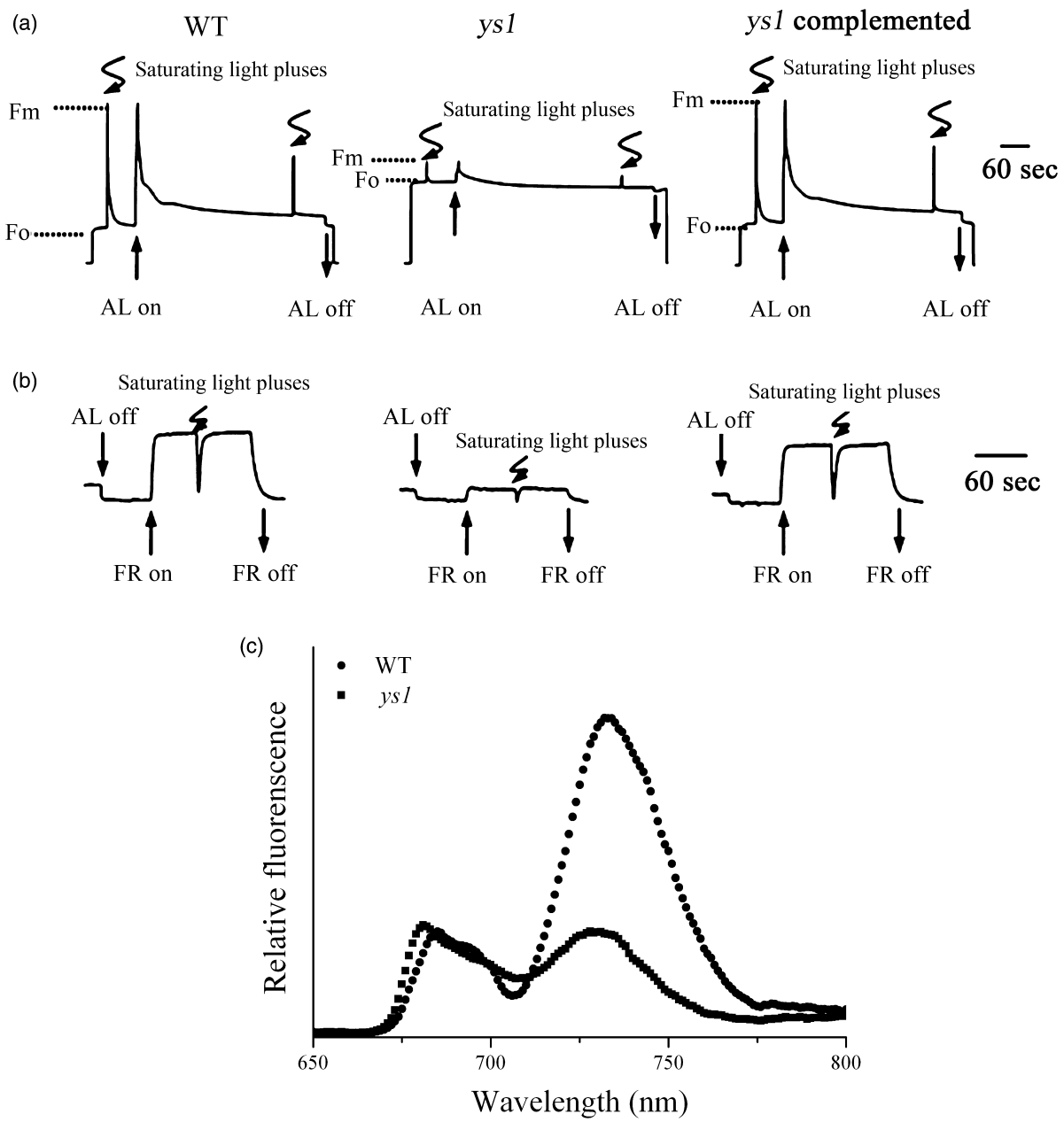


Figure 2. Analysis of photosynthetic activity of 8-day-old seedlings.
(a) Induction of chlorophyll fluorescence. F_0, minimal fluorescence. F_m, maximum fluorescence. AL, actinic light (85 μmol photons m^{-2} sec^{-1}).
(b) P700 redox kinetics under far-red light (FR, 720 nm).
(c) Low-temperature (77 K) fluorescence emission spectra. WT, wild type. *ys1* complemented, the full genomic sequence of *YS1* was transformed into the *ys1* mutant.

(Figure 3c). Chloroplasts with abnormal thylakoid membranes were still dominant in 7-day-old leaves (Figure 3d). At day 11, chloroplasts with well-organized grana thylakoids appeared (Figure 3e), and at day 16 most chloroplasts were oval with a well-developed thylakoid structure (Figure 3f). The ultrastructure of chloroplasts was indistinguishable between *ys1* and the wild type after 3 weeks' growth (Figure 3g). Etioplasts in dark-grown cotyledons were indistinguishable between *ys1* and the wild type (Figure 3h,i). These results confirm that YS1 plays an important role in the early stages of chloroplast biogenesis during leaf greening.

Accumulation of δ-aminolevulinic acid is reduced in the light-grown ys1 *mutants*

Since no difference in etioplast ultrastructure was observed between *ys1* and the wild type (Figure 3h,i), we reasoned that cotyledons of etiolated *ys1* seedlings might be able to

Table 1 Spectroscopic and fluorimetric data for wild-type (WT) and *ys1 s*eedlings. Values represent means of three to six independent measurements ($n \pm$ SE)

	F_0	F_v/F_m	ΦPSII	$1 - qP$	ETR	$\Delta A/\Delta A_{max}$	F_{PSI}/F_{PSII}
WT	372.7 ± 62.5^a	0.78 ± 0.01^a	0.56 ± 0.02^a	0.12 ± 0.04^a	41.6 ± 1.69^a	0.27 ± 0.12^a	2.69 ± 0.47^a
ys1	727.3 ± 195.9^b	0.20 ± 0.02^b	0.12 ± 0.01^b	0.28 ± 0.01^b	6.93 ± 1.97^b	0.69 ± 0.13^b	1.36 ± 0.13^b

F_0, minimal chlorophyll *a* fluorescence of dark-adapted leaves; *qP*, photochemical quenching; F_v/F_m, ratio of variable to the maximum fluorescence; ΦPSII, effective quantum yield of PSII; ETR (electron transport rate), calculated as photosynthetically active radiation × 0.5 × 0.84 × ΦPSII (μmol photons m^{-2} sec^{-1}); $\Delta A/\Delta A_{max}$, P700 oxidation ratio; F_{PSI}/F_{PSII}, ratio of 77 K fluorescence emitted by PSI and PSII.
[a,b]Entries with different superscript letters in the same column are significantly different ($P < 0.05$, *t*-test).

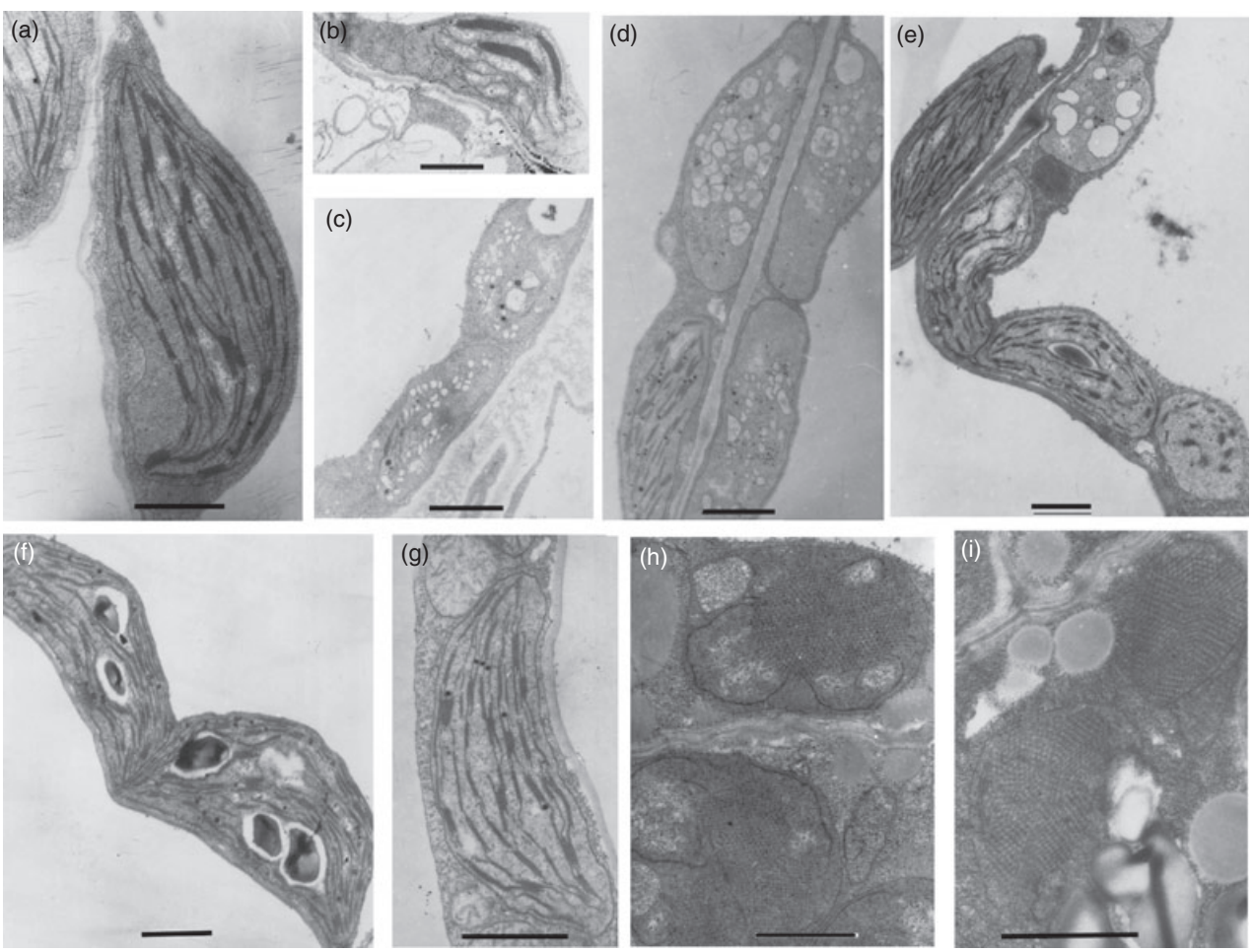


Figure 3. Ultrastructure of chloroplasts from wild-type and *ys1* leaves.
Representative plastids are shown in: (a) 4-day-old light-grown wild-type leaves; (b) and (c) 4-day-old light-grown *ys1* leaves; (d) 7-day-old light-grown *ys1* leaves; (e) 11-day-old light-grown *ys1* leaves; (f) 16-day-old light-grown *ys1* leaves; (g) 22-day-old light-grown *ys1* leaves; (h) 5-day-old etiolated wild-type cotyledons; (i) 5-day-old etiolated *ys1* cotyledons. Bars = 1 μm.

turn green like those of the wild type. To test this, 5-day-old seedlings grown in darkness were transferred to the light. The cotyledons of *ys1* seedlings were able to green almost as much as the wild type 2 days after exposure to light (Figure 4a,b). However, new leaves emerging from the seedlings were still yellow (Figure 4c). These results are consistent with the time-course of changes in chlorophyll content (Figure 4d). Mutant seedlings had accumulated slightly lower levels of chlorophyll than the wild type at 12 h but contained only 59.2% of the chlorophyll of the wild type at 48 h after transfer to light (Figure 4d). We hypothesized that delayed greening of *ys1* seedlings in light might result from a low level of δ-aminolevulinic acid (ALA), a precursor for chlorophyll synthesis. As shown in Figure 4(e), the level

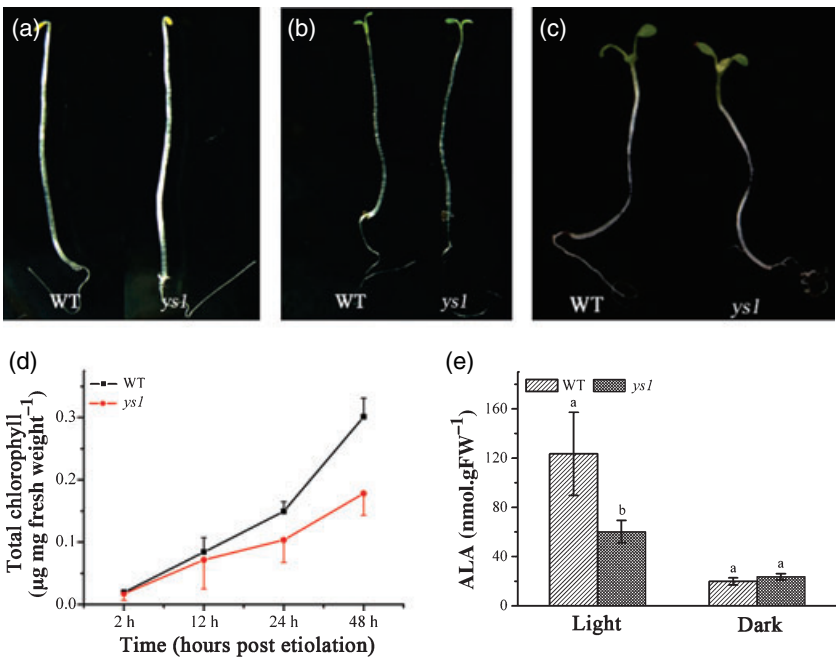


Figure 4. Dark treatment rescues the virescent phenotype via promotion of δ-aminolevulinic acid (ALA) biosynthesis in *ys1*.
(a) Five-day-old etiolated seedlings.
(b, c) Five-day-old etiolated seedlings transferred to light for 2 days and for 5 days, respectively.
(d) Change in chlorophyll content of plants after 5-day-old etiolated seedlings were shifted to continuous light.
(e) Accumulation of ALA in 5-day-old plants continuously incubated in light or in darkness, respectively. Those with different letters (a and b) on the error bars are significantly different ($P < 0.05$, *t*-test). Data are means ± SD ($n = 3$–4).

of ALA was indistinguishable between *ys1* and the wild type in 5-day-old etiolated plants, whereas accumulation of ALA was significantly reduced in the mutant compared with the wild type in light-grown 5-day-old plants. These results suggest that accumulation of chlorophyll may be limited by ALA content in light-grown *ys1* plants.

Assembly of photosynthetic complexes is impaired in ys1

We examined the level of photosynthetic pigment-containing complexes by blue native (BN)-PAGE and sucrose gradient centrifugation. Our data showed that PSII/light-harvesting complex (LHC) II supercomplexes, the PSI/LHCI complex and the Cyt b_6/f complex were all strongly reduced in 8-day-old mutant seedlings compared with the wild type (Figure 5a,c). Additionally, *ys1* accumulated less LHCII trimer but a higher level of LHCII monomer relative to the wild type (Figure 5a,c), suggesting that assembly of LHCII trimer is blocked. In the 30-day-old plants, however, no significant differences in the pigment-containing complexes were found between the mutant and the wild type (Figure 5b,d).

To understand how mutation of YS1 led to reduced photosynthetic complexes, we assayed the accumulation of representative subunits of thylakoid membrane complexes by immunoblots. Except for AtpB, the levels of plastid-encoded proteins such as D1, D2, PsaA, PsaC, Cyt b_6 and Cyt *f* were significantly reduced in *ys1* compared with the wild type (Figure 6). Compared with the PSII reaction centre core subunits (D1 and D2), which were reduced to 50% of the wild type, the relative level of the PSI core protein (PsaA) was more drastically reduced in *ys1*. This was consistent with the results of chlorophyll fluorescence analyses. Moreover, the reduction of Cyt b_6/f subunits was also in agreement with the result from the analysis of ETR. In contrast, the nuclear-encoded subunits such as PsbO, PsaD, Lhca1 and Lhcb1 were unaltered in the mutant (Figure 6). These results suggest that the loss-of-function of *YS1* specifically results in a decrease in plastid-encoded proteins.

The reduction of plastid-encoded proteins in ys1 *is due to reduced levels of transcripts*

Protein accumulation could be controlled at a transcriptional or post-transcriptional level. To distinguish between these two possibilities, the expression of a wide variety of plastid genes was analyzed by northern blot and quantitative RT-PCR. Results showed that mutation of YS1 caused a marked decrease in transcript accumulation of PEP-dependent photosynthesis-related (class I) genes (Figure 7a), but enhanced accumulation of transcripts of NEP-dependent housekeeping (class III) genes (Figure 7b). The 4.8-kb transcript detected with the *atpB* probe has been identified to be the product transiently transcribed by NEP in *sig6*, where NEP activity was upregulated (Schweer *et al.*, 2006). This might explain why AtpB accumulated at the same level as that of the wild type in *ys1* (Figure 6).

Consistent with protein blot analysis, nuclear gene expression was not altered in *ys1* (Figure 7c). Interestingly, expression of a group of PEP-dependent *trn* genes including *trnE* significantly decreased in *ys1*, both in light-grown (Figure 7d) and etiolated plants (Figure 7e). The expression in etiolated plants of other PEP- or NEP-dependent genes such as *psbA*, *psaA*, *petA*, *atpB*, *rbcL* and *accD* was not radically different between *ys1* and the wild type (Figure 7e).

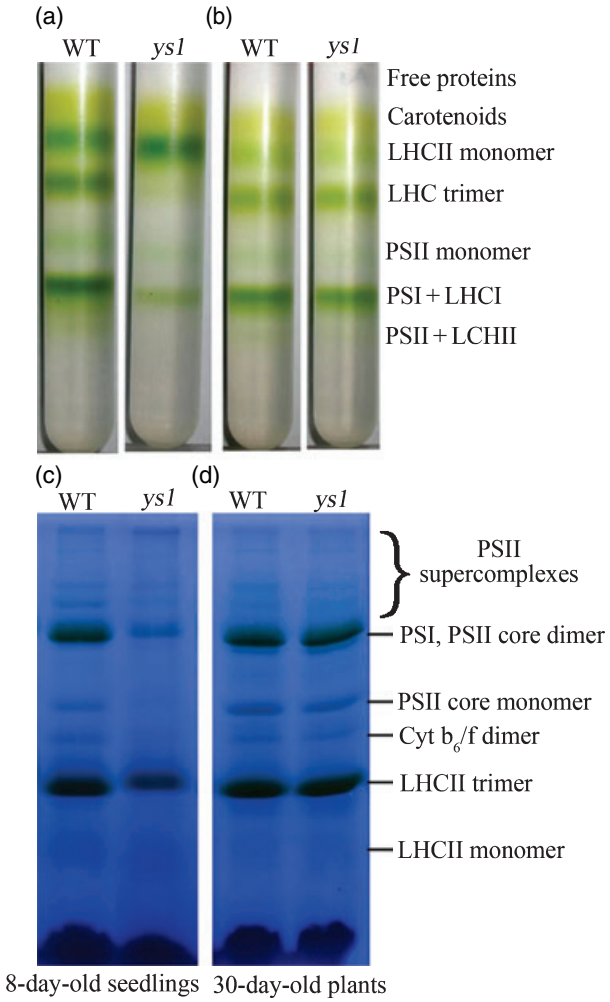


Figure 5. Photosynthetic pigment-containing complexes extracted from 8-day-old and 30-day-old plants were separated by sucrose gradient centrifugation (40 μg of chlorophyll per lane) or by blue native (BN)-PAGE (10 μg of chlorophyll per lane). The identities of the resolved colour bands are indicated.

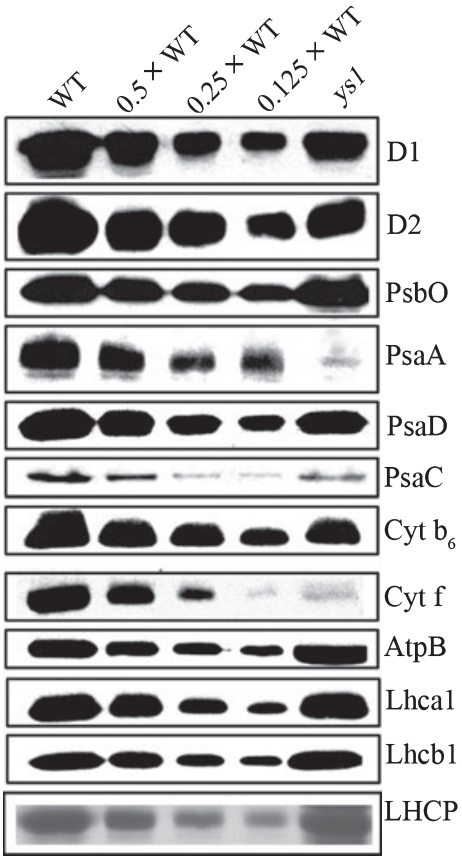


Figure 6. Immunoblot analysis of photosynthetic membrane proteins from *ys1* and wild-type seedlings.
Thylakoid membranes were isolated from 8-day-old seedlings. A dilution series containing total protein corresponding to 2, 1, 0.5, 0.25 and 2 μg of chlorophyll was loaded in the lanes marked by WT, 0.5×, 0.25×, 0.125× and *ys1*, respectively. Western blot analysis was carried out with the antisera indicated. The light-harvesting complex protein (LHCP) was used as a loading control.

To test whether YS1 also translationally regulates plastid gene expression, polysome-bound mRNAs were fractionated by sedimentation. The extent to which transcripts are associated with polysomes reflects the efficiency of translation initiation and elongation (Barkan, 1993). The distribution of both ribosomal RNAs and examined mRNAs was unaltered between the wild type and the mutant (Figure S4). Thus, YS1 is unlikely to mediate plastid gene expression via a translational mechanism, and we conclude that the phenotype in *ys1* is due to decreased PEP activity.

YS1 *is required for editing of rpoB transcripts*

Several PPR proteins with C-terminal E domains have been shown to be required for editing of specific sites in Arabidopsis chloroplasts (Kotera *et al.*, 2005; Okuda *et al.*, 2007; Chateigner-Boutin *et al.*, 2008). Related proteins of the DYW subgroup have also been predicted to function in RNA editing on phylogenetic grounds (Salone *et al.*, 2007). We therefore screened for defects in RNA editing in *ys1* at all 34 sites known to be edited in wild-type chloroplasts (Chateigner-Boutin and Small, 2007). A single site, at site 25992 in *rpoB* transcripts, was not detectably edited in light- or dark-grown *ys1* seedlings or in adult (and fully green) *ys1* plants (Figure 8). The two other editing sites in *rpoB* transcripts (and all 31 sites in other transcripts) are edited normally in *ys1* (Figure S5).

Discussion

YS1 is an editing specificity factor

The loss of YS1 in *ys1* mutants results in the loss of a single editing event in the *rpoB* transcript. Although *rpoB* transcripts are more abundant in *ys1*, and increased transcript

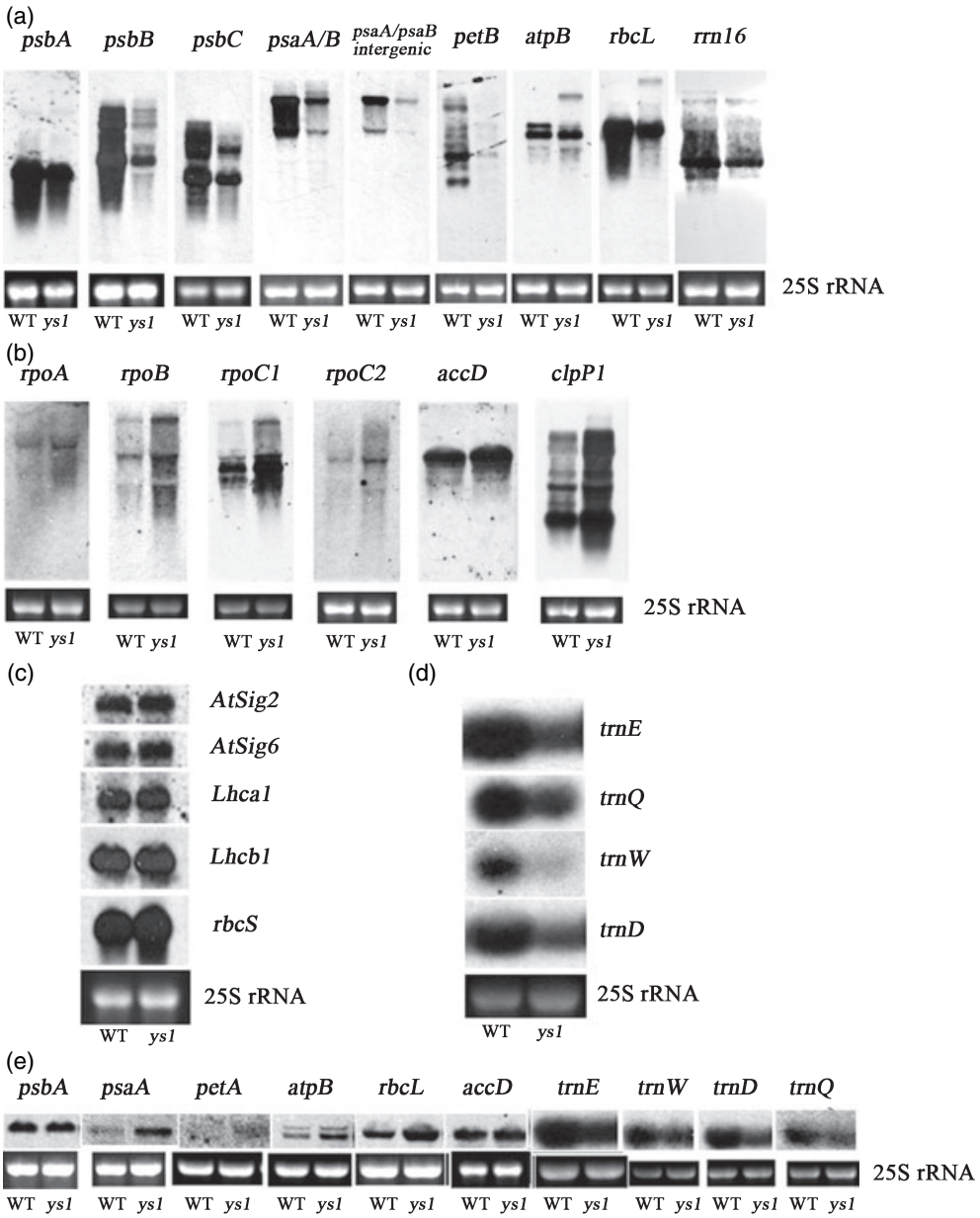


Figure 7. Messenger RNA levels of plastid and nuclear genes encoding plastid-localized proteins in *ys1* and wild-type seedlings.
Ten μg or 5 μg of total RNA was isolated from 8-old-day seedlings grown in light or from 5-day-old seedlings grown in darkness, respectively.
(a) Plastid-encoded RNA polymerase (PEP)-dependent gene transcripts in 8-old-day seedlings.
(b) Nucleus-encoded RNA polymerase (NEP)-dependent gene transcripts in 8-old-day seedlings.
(c) Nuclear gene transcripts.
(d) Transcript levels of a group of plastid *trn* genes.
(e) Messenger RNA levels of representative genes in each gene class in 5-day-old etiolated seedlings.
The bottom of each lane shows loading controls (25S rRNA) stained with ethidium bromide.

levels can titrate out editing factors (Chaudhuri *et al.*, 1995), editing of this site is undetectable in *ys1* mutants (Figure 8), strongly suggesting that YS1 is absolutely required to edit this site. Furthermore, neither of the other two *rpoB* editing sites is affected, indicating that the 25992 site is specifically targeted by YS1. Three other protein factors have been described that are similarly required for editing of specific sites: CRR4 (Kotera *et al.*, 2005) and CRR21 (Okuda *et al.*, 2007), required for editing of different sites in *ndhD*, and CLB19, required for editing of two sites in *clpP* and *rpoA* (Chateigner-Boutin *et al.*, 2008). The three previous factors are all similar PPR proteins within the E subclass (so-named as these proteins carry a C-terminal E domain after the PPR repeats), whereas YS1 differs in having a further C-terminal

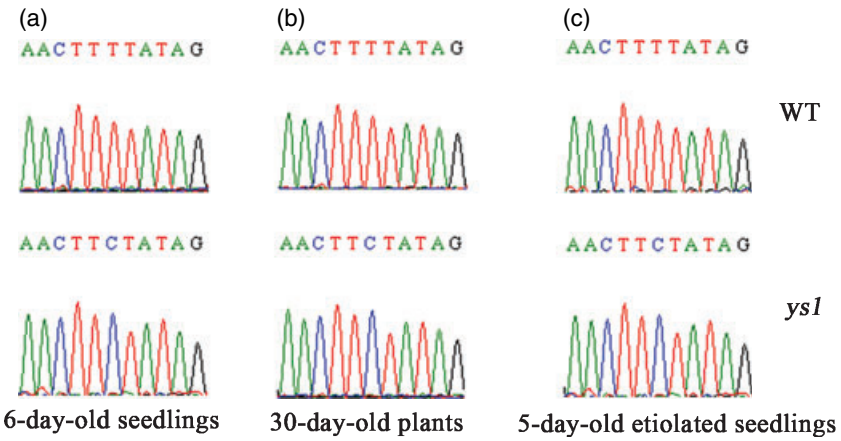


Figure 8. *ys1* is defective in editing of *rpoB* transcripts.
Bulk RT-PCR products of *rpoB* transcripts from wild-type and *ys1* plants of various ages were sequenced. The regions of the chromatograms covering the editing site at 25992 (central peak in all panels) are shown. In all three sets of samples, the base at 25992 is a T (indicating the cDNA was derived from edited mRNA) in wild-type sequence and an unedited C in *ys1* sequences.

DYW domain (Figure S6). The DYW domain was hypothesized to carry RNA editing activity based on its phylogenetic distribution (it has only ever been found in plants that carry out RNA editing, and is absent from *Marchantia* and green algae that do not edit organellar transcripts) and based on the presence of conserved residues that match the active site of known editing deaminases from other organisms (Salone *et al.*, 2007). However, no genetic evidence in favour of a link between DYW proteins and editing has been published apart from a broad co-localisation between two DYW-class genes and a quantitative trait locus (QTL) for editing efficiency in a cross between two Arabidopsis accessions (Bentolila *et al.*, 2005). The results presented here demonstrate that the DYW-containing PPR protein YS1 is an editing specificity factor for *rpoB* site 25992, but do not show that YS1 carries the editing activity itself; other essential factors may be required.

ys1 is an RpoB mutant

Editing at position 25992 changes codon 113 of *rpoB* to encode phenylalanine rather than serine. A phenylalanine codon is conserved at this position in many plant lineages including the liverwort *Marchantia polymorpha* that does not edit its plastid transcripts (Figure S7). This phenylalanine is located within the essential B domain of the beta subunit of the *Escherichia coli* RNA polymerase. Mutations in this domain are lethal in *E. coli* (Severinov *et al.*, 1993). Therefore a defect in editing at this position is likely to affect RpoB function. The transcript patterns in *ys1* mutants are characteristic of mutants in which PEP activity is reduced (Pfalz *et al.*, 2006; Chateigner-Boutin *et al.*, 2008), consistent with the primary defect being in RpoB. Deletion of *rpoB* in tobacco (Allison *et al.*, 1996; De Santis-Maclossek *et al.*, 1999) gives rise to a much stronger phenotype (total lack of photosynthetic ability and green pigmentation), so the RpoB defect in *ys1* is likely to be a partial, not total, loss of activity. As transformation of Arabidopsis plastids is far from routine (Dhingra and Daniell, 2006), creating mutants of plastid genes is extremely difficult. The only similar mutant described, *clb19* (Chateigner-Boutin *et al.*, 2008), has editing defects in *clpP* as well as *rpoA*, complicating the analysis of the phenotype. *ys1* mutants therefore offer a currently unique opportunity to examine the role of the PEP polymerase during Arabidopsis development.

YS1 *is required for correct plastid gene expression during early chloroplast development*

It is well documented that chloroplast development upon shifting from dark to light is supported by a huge increase in plastid gene expression. Both NEP and PEP are already present in dry seeds and are probably active immediately after seed imbibition (Demarsy *et al.*, 2006). The level of NEP reaches a peak earlier than that of PEP during seed germination and early chloroplast development. The activity of NEP is reduced and maintained at a basal level during the later stages of chloroplast development, whereas at this stage PEP activity increases and is kept at a high level. PEP and NEP activities are highly coordinated to optimize plastid gene expression during light-triggered chloroplast development. Disturbance of the coordinated activities of NEP and PEP usually results in delayed or completely inhibited chloroplast development (Hajdukiewicz *et al.*, 1997; Kusumi *et al.*, 1997; Kanamaru *et al.*, 2001; Ishizaki *et al.*, 2005; Loschelder *et al.*, 2006; Pfalz *et al.*, 2006). Recently, Hanaoka *et al.* (2005) reported that $tRNA^{Glu}$ is a direct mediator to coordinate the activities of PEP and NEP through a feedback mechanism.

In *ys1*, PEP transcripts are lower in the light than in the wild type (Figure 6a,d), whilst NEP transcripts are maintained at a high level (Figure 6b). The high activity of NEP observed in *ys1* is probably induced by the low accumulation of $tRNA^{Glu}$, which fails to inhibit the activity of NEP. In the dark, however, NEP and PEP transcripts, except for those of a group of *trn* genes, accumulate to similar levels in *ys1* to those in the wild type. No obvious visible, ultrastructural or molecular phenotypic defects were noted in etiolated *ys1* seedlings, indicating that although *rpoB* transcripts are

edited in etiolated wild-type seedlings, this editing event is unlikely to be important for etioplast development, as opposed to the requirement for YS1-mediated *rpoB* editing in light-grown seedlings.

YS1 *helps to coordinate biosynthesis of chlorophyll and protein in light*

Although enzymes catalysing all reactions in the chlorophyll biosynthetic pathway have been identified in higher plants, we are still facing a challenge to understand how plants incorporate environmental and developmental cues to coordinate chlorophyll synthesis with functional chloroplast development (Tanaka and Tanaka, 2007). In the chlorophyll biosynthetic pathway, formation of ALA is the first limiting step and is regulated by multiple mechanisms. The first reaction required for ALA biosynthesis is catalysed by glutamyl-tRNA synthetase, which ligates $tRNA^{Glu}$ to glutamate, and is shared with protein synthesis in plastids. In light-grown plants, chlorophyll biosynthesis and plastid protein synthesis are both strongly activated simultaneously. We show that under these conditions, loss of YS1 leads to decreased $tRNA^{Glu}$ accumulation, decreased ALA and decreased chlorophyll biosynthesis. It is plausible to infer that glutamyl-$tRNA^{Glu}$ is a rate-limiting factor for chlorophyll biosynthesis in *ys1* in developing chloroplasts when protein biosynthesis competes strongly for the substrate. In etiolated seedlings the situation is different. Kleffmann *et al.* (2007) demonstrated that the tetrapyrrole pathway is very active in the dark, whereas photosynthesis-related metabolic pathways and plastid gene expression are only significantly activated after exposure to light. The reduced level of $tRNA^{Glu}$ in *ys1* has no effect on accumulation of ALA in etiolated seedlings, presumably because plastid protein synthesis in these conditions is greatly reduced, and no longer competes. When the seedlings are transferred to the light, chlorophyll biosynthesis can proceed without further requirement for ALA and protein synthesis can proceed rapidly without competition for glutamyl-$tRNA^{Glu}$. This temporal uncoupling between ALA synthesis and plastid protein synthesis in etiolated seedlings transferred to the light probably explains how *ys1* plants can green rapidly under these conditions.

We propose a model to explain the role of YS1 during the early development of chloroplasts (Figure 9). In this model, YS1 is required to edit *rpoB* transcripts to produce active PEP capable of transcribing a subclass of *trn* genes. The expression of the plastid *trnE* gene and/or other *trn* genes is required to meet the demand of developing chloroplasts for biosynthesis of protein and chlorophyll. Transfer RNA^{Glu} also plays an important role in feedback inhibition of NEP activity. Loss of YS1 leads to a low level of $tRNA^{Glu}$, which limits protein and chlorophyll biosynthesis in the developing chloroplasts, and lengthens the course of chloroplast development. It will be important in the future to examine whether the extent of *rpoB* editing by YS1 varies according to cell type, developmental stage or environmental conditions. Such variation might indicate that the rate of *rpoB* editing controls expression of active PEP.

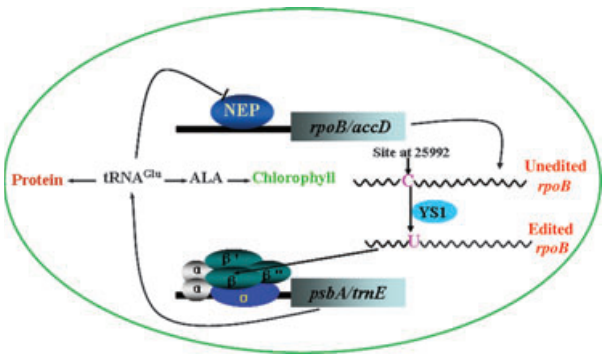


Figure 9. A working model for the role of YS1 in rapid chloroplast development in light-grown Arabidopsis seedlings.
In developing chloroplasts, nucleus-encoded RNA polymerase (NEP) together with plastid-encoded RNA polymerase (PEP) sigma factors and YS1 are actively expressed and imported into plastids. The NEP transcribes housekeeping genes such as *rpoB*, a subunit of PEP. YS1 edits *rpoB* transcripts, allowing fully active RpoB PEP subunits to be translated. Newly assembled PEP is activated by sigma factors to regulate transcription of photosynthesis-related protein-coding and *trn* genes, notably *trnE* encoding $tRNA^{Glu}$. In the absence of YS1, PEP activity stays low, and the level of $tRNA^{Glu}$, which couples biosynthesis of chlorophyll and proteins, is significantly reduced at the seedling stage, leading to the virescent phenotype. The lowered level of $tRNA^{Glu}$ is unable to suppress NEP activity via the feedback regulatory mechanism, thus NEP activity remains high in *ys1* seedlings.

Experimental procedures

Plant growth, plasmids, identification of mutants and transgenic plants

All wild-type and mutant Arabidopsis plants were in the ecotype Columbia background. Surface-sterilized seeds were stratified for 3 days, and then sown onto half-strength Murashige and Skoog (MS) medium with 1% sucrose at 22°C under long-day (16 h light/8 h dark) conditions at 120 μmol photons m^{-2} sec^{-1}.

The various plasmids were constructed using the GATEWAY cloning system (Invitrogen, http://www.invitrogen.com/). To conduct the complementation experiment, a 3778-bp *YS1* genomic fragment from the ATG start codon to 472 bp downstream of the 3′-untranslated (UTR) was amplified by PCR using the primers: 5′-CACCATGGCGATGTTGGGTAATGTTCT-3′ and 5′-CTAAGTTGCTTGCTCTATGGTGAGCATTGCT-3′. For analysis of subcellular localization of YS1, the genomic fragment encoding the N-terminal 100 amino acids (Met-1 to Phe-100) of YS1 was amplified. The resulting DNA fragments were cloned into pENTR™/SD/D-TOPO vector (Invitrogen). After sequence verification, the fragments were recombined into pGWB destination binary vectors (Research Institute of Molecular Genetics, Shimane University, Japan). The constructs were transformed into *A. tumefaciens* GV3101 strain and introduced into Arabidopsis mutant plants by the floral dip method (Clough and Bent, 1998). Transgenic T_1 plants were identified by kanamycin resistance.

The T-DNA insertion lines SALK_123515 (*ys1-1*), SALK_144420 (*ys1-2*) and SALK_073746 (*ys1-3*) were obtained from the Arabidopsis Biological Resource Center (ABRC; Ohio State University). Plasmid rescue was conducted as described by (Ichikawa *et al.*, 2003). Primers used for genomic PCR verification of T-DNA insertion in the mutants were 5′-CGCACTCAACATCGAAATCCTATC-3′ and 5′-GGCGATGTTGGGTAATGTTCTTC-3′. The left border primer used was 5′-GCGTGGACCGCTTGCTGCAACT-3′. The plants homozygous for T-DNA insertion were checked for lack of *YS1* transcripts using primers 5′-TCTGGTGTTAGACCTGATAGGATTTC-3′ and 5′-AGTCTGACATCAAGCTGAATTCCGTTC-3′ by RT-PCR.

Pigments and ALA analysis

Chlorophylls and total carotenoids were estimated spectrophotometrically according to Lichtenthaler (1987). For ALA determination, seedlings grown for 7 days on 1/2 MS medium containing 1% sucrose were incubated with 50 mM levulinic acid (LA) in 0.1 M sodium phosphate buffer (pH 7.0) in the light or the dark for 12 h. The method of ALA measurement was as previously described (Kruse *et al.*, 1997). Measurements were repeated in three independent experiments.

Chlorophyll fluorescence induction, P700 redox kinetics and 77 K fluorescence emission spectra

In vivo chlorophyll *a* fluorescence measurements were performed with 8-day-old Arabidopsis seedlings grown in the light at room temperature using a pulse amplitude-modulated fluorometer equipped with a PDA100 data acquisition system to record fast changes (PAM 101, Walz, Germany). Prior to measurements, the fibre optic of the emitter/detector unit (101-ED) was positioned close to the upper surface of the plants. Consecutive saturating light pulses of 800 msec were applied by halogen lamps. The connected emitter–detector subunit allowed us to measure the light-induced P700 redox kinetics by absorbance changes at 830 nm. Chlorophyll fluorescence emission spectra of ground leaves at 77 K were recorded as described (Meurer *et al.*, 1996a,b).

Microscopy

Wild-type and mutant leaves from different growth stages of plants were collected For transmission electron microscopy processing. The tissue was cut into small pieces and fixed in 2.5% glutaraldehyde in phosphate buffer (pH 7.2) for 4 h at 0°C. After fixation, the tissue was rinsed and post-fixed overnight at 4°C in 1% OsO_4. After rinsing in phosphate buffer, the samples were dehydrated in an ethanol series, infiltrated with a graded series of epoxy resin in epoxy propane, and embedded in Epon 812 resin. Thin sections were stained in uranium acetate followed by lead citrate and viewed with a transmission electron microscope (Phillips CM120, http://www.fei.com). The GFP fluorescence was observed with a confocal laser scanning microscope (FITC488, Zeiss LSM500, http://www.zeiss.com/). The filter sets used were BP505-545 (excitation 488 nm; emission 505–545 nm) and LP585 (excitation 488 nm; emission 585 nm) to detect GFP and the chlorophyll autofluorescence.

Separation of thylakoid membrane complexes

The leaves were homogenized in an ice-cold isolation buffer containing 400 mM sucrose, 50 mM 2-amino-2-(hydroxymethyl)-1,3-propanediol (TRIS)-HCl, pH 7.6, and 10 mM NaCl with a chilled mortar and pestle and filtrated through two layers of nylon cloth. The filtrate was centrifuged at 5000 ***g*** for 10 min. The thylakoid pellets were washed with isolation buffer, recentrifuged and suspended in isolation buffer. The resulting thylakoid membrane pellets were either used freshly or frozen in liquid N_2 and stored at −70°C before use. The chlorophyll content was determined spectrophotometrically according to Porra *et al.* (1989). Thylakoid membrane complexes were separated by BN-PAGE (Schägger *et al.*, 1994; Cline and Mori, 2001) or by sucrose density gradient centrifugation (Swiatek *et al.*, 2003).

Immunoblot analysis

Isolated thylakoid samples were separated by 15% SDS-PAGE, transferred to Hybond-ECL Nitrocellulose membrane (Amersham Biosciences, http://www.amersham.com/), and immunoblotted with various antibodies (Agrisera, http://www.agrisera.com/). Antibodies were detected using a chemiluminescence detection system (ECL; Amersham Biosciences) according to the manufacturer's instructions.

Analyses by RT-PCR and northern blot

Total RNA was isolated using the RNAgents total RNA isolation system (Promega, http://www.promega.com/). Semi-quantitative RT-PCR was carried out using a reverse transcription system (Promega). The following gene-specific primers were used: the Arabidopsis *Actin2* gene (*At3g18780*) was amplified using primers 5′-TCTTCTTCCGCTCTTTCTTTCC-3′ and 5′-TCTTACAATTTCCCGCTCTGC-3′ as an internal positive control. Amplified DNA products were separated using an agarose gel followed by image analysis with GIS-2010 (Tanon, http://www.bio-tanon.com.cn). Northern blot analysis was carried out as described (Huang *et al.*, 2000). The sequences of the PCR primers used for amplifying the genes are presented in Table S1. The plastome-wide screen of RNA editing was carried out as described by Chateigner-Boutin and Small (2007). The results of screening for RNA editing and splicing were confirmed by sequencing. The editing sites were amplified by PCR with primers *rpoB*-F2 (5′-CACCATGCTTGGGGATGAAAAAGA-3′) and *rpoB*-R2 (5′-TTAAACTTCCTTCCTATTAATCTGGA-3′). The RT-PCR products were sequenced directly.

Acknowledgements

The authors thank Dr A. Barkan for suggestions on experiments and helpful discussion. This work was supported by grants to I.S. from the Australian Research Council (CE0561495), the Western Australian State Government and the University of Western Australia, and to J.H. from the Natural Scientific Foundation of China (30670161), the Chinese Academy of Sciences (KSCX2-YW-N-059), the Bairen Project of Chinese Academy of Science and the Ministry of Science and Technology of China (2007CB108800).

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. Open reading frame of *YS1* deduced from 3′-rapid amplification of cDNA ends (RACE) experiments.

Figure S2. Alignment of YS1 homologues from other species.

Figure S3. Subcellular localization of YS1 estimated by GFP-tagging in mesophyll cells.

Figure S4. Association of several chloroplast mRNAs with polysomes in *ys1* and wild-type seedlings.
Figure S5. Complete screen of editing sites in *ys1* mutants.
Figure S6. Alignment of the YS1 sequence with that of the editing factors CRR4, CRR21 and CLB19.
Figure S7. Alignment of RpoB sequences to show conservation of the amino acid affected by editing.
Table S1. Gene-specific primers used for amplifying RNA gel blot probes.
Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

References

Allison, L.A. (2000) The role of sigma factors in plastid transcription. *Biochimie*, **82**, 537–548.
Allison, L.A., Simon, L.D. and Maliga, P. (1996) Deletion of *rpoB* reveals a second distinct transcription system in plastids of higher plants. *EMBO J.* **15**, 2802–2809.
Alonso, J.M., Stepanova, A.N., Leisse, T.J. *et al.* (2003) Genome-wide insertional mutagenesis of Arabidopsis thaliana. *Science*, **301**, 653–657.
Archer, E.K. and Bonnett, H.T. (1987) Characterization of a virescent chloroplast mutant of tobacco. *Plant Physiol.* **83**, 920–925.
Aubourg, S., Boudet, N., Kreis, M. and Lecharny, A. (2000) In *Arabidopsis thaliana*, 1% of the genome codes for a novel protein family unique to plants. *Plant Mol. Biol.* **42**, 603–613.
Barkan, A. (1993) Nuclear mutants of maize with defects in chloroplast polysome assembly have altered chloroplast RNA metabolism. *Plant Cell*, **5**, 389–402.
Barkan, A. and Goldschmidt-Clermont, M. (2000) Participation of nuclear genes in chloroplast gene expression. *Biochimie*, **82**, 559–572.
Beick, S., Schmitz-Linneweber, C., Williams-Carrier, R., Jensen, B. and Barkan, A. (2008) The pentatricopeptide repeat protein PPR5 stabilizes a specific tRNA precursor in maize chloroplasts. *Mol. Cell. Biol.* **28**, 5337–5347.
Benedict, C.R. and Ketring, D.L. (1972) Nuclear gene affecting greening in virescent peanut leaves. *Plant Physiol.* **49**, 972–976.
Bentolila, S., Chateigner-Boutin, A.L. and Hanson, M.R. (2005) Ecotype allelic variation in C-to-U editing extent of a mitochondrial transcript identifies RNA-editing quantitative trait loci in Arabidopsis. *Plant Physiol.* **139**, 2006–2016.
Chateigner-Boutin, A.L. and Small, I. (2007) A rapid high-throughput method for the detection and quantification of RNA editing based on high-resolution melting of amplicons. *Nucleic Acids Res.* **35**, e114.
Chateigner-Boutin, A.L., Ramos-Vega, M., Guevara-Garcia, A. *et al.* (2008) CLB19, a pentatricopeptide repeat protein required for editing of *rpoA* and *clpP* chloroplast transcripts. *Plant J.* **56**, 590–602.
Chaudhuri, S., Carrer, H. and Maliga, P. (1995) Site-specific factor involved in the editing of the psbL mRNA in tobacco plastids. *EMBO J.* **14**, 2951–2957.
Cline, K. and Mori, H. (2001) Thylakoid delta pH-dependent precursor proteins bind to a cpTatC-Hcf106 complex before Tha4-dependent transport. *J. Cell Biol.* **154**, 719–730.
Clough, S.J. and Bent, A.F. (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735–743.
da Costa e Silva, O., Lorbiecke, R., Garg, P. *et al.* (2004) The *Etched1* gene of *Zea mays* (L.) encodes a zinc ribbon protein that belongs to the transcriptionally active chromosome (TAC) of plastids and is similar to the transcription factor TFIIS. *Plant J.* **38**, 923–939.
De Santis-Maclossek, G., Kofer, W., Bock, A., Schoch, S., Maier, R.M., Wanner, G., Rudiger, W., Koop, H.U. and Herrmann, R.G. (1999) Targeted disruption of the plastid RNA polymerase genes *rpoA*, *B* and *C1*: molecular biology, biochemistry and ultrastructure. *Plant J.* **18**, 477–489.
Delannoy, E., Stanley, W.A., Bond, C.S. and Small, I.D. (2007) Pentatricopeptide repeat (PPR) proteins as sequence-specificity factors in post-transcriptional processes in organelles. *Biochem. Soc. Trans.* **35**, 1643–1647.
Demarsy, E., Courtois, F., Azevedo, J., Buhot, L. and Lerbs-Mache, S. (2006) Building up of the plastid transcriptional machinery during germination and early plant development. *Plant Physiol.* **142**, 993–1003.
Dhingra, A. and Daniell, H. (2006) Chloroplast genetic engineering via organogenesis or somatic embryogenesis. *Methods Mol. Biol.* **323**, 245–262.
Dyall, S.D., Brown, M.T. and Johnson, P.J. (2004) Ancient invasions: from endosymbionts to organelles. *Science*, **304**, 253–257.
Emanuelsson, O., Nielsen, H. and von Heijne, G. (1999) ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci.* **8**, 978–984.
Falcon de Longevialle, A.F., Meyer, E.H., Andrés, C., Taylor, N.L., Lurin, C., Millar, A.H. and Small, I.D. (2007) The pentatricopeptide repeat gene *OTP43* is required for *trans*-splicing of the mitochondrial *nad1* Intron 1 in *Arabidopsis thaliana*. *Plant Cell*, **19**, 3256–3265.
Falcon de Longevialle, A., Hendrickson, L., Taylor, N.L., Delannoy, E., Lurin, C., Badger, M., Millar, A.H. and Small, I. (2008) The pentatricopeptide repeat gene *OTP51* with two LAGLIDADG motifs is required for the *cis*-splicing of plastid *ycf3* intron 2 in *Arabidopsis thaliana*. *Plant J.* **56**, 157–168.
Hajdukiewicz, P.T., Allison, L.A. and Maliga, P. (1997) The two RNA polymerases encoded by the nuclear and the plastid compartments transcribe distinct groups of genes in tobacco plastids. *EMBO J.* **16**, 4041–4048.
Hanaoka, M., Kanamaru, K., Fujiwara, M., Takahashi, H. and Tanaka, K. (2005) Glutamyl-tRNA mediates a switch in RNA polymerase use during chloroplast biogenesis. *EMBO Rep.* **6**, 545–550.
Hattori, M., Miyake, H. and Sugita, M. (2007) A pentatricopeptide repeat protein is required for RNA processing of *clpP* Pre-mRNA in moss chloroplasts. *J. Biol. Chem.* **282**, 10773–10782.
Hedtke, B., Börner, T. and Weihe, A. (1997) Mitochondrial and chloroplast phage type RNA polymerases in *Arabidopsis*. *Science*, **277**, 809–811.
Hess, W.R. and Börner, T. (1999) Organellar RNA polymerases of higher plants. *Int. Rev. Cytol.* **190**, 1–59.
Hricová, A., Quesada, V. and Micol, J.L. (2006) The *SCABRA3* nuclear gene encodes the plastid RpoTp RNA polymerase, which is required for chloroplast biogenesis and mesophyll cell proliferation in Arabidopsis. *Plant Physiol.* **141**, 942–956.
Huang, J., Takano, T. and Akita, S. (2000) Expression of α-expansin genes in young seedlings of rice (*oryza sativa* L.). *Planta*, **211**, 467–473.
Ichikawa, T., Nakazawa, M., Kawashima, M. *et al.* (2003) Sequence database of 1172 T-DNA insertion sites in Arabidopsis activation-tagging lines that showed phenotypes in T1 generation. *Plant J.* **36**, 421–429.
Ikeda, T.M. and Gray, M.W. (1999) Characterization of a DNA-binding protein implicated in transcription in wheat mitochondria. *Mol. Cell. Biol.* **19**, 8113–8122.

Ishizaki, Y., Tsunoyama, Y., Hatano, K., Ando, K., Kato, K., Shinmyo, A., Kobori, M., Takeba, G., Nakahira, Y. and Shiina, T. (2005) A nuclear-encoded sigma factor, Arabidopsis SIG6, recognizes sigma-70 type chloroplast promoters and regulates early chloroplast development in cotyledons. *Plant J.* **42**, 133–144.

Kanamaru, K., Nagashima, A., Fujiwara, M., Shimada, H., Shirano, Y., Nakabayashi, K., Shibata, D., Tanaka, K. and Takahashi, H. (2001) An Arabidopsis sigma factor (SIG2)-dependent expression of plastid-encoded tRNAs in chloroplasts. *Plant Cell Physiol.* **42**, 1034–1043.

Kasai, K., Kawagishi-Kobayashi, M., Teraishi, M., Ito, Y., Ochi, K., Wakasa, K. and Tozawa, Y. (2004) Differential expression of three plastidial sigma factors, OsSIG1, OsSIG2A, and OsSIG2B, during leaf development in rice. *Biosci. Biotechnol. Biochem.* **68**, 973–977.

Kleffmann, T., von Zychlinski, A., Russenberger, D., Hirsch-Hoffmann, M., Gehrig, P., Gruissem, W. and Baginsky, S. (2007) Proteome dynamics during plastid differentiation in rice. *Plant Physiol.* **143**, 912–923.

Kotera, E., Tasaka, M. and Shikanai, T. (2005) A pentatricopeptide repeat protein is essential for RNA editing in chloroplasts. *Nature*, **433**, 326–330.

Koussevitzky, S., Stanne, T.M., Peto, C.A., Giap, T., Sjögren, L.L., Zhao, Y., Clarke, A.K. and Chory, J. (2007) An Arabidopsis thaliana virescent mutant reveals a role for ClpR1 in plastid development. *Plant Mol. Biol.* **63**, 85–96.

Kruse, E., Grimm, B., Beator, J. and Kloppstech, K. (1997) Developmental and circadian control of the capacity for δ-aminolevulinic acid synthesis in green barley. *Planta*, **202**, 235–241.

Kusumi, K., Mizutani, A., Nishimura, M. and Iba, K. (1997) A virescent gene *V1* determines the expression timing of plastid genes for transcription/translation apparatus during early leaf development in rice. *Plant J.* **12**, 1241–1250.

Lichtenthaler, H.K. (1987) Chlorophylls and carotenoids: pigments of photosynthetic biomembranes. *Methods Enzymol.* **148**, 350–382.

Liere, K. and Maliga, P. (1999) *In vitro* characterization of the tobacco *rpoB* promoter reveals a core sequence motif conserved between phage-type plastid and plant mitochondrial promoters. *EMBO J.* **18**, 249–257.

Loschelder, H., Schweer, J., Link, B. and Link, G. (2006) Dual temporal role of plastid sigma factor 6 in Arabidopsis development. *Plant Physiol.* **142**, 642–650.

Lurin, C., Andrés, C., Aubourg, S. *et al.* (2004) Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell*, **16**, 2089–2103.

Meierhoff, K., Felder, S., Nakamura, T., Bechtold, N. and Schuster, G. (2003) HCF152, an Arabidopsis RNA binding pentatricopeptide repeat protein involved in the processing of chloroplast *psbB-psbT-psbH-petB-petD* RNAs. *Plant Cell*, **15**, 1480–1495.

Meurer, J., Berger, A. and Westhoff, P. (1996a) A nuclear mutant of *Arabidopsis* with impaired stability on distinct transcripts of the plastid *psbB, psbD/C, ndhH,* and *ndhC* operons. *Plant Cell*, **8**, 1193–1207.

Meurer, J., Meierhoff, K. and Westhoff, P. (1996b) Isolation of high chlorophyll-fluorescence mutants of Arabidopsis thaliana and their characterisation by spectroscopy, immunoblotting and northern hybridisation. *Planta*, **198**, 385–396.

Mullet, J.E. (1988) Chloroplast development and gene expression. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **39**, 475–502.

O'Toole, N., Hattori, M., Andres, C., Iida, K., Lurin, C., Schmitz-Linneweber, C., Sugita, M. and Small, I. (2008) On the expansion of the pentatricopeptide repeat gene family in plants. *Mol. Biol. Evol.* **25**, 1120–1128.

Okuda, K., Myouga, F., Motohashi, R., Shinozaki, K. and Shikanai, T. (2007) Conserved domain structure of pentatricopeptide repeat proteins involved in chloroplast RNA editing. *Proc. Natl Acad. Sci. USA*, **104**, 8178–8183.

Pfalz, J., Liere, K., Kandlbinder, A., Dietz, K.J. and Oelmüller, R. (2006) pTAC2, -6, and -12 are components of the transcriptionally active plastid chromosome that are required for plastid gene expression. *Plant Cell*, **18**, 176–197.

Porra, R.J., Thompson, W.A. and Kriedemann, P.E. (1989) Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectrometry. *Biochim. Biophys. Acta* **975**, 384–394.

Rhyne, C.L. (1955) The inheritance of Yellow-Green, a possible mutation in cotton. *Genetics*, **40**, 235–245.

Rivals, E., Bruyère, C., Toffano-Nioche, C. and Lecharny, A. (2006) Formation of the Arabidopsis pentatricopeptide repeat family. *Plant Physiol.* **141**, 825–839.

Salone, V., Rüdinger, M., Polsakiewicz, M., Hoffmann, B., Groth-Malonek, M., Szurek, B., Small, I., Knoop, V. and Lurin, C. (2007) A hypothesis on the identification of the editing enzyme in plant organelles. *FEBS Lett.* **581**, 4132–4138.

Sato, S., Nakamura, Y., Kaneko, T., Asamizu, E. and Tabata, S. (1999) Complete structure of the chloroplast genome of *Arabidopsis thaliana*. *DNA Res.* **6**, 283–290.

Schägger, H., Cramer, W.A. and von Jagow, G. (1994) Analysis of molecular masses and oligomeric states of protein complexes by blue native electrophoresis and isolation of membrane protein complexes by two-dimensional native electrophoresis. *Anal. Biochem.* **217**, 220–230.

Schmitz-Linneweber, C., Williams-Carrier, R. and Barkan, A. (2005) RNA immunoprecipitation and microarray analysis show a chloroplast pentatricopeptide repeat protein to be associated with the 5′ region of mRNAs whose translation it activates. *Plant Cell*, **17**, 2791–2804.

Schmitz-Linneweber, C., Williams-Carrier, R.E., Williams-Voelker, P.M., Kroeger, T.S., Vichas, A. and Barkan, A. (2006) A pentatricopeptide repeat protein facilitates the *trans*-splicing of the maize chloroplast *rps12* pre-mRNA. *Plant Cell*, **18**, 2650–2663.

Schweer, J., Loschelder, H. and Link, G. (2006) A promoter switch that can rescue a plant sigma factor mutant. *FEBS Lett.* **580**, 6617–6622.

Severinov, K., Soushko, M., Goldfarb, A. and Nikiforov, V. (1993) Rifampicin region revisited. New rifampicin-resistant and streptolydigin-resistant mutants in the beta subunit of Escherichia coli RNA polymerase. *J. Biol. Chem.* **268**, 14820–14825.

Small, I.D. and Peeters, N. (2000) The PPR motif - A TPR-related motif prevalent in plant organellar proteins. *Trends Biochem. Sci.* **25**, 46–47.

Sugimoto, H., Kusumi, K., Noguchi, K., Yano, M., Yoshimura, A. and Iba, K. (2007) The rice nuclear gene, *VIRESCENT 2*, is essential for chloroplast development and encodes a novel type of guanylate kinase targeted to plastids and mitochondria. *Plant J.* **52**, 512–527.

Swiatecka-Hagenbruch, M., Emanuel, C., Hedtke, B., Liere, K. and Börner, T. (2008) Impaired function of the phage-type RNA polymerase RpoTp in transcription of chloroplast genes is

compensated by a second phage-type RNA polymerase. *Nucleic Acids Res.* **36**, 785–792.

Swiatek, M., Regel, R.E., Meurer, J., Wanner, G., Pakarasi, H.B., Ohad, I. and Herrmann, R.G. (2003) Effects of selective inactivation of individual genes for low-molecular-mass subunits on the assembly of photosystem II, as revealed by chloroplast transformation: the *psbEFLJ* operon in *Nicotiana tabacum*. *Mol. Genet. Genomics*, **268**, 699–710.

Tanaka, R. and Tanaka, A. (2007) Tetrapyrrole biosynthesis in higher plants. *Annu. Rev. Plant Biol.* **58**, 321–346.

Tavares-Carreón, F., Camacho-Villasana, Y., Zamudio-Ochoa, A., Shingú-Vázquez, M., Torres-Larios, A. and Pérez-Martínez, X. (2008) The pentatricopeptide repeats present in Pet309 are necessary for translation but not for stability of the mitochondrial *COX1* mRNA in yeast. *J. Biol. Chem.* **283**, 1472–1479.

4.3.2. PUBLICATION 5: PPR336 IS ASSOCIATED WITH RIBOSOMES IN PLANT MITOCHONDRIA. JOURNAL OF MOLECULAR BIOLOGY (2009)

doi:10.1016/j.jmb.2007.11.011 J. Mol. Biol. (2008) 375, 626–636

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

PPR336 is Associated with Polysomes in Plant Mitochondria

Magalie Uyttewaal[1]†, Hakim Mireau[1]†, Michal Rurek[2], Kamel Hammani[2], Nadège Arnal[1], Martine Quadrado[1] and Philippe Giegé[2]*

[1]*Station de Génétique et d'Amélioration des Plantes, INRA, Route de Saint-Cyr, 78026 Versailles Cedex, France*

[2]*Institut de Biologie Moléculaire des Plantes, CNRS, 12 rue du Général Zimmer, 67084 Strasbourg, France*

Received 11 July 2007; received in revised form 2 November 2007; accepted 5 November 2007
Available online 13 November 2007

The function of pentatricopeptide repeat (PPR) proteins has been associated with various post-transcriptional steps of organelle gene expression. Among them, translation and its regulation are essential processes. However, in plant mitochondria, they are also the steps of gene expression that are the least understood. In this study, PPR336 was identified as part of a high-molecular-weight complex in *Arabidopsis* mitochondria. PPR336 is an unusual representative of the large PPR family because it is relatively short and is characterised by a high expression level compared with other PPR proteins. PPR336 defines a small subgroup of eight class P PPR proteins that are similar in terms of motif organization. Among them, PPR336-like is the closest homolog of PPR336. Biochemical analysis has indicated that PPR336 is a strictly mitochondrial protein, extrinsically attached to the inner mitochondrial membrane and part of an RNase-sensitive complex. Sucrose gradients and polysome destabilisation experiments show that PPR336 is associated with ribosomes in plant mitochondria. Moreover, in *Ppr336/336-like* mutants, mitochondrial polysomes of lower molecular weight accumulate compared with wild-type plants. Polysome association and these unusual features suggest that PPR336 could be involved in a distinctive process, possibly translation in plant mitochondria.

Keywords: pentatricopeptide repeat proteins; ribosomes; translation; plant organelles; blue native

Edited by J. Karn

Introduction

Pentatricopeptide repeat (PPR) proteins are encoded by huge gene families in plant genomes (e.g., they are composed of 441 representatives in *Arabidopsis*).[1] The genes coding for PPR proteins are characteristic of eukaryotes because they are totally absent from prokaryotes. They are found in yeast and animals, albeit in very small numbers (i.e., on average, there are 80 times more PPR genes in plants).[2] The explosion of the number of PPR genes occurred during the evolution of terrestrial plants.[3] The discovery of PPR genes in the bryophyte *Physcomitrella patens* indicates that a large group of PPR genes already existed before the evolutionary separation of mosses and vascular plants.[4] As indicated by their name, PPR proteins are composed of 35 amino acid motifs repeated in tandem. These proteins have, on average, 12 PPR motifs. The sequences of these motifs are very degenerated, but their structure is likely to be conserved. Each PPR motif is thought to be configured as two antiparallel α-helices, and the tandem repeats of PPR motifs would form a superhelix enclosing a central groove involved in target binding. Moreover, PPR proteins have been divided into several categories based on the presence or absence of sequence variants of PPR motifs and various additional C-terminal domains.[1]

**Corresponding author.* E-mail address: Philippe.Giege@ibmp-ulp.u-strasbg.fr.

† M.U. and H.M. contributed equally to the work.

Present address: M. Rurek, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, 5 Miedzychodzka Str., 60-371 Poznan, Poland.

Abbreviations used: PPR, pentatricopeptide repeat; MALDI–TOF, matrix-assisted laser desorption/ionization time of flight; T-DNA, transferred DNA; ORF, open reading frame; MnSOD, manganese superoxide dismutase; rRNA, ribosomal RNA.

PPR proteins are putative RNA binding proteins, and the vast majority are predicted to be targeted to organelles.[1] The expansion of PPR proteins in plants could have evolved to respond to plant-specific processes involving RNA in organelles. Indeed, genetic expression in plant mitochondria and that in chloroplasts have many distinctive processes for, as examples, the transcription of their genes, the splicing of group II introns, RNA editing, the maturation of transcript ends and a poly A-dependent RNA degradation system.[5–7] Similarly, the translation apparatus is original in plant organelles,[8] and translation initiation and its control are poorly understood. Particularly, in mitochondria, hardly anything is known. No Shine–Dalgarno-like sequence is present in the 5′-untranslated region of transcripts, and some genes are translated without any AUG initiation codon or classical termination codon.[5,9] PPR proteins are predicted to be involved in all the plant-specific gene expression processes.

PPR proteins have been described as being able to bind RNA *in vitro* and *in vivo*,[1,10–14] although some of them were shown to bind DNA.[15] Genetic approaches have also identified PPR proteins as restorers of cytoplasmic male sterility (e.g., in petunia,[14] radish,[16] sorghum[17] and rice[18,19]). The genes *CRP1* and *PPR4* from maize and genes *HCF152*, *CRR2* and *PGR3* from *Arabidopsis* code for PPR proteins involved in the processing, stabilisation, splicing and translation of chloroplastic transcripts.[12,13,20,21] CRR4 and CCR21 are both E+-type PPR proteins essential for RNA editing in chloroplasts.[22,23] Interestingly, in *Trypanosoma brucei*, a number of PPR proteins were recently found to be essential for the biogenesis of mitochondrial ribosomes.[24]

Thus, studies on PPR proteins indeed connect these proteins with post-transcriptional processes in organelles. However, reports describing the biochemical properties and the mechanism of action of PPR proteins are still rare. Among them, HCF152 was described as forming a homodimer in chloroplasts.[10] PPR2 was found in a large heterogeneous protein complex in the stroma of chloroplasts.[25] Similarly, CRP1 was found in a multi-subunit complex but was not found associated with membranes and/or ribosomes.[20] In the mitochondria of petunia, a gel filtration approach identified PPR592 in fractions containing complexes of high molecular weight.[26]

Here we report a biochemical study on PPR336, an unusual representative of the PPR protein family. PPR336 defines a small subfamily of shorter PPR proteins. It is associated with polysomes on the inner membrane of *Arabidopsis* mitochondria and could thus play a role in mitochondrial translation.

Results

PPR336 is identified in a high-molecular-weight membrane complex

As part of a general survey of proteins present in membrane-associated complexes in *Arabidopsis* mitochondria, PPR336 (At1g61870) was found in a high-molecular-weight complex. For this analysis, the equivalent of 4 mg of *Arabidopsis* mitochondrial membrane proteins was treated with digitonin and the solubilised complexes obtained were separated on preparative blue native gels.[27,28] After migration, gel pieces covering the entire length of the blue native gels were excised, and the proteins they contained were electroeluted and concentrated. Proteins were identified by matrix-assisted laser desorption/ionization time-of-flight (MALDI–TOF) mass spectrometry. Numerous respiratory proteins were identified in gel bands corresponding to respiratory complexes IV, III, V as well as I and the supercomplex $I+III_2$ (data not shown), as already identified in previous studies.[29–31] The most interesting finding came from a gel band of 800 kDa, where no apparent band was visible on the stained blue native gels. In this band, PPR336 was identified (Fig. 1a). This protein is composed of conventional PPR repeats only and therefore belongs to the P subfamily of PPR proteins.[1] It is predicted to be localised in mitochondria by Predotar[32] and to be mitochondrial as well as chloroplastic by Mitoprot2[33] and TargetP.[34] PPR336 is an unusual representative of the PPR protein family because it is a relatively smaller protein that is composed of 408 amino acids as compared with the average 673 amino acids of the overall PPR family.[1] It also contains less PPR repeats, with 6 PPR repeats as compared with the average 12 PPR repeats observed for the proteins of the family. PPR336 is also characterised by its relatively high expression level, with, for example, 16 expressed sequence tags found for this gene as compared with the average 2 expressed sequence tags found for the other PPR genes.[1] A more precise investigation of microarray expression data, available from GENEVESTIGATOR‡, reveals that PPR336 is expressed at comparable levels throughout plant development (Supplementary Figure 1).[35] It is expressed at comparatively higher levels in inflorescences and roots, particularly in root tips and at lower levels in seedlings and rosettes (Supplementary Figure 2).

The subfamily of PPR336

Phylogenic analyses reveal that PPR336 defines a small subfamily of PPR proteins (Fig. 1b). In the global phylogenetic tree of class P PPR proteins,[1] a small subset of eight PPR proteins is composed of proteins resembling PPR336. They are all short (423 amino acids on average), have an average of six PPR repeats and share an unusually high percentage of sequence identity with PPR336 (average of 35%). Among them, At1g11630 is particularly similar to PPR336, with 70% of identical amino acids; we therefore named this protein PPR336-like. However, the eight proteins of the subfamily are not all expressed at high levels. GENEVESTIGATOR data indicate that expression levels vary 10-fold across the

‡ https://www.genevestigator.ethz.ch

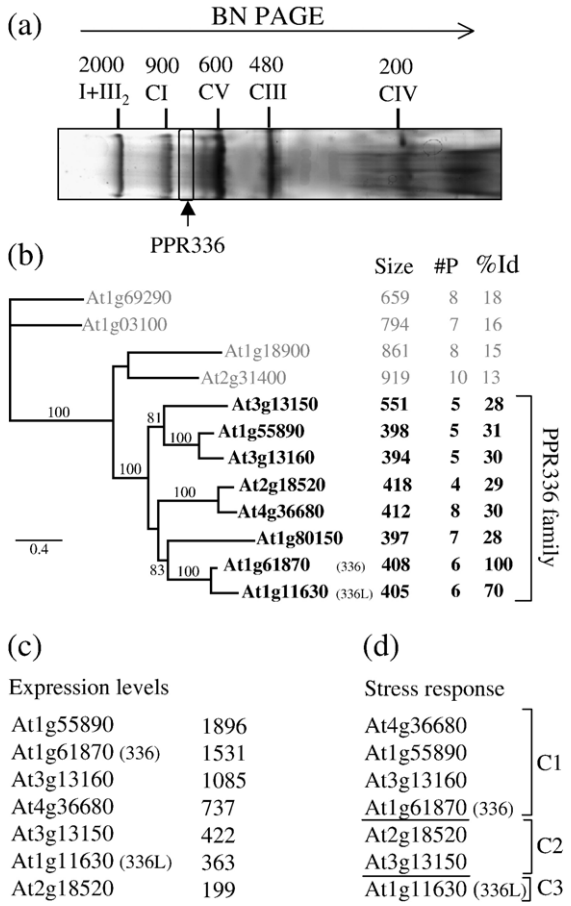


Fig. 1. (a) Blue native gel track with *Arabidopsis* mitochondrial membrane complexes. The size of the major bands corresponding to respiratory complexes and supercomplexes is given in kilodaltons. The box identifies the gel band excised from this track and pooled with bands from seven other identical gel tracks for elution. PPR336 was identified from the eluted proteins by MALDI–TOF mass spectrometry. Results were analysed by MASCOT. PPR336 was identified by a score of 152, with significant scores above 78 ($p<0.05$). (b) Unrooted maximum likelihood phylogenetic tree, part of the global tree for all the class P PPR proteins in *Arabidopsis*. It shows the eight proteins resembling PPR336 in bold. Bootstrap values above 80 (using 100 replications) are indicated along branches. The scale bar indicates the evolution distance (amino acid substitutions per site). The lists indicate the length of the proteins in amino acids (Size), the number of PPR repeats (#P) and the percentage of amino acid sequence identity with PPR336 (%Id). (c) PPR336 subfamily genes are ranked according to expression levels, exemplary during rosette growth here. Values are given for seven of eight genes for which information is available in GENEVESTIGATOR. They are Affymetrix MAS5-normalised signal intensities. (d) Result of a *K*-means clustering analysis of stress response data available in GENEVESTIGATOR. The genes of PPR336 family are sorted into three clusters (C1–C3) according to their stress response pattern.

family members (Fig. 1c). At1g55890, PPR336 and At3g13160 are the most highly expressed members of the subgroup, whereas PPR336-like is among the two least expressed genes (detailed results are shown in Supplementary Figures 1 and 2). The functional diversity or redundancy that might exist between the family members was first assessed by looking at their respective responses to stress. GENEVESTIGATOR stress response data were subjected to a *K*-means clustering analysis.[36] Genes were sorted into three clusters. PPR336, but not PPR336-like, was found in the same cluster as were At1g55890 and At3g13160. This could mean that they share similar functions (Fig. 1d). Yet, the results of this stress response analysis must be interpreted with care because the response ratios given by GENEVESTIGATOR were overall low for the genes of the PPR336 subfamily (detailed results are shown in Supplementary Figure 3). A similar small subfamily of PPR proteins is also found in the genome of rice, with six proteins resembling the ones from the *Arabidopsis* family. The conserved structure of the family members, together with the presence of a set of potential orthologs in rice, suggests that PPR336 and its family could have a distinctive function.

Characterisation of PPR336 mutants

In order to investigate the function of PPR336, we analysed transferred DNA (T-DNA) insertion mutants. Mutant lines were identified for *Ppr336* (N639562 and N037390) and *Ppr336-like* (CS808272), with insertions interrupting open reading frames (ORFs) in all cases. Locations of T-DNA insertions were verified for each line, and homozygous mutant plants could be obtained for both genes. The mutant plants were initially analysed by reverse transcription (RT)–PCR, which indicated that no transcript comprising the entire ORFs corresponding to *PPR336* and *PPR336-like* genes accumulates in the respective mutant plants (Supplementary Figure 4A). RT–PCR products corresponding to truncated transcripts could be amplified using primer pairs located upstream or downstream of T-DNA insertions (not shown). Likely initiated from promoters present in the T-DNA, such transcripts have no chance to code for functional PPR336 or PPR336-like proteins. Regarding the insertion in *PPR336-like*, these truncated transcripts would at most code a protein deprived of its mitochondrial targeting sequence as the insertion in this gene is located 40 bases downstream of the ATG codon. We then analysed the plants with antibodies raised against the full-length PPR336. Western blot analysis revealed a band of an apparent molecular mass of 45 kDa corresponding to PPR336 in an extract prepared from wild-type plants (Supplementary Figure 4B). This is in agreement with the calculated size of PPR336 (i.e., 45.7 kDa). The signal was absent in PPR336 mutant plants derived from the N037390 line, which indicated that they correspond to null mutants. In PPR336-like mutants, the 45-kDa signal was observed and was likely due to the expression of the functional *PPR336* gene. We expressed PPR336-like in *Escherichia coli* and found that despite

the high sequence similarity with PPR336, the PPR336 antibodies do not recognise PPR336-like (data not shown). Both mutant lines were analysed for macroscopic phenotypes. They do not significantly differ from the wild type regarding rosette size, flowering induction time, size of plants, number of siliques per plant, size of siliques and root length in all growth conditions tested (data not shown). Given that PPR proteins are predicted to be involved in gene expression processes in organelles, we had expected the mutations in these PPR genes to either be lethal or have serious deleterious effects. However, the two homozygous mutant plants did not show any apparent alteration in their phenotype compared with wild-type plants. It was therefore plausible that the function of one of the proteins could compensate for the function of the other. We crossed the mutant plants to obtain double mutants. We could obtain plants that were homozygous for the two insertions, and again that did not show any apparent deleterious phenotype. We deduce from this that either the functions of PPR336 and PPR336-like are not essential or their activity can be functionally complemented by additional PPR proteins, possibly of the small subfamily of PPR336.

PPR336 is an extrinsic protein of the mitochondrial inner membrane

The different subcellular localisation prediction programs did not give homogeneous results for the targeting of PPR336. We therefore analysed whether PPR336 was a strictly mitochondrial protein or dually targeted to mitochondria and chloroplasts. *Arabidopsis* cell extracts were separated into cytosolic, chloroplastic and mitochondrial fractions. Western blot analysis showed that PPR336-specific antibodies detected a 45-kDa signal in the mitochondrial fraction but not in the chloroplastic one (Fig. 2a). Then, the submitochondrial localisation of PPR336 was investigated. Mitochondria were fractionated into mitoplasts (i.e., matrix and inner membrane), a soluble fraction and a membrane fraction. Mitochondrial membrane proteins were further fractionated into extrinsic and intrinsic proteins by carbonate treatment. PPR336 was found in the mitoplast, membrane and extrinsic protein fractions (Fig. 2b). These results indicate that PPR336 is a peripheral protein of the mitochondrial inner membrane. The quality of the fractionations was controlled with antibodies directed against NAD9, an extrinsic mitochondrial membrane protein, against the large subunit of the chloroplast RuBisCo, against manganese superoxide dismutase (MnSOD), an enzyme of the mitochondrial matrix, and against cytochrome *c*1, an intrinsic mitochondrial inner membrane protein.

PPR336 can interact with RNA

The capacity of PPR336 to bind RNA was investigated by Northwestern experiments. We used the Northwestern method instead of band shift assays

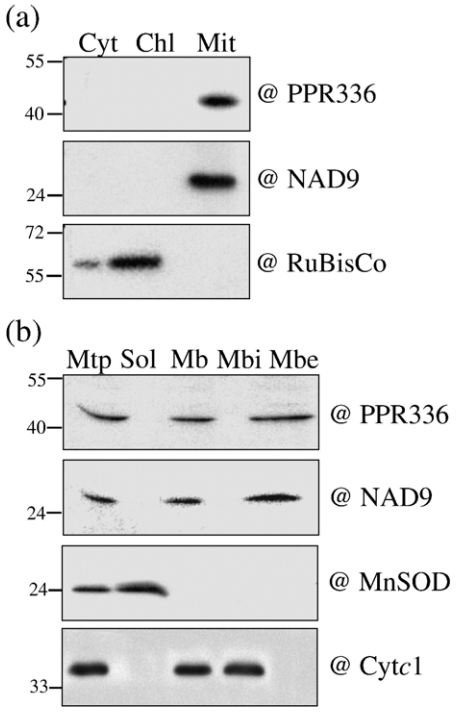


Fig. 2. (a) Immunodetection of PPR336 on subcellular protein fractions. Cyt indicates the cytosolic fraction; Chl, the chloroplastic fraction; and Mit, the mitochondrial fraction. (b) Immunodetection of PPR336 on submitochondrial protein fractions. Mtp shows the mitoplast fraction; Sol, the soluble fraction; Mb, the membrane fraction; Mbi, the intrinsic membrane protein fraction; and Mbe, the extrinsic membrane protein fraction. Fractionations are controlled with antibodies specific for the respiratory complex I subunit 9 (NAD9), the large subunit of RuBisCo, MnSOD and cytochrome *c*1 (Cyt *c*1). Molecular masses are given in kilodaltons.

because we were unable to express PPR336 as a soluble protein. Denatured recombinant proteins were separated by SDS-PAGE, transferred on blots and renatured directly on blots that were probed with radiolabelled RNA. Given the strict mitochondrial localisation of PPR336, we first probed the membranes with total mitochondrial RNA. We detected the 45-kDa signal corresponding to PPR336; thus, it proved that PPR336 is indeed able to bind RNA. We used purified green fluorescent protein as a negative control to show that the signal observed was not due to unspecific binding of RNA to renatured proteins. We also used total chloroplast RNA to probe the blots. Surprisingly, we observed a signal in these experiments. This suggested that PPR336 is able to bind RNA but shows no apparent specificity at least by this method (Supplementary Figure 5).

PPR336 is found in complexes containing RNA

The PPR336-specific antibodies were used to probe protein blots separated on one-dimensional blue native gels and two-dimensional blue native/SDS-PAGE gels where *Arabidopsis* mitochondrial

membrane complexes had been separated and their subunits resolved. For one-dimensional blue native gels, the antibodies detected a signal corresponding to complexes ranging from 500 to 1200 kDa (Fig. 3a). In two-dimensional gels, a similar signal was detected with an apparent molecular mass of 45 kDa for the second dimension, showing that the signal detected corresponded to a complex containing PPR336 (Fig. 3b). This result complemented the previous observation (i.e., PPR336 in a complex of 800 kDa) and suggests that this protein could be found in an array of complexes of different sizes. However, it is also possible that the smear-like signal detected here does not reflect the association with complexes of various sizes but rather the biochemical nature of the complex (i.e., its inability to be properly solubilised and/or resolved with the method used here). Smear-like signals are not an uncommon feature with blue native PAGE, as, for example, observed for the plant mitochondrial protein AtCCMH.[37] This kind of signals, typically not observed with highly structured respiratory complexes,[29] could reflect the transitory nature of interactions between the investigated proteins and their interactors. Finally, this analysis suggested that PPR336 is in any case not entirely embedded in the complexes detected because some epitopes of PPR336 were not masked in the native complexes separated here by blue native PAGE.

Since PPR336 was found to be able to bind RNA, we investigated whether RNA was as well present in the complexes detected here. For this, we treated solubilised complexes with RNase A before separation by blue native PAGE. This showed that the complexes containing PPR336 were indeed sensitive to RNase treatment (Fig. 3a). The signal detected previously was reduced to a 500- to 800-kDa signal

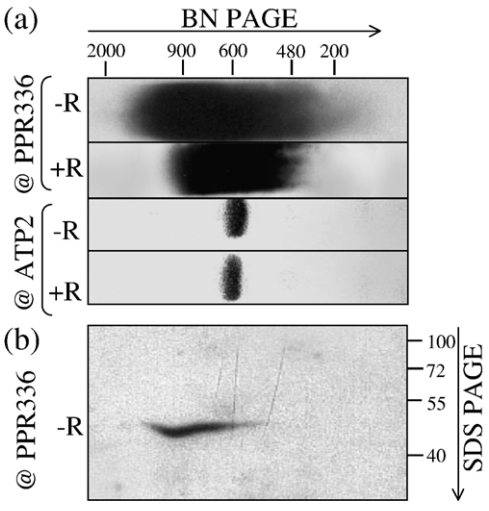


Fig. 3. One-dimensional blue native gel (a) and two-dimensional blue native/SDS-PAGE gels (b) of *Arabidopsis* mitochondrial membrane complexes blotted and reacted with PPR336 and ATP2 antibodies. Mitochondrial membrane complexes were treated with (+R) or without (−R) RNase before blue native gel electrophoresis. Molecular masses are given in kilodaltons.

after treatment with RNase. This could mean that some of the highest molecular-weight-complexes where PPR336 occurs are held together by RNA molecules and subsequently broken down after RNase treatment. This could also mean that the biochemical nature and/or structure of the complexes detected has been altered by RNase treatment, therefore inducing a change in the migration pattern of the complexes on blue native gels. As a control, the blots were probed with antibodies specific for ATP2, a subunit of the ATPase. It confirmed that RNase treatment did not affect complexes merely composed of proteins. Altogether, this analysis evidenced an involvement of RNA in the complexes where PPR336 was detected.

PPR336 is associated with mitochondrial polysomes

PPR336 was found by a blue native PAGE approach to be part of high-molecular-mass complexes (i.e., up to 1200 kDa) containing RNA. Technical limitations inherent in this gel-based approach do not authorize investigation of the possible presence of PPR336 in complexes of even much higher molecular weight, such as polysomes. For this reason, mitochondrial proteins and complexes were separated on sucrose gradients.[38] For this analysis, mitochondria were extracted from *Solanum tuberosum* and not from *Arabidopsis* for the former's suitability to obtain large amounts of pure mitochondria necessary for this biochemical approach (i.e., the equivalent of 100 mg of mitochondrial proteins). Prior to this, we controlled that PPR336 was also detected in the potato extracts with our antibody (data not shown). Mitochondrial extracts were separated on sucrose gradients, and fractions representing the entire gradients were collected. RNA was prepared for all the fractions, and equivalent amounts of proteins from each fraction were separated on SDS-PAGE, blotted and reacted with PPR336 antibodies. The 45-kDa signal was detected for fractions at the top and at the bottom of the gradients (Fig. 4). The same protein blots were reacted with antibodies specific for the ribosomal protein RPS1 and NAD9 from respiratory complex I. NAD9 was only detected in the top fractions of the gradients, suggesting that they contain complexes of sizes up to 2000 kDa. On the other hand, RPS1 and the 5S ribosomal RNA (rRNA) were detected in the same top and bottom fractions of the gradient as was PPR336. These top fractions were already proposed to contain ribosomes, whereas the bottom ones were proposed to contain polysomes.[39] This suggested that PPR336 could be associated with ribosomes and polysomes. Mitochondrial samples were subjected to various treatments, leading to the destabilisation of polysomes, in order to demonstrate this assumption. We treated samples with RNases, ethylenediaminetetraacetic acid (EDTA) and puromycin. Puromycin specifically destabilises ribosomes,[40] whereas EDTA and RNase result in ribosome destabilisation by either chelating

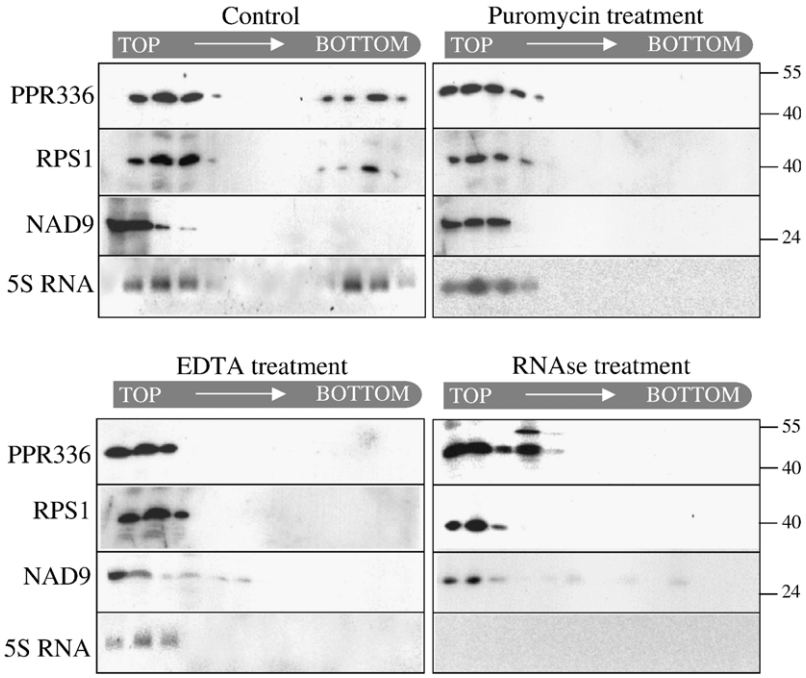


Fig. 4. Mitochondrial protein fractions were collected from sucrose gradients, separated by SDS-PAGE, blotted and reacted with antibodies specific for PPR336, NAD9 and the ribosomal protein RPS1. RNA was extracted from the fractions, and 5S rRNA was detected by Northern blot. Control sample was untreated, whereas other samples were treated to destabilise polysomes with puromycin or EDTA or RNases. Molecular masses are given in kilodaltons.

Mg^{2+} ions or degrading RNAs necessary for their integrity.[40,41] In all the cases, after treatments and separation on gradients, PPR336, similar to RPS1 and the 5S RNA, was not detected anymore in the fractions at the bottom of the gradients (Fig. 4). This confirms that PPR336 is indeed associated with polysomes in *Arabidopsis* mitochondria. However, the precise nature of this association with polysomes is unknown. It is possible that this association is indirect—that PPR336 is part of an RNA maturation complex that could be attached to polysomes. Alternatively, it is possible that PPR336 interacts directly with ribosomes or other components of the translation apparatus, thus having a direct role in mitochondrial translation or its control.

Mitochondrial polysomes in *Ppr336/336-like* mutants

The molecular phenotype of the *Ppr336/336-like* double mutants was first investigated by comparing steady-state levels of rRNAs in wild-type and mutant plants. Northern blots were analysed with probes specific for *Arabidopsis* 18S and 5S rRNAs. No significant difference in rRNA levels was observed. Similarly, no difference was observed for the exemplary chosen mRNAs nad9 and atp9 (Fig. 5a). Then, since PPR336 was found to be associated with polysomes, the sedimentation of polysomes was compared in wild-type and mutant plants. The occurrence of mitochondrial polysomes in sucrose gradient fractions was tested by Northern blots with a probe representing 5S RNA. In this analysis, 5S RNA was detected essentially in the same fractions (3–7) of the gradients for wild-type and mutant plants. However, in three independent experiments, the strongest signals were consistently detected in lower-molecular-weight fractions for the double mutant (i.e., in fractions 5 and 6 for the wild-type plant and fractions 3 and 4 for the double mutant) (Fig. 5b). This result suggests that polysomes with lower molecular weight are to some extent more abundant in the mutants as compared with the wild-type plants.

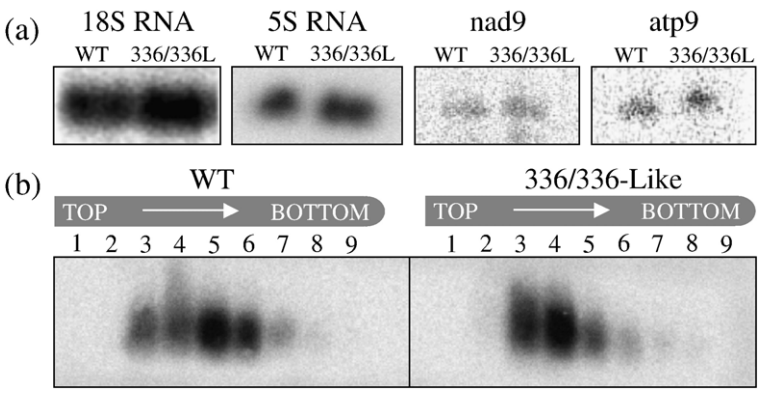


Fig. 5. (a) Expression levels of mitochondrial transcripts in wild-type and *Ppr336/336-like* double-mutant plants analysed by Northern blot with total RNA. Blots were hybridised with probes representing the 18S and 5S rRNAs as well as nad9 and atp9 mRNAs. (b) Extracts from wild-type and *Ppr336/336-like* mutant plants separated on sucrose gradients. Northern blots of RNA extracted from the fractions were probed with 5S RNA. Signals indicate the fractions where mitochondrial polysomes are found.

Discussion

In this study, PPR336 was first identified by blue native PAGE. Very few PPR proteins have been identified by other proteomic approaches,[42,43] in particular with the classical IEF/SDS-PAGE two-dimensional gels.[44] It is well known that hydrophobic membrane proteins are difficult to be identified from this type of gels (e.g., mitochondrial encoded proteins).[44] On the contrary, blue native gels have been used successfully to analyse membrane-bound respiratory complexes and supercomplexes (e.g., Ref. 31) and to identify new or plant-specific subunits.[30,45] The absence of PPR proteins in most previous proteomic studies had often been attributed to their low abundance. However, it is possible that this was rather due to the biochemical properties of PPR proteins. Indeed, after expression in *E. coli*, difficulties have been reported to obtain soluble PPR proteins (e.g., Ref. 12 and this study). Even if PPR proteins are predicted to be soluble proteins, many PPR proteins might be relatively insoluble outside their *in vivo* context (i.e., when they are not in complex with RNA and/or proteins, with some of these complexes attached to membranes).

PPR proteins targeted to mitochondria are all predicted to be involved in mitochondrial gene expression.[1] Mutations in the corresponding PPR genes were expected to either be lethal or have serious deleterious effects because mitochondrial genes are all directly or indirectly necessary for the biogenesis of the respiratory chain. Respiration is essential, and indeed mitochondrial mutations generally result in serious disorders, the consequence being, for example, male sterility, with cytoplasmic male sterile mutants[46] or deeply altered growth with non-chromosomal stripe mutants.[47] Similarly, mutants from genes involved in mitochondrial gene expression (e.g., PNPase) show very severe disorders.[48] In this light, it would be surprising that PPR genes would have non-essential functions. However, in this analysis, we observed that mutants for two similar PPR genes and the corresponding double mutants do not show any abnormal macroscopic phenotype. *PPR336* and *PPR336-like* are part of a small subfamily of eight PPR genes of conserved structure. It is then possible that, despite sequence divergence, other proteins of the subfamily can compensate for the function of PPR336, particularly At1g55890 and At3g13160, which share common expression features with PPR336. Similarly, in the previous global analysis of the PPR gene family,[1] 12 of 22 homozygous mutant lines analysed did not have a visible phenotype. Thus, it raises the question of the functional redundancy that exists in the huge family of PPR genes. It even questions the reason for the sheer high number of PPR genes in plants as compared with other eukaryotes. It is possible that this number is due not only to the high number of plant-specific RNA-processing events but also to a high functional redundancy of PPR genes in plants.

During mitochondrial gene expression, there is little control at the RNA level to obtain the stoichiometric amount of proteins necessary for the biogenesis of respiratory complexes.[7,49] A lot of the required regulation appears to be done at the protein level,[50] most likely during translation and its control. The polysome destabilisation experiments show that PPR336 is associated with polysomes. What is the functional reason for this association? As previously proposed, PPR proteins are most probably in complex with transcripts for their maturation or processing.[1] Thus, PPR336 could be part of a transcript maturation complex that would be recruited by ribosomes before the end of the maturation process (i.e., with PPR336 and other maturation *trans*-acting factors still attached to it). This would mean that RNA maturation and translation could be coupled in plant mitochondria. In favour of this hypothesis, partially edited and unspliced transcripts were found associated with polysomes.[51,52] As a second hypothesis, PPR336 could have a direct function in mitochondrial translation.

In chloroplasts, the function of PPR proteins has also been related to translation, with CRP1 described as an activator of translation that binds the 5′-untranslated region of two transcripts.[12] In plant mitochondria, little is known on the mechanism of translation and its control. One of its specificities is that polysomes are attached to membranes.[39] Contrary to chloroplasts, no Shine–Dalgarno-like sequence is found on mitochondrial mRNAs and some transcripts are translated without AUG initiation codons.[53] Similarly, translation can be terminated without a stop codon on mRNAs.[9] PPR proteins could be involved in the specific initiation and termination of translation in plant mitochondria.

In *T. brucei*, the function of several PPR proteins has recently been associated with the stability of mitochondrial rRNA.[24] Although the precise activity of these proteins for ribosomal function is unclear in *T. brucei*, it is possible that a similar activity could exist in plant mitochondria. In this study, in contrast to *T. brucei*, no decrease in the steady-state levels of rRNA was observed in *Ppr336* and *Ppr336-like* mutants. Again, it is possible that the functions of PPR336 and/or PPR336-like were compensated by the function of other proteins of the subfamily. Still, the analysis of mitochondrial polysome sedimentation revealed that lower-molecular-weight polysomes are more abundant in *Ppr336/336-like* mutant plants than in wild-type plants. The observation of this molecular phenotype suggests that even if the proteins of the PPR336 subfamily may share similar functions, these functions are not entirely redundant. Yet, the role of PPR336 remains elusive. It could be involved in RNA processing; for example, with it as an abundant protein bound to mRNA already attached to ribosomes, the discrepancy in polysome sedimentation in mutants could then be explained by the loss of PPR336 multimers from the transcript-processing complexes. Alternatively, PPR336 could be involved in a specific aspect of mitochondrial translation. Then, the difference in polysome sedimentation could be explained by a role for, as an example, the structure and/or the

stabilisation of polysomes, possibly together with other proteins. Thus, in mutants, polysomes would be at least partially destabilised. Further investigations of translation patterns in mutant plants and the determination of protein and RNA partners of PPR336, for example, will be necessary to understand the functional basis for the association of PPR336 with mitochondrial polysomes.

Materials and Methods

Cell suspension, mitochondrial isolation and fractionation

Arabidopsis thaliana var. *Landsberg erecta* suspension cultures were maintained in a Gamborg G0210 basal medium containing 1 mg/l of 2,4-D and 2% (w/v) sucrose (pH 5.8). One hundred milliliters of cell cultures was maintained in 250 ml of conical flasks in the dark at 22 °C on a rotary shaker at 150 rpm. Every 7 days, 10 ml of the culture was subcultured into 100 ml of fresh media. Five-day-old cultures were used for the preparation of subcellular fractions as described previously.[54] For mitochondrial extraction, cells were harvested by filtration through 100 μM nylon mesh and mitochondria extracted and purified by Percoll gradient centrifugation as described previously.[55] Mitochondria were fractionated into mitoplast, membrane and soluble fractions as described previously.[54] The membrane fraction was subjected to alkaline treatment (0.1 M Na_2CO_3, pH 11.5, for 30 min at 4 °C) to extract peripheral proteins.

Mitochondrial membrane complex solubilisation, RNase treatment and fractionation

The equivalent of 500 μg of mitochondrial membrane proteins was solubilised with digitonin, 5:1 detergent/protein (w/w) for 30 min on ice. For RNase sensitivity experiments, the solubilised complexes were treated with 5 μg of RNase A for 15 min at room temperature. Mitochondrial membrane complexes were resolved by blue native PAGE in the first dimension and/or by SDS-PAGE in the second dimension as described previously.[55] For protein identification, proteins contained in gel bands coming from eight one-dimensional blue native gels (i.e., corresponding to 4 mg of mitochondrial membrane complexes) were eluted in 25 mM *N*-[2-hydroxy-1,1-bis (hydroxymethyl)ethyl]glycine, 7.5 mM Bis–Tris and 1 mM PMSF overnight at 150 V and 4 °C with an ECU-040 electroeluter concentrator (C.B.S. Scientific, USA).

Protein identification

Proteins eluted from the blue native gels were subjected to reduction with DTT and alkylation with iodoacetamide in order to break disulphide bridges between cysteine residues. Proteins were digested with a trypsin solution (5–15 ng/μl in 25 mM Na bicarbonate) overnight at 37 °C. The resulting solution was analysed using an automated Biflex III MALDI–TOF mass spectrometer (Bruker, Germany). The list of masses obtained was used via the MASCOT software§ to identify hits in a translated NCBInr protein database. Matching was performed at ±50 ppm from the input masses, and hits were assessed by peptide number matching (6–15), coverage (typically, >25%) and the molecular weight search score. In all cases, the molecular weight search score for identified proteins gave a significant probability of identity ($P<0.05$).

§ http://www.matrixscience.com

Phylogenetic analyses

PPR gene sequences were aligned with the MUSCLE program (version 3.52).[56] The poorly aligned and too divergent positions were removed from alignment using the Gblocks 0.91b program.[57] Gblocks parameters were defined as follows: minimum number of sequences for a conserved position, 7; minimum number of sequences for a flanking position, 7; maximum number of contiguous non-conserved positions, 8; minimum length of a block, 5; and allowed gap positions, with half. The resulting alignment contained 17% of the original positions. A phylogenetic analysis was performed using Phyml[58] and 100-bootstrap support. The unrooted tree was drawn with TreeDyn.[59]

Antibody production and immunodetection

PPR336 ORF without the region coding the mitochondrial targeting presequence was amplified by PCR and cloned into pet14b plasmid (Novagen, USA). After expression in BL21pLysS *E. coli* strain (Novagen, USA), PPR336 appeared to be extremely insoluble and was found almost pure in the pellet obtained after solubilising bacterial proteins in 8 M urea. Ten milligrams of the protein was resuspended in 6 M guanidine hydrochloride, dialyzed against 5 l of phosphate-buffered saline buffer and used to immunize rabbits (Eurogentech, Belgium).

SDS-PAGE and blue native gels were transferred to polyvinylidene difluoride Immobilon-P membranes (Millipore). After blocking, membranes were incubated overnight with rabbit polyclonal antibodies at a dilution of 1/5000 for *Arabidopsis* PPR336, that of 1/100,000 for wheat NAD9,[60] that of 1/10,000 for spinach RuBisCo large subunit (obtained from B. Camara, Strasbourg, France), that of 1/10,000 for tobacco MnSOD,[61] that of 1/50,000 for yeast cytochrome *c*1 (obtained from G. Schatz, Basel, Switzerland), that of 1/10,000 for wheat RPS1, as well as that of 1/5000 for potato porin (obtained from H.P. Braun, Hannover, Germany) and with mouse monoclonal antibodies at a dilution of 1/1000 for maize ATP2 (obtained from T.E. Elthon, Lincoln, USA). Sheep anti-mouse and goat anti-rabbit antibodies conjugated with horseradish peroxidase (Amersham, UK) were used as secondary antibodies and visualised with enhanced chemiluminescent reagents (Amersham, UK).

Preparation of mitochondrial polysome-enriched fractions

Polysome-enriched fractions were prepared as described in Ref. 39, derived from the work of Maffey *et al.*[38] and Lu *et al.*[51] Succinctly, a lysate obtained from 100 mg of *S. tuberosum* mitochondria was centrifuged through a 61% sucrose cushion at 235,000*g* for 2 h. The polysome-containing pellet was further separated on 10–35% sucrose gradients centrifuged at 155,000*g* for 2 h. Fractions 1–17 (from top to bottom) were collected on the gradients and analysed on 10–17% SDS-PAGE. For Northern blot analysis, total RNA was extracted from 100 μl of each

fraction, separated on 15% polyacrylamide gel, electrotransferred onto Hybond-N nylon membrane (Amersham Biosciences) and hybridised to an oligonucleotide probe specific for potato mt 5S rRNA (CATATGGCGCAAGACGATTC). For polysome destabilisation experiments, lysates were treated with 10 mM puromycin, 50 mM EDTA or 600 U/ml of RNase A.

Preparation and analysis of fractions containing total polysomes

The procedure for total polysome preparation was derived from previous work.[62] Briefly, *Arabidopsis* wild-type and mutant plants were grown *in vitro* for 3 weeks. Entire plants were frozen and ground in liquid nitrogen. The equivalents of 500 μl of powder obtained were resuspended in polysome buffer containing 50 mM Tris–HCl, pH 8, 5 mM DTT, 25 mM $MgCl_2$, 100 mM KCl, 50 μg/ml of cycloheximide and 50 μg/ml of chloramphenicol. A total of 2% (v/v) Triton X-100 was added, and the extracts were vortexed for 1 min at room temperature and centrifuged at 3000*g* for 3 min at 4 °C. Supernatants were loaded on 3.5 ml of sucrose gradients composed of 0.75-ml 15%, 1-ml 30%, 1-ml 45% and 0.75-ml 60% sucrose (w/v) in polysome buffer. Gradients were centrifuged at 50,000*g* for 80 min at 4 °C. Fractions were collected, and total RNA was prepared and analysed by Northern blot according to standard methods. Blots were hybridised with probes already used in previous work,[49] representing *Arabidopsis* 18S and 5S rRNA, nad9 and atp9.

Mutant analysis

N639562 (*Ppr336*), N037390 (*Ppr336*) and CS808272 (*Ppr336-like*) mutant lines were obtained from the Signal mutant collection.[63] The presence of the T-DNA insertions in progenies was monitored by PCR using the 336.5 (AAGAAAACAACAATGGCGTTACTCTCT) and LBSALK1 (CATCAAACAGGATTTTCGCC) primers for N639562, the 336-3 (TCCGGTTCAGCGAGTTCGTCGTACT) and LBSALK1 primers for N037390 and the 336L1 (TGATACGGTCGACGTGA) and LB3SAIL (TAGCATCTGAATTTCATAACCAATCTCGATACAC) primers for CS808272. Wild-type alleles corresponding to each insertion were revealed with another PCR using primers surrounding T-DNA insertion sites. For this purpose, the primers 336.5 and 336.3 (TTCATGATACTAAAACTCGGAACCC) were employed for N639562, 336-5 and 336-3 (TTCATGATACTAAAACTCGGAACCC) for N037390 and 336L1 and 336L2 (GTTCCTCTTCCGCATTCG) for CS808272. *Ppr336* and *Ppr336-like* double mutants were generated by crossing simple homozygous mutants for each gene derived from N037390 and CS808272 lines. Double-homozygous mutant plants were identified by PCR analysis in the subsequent progenies of the crosses. RT–PCR was performed as described previously.[64] PCR amplifications aiming at revealing full-length mRNA originating from insertion loci employed the 336-5 and 336-3 oligonucleotides for *PPR336* and 336L1 and 336L2 for *PPR336-like*.

Acknowledgements

This project was supported by the "Centre National de la Recherche Scientifique" and the "Institut National de Recherche Agronomique." P.G. was supported by an "ANR Jeune Chercheur" research grant; K.H., by a PhD thesis grant from the Région Alsace; and M.R., by a foreign post-doc fellowship from the Foundation for Polish Science and EU ToK grant, MTKD-CT-2004-517068.

We thank V. Cognat and the "Plateforme Protéomique" (Esplanade, Strasbourg) for providing technical assistance. We also thank the reviewers for their constructive comments.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2007.11.011

References

1. Lurin, C., Andres, C., Aubourg, S., Bellaoui, M., Bitton, F., Bruyere, C. *et al.* (2004). Genome-wide analysis of *Arabidopsis* pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell*, **16**, 2089–2103.
2. Mootha, V. K., Lepage, P., Miller, K., Bunkenborg, J., Reich, M., Hjerrild, M. *et al.* (2003). Identification of a gene causing human cytochrome *c* oxidase deficiency by integrative genomics. *Proc. Natl. Acad. Sci. USA*, **100**, 605–610.
3. Aubourg, S., Boudet, N., Kreis, M. & Lecharny, A. (2000). In *Arabidopsis thaliana*, 1% of the genome codes for a novel protein family unique to plants. *Plant Mol. Biol.* **42**, 603–613.
4. Hattori, M., Hasebe, M. & Sugita, M. (2004). Identification and characterization of cDNAs encoding pentatricopeptide repeat proteins in the basal land plant, the moss *Physcomitrella patens*. *Gene*, **343**, 305–311.
5. Giegé, P. & Brennicke, A. (2001). From gene to protein in higher plant mitochondria. *C. R. Acad. Sci., Ser. III*, **324**, 209–217.
6. Giegé, P. & Brennicke, A. (1999). RNA editing in *Arabidopsis* mitochondria effects 441 C to U changes in ORFs. *Proc. Natl. Acad. Sci. USA*, **96**, 15324–15329.
7. Holec, S., Lange, H., Kuhn, K., Alioua, M., Borner, T. & Gagliardi, D. (2006). Relaxed transcription in *Arabidopsis* mitochondria is counterbalanced by RNA stability control mediated by polyadenylation and polynucleotide phosphorylase. *Mol. Cell. Biol.* **26**, 2869–2876.
8. Choquet, Y. & Wollman, F. A. (2002). Translational regulations as specific traits of chloroplast gene expression. *FEBS Lett.* **529**, 39–42.
9. Raczynska, K. D., Le Ret, M., Rurek, M., Bonnard, G., Augustyniak, H. & Gualberto, J. M. (2006). Plant mitochondrial genes can be expressed from mRNAs lacking stop codons. *FEBS Lett.* **580**, 5641–5646.
10. Nakamura, T., Meierhoff, K., Westhoff, P. & Schuster, G. (2003). RNA-binding properties of HCF152, an *Arabidopsis* PPR protein involved in the processing of chloroplast RNA. *Eur. J. Biochem.* **270**, 4070–4081.
11. Lahmy, S., Barneche, F., Derancourt, J., Filipowicz, W., Delseny, M. & Echeverria, M. (2000). A chloroplastic RNA-binding protein is a new member of the PPR family. *FEBS Lett.* **480**, 255–260.

12. Schmitz-Linneweber, C., Williams-Carrier, R. & Barkan, A. (2005). RNA immunoprecipitation and microarray analysis show a chloroplast pentatricopeptide repeat protein to be associated with the 5′ region of mRNAs whose translation it activates. *Plant Cell*, **17**, 2791–2804.
13. Schmitz-Linneweber, C., Williams-Carrier, R. E., Williams-Voelker, P. M., Kroeger, T. S., Vichas, A. & Barkan, A. (2006). A pentatricopeptide repeat protein facilitates the *trans*-splicing of the maize chloroplast rps12 pre-mRNA. *Plant Cell*, **18**, 2650–2663.
14. Bentolila, S., Alfonso, A. A. & Hanson, M. R. (2002). A pentatricopeptide repeat-containing gene restores fertility to cytoplasmic male-sterile plants. *Proc. Natl. Acad. Sci. USA*, **99**, 10887–10892.
15. Ikeda, T. M. & Gray, M. W. (1999). Characterization of a DNA-binding protein implicated in transcription in wheat mitochondria. *Mol. Cell. Biol.* **19**, 8113–8122.
16. Desloire, S., Gherbi, H., Laloui, W., Marhadour, S., Clouet, V., Cattolico, L. *et al.* (2003). Identification of the fertility restoration locus, Rfo, in radish, as a member of the pentatricopeptide-repeat protein family. *EMBO Rep.* **4**, 588–594.
17. Klein, R. R., Klein, P. E., Mullet, J. E., Minx, P., Rooney, W. L. & Schertz, K. F. (2005). Fertility restorer locus Rf1 of sorghum (*Sorghum bicolor* L.) encodes a pentatricopeptide repeat protein not present in the colinear region of rice chromosome 12. *Theor. Appl. Genet.* **111**, 994–1012.
18. Komori, T., Ohta, S., Murai, N., Takakura, Y., Kuraya, Y., Suzuki, S. *et al.* (2004). Map-based cloning of a fertility restorer gene, Rf-1, in rice (*Oryza sativa* L.). *Plant J.* **37**, 315–325.
19. Wang, Z., Zou, Y., Li, X., Zhang, Q., Chen, L., Wu, H. *et al.* (2006). Cytoplasmic male sterility of rice with boro II cytoplasm is caused by a cytotoxic peptide and is restored by two related PPR motif genes via distinct modes of mRNA silencing. *Plant Cell*, **18**, 676–687.
20. Fisk, D. G., Walker, M. B. & Barkan, A. (1999). Molecular cloning of the maize gene crp1 reveals similarity between regulators of mitochondrial and chloroplast gene expression. *EMBO J.* **18**, 2621–2630.
21. Nakamura, T., Schuster, G., Sugiura, M. & Sugita, M. (2004). Chloroplast RNA-binding and pentatricopeptide repeat proteins. *Biochem. Soc. Trans.* **32**, 571–574.
22. Kotera, E., Tasaka, M. & Shikanai, T. (2005). A pentatricopeptide repeat protein is essential for RNA editing in chloroplasts. *Nature*, **433**, 326–330.
23. Okuda, K., Myouga, F., Motohashi, R., Shinozaki, K. & Shikanai, T. (2007). Conserved domain structure of pentatricopeptide repeat proteins involved in chloroplast RNA editing. *Proc. Natl. Acad. Sci. USA*, **104**, 8178–8183.
24. Pusnik, M., Small, I., Read, L. K., Fabbro, T. & Schneider, A. (2007). Pentatricopeptide repeat proteins in *Trypanosoma brucei* function in mitochondrial ribosomes. *Mol. Cell. Biol.* **27**, 6876–6888.
25. Williams, P. M. & Barkan, A. (2003). A chloroplast-localized PPR protein required for plastid ribosome accumulation. *Plant J.* **36**, 675–686.
26. Gillman, J. D., Bentolila, S. & Hanson, M. R. (2007). The petunia restorer of fertility protein is part of a large mitochondrial complex that interacts with transcripts of the CMS-associated locus. *Plant J.* **49**, 217–227.
27. Schägger, H. & von Jagow, G. (1991). Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Anal. Biochem.* **199**, 223–231.
28. Schägger, H., Cramer, W. A. & von Jagow, G. (1994). Analysis of molecular masses and oligomeric states of protein complexes by blue native electrophoresis and isolation of membrane protein complexes by two-dimensional native electrophoresis. *Anal. Biochem.* **217**, 220–230.
29. Jansch, L., Kruft, V., Schmitz, U. K. & Braun, H. P. (1996). New insights into the composition, molecular mass and stoichiometry of the protein complexes of plant mitochondria. *Plant J.* **9**, 357–368.
30. Eubel, H., Jansch, L. & Braun, H. P. (2003). New insights into the respiratory chain of plant mitochondria. Supercomplexes and a unique composition of complex II. *Plant Physiol.* **133**, 274–286.
31. Eubel, H., Heinemeyer, J. & Braun, H. P. (2004). Identification and characterization of respirasomes in potato mitochondria. *Plant Physiol.* **134**, 1450–1459.
32. Small, I., Peeters, N., Legeai, F. & Lurin, C. (2004). Predotar: a tool for rapidly screening proteomes for N-terminal targeting sequences. *Proteomics*, **4**, 1581–1590.
33. Claros, M. G. & Vincens, P. (1996). Computational method to predict mitochondrially imported proteins and their targeting sequences. *Eur. J. Biochem.* **241**, 779–786.
34. Emanuelsson, O., Nielsen, H., Brunak, S. & von Heijne, G. (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol.* **300**, 1005–1016.
35. Zimmermann, P., Hirsch-Hoffmann, M., Hennig, L. & Gruissem, W. (2004). GENEVESTIGATOR. *Arabidopsis* microarray database and analysis toolbox. *Plant Physiol.* **136**, 2621–2632.
36. Soukas, A., Cohen, P., Socci, N. D. & Friedman, J. M. (2000). Leptin-specific patterns of gene expression in white adipose tissue. *Genes Dev.* **14**, 963–980.
37. Meyer, E. H., Giegé, P., Gelhaye, E., Rayapuram, N., Ahuja, U., Thony-Meyer, L. *et al.* (2005). AtCCMH, an essential component of the *c*-type cytochrome maturation pathway in *Arabidopsis* mitochondria, interacts with apocytochrome *c*. *Proc. Natl. Acad. Sci. USA*, **102**, 16113–16118.
38. Maffey, L., Degand, H. & Boutry, M. (1997). Partial purification of mitochondrial ribosomes from broad bean and identification of proteins encoded by the mitochondrial genome. *Mol. Gen. Genet.* **254**, 365–371.
39. Delage, L., Giegé, P., Sakamoto, M. & Marechal-Drouard, L. (2007). Four paralogues of RPL12 are differentially associated to ribosome in plant mitochondria. *Biochimie*, **89**, 658–668.
40. Lu, M. & Draper, D. E. (1994). Bases defining an ammonium and magnesium ion-dependent tertiary structure within the large subunit ribosomal RNA. *J. Mol. Biol.* **244**, 572–585.
41. Ott, M., Prestele, M., Bauerschmitt, H., Funes, S., Bonnefoy, N. & Herrmann, J. M. (2006). Mba1, a membrane-associated ribosome receptor in mitochondria. *EMBO J.* **25**, 1603–1610.
42. Heazlewood, J. L., Howell, K. A., Whelan, J. & Millar, A. H. (2003). Towards an analysis of the rice mitochondrial proteome. *Plant Physiol.* **132**, 230–242.
43. Heazlewood, J. L., Tonti-Filippini, J. S., Gout, A. M., Day, D. A., Whelan, J. & Millar, A. H. (2004). Experimental analysis of the *Arabidopsis* mitochondrial proteome highlights signaling and regulatory components, provides assessment of targeting prediction programs, and indicates plant-specific mitochondrial proteins. *Plant Cell*, **16**, 241–256.

44. Millar, A. H., Sweetlove, L. J., Giegé, P. & Leaver, C. J. (2001). Analysis of the *Arabidopsis* mitochondrial proteome. *Plant Physiol.* **127**, 1711–1727.
45. Millar, A. H., Eubel, H., Jansch, L., Kruft, V., Heazlewood, J. L. & Braun, H. P. (2004). Mitochondrial cytochrome *c* oxidase and succinate dehydrogenase complexes contain plant specific subunits. *Plant Mol. Biol.* **56**, 77–90.
46. Budar, F., Touzet, P. & De Paepe, R. (2003). The nucleo-mitochondrial conflict in cytoplasmic male sterilities revisited. *Genetica*, **117**, 3–16.
47. Newton, K. J., Mariano, J. M., Gibson, C. M., Kuzmin, E. & Gabay-Laughnan, S. (1996). Involvement of S2 episomal sequences in the generation of NCS4 deletion mutation in maize mitochondria. *Dev. Genet.* **19**, 277–286.
48. Perrin, R., Meyer, E. H., Zaepfel, M., Kim, Y. J., Mache, R., Grienenberger, J. M. *et al.* (2004). Two exoribonucleases act sequentially to process mature 3′-ends of atp9 mRNAs in *Arabidopsis* mitochondria. *J. Biol. Chem.* **279**, 25440–25446.
49. Giegé, P., Hoffmann, M., Binder, S. & Brennicke, A. (2000). RNA degradation buffers asymmetries of transcription in *Arabidopsis* mitochondria. *EMBO Rep.* **1**, 164–170.
50. Giegé, P., Sweetlove, L. J., Cognat, V. & Leaver, C. J. (2005). Coordination of nuclear and mitochondrial genome expression during mitochondrial biogenesis in *Arabidopsis*. *Plant Cell*, **17**, 1497–1512.
51. Lu, B. & Hanson, M. R. (1996). Fully edited and partially edited nad9 transcripts differ in size and both are associated with polysomes in potato mitochondria. *Nucleic Acids Res.* **24**, 1369–1374.
52. Yang, A. J. & Mulligan, R. M. (1993). Distribution of maize mitochondrial transcripts in polysomal RNA: evidence for non-selectivity in recruitment of messenger RNAs. *Curr. Genet.* **23**, 532–536.
53. Unseld, M., Marienfeld, J. R., Brandt, P. & Brennicke, A. (1997). The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. *Nat. Genet.* **15**, 57–61.
54. Spielewoy, N., Schulz, H., Grienenberger, J. M., Thony-Meyer, L. & Bonnard, G. (2001). CCME, a nuclear-encoded heme-binding protein involved in cytochrome *c* maturation in plant mitochondria. *J. Biol. Chem.* **276**, 5491–5497.
55. Giegé, P., Sweetlove, L. & Leaver, C. (2003). Identification of mitochondrial protein complexes in *Arabidopsis* using two-dimensional blue-native polyacrylamide gel electrophoresis. *Plant Mol. Biol. Rep.* **21**, 133–144.
56. Edgar, R. C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, **5**, 113.
57. Talavera, G. & Castresana, J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* **56**, 564–577.
58. Guindon, S. & Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**, 696–704.
59. Chevenet, F., Brun, C., Banuls, A. L., Jacq, B. & Christen, R. (2006). TreeDyn: towards dynamic graphics and annotations for analyses of trees. *BMC Bioinformatics*, **7**, 439.
60. Lamattina, L., Gonzalez, D., Gualberto, J. & Grienenberger, J. M. (1993). Higher plant mitochondria encode an homologue of the nuclear-encoded 30-kDa subunit of bovine mitochondrial complex I. *Eur. J. Biochem.* **217**, 831–838.
61. Bowler, C., Alliotte, T., Van den Bulcke, M., Bauw, G., Vandekerckhove, J., Van Montagu, M. & Inze, D. (1989). A plant manganese superoxide dismutase is efficiently imported and correctly processed by yeast mitochondria. *Proc. Natl. Acad. Sci. USA*, **86**, 3237–3241.
62. Slaymaker, D. H. & Hoppey, C. M. (2006). Reduced polysome levels and preferential recruitment of a defense gene transcript into polysomes in soybean cells treated with the syringolide elicitor. *Plant Sci.* **170**, 54–60.
63. Alonso, J. M., Stepanova, A. N., Leisse, T. J., Kim, C. J., Chen, H., Shinn, P. *et al.* (2003). Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science*, **301**, 653–657.
64. Arnal, N., Alban, C., Quadrado, M., Grandjean, O. & Mireau, H. (2006). The *Arabidopsis* Bio2 protein requires mitochondrial targeting for activity. *Plant Mol. Biol.* **62**, 471–479.

5. DISCUSSION

Lors de ma thèse, j'ai réalisé des études fonctionnelles de protéines PPR chez *Arabidopsis thaliana*. Pour cela, nous avons utilisé une stratégie de génétique inverse en sélectionnant des gènes *PPR* de fonctions inconnues pour lesquels des mutants d'insertion étaient disponibles dans des banques publiques. Cette stratégie s'est avérée fructueuse; nous avons avec succès pu relier la fonction de plusieurs de ces gènes à un mécanisme moléculaire précis ou hypothétique au sein des organites d'*Arabidopsis*. Ces travaux ont permis une avancée considérable dans l'élucidation fonctionnelle de cette vaste famille de gènes (comptant plus de 400 membres chez les plantes supérieures); 6 gènes *PPR* (*OTP80, OTP81, OTP82, OTP84, OTP85, OTP86*) ont pu être associé aux mécanismes d'édition de l'ARN chloroplastique et le gène *PPR9* a été associé à la régulation potentielle de l'expression des génomes nucléaire et mitochondrial.

5.1. NECESSITE DE L'EDITION DE L'ARN DANS LES ORGANITES

L'édition de l'ARN dans les organites des plantes est un mécanisme très étudié dans le domaine de la biologie végétale. Ce mécanisme est apparu comme essentiel à l'expression des gènes des organites des plantes terrestres. En effet, l'édition de l'ARN entraîne un changement de la nature du résidu codé par le gène et corrigerait les mutations des génomes des organites apparues au cours de l'évolution.

De nombreuses études génétiques de mutants d'*Arabidopsis* affectés dans l'édition de transcrits chloroplastiques ou mitochondriaux ont montré l'importance de l'édition pour la fonction des protéines codées (*clb19, lpa66, ys1, crr4, crr21, crr22, crr28, ogr1*).

Toutefois, les études biochimiques effectuées sur certains des mutants *Arabidopsis* présentant un ou plusieurs défauts d'édition chloroplastique (*otp82* et *otp85*), ont montré qu'une mutation ponctuelle au sein de certaines protéines chloroplastiques engendrée par un défaut d'édition ne conduit pas forcément à un dysfonctionnement de la protéine dans des conditions de culture optimales des plantes.

L'absence de défaut fonctionnel des protéines résultant de la traduction de certains transcrits non édités suggère que le mécanisme d'édition n'est pas toujours essentiel à l'expression correcte de certains gènes chloroplastiques. Ceci pose la question de la conservation de ces sites d'édition au cours de l'évolution. En effet, si un site d'édition n'est pas essentiel, on peut penser que celui-ci pourrait être fixé au niveau de la séquence génomique par une mutation reverse qui éviterait à l'organite une dépense énergétique inutile en éditant le site. La conservation de l'édition de ces transcrits suggérerait que le mécanisme d'édition ait un rôle autre que la correction des mutations. Les gènes nucléaires d'origines mitochondriale et chloroplastique ne nécessitent pas d'étape d'édition pour leur expression. L'édition de certains transcrits des organites pourrait être un moyen de pression exercé par l'organite dans le but de conserver ses gènes en empêchant leur transfert vers le noyau.

Dans le cas des plantes mutantes défectives dans l'édition des protéines ribosomales chloroplastiques (*otp80, otp86*), il est difficile d'évaluer l'impact du défaut d'édition sur les fonctions biologiques des protéines ribosomales Rps14 et Rpl23. En effet, nous avons observé que des défauts d'édition des transcrits codant ces protéines n'engendrent pas un dysfonctionnement de la traduction chloroplastique chez les plantes mutantes qui est généralement caractérisé par un phénotype sévère chez *Arabidopsis*. Toutefois, même si la traduction chloroplastique ne semble pas affectée chez les mutants *otp80* et *otp86*, nous ne pouvons exclure le fait que les protéines Rps14 et Rpl23 ne soient pas fonctionnelles. En effet, un mutant de tabac "knock out" pour le gène chloroplastique codant la protéine ribosomale Rpl33 n'est pas affecté dans la traduction chloroplastique dans des conditions de culture optimale (Rogalski et *al*., 2008). La traduction chloroplastique de la plante mutante n'est affectée que dans des conditions de culture à basses températures. Ces résultats indiquent que dans des conditions de culture optimale des plantes, certaines protéines ribosomales ne sont pas essentielles à la fonction du ribosome.

Il est possible que la nécessité de l'édition de certains transcrits ne puisse être révélée que sous certaines conditions de culture des plantes pour lesquelles la fonction des protéines codées est essentielle.

Ainsi, il serait intéressant de tester différentes conditions de culture induisant un stress chez les différents mutants ne présentant pas de phénotype macroscopique et moléculaire (*otp80, otp81, otp82, otp85* et *otp86*).

Le rôle exact du mécanisme complexe d'édition et la question de sa conservation au sein des organites végétaux restent toujours des énigmes.

5.2. SPECIFICITE DE RECONNAISSANCE MULTIPLE D'UN *TRANS*-FACTEUR PPR

Une analyse bioinformatique des éléments *cis* (15 nucléotides en amont du site d'édition) reconnus par un même *trans*-facteur PPR a montré que la spécificité de reconnaissance des éléments *cis* au sein du transcriptome chloroplastique peut être expliquée si la protéine reconnait les nucléotides conservées et distingue également les résidus pyrimidines et purines.

Les résultats de cette analyse bioinformatique reposent sur le principe que les protéines PPR *trans*-facteurs lient l'ARN par extrapolation des résultats *in vitro* obtenus pour les facteurs d'édition CRR4 et PpPPR_71 et sur le principe que le motif PPR est un domaine putatif de liaison à l'ARN. Toutefois, si ces principes sont vrais pour des *trans*-facteurs liant un seul site d'édition tels que CRR4 et PpPPR_71, les liaisons multiples d'un *trans*-facteur PPR à différents éléments *cis* restent à démontrer *in vivo* (par exemple par co-immunoprécipitation de l'ARN) ou *in vitro* (expériences de gel retard).

Dans notre étude, nous nous sommes appuyés sur la comparaison de certains facteurs PPR du chloroplaste d'*Arabidopsis* (OTP84 et YS1) à des facteurs d'édition chloroplastiques identifiés chez le tabac (qui sont vraisemblablement les orthologues de ces facteurs PPR)

pour lesquels des expériences biochimiques ont montré leur liaison directe aux éléments *cis* des sites d'édition qu'ils reconnaissent (ces sites d'édition étant les équivalents de ceux reconnus par OTP84 et YS1 chez *Arabidopsis*).

Il serait intéressant et nécessaire de confirmer notre hypothèse de reconnaissance multiple des facteurs d'édition PPR par des expériences *in vitro* de gel retard combinées à de la mutagénèse dirigée sur les ligands ARN. L'étape limitant ces approches biochimiques et l'obtention de protéines PPR recombinantes à l'état soluble.

Même si nos résultats restent préliminaires, l'hypothèse propose un modèle de reconnaissance de l'ARN par les protéines PPR de la sous-famille PLS intéressant à tester biochimiquement.

5.3. LES PROTEINES PPR, FACTEURS PUTATIFS D'EDITION DANS LES ORGANITES

Au total, les transcriptomes des organites d'*Arabidopsis* présentent environ 500 sites d'édition (environ 450 et 34 sites dans le transcriptome mitochondrial et chloroplastique, respectivement). Ces sites nécessitent d'être reconnus par la machinerie d'édition de manière spécifique. Pour cela, des groupes de sites d'édition partagent une information commune au sein de leurs élements *cis* qui seraient reconnue par des *trans*-facteurs d'origine protéique. Durant les dernières années, les membres de la famille des protéines PPR se sont imposés comme des acteurs majeurs de cette reconnaissance.

Jusqu'à ce jour, les études publiées ont montré que dans les chloroplastes, une même protéine PPR (OTP84, CRR22) peut jouer le rôle de *trans*-facteur au niveau de 3 sites maximum chez *Arabidopsis*. Dans les mitochondries de plantes, des études ont démontré l'implication d'une même protéine PPR dans l'édition d'au moins 3 sites chez *Arabidopsis* (MEF11) et jusqu'à 7 sites chez le riz (OGR1). Si l'on considère que les *trans*-facteurs putatifs de chaque site d'édition seraient des protéines PPR et que ceux-ci puissent être impliqués dans l'édition de plusieurs sites, les organites d'*Arabidopsis* nécessiteraient environ 150 protéines PPR adressés aux organites. Nous avons vu, que les protéines PPR impliquées dans l'édition appartiennent principalement aux sous-groupes E, E+ et DYW. Les 194 membres que composent ces sous-groupes chez *Arabidopsis*, pourraient donc assurer les fonctions de *trans*-facteur de l'ensemble des sites d'édition dans les organites.

Il est également important de noter que contrairement aux facteurs d'édition de la famille des protéines RRM dont l'absence chez *Arabidopsis* entraîne des défauts d'édition partiels de transcrits chloroplastiques, l'absence des facteurs d'édition PPR entraîne des défauts d'édition totaux. Ces résultats indiquent l'indispensabilité de ces facteurs PPR au sein de la machinerie d'édition contrairement aux facteurs RRM qui peuvent être qualifiés d'accessoires.

5.4. L'ADRESSAGE DE PROTEINES PPR EN DEHORS DES ORGANITES

Durant ces travaux de thèse, nous avons identifié une protéine PPR, PPR9 présentant deux signaux d'adressage chez *Arabidopsis*. Ces deux signaux d'adressage prédisent pour la protéine une double localisation dans les mitochondries et également en dehors de ceux-ci dans le noyau des cellules. Ceci a pu être confirmé par différentes techniques de biologie cellulaire et de biochimie.

Il est intéressant de noter que les protéines détectées dans les mitochondries et les noyaux par immunodétection ont la même masse apparente après migration sur gel SDS-PAGE. Cette masse correspond à la masse prédite de la protéine mature, libérée de sa préséquence d'adressage mitochondrial. Ceci suggère que l'adressage de la protéine dans le noyau serait postérieur à son import dans la mitochondrie ou qu'un événement alternatif de traduction ou maturation du transcrit conduise à la synthèse de deux protéines de même masse apparente. Deux possibilités peuvent être proposées:

- une initiation interne de la traduction, au niveau du site prédit de clivage de la préséquence mitochondriale. L'initiation de la traduction chez les plantes se réalise généralement au niveau d'un codon AUG codant une méthionine mais peut également se faire au niveau du codon alternatif CTG codant une leucine (Christensen et *al.*, 2005). L'analyse de la séquence protéique de PPR9 (prédite à partir de la séquence génomique) ne révèle aucune méthionine ou leucine internes pouvant conduire à l'obtention d'une protéine de masse équivalente à celle de la protéine maturée (Figure 2, manuscrit en préparation).

- un épissage différentiel au niveau de l'extrémité 5' du transcrit *PPR9* conduisant à l'obtention de deux transcrits aux extrémités 5' possédant des sites d'initiation de la traduction différents. Ce phénomène peut introduire un site AUG initiateur non détectable par analyse de la séquence génomique. La cartographie de l'extrémité 5' du transcrit *PPR9* par 5'-RACE (pour Rapid Amplification of 5' Complementary DNA Ends) n'a pas révélé d'épissage différentiel.

Au vu de ces résultats, il est possible que les formes nucléaire et mitochondriale de PPR9 soient le même produit de traduction.

Chez l'orge, la protéine Whirly1 à une localisation nucléaire et chloroplastique et des expériences immunologiques ont également démontré la même masse apparente de la protéine dans les deux compartiments suggérant également un produit de traduction unique (Grabowski et *al*., 2008). Dans le cas de Whirly1 et PPR9, la protéine des organites pourrait être relocalisée au noyau.

Dans les levures, la fumarase est une enzyme doublement localisée dans le cytosol et les mitochondries. Les protéines des deux compartiments dérivent du même précurseur maturé par une peptidase mitochondriale (Sass et *al.*, 2001). Dans le cas de l'adressage cytosolique, la protéine n'est pas importée dans la matrice mitochondriale, mais son repliement permet son largage dans le cytosol après clivage de la préséquence mitochondriale (Sass et *al*., 2003).

Le même scénario peut être imaginé pour PPR9 en plus duquel le largage cytosolique de la protéine serait suivi de son adressage au noyau des cellules.
La possibilité de translocation de protéines des organites vers le noyau des cellules de plantes nécessite des preuves expérimentales. Ceci pourrait être achevé par des expériences de transformation stable du génome chloroplastique à l'aide de constructions codant une protéine d'intérêt fusionnée à une étiquette. La relocalisation de la protéine synthétisée dans le chloroplaste pourrait être suivie par des analyses d'immunodétection.

Les protéines comme PPR9 avec une double localisation dans le noyau et un organite sont des candidats idéaux à la transduction de signaux impliqués dans la coordination de l'expression des génomes entre les deux compartiments.

Il serait intéressant de réaliser une étude bioinformatique systématique sur les protéines PPR afin de mettre en évidence d'autres membres présentant des séquences putatives d'adressage au noyau et aux organites, de vérifier leur localisation et d'entreprendre leur étude fonctionnelle.

5.5. INTERACTIONS PROTEIQUES DES PROTEINES PPR

Au cours de l'étude fonctionnelle de PPR9, nous avons mis en évidence l'interaction de la protéine avec deux partenaires de fonction connue dans le noyau: NAP1 et TCP8. Cette interaction pourrait impliqué certains motifs PPR majoritaires dans la séquence protéique ou les domaines terminaux qui comportent 115 acides aminés à l'extrémité N-terminale dépourvue de la séquence d'adressage mitochondriale et 57 acides aminés à l'extrémité C-terminale (Figure 2, manuscrit en préparation). Le fait qu'une protéine à motifs PPR interagisse physiquement avec d'autres protéines n'est pas sans rappeler la fonction du motif protéique TPR auquel s'apparente le motif PPR. Le motif TPR est connu pour être un domaine d'interaction protéine-protéine (Blatch et Lassle, 1999) et il est possible que le domaine PPR soit également capable de jouer un rôle similaire. Cette possibilité est également suggérée par de nombreuses études qui ont identifiées des protéines PPR engagées au sein de complexes protéiques de haut poids moléculaire dans les mitochondries et chloroplastes (Dominic et *al*., 2010; Gillman et *al.*, 2007). Néanmoins, les études n'ont pas démontré l'implication directe des motifs PPR de la protéine dans l'interaction.

Des réponses à la question peuvent être apportées en réalisant les expériences de double hybride pour lesquelles différentes versions tronquées de la protéine PPR9 seraient individuellement exprimées dans les levures en partenariat avec NAP1 ou TCP8. Ceci permettrait de cartographier la région minimale d'interaction de PPR9.

5.6. LE DOMAINE PPR, UNE PLATEFORME MODULABLE D'INTERACTION AVEC L'ARN

Dans les cellules, les protéines de liaison à l'ARN dictent le devenir des ARNs et leur fonction en modulant leur maturation, localisation, traduction et stabilité. Les conséquences de la liaison d'une protéine à l'ARN dépendent du site de fixation de la

protéine sur l'ARN, l'affinité de cette interaction et les tierces interactions engagées par la protéine. La liaison d'une protéine à l'ARN est médiée par des domaines de liaison à l'ARN. Le domaine de liaison à l'ARN, PUM-HD (pour Pumilio Homology Domain) a suscité une attention considérable. Le domaine PUM-HD définit la famille protéique PUF. Les membres de cette famille régulent l'expression des gènes chez les eucaryotes en influençant la stabilité et traduction des transcrits en liant de manière spécifique les régions 3'-UTR (Wharton et *al.*, 2006). Les protéines PUF sont codées chez les eucaryotes, par une petite famille de gènes et reconnaissent une séquence d'environ 8 nucléotides. Les analyses structurales de protéines PUF ont révélé que le domaine PUM-HD comprend 8 répétitions en tandem de 36 acides aminés formant chacune 3 hélices alpha qui s'empilent pour former une surface d'interaction avec les ARNs. Chaque répétition au sein de la protéine lie un nucléotide de la séquence ARN (Wang et *al.*, 2002). Cette reconnaissance de type un motif/un nucléotide suggère que la spécificité de reconnaissance de l'ARN par le domaine PUM-HD pourrait être décodée.

Plus récemment, les protéines à motifs PPR ont été reconnues comme une nouvelle classe de protéines de liaison à l'ARN dont le potentiel fonctionnel n'est pas totalement résolu. Ces protéines sont définies par la présence de motifs PPR correspondant à une séquence dégénérée de 35 acides aminés. Chez les plantes supérieures, cette famille de gènes comporte plus de 400 membres et les analyses génétiques de certains de ces membres ont montré leur vaste éventail fonctionnel associé au métabolisme de l'ARN: édition, épissage, stabilisation, définition des extrémités 5' et 3' et traduction. L'implication des protéines PPR dans ces étapes est arbitrée par la liaison directe des motifs PPR à une séquence spécifique du transcrit ARN. Les meilleures preuves biochimiques de la liaison spécifique de ces motifs PPR à l'ARN ont été apportées par des expériences de coimmunoprécipitation qui permettent l'identification des séquences ARN liées aux protéines PPR et surtout par des expériences de gel retard en utilisant des protéines PPR recombinantes. La combinaison de ces deux méthodes a permis de caractériser les sites de liaison à l'ARN de certaines PPR de la sous-famille P (présentant uniquement des motifs PPR répétés en tandem). La meilleure étude concerne la protéine PPR10 qui lie spécifiquement une séquence conservée minimale de 19 nucléotides au niveau de deux transcrits chloroplastiques distincts (*atpH* et *psaJ*) (Pfalz et *al.*, 2009). Contrairement aux protéines PUF liant environ 8 nucléotides, cette séquence minimale est particulièrement longue et suggère que les protéines PPR sont capables de lier l'ARN sur une longue surface qui reconnait plusieurs nucléotides successifs. La protéine PPR10 présente une répétition de 19 motifs PPR suggérant tout comme les protéines PUM-HD, une reconnaissance modulaire de type un motif PPR/un nucléotide. L'évolution supporte le fait que les protéines PPR peuvent être modifiée pour reconnaitre une diversité de substrats ARN. L'extrapolation des données actuelles implique que la plupart des ~450 protéines PPR chez les plantes supérieures lient des séquences distinctes d'ARN. Contrairement aux protéines PUF, les protéines PPR présentent un nombre de motifs PPR variable avec une moyenne de 12 chez *Arabidopsis* qui permettrait de lier une séquence ARN plus ou moins grande.

Alors que les fonctions des protéines PUM-HD sont limitées et la famille PUM-HD petite avec 26 membres chez *Arabidopsis* (Tam et *al*., 2010; Francischini et Quaggio, 2009), les protéines PPR ont été sélectionnées pour remplir différentes niches fonctionnelles. Ces observations suggèrent une évolution rapide et une plus grande maléabilité du motif PPR dans la reconnaissance spécifique de substrats ARN: la spécificité et l'affinité de reconnaissance d'un ARN pourrait être modulées par le nombre de motifs PPR et par la nature des acides aminés des motifs PPR entrant en contact avec l'ARN. Les études cristallographiques de protéines PPR en complexe avec leur substrat ARN permettront dans un futur proche de caractériser la nature précise de l'interaction entre les motifs PPR avec l'ARN.

Plus nous accumulons de données génétiques et biochimiques sur cette large famille de protéines et plus l'intérêt qu'elle suscite grandit. Plus que tout autre motif de liaison à l'ARN, les motifs PPR semblent constituer une plateforme unique pour l'ingénierie des protéines et leurs applications biotechnologiques dans la modulation de l'expression des génomes.

6. BIBLIOGRAPHIE

Akagi, H., Nakamura, A., Yokozeki-Misono, Y., Inagaki, A., Takahashi, H., Mori, K., and Fujimura, T. (2004). Positional cloning of the rice Rf-1 gene, a restorer of BT-type cytoplasmic male sterility that encodes a mitochondria-targeting PPR protein. Theor Appl Genet **108,** 1449-1457.

Allison, L.A. (2000). The role of sigma factors in plastid transcription. Biochimie **82,** 537-548.

Andres, C., Lurin, C., and Small, I.D. (2007). The multifarious roles of PPR proteins in plant mitochondrial gene expression. Physiologia Plantarum **129,** 14-22.

Asakura, Y., and Barkan, A. (2006). Arabidopsis orthologs of maize chloroplast splicing factors promote splicing of orthologous and species-specific group II introns. Plant Physiol **142,** 1656-1663.

Asakura, Y., and Barkan, A. (2007). A CRM domain protein functions dually in group I and group II intron splicing in land plant chloroplasts. Plant Cell **19,** 3864-3875.

Asakura, Y., Bayraktar, O.A., and Barkan, A. (2008). Two CRM protein subfamilies cooperate in the splicing of group IIB introns in chloroplasts. RNA **14,** 2319-2332.

Aubourg, S., Boudet, N., Kreis, M., and Lecharny, A. (2000). In Arabidopsis thaliana, 1% of the genome codes for a novel protein family unique to plants. Plant Mol Biol **42,** 603-613.

Barkan, A., Klipcan, L., Ostersetzer, O., Kawamura, T., Asakura, Y., and Watkins, K.P. (2007). The CRM domain: an RNA binding module derived from an ancient ribosome-associated protein. RNA **13,** 55-64.

Barkan, A., Walker, M., Nolasco, M., and Johnson, D. (1994). A nuclear mutation in maize blocks the processing and translation of several chloroplast mRNAs and provides evidence for the differential translation of alternative mRNA forms. EMBO J **13,** 3170-3181.

Barkan, A. (2004). Intron splicing in plant organelles. In Molecular Biology and Biotechnology of Plant Organelles, H. Daniell and C. Chase, eds (Dordrecht, The Netherlands: Kluwer Academic Publishers), pp. 281–308.

Beick, S., Schmitz-Linneweber, C., Williams-Carrier, R., Jensen, B., and Barkan, A. (2008). The pentatricopeptide repeat protein PPR5 stabilizes a specific tRNA precursor in maize chloroplasts. Mol Cell Biol **28,** 5337-5347.

Benne, R., Van den Burg, J., Brakenhoff, J.P., Sloof, P., Van Boom, J.H., and Tromp, M.C. (1986). Major transcript of the frameshifted coxII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. Cell **46,** 819-826.

Bentolila, S., Elliott, L.E., and Hanson, M.R. (2008). Genetic architecture of mitochondrial editing in Arabidopsis thaliana. Genetics **178,** 1693-1708.

Binder, S., Thalheim, C., and Brennicke, A. (1994). Transcription of potato mitochondrial 26S rRNA is initiated at its mature 5' end. Curr Genet **26,** 519-523.

Blatch, G.L., and Lassle, M. (1999). The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. Bioessays **21,** 932-939.

Bligny, M., Courtois, F., Thaminy, S., Chang, C.C., Lagrange, T., Baruah-Wolff, J., Stern, D., and Lerbs-Mache, S. (2000). Regulation of plastid rDNA transcription by interaction of CDF2 with two different RNA polymerases. EMBO J **19,** 1851-1860.

Bock, H., Brennicke, A., and Schuster, W. (1994). Rps3 and rpl16 genes do not overlap in Oenothera mitochondria: GTG as a potential translation initiation codon in plant mitochondria? Plant Mol Biol **24,** 811-818.

Bock, R., and Koop, H.U. (1997). Extraplastidic site-specific factors mediate RNA editing in chloroplasts. EMBO J **16,** 3282-3288.

Bock, R., Kossel, H., and Maliga, P. (1994). Introduction of a heterologous editing site into the tobacco plastid genome: the lack of RNA editing leads to a mutant phenotype. EMBO J **13,** 4623-4628.

Bollenbach, T., Schuster, G., Portnoy V., and Stern, DB. (2007). Polyadenylation, processing and degradation of chloroplast RNA. Topics in Current Genetics **19,** 175-211

Bonen, L., and Vogel, J. (2001). The ins and outs of group II introns. Trends Genet **17,** 322-331.

Brown, G.G., Formanova, N., Jin, H., Wargachuk, R., Dendy, C., Patil, P., Laforest, M., Zhang, J., Cheung, W.Y., and Landry, B.S. (2003). The radish Rfo restorer gene of Ogura cytoplasmic male sterility encodes a protein with multiple pentatricopeptide repeats. Plant J **35,** 262-272.

Bruce Cahoon, A., and Stern, D.B. (2001). Plastid transcription: a menage a trois? Trends Plant Sci **6,** 45-46.

Budar, F., and Pelletier, G. (2001). Male sterility in plants: occurrence, determinism, significance and use. C R Acad Sci III **324,** 543-550.

Cai, W., Ji, D., Peng, L., Guo, J., Ma, J., Zou, M., Lu, C., and Zhang, L. (2009). LPA66 is required for editing psbF chloroplast transcripts in Arabidopsis. Plant Physiol **150,** 1260-1271.

Carrillo, C., and Bonen, L. (1997). RNA editing status of nad7 intron domains in wheat mitochondria. Nucleic Acids Res **25,** 403-409.

Chang, C.C., Sheen, J., Bligny, M., Niwa, Y., Lerbs-Mache, S., and Stern, D.B. (1999). Functional analysis of two maize cDNAs encoding T7-like RNA polymerases. Plant Cell **11,** 911-926.

Chapdelaine, Y., and Bonen, L. (1991). The wheat mitochondrial gene for subunit I of the NADH dehydrogenase complex: a trans-splicing model for this gene-in-pieces. Cell **65,** 465-472.

Chateigner-Boutin, A.L., and Hanson, M.R. (2002). Cross-competition in transgenic chloroplasts expressing single editing sites reveals shared cis elements. Mol Cell Biol **22,** 8448-8456.

Chateigner-Boutin, A.L., and Hanson, M.R. (2003). Developmental co-variation of RNA editing extent of plastid editing sites exhibiting similar cis-elements. Nucleic Acids Res **31,** 2586-2594.

Chateigner-Boutin, A.L., Ramos-Vega, M., Guevara-Garcia, A., Andres, C., de la Luz Gutierrez-Nava, M., Cantero, A., Delannoy, E., Jimenez, L.F., Lurin, C., Small, I., and Leon, P. (2008). CLB19, a pentatricopeptide repeat protein required for editing of rpoA and clpP chloroplast transcripts. Plant J **56,** 590-602.

Chateigner-Boutin, A.L., and Small, I. (2007). A rapid high-throughput method for the detection and quantification of RNA editing based on high-resolution melting of amplicons. Nucleic Acids Res **35,** e114.

Chaudhuri, S., Carrer, H., and Maliga, P. (1995). Site-specific factor involved in the editing of the psbL mRNA in tobacco plastids. EMBO J **14,** 2951-2957.

Cho, Y., Qiu, Y.L., Kuhlman, P., and Palmer, J.D. (1998). Explosive invasion of plant mitochondria by a group I intron. Proc Natl Acad Sci U S A **95,** 14244-14249.

Choquet, Y., Stern, D.B., Wostrikoff, K., Kuras, R., Girard-Bascou, J., and Wollman, F.A. (1998). Translation of cytochrome f is autoregulated through the 5' untranslated region of petA mRNA in Chlamydomonas chloroplasts. Proc Natl Acad Sci U S A **95,** 4380-4385.

Choquet, Y., and Wollman, F.A. (2002). Translational regulations as specific traits of chloroplast gene expression. FEBS Lett **529,** 39-42.

Choquet, Y., Wostrikoff, K., Rimbault, B., Zito, F., Girard-Bascou, J., Drapier, D., and Wollman, F.A. (2001). Assembly-controlled regulation of chloroplast gene translation. Biochem Soc Trans **29,** 421-426.

Choury, D., Farre, J.C., Jordana, X., and Araya, A. (2004). Different patterns in the recognition of editing sites in plant mitochondria. Nucleic Acids Res **32,** 6397-6406.

Christensen, A.C., Lyznik, A., Mohammed, S., Elowsky, C.G., Elo, A., Yule, R., and Mackenzie, S.A. (2005). Dual-domain, dual-targeting organellar protein presequences in Arabidopsis can use non-AUG start codons. Plant Cell **17,** 2805-2816.

Coffin, J.W., Dhillon, R., Ritzel, R.G., and Nargang, F.E. (1997). The Neurospora crassa cya-5 nuclear gene encodes a protein with a region of homology to the Saccharomyces cerevisiae PET309 protein and is required in a post-transcriptional step for the expression of the mitochondrially encoded COXI protein. Curr Genet **32,** 273-280.

Covello, P.S., and Gray, M.W. (1989). RNA editing in plant mitochondria. Nature **341,** 662-666.

Cummings, M.P., and Myers, D.S. (2004). Simple statistical models predict C-to-U edited sites in plant mitochondrial RNA. BMC Bioinformatics **5,** 132.

Davies, S.M., Rackham, O., Shearwood, A.M., Hamilton, K.L., Narsai, R., Whelan, J., and Filipovska, A. (2009). Pentatricopeptide repeat domain protein 3 associates with the mitochondrial small ribosomal subunit and regulates translation. FEBS Lett **583,** 1853-1858.

de Longevialle, A.F., Hendrickson, L., Taylor, N.L., Delannoy, E., Lurin, C., Badger, M., Millar, A.H., and Small, I. (2008). The pentatricopeptide repeat gene OTP51 with two LAGLIDADG motifs is required for the cis-splicing of plastid ycf3 intron 2 in Arabidopsis thaliana. Plant J **56,** 157-168.

de Longevialle, A.F., Meyer, E.H., Andres, C., Taylor, N.L., Lurin, C., Millar, A.H., and Small, I.D. (2007). The pentatricopeptide repeat gene OTP43 is required for trans-splicing of the mitochondrial nad1 Intron 1 in Arabidopsis thaliana. Plant Cell **19,** 3256-3265.

Del Campo, E.M. (2009). Post-transcriptional control of chloroplast gene expression. Gene Regul Syst Bio **3,** 31-47.

Del Campo, E.M., Sabater, B., and Martin, M. (2002). Post-transcriptional control of chloroplast gene expression. Accumulation of stable psaC mRNA is due to downstream RNA cleavages in the ndhD gene. J Biol Chem **277,** 36457-36464.

Ding, Y.H., Liu, N.Y., Tang, Z.S., Liu, J., and Yang, W.C. (2006). Arabidopsis GLUTAMINE-RICH PROTEIN23 is essential for early embryogenesis and encodes a novel nuclear PPR motif protein that interacts with RNA polymerase II subunit III. Plant Cell **18,** 815-830.

Dombrowski, S., Brennicke, A., and Binder, S. (1997). 3'-Inverted repeats in plant mitochondrial mRNAs are processing signals rather than transcription terminators. EMBO J **16,** 5069-5076.

Dong, F.G., Wilson, K.G., and Makaroff, C.A. (1998). The radish (Raphanus sativus L.) mitochondrial cox2 gene contains an ACG at the predicted translation initiation site. Curr Genet **34,** 79-87.

Duchene, A.M., Giritch, A., Hoffmann, B., Cognat, V., Lancelin, D., Peeters, N.M., Zaepfel, M., Marechal-Drouard, L., and Small, I.D. (2005). Dual targeting is the rule for organellar aminoacyl-tRNA synthetases in Arabidopsis thaliana. Proc Natl Acad Sci U S A **102,** 16484-16489.

Duff, R.J., and Moore, F.B. (2005). Pervasive RNA editing among hornwort rbcL transcripts except Leiosporoceros. J Mol Evol **61,** 571-578.

Ellis, T.P., Helfenbein, K.G., Tzagoloff, A., and Dieckmann, C.L. (2004). Aep3p stabilizes the mitochondrial bicistronic mRNA encoding subunits 6 and 8 of the H+-translocating ATP synthase of Saccharomyces cerevisiae. J Biol Chem **279,** 15728-15733.

Faivre-Nitschke, S.E., Grienenberger, J.M., and Gualberto, J.M. (1999). A prokaryotic-type cytidine deaminase from Arabidopsis thaliana gene expression and functional characterization. Eur J Biochem **263,** 896-903.

Fargo, D.C., Zhang, M., Gillham, N.W., and Boynton, J.E. (1998). Shine-Dalgarno-like sequences are not required for translation of chloroplast mRNAs in Chlamydomonas reinhardtii chloroplasts or in Escherichia coli. Mol Gen Genet **257,** 271-282.

Farre, J.C., Leon, G., Jordana, X., and Araya, A. (2001). cis Recognition elements in plant mitochondrion RNA editing. Mol Cell Biol **21,** 6731-6737.

Fedorova, O., and Zingler, N. (2007). Group II introns: structure, folding and splicing mechanism. Biol Chem **388,** 665-678.

Ferro, M., Brugiere, S., Salvi, D., Seigneurin-Berny, D., Court, M., Moyet, L., Ramus, C., Miras, S., Mellal, M., Le Gall, S., Kieffer-Jaquinod, S., Bruley, C., Garin, J., Joyard, J., Masselon, C., and Rolland, N. AT_CHLORO: A comprehensive chloroplast proteome database with subplastidiallocalization and curated information on envelope proteins. Mol Cell Proteomics.

Fey, J., and Marechal-Drouard, L. (1999). Compilation and analysis of plant mitochondrial promoter sequences: An illustration of a divergent evolution between monocot and dicot mitochondria. Biochem Biophys Res Commun **256,** 409-414.

Fisk, D.G., Walker, M.B., and Barkan, A. (1999). Molecular cloning of the maize gene crp1 reveals similarity between regulators of mitochondrial and chloroplast gene expression. EMBO J **18,** 2621-2630.

Forner, J., Weber, B., Thuss, S., Wildum, S., and Binder, S. (2007). Mapping of mitochondrial mRNA termini in Arabidopsis thaliana: t-elements contribute to 5' and 3' end formation. Nucleic Acids Res **35,** 3676-3692.

Francischini, C.W., and Quaggio, R.B. (2009). Molecular characterization of Arabidopsis thaliana PUF proteins--binding specificity and target candidates. FEBS J **276,** 5456-5470.

Freyer, R., Hoch, B., Neckermann, K., Maier, R.M., and Kossel, H. (1993). RNA editing in maize chloroplasts is a processing step independent of splicing and cleavage to monocistronic mRNAs. Plant J **4,** 621-629.

Giege, P., and Brennicke, A. (1999). RNA editing in Arabidopsis mitochondria effects 441 C to U changes in ORFs. Proc Natl Acad Sci U S A **96,** 15324-15329.

Giege, P., Hoffmann, M., Binder, S., and Brennicke, A. (2000). RNA degradation buffers asymmetries of transcription in Arabidopsis mitochondria. EMBO Rep **1,** 164-170.

Glanz, S., and Kuck, U. (2009). Trans-splicing of organelle introns--a detour to continuous RNAs. Bioessays **31,** 921-934.

Gobert, A., Gutmann, B., Taschner, A., Gossringer, M., Holzmann, J., Hartmann, R.K., Rossmanith, W., and Giege, P. (2010). A single Arabidopsis organellar protein has RNase P activity. Nat Struct Mol Biol **17,** 740-744.

Gohil, V.M., Nilsson, R., Belcher-Timme, C.A., Luo, B., Root, D.E., and Mootha, V.K. (2010). Mitochondrial and nuclear genomic responses to loss of LRPPRC expression. J Biol Chem **285,** 13742-13747.

Grabowski, E., Miao, Y., Mulisch, M., and Krupinska, K. (2008). Single-stranded DNA-binding protein Whirly1 in barley leaves is located in plastids and the nucleus of the same cell. Plant Physiol **147,** 1800-1804.

Gray, J.C. (2003). Chloroplast-to-nucleus signalling: a role for Mg-protoporphyrin. Trends Genet **19,** 526-529.

Gray, M.W. (1999). Evolution of organellar genomes. Curr Opin Genet Dev **9,** 678-687.

Grohmann, L., Thieck, O., Herz, U., Schroder, W., and Brennicke, A. (1994). Translation of nad9 mRNAs in mitochondria from Solanum tuberosum is restricted to completely edited transcripts. Nucleic Acids Res **22,** 3304-3311.

Gualberto, J.M., Lamattina, L., Bonnard, G., Weil, J.H., and Grienenberger, J.M. (1989). RNA editing in wheat mitochondria results in the conservation of protein sequences. Nature **341,** 660-662.

Gutierrez-Marcos, J.F., Dal Pra, M., Giulini, A., Costa, L.M., Gavazzi, G., Cordelier, S., Sellam, O., Tatout, C., Paul, W., Perez, P., Dickinson, H.G., and Consonni, G. (2007). Empty pericarp4 encodes a mitochondrion-targeted pentatricopeptide repeat protein necessary for seed development and plant growth in maize. Plant Cell **19,** 196-210.

Hammani, K., Okuda, K., Tanz, S.K., Chateigner-Boutin, A.L., Shikanai, T., and Small, I. (2009). A study of new Arabidopsis chloroplast RNA editing mutants reveals general features of editing factors and their target sites. Plant Cell **21,** 3686-3699.

Hao, W., and Palmer, J.D. (2009). Fine-scale mergers of chloroplast and mitochondrial genes create functional, transcompartmentally chimeric mitochondrial genes. Proc Natl Acad Sci U S A **106,** 16728-16733.

Hashimoto, M., Endo, T., Peltier, G., Tasaka, M., and Shikanai, T. (2003). A nucleus-encoded factor, CRR2, is essential for the expression of chloroplast ndhB in Arabidopsis. Plant J **36,** 541-549.

Hattori, M., Miyake, H., and Sugita, M. (2007). A Pentatricopeptide repeat protein is required for RNA processing of clpP Pre-mRNA in moss chloroplasts. J Biol Chem **282,** 10773-10782.

Hayes, M.L., Reed, M.L., Hegeman, C.E., and Hanson, M.R. (2006). Sequence elements critical for efficient RNA editing of a tobacco chloroplast transcript in vivo and in vitro. Nucleic Acids Res **34,** 3742-3754.

Hayes, R., Kudla, J., Schuster, G., Gabay, L., Maliga, P., and Gruissem, W. (1996). Chloroplast mRNA 3'-end processing by a high molecular weight protein complex is regulated by nuclear encoded RNA binding proteins. EMBO J **15,** 1132-1141.

Hedtke, B., Borner, T., and Weihe, A. (2000). One RNA polymerase serving two genomes. EMBO Rep **1,** 435-440.

Hegeman, C.E., Halter, C.P., Owens, T.G., and Hanson, M.R. (2005). Expression of complementary RNA from chloroplast transgenes affects editing efficiency of transgene and endogenous chloroplast transcripts. Nucleic Acids Res **33,** 1454-1464.

Hegeman, C.E., Hayes, M.L., and Hanson, M.R. (2005). Substrate and cofactor requirements for RNA editing of chloroplast transcripts in Arabidopsis in vitro. Plant J **42,** 124-132.

Hiesel, R., Wissinger, B., Schuster, W., and Brennicke, A. (1989). RNA editing in plant mitochondria. Science **246,** 1632-1634.

Hirose, T., and Sugiura, M. (2001). Involvement of a site-specific trans-acting factor and a common RNA-binding protein in the editing of chloroplast mRNAs: development of a chloroplast in vitro RNA editing system. EMBO J **20,** 1144-1152.

Hoch, B., Maier, R.M., Appel, K., Igloi, G.L., and Kossel, H. (1991). Editing of a chloroplast mRNA by creation of an initiation codon. Nature **353,** 178-180.

Hoch, B., Maier, R.M., Appel, K., Igloi, G.L., and Kossel, H. (1991). Editing of a chloroplast mRNA by creation of an initiation codon. Nature **353,** 178-180.

Hoffmann, M., Dombrowski, S., Guha, C., and Binder, S. (1999). Cotranscription of the rpl5-rps14-cob gene cluster in pea mitochondria. Mol Gen Genet **261,** 537-545.

Holec, S., Lange, H., Kuhn, K., Alioua, M., Borner, T., and Gagliardi, D. (2006). Relaxed transcription in Arabidopsis mitochondria is counterbalanced by RNA stability control mediated by polyadenylation and polynucleotide phosphorylase. Mol Cell Biol **26,** 2869-2876.

Ikeda, T.M., and Gray, M.W. (1999). Characterization of a DNA-binding protein implicated in transcription in wheat mitochondria. Mol Cell Biol **19,** 8113-8122.

Ingman, M., Kaessmann, H., Paabo, S., and Gyllensten, U. (2000). Mitochondrial genome variation and the origin of modern humans. Nature **408,** 708-713.

Isono, K., Shimizu, M., Yoshimoto, K., Niwa, Y., Satoh, K., Yokota, A., and Kobayashi, H. (1997). Leaf-specifically expressed genes for polypeptides destined for chloroplasts with domains of sigma70 factors of bacterial RNA polymerases in Arabidopsis thaliana. Proc Natl Acad Sci U S A **94,** 14948-14953.

Jenkins, B.D., and Barkan, A. (2001). Recruitment of a peptidyl-tRNA hydrolase as a facilitator of group II intron splicing in chloroplasts. EMBO J **20,** 872-879.

Jenkins, B.D., Kulhanek, D.J., and Barkan, A. (1997). Nuclear mutations that block group II RNA splicing in maize chloroplasts reveal several intron classes with distinct requirements for splicing factors. Plant Cell **9,** 283-296.

Jinek, M., Rehwinkel, J., Lazarus, B.D., Izaurralde, E., Hanover, J.A., and Conti, E. (2004). The superhelical TPR-repeat domain of O-linked GlcNAc transferase exhibits structural similarities to importin alpha. Nat Struct Mol Biol **11,** 1001-1007.

Johnson, X., Wostrikoff, K., Finazzi, G., Kuras, R., Schwarz, C., Bujaldon, S., Nickelsen, J., Stern, D.B., Wollman, F.A., and Vallon, O. (2010). MRL1, a conserved Pentatricopeptide repeat protein, is required for stabilization of rbcL mRNA in Chlamydomonas and Arabidopsis. Plant Cell **22,** 234-248.

Jonietz, C., Forner, J., Holzle, A., Thuss, S., and Binder, S. (2010). RNA PROCESSING FACTOR2 is required for 5' end processing of nad9 and cox3 mRNAs in mitochondria of Arabidopsis thaliana. Plant Cell **22,** 443-453.

Karpinski, S., Reynolds, H., Karpinska, B., Wingsle, G., Creissen, G., and Mullineaux, P. (1999). Systemic signaling and acclimation in response to excess excitation energy in Arabidopsis. Science **284,** 654-657.

Keren, I., Bezawork-Geleta, A., Kolton, M., Maayan, I., Belausov, E., Levy, M., Mett, A., Gidoni, D., Shaya, F., and Ostersetzer-Biran, O. (2009). AtnMat2, a nuclear-encoded maturase required for splicing of group-II introns in Arabidopsis mitochondria. RNA **15,** 2299-2311.

Keren, I., Bezawork-Geleta, A., Kolton, M., Maayan, I., Belausov, E., Levy, M., Mett, A., Gidoni, D., Shaya, F., and Ostersetzer-Biran, O. (2009). AtnMat2, a nuclear-encoded maturase required for splicing of group-II introns in Arabidopsis mitochondria. RNA **15,** 2299-2311.

Keren, I., Klipcan, L., Bezawork-Geleta, A., Kolton, M., Shaya, F., and Ostersetzer-Biran, O. (2008). Characterization of the molecular basis of group II intron RNA recognition by CRS1-CRM domains. J Biol Chem **283,** 23333-23342.

Kim, J., and Mullet, J.E. (1994). Ribosome-binding sites on chloroplast rbcL and psbA mRNAs and light-induced initiation of D1 translation. Plant Mol Biol **25,** 437-448.

Kim, S.R., Yang, J.I., Moon, S., Ryu, C.H., An, K., Kim, K.M., Yim, J., and An, G. (2009). Rice OGR1 encodes a pentatricopeptide repeat-DYW protein and is essential for RNA editing in mitochondria. Plant J **59,** 738-749.

Kishine, M., Takabayashi, A., Munekage, Y., Shikanai, T., Endo, T., and Sato, F. (2004). Ribosomal RNA processing and an RNase R family member in chloroplasts of Arabidopsis. Plant Mol Biol **55,** 595-606.

Klein, M., Eckert-Ossenkopp, U., Schmiedeberg, I., Brandt, P., Unseld, M., Brennicke, A., and Schuster, W. (1994). Physical mapping of the mitochondrial genome of Arabidopsis thaliana by cosmid and YAC clones. Plant J **6,** 447-455.

Klodmann, J., Sunderhaus, S., Nimtz, M., Jansch, L., and Braun, H.P. Internal architecture of mitochondrial complex I from Arabidopsis thaliana. Plant Cell **22,** 797-810.

Knoop, V. (2004). The mitochondrial DNA of land plants: peculiarities in phylogenetic perspective. Curr Genet **46,** 123-139.

Knoop, V., Schuster, W., Wissinger, B., and Brennicke, A. (1991). Trans splicing integrates an exon of 22 nucleotides into the nad5 mRNA in higher plant mitochondria. EMBO J **10,** 3483-3493.

Kobayashi, Y., Kanesaki, Y., Tanaka, A., Kuroiwa, H., Kuroiwa, T., and Tanaka, K. (2009). Tetrapyrrole signal as a cell-cycle coordinator from organelle to nuclear DNA replication in plant cells. Proc Natl Acad Sci U S A **106,** 803-807.

Kobayashi, Y., Matsuo, M., Sakamoto, K., Wakasugi, T., Yamada, K., and Obokata, J. (2008). Two RNA editing sites with cis-acting elements of moderate sequence identity are recognized by an identical site-recognition protein in tobacco chloroplasts. Nucleic Acids Res **36,** 311-318.

Koizuka, N., Imai, R., Fujimoto, H., Hayakawa, T., Kimura, Y., Kohno-Murase, J., Sakai, T., Kawasaki, S., and Imamura, J. (2003). Genetic characterization of a pentatricopeptide repeat protein gene, orf687, that restores fertility in the cytoplasmic male-sterile Kosena radish. Plant J **34,** 407-415.

Kotera, E., Tasaka, M., and Shikanai, T. (2005). A pentatricopeptide repeat protein is essential for RNA editing in chloroplasts. Nature **433,** 326-330.

Krause, K., Herrmann, U., Fuß, J., Miao, Y., Krupinska, K. (2009) Whirly proteins as communicators between plant organelles and the nucleus. Endocytobiology **19,** 51-62

Krause, K., and Krupinska, K. (2009). Nuclear regulators with a second home in organelles. Trends Plant Sci **14,** 194-199.

Kroeger, T.S., Watkins, K.P., Friso, G., van Wijk, K.J., and Barkan, A. (2009). A plant-specific RNA-binding domain revealed through analysis of chloroplast group II intron splicing. Proc Natl Acad Sci U S A **106,** 4537-4542.

Kuhn, K., Richter, U., Meyer, E.H., Delannoy, E., de Longevialle, A.F., O'Toole, N., Borner, T., Millar, A.H., Small, I.D., and Whelan, J. (2009). Phage-type RNA polymerase RPOTmp performs gene-specific transcription in mitochondria of Arabidopsis thaliana. Plant Cell **21,** 2762-2779.

Kuhn, K., Weihe, A., and Borner, T. (2005). Multiple promoters are a common feature of mitochondrial genes in Arabidopsis. Nucleic Acids Res **33,** 337-346.

Kunzmann, A., Brennicke, A., and Marchfelder, A. (1998). 5' end maturation and RNA editing have to precede tRNA 3' processing in plant mitochondria. Proc Natl Acad Sci U S A **95,** 108-113.

Kuroda, H., Suzuki, H., Kusumegi, T., Hirose, T., Yukawa, Y., and Sugiura, M. (2007). Translation of psbC mRNAs starts from the downstream GUG, not the upstream AUG, and requires the extended Shine-Dalgarno sequence in tobacco chloroplasts. Plant Cell Physiol **48,** 1374-1378.

Lambowitz, A.M., and Zimmerly, S. (2004). Mobile group II introns. Annu Rev Genet **38,** 1-35.

Larkin, R.M., Alonso, J.M., Ecker, J.R., and Chory, J. (2003). GUN4, a regulator of chlorophyll synthesis and intracellular signaling. Science **299,** 902-906.

Liere, K., and Maliga, P. (1999). In vitro characterization of the tobacco rpoB promoter reveals a core sequence motif conserved between phage-type plastid and plant mitochondrial promoters. EMBO J **18,** 249-257.

Logan, D.C., and Knight, M.R. (2003). Mitochondrial and cytosolic calcium dynamics are differentially regulated in plants. Plant Physiol **133,** 21-24.

Loiselay, C., Gumpel, N.J., Girard-Bascou, J., Watson, A.T., Purton, S., Wollman, F.A., and Choquet, Y. (2008). Molecular identification and function of cis- and trans-acting determinants for petA transcript stability in Chlamydomonas reinhardtii chloroplasts. Mol Cell Biol **28,** 5529-5542.

Lu, B., and Hanson, M.R. (1994). A single homogeneous form of ATP6 protein accumulates in petunia mitochondria despite the presence of differentially edited atp6 transcripts. Plant Cell **6,** 1955-1968.

Lupold, D.S., Caoile, A.G., and Stern, D.B. (1999). The maize mitochondrial cox2 gene has five promoters in two genomic regions, including a complex promoter consisting of seven overlapping units. J Biol Chem **274,** 3897-3903.

Lurin, C., Andres, C., Aubourg, S., Bellaoui, M., Bitton, F., Bruyere, C., Caboche, M., Debast, C., Gualberto, J., Hoffmann, B., Lecharny, A., Le Ret, M., Martin-Magniette, M.L., Mireau, H., Peeters, N., Renou, J.P., Szurek, B., Taconnat, L., and Small, I. (2004). Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. Plant Cell **16,** 2089-2103.

Mackenzi, S.A. (2007). The unique biology of mitochondrial genome instability in plants, pp. 36-46, in Plant Mitochondria. Edited by D. LOGAN, Blackwell Publ., Oxford.

Malek, O., Lattig, K., Hiesel, R., Brennicke, A., and Knoop, V. (1996). RNA editing in bryophytes and a molecular phylogeny of land plants. EMBO J **15,** 1403-1411.

Mancebo, R., Zhou, X., Shillinglaw, W., Henzel, W., and Macdonald, P.M. (2001). BSF binds specifically to the bicoid mRNA 3' untranslated region and contributes to stabilization of bicoid mRNA. Mol Cell Biol **21,** 3462-3471.

Manthey, G.M., and McEwen, J.E. (1995). The product of the nuclear gene PET309 is required for translation of mature mRNA and stability or production of intron-containing RNAs derived from the mitochondrial COX1 locus of Saccharomyces cerevisiae. EMBO J **14,** 4031-4043.

Marechal, A., Parent, J.S., Veronneau-Lafortune, F., Joyeux, A., Lang, B.F., and Brisson, N. (2009). Whirly proteins maintain plastid genome stability in Arabidopsis. Proc Natl Acad Sci U S A **106,** 14693-14698.

Marechal-Drouard, L., Cosset, A., Remacle, C., Ramamonjisoa, D., and Dietrich, A. (1996). A single editing event is a prerequisite for efficient processing of potato mitochondrial phenylalanine tRNA. Mol Cell Biol **16,** 3504-3510.

Marechal-Drouard, L., Kumar, R., Remacle, C., and Small, I. (1996). RNA editing of larch mitochondrial tRNA(His) precursors is a prerequisite for processing. Nucleic Acids Res **24,** 3229-3234.

Marienfeld, J., Unseld, M., and Brennicke, A. (1999). The mitochondrial genome of Arabidopsis is composed of both native and immigrant information. Trends Plant Sci **4,** 495-502.

Martin, W. (2003). Gene transfer from organelles to the nucleus: frequent and in big chunks. Proc Natl Acad Sci U S A **100,** 8612-8614.

Martin, W., and Herrmann, R.G. (1998). Gene transfer from organelles to the nucleus: how much, what happens, and Why? Plant Physiol **118,** 9-17.

Mathy, N., Benard, L., Pellegrini, O., Daou, R., Wen, T., and Condon, C. (2007). 5'-to-3' exoribonuclease activity in bacteria: role of RNase J1 in rRNA maturation and 5' stability of mRNA. Cell **129,** 681-692.

Mayfield, S.P., Cohen, A., Danon, A., and Yohn, C.B. (1994). Translation of the psbA mRNA of Chlamydomonas reinhardtii requires a structured RNA element contained within the 5' untranslated region. J Cell Biol **127,** 1537-1545.

Meierhoff, K., Felder, S., Nakamura, T., Bechtold, N., and Schuster, G. (2003). HCF152, an Arabidopsis RNA binding pentatricopeptide repeat protein involved in the processing of chloroplast psbB-psbT-psbH-petB-petD RNAs. Plant Cell **15,** 1480-1495.

Meng, Q., Wang, Y., and Liu, X.Q. (2005). An intron-encoded protein assists RNA splicing of multiple similar introns of different bacterial genes. J Biol Chem **280,** 35085-35088.

Michel, F., and Ferat, J.L. (1995). Structure and activities of group II introns. Annu Rev Biochem **64,** 435-461.

Mili, S., and Pinol-Roma, S. (2003). LRP130, a pentatricopeptide motif protein with a noncanonical RNA-binding domain, is bound in vivo to mitochondrial and nuclear RNAs. Mol Cell Biol **23,** 4972-4982.

Millar, A.H., Day, D.A., and Whelan, J. (2004). Mitochondrial Biogenesis and Function in Arabidopsis. American Society of Plant Biologists.

Millar, A.H., Whelan, J., and Small, I. (2006). Recent surprises in protein targeting to mitochondria and plastids. Curr Opin Plant Biol **9,** 610-615.

Miyamoto, T., Obokata, J., and Sugiura, M. (2002). Recognition of RNA editing sites is directed by unique proteins in chloroplasts: biochemical identification of cis-acting elements and trans-acting factors involved in RNA editing in tobacco and pea chloroplasts. Mol Cell Biol **22,** 6726-6734.

Miyamoto, T., Obokata, J., and Sugiura, M. (2004). A site-specific factor interacts directly with its cognate RNA editing site in chloroplast transcripts. Proc Natl Acad Sci U S A **101,** 48-52.

Mochizuki, N., Brusslan, J.A., Larkin, R., Nagatani, A., and Chory, J. (2001). Arabidopsis genomes uncoupled 5 (GUN5) mutant reveals the involvement of Mg-chelatase H subunit in plastid-to-nucleus signal transduction. Proc Natl Acad Sci U S A **98,** 2053-2058.

Mohr, G., and Lambowitz, A.M. (2003). Putative proteins related to group II intron reverse transcriptase/maturases are encoded by nuclear genes in higher plants. Nucleic Acids Res **31,** 647-652.

Mullet, J.E. (1993). Dynamic regulation of chloroplast transcription. Plant Physiol **103,** 309-313.

Nakagawa, N., and Sakurai, N. (2006). A mutation in At-nMat1a, which encodes a nuclear gene having high similarity to group II intron maturase, causes impaired splicing of mitochondrial NAD4 transcript and altered carbon metabolism in Arabidopsis thaliana. Plant Cell Physiol **47,** 772-783.

Nakamura, T., Meierhoff, K., Westhoff, P., and Schuster, G. (2003). RNA-binding properties of HCF152, an Arabidopsis PPR protein involved in the processing of chloroplast RNA. Eur J Biochem **270,** 4070-4081.

Nakamura, T., and Sugita, M. (2008). A conserved DYW domain of the pentatricopeptide repeat protein possesses a novel endoribonuclease activity. FEBS Lett **582,** 4163-4168.

Neuhaus, H.E., and Emes, M.J. (2000). Nonphotosynthetic Metabolism in Plastids. Annu Rev Plant Physiol Plant Mol Biol **51,** 111-140.

Neuwirt, J., Takenaka, M., van der Merwe, J.A., and Brennicke, A. (2005). An in vitro RNA editing system from cauliflower mitochondria: editing site recognition parameters can vary in different plant species. RNA **11,** 1563-1570.

Nishikura, K. (2009). Functions and Regulation of RNA Editing by ADAR Deaminases. Annu Rev Biochem.

Notsu, Y., Masood, S., Nishikawa, T., Kubo, N., Akiduki, G., Nakazono, M., Hirai, A., and Kadowaki, K. (2002). The complete sequence of the rice (Oryza sativa L.) mitochondrial genome: frequent DNA sequence acquisition and loss during the evolution of flowering plants. Mol Genet Genomics **268,** 434-445.

O'Toole, N., Hattori, M., Andres, C., Iida, K., Lurin, C., Schmitz-Linneweber, C., Sugita, M., and Small, I. (2008). On the expansion of the pentatricopeptide repeat gene family in plants. Mol Biol Evol **25,** 1120-1128.

Okuda, K., Chateigner-Boutin, A.L., Nakamura, T., Delannoy, E., Sugita, M., Myouga, F., Motohashi, R., Shinozaki, K., Small, I., and Shikanai, T. (2009). Pentatricopeptide repeat proteins with the DYW motif have distinct molecular functions in RNA editing and RNA cleavage in Arabidopsis chloroplasts. Plant Cell **21,** 146-156.

Okuda, K., Hammani, K., Tanz, S.K., Peng, L., Fukao, Y., Myouga, F., Motohashi, R., Shinozaki, K., Small, I., and Shikanai, T. The pentatricopeptide repeat protein OTP82 is required for RNA editing of plastid ndhB and ndhG transcripts. Plant J **61,** 339-349.

Okuda, K., Myouga, F., Motohashi, R., Shinozaki, K., and Shikanai, T. (2007). Conserved domain structure of pentatricopeptide repeat proteins involved in chloroplast RNA editing. Proc Natl Acad Sci U S A **104,** 8178-8183.

Okuda, K., Nakamura, T., Sugita, M., Shimizu, T., and Shikanai, T. (2006). A pentatricopeptide repeat protein is a site recognition factor in chloroplast RNA editing. J Biol Chem **281,** 37661-37667.

Ostersetzer, O., Cooke, A.M., Watkins, K.P., and Barkan, A. (2005). CRS1, a chloroplast group II intron splicing factor, promotes intron folding through specific interactions with two intron domains. Plant Cell **17,** 241-255.

Ostheimer, G.J., Williams-Carrier, R., Belcher, S., Osborne, E., Gierke, J., and Barkan, A. (2003). Group II intron splicing factors derived by diversification of an ancient RNA-binding domain. EMBO J **22,** 3919-3929.

Peebles, C.L., Perlman, P.S., Mecklenburg, K.L., Petrillo, M.L., Tabor, J.H., Jarrell, K.A., and Cheng, H.L. (1986). A self-splicing RNA excises an intron lariat. Cell **44,** 213-223.

Perrin, R., Lange, H., Grienenberger, J.M., and Gagliardi, D. (2004). AtmtPNPase is required for multiple aspects of the 18S rRNA metabolism in Arabidopsis thaliana mitochondria. Nucleic Acids Res **32,** 5174-5182.

Perrin, R., Meyer, E.H., Zaepfel, M., Kim, Y.J., Mache, R., Grienenberger, J.M., Gualberto, J.M., and Gagliardi, D. (2004). Two exoribonucleases act sequentially to process mature 3'-ends of atp9 mRNAs in Arabidopsis mitochondria. J Biol Chem **279,** 25440-25446.

Pfalz, J., Bayraktar, O.A., Prikryl, J., and Barkan, A. (2009). Site-specific binding of a PPR protein defines and stabilizes 5' and 3' mRNA termini in chloroplasts. EMBO J **28,** 2042-2052.

Pfalz, J., Liere, K., Kandlbinder, A., Dietz, K.J., and Oelmuller, R. (2006). pTAC2, -6, and -12 are components of the transcriptionally active plastid chromosome that are required for plastid gene expression. Plant Cell **18,** 176-197.

Pfannschmidt, T., Ogrzewalla, K., Baginsky, S., Sickmann, A., Meyer, H.E., and Link, G. (2000). The multisubunit chloroplast RNA polymerase A from mustard (Sinapis alba L.). Integration of a prokaryotic core into a larger complex with organelle-specific functions. Eur J Biochem **267,** 253-261.

Phreaner, C.G., Williams, M.A., and Mulligan, R.M. (1996). Incomplete editing of rps12 transcripts results in the synthesis of polymorphic polypeptides in plant mitochondria. Plant Cell **8,** 107-117.

Prikryl, J., Watkins, K.P., Friso, G., van Wijk, K.J., and Barkan, A. (2008). A member of the Whirly family is a multifunctional RNA- and DNA-binding protein that is essential for chloroplast biogenesis. Nucleic Acids Res **36,** 5152-5165.

Pring, D.R., Mullen, J.A., and Kempken, F. (1992). Conserved sequence blocks 5' to start codons of plant mitochondrial genes. Plant Mol Biol **19,** 313-317.

Quesada-Vargas, T., Ruiz, O.N., and Daniell, H. (2005). Characterization of heterologous multigene operons in transgenic chloroplasts: transcription, processing, and translation. Plant Physiol **138,** 1746-1762.

Rackham, O., Davies, S.M., Shearwood, A.M., Hamilton, K.L., Whelan, J., and Filipovska, A. (2009). Pentatricopeptide repeat domain protein 1 lowers the levels of mitochondrial leucine tRNAs in cells. Nucleic Acids Res **37,** 5859-5867.

Rajasekhar, V.K., and Mulligan, R.M. (1993). RNA Editing in Plant Mitochondria: [alpha]-Phosphate Is Retained during C-to-U Conversion in mRNAs. Plant Cell **5,** 1843-1852.

Raven, J.A., and Allen, J.F. (2003). Genomics and chloroplast evolution: what did cyanobacteria do for plants? Genome Biol **4,** 209.

Reed, M.L., Peeters, N.M., and Hanson, M.R. (2001). A single alteration 20 nt 5' to an editing target inhibits chloroplast RNA editing in vivo. Nucleic Acids Res **29,** 1507-1513.

Remacle, C., and Marechal-Drouard, L. (1996). Characterization of the potato mitochondrial transcription unit containing 'native' trnS (GCU), trnF (GAA) and trnP (UGG). Plant Mol Biol **30,** 553-563.

Richly, E., and Leister, D. (2004). NUMTs in sequenced eukaryotic genomes. Mol Biol Evol **21,** 1081-1084.

Rivals, E., Bruyere, C., Toffano-Nioche, C., and Lecharny, A. (2006). Formation of the Arabidopsis pentatricopeptide repeat family. Plant Physiol **141,** 825-839.

Robbins, J.C., Heller, W.P., and Hanson, M.R. (2009). A comparative genomics approach identifies a PPR-DYW protein that is essential for C-to-U editing of the Arabidopsis chloroplast accD transcript. RNA **15,** 1142-1153.

Rogalski, M., Schottler, M.A., Thiele, W., Schulze, W.X., and Bock, R. (2008). Rpl33, a nonessential plastid-encoded ribosomal protein in tobacco, is required under cold stress conditions. Plant Cell **20,** 2221-2237.

Rudinger, M., Polsakiewicz, M., and Knoop, V. (2008). Organellar RNA editing and plant-specific extensions of pentatricopeptide repeat proteins in jungermanniid but not in marchantiid liverworts. Mol Biol Evol **25,** 1405-1414.

Ruf, S., Zeltz, P., and Kossel, H. (1994). Complete RNA editing of unspliced and dicistronic transcripts of the intron-containing reading frame IRF170 from maize chloroplasts. Proc Natl Acad Sci U S A **91,** 2295-2299.

Sakamoto, W., Sturm, N.R., Kindle, K.L., and Stern, D.B. (1994). petD mRNA maturation in Chlamydomonas reinhardtii chloroplasts: role of 5' endonucleolytic processing. Mol Cell Biol **14,** 6180-6186.

Salinas, T., Duchene, A.M., and Marechal-Drouard, L. (2008). Recent advances in tRNA mitochondrial import. Trends Biochem Sci **33,** 320-329.

Salone, V., Rudinger, M., Polsakiewicz, M., Hoffmann, B., Groth-Malonek, M., Szurek, B., Small, I., Knoop, V., and Lurin, C. (2007). A hypothesis on the identification of the editing enzyme in plant organelles. FEBS Lett **581,** 4132-4138.

Sasaki, T., Yukawa, Y., Wakasugi, T., Yamada, K., and Sugiura, M. (2006). A simple in vitro RNA editing assay for chloroplast transcripts using fluorescent dideoxynucleotides: distinct types of sequence elements required for editing of ndh transcripts. Plant J **47,** 802-810.

Sasaki, Y., Kozaki, A., Ohmori, A., Iguchi, H., and Nagano, Y. (2001). Chloroplast RNA editing required for functional acetyl-CoA carboxylase in plants. J Biol Chem **276,** 3937-3940.

Sasarman, F., Brunel-Guitton, C., Antonicka, H., Wai, T., and Shoubridge, E.A. (2010). LRPPRC and SLIRP interact in a ribonucleoprotein complex that regulates posttranscriptional gene expression in mitochondria. Mol Biol Cell **21,** 1315-1323.

Sass, E., Blachinsky, E., Karniely, S., and Pines, O. (2001). Mitochondrial and cytosolic isoforms of yeast fumarase are derivatives of a single translation product and have identical amino termini. J Biol Chem **276,** 46111-46117.

Sass, E., Karniely, S., and Pines, O. (2003). Folding of fumarase during mitochondrial import determines its dual targeting in yeast. J Biol Chem **278,** 45109-45116.

Sato, S., Nakamura, Y., Kaneko, T., Asamizu, E., and Tabata, S. (1999). Complete structure of the chloroplast genome of Arabidopsis thaliana. DNA Res **6,** 283-290.

Schein, A., Sheffy-Levin, S., Glaser, F., and Schuster, G. (2008). The RNase E/G-type endoribonuclease of higher plants is located in the chloroplast and cleaves RNA similarly to the E. coli enzyme. RNA **14,** 1057-1068.

Schmelzer, C., and Schweyen, R.J. (1986). Self-splicing of group II introns in vitro: mapping of the branch point and mutational inhibition of lariat formation. Cell **46,** 557-565.

Schmitz-Linneweber, C., Williams-Carrier, R., and Barkan, A. (2005). RNA immunoprecipitation and microarray analysis show a chloroplast Pentatricopeptide repeat protein to be associated with the 5' region of mRNAs whose translation it activates. Plant Cell **17,** 2791-2804.

Schmitz-Linneweber, C., Williams-Carrier, R.E., Williams-Voelker, P.M., Kroeger, T.S., Vichas, A., and Barkan, A. (2006). A pentatricopeptide repeat protein facilitates the trans-splicing of the maize chloroplast rps12 pre-mRNA. Plant Cell **18,** 2650-2663.

Schuster, W. (1993). Ribosomal protein gene rpl5 is cotranscribed with the nad3 gene in Oenothera mitochondria. Mol Gen Genet **240,** 445-449.

Schuster, W., Wissinger, B., Unseld, M., and Brennicke, A. (1990). Transcripts of the NADH-dehydrogenase subunit 3 gene are differentially edited in Oenothera mitochondria. EMBO J **9,** 263-269.

Scott, I., and Logan, D.C. (2008). Mitochondria and cell death pathways in plants: Actions speak louder than words. Plant Signal Behav **3,** 475-477.

Shadel, G.S., and Clayton, D.A. (1995). A Saccharomyces cerevisiae mitochondrial transcription factor, sc-mtTFB, shares features with sigma factors but is functionally distinct. Mol Cell Biol **15,** 2101-2108.

Shiina, T., Tsunoyama, Y., Nakahira, Y., and Khan, M.S. (2005). Plastid RNA polymerases, promoters, and transcription regulators in higher plants. Int Rev Cytol **244,** 1-68.

Shikanai, T. (2006). RNA editing in plant organelles: machinery, physiological function and evolution. Cell Mol Life Sci **63,** 698-708.

Shutt, T.E., and Gray, M.W. (2006). Bacteriophage origins of mitochondrial replication and transcription proteins. Trends Genet **22,** 90-95.

Siculella, L., Pacoda, D., Treglia, S., Gallerani, R., and Ceci, L.R. (1996). GTG as translation initiation codon in the apocytochrome b gene of sunflower mitochondria. DNA Seq **6,** 365-369.

Small, I.D., and Peeters, N. (2000). The PPR motif - a TPR-related motif prevalent in plant organellar proteins. Trends Biochem Sci **25,** 46-47.

Sriraman, P., Silhavy, D., and Maliga, P. (1998). Transcription from heterologous rRNA operon promoters in chloroplasts reveals requirement for specific activating factors. Plant Physiol **117,** 1495-1499.

Staudinger, M., Bolle, N., and Kempken, F. (2005). Mitochondrial electroporation and in organello RNA editing of chimeric atp6 transcripts. Mol Genet Genomics **273,** 130-136.

Stegemann, S., Hartmann, S., Ruf, S., and Bock, R. (2003). High-frequency gene transfer from the chloroplast genome to the nucleus. Proc Natl Acad Sci U S A **100,** 8828-8833.

Stern, D.B., Goldschmidt-Clermont, M., and Hanson, M.R. Chloroplast RNA metabolism. Annu Rev Plant Biol **61,** 125-155.

Stern, D.B., Hanson, M.R., and Barkan, A. (2004). Genetics and genomics of chloroplast biogenesis: maize as a model system. Trends Plant Sci **9,** 293-301.

Stern, D.B., and Lonsdale, D.M. (1982). Mitochondrial and chloroplast genomes of maize have a 12-kilobase DNA sequence in common. Nature **299,** 698-702.

Stern, D.B., and Palmer, J.D. (1984). Extensive and widespread homologies between mitochondrial DNA and chloroplast DNA in plants. Proc Natl Acad Sci U S A **81,** 1946-1950.

Strand, A., Asami, T., Alonso, J., Ecker, J.R., and Chory, J. (2003). Chloroplast to nucleus communication triggered by accumulation of Mg-protoporphyrinIX. Nature **421,** 79-83.

Stuart, K.D., Schnaufer, A., Ernst, N.L., and Panigrahi, A.K. (2005). Complex management: RNA editing in trypanosomes. Trends Biochem Sci **30,** 97-105.

Sturm, N.R., Kuras, R., Buschlen, S., Sakamoto, W., Kindle, K.L., Stern, D.B., and Wollman, F.A. (1994). The petD gene is transcribed by functionally redundant promoters in Chlamydomonas reinhardtii chloroplasts. Mol Cell Biol **14,** 6171-6179.

Sugiura, M. (1992). The chloroplast genome. Plant Mol Biol **19,** 149-168.

Sugiura, M., Hirose, T., and Sugita, M. (1998). Evolution and mechanism of translation in chloroplasts. Annu Rev Genet **32,** 437-459.

Susek, R.E., Ausubel, F.M., and Chory, J. (1993). Signal transduction mutants of Arabidopsis uncouple nuclear CAB and RBCS gene expression from chloroplast development. Cell **74,** 787-799.

Suzuki, J.Y., Ytterberg, A.J., Beardslee, T.A., Allison, L.A., Wijk, K.J., and Maliga, P. (2004). Affinity purification of the tobacco plastid RNA polymerase and in vitro reconstitution of the holoenzyme. Plant J **40,** 164-172.

Takenaka, M. MEF9, an E-subclass pentatricopeptide repeat protein, is required for an RNA editing event in the nad7 transcript in mitochondria of Arabidopsis. Plant Physiol **152,** 939-947.

Takenaka, M., and Brennicke, A. (2003). In vitro RNA editing in pea mitochondria requires NTP or dNTP, suggesting involvement of an RNA helicase. J Biol Chem **278,** 47526-47533.

Takenaka, M., and Brennicke, A. (2007). RNA editing in plant mitochondria: assays and biochemical approaches. Methods Enzymol **424,** 439-458.

Takenaka, M., Neuwirt, J., and Brennicke, A. (2004). Complex cis-elements determine an RNA editing site in pea mitochondria. Nucleic Acids Res **32,** 4137-4144.

Takenaka, M., Verbitskiy, D., van der Merwe, J.A., Zehrmann, A., and Brennicke, A. (2008). The process of RNA editing in plant mitochondria. Mitochondrion **8,** 35-46.

Tam, P.P., Barrette-Ng, I.H., Simon, D.M., Tam, M.W., Ang, A.L., and Muench, D.G. The Puf family of RNA-binding proteins in plants: phylogeny, structural modeling, activity and subcellular localization. BMC Plant Biol **10,** 44.

Tasaki, E., Hattori, M., and Sugita, M. The moss pentatricopeptide repeat protein with a DYW domain is responsible for RNA editing of mitochondrial ccmFc transcript. Plant J.

Tillich, M., Hardel, S.L., Kupsch, C., Armbruster, U., Delannoy, E., Gualberto, J.M., Lehwark, P., Leister, D., Small, I.D., and Schmitz-Linneweber, C. (2009). Chloroplast ribonucleoprotein CP31A is required for editing and stability of specific chloroplast mRNAs. Proc Natl Acad Sci U S A **106,** 6002-6007.

Tillich, M., and Krause, K. The ins and outs of editing and splicing of plastid RNAs: lessons from parasitic plants. N Biotechnol.

Timmis, J.N., Ayliffe, M.A., Huang, C.Y., and Martin, W. (2004). Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nat Rev Genet **5,** 123-135.

Tracy, R.L., and Stern, D.B. (1995). Mitochondrial transcription initiation: promoter structures and RNA polymerases. Curr Genet **28,** 205-216.

Tseng, C.C., Sung, T.Y., Li, Y.C., Hsu, S.J., Lin, C.L., and Hsieh, M.H. Editing of accD and ndhF chloroplast transcripts is partially affected in the Arabidopsis vanilla cream1 mutant. Plant Mol Biol **73,** 309-323.

Unseld, M., Marienfeld, J.R., Brandt, P., and Brennicke, A. (1997). The mitochondrial genome of Arabidopsis thaliana contains 57 genes in 366,924 nucleotides. Nat Genet **15,** 57-61.

Uyttewaal, M., Mireau, H., Rurek, M., Hammani, K., Arnal, N., Quadrado, M., and Giege, P. (2008). PPR336 is associated with polysomes in plant mitochondria. J Mol Biol **375,** 626-636.

van der Veen, R., Arnberg, A.C., van der Horst, G., Bonen, L., Tabak, H.F., and Grivell, L.A. (1986). Excised group II introns in yeast mitochondria are lariats and can be formed by self-splicing in vitro. Cell **44,** 225-234.

Verbitskiy, D., van der Merwe, J.A., Zehrmann, A., Brennicke, A., and Takenaka, M. (2008). Multiple specificity recognition motifs enhance plant mitochondrial RNA editing in vitro. J Biol Chem **283,** 24374-24381.

Verbitskiy, D., Zehrmann, A., van der Merwe, J.A., Brennicke, A., and Takenaka, M. (2009). The PPR protein encoded by the LOVASTATIN INSENSITIVE 1 gene is involved in RNA editing at three sites in mitochondria of Arabidopsis thaliana. Plant J.

Vogel, J., and Borner, T. (2002). Lariat formation and a hydrolytic pathway in plant chloroplast group II intron splicing. EMBO J **21,** 3794-3803.

Walter, M., Kilian, J., and Kudla, J. (2002). PNPase activity determines the efficiency of mRNA 3'-end processing, the degradation of tRNA and the extent of polyadenylation in chloroplasts. EMBO J **21,** 6905-6914.

Wang, X., McLachlan, J., Zamore, P.D., and Hall, T.M. (2002). Modular recognition of RNA by a human pumilio-homology domain. Cell **110,** 501-512.

Wang, Z., Zou, Y., Li, X., Zhang, Q., Chen, L., Wu, H., Su, D., Chen, Y., Guo, J., Luo, D., Long, Y., Zhong, Y., and Liu, Y.G. (2006). Cytoplasmic male sterility of rice with boro II cytoplasm is caused by a cytotoxic peptide and is restored by two related PPR motif genes via distinct modes of mRNA silencing. Plant Cell **18,** 676-687.

Wank, H., SanFilippo, J., Singh, R.N., Matsuura, M., and Lambowitz, A.M. (1999). A reverse transcriptase/maturase promotes splicing by binding at its own coding segment in a group II intron RNA. Mol Cell **4,** 239-250.

Watkins, K.P., Kroeger, T.S., Cooke, A.M., Williams-Carrier, R.E., Friso, G., Belcher, S.E., van Wijk, K.J., and Barkan, A. (2007). A ribonuclease III domain protein functions in group II intron splicing in maize chloroplasts. Plant Cell **19,** 2606-2623.

Wharton, R.P., and Aggarwal, A.K. (2006). mRNA regulation by Puf domain proteins. Sci STKE **2006,** pe37.

Williams, M.A., Johzuka, Y., and Mulligan, R.M. (2000). Addition of non-genomically encoded nucleotides to the 3'-terminus of maize mitochondrial mRNAs: truncated rps12 mRNAs frequently terminate with CCA. Nucleic Acids Res **28,** 4444-4451.

Williams, P.M., and Barkan, A. (2003). A chloroplast-localized PPR protein required for plastid ribosome accumulation. Plant J **36,** 675-686.

Williams-Carrier, R., Kroeger, T., and Barkan, A. (2008). Sequence-specific binding of a chloroplast pentatricopeptide repeat protein to its native group II intron ligand. RNA **14,** 1930-1941.

Wintz, H., and Hanson, M.R. (1991). A termination codon is created by RNA editing in the petunia atp9 transcript. Curr Genet **19,** 61-64.

Wolf, P.G., Rowe, C.A., and Hasebe, M. (2004). High levels of RNA editing in a vascular plant chloroplast genome: analysis of transcripts from the fern Adiantum capillus-veneris. Gene **339,** 89-97.

Woodson, J.D., and Chory, J. (2008). Coordination of gene expression between organellar and nuclear genomes. Nat Rev Genet **9,** 383-395.

Wostrikoff, K., and Stern, D. (2007). Rubisco large-subunit translation is autoregulated in response to its assembly state in tobacco chloroplasts. Proc Natl Acad Sci U S A **104,** 6466-6471.

Xu, F., Ackerley, C., Maj, M.C., Addis, J.B., Levandovskiy, V., Lee, J., Mackay, N., Cameron, J.M., and Robinson, B.H. (2008). Disruption of a mitochondrial RNA-binding protein gene results in decreased cytochrome b expression and a marked reduction in ubiquinol-cytochrome c reductase activity in mouse heart mitochondria. Biochem J **416,** 15-26.

Xu, F., Morin, C., Mitchell, G., Ackerley, C., and Robinson, B.H. (2004). The role of the LRPPRC (leucine-rich pentatricopeptide repeat cassette) gene in cytochrome oxidase assembly: mutation causes lowered levels of COX (cytochrome c oxidase) I and COX III mRNA. Biochem J **382,** 331-336.

Yamazaki, H., Tasaka, M., and Shikanai, T. (2004). PPR motifs of the nucleus-encoded factor, PGR3, function in the selective and distinct steps of chloroplast gene expression in Arabidopsis. Plant J **38,** 152-163.

Yoo, H.H., Kwon, C., Lee, M.M., and Chung, I.K. (2007). Single-stranded DNA binding factor AtWHY1 modulates telomere length homeostasis in Arabidopsis. Plant J **49,** 442-451.

Yu, Q.B., Jiang, Y., Chong, K., and Yang, Z.N. (2009). AtECB2, a pentatricopeptide repeat protein, is required for chloroplast transcript accD RNA editing and early chloroplast biogenesis in Arabidopsis thaliana. Plant J **59,** 1011-1023.

Yu, W., Fester, T., Bock, H., and Schuster, W. (1995). RNA editing in higher plant mitochondria: analysis of biochemistry and specificity. Biochimie **77,** 79-86.

Zehrmann, A., van der Merwe, J.A., Verbitskiy, D., Brennicke, A., and Takenaka, M. (2008). Seven large variations in the extent of RNA editing in plant mitochondria between three ecotypes of Arabidopsis thaliana. Mitochondrion **8,** 319-327.

Zehrmann, A., Verbitskiy, D., van der Merwe, J.A., Brennicke, A., and Takenaka, M. (2009). A DYW domain-containing pentatricopeptide repeat protein is required for RNA editing at multiple sites in mitochondria of Arabidopsis thaliana. Plant Cell **21,** 558-567.

Zhou, W., Cheng, Y., Yap, A., Chateigner-Boutin, A.L., Delannoy, E., Hammani, K., Small, I., and Huang, J. (2008). The Arabidopsis gene YS1 encoding a DYW protein is required for editing of rpoB transcripts and the rapid development of chloroplasts during early growth. Plant J.

Zoschke, R., Nakamura, M., Liere, K., Sugiura, M., Borner, T., and Schmitz-Linneweber, C. An organellar maturase associates with multiple group II introns. Proc Natl Acad Sci U S A **107,** 3245-3250.