

**ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTE**

**Institut de biologie moléculaire des plantes – UPR2357 - CNRS**

**THÈSE** présentée par :

**Nicolas CHEVIGNY**

soutenue le : 14 juin 2019

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Science du vivant/ Aspect moléculaire et cellulaire de biologie

**Caractérisation des fonctions de TRCF et de RADA dans l'expression et la maintenance des génomes des organelles chez *Arabidopsis thaliana***

**THÈSE dirigée par :**

**Dr. GUALBERTO José**

Directeur de thèse, université de Strasbourg - CNRS

**RAPPORTEURS :**

[Dr. LURIN Claire]

[Dr. POLARD Patrice]

Rapporteuse externe, université Paris-Saclay - INRA

Rapporteur externe, université Toulouse III - CNRS

---

**AUTRES MEMBRES DU JURY :**

[Pr. KUHN Kristina]

[Dr. BLEVINS Todd]

[Dr. DIETRICH André]

Examinatrice externe, université de Halle-Wittenberg

Examinateur interne, université de Strasbourg - CNRS

Examinateur interne, université de Strasbourg - CNRS

## Remerciements

Tout d'abord, je tiens à remercier les docteurs Claire Lurin, Patrice Polard, Todd Blevins et la professeure Kristina Kühn d'avoir accepté d'évaluer et de discuter le travail que j'ai effectué durant ma thèse.

Je remercie mes deux directeurs de thèse de m'avoir accueilli et fait confiance durant ces quatre années. Vous m'avez fait découvrir le fabuleux monde des organelles de plante dont je ne soupçonnais rien. Je remercie José Gualberto de m'avoir aiguillé tout au long de ma thèse. Merci pour nos discussions, scientifiques ou non, toujours enrichissantes et merci pour ton aide lors de la rédaction. Je remercie André Dietrich pour ses connaissances exhaustives en biologie des organelles et en jeux de mots.

Je remercie le Labex MitoCross et l'Université de Strasbourg d'avoir financé ma thèse et mes missions d'enseignement. Merci à tous mes collègues de biochimie pour ces trois chouettes années d'enseignement. Merci aux étudiants d'être ce qu'ils sont.

Je remercie tous les membres du labo 312, actuels comme anciens. Merci à Clémentine d'avoir débroussaillé le sujet TRCF et de m'avoir initié aux techniques secrètes du laboratoire. Merci à Frédérique pour ton aide lors de la rédaction de la thèse. Merci à Déborah de m'avoir aidé en me déchargeant de nombreuses expériences, souvent laborieuses. Merci aussi pour tes blagues novatrices et ta répartie légendaire et implacable. Merci à Arnaud, Pierre et Déborah (encore) pour la bonne ambiance du box. Merci à Mathilde et Marjorie d'avoir contribué à mes travaux.

Je remercie également toutes les personnes qui se sont impliquées dans mes projets de thèse. Merci à Lauriane, Johana et Philippe pour l'analyse par spectrométrie de masse et pour vos conseils. Merci à Hélène Scheer pour m'avoir montré comment faire de bonnes Co-IP. Merci à Marc Bergdoll pour la modélisation de la structure de TRCF. Merci à Nicolas Baumberger et Laurence Herrgott pour la purification des protéines TRCF et RADA. Merci à Marc Bichara de m'avoir accueilli quelques semaines à l'ESBS pour réaliser les complémentations bactériennes. Merci à Malek et à Sandrine pour ces centaines de séquençages que vous avez fait pour moi. Merci à Corinne Keichinger de m'avoir aidé et conseillé pour le double hybride de levure. Merci à toute l'équipe de production de plante. Merci au labo 409 et à Claudine Bleykasten de m'avoir donné envie de faire de la recherche. Enfin, merci aussi à tous les membres de l'IBMP pour leurs conseils, coups de pouce et discussions.

Je remercie aussi mes proches et ma famille qui m'ont soutenu durant toute ma thèse, chacun à sa manière. Merci à mes autres financeurs, mes parents, d'avoir tant investi pour que je puisse me débrouiller seul. Merci Maman d'avoir cru en moi durant toutes ces années et cela malgré mes échecs. C'est surtout grâce à toi que j'en suis arrivé là. Merci Sarah de m'aider à me motiver chaque mercredi et de m'écouter. Merci Bruno pour nos discussions improbables et pour tes précieuses corrections. Et par-dessus tout, merci Iza d'avoir toujours été là pour moi et de m'avoir tant aidé. Merci pour tout le temps que tu m'as donné. Je te dois beaucoup.

## Abréviations

aaRS	aminoacyl-ARNt synthetase ( <i>aminoacyl tRNA synthetase</i> )
ADN	acide désoxyribonucléide
ADNc	ADN complémentaire
ADNcp	ADN chloroplastique
ADNdb	ADN double brin
ADNmt	ADN mitochondrial
ADNsb	ADN simple brin
ADP	adénosine diphosphate
AP	apurinique/apyrimidique ( <i>apurinic/apyrimidinic</i> )
APS	persulfate d'ammonium
ARN	acide ribonucléide
ARNt	ARN de transfert
ARNr	ARN ribosomique
ATP	adénosine triphosphate
BER	réparation par excision de base ( <i>base excision repair</i> )
BET	bromure d'éthidium
BiFC	complémentation bimoléculaire de la fluorescence
BIR	réplication induite par cassure ( <i>break induced replication</i> )
BSA	sérum d'albumine bovin ( <i>bovin serum albumin</i> )
CaMV	virus de la mosaïque du chou-fleur ( <i>cauliflower mosaic virus</i> )
Chi	<i>crossover hotspot instigator</i>
Col-0	écotype Columbia
CTAB	<i>cetyl-trimethylammonium bromide</i>
ddNTP	didésoxribonucléotide
D-loop	<i>displacement loop</i>
DO	densité optique
DOC	déoxycholate de sodium
DSB	cassure double brin ( <i>double-stranded break</i> )
DSBR	réparation des cassures double brin ( <i>double-stranded breakrepair</i> )
DTT	dithiothreitol
ECL	<i>enhanced chemical luminescence</i>
EDTA	éthylène diamine tétra-acétate de sodium
EMSA	retard sur gel ( <i>electro mobility shift assay</i> )
GFP	<i>green fluorescent protein</i>
GG-NER	<i>global genomic NER</i>
HA	hémagglutinine
HH	<i>hook helix</i>
HJ	jonction de Holliday ( <i>Holliday junction</i> )
HR	recombinaison homologue ( <i>homologous recombination</i> )
HSP	<i>heat shock protein</i>
IPTG	Isopropyl $\beta$ -D-1thiogalactopyranoside



IR(s)	séquence(s) répétée(s) de taille intermédiaire ( <i>inverted repeat(s)</i> )
KO	<i>knock-out</i>
LB	milieu de Luria Bertani (ou <i>lysogeny broth</i> )
LECA	dernier ancêtre commun aux eucaryotes ( <i>last eukaryotic common ancestor</i> )
LSC	<i>large single copy</i>
MDS	<i>multidimensional scaling</i>
Mfd	<i>mutation frequency decline</i>
miARN	micro-ARN
min	minute
MMBIR	<i>microhomology-mediated BIR</i>
MMEJ	<i>microhomology-mediated end-joining</i>
MMF	<i>massive mitochondrial fusion</i>
MMR	réparation des mésappariement ( <i>mismatch repair</i> )
MMS	<i>methyl methanesulfonate</i>
MS	milieu Murashige et Skoog
NADH	nicotinamide adénine dinucléotide
NASC	Nottingham Arabidopsis Stock Center
NEP	polymérase codée par le noyau ( <i>nuclear encoded polymerase</i> )
NER	réparation par excision de nucléotide ( <i>nucleotide excision repair</i> )
NHEJ	<i>non-homologous end-joining</i>
NLS	signal d'adressage nucléaire ( <i>nuclear localization signal</i> )
nt(s)	nucléotide(s)
ODB	<i>organellar DNA-binding protein</i>
OSB	<i>organellar single-stranded DNA-binding protein</i>
OTS	<i>signal d'adressage organellaire (organellar targeting signal)</i>
PAM	<i>presequence translocaste-associated motor</i>
PAP	protéine associée à la PEP ( <i>PEP associated protein</i> )
PCOR	protochlorophyllide oxidoréductase
PCR	réaction en chaîne par polymérase ( <i>polymerase chain reaction</i> )
PEG	polyéthylène glycol
PEP	polymérase codée par le plaste ( <i>plastidial encoded polymerase</i> )
Pi	phosphate inorganique
PMSF	<i>phenylmethylsulfonyl fluoride</i>
PPR	<i>pentatricopeptide repeat</i>
PRORP	<i>proteinaceous RNase P</i>
PSI	photosystème I
PSII	photosystème II
pTAC	protéine des TAC
PVDF	polyvinylidene fluoride
PVP	polyvinylpyrrolidone
qPCR	<i>PCR quantitative</i>
RH	<i>relay helix</i>
RID	domaine d'interaction avec la RNAP ( <i>RNA polymerase interacting domain</i> )

RNAP	ARN polymérase ( <i>RNA polymerase</i> )
ROS	espèce réactive de l'oxygène ( <i>reactive oxygen specie</i> )
rpm	rotation par minute
RPOT	<i>RNA polymerase of the T3/T7 type</i>
RPOTm	RPOT mitochondriale
RPOTmp	RPOT mitochondriale et plastidiale
RPOTp	RPOT plastidiale
Rubisco	ribulose biphosphate carboxylase
s	seconde
SD	<i>synthetic defined</i>
SDS	<i>sodium dodécyl sulphate</i>
SDSA	<i>synthesis-dependant strand-annealing</i>
sms	<i>sensitivity to MMS</i>
SSA	<i>single-stranded annealing</i>
SSB	<i>protéine de liaison à l'ADNsb (single-stranded binding protein)</i>
SSC	<i>small single copy</i>
SSG	lacune simple brin ( <i>single-stranded gap</i> )
SSS	shift stœchiométrique ( <i>stoichiometric shift</i> )
SSU	petite sous unité de la Rubisco ( <i>small subunit</i> )
SV40	virus simien 40 ( <i>simian virus 40</i> )
TAC	<i>transcriptionally active chromosome</i>
TAE	Tris acétate EDTA
TBS	<i>Tris-buffered saline</i>
TBSV	<i>tomato bushy stunt virus</i>
TC-NER	<i>transcription-coupled NER</i>
TCA	acide trichloracétique
TEMED	tétraméthylethylènediamine
TIC	<i>translocon of the inner chloroplast membrane</i>
TIM	<i>translocon of the inner mitochondrion membrane</i>
Tm	température de fusion ( <i>melting temperature</i> )
TOC	<i>translocon of the outer chloroplast membrane</i>
TOM	<i>translocon of the outer mitochondrion membrane</i>
TPR	<i>tetratricopeptide repeat</i>
TRCF	<i>transcription repair coupling factor</i>
TRG	<i>translocation in RecG</i>
UDG	uracil-DNA <i>glycosylase</i>
UV	ultraviolet
VDAC	<i>voltage-dependent anion channel</i>
X-Gal	5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactopyranoside
Y2H	<i>double hybride de levure (yeast two hybrids)</i>
YFP	<i>yellow fluorescent protein</i>
YPD	<i>yeast extract peptone dextrose</i>

# Table des matières

<b>TABLE DES MATIERES .....</b>	<b>1</b>
<b>INTRODUCTION GÉNÉRALE.....</b>	<b>6</b>
1. <i>L'origine des organelles .....</i>	7
2. <i>La mitochondrie et le chloroplaste .....</i>	10
2.1. Structure et dynamique .....	10
2.2. Les complexes de transport d'électrons.....	12
2.3. Import de macromolécules dans les organelles .....	13
3. <i>Les génomes organellaires.....</i>	16
4. <i>Les mécanismes de recombinaison.....</i>	20
4.1. La recombinaison homologue procaryotique.....	21
4.2. La recombinaison homologue eucaryotique .....	23
4.3. La recombinaison homologue organellaire .....	25
4.4. Évolution de la structure des génomes organellaires des plantes .....	29
5. <i>Les mécanismes de réparation de l'ADN organellaire .....</i>	31
5.1. La réparation par excision de base .....	32
5.2. La réparation des mésappariements .....	32
5.3. La réversion directe des dommages par l'activité photolyase.....	33
5.4. La réparation par excision de nucléotide .....	34
6. <i>Le contrôle de l'expression des gènes dans les organelles .....</i>	35
6.1. Les ARN polymérases organellaires .....	35
6.2. La transcription organellaire .....	37
<b>OBJECTIFS DE MA THESE .....</b>	<b>39</b>
<b>CHAPITRE 1 : ÉTUDE DE RADA .....</b>	<b>40</b>
INTRODUCTION.....	41
1. <i>Structure de RadA.....</i>	41
2. <i>Rôle de RadA.....</i>	42
RESULTATS .....	43
1. <i>Résumé de la publication.....</i>	43
2. <i>Article .....</i>	44
<b>CHAPITRE 2 : ÉTUDE DE TRCF.....</b>	<b>70</b>
INTRODUCTION.....	71
1. <i>Structure de TRCF bactérienne.....</i>	71

2. Rôles de TRCF.....	72
RESULTATS .....	74
1. Localisation subcellulaire de TRCF.....	74
2. Étude des séquences et modélisation de la structure de TRCF.....	75
2.1. Analyse de l'alignement des séquences protéiques de TRCF .....	75
2.2. Modélisation de la structure de TRCF d'Arabidopsis.....	75
3. Production d'anticorps reconnaissant la partie N-terminale de TRCF .....	76
4. Caractérisation de mutants par insertion de T-DNA.....	77
5. Sensibilité des mutants à différents stress génotoxiques.....	78
6. Sensibilité du mutant <i>trcf-4</i> à d'autres stress abiotiques non génotoxiques.....	79
7. Analyse de l'expression de TRCF.....	80
7.1. Localisation tissulaire de l'expression du promoteur de TRCF.....	80
7.2. Régulation par la lumière du jour .....	81
8. Recherche des partenaires d'interaction de TRCF .....	82
8.1. Etude de l'interaction entre TRCF et les ARN polymérase organellaires.....	82
8.2. Recherche des partenaires d'interaction par Co-IP et spectrométrie de masse .....	83
8.3. Test des interactions entre TRCF et la PEP par BiFC.....	90
9. Caractérisation des phénotypes des mutants complémentés.....	91
10. Étude des fonctions de TRCF dans la transcription chloroplastique.....	92
10.1. Impact de la surexpression de TRCF:HA sur la transcription .....	92
10.2. Augmentation de la transcription en absence de TRCF.....	93
DISCUSSION ET PERSPECTIVES .....	94
<b>MATERIEL ET MÉTHODES .....</b>	<b>99</b>
MATERIEL .....	100
1. Matériel bactérien.....	100
1.1. Agrobactérium tumefaciens GV3101 pMP90 .....	100
1.2. Escherichia coli BW25113 .....	100
1.3. Escherichia coli DH5 $\alpha$ .....	100
1.4. Escherichia coli TOP10 .....	100
1.5. Escherichia coli JW4352 .....	101
1.6. Escherichia coli Rosetta 2 (DE3) pLysS .....	101
1.7. Escherichia coli DB3.1 .....	101
2. Levures.....	101
2.1. Saccharomyces cerevisiae AH109.....	101
2.2. Saccharomyces cerevisiae Y187.....	102
3. Matériel végétal.....	102
3.1. Arabidopsis thaliana écotype Columbia.....	102

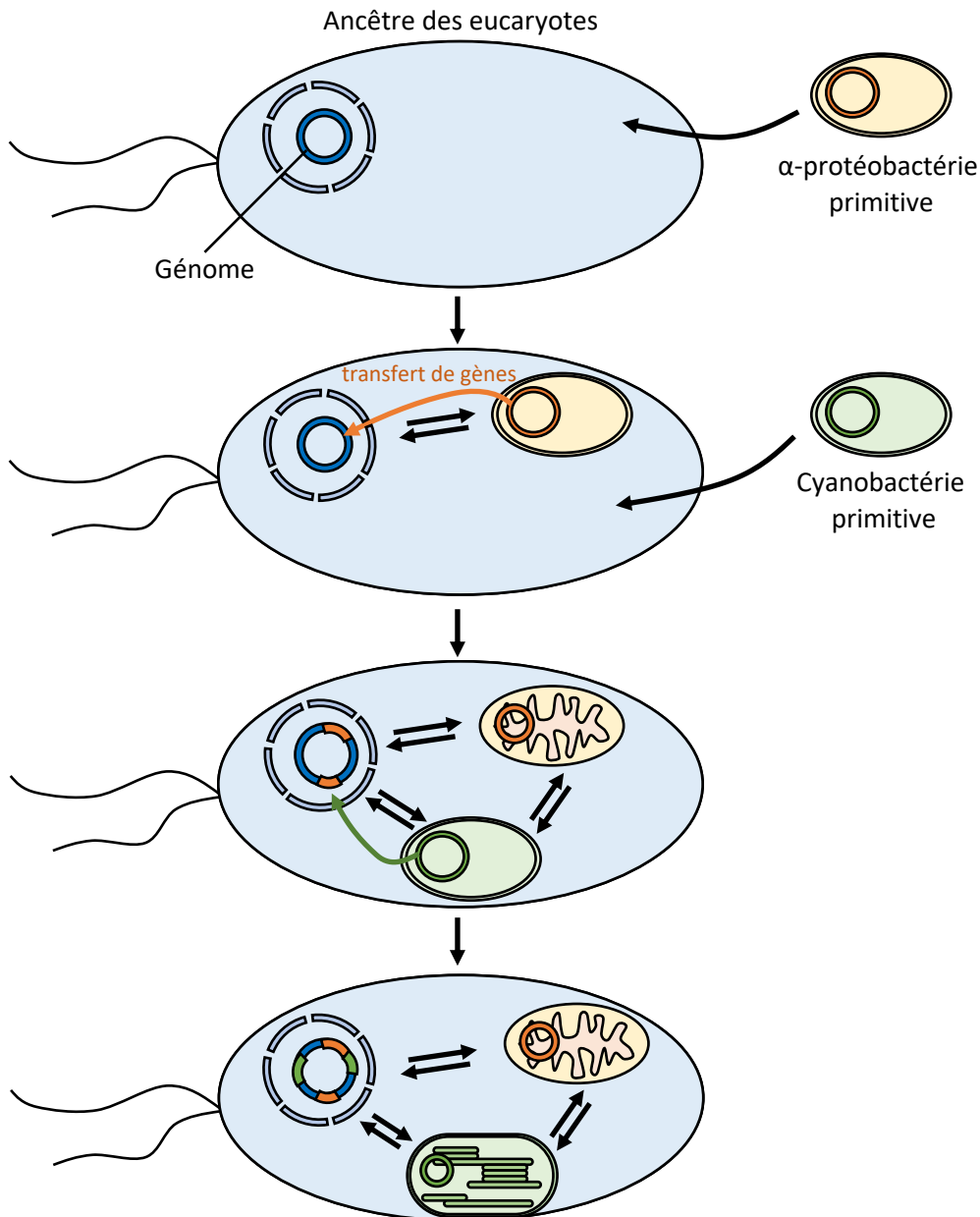
3.2.	Nicotiana benthamiana.....	102
4.	Milieux de cultures .....	103
4.1.	Milieu LB .....	103
4.2.	Milieu M9.....	103
4.3.	Milieu YPD.....	103
4.4.	Milieu SD.....	103
4.5.	Milieu MS solide .....	103
5.	Vecteurs plasmidiques.....	104
5.1.	pUCAP35S .....	104
5.2.	pGEM-T Easy .....	104
5.3.	pBluescript II SK+ .....	104
5.4.	pDONR/Zeo .....	104
5.5.	pGBW13 et pGWB14 .....	105
5.6.	pGADT7 et pGADgwT7.....	105
5.7.	pGBKT7, pGBKgwT7 et pGCKCg .....	105
5.8.	pACYC184LacZ.....	106
5.9.	pET28A.....	106
5.10.	pGW-YN .....	106
5.11.	pGW-YC.....	107
5.12.	pMDC162 .....	107
6.	Gène de synthèse .....	107
7.	Oligonucléotides.....	107
8.	Outils informatiques.....	107
METHODES	.....	108
1.	Préparation de bactéries électro-compétentes .....	108
2.	Préparation de levures chimio-compétentes.....	108
3.	Test de complémentation du mutant bactérien radA par RADA .....	108
3.1.	Complémentation du mutant bactérien .....	108
3.2.	Sensibilité aux UV .....	109
3.3.	Sensibilité à la CIP.....	109
4.	Méthodes relatives aux acides nucléiques.....	109
4.1.	Quantification .....	109
4.2.	Électrophorèse sur gel d'agarose .....	109
5.	Méthodes relatives à l'ARN.....	110
5.1.	Extraction d'ARN .....	110
5.2.	Traitement à la DNase .....	110
5.3.	Réaction de transcription inverse.....	110
6.	Méthodes relatives à l'ADN.....	111

6.1.	Amplification par PCR .....	111
6.2.	Techniques de clonage.....	111
6.3.	Transformation des bactéries.....	114
6.4.	Transformation de levure par choc thermique.....	115
6.5.	Préparation d'ADN plasmidique .....	115
6.6.	Séquençage.....	115
6.7.	Extraction d'ADN total de plante.....	116
6.8.	Extraction de l'ADN chromosomique bactérien .....	116
6.9.	PCR quantitative.....	116
7.	<i>Méthodes relatives aux protéines</i> .....	117
7.1.	Extraction de protéines à partir de matériel végétal .....	117
7.2.	Extraction de protéines à partir de bactérie.....	117
7.3.	Extraction de protéines à partir de levure.....	117
7.4.	Précipitation des protéines .....	117
7.5.	Séparation par électrophorèse sur gel de polyacrylamide en conditions dénaturantes .....	118
7.6.	Quantification .....	118
7.7.	Surproduction dans E. coli.....	119
7.8.	Transfert semi-sec sur membrane de PVDF .....	120
7.9.	Immunodétection.....	121
7.10.	Co-immunoprécipitation de partenaires protéiques .....	121
7.11.	Etude de l'activation de la transcription par double hybride de levure .....	122
7.12.	Test d'échange de brin in vitro .....	122
7.13.	Test d'interaction avec l'ADN par retard sur gel.....	123
8.	<i>Méthodes relatives aux plantes</i> .....	124
8.1.	Transformation par la méthode biolistique.....	124
8.2.	Transformation transitoire par agro-infiltration.....	124
8.3.	Transformation par floral dip .....	125
8.4.	Sélection des transformants obtenus par floral dip.....	125
8.5.	Croisements .....	126
8.6.	Culture de plantes en milieu in vitro.....	126
8.7.	Fusion d'un promoteur avec le gène rapporteur GUS.....	126
8.8.	Test de sensibilité aux UV-C .....	127
8.9.	Test de sensibilité au NaCl, à la roténone et à la lincomycine .....	127
8.10.	Test de sensibilité au H <sub>2</sub> O <sub>2</sub> .....	127
8.11.	Complémentation bimoléculaire de la fluorescence (BiFC).....	127
8.12.	Observation du développement des embryons de plantes fertilisées .....	128
8.13.	Observation de la forme des cellules et de l'état de pollinisation des fleurs.....	128
8.14.	Mesure de la chlorophylle.....	128

9. Méthodes relatives aux organelles.....	128
9.1. Extraction de mitochondrie à partir de plantules étiolées .....	128
9.2. Extraction de chloroplaste à partir de plantes et plantules.....	129
<b>BIBLIOGRAPHIE .....</b>	<b>130</b>
<b>ANNEXES .....</b>	<b>148</b>

# **INTRODUCTION GÉNÉRALE**





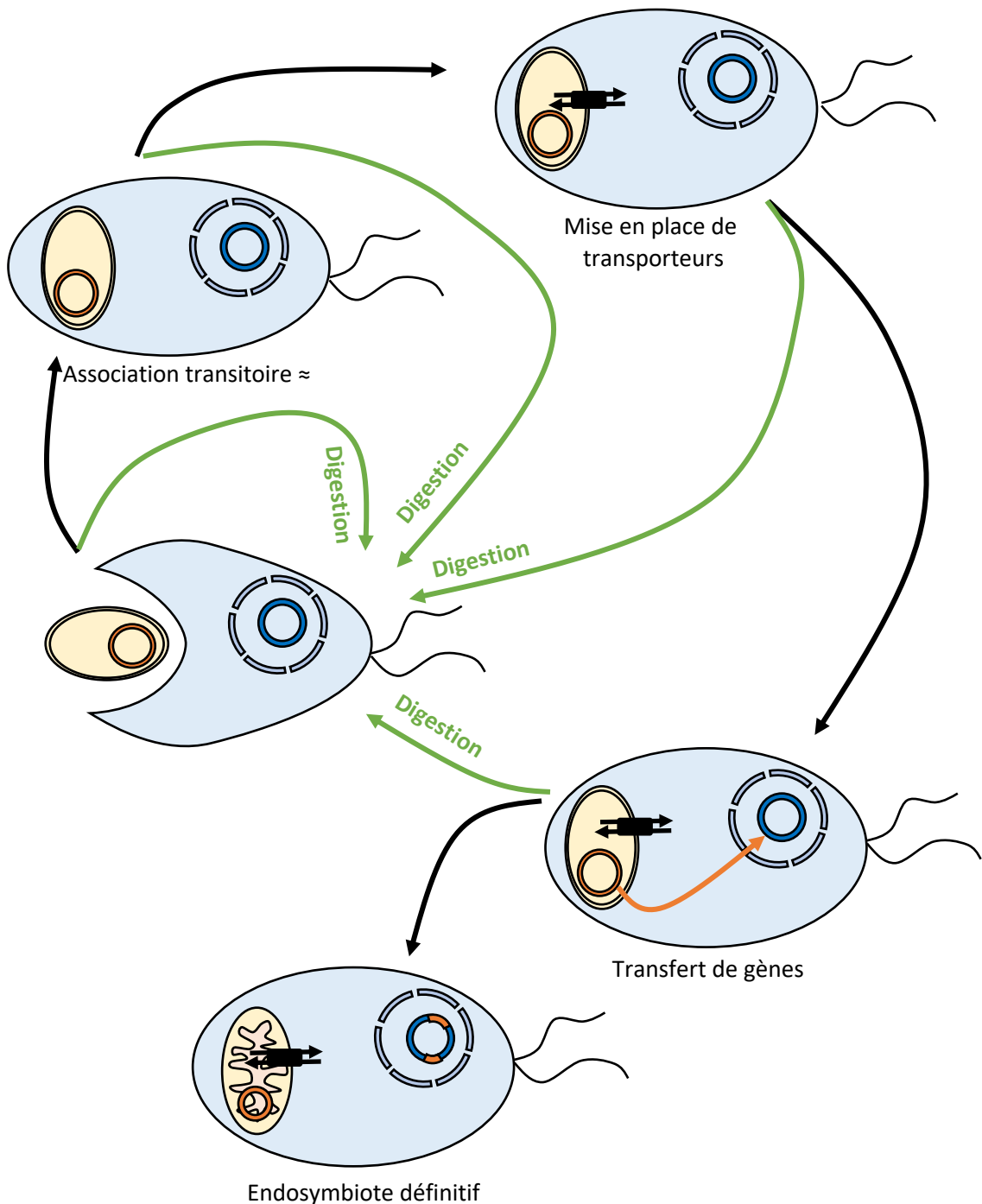
**Figure 1 : Origine endosymbiotique de la mitochondrie et du chloroplaste.** La mitochondrie a été acquise initialement et provient de l'interaction entre une  $\alpha$ -protéobactérie primitive et une cellule proto-eucaryotique ou archéenne. Le chloroplaste dérive d'une cyanobactérie primitive, endocytée par une cellule eucaryotique possédant déjà une mitochondrie plus ou moins développée. Des échanges se sont alors mis en place entre la cellule hôte et ses endosymbiotes, conduisant à la coévolution des trois organismes. Peu à peu, des gènes provenant des endosymbiotes ont migré vers le génome nucléaire, permettant une meilleure coordination des processus cellulaires entre l'hôte et ses symbiotes. Le génome nucléaire présente ainsi de nombreuses séquences d'origine procaryotique. Les deux endosymbiotes ont finalement évolué en organelles, complètement dépendants de la cellule hôte pour toutes leurs fonctions, mais possédant toujours un génome codant un nombre restreint de facteurs.

## 1. L'origine des organelles

La mitochondrie est un organite cytoplasmique présent chez presque tous les eucaryotes. Elle joue un rôle important dans le métabolisme énergétique de la cellule *via* la production d'ATP (adénosine triphosphate) lors de la phosphorylation oxydative. Elle est également impliquée dans la synthèse des phospholipides et de l'hème, dans l'homéostasie du calcium, la signalisation cellulaire, la mort cellulaire programmée et la formation des *clusters* fer-soufre (Mackenzie et McIntosh, 1999; Sousa *et al.*, 2018). Le chloroplaste est aussi un organite cytoplasmique qui n'est retrouvé que chez les cellules végétales et les algues. Il permet la production d'ATP et de sucre à partir du CO<sub>2</sub> inorganique *via* la photosynthèse ainsi que la biosynthèse des pigments, de certains acides aminés et lipides, et la formation des *clusters* fer-soufre (Pyke, 1999; Johnson, 2016).

Ces deux organelles sont entourées par une double membrane lipidique et possèdent leurs propres génomes et systèmes de transcription et traduction, permettant l'expression de protéines impliquées dans les chaînes de transport d'électrons. Ces caractéristiques particulières sont aujourd'hui largement acceptées comme l'une des preuves de l'origine endosymbiotique de la mitochondrie et du chloroplaste (Sagan, 1967). Selon la théorie endosymbiotique (**figure 1**), la mitochondrie a été acquise il y a plus de 1,5 milliard d'années par l'association d'une  $\alpha$ -protéobactérie primitive avec un micro-organisme ancestral. Le chloroplaste aura ensuite été formé par l'association de l'eucaryote ancestral possédant une mitochondrie (dérivant de l' $\alpha$ -protéobactérie) avec une cyanobactérie primitive. Mais si la communauté scientifique s'accorde sur l'origine bactérienne de ces organelles, de nombreuses zones d'ombre persistent autour du *timing* et du mécanisme d'acquisition de la mitochondrie et de l'identité de l'hôte l'ayant acquise. De nombreuses théories coexistent pour tenter d'expliquer ces différents points, mais aucune n'est unanimement acceptée.

Selon la théorie dite *mito-early*, l'acquisition de la mitochondrie est l'une des premières étapes de l'évolution des eucaryotes. L' $\alpha$ -protéobactérie primitive se serait associée avec une archée primitive (Poole et Gribaldo, 2014), donnant ainsi naissance à la branche évolutive des eucaryotes. Les possibilités énergétiques conférées par l'association des deux organismes auraient ensuite permis la mise en place des autres traits caractéristiques des eucaryotes. Selon la théorie opposée, dite *mito-late*, la mitochondrie aurait été acquise par une cellule proto-eucaryotique possédant déjà plusieurs traits caractéristiques des eucaryotes, dont la phagocytose nécessaire à l'absorption de l' $\alpha$ -protéobactérie (Pittis et Gabaldon, 2016). Ces deux théories s'accordent néanmoins sur le fait que le dernier ancêtre commun à tous les eucaryotes (*last eukaryotic common ancestor*: LECA) possédait une mitochondrie. Elles s'accordent également sur la proximité de la cellule hôte avec les archées, les eucaryotes partageant un ancêtre commun plus proche avec les archées qu'avec les eubactéries (Eme *et*

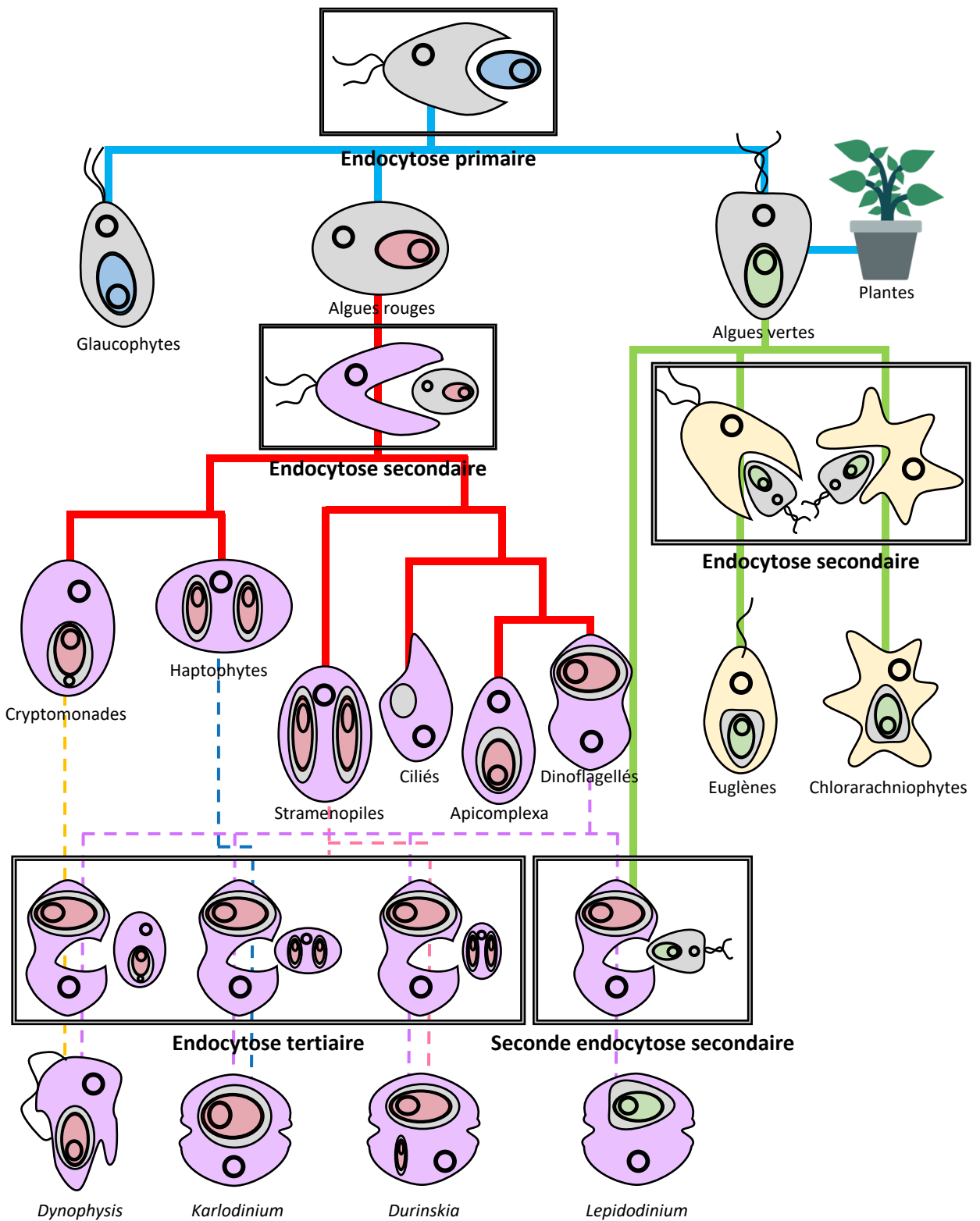


**Figure 2 : Modèle hypothétique du *targeting ratchet*.** Dans le cas où l'endosymbiose proviendrait d'un acte de prédation, il paraît probable que de nombreux organismes aient été endocytés avant la mise en place de l'endosymbiote définitif. Dans ce modèle, les éléments nécessaires à la mise en place de l'endosymbiose se sont mis en place petit à petit. Les bactéries endocytées auraient d'abord été conservées plus longtemps dans le cytoplasme de l'hôte avant digestion, probablement en apportant un avantage à l'hôte. Les interactions entre les deux organismes auraient ensuite été favorisées par l'adressage de canaux et transporteurs aux membranes des bactéries endocytées. Certains gènes procaryotiques auraient ensuite migré dans le génome nucléaire, expliquant la présence de nombreuses séquences bactériennes ne provenant pas d' $\alpha$ -protéobactérie ni de cyanobactérie dans les génomes eucaryotiques actuels. Les relations symbiotiques entre les deux organismes auraient conduit au maintien du symbiote, évoluant ensuite en organelle. Un modèle similaire peut être appliqué pour l'endosymbiose avec le chloroplaste, ou pour les endosymbioses multiples. D'après Keeling, 2013.

*al.*, 2017). Malgré leurs divergences, il faut noter que tous les points des deux théories ne sont pas mutuellement exclusifs. L'absorption de l' $\alpha$ -protéobactérie ancestrale par phagocytose, par exemple, est tout à fait envisageable pour des archées ayant acquis ce trait (Poole et Gribaldo, 2014). Ces archées auraient ensuite évolué pour devenir les eucaryotes. Néanmoins, aucune archée présentant un endosymbiote bactérien n'a été trouvée à ce jour. Dans le cas d'une phagocytose, la membrane du phagosome aurait néanmoins été perdue par la suite. Une autre hypothèse serait que l'hôte proto-eucaryotique ou archéal et l' $\alpha$ -protéobactérie primitive auraient pu être impliqués dans une relation de syntrophie. L'interdépendance des deux organismes aurait ensuite mené à l'englobement de la bactérie par son partenaire menant ainsi à l'endosymbiose (Moreira et Lopez-Garcia, 1998). Mais de telles endosymbioses n'ont jamais été observées dans la nature chez d'autres organismes. Enfin une autre hypothèse implique un événement de prédation ou de parasitisme dans la formation de la mitochondrie (Davidov et Jurkevitch, 2009). L' $\alpha$ -protéobactérie primitive se serait infiltrée dans le cytoplasme de son hôte comme parasite. Le mécanisme aurait cependant échoué et le parasite aurait été conservé dans la cellule hôte, sans la détruire, jusqu'à la mise en place d'une relation symbiotique.

L'endosymbiose implique une capacité d'interaction entre les deux organismes et nécessite donc la présence de canaux et de transporteurs à travers les membranes de l'endosymbiote et de signaux d'adressage pour les facteurs de l'hôte. Ces mécanismes ont pu être développés par l'hôte lors de l'association avec d'autres organismes précédant celle avec la proto-mitochondrie. La proto-mitochondrie étant seulement le dernier endosymbiote retenu par l'hôte. Il est possible d'imaginer que les premières interactions entre l'hôte et ses partenaires étaient de type prédateur (phagocytose). Le partenaire aurait pu être retenu de plus en plus longtemps chez son hôte avant digestion, permettant ainsi la mise en place de mécanismes d'interaction profitables aux deux organismes (**figure 2**) (Keeling, 2013; Gabaldon, 2018). Cette hypothèse permet également d'expliquer la présence de nombreux gènes dans le génome nucléaire d'origine procaryotique ne provenant pas de l'ancêtre de la mitochondrie (ou du chloroplaste), (Gabaldon, 2018).

Après l'engloutissement de l' $\alpha$ -protéobactérie primitive par son hôte, le génome de l'endosymbiote s'est peu à peu réduit. Les gènes nécessaires à la vie sous forme libre et aux fonctions déjà supportées par l'hôte n'étant plus nécessaires, ils ont été perdus (Adams et Palmer, 2003). De nombreux gènes ont également migré vers le génome de l'hôte pour permettre une meilleure coordination des activités cellulaires tout en éloignant l'ADN de la chaîne mitochondriale de transport d'électrons, source d'espèces réactives de l'oxygène (ROS) pouvant endommager le génome (Allen et Raven, 1996). Dans la majorité des cas, seuls les gènes permettant l'expression d'ARN de transfert (ARNt), d'ARN ribosomiques (ARNr) et de protéines impliquées dans les complexes de respiration ont été conservés dans le génome

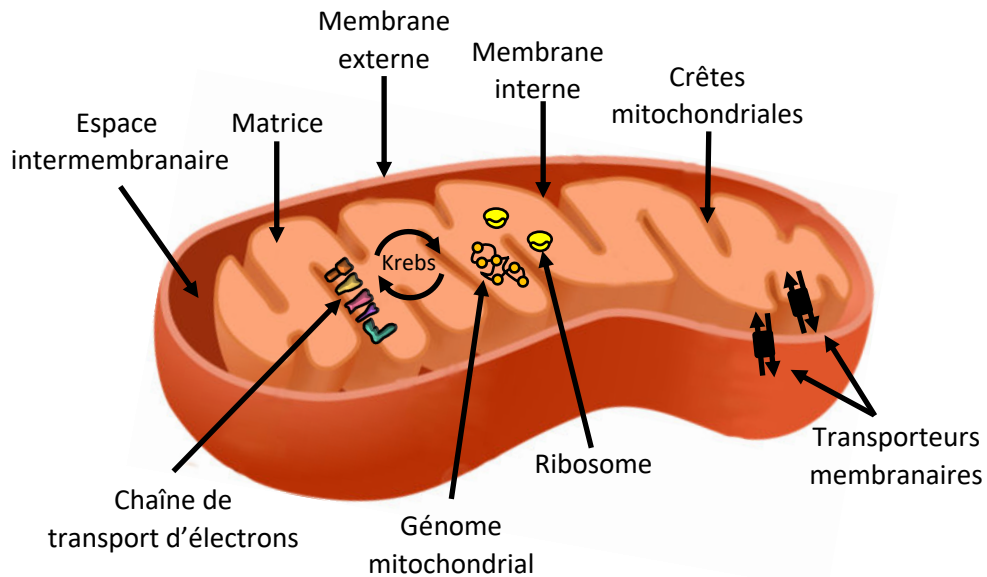


**Figure 3 : Evolution des plastes à travers les endosymbioses multiples.** Le premier chloroplaste est apparu par l'endosymbiose d'une cyanobactérie primitive (en bleu). Les eucaryotes photosynthétiques ont ensuite évolué en trois branches : les glaucophytes, les algues rouges, et les algues vertes (desquelles se sont développées les plantes terrestres). Des événements d'endosymbiose secondaire entre des eucaryotes sans chloroplaste et des algues rouges ou vertes ont conduit aux taxons représentés après les lignes pleines rouges et vertes. Les lignes pointillées représentent les événements d'endosymbiose tertiaire. D'après Keeling, 2013.

mitochondrial. La conservation de ces gènes serait dû aux difficultés d'import des protéines très hydrophobes à travers la membrane interne de la mitochondrie (Daley *et al.*, 2002), ainsi qu'au besoin d'une régulation rapide de ces protéines en fonction de l'état d'oxydo-réduction (Allen, 1993). Cependant, certains eucaryotes anaérobies possèdent des mitochondries dégénérées ayant complètement perdu leur génome (Muller *et al.*, 2012). Ces organelles utilisent des composés organiques comme accepteur terminal d'électrons plutôt que l'oxygène. Certaines de ces organelles ne possèdent plus de chaîne de transport d'électrons, comme les hydrogénosomes qui produisent de l'ATP par fermentation, ou les mitosomes qui ne servent qu'à l'assemblage des *clusters* fer-soufre (Goldberg *et al.*, 2008). Un groupe d'eucaryotes, *Monocercomonoid spp.*, ayant complètement perdu leurs mitochondries a récemment été identifié, révélant que la formation des *clusters* fer-soufre, considéré comme étant la fonction minimale requise pour une mitochondrie, pouvait être réalisée en absence d'organelle chez un eucaryote (Karnkowska *et al.*, 2016).

Contrairement à la mitochondrie ayant été acquise très tôt lors de l'évolution des eucaryotes, le chloroplaste est apparu plus tardivement et n'est présent que chez une partie des eucaryotes. Il est admis qu'une cyanobactérie primitive s'est associée avec un eucaryote ancestral, probablement à la suite d'un événement de phagocytose (Keeling, 2013). Cet événement est connu comme une endosymbiose primaire, responsable de la formation des premiers eucaryotes photosynthétiques. Mais certains protistes possèdent un chloroplaste sans avoir directement endocyté de cyanobactérie. Ainsi, ces eucaryotes non photosynthétiques ont phagocyté d'autres eucaryotes qui possédaient des chloroplastes. Ces événements sont connus sous le nom d'endosymbiose secondaire, voire tertiaire (**figure 3**) (McFadden, 2001). Ces endosymbioses multiples ont généralement conservé les membranes du phagosome autour de l'endosymbiote, nécessitant l'ajout de transporteurs membranaires supplémentaires. De plus, l'endosymbiose avec une cyanobactérie est un événement qui est arrivé plusieurs fois dans l'histoire évolutive des eucaryotes photosynthétiques : l'amibe *Paulinella chromatophora* présente un cas de seconde endosymbiose primaire la convertissant de l'hétérotrophie à l'autotrophie (Nakayama et Ishida, 2009; Nowack *et al.*, 2011), tandis que la diatomée *Rhopalodia gibba* possède un plaste dérivant d'une algue rouge (endosymbiose secondaire) et un nouveau symbiote d'origine cyanobactérienne (Gould *et al.*, 2008). Ces endosymbioses plus récentes permettent de mieux appréhender les mécanismes à leurs origines, bien qu'ils soient encore largement inconnus.

Au cours de l'évolution du chloroplaste, le génome de la cyanobactérie primitive a été considérablement réduit, comme pour celui de la mitochondrie. De nombreux gènes ont été perdus tandis que d'autres ont migré, ou continuent de migrer (Huang *et al.*, 2003), vers le génome nucléaire de l'hôte. Ainsi, le génome des eucaryotes photosynthétiques, tels que les



**Figure 4 : Représentation schématique de la structure interne de la mitochondrie isolée.** Les complexes respiratoires sont retrouvés au niveau des crêtes mitochondriales, séparés de l'espace intermembranaire par des jonctions (non représentées). L'espace intermembranaire est légèrement acide à cause du pompage des protons en dehors de la matrice par les complexes I, II et IV de la chaîne respiratoire. Chez les plantes, les mitochondries renferment moins d'un ADNmt complet par mitochondrie. Le génome est associé à de nombreuses protéines impliquées dans le métabolisme des acides nucléiques et forment le nucléoïde.

plantes, est une mosaïque de gènes de diverses origines. De nombreux facteurs cellulaires sont des homologues de facteurs procaryotiques.

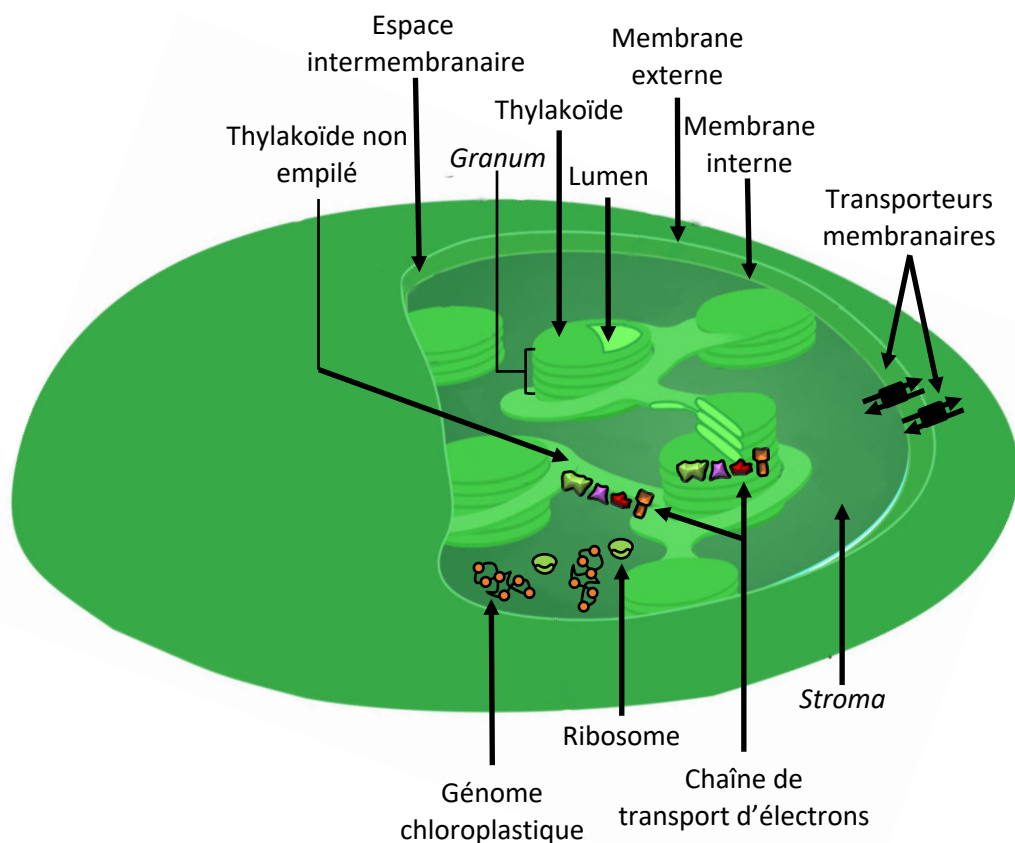
## 2. La mitochondrie et le chloroplaste

### 2.1. Structure et dynamique

Après l'endosymbiose, la structure des endosymbiotes a considérablement évolué. Les organismes d'abord libres se sont peu à peu transformés en compartiments spécialisés. Libérés des contraintes prises en charge par l'hôte, ils ont ainsi développé des structures particulières afin d'accroître leur spécificité.

Le chondriome des plantes est organisé en organelles pléomorphes et discrètes tandis que celui des animaux et des levures s'organisent en longs tubules et réticules. Le nombre de mitochondrie d'une cellule dépend du tissu d'où provient la cellule. On considère qu'une cellule de mésophylle d'*Arabidopsis thaliana* contient 200 à 300 mitochondries alors qu'une levure présente cinq à dix réticules de mitochondries et qu'une cellule humaine contient entre 500 et 2 000 mitochondries (Scott et Logan, 2008). Isolée, une mitochondrie est constituée de deux membranes séparant le cytosol à l'extérieur de l'espace intermembranaire et de la matrice interne (**figure 4**). Ces deux membranes sont traversées de canaux et contrôlent l'échange des ions et molécules entre le cytosol et la matrice. Cinq familles de transporteurs membranaires sont utilisées pour l'import des protéines dans les différents compartiments de la mitochondrie en fonction de leur adressage (Wiedemann et Pfanner, 2017). La membrane externe est composée de pores permettant le passage passif des molécules de taille inférieure à 6 kDa, tandis que la membrane interne est beaucoup moins perméable (O'Rourke, 2007). La composition de la membrane interne est assez différente des membranes biologiques habituelles car elle est composée de lipides différents et d'un grand nombre de protéines (impliquées notamment dans la phosphorylation oxydative) (Comte *et al.*, 1976; Horvath et Daum, 2013). L'espace intermembranaire joue le rôle d'un *hub* liant les composants produits à l'intérieur et à l'extérieur de la mitochondrie (Herrmann et Riemer, 2010) et présente une composition ionique similaire à celle du cytosol. Cet espace est néanmoins plus acide que le cytosol à cause du pompage des protons hors de la matrice par les complexes de transport des électrons. La membrane interne forme de nombreuses invaginations appelées crêtes mitochondriales. Ces crêtes sont reliées par des jonctions qui compartimentent la respiration aérobie en limitant la diffusion des composants (protons, métabolites) (Gilkerson *et al.*, 2003). Enfin, la matrice contient les éléments nécessaires aux nombreuses voies métaboliques (cycle de Krebs,  $\beta$ -oxydation des acides gras, oxydation du pyruvate, cycle de l'urée, etc.), ainsi que le génome et la machinerie traductionnelle de la mitochondrie. Ce génome est empaqueté dans des complexes protéiques sphériques, appelés nucléoïdes, contenant toutes les protéines nécessaires pour sa transcription, répllication et ségrégation (Gualberto et Kühn, 2014).

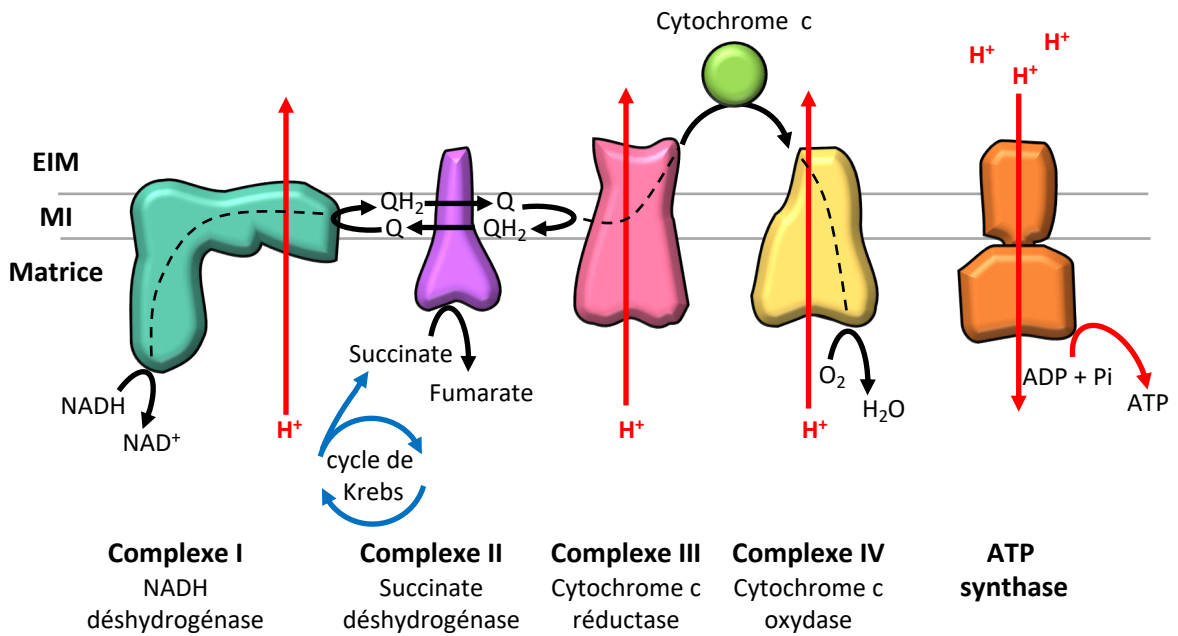




**Figure 5 : Représentation schématique de la structure interne du chloroplaste.** Les complexes photosynthétiques se situent au niveau des membranes des thylakoïdes empilés et non empilés. Chaque chloroplaste possède plusieurs copies de l'ADNcp associées à de nombreuses protéines impliquées dans le métabolisme des acides nucléiques, formant les nucléoïdes.

Les chloroplastes sont plus gros que les mitochondries. Une cellule de mésophylle d'*Arabidopsis thaliana* contient une centaine de chloroplastes de forme sphéroïde, tandis que les algues ne contiennent généralement qu'un seul chloroplaste, pouvant occuper la majeure partie du cytoplasme et présentant des formes plus variées. Les chloroplastes se développent à partir de proplastides indifférenciés en présence de lumière. Ces proplastides peuvent également se différencier en d'autres types de plastides tels que les amyloplastides (stockage de l'amidon), les chromoplastides (stockage des pigments) ou les oléoplastides (stockage des lipides) (Pogson *et al.*, 2015), tous incapables de réaliser la photosynthèse. En absence de lumière, les chloroplastes peuvent se dédifférencier en étioplastes qui n'assurent plus la photosynthèse. Comme la mitochondrie, le chloroplaste est entouré par une double membrane lipidique, appelée l'enveloppe, séparant le cytosol externe du *stroma* interne (**figure 5**). Un troisième système de membranes est situé dans le *stroma* : les thylakoïdes, entourant le *lumen*. Ces thylakoïdes sont des structures plates et circulaires appelées *granum*, regroupées en piles ou *grana* interconnectées les unes avec les autres par certains thylakoïdes non empilés. Les photosystèmes I (PSI) et le photosystème II (PSII) sont imbriqués dans les membranes des thylakoïdes non empilés et les *grana*, respectivement (Albertsson, 2001). Les thylakoïdes sont ainsi le siège de la photosynthèse. Les membranes de l'enveloppe sont traversées par de nombreux canaux pour le transport des ions ainsi que par de larges complexes protéiques permettant l'import des protéines possédant une séquence d'adressage chloroplastique. Le *stroma* contient la machinerie traductionnelle et le génome du chloroplaste. Il y a généralement plusieurs copies du génome par chloroplaste. Ces génomes sont empaquetés dans des nucléoïdes dont la composition diffère des nucléoïdes mitochondriaux (Pfalz et Pfannschmidt, 2013).

Les chloroplastes et les mitochondries sont liés au cytosquelette et sont capables de se déplacer dans la cellule. Le cytosquelette joue également un rôle dans le maintien de la morphologie des organelles. Le déplacement le long des filaments d'actine permet la relocalisation des chloroplastes en fonction de la fluctuation de la lumière afin d'optimiser la photosynthèse ou limiter l'impact délétère de rayonnements lumineux trop puissants (Davis et Hangarter, 2012). Les mitochondries sont associées avec les microtubules et les filaments d'actine qui les positionnent aux endroits où elles sont le plus nécessaires dans la cellule (exemple : proche des chloroplastes lors de la période lumineuse (Islam *et al.*, 2009) ou dans les parties distales des neurones chez les animaux (Ruthel et Hollenbeck, 2003). Le mouvement des organelles permet également leur ségrégation lors de la division cellulaire. Le chloroplaste et la mitochondrie se divisent en utilisant des systèmes d'anneaux contractiles formés au niveau des membranes interne et externe (Arimura et Tsutsumi, 2002). Ces mécanismes sont proches de la division procaryotique et utilisent des facteurs procaryotiques et eucaryotiques. Avant la division cellulaire, les mitochondries subissent un événement de fusion massive



**Figure 6 : Représentation schématique de la chaîne de transport d'électrons mitochondriale des mammifères.** Le transfert d'électrons est représenté par les flèches noires et les traits en pointillés. Les complexes I, III et IV créent un gradient en transportant les protons de la matrice vers l'espace intermembranaire. L'ATP synthase utilise le flux de protons pour produire de l'ATP à partir d'ADP et de phosphate inorganique. D'après Sousa *et al.*, 2018. EIM : espace intermembranaire ; MI : membrane interne ; Q : quinone ; QH<sub>2</sub> : quinol.

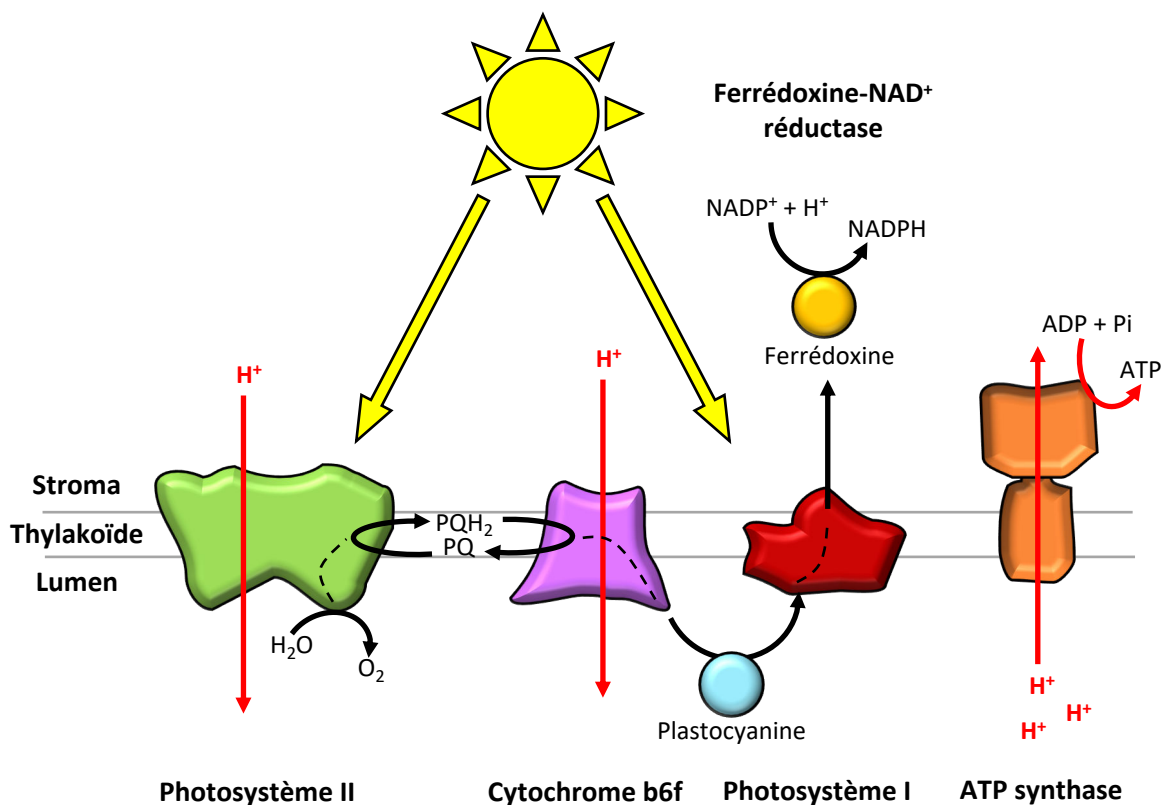
(MMF : *massive mitochondrial fusion*) permettant la réplication du génome par homogénéisation de ses composants (Logan, 2006b; Logan, 2006a). Les événements de fusion et fission de mitochondries sont fréquents en dehors de phase de division cellulaire. Chez les plantes, cela permet de rapprocher les génomes de différentes mitochondries pour permettre la recombinaison homologue intermoléculaire. Il existe ainsi une balance entre la fusion et la fission des mitochondries.

## 2.2. Les complexes de transport d'électrons

La mitochondrie et le chloroplaste sont principalement connus pour être responsables de la production de l'énergie cellulaire, qui repose sur l'utilisation de chaînes de transport d'électrons permettant la formation d'un gradient de protons utilisé pour la conversion d'ADP (adénosine diphosphate) en ATP (Johnson, 2016; Sousa *et al.*, 2018).

La mitochondrie produit l'ATP lors de la phosphorylation oxydative. Ce processus repose sur l'oxydation du NADH et du FADH<sub>2</sub> *via* la chaîne de transport d'électrons située dans la membrane interne des crêtes mitochondriales (Sousa *et al.*, 2018). Cette chaîne de transport d'électrons est constituée de quatre complexes qui permettent la formation d'un gradient de protons en pompant les protons de la matrice vers l'espace intermembranaire (**figure 6**). Les complexes I (NADH : ubiquinone oxydoréductase) et II (succinate déshydrogénase) oxydent le NADH et le succinate issus du cycle de Krebs en NAD<sup>+</sup> et fumarate, respectivement. Les électrons ainsi récupérés servent à réduire l'ubiquinone en ubiquinol qui sera oxydé par le complexe III (ubiquinol-cytochrome c réductase) pour réduire le cytochrome c. Le complexe IV (cytochrome c oxydase) récupère les électrons du cytochrome c réduit pour les donner à une molécule d'oxygène, accepteur final d'électron, et former une molécule d'eau. Lors de ces réactions d'oxydo-réduction les complexes I, III et IV permettent le pompage de deux (complexe IV) ou quatre (complexes I et III) protons vers l'espace intermembranaire par paire d'électrons traversant le complexe. Ces protons seront ensuite utilisés par l'ATP synthase pour produire l'ATP à partir d'ADP et de phosphate inorganique (Pi). Le gradient de protons met en marche le rotor de l'ATP synthase lorsque les protons retournent dans la matrice, induisant un changement conformationnel des sites de liaison aux nucléotides permettant la formation de l'ATP (Stock *et al.*, 2000).

La photosynthèse est l'un des processus les plus importants de la vie sur Terre. Elle permet la conversion de l'énergie lumineuse émise par le soleil en énergie chimique (ATP), mais aussi la formation des composés carbonés nécessaires au développement de l'organisme lors des phases d'obscurité. Ainsi, tous les organismes vivants utilisent directement ou indirectement (*via* l'ingestion d'organismes photosynthétiques) la photosynthèse dans leur développement. Les cyanobactéries libres sont responsables de 50 % de la photosynthèse



**Figure 7 : Représentation schématique de la chaîne de transport d'électrons chloroplastique.** Le transfert des électrons est représenté par les flèches noires et les tirets pointillés. Les photosystèmes I et II utilisent l'énergie du soleil pour permettre le transfert d'électrons. Le photosystème II et le cytochrome b6f créent un gradient de protons en important des protons dans le lumen. L'ATP synthase utilise le gradient de protons pour produire de l'ATP. D'après Johnson, 2016.  
PQH<sub>2</sub> : plastoquinol ; PQ : plastoquinone.

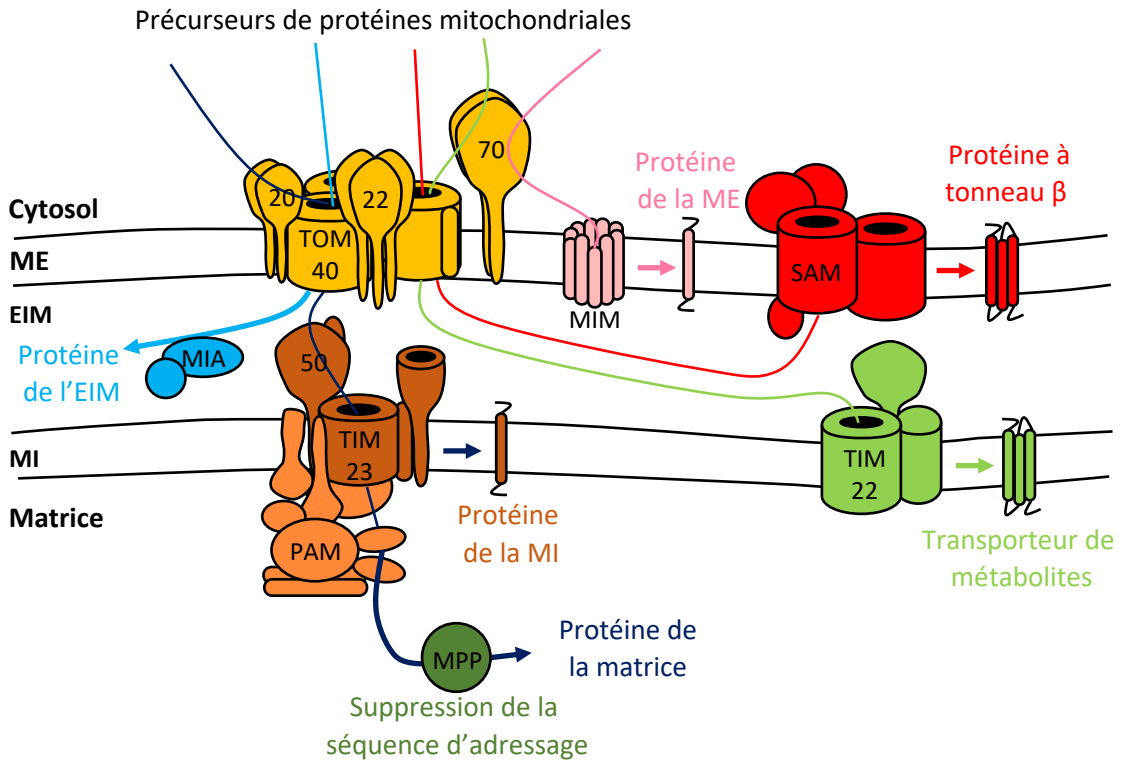
mondiale (Johnson, 2016). L'autre moitié est réalisée par les eucaryotes photosynthétiques possédant des plastes dérivant de cyanobactéries ancestrales.

La production d'énergie chimique à partir d'énergie lumineuse utilise des complexes protéiques imbriqués dans la membrane des thylakoïdes (Johnson, 2016). La première étape est réalisée par le photosystème II (PSII) qui oxyde une molécule d'eau grâce à l'énergie lumineuse et réduit la plastoquinone en plastoquinol (**figure 7**). L'absorption de la lumière est réalisée par la chlorophylle P680, qui aura ensuite le potentiel redox nécessaire pour oxyder l'eau. Le plastoquinol est à nouveau oxydé en plastoquinone par le cytochrome  $b_6f$ , récupérant les électrons pour réduire la plastocyanine. La plastocyanine réduite se déplace dans le *lumen* jusqu'au photosystème I (PSI), où elle est oxydée grâce à l'énergie lumineuse captée par la paire de chlorophylle a et b P700. Les électrons récupérés permettent alors la réduction de la ferrédoxine. Enfin, la ferrédoxine réduite est utilisée par la ferrédoxine-NAD<sup>+</sup> réductase pour produire du NADPH. Le NADPH ainsi formé sera utilisé dans la formation de composés carbonés à partir de CO<sub>2</sub> inorganique lors de phase nocturne de la photosynthèse. Le PSII et cytochrome  $b_6f$  pompent des protons du *stroma* vers le *lumen* lors du transfert des électrons. Ces protons permettront ensuite la production d'ATP à partir d'ADP + Pi grâce à l'ATP synthase (Stock *et al.*, 2000).

La formation de carbohydrates lors de la phase sombre de la photosynthèse nécessite le NADPH et l'ATP produits lors de la phase lumineuse, dans un processus appelé le cycle de Calvin-Beson (Johnson, 2016). Le CO<sub>2</sub> inorganique capté par les stomates est combiné avec le ribulose 1,5-biphosphate grâce à l'activité de la ribulose-1,5-biphosphate carboxylase/oxygénase, plus communément appelée Rubisco. La molécule à six carbones ainsi formée est instable et forme rapidement deux molécules de 3-phosphoglycérate à trois carbones. L'ATP et le NADPH sont utilisés pour former le 1,3-biphosphoglycérate, puis le glycéraldéhyde-3-phosphate à l'aide d'une kinase et d'une déshydrogénase, respectivement. Sur six molécules de glycéraldéhyde-3-phosphate produites, cinq serviront à la régénération du ribulose 1,5-biphosphate nécessaire au cycle. La sixième sera convertie dans différentes voies métaboliques permettant la croissance de l'organisme et la production d'énergie chimique.

### 2.3. Import de macromolécules dans les organelles

La mitochondrie et le chloroplaste sont constitués de milliers de protéines différentes alors que leur génome permet l'expression de seulement quelques dizaines de facteurs. La majorité des protéines organellaires est donc codée par le génome nucléaire, synthétisées dans le cytoplasme puis importée dans la mitochondrie ou le chloroplaste. Ainsi, toutes les activités des organelles sont contrôlées par le noyau, y compris l'expression des quelques gènes organellaires. Afin d'assurer l'import des différentes protéines dans le bon compartiment, la plupart des facteurs mitochondriaux et chloroplastiques codés par le noyau sont synthétisés



**Figure 8 : Représentation schématique des voies d'import de protéines dans la mitochondrie.**

La majorité des protéines importées dans la mitochondrie traversent le complexe TOM (*translocase of the outer membrane*). Les protéines adressées à la matrice qui possèdent une séquence d'adressage mitochondriale (en bleu foncé) passent ensuite par le complexe TIM (*translocase of the inner membrane*) et son moteur PAM. Une fois dans la matrice, la séquence d'adressage est clivée par la peptidase MPP. Les protéines riches en cystéine adressées à l'espace intermembranaire (en bleu clair) utilisent le complexe MIA. Les protéines membranaires utilisent les MIM, SAM, TIM 22 et TIM23 (en rose, rouge, bleu foncé et vert, respectivement). D'après Wiedemann et Pfanner, 2017

ME : membrane externe ; EIM : espace intermembranaire ; MI : membrane interne.

avec une séquence d'adressage ou un signal interne. Les séquences d'adressage des protéines de la matrice mitochondriale et du stroma du chloroplastiques se situent dans la partie N-terminale de la protéine. Ces séquences sont très peu conservées et ne présentent pas de consensus (Kunze et Berger, 2015). Néanmoins, elles présentent des propriétés physico-chimiques partagées. Les séquences d'adressage mitochondriales forment généralement des hélices- $\alpha$  amphiphiles et présentent une grande proportion d'acides aminés chargés positivement, notamment l'arginine, mais une faible proportion (voire l'absence) d'acide aminés chargés négativement (von Heijne *et al.*, 1989). Les séquences d'adressage chloroplastiques forment aussi des hélices- $\alpha$  amphiphiles avec une faible proportion d'acide aminés chargés négativement (néanmoins présents), mais une plus forte proportion d'acides aminés hydroxylés comme la sérine et la thréonine (Bruce, 2000). Chez les plantes, un grand nombre de protéines sont à la fois adressées aux mitochondries et aux chloroplastes. Ces protéines peuvent posséder la séquence d'adressage de l'une ou l'autre des organelles *via* des mécanismes d'épissage alternatif (Teardo *et al.*, 2015), ou des sites d'initiation de la transcription ou de traduction alternatifs (Obara *et al.*, 2002; Wamboldt *et al.*, 2009)). Elles peuvent aussi posséder une séquence d'adressage dite ambiguë possédant les caractéristiques des deux organelles (Silva-Filho, 2003). Malgré l'adressage de protéines aux deux organelles, un grand nombre d'entre elles n'ont de rôle identifié que dans un des deux compartiments (Rödiger *et al.*, 2011).

Les protéines adressées aux mitochondries et aux chloroplastes s'associent avec des protéines chaperonnes de la famille des HSP70 (*heat shock protein*) (Komiya *et al.*, 1996; Flores-Perez et Jarvis, 2013) afin de rester en conformation non repliée et de favoriser la reconnaissance par les systèmes d'import présents dans les membranes des organelles. Ces systèmes d'import sont composés des complexes de translocation à travers la membrane externe, TOM et TOC, et de translocation à travers la membrane interne, TIM et TIC, pour la mitochondrie et le chloroplaste, respectivement.

Le complexe TOM (*translocon of the outer mitochondrial membrane*) est le site d'entrée principal des protéines dans la mitochondrie (**figure 8**). Il est composé des récepteurs tom20 et tom22 reconnaissant les séquences d'adressage, et du canal membranaire tom40 permettant le passage de la protéine dans l'espace intermembranaire (Wiedemann et Pfanner, 2017). La protéine y est ensuite reconnue par le récepteur du complexe TIM (*translocon of the inner mitochondrial membrane*) appelé tim50 permettant le passage de la protéine dans le canal tim23 grâce au moteur PAM (*presequence translocase-associated motor*). Une fois dans la matrice, la séquence d'adressage subit un clivage protéolytique par la peptidase MPP. Quatre autres voies permettent l'import de 40 % des protéines mitochondriales qui ne possèdent pas de séquence d'adressage. Ces voies sont surtout utilisées pour l'import des protéines qui seront ensuite insérées dans les membranes mitochondriales (Wiedemann et Pfanner, 2017). Pour le

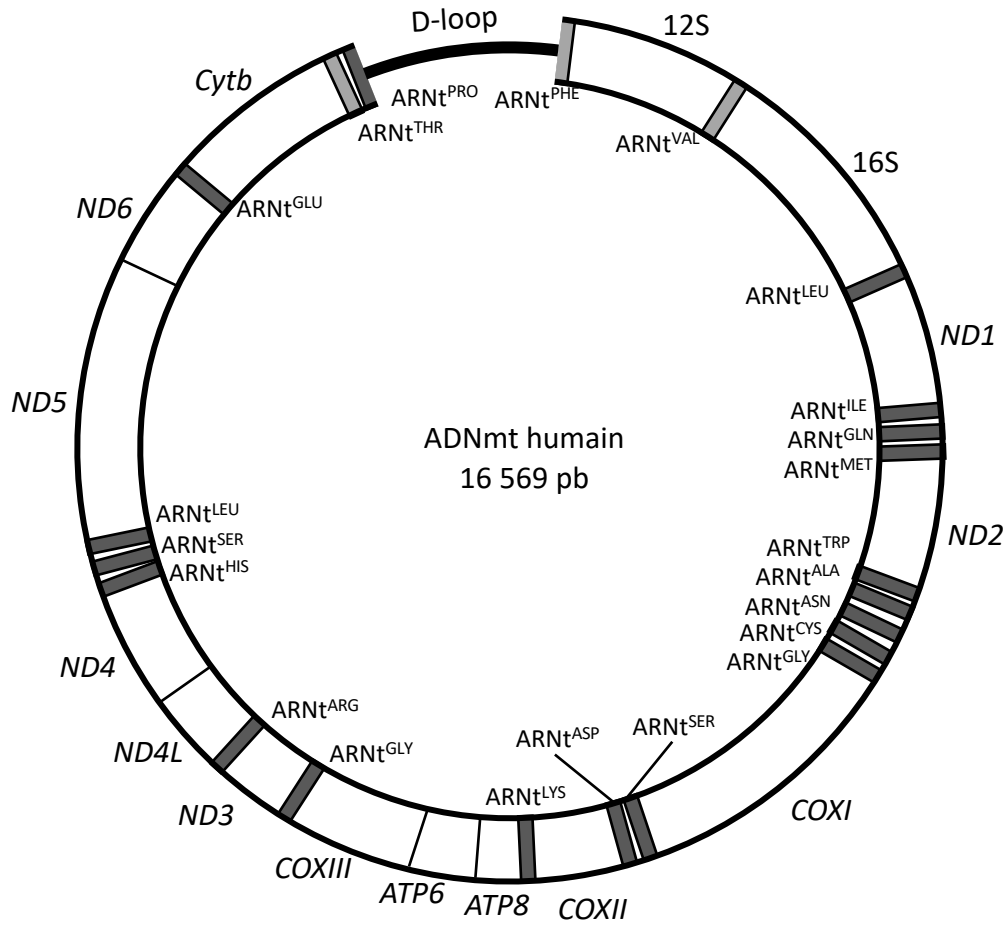


chloroplaste, le complexe TOC (*translocon of the outer chloroplast membrane*) est composé des GTPases Toc159 et Toc34, reconnaissant les séquences d'adressage chloroplastiques, et du canal membranaire Toc75 (Nakai, 2018). La séquence d'adressage est reconnue par la protéine Tic22 dans l'espace intermembranaire qui lui permet de passer à travers le complexe TIC (*translocon of the inner chloroplast membrane*) dans la membrane interne vers le *stroma*. Les protéines adressées au lumen ou aux membranes des thylakoïdes présentent une seconde séquence d'adressage en aval de la première (Albiniak *et al.*, 2012).

Le génome mitochondrial code un certain nombre d'ARNt. Ce nombre varie en fonction des organismes, pouvant aller du jeu complet de 22 ou 24 ARNt mitochondriaux chez *Homo sapiens* ou *Saccharomyces cerevisiae*, respectivement, à leur absence totale chez certains protistes comme *Trypanosoma brucei* (Salinas-Giege *et al.*, 2015). Les organismes dont le génome mitochondrial ne code pas la totalité des ARNt nécessitent donc l'import des ARNt manquant depuis le cytosol. Les mécanismes d'import de ces ARNt ne sont pas encore complètement résolus, et semblent varier selon l'organisme. Ils font intervenir un signal d'import permettant l'exclusion de ces ARNt de la machinerie traductionnelle cytosolique pour les envoyer vers la membrane externe mitochondriale. Chez la levure, l'ARNt<sup>Lys</sup> amino-acylé est pris en charge par une énoïase et le précurseur de la lysyl-ARNt-synthétase (pré-LysRS) qui est adressé à la mitochondrie. L'ARNt est ainsi importé dans la matrice mitochondriale *via* les complexes TOM et TIM en même temps que la pré-LysRS (Kamenisch *et al.*, 2010). Chez les plantes, le mécanisme fait également intervenir les aaRS (*aminoacyl tRNA synthetase*) ainsi que les protéines TOM20 et TOM40 comme récepteurs membranaires. La translocation de l'ARNt dans la mitochondrie est cependant réalisée à travers la porine, ou VDAC (*voltage-dependant anion channel*) (Salinas *et al.*, 2006). Les organismes dont la mitochondrie code la totalité des ARNt nécessaire à la traduction importent tout de même des ARNt cytosoliques dans un but de régulation (Kamenski *et al.*, 2007).

D'autres types d'ARN non codants sont également importés dans la mitochondrie. Chez les mammifères, l'ARNr 5S codé par le génome nucléaire est adressé à la mitochondrie (Magalhaes *et al.*, 1998). Il pourrait y être incorporé dans la grande sous-unité du ribosome mitochondrial ou mitoribosome (Entelis *et al.*, 2001). Cependant, l'élucidation de la structure du mitoribosome a montré qu'à la place de l'ARNr 5S on retrouve un ARNt qui le remplace au niveau structural, l'ARNt<sup>Val</sup> ou l'ARNt<sup>Phe</sup> dans le cas du mitoribosome humain (Chrzanowska-Lightowlers *et al.*, 2017). La raison de l'import de l'ARNr 5S reste mystérieuse.

Le long ARN non codant *RMRP*, impliqué dans la maturation nucléaire de l'ARNr 5,8S et le cycle cellulaire de la levure, est importé dans la mitochondrie où il pourrait jouer un rôle dans la réplication du génome mitochondrial ou dans la respiration oxydative (Noh *et al.*, 2016). L'ARN de la RNase P d'origine bactérienne possédant une activité catalytique est également importé dans la mitochondrie humaine, où ses fonctions sont controversées (Noh *et al.*, 2016;



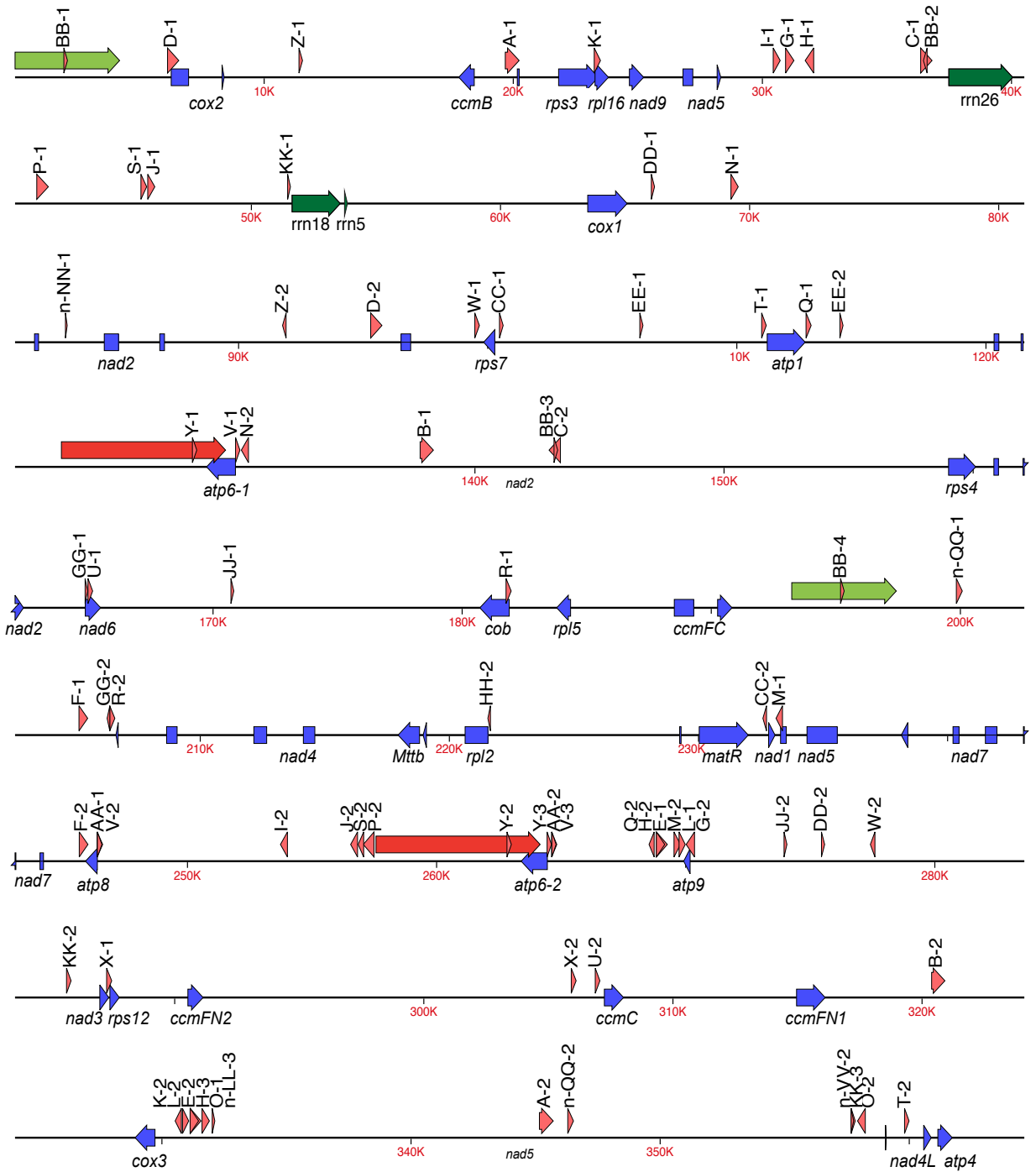
**Figure 9 : Représentation du génome mitochondrial humain.** Le ADNmt des mammifères est de petite taille et assez compact, avec peu de séquences non codantes. Les larges bandes blanches représentent les 15 gènes codant pour les 13 protéines et les deux ARN ribosomiques. Les petites bandes grises représentent les 22 gènes codant des ARN de transfert. La *D-loop* est une région non codante impliquée dans l'expression et le maintien de l'ADNmt.

Kim *et al.*, 2017). Chez les plantes et beaucoup d'autres eucaryotes, l'activité RNase P est assurée par une RNase P seulement protéique (PRORP) (Gutmann *et al.*, 2012; Lechner *et al.*, 2015). Enfin, certains miARN (micro ARN) nucléaires ont été retrouvés dans la mitochondrie humaine (Barrey *et al.*, 2011).

### 3. Les génomes organellaires

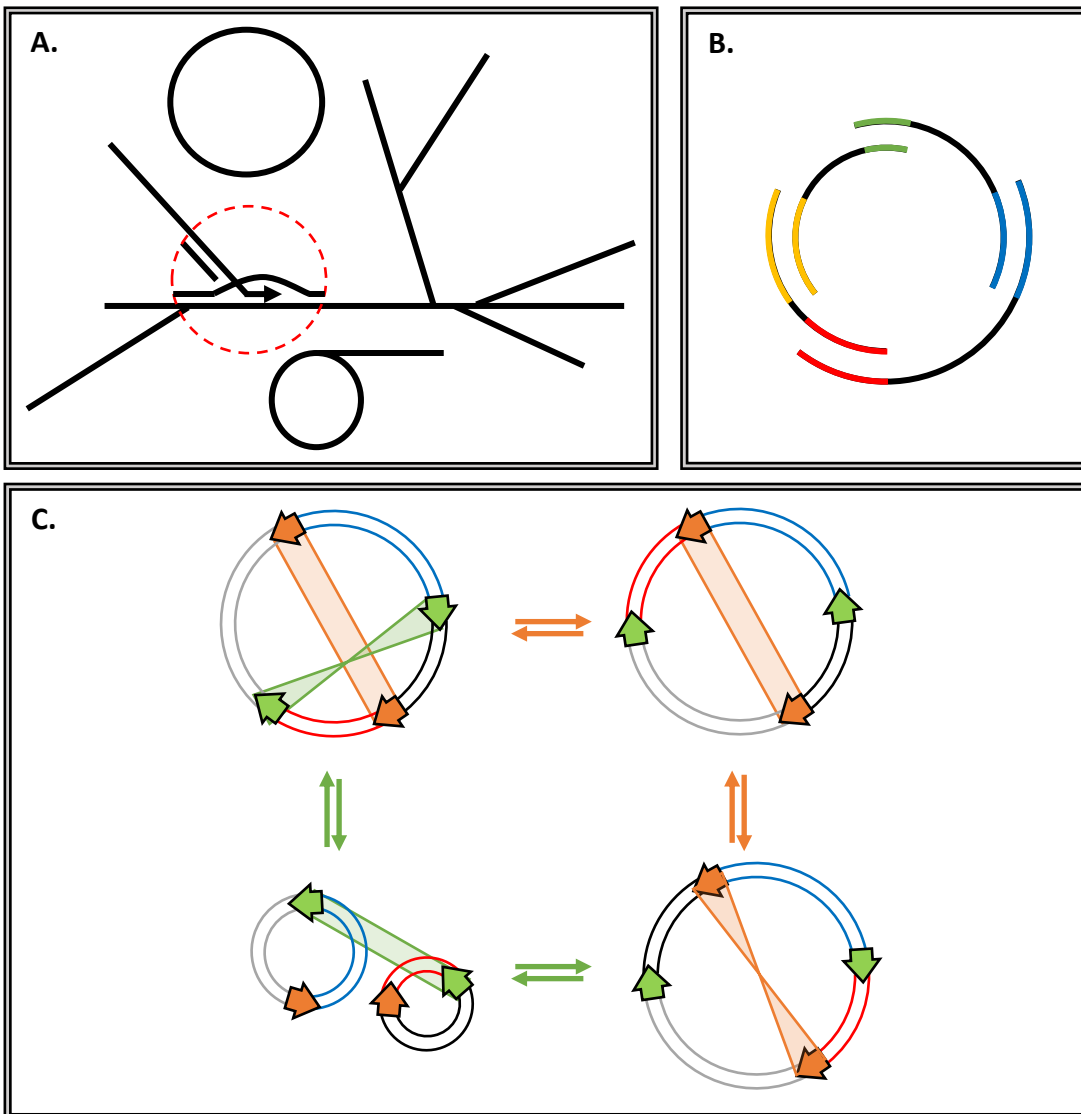
Les génomes mitochondriaux (ADNmt : ADN mitochondrial) et chloroplastiques (ADNcp : ADN chloroplastique) dérivent des génomes de l' $\alpha$ -protéobactérie et de la cyanobactérie ancêtres symbiotiques des deux organelles, qui se sont peu à peu réduits au cours de leur évolution, *via* la perte ou la migration de leurs gènes vers le génome nucléaire. Les ADNmt actuels sont très variables en taille et en structure mais présentent globalement la même capacité codante. En considérant *Rickettsia prowazekii* comme le procaryote actuel le plus proche de la mitochondrie, l'ADNmt des animaux présente une réduction de 98,5 % en taille et de 95 % en capacité codante, alors même que la bactérie est un parasite obligatoire (Gray, 1999). L'ADNmt des levures peut avoir une taille relativement proche de l'ADNmt des animaux (18,5 kb pour *Hanseniaspora uvarum*) à plus de 105 kb (chez *Nakasemyces bacillisporus*) pour une capacité codante similaire aux animaux (30 à 40 gènes) (Pramateftaki *et al.*, 2006; Bouchier *et al.*, 2009; Freel *et al.*, 2015). Les plantes ont des ADNmt beaucoup plus grands, généralement autour des 200-300 kb mais pouvant aller jusqu'à 11,3 Mb pour *Silene conica* (Sloan *et al.*, 2012). Malgré leur taille, les ADNmt de plantes ne présentent qu'une vingtaine de gènes supplémentaires par rapports aux animaux et aux levures, et sont donc majoritairement composés de séquences non codantes. Les organismes plus simples présentent généralement un nombre plus élevé de gènes mitochondriaux. Le protiste *Reclinomonas americana*, considéré comme l'organisme avec l'ADNmt le plus proche de *Rickettsia prowazekii*, a une taille de 69 kb et contient une centaine de gènes (Lang *et al.*, 1997). D'autres protistes, tels que *Trypanosoma brucei* ou *Diplonema papillatum* possèdent d'énormes ADNmt (4,2 Mb et 260 Mb, respectivement), sans coder plus de protéines que les animaux (Lukeš *et al.*, 2018). L'ADNcp des plantes supérieures est moins variable, généralement d'une taille de 150 kb pour 120 gènes (contre 4 Mb et 3 000 gènes pour les cyanobactéries) (Timmis *et al.*, 2004).

Le génome mitochondrial humain est une molécule d'ADN circulaire de 16,6 kb présentant 37 gènes codant 13 protéines impliquées dans la chaîne de transport d'électrons, 22 ARNt et les deux ARNr 12S et 16S de la petite et de la grande sous-unité du mitoribosome, respectivement (**figure 9**). Une cellule contient entre 1 000 et 5 000 copies de l'ADNmt empaquetées généralement individuellement ou deux par deux dans des nucléoïdes (Kukat *et al.*, 2011). Ces nucléoïdes sont situés dans la matrice et peuvent être échangés entre les mitochondries lors d'événement de fusion et de fission. L'ADNmt est compact, avec une forte densité de gènes parfois chevauchants et peu de séquences intergéniques en dehors de la



**Figure 10 : Représentation du génome mitochondrial d'*Arabidopsis thaliana* ecotype Col-0.**

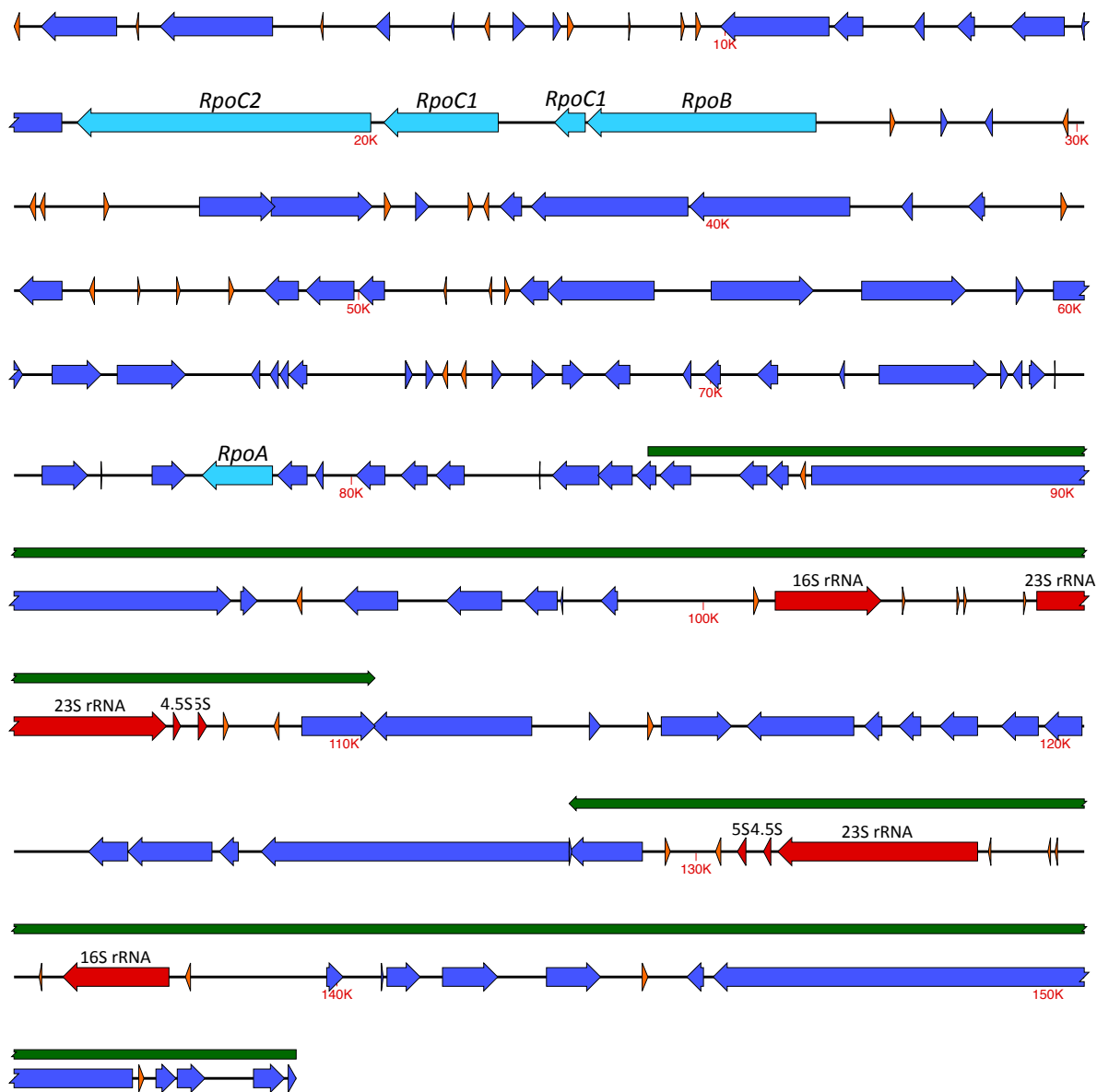
Le génome est représenté sous une forme linéaire regroupant les éventuels sous-génomes. Les gènes codant les protéines et les ARNr sont représentés par des flèches bleues et vertes foncées sur la séquence, respectivement. Les séquences répétées sont indiquées au dessus de la séquence : les deux paires de grandes séquences répétées sont en rouge et en vert clair tandis que les IRs de plus de 100 pb sont en rose.



**Figure 11 : Structure de l'ADN mitochondrial des plantes.** A. Observé par microscopie, le ADNmt des plantes présente un mélange de structures circulaire, linéaire et branchée. Le branchement des molécules linéaires proviendrait de l'invasion d'une molécule d'ADN par une autre lors d'un évènement de recombinaison homologue. La *D-loop* ainsi formée est représentée dans le cercle de pointillé rouge. B. La représentation sous forme circulaire de l'ADNmt (*master circle*) est due à la permutation circulaire de molécules d'ADN linéaires présentant des séquences répétées (en couleur) entre elles. C. La recombinaison entre les séquences répétées (en orange et vert) entraîne des réarrangements dans le ADNmt et peut conduire à la formation de sous-génomes. D'après Gualberto *et al.*, 2014.

région régulatrice. Cette région possède une structure triple brin, appelée D-loop (displacement loop), est nécessaire pour la transcription et la réplication du génome, par l'ARN polymérase POLRMT et l'ADN polymérase Pol $\gamma$ , respectivement, codées par le génome nucléaire (Ciesielski *et al.*, 2016; Barshad *et al.*, 2018). Les transcrits mitochondriaux sont polycistroniques, débutant par deux promoteurs au niveau de la région régulatrice, et ne contiennent pas d'intron.

Bien qu'il soit 45 fois plus grand que l'ADNmt humain, l'ADNmt d'*Arabidopsis thaliana* (**figure 10**) ne code que 32 protéines (impliquées dans la chaîne de transport d'électrons et la structure du mitoribosome), 22 ARNt et trois ARNr (5S, 18S et 26S), soit seulement 20 facteurs supplémentaires (Unsel *et al.*, 1997). La grande majorité de l'ADNmt des plantes (*Arabidopsis* comme exemple) est composée de séquences non codantes dérivant de séquences nucléaires (4 %), chloroplastiques (1 %), virales, ou inconnues pour la plus grande partie (62 %). Ces proportions sont bien plus grandes pour les très grands génomes comme ceux de *Silene*, qui ne codent pas plus de facteurs. Les séquences de l'ADNmt ne sont pas uniformément conservées, et certaines parties sont stœchiométriquement plus représentées que d'autres dans une cellule (Preuten *et al.*, 2010). Une cellule possède également plus de mitochondries que de copies de l'ADNmt, indiquant qu'une partie de ses mitochondries ne contient pas de génome entier (Preuten *et al.*, 2010). Enfin, il est montré que plusieurs populations différentes de l'ADNmt, appelées sublimons ou sous-génomes, sont présentes au sein d'une même mitochondrie (Small *et al.*, 1987; Small *et al.*, 1989). L'ADNmt des plantes est ainsi divisé en plusieurs molécules, obtenues par des événements de recombinaison, qui sont dispersées dans les différentes mitochondries d'une cellule au gré de leurs fusions et fissions. Ces événements de recombinaison sont fréquents et impliquent de grandes séquences répétées. L'ADNmt d'*Arabidopsis* contient deux paires de grandes séquences répétées, d'une taille de 6,5 kb et 4,2 kb (**figure 10**). D'autres événements de recombinaison, moins fréquents, impliquent des séquences répétées de taille inférieure (50-500 pb) voire des micro-homologies (< 50 pb), (Woloszynska et Trojanowski, 2009; Maréchal et Brisson, 2010; Janicka *et al.*, 2012). Ces recombinaisons ectopiques sont généralement inhibées par les facteurs de régulation de la recombinaison homologue car elles entraînent des réarrangements génomiques plus importants, parfois délétères pour la plante (Maréchal et Brisson, 2010; Woloszynska, 2010). Ces recombinaisons rendent difficiles la conception de la structure de l'ADNmt des plantes. Bien qu'il soit représenté sous la forme d'un ADN circulaire appelé le master circle, un tel génome n'a jamais été observé. Des études de microscopie ont montré que l'ADNmt présentait plutôt une structure hétérogène avec un mélange de sous-génomes circulaires et de molécules linéaires et branchées (**figure 11**) (Backert *et al.*, 1997). Ces ADNmt linéaires pourraient être concatémériques et peuvent ainsi être représentés par permutation circulaire. L'existence de ADNmt concatémériques est appuyée par la présence d'une insertion dérivant d'un concatémère de génome mitochondrial dans le génome nucléaire d'*Arabidopsis* (Stupar *et al.*,



**Figure 12 : Représentation du génome chloroplastique d'*Arabidopsis thaliana* écotype Col-0.**

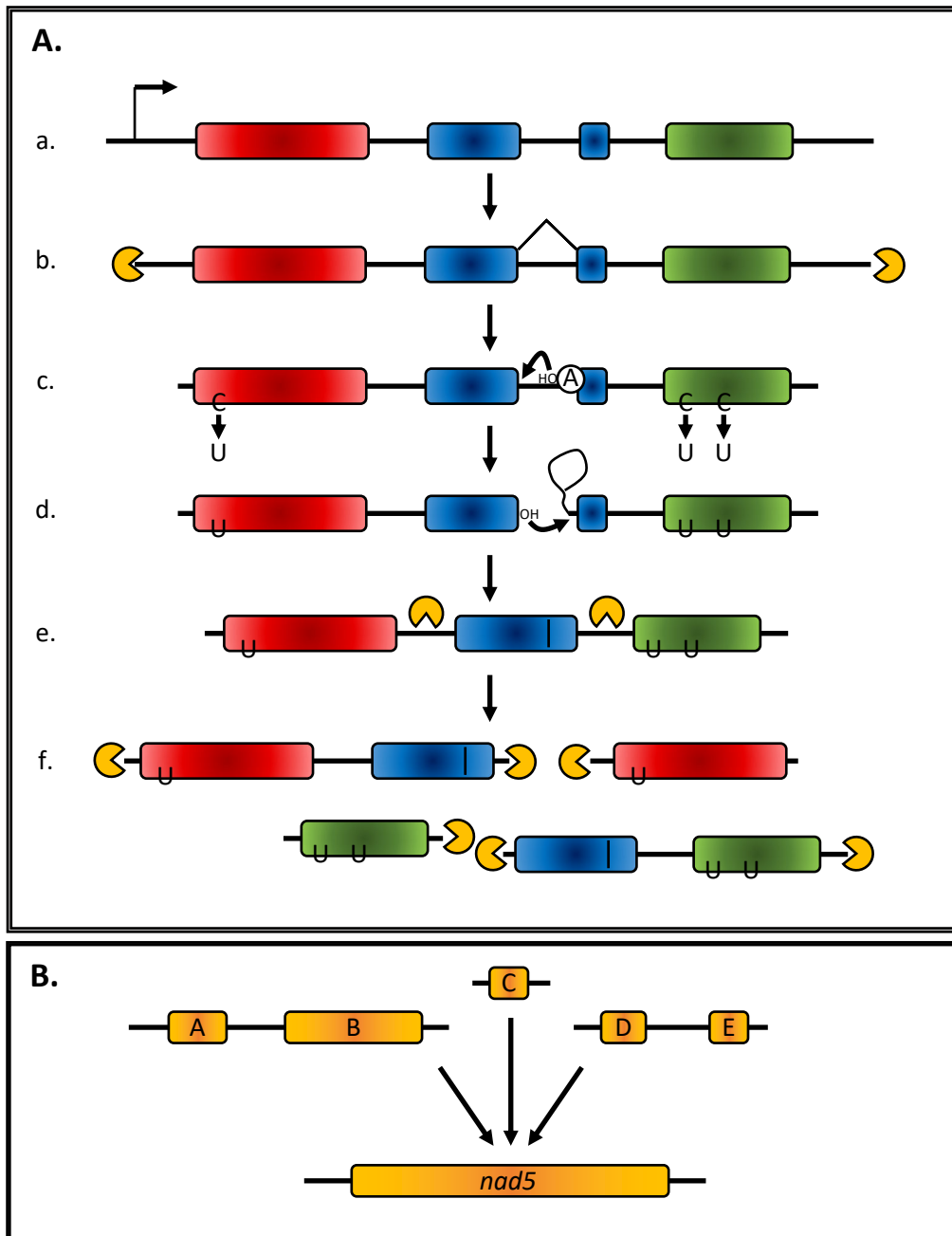
Les gènes codant les protéines, les ARNr et les ARNt sont représentés sur la séquence par des flèches bleues, rouges et oranges, respectivement. Les gènes codant les quatre sous-unités de la PEP sont en bleu clair. Les deux séquences répétées en orientation inverse sont représentées par les flèches vertes foncées au dessus de la séquence.

2001).

Le génome chloroplastique est relativement petit par rapport au génome mitochondrial des plantes, allant de 120 à 190 kb pour les algues comme pour les plantes terrestres. L'ADNcp est représenté sous une forme circulaire et double brin, certainement identique pour chaque copie. Néanmoins, il a été montré que la majorité des molécules de l'ADNcp était arrangée en concatémères et présentait une structure linéaire et branchée (Lilly *et al.*, 2001; Bendich, 2004). Le nombre de copies de l'ADNcp dépend du type de cellule et de son stade de développement, variant d'une centaine dans une cellule méristématique à plusieurs milliers de copies dans une cellule de parenchyme de feuille (Golczyk *et al.*, 2014). L'ADNcp a une densité génique plus importante que l'ADNmt, généralement 50 % (contre 10 % pour l'ADNmt d'*Arabidopsis thaliana*), et code ainsi un plus grand nombre de facteurs (100 à 200 gènes). L'ADNcp d'*Arabidopsis thaliana* a une taille de 154 kb avec 129 gènes permettant l'expression de quatre ARNr (4,5S ; 5S ; 16S et 23S), 30 ARNt et 77 protéines impliquées dans le métabolisme photosynthétique, la machinerie traductionnelle et la transcription de l'ADNcp (**figure 12**) (Sato *et al.*, 1999). Dans presque toutes les plantes, l'ADNcp est (virtuellement) divisé en quatre parties. Deux grande régions répétées (IRs : *inverted repeats*) séparent une petite région (SSC : small single copy) et une grande région (LSC : large single copy) non répétée. Les deux séquences répétées sont la partie la plus conservée de l'ADNcp grâce à la correction fidèle des erreurs par recombinaison homologue intramoléculaire et contiennent des gènes en orientation inverse.

Les génomes mitochondriaux et chloroplastiques sont empaquetés dans des complexes nucléo-protéiques similaires à ceux des chromosomes bactériens. Bien que la composition exacte de ces nucléoïdes ne soit pas encore complètement résolue, de nombreux facteurs impliqués dans l'expression, le maintien et la ségrégation des génomes organellaires ont été associés avec ces complexes (Pfalz *et al.*, 2006; Parent *et al.*, 2011; Majeran *et al.*, 2012). Les nucléoïdes mitochondriaux de plantes pourraient être composés des ADN polymérase mitochondriales POL1A et POL1B, d'un homologue de l'hélicase réplicative animale TWINKLE, de la topoisomérase de type II (sous-unités GYRA et GYRB), des recombinases RECA2 et RECA3, d'homologues de la protéine de liaison à l'ADN mitochondrial mtSSB (SSB1 et SSB2), de nombreux autres facteurs de liaison à l'ADN simple brin (OSB1, OSB3, WHY2 et ODB1), de la ligase LIG1, des ARN polymérase de type phagique RPOTm et RPOTmp et de l'homologue du facteur de réparation MSH1 (Xu *et al.*, 2011; Gualberto et Kühn, 2014). Plusieurs de ces protéines sont également impliquées dans la réparation de l'ADN et la recombinaison homologue. Certains facteurs composant le nucléoïde mitochondrial sont aussi adressés aux chloroplastes. Les nucléoïdes chloroplastiques ont pu être purifiés chez plusieurs espèces et leur composition est moins hypothétique (Pfalz *et al.*, 2006; Majeran *et al.*, 2012; Melonek *et al.*, 2016). Ils sont composés de nombreux facteurs impliqués dans la réplication, la





**Figure 13 : Modifications post-transcriptionnelles des transcrits organellaires de plante. A.** Les gènes des organelles sont généralement organisés en opérons avec un même promoteur (a) et transcrits en un ARN polycistronique (b). Les extrémités des ARN sont maturées par des nucléases aidées par des facteurs protéiques se liant à l'ARN, telles que les protéines PPR, qui définissent les séquences à cliver (b). La séquence des gènes peut être éditée et les introns sont épissés (c et d). Les transcrits organellaires présentent généralement des introns de groupe II, impliquant la formation d'un lasso lors des attaques nucléophiles. Les différents cistrons sont clivés (e) et les extrémités sont à nouveau maturées par des nucléases (f). **B.** La maturation de certains transcrits comme *nad5* nécessite plusieurs précurseurs d'ARN transcrits indépendamment depuis plusieurs *loci* et épissés en *cis* (entre les exons A et B, et D et E) et en *trans* (entre les exons B, C et D).

transcription, l'organisation et la réparation de l'ADNcp, mais aussi dans la maturation des transcrits (épissage et édition) et la biogenèse du ribosome. La présence de ces facteurs au sein des nucléoïdes pourrait indiquer une maturation co-transcriptionnelle des transcrits couplée directement à la traduction, de la même façon que chez les procaryotes. L'ARN polymérase de type bactérienne PEP (plastidial encoded polymerase) codée par l'ADNcp et ses facteurs associés pTAC sont largement retrouvés au sein de ces nucléoïdes, alors que les ARN polymérases de type phagique NEP (nuclear encoded polymerase) en sont absents ou en trop faible concentration pour être détectées (Pfalz *et al.*, 2006; Majeran *et al.*, 2012). Des facteurs d'ancrage des nucléoïdes avec les membranes du chloroplaste ont également été retrouvés (Melonek *et al.*, 2012; Krupinska *et al.*, 2013).

Chez les plantes, les mécanismes d'expression des génomes mitochondriaux et chloroplastiques dérivent en grande partie de ceux retrouvés chez leurs ancêtres procaryotiques, bien qu'ils soient tous codés par le génome nucléaire. Mais ces mécanismes présentent également des caractéristiques provenant de leur hôte eucaryotique. Ainsi, de nombreux gènes organellaires (particulièrement chez le chloroplaste) sont organisés en opérons exprimés sous forme de transcrits polycistroniques comme chez l'ancêtre procaryotique. Chez la mitochondrie, les gènes des ARNr 5S et 18S, et souvent *nad3* et *rps12* sont co-transcrits afin d'être produits en quantité stœchiométrique bien qu'ils puissent agir dans des complexes différents (Bonen et Gray, 1980; Gualberto *et al.*, 1988; Giegé et Brennicke, 2001). A la différence des organelles, les opérons bactériens sont toujours constitués de gènes présentant les mêmes fonctions. Les gènes mitochondriaux *rps3* et *rp16* sont également souvent co-transcrits. Les transcrits mono- et polycistroniques sont ensuite maturés avant d'être traduits (**figure 13**). Les transcrits polycistroniques sont ensuite clivés en ARN monocistroniques par des exo- et endonucléases, comme la RNase J et la RNase Z dans le chloroplaste (Luro *et al.*, 2013). Dans la mitochondrie, la majorité des gènes est d'abord transcrite en ARN précurseurs plus longs qui sont maturés en 5' et en 3' par des nucléases, comme la PNPase, la RNase II et la RNase Z (Perrin *et al.*, 2004; Small *et al.*, 2013). Certains transcrits organellaires possèdent également des introns qui doivent être épissés. Ces introns appartiennent généralement au groupe II. De rares introns du groupe I sont retrouvés dans certains gènes mais ont généralement perdu leur activité autocatalytique et nécessitent l'assistance de facteurs protéiques additionnels (Cho *et al.*, 1998; Brown *et al.*, 2014; Massel *et al.*, 2016). Les introns se retrouvent surtout dans les gènes codant des protéines, mais peuvent aussi être présent dans les gènes des ARNt chez le chloroplaste (Steinmetz *et al.*, 1982). Ils peuvent avoir une organisation cis, lorsqu'introns et exons sont présents sur le même ARN précurseur, ou une organisation trans, lorsque les exons sont fragmentés sur plusieurs transcrits devant être assemblés. Par exemple, le gène mitochondrial *nad5* possède cinq exons fragmentés sur trois unités transcriptionnelles (**figure 13**) (Pereira de Souza *et al.*, 1991). Enfin,

de nombreux transcrits organellaires sont également édités. Chez *Arabidopsis thaliana*, 34 cytidines chloroplastiques et 400 à 500 cytidines mitochondriales peuvent être éditées en uridines lors de la maturation post-transcriptionnelle des ARN, le plus souvent messagers (Handa, 2003; Knoop, 2011; Takenaka *et al.*, 2013). L'édition d'une base entraîne généralement une modification de la séquence protéique traduite, pouvant permettre la restauration d'un acide aminé conservé chez les homologues bactériens (Covello et Gray, 1989; Gualberto *et al.*, 1989; Hiesel *et al.*, 1989). L'édition des ARN permet également d'éviter la traduction précoce des précurseurs ARN polycistroniques, de modifier l'expression de certains gènes par l'édition de codons d'initiation, ou encore de réguler l'épissage (Hoch *et al.*, 1991; Carrillo et Bonen, 1997; Del Campo *et al.*, 2002).

Les protéines PRR (*pentatricopeptide repeat*) jouent un rôle important dans toutes les étapes de maturation post-transcriptionnelles évoquées (Lurin *et al.*, 2004; de Longevialle *et al.*, 2010; Barkan et Small, 2014). Elles sont majoritairement retrouvées chez les organelles de plantes, avec plus de 400 protéines identifiées chez *Arabidopsis thaliana* et *Oryza sativa* pour moins d'une dizaine chez les animaux (Lurin *et al.*, 2004). Ces protéines possèdent des séquences répétées en tandem dégénérées de 35 acides aminés appelées motifs PPR (Small et Peeters, 2000). Individuellement, les motifs PPR adoptent une structure en paire d'hélices- $\alpha$  antiparallèles similaire à celle des motifs TPR (*tetratricopeptide repeat*). A l'échelle de la protéine, les motifs PPR forment une super-hélice droite capable d'interagir avec l'ARN de manière séquence spécifique (Barkan *et al.*, 2012; Yin *et al.*, 2013; Yan *et al.*, 2019). La séquence de chaque motif PPR permet la reconnaissance de nucléotide spécifique. Ainsi, les protéines PPR sont largement utilisées dans les mécanismes de maturation des ARN organellaires comme facteur de reconnaissance des séquences à maturer (Kotera *et al.*, 2005). Certaines protéines PPR possèdent également des domaines présentant une activité catalytique permettant la maturation d'un ARN sans l'assistance d'autres facteurs supplémentaires (Oldenkott *et al.*, 2019). Globalement, les protéines PPR sont impliquées dans la maturation des extrémités des ARNm et des ARNt, dans l'épissage et dans l'édition des transcrits organellaires (Meierhoff *et al.*, 2003; Kotera *et al.*, 2005; Gobert *et al.*, 2010; Boussardon *et al.*, 2014).

#### 4. Les mécanismes de recombinaison

Les génomes mitochondriaux des animaux présentent un taux de mutation plus élevé que leur génome nucléaire conduisant à l'évolution rapide de leur séquence (Brown *et al.*, 1979). Ce haut taux de mutation est dû à l'environnement fortement oxydatif dans la matrice mitochondriale qui peut endommager le génome, à la limitation des mécanismes de réparation de l'ADN dans les mitochondries animales et à la faible fidélité de l'ADN polymérase Pol $\gamma$  mitochondriale (Song *et al.*, 2005). Chez les plantes en revanche, les séquences de l'ADNmt sont généralement très conservées, mais l'organisation du génome est très variable (Wolfe *et al.*,

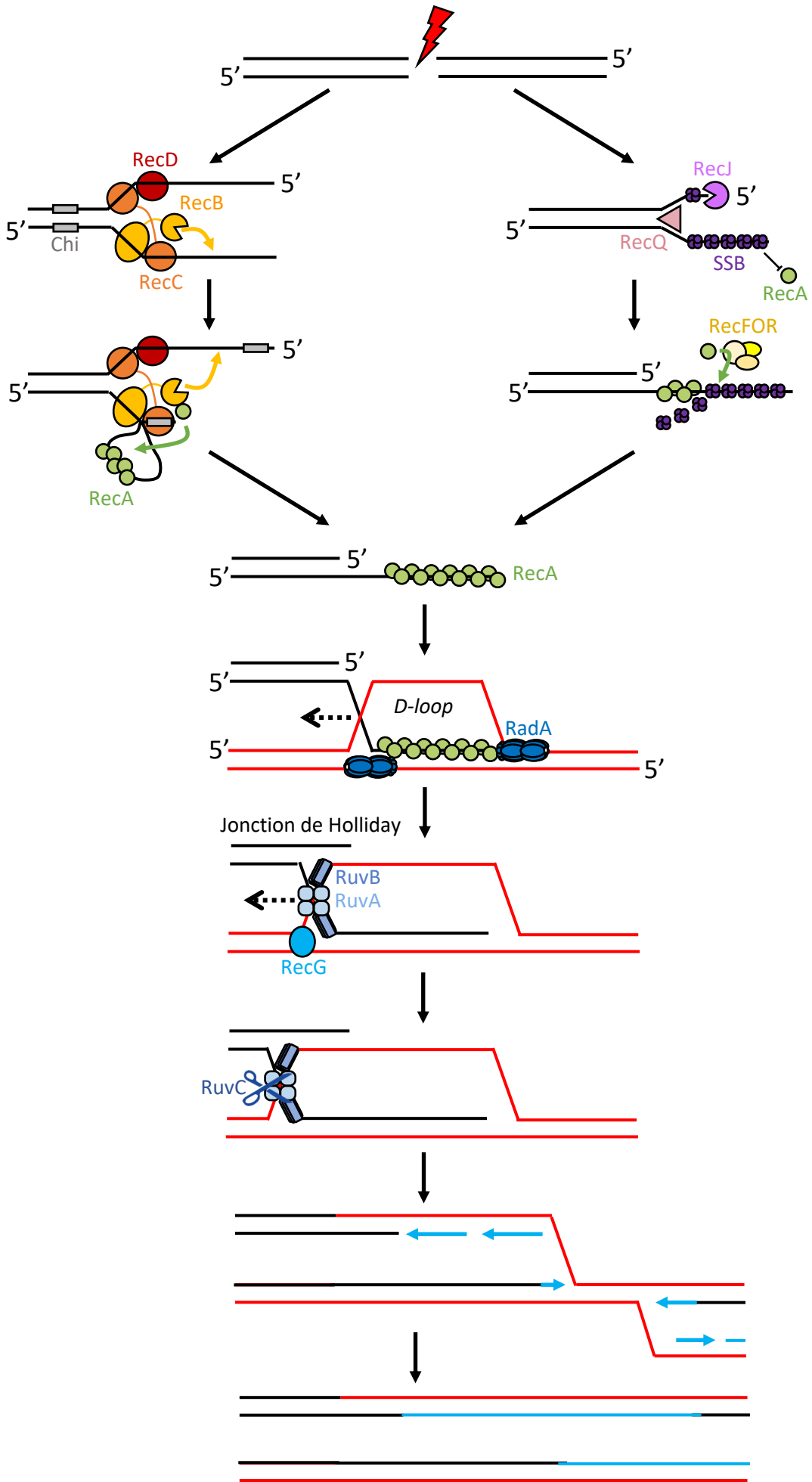
1987; Palmer et Herbon, 1988). La conservation des séquences est néanmoins nettement plus importante pour les séquences codantes que pour les séquences non codantes (Christensen, 2013). L'utilisation de mécanismes fidèles de réparation de l'ADN dans la mitochondrie des plantes a été mis en avant pour expliquer le faible taux de mutation de l'ADNmt, notamment la recombinaison homologue (HR) (Davila *et al.*, 2011).

La HR est un mécanisme de réparation des cassures double brin de l'ADN (DSB : double-strand break) et des lacunes simple brin (SSG : single-strand gap), et de réarrangements du génome contribuant à son évolution. C'est un processus important partagé par une vaste gamme d'organismes, procaryotiques comme eucaryotiques ou même viraux. Les différentes étapes de la HR sont généralement conservées bien que les facteurs puissent varier d'un organisme à l'autre. La première de ces étapes est la formation d'un filament présynaptique capable d'envahir un ADN homologue. Cette invasion, appelée synapse, forme la seconde étape. Enfin, l'étape post-synaptique permet la maturation de la synapse et la séparation des deux molécules d'ADN homologues.

#### 4.1. La recombinaison homologue procaryotique

Chez *Escherichia coli*, la réparation des cassures double brin par HR débute par la résection des extrémités 5'-phosphate de la cassure par le complexe protéique RecBCD (**figure 14**). Les protéines RecB et RecD sont des hélicases 3'-5' et 5'-3', respectivement, dépendantes de l'hydrolyse de l'ATP et permettent l'ouverture de l'hélice d'ADN en se déplaçant sur chacun des brins. RecB possède également une activité exonucléase très processive lui permettant de digérer l'ADN jusqu'à la rencontre d'un site Chi (*crossover hotspot instigator*) (Taylor *et al.*, 2014). Les sites Chi sont des octamères de séquence 5'-GCTGGTGG-3' reconnus par RecC. La reconnaissance d'un site Chi provoque un changement de conformation dans la structure du complexe qui entraîne l'arrêt de l'activité hélicase de RecD. L'activité hélicase de RecB perdure néanmoins et entraîne la formation d'une boucle d'ADN en amont du site Chi (**figure 14**) (Cockram *et al.*, 2015). RecC digère l'extrémité 3' en aval de Chi et clive l'extrémité 5'. Enfin, le complexe RecBCD recrute la recombinase RecA et permet son chargement sur la boucle formée jusqu'à l'arrêt de l'activité hélicase de RecB et la dissociation du complexe. Les protéines RecA s'associent bi-directionnellement avec l'ADN simple brin (ADNsb) à partir d'un noyau de trois à cinq protéines RecA, formant ainsi le complexe présynaptique (**figure 14**) (Galletto *et al.*, 2006; Handa *et al.*, 2009). La nucléation de RecA sur l'ADN est un processus lent. L'extension du noyau en revanche, est beaucoup plus rapide. Chaque protéine RecA se lie avec trois nucléotides en présence d'ATP pour former une hélice droite composée de six protéines RecA par tour d'hélice.

En absence du complexe RecBCD, les extrémités d'une cassure double brin peuvent être prises en charge par l'hélicase 3'-5' RecQ et la nucléase 5'-3' RecJ (**figure 14**) (Morimatsu et Kowalczykowski, 2014). RecQ ouvre la double hélice d'ADN à partir de l'extrémité de la cassure

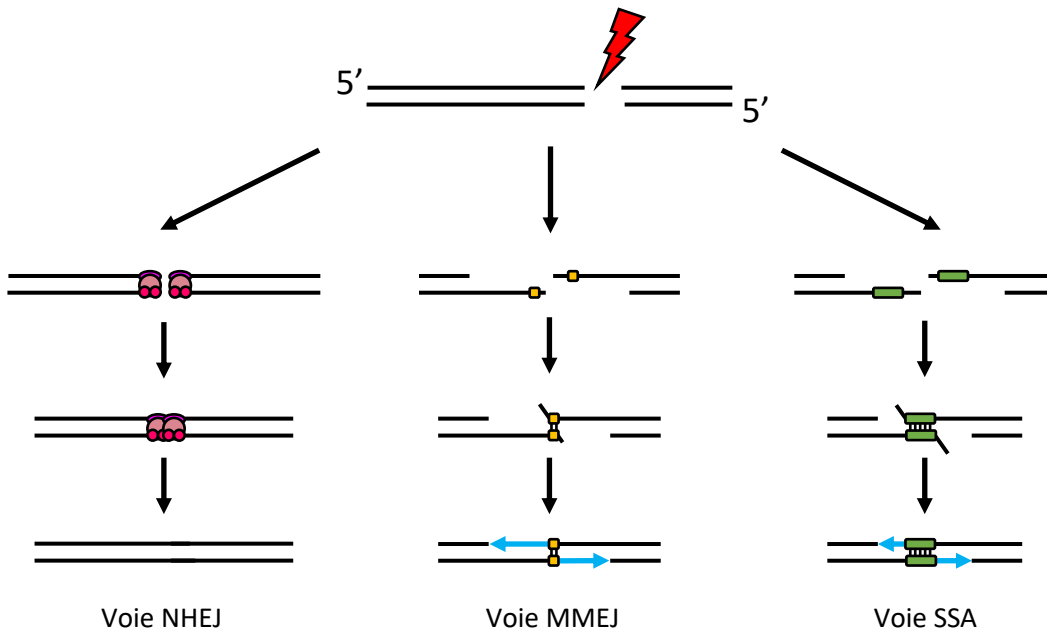


**Figure 14 : Mécanisme de recombinaison homologue bactérien.** La réparation des cassures double brin par recombinaison homologue débute par la résection de ses extrémités. A gauche, la voie RecBCD implique le complexe RecBCD composé de deux hélicases et d'une exonucléase. La nucléase RecB clive l'extrémité 3' jusqu'à la reconnaissance d'un site Chi par RecC. RecB clive alors l'extrémité 5' en amont du site chi tandis que RecD tire l'ADN et forme une boucle permettant le chargement de la recombinase RecA. A droite, la voie RecFOR implique l'hélicase RecQ et l'exonucléase RecJ qui ouvrent et dégradent l'ADN, formant une extrémité d'ADN 3' sortante qui est recouverte de protéine SSB. Le complexe RecFOR est alors recruté pour permettre le remplacement de SSB par RecA. Après la nucléation de RecA, le nucléofilament ADN-RecA s'allonge puis cherche des homologies dans le génome en entrant en contact avec de nombreuses régions de l'ADN de manière aléatoire. Lorsqu'une séquence homologue est trouvée, le nucléofilament l'envahit et forme une structure à trois brin d'ADN (*D-loop*). La migration de branche, précoce *via* RadA, puis plus tardive *via* RecG et RuvAB, permet l'extension de la zone d'homologie et le recrutement du quatrième brin formant ainsi une jonction de Holliday. La jonction de Holliday est résolue, avec ou sans crossover, par la nucléase RuvC recrutée par RuvAB tandis que l'ADN est synthétisé par une ADN polymérase. D'après Cockram et al., 2015.

et permet le chargement de RecJ qui digère l'un des brins d'ADN. Cette résection par RecJ conduit à la formation d'une extrémité simple brin 3' sortante, recouverte de protéines SSB pour protéger l'ADNsb. Les protéines SSB inhibent la nucléation de RecA pour la formation du filament présynaptique et doivent être déplacées pour permettre la HR. Ces protéines SSB sont néanmoins nécessaires à la HR, notamment pour la suppression de structures secondaires (Roy *et al.*, 2009). Le complexe RecFOR agit comme médiateur de la recombinaison, en permettant le chargement de RecA à la place des protéines SSB (**figure 14**) (Umezū et Kolodner, 1994). La protéine RecF permet de localiser les protéines RecO et RecR au niveau de l'ADNsb recouvert de protéines SSB, tandis que le complexe RecOR permet le remplacement des protéines SSB (Sandler et Clark, 1994). Le complexe RecFOR est également impliqué dans le chargement de la protéine RecA au niveau des lacunes simple brin résultant généralement de défauts de réplication (**figure 14**). La réparation des SSG par recombinaison homologue permet de contourner d'éventuels blocages de l'ADN polymérase sur le brin matrice et de restaurer la progression des fourches de réplication bloquées (McGlynn et Lloyd, 2002).

Le complexe présynaptique scanne l'ADN double brin (ADNdb) à la recherche d'une séquence homologue. Pour cela, le nucléofilament de RecA s'associe aléatoirement avec de multiples régions double brin de l'ADN et se déplace jusqu'à la rencontre d'une séquence homologue identifiée par l'appariement Watson-Crick des bases (Forget et Kowalczykowski, 2012; Ragnathan *et al.*, 2012). L'invasion de l'ADNdb hétérologue par le nucléofilament de RecA forme une structure triple brin appelée D-loop (**figure 14**). Lorsque la séquence homologue est identifiée, l'hydrolyse d'ATP permet l'appariement des nucléotides entre le brin d'ADN envahissant avec l'ADN récepteur. La formation ou le maintien de la D-loop peuvent être inhibés par de multiples protéines régulant la recombinaison homologue, telles que UvrD ou RecX, dépolymérisant ou bloquant l'extension du nucléofilament de RecA, respectivement (Cárdenas *et al.*, 2012; Petrova *et al.*, 2015).

Après l'invasion de l'ADN homologue par le nucléofilament ADNsb-RecA, la région d'appariement est étendue lors de la phase de migration de branche. L'extension de la zone d'homologie assure la recombinaison entre les séquences effectivement homologues et non les séquences possédant uniquement une courte région identique. Dans le cas où l'homologie est trop faible, le mécanisme de recombinaison homologue est avorté (Mawer et Leach, 2014). L'étape de migration de branche fait intervenir plusieurs facteurs capables de maturer les différents intermédiaires de recombinaison jusqu'à la résolution du duplex d'ADN (**figure 14**). Ces facteurs sont impliqués dans différentes voies de maturation spécifiques et partiellement redondantes (Cooper et Lovett, 2016). La première voie de maturation de la D-loop fait intervenir la protéine RadA, un paralogue de la protéine RecA. Des hexamères de RadA sont recrutés par RecA et chargés de part et d'autre de la D-loop sur un brin de l'ADNdb récepteur (Marie *et al.*, 2017). Ces hexamères possèdent une activité hélicase dépendante de l'ATP qui



**Figure 15 : Mécanismes de liaison d'extrémités.** Les cassures double brin peuvent être réparées par des mécanismes non fidèles et indépendants d'une recombinaise. **A gauche**, la voie *non homologous end joining* (NHEJ) implique deux hétérotrimères (Ku70, Ku80 et DNA-PKc) qui se lient aux extrémités de la DSB pour les rapprocher l'une de l'autre. Après traitement des extrémités pour les rendre franches par XRCC4, les extrémités sont reliées par une ADN ligase. Ce mécanisme entraîne généralement la suppression de plusieurs nucléotides de part et d'autre de la cassure. **Au milieu**, la voie *microhomology-mediated end joining* (MMEJ) utilise des séquences homologues de petite taille (5-30 pb) pour joindre les deux extrémités d'une DSB et permettre la synthèse des brins complémentaires manquants. Ce mécanisme conduit également à la perte des séquences de part et d'autre du motif répété. La voie *microhomology-mediated break-induced replication* (MMBIR) est une variante du MMEJ où l'ADN est entièrement répliqué à partir de la séquence répétée. **A droite**, la voie *single-strand annealing* (SSA) est similaire au MMEJ mais utilise des séquences répétées de taille supérieure.

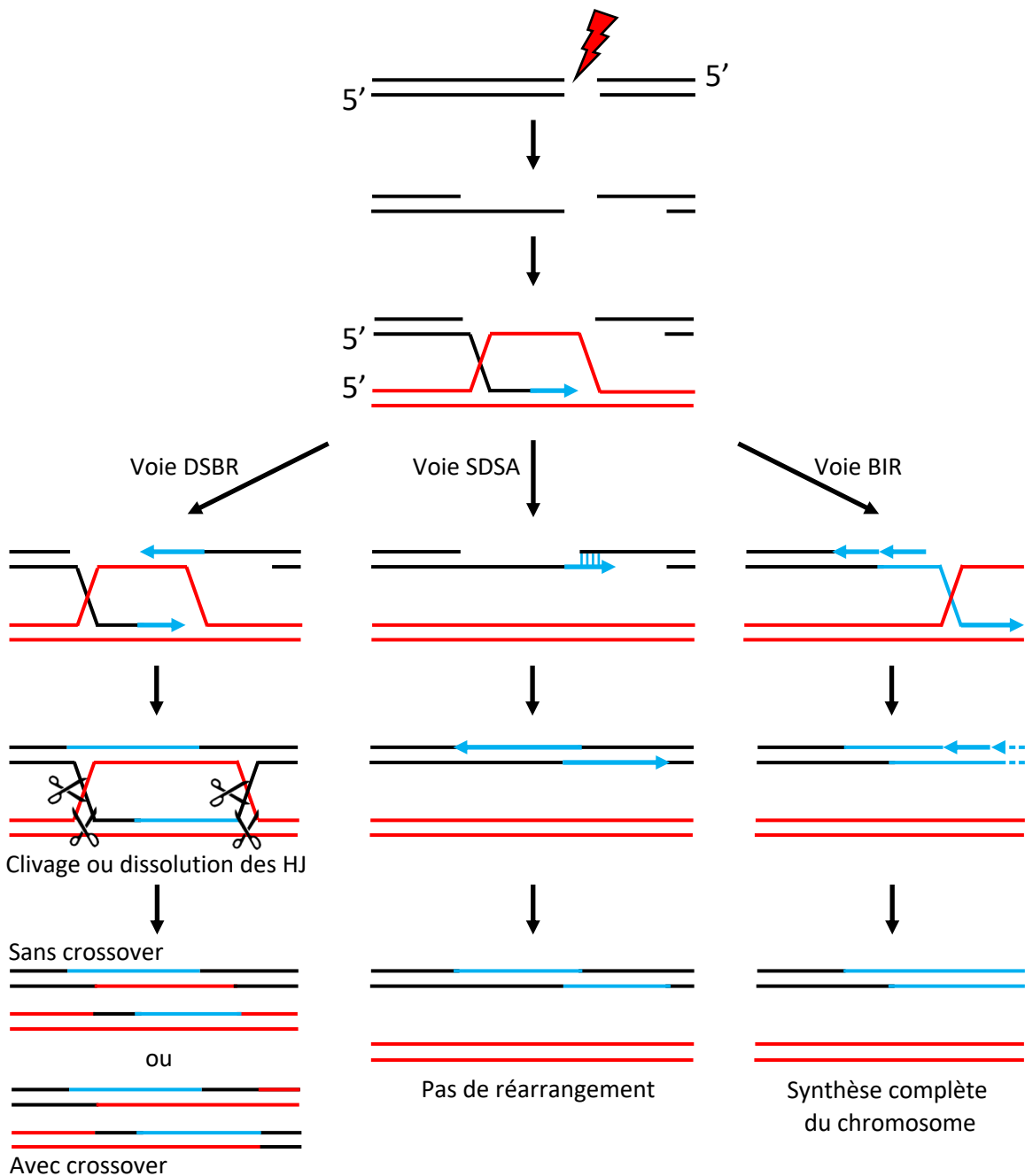


délie l'ADNdb récepteur, étendant la D-loop dans les deux sens sur de longues distances. L'ouverture de l'ADNdb récepteur permet son appariement avec l'ADNsb donneur envahissant. La protéine RadA pourrait également déstabiliser le filament de RecA pour orienter le processus de recombinaison homologues vers ses étapes tardives. La seconde voie de maturation des intermédiaires de recombinaison implique la protéine RecG (Whitby *et al.*, 1994). RecG est recrutée par les protéines SSB et agit sous forme de monomère en déroulant l'ADNdb récepteur pour l'hybrider au brin complémentaire homologue. Contrairement à RadA, RecG est une hélicase dépendante de l'ATP se déplaçant sur l'ADNdb. La protéine RecG pourrait ainsi permettre le recrutement du deuxième brin de l'ADN envahissant, convertissant la D-loop en une structure d'ADN à quatre brins appelée jonction de Holliday (HJ : *Holliday junction*) (Mawer et Leach, 2014). Enfin, les HJ sont maturées par le complexe RuvAB composé d'un tétramère de RuvA et de deux hexamères de RuvB (West, 1997). Le tétramère de RuvA se lie spécifiquement au croisement des quatre brins d'ADN, stabilisant sa structure, et recrute RuvB de part et d'autre de la HJ. Chaque hexamère de RuvB est ancré sur l'un des brins d'ADN et se déplace vers l'autre hexamère grâce à son activité hélicase dépendante de l'ATP. L'ancrage des hexamères entraîne ainsi le déplacement de la HJ dans un sens, ou dans l'autre. La jonction est déplacée jusqu'au recrutement d'un dimère de protéines RuvC qui va cliver la HJ au niveau d'une séquence consensus 5'-<sup>A</sup>/TT<sup>G</sup>/C-3'. Le clivage de la HJ permet la résolution de l'hétéroduplex via la séparation des deux molécules d'ADN (**figure 14**).

#### 4.2. La recombinaison homologue eucaryotique

Les mécanismes nucléaires de HR chez les mammifères et la levure sont assez proches de ceux retrouvés chez les bactéries. Ils sont régulés par le cycle cellulaire qui limite la HR aux phases G2 et S où les chromatides sœurs sont disponibles (Orthwein *et al.*, 2015). Cette limitation favorise la recombinaison entre séquences véritablement homologues. En dehors de ces phases, des mécanismes de jonction d'extrémités tels que le *non homologous end joining* (NHEJ) ou le *microhomology-mediated end joining* (MMEJ) sont privilégiés pour la réparation des DSB bien qu'ils puissent être délétères pour le génome (**figure 15**) (Chang *et al.*, 2017a).

La réparation d'une cassure double brin débute par la résection des extrémités de la lésion par le complexe MRN (MRE11 ; RAD50 ; NBS1) et MRX (Mre11 ; Rad50 ; Xrs2) chez l'homme et la levure, respectivement (Johzuka et Ogawa, 1995; Ranjha *et al.*, 2018). La nucléase MRE11/Mre11, assistée par RAD50/Rad50, clive l'ADN en amont de la lésion puis digère l'ADN de 5' vers 3' vers la lésion afin de contourner les protéines fixées à l'extrémité, formant une extrémité 3' sortante simple brin (Neale *et al.*, 2005). Après cette courte résection, les exonucléases 5'-3' EXO1/Exo1 ou DNA2/Dna2 sont recrutées pour dégrader davantage l'ADN et former une extrémité 3' sortante simple brin de plusieurs milliers de nucléotides. Cette extension de la résection permet d'orienter la réparation vers la recombinaison homologue



**Figure 16 : Voies alternatives de terminaison de la recombinaison homologue.** Après résection des extrémités d'une cassure double brin, et invasion de l'ADN homologue par le nucléofilament de RecA/RAD51, plusieurs voies de résolution du mécanisme peuvent se mettre en place. **A gauche**, la voie *double-strand break repair* (DSBR), chez les eucaryotes, implique le recrutement de la deuxième partie de la DSB et la formation d'une deuxième jonction de Holliday. Ces deux HJ pourront être résolues par dissolution, en migrant l'une vers l'autre avant d'être prises en charge par une topoisomérase, ou bien clivées horizontalement ou verticalement, conduisant à un, deux ou aucun crossover. **Au milieu**, la voie *synthesis-dependant strand annealing* (SDSA) permet de recruter rapidement la seconde extrémité de la DSB après synthèse de la séquence manquante grâce à l'ADN homologue. La synthèse des brins complémentaires utilise directement l'ADN d'origine et n'entraîne pas de réarrangement. **A droite**, la voie *break-induced replication* (BIR) utilise l'ADN homologue pour re-synthétiser entièrement l'ADN ayant subi une DSB. Ce mécanisme peut conduire à des réarrangements importants lorsque les deux molécules d'ADN ne sont pas complètement homologues.

plutôt que vers d'autres mécanismes de jonction d'extrémités (Shibata, 2017). Lors de la résection, l'ADNsb est recouvert de protéines RPA qui le protège des nucléases cellulaires.

Comme la protéine RPA possède une affinité plus élevée pour l'ADNsb que la recombinaise RAD51/Rad51 (homologue fonctionnelle de RecA), la formation du filament présynaptique nécessite l'action d'un médiateur similaire au complexe RecFOR chez *E. coli* (Benson *et al.*, 1998). La protéine BRCA2/Rad52 inhibe l'activité ATPase de la recombinaise, favorisant sa liaison avec l'ADN, et permet le remplacement de RPA par RAD51/Rad51 (Benson *et al.*, 1998; Yang *et al.*, 2005). Tout comme RecA, les protéines RAD51/Rad51 vont former un filament sur l'ADNsb à partir d'un noyau. La nucléation et la formation du nucléofilament de RAD51/Rad51 sont régulées par de nombreuses protéines, comme les complexes BCDX2 et CX3 chez l'Homme ou la translocase Srs2 chez la levure, qui favorisent ou déstabilisent le filament, respectivement (Masson *et al.*, 2001; Krejci *et al.*, 2003). Le mécanisme de recherche de la séquence homologue par le filament présynaptique n'est pas encore élucidé. L'hypothèse actuelle considère qu'il y a une association temporaire du filament présynaptique avec de multiples loci d'ADN jusqu'au décèlement d'une région présentant suffisamment d'homologie. La protéine BRCA1 permettrait alors l'invasion de brin et l'association avec les régions alentours pour augmenter et vérifier l'homologie (Zhao *et al.*, 2017). Une fois la D-loop formée, la recombinaise RAD51/Rad51 est délogée par la translocase à ADNdb RAD54 afin d'augmenter la stabilité de la structure triple brin. Les polymérases  $\delta$  ou  $\epsilon$  synthétisent alors l'ADN au niveau de l'extrémité 3' en utilisant la séquence homologue comme matrice (Solinger *et al.*, 2002).

Le mécanisme de recombinaison homologue peut ensuite évoluer vers trois voies différentes : le *break-induced replication* (BIR), le *synthesis-dependant strand annealing* (SDSA) ou la voie canonique de *double-strand break repair* (DSBR) (**figure 16**). Lors du BIR, la synthèse d'ADN va continuer jusqu'à l'extrémité du chromosome (Sakofsky et Malkova, 2017). Comme la polymérase utilisée n'est pas assistée de facteurs augmentant la fidélité de synthèse, le brin synthétisé présentera un taux de mutations plutôt élevé. Aussi, le BIR est rarement privilégié si d'autres mécanismes sont possibles. Si la D-loop est déstabilisée par certaines hélicases, la recombinaison homologue est dirigée vers la voie du SDSA. Après synthèse d'une courte séquence d'ADN, l'hétéroduplex se sépare et la séquence néosynthétisée s'hybride avec la seconde extrémité de la cassure double brin (Nassif *et al.*, 1994). Le brin d'ADN manquant est ensuite synthétisé en utilisant le brin complémentaire comme matrice. Si la D-loop est stabilisée, le mécanisme est dirigé vers le DSBR (Nimonkar *et al.*, 2009). La seconde extrémité de la cassure double brin est recrutée par l'ADNsb de la D-loop grâce à la protéine BRCA2/Rad52. L'ADN complémentaire du second brin lésé est synthétisé en utilisant le second brin de l'ADN homologue comme matrice. La voie du DSBR mène à la formation de deux HJ qui sont généralement éliminées par dissolution, pour limiter les réarrangements génomiques. Pour cela, les deux HJ sont déplacées l'une vers l'autre par les complexes BTRR

(BLM-TOPOIII $\beta$ -RMI1-RMI2) et STR (Sgs1-Top3-Rmi1) chez l'homme et la levure, respectivement (Chen *et al.*, 2014; Matos et West, 2014). Les contraintes de l'ADN liées à la migration des HJ sont relâchées par la topoisomérase jusqu'à la formation d'un hémicaténane, une structure où un brin de l'ADNdb est enroulé autour du brin du second ADNdb en un point, qui est ensuite résolue par RMI1/Rmi1. Dans certains cas la dissolution des HJ est impossible. Ces structures à quatre brins sont alors résolues par des résolvases, dont la principale est GEN1/Yen1 (Ip *et al.*, 2008). Ces nucléases forment un dimère au centre de la HJ et réalisent deux incisions sur l'un ou l'autre de ses bras, conduisant ou non à la formation de crossover.

Comme pour les autres eucaryotes, la réparation des cassures double brin dans le génome nucléaire des plantes utilise principalement des mécanismes de jonction d'extrémités (NHEJ, MMEJ) plutôt que la recombinaison homologue (Puchta, 1999). La machinerie de recombinaison est très conservée au sein des organismes et les plantes présentent des homologues à pratiquement tous les facteurs énoncés précédemment (Schuermann *et al.*, 2005). La HR chez les plantes est liée à d'autres mécanismes tels que le nucleotide excision repair (NER) ou la photomorphogénèse (Lieberman *et al.*, 2004; Molinier *et al.*, 2008).

#### 4.3. La recombinaison homologue organellaire

Initialement, on pensait que les organelles ne possédaient pas de mécanisme de réparation de l'ADN (Karol et Simpson, 1968). Chaque cellule contenant plusieurs exemplaires des génomes organellaires, l'endommagement d'une copie ne serait pas délétère pour la cellule. A la place de réparer le génome endommagé, celui-ci serait alors plutôt détruit. Depuis, de nombreux mécanismes de réparation de l'ADN ont été identifiés dans les organelles. Parmi ces mécanismes, la recombinaison a bien été caractérisée chez les levures et les plantes, mais semblait absente dans les mitochondries des mammifères (Birky, 2001). Néanmoins, plus récemment, la jonction d'extrémités et la recombinaison homologue ont été décrites chez les mammifères (Tadi *et al.*, 2016; Dahal *et al.*, 2018).

Les mécanismes de jonctions d'extrémités et de recombinaison homologue dans les mitochondries des levures et des mammifères semblent utiliser les mêmes facteurs que les mécanismes nucléaires (Stein et Sia, 2017). De nombreux facteurs sont doublement adressés au noyau et aux mitochondries après leur traduction dans le cytoplasme. S'ils utilisent les mêmes protéines, les mécanismes organellaires semblent néanmoins différents des nucléaires, et sont encore peu caractérisés. Le MMEJ et le DSBR semblent être les voies principales de réparation des DSB, alors que le NHEJ n'est pas détecté dans les mitochondries (Tadi *et al.*, 2016; Dahal *et al.*, 2018). La HR pourrait être liée à la réparation couplée à la réplication, tandis que le MMEJ serait responsable de la délétion de séquences mitochondriales observées dans certains mutants (Fukui et Moraes, 2009).

Tableau 1 : protéines impliquées dans la recombinaison homologues des organelles. D'après Gualberto & Newton, 2017

<b>Fonction moléculaire</b>	<b><i>E. coli</i></b>	<b><i>A. thaliana</i></b>	<b>Localisation</b>
Résection de l'ADN	RecBCD	?	-
Protection de l'ADN	SSB	SSB1	Mt + Cp
		SSB2	Mt
Chargement de la recombinaison	RecFOR	ODB1	Mt
		WHY1	Cp
		WHY2	Mt
		WHY3	Cp
		OSB1	Mt
		OSB2	Cp
		OSB3	Mt + Cp
		OSB4	Mt
Régulation de la recombinaison	RecX	RECX	Mt + Cp
	UvrD	-	-
	MutH	MSH1	Mt + Cp
	MutS		
Formation de la synapse	RecA	RECA1	Cp
		RECA2	Mt + Cp
		RECA3	Mt
Migration de branche	RuvAB	-	-
	RecG	RECG1	Mt + Cp
	RadA	RADA ?	?
Résolution des HJ	RuvC	?	-
Synthèse d'ADN	Poll	POL1B	Mt + Cp
Ligation de l'ADN	LigA	LIG1	Mt + Cp

Dans les mitochondries des plantes, la HR est la voie principale de réparation de l'ADN alors que les mécanismes de jonction d'extrémités sont peu utilisés (Gualberto et Newton, 2017). De par le nombre de séquences répétées dans l'ADNmt, la recombinaison homologue joue un rôle important dans le maintien de la séquence des gènes et dans l'évolution rapide de la structure du génome. Elle doit être finement régulée pour éviter des réarrangements génomiques pouvant être délétères pour la mitochondrie. La mutation de facteurs impliqués dans la HR et sa régulation entraîne généralement d'importants réarrangements de l'ADNmt (Miller-Messmer *et al.*, 2012; Wallet *et al.*, 2015). La HR joue également un rôle important dans le chloroplaste où elle permet de maintenir la structure du génome (Cerutti *et al.*, 1995; Kwon *et al.*, 2010).

Dans les organelles de plantes, la HR implique de nombreux facteurs d'origine procaryotique codés par le noyau, qui ne sont pas impliqués dans les mécanismes de recombinaison nucléaires (Gualberto et Newton, 2017). Le mécanisme de recombinaison dans les organelles est considéré similaire à celui retrouvé chez les bactéries, bien qu'il soit encore assez peu compris. Les étapes précoces et tardives sont particulièrement nébuleuses car les protéines impliquées dans la résection des extrémités des DSB et dans la résolution des HJ n'ont toujours pas été identifiées. Comme plusieurs facteurs adressés aux organelles semblent avoir des activités redondantes, il est difficile de distinguer le rôle exact de chacun. Ainsi, la recombinaison homologue organellaire implique l'existence de nombreuses protéines de liaison à l'ADNsb et plusieurs homologues de la recombinaison bactérienne RecA. Ces protéines pourraient aussi jouer un rôle dans des voies alternatives de recombinaison prenant en charge différents substrats et/ou les menant vers des différents produits. Les protéines ayant un rôle dans la recombinaison homologue des organelles sont résumées dans le **tableau 1**.

Après résection, les extrémités d'ADNsb sont protégées de l'action des nucléases par des protéines de liaison à l'ADN. Cette fonction pourrait être remplie par les protéines SSB1 et SSB2, homologues de la protéine SSB bactérienne. La différence de fonction entre les deux protéines n'est pas encore connue, mais toutes deux sont nécessaires au processus de recombinaison (Edmondson *et al.*, 2005). La protection de l'ADNsb et la promotion de la HR pourraient aussi être assurées par les protéines OSB (organellar single-stranded binding protein) ou les protéines WHY de la famille whirly (Zaegel *et al.*, 2006; Cappadocia *et al.*, 2010). Ces protéines sont adressées à la mitochondrie (OSB1, OSB4, WHY2), au chloroplaste (OSB2, WHY1, WHY3) ou aux deux organelles (OSB3), et possèdent une meilleure affinité pour l'ADNsb que les protéines SSB1 et SSB2. Les protéines OSB et WHY ne peuvent cependant pas interagir avec la polymérase organellaire POL1B, impliquée dans les mécanismes de réparation (Garcia-Medel *et al.*, 2019). Aussi, Garcia-Medel *et al.*, ont proposé des voies alternatives pour les protéines SSB1/SSB2 et OSB/WHY. Lors d'une DSB, les protéines SSB1 et SSB2 pourraient promouvoir l'association de POL1B avec les extrémités de la cassure, favorisant la réparation par le MMEJ,

alors qu'OSB1 et WHY2 inhibent les mécanismes de jonction d'extrémités en limitant l'accès aux SSB et à la polymérase (Garcia-Medel *et al.*, 2019). La mutation des facteurs OSB1 et OSB4 entraîne une augmentation de la recombinaison avec les séquences répétées de taille intermédiaire dans la mitochondrie propice aux réarrangements génomiques, alors que la perte des WHY entraîne une augmentation de la recombinaison impliquant des micro-homologies, autant dans la mitochondrie (mutants *why2*) que dans le chloroplaste (double mutant *why1 why3*) et inhibe la réparation par HR des cassures induites par la ciprofloxacine (Zaegel *et al.*, 2006; Marechal *et al.*, 2009; Cappadocia *et al.*, 2010).

Le chargement de la recombinase sur l'ADNsb pour former le filament présynaptique nécessite le remplacement des protéines de liaison à l'ADNsb. Les protéines ODB1 et ODB2 (*organellar DNA-binding protein*) sont des homologues du facteur Rad52 de la levure et pourraient remplir cette fonction dans la mitochondrie et le chloroplaste, respectivement (Samach *et al.*, 2011; Janicka *et al.*, 2012). La protéine ODB1 favorise l'hybridation des séquences d'ADNsb complémentaires *in vitro*, comme en est capable la protéine Rad52, appuyant le rôle de médiateur de la recombinaison d'ODB1 (Janicka *et al.*, 2012). Dans la mitochondrie, la protéine WHY2 pourraient également jouer le rôle de médiateur en permettant le chargement de la recombinase (Cappadocia *et al.*, 2010). ODB1 et WHY2 pourraient alors être redondants ou permettre le recrutement de différents facteurs impliqués dans différentes voies alternatives de recombinaison. Les protéines ODB et WHY sont co-purifiées avec les génomes organellaires, et sont considérées comme étant des composants importants des nucléoïdes mitochondriaux et chloroplastiques (Pfalz et Pfanschmidt, 2013; Gualberto et Kühn, 2014). Tout comme pour WHY2, la mutation d'ODB1 inhibe la réparation par HR des cassures induites par des traitements génotoxique, entraînant l'augmentation de la recombinaison impliquant des micro-homologies (Janicka *et al.*, 2012).

Chez *Arabidopsis* trois recombinases homologues de la RecA bactérienne sont adressées aux organelles : RECA1 est chloroplastique, RECA3 est mitochondriale et RECA2 est adressée aux deux compartiments. Les orthologues de RECA3 ne se trouvent que chez les angiospermes. Dans la mitochondrie, les protéines RECA2 et RECA3 semblent impliquées dans différentes voies de recombinaison pouvant être partiellement redondantes (Miller-Messmer *et al.*, 2012). RECA2 et RECA3 peuvent partiellement compléter le mutant bactérien *recA* pour la survie après des traitements génotoxiques. Les deux mutants *recA2* et *recA3* montrent une augmentation de la recombinaison ectopique impliquant les IR, mais *recA2* est létale très tôt dans le développement de la plantule alors que *recA3* n'affecte que peu le développement de la plante dans la première génération des mutants homozygotes, mais augmente la sensibilité de la plante aux stress génotoxiques (Miller-Messmer *et al.*, 2012). Cela suggère que RECA2 pourrait être la recombinase principale, impliquée dans la surveillance de la recombinaison homologue, alors que RECA3 participerait surtout à la réparation de l'ADNmt, notamment dans

des voies sans crossover tels que le SDSA (Miller-Messmer *et al.*, 2012). L'absence de RECA2 entraînerait des réarrangements trop importants de l'ADNmt lié à l'augmentation de la recombinaison ectopique. La protéine RECA3 ne possède pas l'extension C-terminale acide typique des protéines RecA, nécessaire pour l'interaction avec les protéines SSB et RecX chez *E. coli* (Eggler *et al.*, 2003; Drees *et al.*, 2004), et pourrait ainsi échapper à la régulation par certains facteurs mitochondriaux. Dans le chloroplaste, RECA1 est impliqué dans un processus limitant l'hybridation de l'ADN avec une micro-homologie en sens inverse lors du blocage d'une fourche de réplication, en favorisant des processus légitimes tels que la régression de fourche (Zampini *et al.*, 2015). Les mutations de RECA1 ou de RECA3 sont très synergétiques avec la mutation des facteurs WHY et OSB. Les plantes double mutantes *odb1 recA3* sont très affectées dans leur développement, suggérant un impact très délétère de la suppression des deux facteurs (résultats non-publiés du laboratoire). Des homologues aux facteurs bactériens MutS et RecX, MSH1 et RECX sont adressés aux deux organelles et pourraient réguler la formation du filament présynaptique (Abdelnoor *et al.*, 2006; Odahara et Sekine, 2018).

Après l'invasion de l'ADN homologue, la région d'appariement est étendue par l'ouverture de l'ADNdb récepteur grâce à des hélicases. La voie de migration de branche bactérienne RuvAB est absente des organelles des plantes. Aucun homologue aux facteurs RuvA et RuvB n'est codé dans le génome des plantes. En revanche, des homologues aux protéines RecG et RadA, décrits comme partiellement redondants avec RuvAB, sont présents dans le génome de toutes les plantes (Cooper *et al.*, 2015; Gualberto et Newton, 2017). La protéine RECG1, homologue de RecG, est adressée aux chloroplastes et aux mitochondries (Wallet *et al.*, 2015). Elle complémente partiellement le mutant bactérien *recG* dans la survie aux UV (ultraviolets), la suppression de la ré-initiation pathologique de la réplication et l'inhibition de la réplication utilisant une amorce ARN. RECG1 doit assurer les mêmes fonctions de réparation et de surveillance de la recombinaison dans les organelles que chez la bactérie. Pourtant, la mutation *recG1* impacte faiblement le développement de la plante, mais augmente la recombinaison ectopique entre IRs dans l'ADNmt (Wallet *et al.*, 2015). Chez les bactéries, la survie à des agents génotoxiques est peu affectée lorsqu'une seule voie de migration est mutée (Beam *et al.*, 2002). La survie diminue cependant drastiquement lorsque deux voies sont affectées. Avec l'absence de RuvAB dans les organelles, l'absence de phénotype développemental pour *recG1* est plutôt surprenante, surtout au regard de la mutation *recA2*, létale au stade de pousse (Miller-Messmer *et al.*, 2012). Cela suggère qu'une autre voie de migration de branche joue un rôle plus important dans la recombinaison homologue organellaire. Cette voie pourrait impliquer la protéine RADA dont un adressage mitochondrial est prédit, mais qui n'a pas encore été étudié.

L'absence de la voie RuvAB pour la migration de branche dans les organelles entraîne également l'absence de la résolvasse RuvC impliquée dans la résolution des HJ. Cependant,



toutes les bactéries n'utilisent pas RuvC, mais des résolvases alternatives telle que YqgF (Aravind *et al.*, 2000). Un homologue d'YqgF est retrouvé chez les plantes et pourrait permettre la résolution des HJ dans les organelles. Il est également possible que les HJ ne soient pas résolues mais dissolues grâce aux facteurs de migration de branche et à la topoisomérase I, comme dans la HR nucléaire. Il a également été proposé que RECG1 pourrait résoudre les HJ en les déplaçant jusqu'aux fourches de réplication (Wardrope *et al.*, 2009; Wallet *et al.*, 2015).

#### 4.4. Évolution de la structure des génomes organellaires des plantes

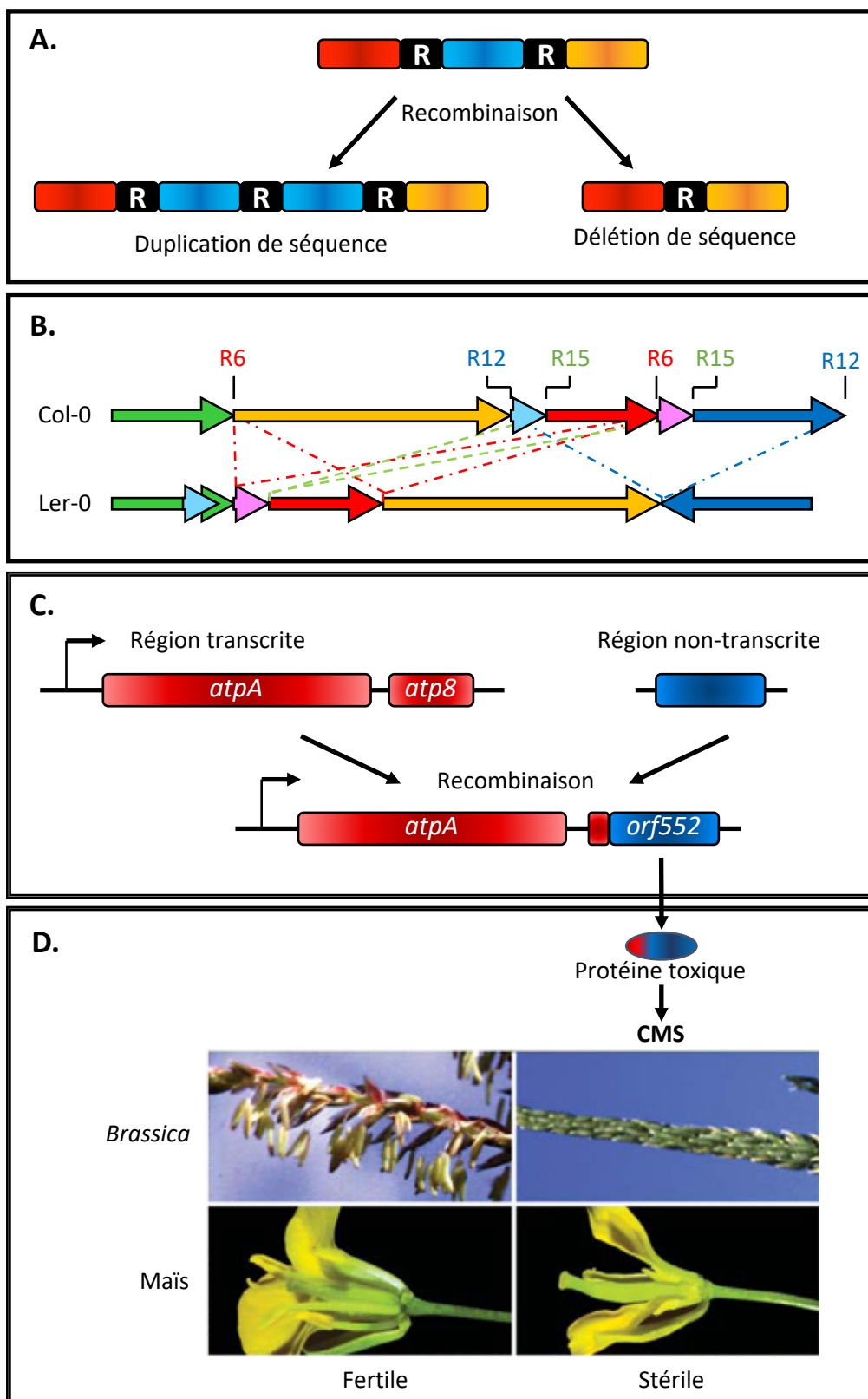
Les génomes mitochondriaux des plantes présentent généralement un taux de mutation très bas par rapport aux génomes nucléaires ou mitochondriaux des animaux et levures (Wolfe *et al.*, 1987). En revanche, ils montrent un taux de réarrangement génomique excessivement haut, conduisant à la faible conservation de la synténie, même au sein d'une même espèce (Palmer et Herbon, 1988). Cette évolution de la structure est très rapide, pouvant apparaître d'une génération à l'autre, et peut conduire à des modifications phénotypiques particulières telles que la stérilité mâle cytoplasmique (CMS : *cytoplasmic male sterility*) (Woloszynska, 2010; Touzet et Meyer, 2014). De par le nombre de séquences répétées dans l'ADNmt, la HR est le processus responsable de cette évolution, aussi bien en corrigeant fidèlement les mutations génétiques qu'en permettant le réarrangement des séquences (Christensen, 2013, 2014).

Si les séquences géniques sont très conservées, les séquences non codantes de l'ADNmt le sont beaucoup moins. Ces régions sont difficilement alignables tant elles varient d'une espèce à l'autre. Cette difficulté d'aligner les séquences non codantes complique l'étude de leur taux de mutation. Néanmoins, en comparant les ADNmt de deux écotypes d'*A. thaliana* ayant récemment divergés, il a été montré que les séquences non codantes présentaient un taux de mutation plus élevé que les séquences codantes (Christensen, 2013). Pourtant, la présence de mécanismes de réparation préférentiellement des régions transcrites n'a encore jamais été démontrée dans les organelles de plantes (Gualberto et Newton, 2017). Cette conservation des régions codantes proviendrait davantage de la sélection spécifique des ADNmt ne présentant pas d'altération de fonction (Christensen, 2013). La pression de sélection sur les régions non codantes étant moindre, voire nulle, les altérations de séquences et de structures ne sont pas forcément contre-sélectionnées. Aussi, il a été proposé que les dommages de l'ADNmt puissent être préférentiellement réparés par HR (Christensen, 2013). Les recombinaisons fréquentes utilisent les grandes séquences répétées et n'entraînent pas d'importants réarrangements génomiques. Cependant, lors de stress, les séquences répétées de plus petites tailles peuvent être utilisées par des mécanismes de réparation moins fidèles et plus propices aux réarrangements tels que le BIR et le MMEJ (Miller-Messmer *et al.*, 2012; Wallet *et al.*, 2015). Ces mécanismes vont induire des délétions ou, plus fréquemment, des duplications de séquences (Christensen, 2013). A moins que ces réarrangements soient délétères pour les

fonctions de la mitochondrie ils seront conservés, conduisant à l'augmentation générale de la taille de l'ADNmt (Christensen, 2013). Chez les animaux, une partie des mitochondries de la lignée germinale femelle ne sont pas utilisées pour éviter les dommages à l'ADN et permettre la transmission d'organelles avec un génome intact (de Paula *et al.*, 2013). Les plantes, ne possédant pas de cellules germinales distinguées, auraient retenu un mécanisme permettant la correction fidèle des séquences géniques en contrepartie de l'augmentation de la taille de l'ADNmt (Christensen, 2013). Ce mécanisme pourrait être à l'origine de la large gamme de taille des ADNmt des plantes de la famille des Silene, allant d'une taille de 253 kb pour *S. latifolia* au plus grand ADNmt connu, de 11,3 Mb pour *S. conica* (Sloan *et al.*, 2012). Ces grands ADNmt présentent aussi un taux de mutation exceptionnellement élevé, dépassant même celui des ADNmt animaux, alors que les plus petits génomes des Silene présentent un taux de mutation assez bas, comme attendu chez les plantes (Sloan *et al.*, 2009; Sloan *et al.*, 2012). Les plus grands ADNmt sont constitués par une multitude de chromosomes indépendants, jusqu'à plus de 50, qui peuvent ne contenir aucun gène identifiable. Comme il n'y a pas de sélection fonctionnelle, ces sous-génomes peuvent être acquis et perdus très rapidement dans la même espèce (Wu *et al.*, 2015).

En absence de stress, les événements de recombinaison avec les régions répétées de taille intermédiaires sont peu fréquents et semblent dépendre de la longueur de la séquence homologue, ainsi que de leur position dans le génome. Plus la taille de la séquence répétée est courte, moins la recombinaison est fréquente et réversible (Lovett *et al.*, 2002). Les versions alternatives générées grâce à de courtes homologues sont tout de même détectées en quantité subliminale, généralement moins d'une par cellule (Small *et al.*, 1989). Ces sublimons participent à l'hétéroplasmie du génome mitochondrial. Néanmoins, ces versions alternatives faiblement présentes peuvent devenir la version dominante de l'ADNmt, lors d'un processus appelé *sub-stoichiometric shift* (SSS) ou expansion clonale (Small *et al.*, 1989). Le SSS peut subvenir rapidement, parfois d'une génération à l'autre seulement, et est responsable de l'évolution rapide de la structure de l'ADNmt des plantes. Les mécanismes responsables du SSS ne sont pas encore élucidés et sont largement incompris. Les mutants de facteurs impliqués dans la HR et sa régulation montrent une proportion supérieure de SSS pouvant être liée à l'augmentation de la recombinaison ectopique (Zaegel *et al.*, 2006; Wallet *et al.*, 2015). Le SSS pourrait provenir de la réplication ou de la ségrégation préférentielle du sublimon, dont les mécanismes pourraient être liés à la recombinaison homologue. La réplication et la ségrégation de l'ADNmt ne sont également pas élucidées chez les plantes et leurs mécanismes restent purement spéculatifs.

Les réarrangements génomiques résultant de recombinaisons ectopiques peuvent être sans conséquence pour la mitochondrie lorsqu'ils affectent uniquement des régions non codantes et non impliquées dans la régulation. Ils peuvent aussi causer l'interruption de



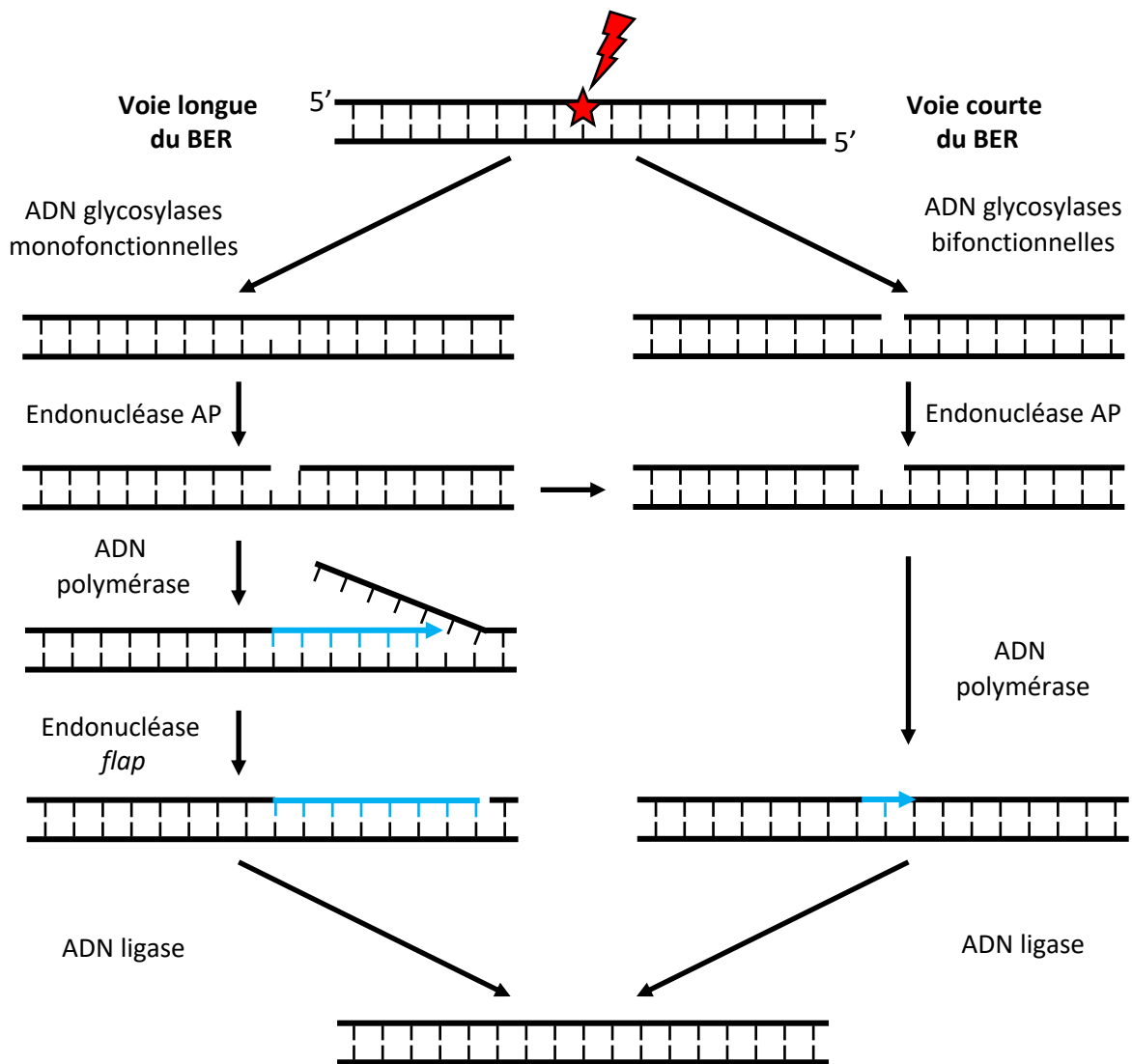
**Figure 17 : Evolution de la structure du génome mitochondrial des plantes.** **A.** La recombinaison entre des séquences répétées peut entraîner la duplication ou la suppression de séquence lors de l'utilisation de mécanismes non fidèles et propices aux réarrangements tels que le BIR. **B.** Le réarrangement des séquences peut se produire très rapidement et modifier grandement la synténie au sein d'une même espèce (ici deux écotypes d'*A. thaliana*). **C.** Les réarrangements peuvent conduire à la formation d'ORF chimériques en liant une région transcrite avec une région non transcrite. **D.** Les ORF chimériques formées par recombinaison peuvent coder des protéines délétères pour la mitochondrie. En affectant les capacités énergétiques de la mitochondrie, les plantes peuvent devenir incapables de produire du pollen viable (CMS). D'après Hanson & Bentolila, 2004

séquences codantes, la création de gènes chimériques, le déplacement d'introns ou de régions régulatrices permettant l'expression de séquences normalement non transcrites (Hanson et Bentolila, 2004). Ces réarrangements peuvent causer des problèmes développementaux ou la létalité de la plante mais peuvent aussi apporter de nouvelles fonctions avantageuses (Kubo et Newton, 2008). Les gènes chimériques sont notamment responsables de la création de lignées CMS recherchées par l'industrie agronomique pour leur incapacité à produire du pollen viable (Budar et Pelletier, 2001; Touzet et Meyer, 2014) (**figure 17**). Ces gènes chimériques permettent l'expression d'une protéine délétère pour la fonction de la mitochondrie, appelée facteur stérilisant, affectant directement la formation des anthères ou la production de pollen, des processus coûteux en énergie pour la cellule (Touzet et Meyer, 2014). Dans la nature, la stérilité mâle cytoplasmique permet de favoriser la reproduction hétérogame et de sauvegarder la dépense énergétique liée à la formation du système reproducteur mâle (Touzet et Meyer, 2014).

De par son nombre de séquences répétées plus restreint, l'ADNcp est nettement moins sujet aux réarrangements que l'ADNmt. Les DSB sont réparées par des recombinaisons intermoléculaires (*gene conversion*) conduisant à une très bonne conservation des séquences chloroplastiques (Khakhlova et Bock, 2006). Le MMEJ et le NHEJ pourraient également participer à la réparation des DSB dans les régions dépourvues de séquences répétées, mais n'entraînent pas d'importants réarrangements ni délétion de séquences (Kwon *et al.*, 2010). Certaines courtes séquences répétées sont néanmoins responsables du glissement de l'ADN polymérase lors de la réplication de l'ADNcp (Massouh *et al.*, 2016). Ces erreurs de réplication semblent être plus fréquemment retrouvées dans l'ADNcp (Massouh *et al.*, 2016). L'ADNcp des mousses, telles que *Physcomitrella patens*, ou des algues, telles que *Chlamydomonas*, présente un nombre plus important de séquences répétées, le rendant plus susceptible aux réarrangements (Higgs *et al.*, 1998; Odahara *et al.*, 2015).

## 5. Les mécanismes de réparation de l'ADN organellaire

En plus des mécanismes de recombinaison (DSBR, SDSA) et de liaison d'extrémités (MMEJ), les organelles possèdent d'autres voies de réparation de leur génome. Ces voies supplémentaires permettent de maintenir l'intégrité de l'ADNmt et de l'ADNcp dans les environnements propices à la formation de ROS que sont la mitochondrie et le chloroplaste (Maxwell *et al.*, 1999; Moller, 2001; Dutilleul *et al.*, 2003; Foyer et Noctor, 2005). Comme pour la recombinaison homologue, les autres voies de réparation de l'ADN dérivent de celles retrouvées dans le noyau et utilisent des facteurs codés par le génome nucléaire puis importés dans l'organelle depuis le cytoplasme. Toutes les voies de réparation nucléaires de l'ADN n'ont cependant pas été caractérisées dans les organelles, et certaines semblent agir différemment dans les organelles et le noyau (Boesch *et al.*, 2011; Gualberto et Newton, 2017).



**Figure 18 : Mécanisme général du *base excision repair*.** Le BER répare les bases modifiées et adduits ne déformant pas l'hélice d'ADN. La base lésée est d'abord retirée par des ADN glycosylases spécifiques du type de dommage et laissent un site abasique. Les ADN glycosylases bifonctionnelles (en haut à droite) clivent également le désoxyribose abasique en 3' du site AP. Une endonucléase AP prend en charge le site AP pour créer une extrémité 5'-phosphate utilisable par une ADN polymérase. Dans la voie longue du BER (à gauche), l'ADN polymérase synthétise plusieurs nucléotides en déplaçant les anciens qui devront être clivés par une endonucléase AP. Dans la voie courte du BER, un seul nucléotide est synthétisé.

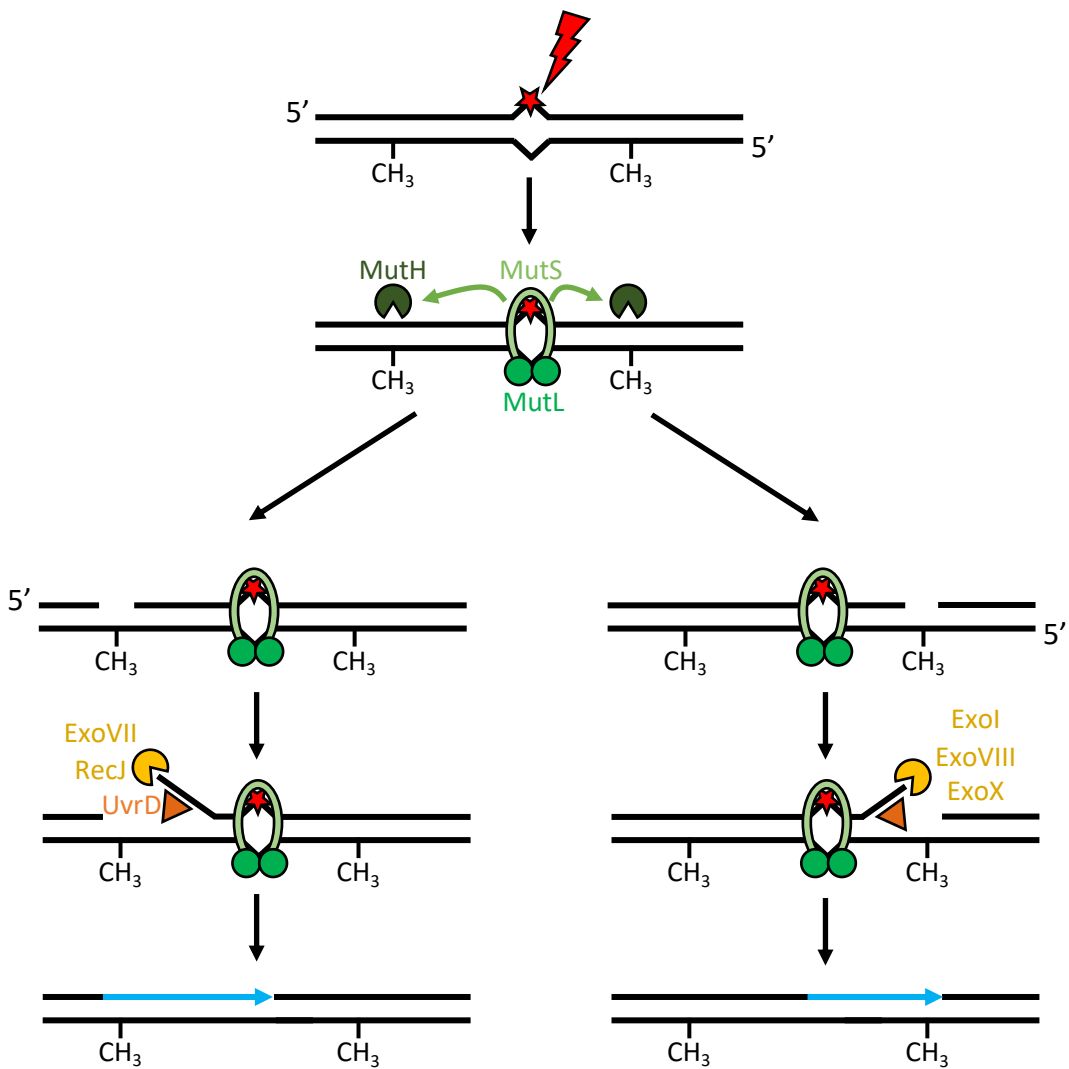
## 5.1. La réparation par excision de base

La réparation par excision de base (BER : *base excision repair*) permet la suppression des bases modifiées de l'ADN pouvant être incorrectement reconnues lors de la réplication et conduire à des modifications de séquence. Le BER utilise des ADN glycosylases qui reconnaissent et clivent spécifiquement certaines lésions de l'ADN, engendrant des sites apurinique/apyrimidique (AP) (**figure 18**). Le site abasique est ensuite pris en charge par une AP endonucléase qui clive le désoxyribose abasique, formant une extrémité 5'--désoxyribose phosphate ou supprimant complètement le nucléotide, si la glycosylase utilisée est monofonctionnelle ou bi-fonctionnelle, respectivement (Dodson et Lloyd, 2002). Enfin, une ADN polymérase synthétise le nucléotide manquant dans la voie courte du BER (short-patch BER), ou plusieurs nucléotides en déplaçant la séquence en aval de la lésion (qui sera ensuite clivée par une endonucléase flap) dans la voie longue du BER (long-patch BER). La liaison phosphodiester entre le(s) nucléotide(s) ajouté(s) et la séquence d'origine est rétablie par une ADN ligase.

Le BER est considéré comme étant la voie de réparation principale des mitochondries animales et de levure (Boesch *et al.*, 2011). Chez les organelles de plantes, une uracil-DNA glycosylase (UDG) impliquée dans la voie courte de BER a été détectée (Boesch *et al.*, 2009). Cette enzyme permet la suppression des uraciles incorporés par erreur dans le génome après déamination d'une cytosine. Récemment, Ferrando et al. ont détecté l'activité des glycosylases, Neil1/2 et OGG1 dans les mitochondries de tubercules de pomme de terre (Ferrando *et al.*, 2018). Neil1 et Neil 2 sont deux glycosylases utilisées dans la suppression des pyrimidines oxydées, de l'hydroxyuracile et des lésions induites par l'urée dans l'ADN, tandis que OGG1 permet d'exciser les produits d'oxydation de la guanine (Ferrando *et al.*, 2018). Une autre glycosylase impliquée dans le remplacement des 8-oxoguanines, FPG, pourrait également être utilisée dans les mitochondries de plantes (Macovei *et al.*, 2011). S'il est certain que la voie courte du BER est présente dans les organelles de plantes, la voie longue n'a pas encore été caractérisée (Boesch *et al.*, 2009; Ferrando *et al.*, 2018). Chez le chloroplaste d'*Arabidopsis thaliana*, les glycosylases NTH1 et NTH2, et l'AP endonucléase ARP ont été identifiées (Gutman et Niyogi, 2009). L'ADN polymérase et l'ADN ligase impliquée dans le BER restent néanmoins à identifier. Il pourrait s'agir de POL1B et LIG1, toutes deux adressées aux mitochondries et aux chloroplastes (Gualberto et Newton, 2017).

## 5.2. La réparation des mésappariements

La réparation des mésappariements (MMR : *mismatch repair*) permet la reconnaissance de nucléotides incorrects ou modifiés incorporés dans l'ADN lors de la réplication (**figure 19**). Chez les procaryotes, la protéine MutS reconnaît spécifiquement la base mutée sur le brin d'ADN néoformé et s'y lie. La protéine MutL stabilise ensuite le complexe MutS-ADN pour



**Figure 19 : Mécanisme du *mismatch repair* procaryotique.** Le MMR permet de retirer les bases mésappariées lors de la réplication. Le brin néosynthétisé sur lequel se trouve le mauvais nucléotide est discriminé par la méthylation de l'ADN. La protéine MutS reconnaît la lésion et recrute la protéine MutL pour stabiliser l'ensemble du complexe. MutH est ensuite recrutée et clive l'ADN au niveau d'un site de méthylation GATC d'un côté ou de l'autre de la lésion. L'hélicase UvrD et des exoribonucléases 5'-3' ou 3'-5' sont recrutées pour digérer l'ADN jusqu'à la lésion. Une ADN polymérase et une ADN ligase permettent de synthétiser et relier l'ADN manquant.

permettre l'action de l'endonucléase MutH, qui incise le brin d'ADN en amont ou en aval de la lésion. Des exonucléases sont ensuite recrutées pour digérer l'ADN jusqu'au mésappariement. L'ADN manquant est ensuite synthétisé par une ADN polymérase (Morita *et al.*, 2010).

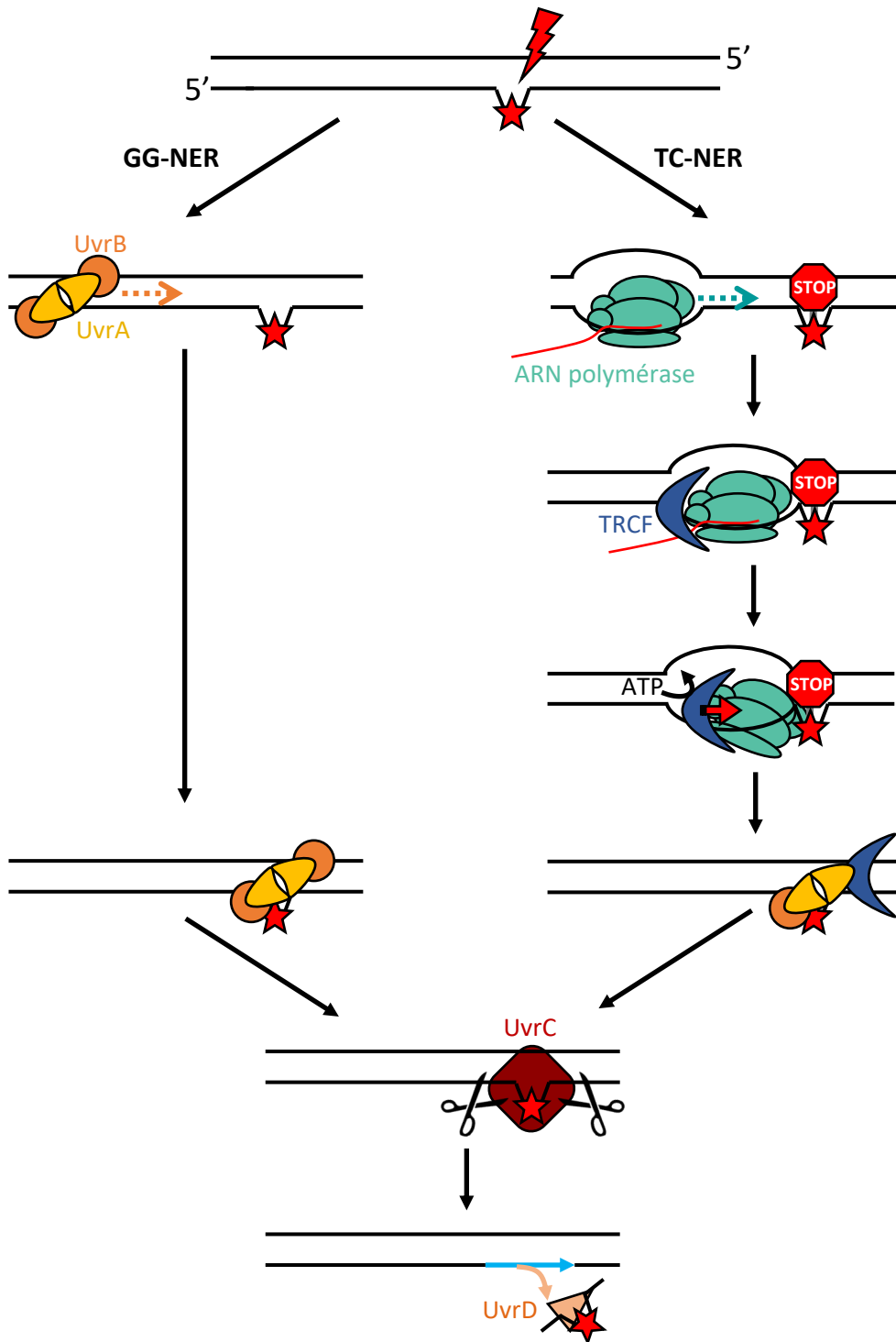
Chez les plantes, un homologue de MutS, MSH1, est adressé aux mitochondries et aux chloroplastes. Aucune autre protéine de la voie du MMR n'a été identifiée dans les organelles, mais MSH1 possède un domaine C-terminal GIY-YIG *homing endonuclease* qui pourrait lui permettre d'inciser l'ADN comme MutH (Abdelnoor *et al.*, 2006). La mutation de MSH1 entraîne des réarrangements de l'ADNmt et de l'ADNcp, suggérant un rôle de la protéine dans la recombinaison (Shedge *et al.*, 2007; Xu *et al.*, 2011). MSH1 pourrait déstabiliser les intermédiaires de recombinaison lorsque l'homologie de séquence n'est pas suffisamment importante (Abdelnoor *et al.*, 2006; Xu *et al.*, 2011). Les cellules déficientes dans le MMR sont hyper-recombinogènes (Elliott et Jasin, 2001). Selon Christensen, MSH1 pourrait permettre la réparation des mésappariements des bases en induisant des DSB en clivant les deux brins d'ADN au niveau de la base mutée (Gualberto et Newton, 2017; Christensen, 2018). Ces DSB seraient ensuite pris en charge par la machinerie de recombinaison des organelles, permettant ainsi la correction fidèle de la lésion, ce qui expliquerait le faible taux de mutation retrouvé dans les organelles.

### 5.3. La réversion directe des dommages par l'activité photolyase

Certaines lésions causées par les UV entraînent le pontage de deux pyrimidines au sein d'un même brin d'ADN et causent une déformation importante de l'hélice d'ADN (Sancar et Sancar, 1988). Cette déformation bloque les enzymes se déplaçant le long de l'ADN, notamment des ARN polymérases, et peut causer des DSB lors de la collision de la machinerie de réplication avec le complexe d'élongation de la transcription bloqué. La réversion directe des dommages, ou photo-réactivation, permet de rétablir la structure de l'hélice d'ADN sans exciser les nucléotides impliqués dans la lésion (Morita *et al.*, 2010). Elle repose sur l'utilisation de la lumière bleue par les photolyases qui suppriment le pontage entre les deux pyrimidines par un transfert d'électrons (Brettel et Byrdin, 2010).

Les UV peuvent également induire la formation de dimères de pyrimidines dans les génomes des organelles (Chen *et al.*, 1996). L'existence de mécanismes de réversion directe des dommages est toujours controversée dans les organelles de plantes, bien qu'ils soient présents dans la mitochondrie de *S. cerevisiae* et dans le chloroplaste de *Chlamydomonas reinhardtii* (Small, 1987; Yasui *et al.*, 1992). Plusieurs études réalisées avec différentes plantes n'ont pas montré d'activité de photo-réactivation (Chen *et al.*, 1996; Hada *et al.*, 2000). Pourtant, Takahashi *et al.* ont montré qu'une photolyase nucléaire était également adressée à la mitochondrie et au chloroplaste chez *Oryza sativa* (riz), tandis qu'un autre groupe a identifié une photolyase dans les organelles d'*Arabidopsis thaliana* (Kleine *et al.*, 2003; Takahashi *et al.*,





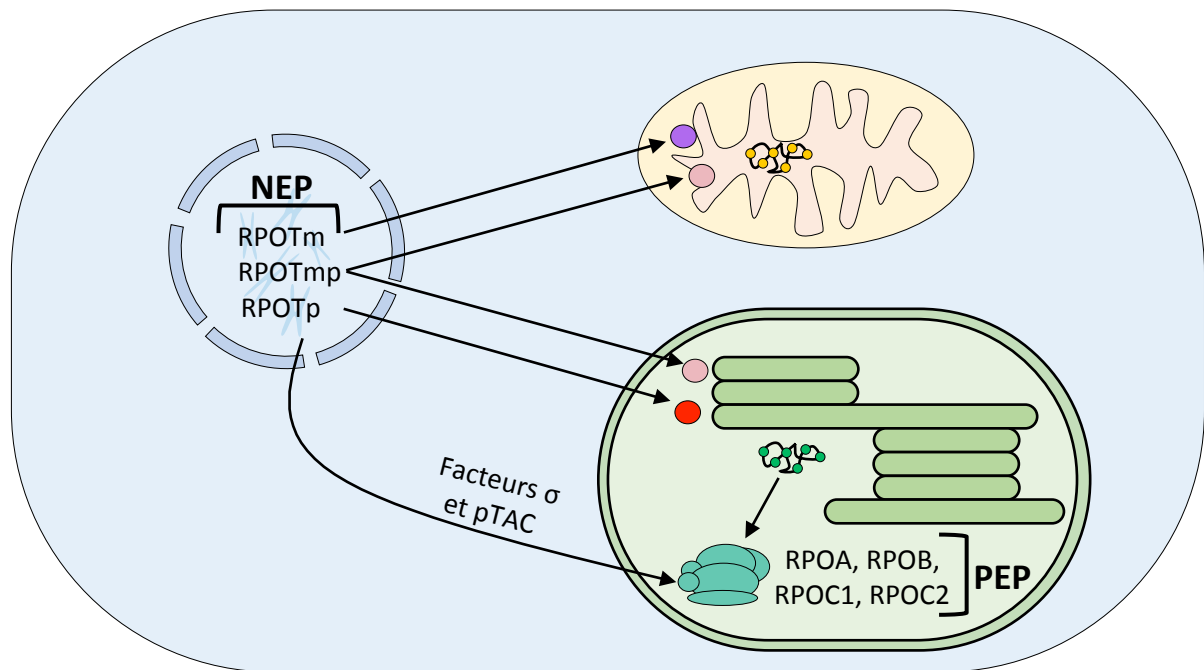
**Figure 20 : Mécanisme du *nucleotide excision repair* procaryotique.** Le NER permet la réparation des lésion et adduits distordant la double hélice d'ADN. Le NER est divisé en deux voies : le *global genomic-NER* (GG-NER, à gauche) réparant aléatoirement le génome et le *transcription coupled-NER* (TC-NER, à droite) réparant préférentiellement les brins transcrits. Pour le GG-NER, les hétérotétramères UvrA<sub>2</sub>UvrB<sub>2</sub> scannent aléatoirement l'ADN à la recherche de lésions. Le TC-NER repère les ARN polymérases bloquées par la lésion distordant l'ADN grâce à la protéine TRCF qui se déplace sur l'ADN. Lorsque TRCF rencontre une ARN polymérase bloquée, il subit un changement de conformation et pousse la polymérase à travers la lésion, déstabilisant tout le complexe d'élongation. TRCF recrute ensuite un hétérotrimère UvrA<sub>2</sub>UvrB<sub>1</sub> pour réparer la lésion. Une fois la lésion repérée par UvrA et UvrB, l'endonucléase UvrC est recrutée pour cliver de part et d'autre de la lésion. L'ADN clivé est retiré par l'hélicase UvrD et l'ADN manquant est synthétisé par une ADN polymérase.

2011). L'activité de réparation de l'ADN des deux protéines n'a cependant pas été testée, rendant impossible toute conclusion quant à la présence de réversion directe dans les organelles de plante.

#### 5.4. La réparation par excision de nucléotide

Les lésions déformant l'hélice d'ADN peuvent aussi être réparées par des mécanismes d'excision. La réparation par excision de nucléotide (NER : *nucleotide excision repair*) permet la reconnaissance et la suppression des lésions et adduits contraignant l'ADN et empêchant la circulation sur l'ADN de machineries essentielles pour le fonctionnement cellulaire (Morita *et al.*, 2010). Le NER est divisé en deux voies : le GG-NER (*global genomic NER*) permettant la réparation non orientée de tout le génome, et le TC-NER (*transcription-coupled NER*) utilisant les ARN polymérases pour réparer préférentiellement les régions transcrites. Seule l'étape de reconnaissance des dommages diffère entre les deux voies, qui utilisent ensuite les mêmes facteurs. Chez les procaryotes, des hétéro-tétramères UvrA<sub>2</sub>-UvrB<sub>2</sub> scannent aléatoirement le génome en permanence à la recherche de lésions distordant l'hélice d'ADN lors du GG-NER (**figure 20**). Lorsqu'une lésion est détectée, UvrA positionne UvrB sur la lésion et quitte le complexe. UvrB recrute alors l'endonucléase UvrC qui clive un brin d'ADN de part et d'autre du dommage. La région d'ADN clivée est retirée par l'hélicase UvrD, puis la région manquante est synthétisée et liguée par l'ADN polymérase I et une ligase. Pour le TC-NER, le mécanisme tire parti de l'incapacité de l'ARN polymérase (RNAP) à traverser les lésions distordant l'hélice de l'ADN et les adduits trop volumineux. Les ARN polymérases s'y retrouvent bloquées et deviennent un signal de recrutement de la protéine TRCF (*transcription repair coupling factor*) (Westblade *et al.*, 2010; Selby, 2017). L'interaction de TRCF avec une ARN polymérase induit un changement de conformation de TRCF, alors capable de déstabiliser tout le complexe d'élongation en tirant l'ADN à travers l'ARN polymérase. Une fois l'ARN polymérase dégagée et la lésion découverte, TRCF recrute un hétéro-trimère UvrA<sub>2</sub>UvrB<sub>1</sub> pour supprimer le dommage, de la même façon que dans la voie du GG-NER (Fan *et al.*, 2016).

La voie du NER semble absente des organelles. Aucun homologue des facteurs UvrA, UvrB et UvrC n'est codé par les génomes des plantes à fleur, ni aucun des facteurs du NER intervenant sur le génome nucléaire n'est prédit comme adressé à la mitochondrie ou au chloroplaste. Un homologue du facteur TRCF bactérien est néanmoins présent dans le génome d'*A. thaliana*, et son adressage est prédit dans le chloroplaste (Gualberto et Newton, 2017). Il pourrait cependant être impliqué dans le BER comme cela a été proposé pour certains facteurs du NER nucléaire adressés dans les mitochondries animales (Stevnsner *et al.*, 2002; Kamenisch *et al.*, 2010; Liu *et al.*, 2015). La protéine TRCF étant un des sujets principaux de ma thèse, une introduction spécifique plus complète est donnée plus loin.



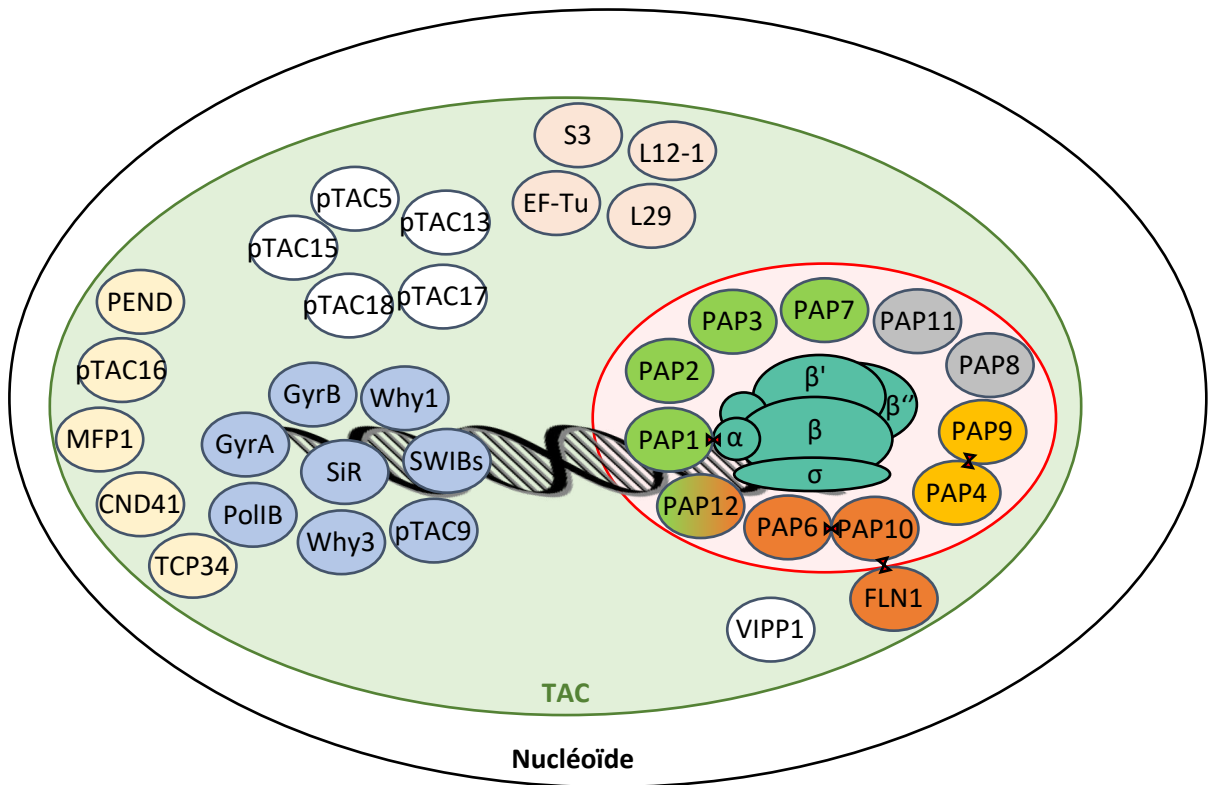
**Figure 21 : Répartition des ARN polymérase dans les organelles des dicotylédones.** Les organelles de plantes possèdent plusieurs ARN polymérase impliquées dans la transcription de leur génome. Le génome nucléaire code pour trois ARN polymérase monomériques de type phagique (NEP) adressées à la mitochondrie (RPOtm et RPOtMp), ou au chloroplaste (RPOtp et RPOtMp). Dans le chloroplaste, les NEP permettent l'expression de la PEP, une ARN polymérase multimérique (sous-unités RPOA, RPOB, RPOC1 et RPOC2) de type bactérienne. Cette PEP nécessite également des facteurs  $\sigma$  pour reconnaître ses promoteurs ainsi que de nombreux facteurs régulant ses fonctions (pTAC), tous codés par le génome nucléaire.

## 6. Le contrôle de l'expression des gènes dans les organelles

L'expression des gènes codés par le génome des organelles est entièrement contrôlée par le noyau, dans un système alliant des caractéristiques procaryotiques, eucaryotiques et phagiques. Le génome nucléaire code la ou les ARN polymérases mitochondriales et chloroplastiques, appelées NEP (nuclear encoded polymerase). Les animaux et les levures possèdent une seule NEP, tandis que les plantes peuvent en avoir jusqu'à trois (**figure 21**) (Hedtke *et al.*, 1997, 2000). Les dicotylédones possèdent une NEP importée à la mitochondrie (RPOTm), une autre importée au chloroplaste (RPOTp), et une dernière doublement adressée aux deux organelles (RPOTmp). Les monocotylédones ne possèdent pas RPOTmp, et les algues ne possèdent ni RPOTp ni RPOTmp (Liere *et al.*, 2011; Borner *et al.*, 2015). Enfin, l'ADNcp code également une ARN polymérase appelée PEP (plastidial encoded polymerase), dont l'expression et l'activité sont dépendantes de facteurs nucléaires (**figure 21**) (Steiner *et al.*, 2011; Borner *et al.*, 2015).

### 6.1. Les ARN polymérases organellaires

Les NEP sont des ARN polymérases monomériques de type phagique, s'apparentant aux ARN polymérases des phages T3, T7 et SP6, d'où provient leur nom RPOT (RNA polymerase of the T3/T7 type) (Hess et Borner, 1999). La NEP des vertébrés et des levures nécessite des facteurs de transcription, mtTFA et mtTFB, pour reconnaître et accéder aux régions promotrices (Fisher et Clayton, 1988; Falkenberg *et al.*, 2002). Chez les plantes, il n'y a aucun homologue de mtTFA. La protéine Dim1B similaire au facteur mtTFB, appartenant à la famille des ARNr diméthylases, est importée dans la mitochondrie des plantes. Cependant elle semble seulement impliquée dans la méthylation de l'ARNr 18S, et n'affecte pas la transcription (Richter *et al.*, 2010). Les NEP des plantes ne semblent pas avoir besoin de facteur de transcription particulier pour débiter la transcription, et il a été montré que ces NEP reconnaissent intrinsèquement leur promoteur *in vitro* (Kühn *et al.*, 2007). L'implication de certaines protéines liant l'ADN, constituantes des nucléoïdes des organelles, est cependant suggérée pour ouvrir l'ADN et faciliter la mise en place de la bulle de transcription (Gualberto et Kühn, 2014; Bohne *et al.*, 2016). Les régions promotrices des NEP de plantes sont assez variables, ne présentant pas le consensus retrouvé dans l'ADNmt des levures. Les promoteurs de plantes ont été divisés en trois catégories en fonction de leur séquence : les promoteurs de classe Ia présentent une séquence 5'-YRTA-3' (avec Y une base pyrimidique et R une base purique), les promoteurs de classe Ib ont une boîte 5'-GAA-3' de 18 à 20 nucléotides en amont de la séquence 5'-YRTA-3', et les promoteurs de classe II ne possèdent pas le consensus 5'-YRTA-3' (Borner *et al.*, 2015). Les NEP RPOTm et RPOTp reconnaissent assez bien la majorité des promoteurs et sont considérées comme les NEP principales dans la mitochondrie et le chloroplaste, tandis que RPOTmp reconnaît assez peu de promoteurs *in vitro* (Kühn *et al.*, 2007). Cette reconnaissance



- |                                                               |                                                                      |                                                                  |                                                       |
|---------------------------------------------------------------|----------------------------------------------------------------------|------------------------------------------------------------------|-------------------------------------------------------|
| <span style="color: green;">●</span> Métabolisme De l'ADN/ARN | <span style="color: orange;">●</span> Régulation de la transcription | <span style="color: yellow;">●</span> Protection contre les ROS  | <span style="color: grey;">●</span> Fonction inconnue |
| <span style="color: blue;">●</span> Réplication               | <span style="color: pink;">●</span> Interaction avec le ribosome     | <span style="color: lightyellow;">●</span> Ancrage aux membranes |                                                       |

**Figure 22 : Représentation schématique du complexe de transcription liée à la PEP.** La PEP-B n'est constituée que des quatre sous-unités RPOA, RPOB, RPOC1 et RPOC2 représentées en turquoise. Dans les chloroplastes, la forme active de la PEP (PEP-A, entourée en rouge) nécessite 12 PAP impliquées dans diverses fonctions. Le complexe de transcription (TAC, en vert) implique également des protéines liées à la réplication et à la traduction de l'ADNcp. L'ensemble du complexe est un composant des nucléoïdes.  
D'après Pfalz et Pfanschmidt, 2013 et Kindren et Åsa, 2015.

peut être augmentée en modifiant la séquence de RPO<sub>Tmp</sub> responsable de la reconnaissance des nucléotides, formant une boucle très conservée au sein des polymérases phagiques (Bohne *et al.*, 2016). La raison de l'existence d'une polymérase adressée aux deux organelles, mais moins efficace que les NEP originales est toujours débattue. Kühn *et al.* proposent que RPO<sub>Tmp</sub> pourrait ajuster le niveau des transcrits en augmentant spécifiquement la transcription de certains gènes comme *nad6* et *cox1* (Kühn *et al.*, 2009). Quel que soit son rôle, RPO<sub>Tmp</sub> est tout de même nécessaire à la cellule car sa mutation entraîne un défaut de transcription limitant la phosphorylation oxydative et conduisant à un important retard de croissance de la plante (Kühn *et al.*, 2009). Le mutant *rpoTm* n'est pas viable, indiquant que RPO<sub>Tmp</sub> n'est pas en mesure de compléter son absence (Kühn *et al.*, 2009). Dans le chloroplaste, RPO<sub>Tp</sub> et RPO<sub>Tmp</sub> sont considérés comme étant partiellement redondants. Les mutants *rpoTp* et *rpoTmp* individuels sont affectés dans l'expression des gènes et le développement des feuilles, tandis que le double mutant *rpoTp rpoTmp* n'est pas viable et cela dès les premiers stades du développement (Hricova *et al.*, 2006).

La PEP est une ARN polymérase multimérique de type bactérienne. Elle est composée des protéines RPOA, RPOB, RPOC1 et RPOC2, représentant les sous-unités  $\alpha$ ,  $\beta$ ,  $\beta'$  et  $\beta''$ , respectivement (Sato *et al.*, 1999). Les sous-unités  $\beta$ ,  $\beta'$  de la PEP peuvent se substituer aux mêmes unités de la RNAP bactérienne (Severinov *et al.*, 1996). La PEP est considérée comme l'ARN polymérase principale des chloroplastes, transcrivant les gènes codant pour les protéines impliquées dans les complexes photosynthétiques. Les promoteurs de la PEP sont de type procaryotique, avec les boîtes -35 et -10 présentant les séquences consensus TTGACA et TATAAT, respectivement (Yagi et Shiina, 2014). Pour reconnaître ses promoteurs, la PEP nécessite des facteurs sigma ( $\sigma$ ) codés par le génome nucléaire. *Arabidopsis thaliana* possède six facteurs sigma chloroplastiques (SIG1-SIG6) de type  $\sigma^{70}$ , qui ont des rôles partiellement redondants (Liere *et al.*, 2011). Les quatre sous-unités de la PEP forment la PEP-B, principalement retrouvée dans les étioplastes et les chloroplastes en développement. Dans les chloroplastes matures, la PEP est associée avec une douzaine de facteurs, les PAP (*PEP associated proteins*), nécessaires à son fonctionnement (Steiner *et al.*, 2011). La PEP et ses PAP sont appelés PEP-A. En plus des PAP, d'autres facteurs sont également requis pour la transcription par la PEP-A, mais n'interagissent pas directement avec les sous-unités de la PEP. Au total, plus de quarante facteurs sont impliqués avec la PEP et forment le TAC (*transcriptionally active chromosome*) (**figure 22**). Ces facteurs sont appelés pTAC et sont des composants des nucléoïdes chloroplastiques (Pfalz *et al.*, 2006; Pfalz et Pfannschmidt, 2013).

Les douze PAP peuvent être classées en plusieurs catégories en fonction de leurs propriétés, bien que leur rôle individuel soit souvent encore assez flou (Yu *et al.*, 2014; Kindgren et Strand, 2015). Plusieurs protéines possèdent un domaine de liaison avec les acides nucléiques (PAP1, 2, 3, 5 et 7) et pourraient être impliquées dans la reconnaissance des

promoteurs, l'initiation, l'élongation ou la terminaison de la transcription (Pfalz *et al.*, 2006; Yagi *et al.*, 2012; Wang *et al.*, 2016; Chang *et al.*, 2017b; Grübler *et al.*, 2017). Un second groupe de protéines serait plutôt impliqué dans la régulation de l'activité de la PEP (PAP5, 6, 7, 10), potentiellement en fonction de l'état des complexes photosynthétiques ou de la lumière perçue (Grübler *et al.*, 2017; Nevarez *et al.*, 2017; He *et al.*, 2018). D'autres facteurs pourraient aussi jouer un rôle dans la protection de l'ARN polymérase contre les ROS (PAP 4 et 9) (Myouga *et al.*, 2008). Enfin, certaines protéines identifiées ont toujours un rôle complètement inconnu (PAP 8 et 11). La mutation des sous-unités de la PEP entraîne un important défaut de transcription des gènes codant les protéines des complexes photosynthétiques, conduisant à un phénotype albinos (De Santis-Maciossek *et al.*, 1999). La mutation d'une PAP conduit à un phénotype développemental similaire, avec des feuilles vertes pâles, ivoires, voire albinos, reflétant l'impossibilité de la PEP d'exprimer correctement les gènes de l'ADNcp en absence de ses facteurs associés (Pfalz et Pfannschmidt, 2013).

## 6.2. La transcription organellaire

Dans la mitochondrie, la transcription est effectuée uniquement par les NEP. Chez les plantes, l'initiation de la transcription ne semble pas particulièrement régulée (Holec *et al.*, 2006). Les gènes mitochondriaux possèdent souvent plusieurs promoteurs peu spécifiques en amont (Kühn *et al.*, 2005). Il est considéré qu'une grande partie du génome mitochondrial est transcrite, y compris les régions non codantes. Le contrôle de la transcription mitochondriale serait davantage post-transcriptionnel, via les mécanismes de maturation des transcrits et la modulation de leur stabilité (Holec *et al.*, 2006). La transcription des régions non codantes pourrait néanmoins avoir comme conséquence l'accumulation d'ARN non codants jouant un rôle dans la régulation de la transcription, comme c'est le cas dans la mitochondrie des mammifères (Dietrich *et al.*, 2015).

Dans le chloroplaste, l'expression des gènes est partagée entre les deux NEP (RPOTp et RPOTmp) et la PEP. Les gènes peuvent posséder un promoteur de la PEP (classe I), un promoteur de la NEP (classe III) ou un promoteur PEP et NEP (classe II) (Borner *et al.*, 2015). La majorité des gènes sont transcrits par les deux types d'ARN polymérases, bien que l'une ou l'autre soit privilégiée en fonction du stade de développement du chloroplaste. En général, la NEP est plutôt active lors des stades précoces du développement chloroplastique pour permettre l'expression des gènes de ménage, dont l'opéron *rpoB* codant les sous-unités de la PEP (*rpoB*, *rpoC1*, *rpoC2*) (Hajdukiewicz *et al.*, 1997). L'opéron *rpoB* est d'ailleurs le seul à posséder uniquement un promoteur NEP (Borner *et al.*, 2015). A l'inverse, la PEP devient l'ARN polymérase principale du chloroplaste mature, transcrivant principalement les gènes codant pour les sous-unités des complexes photosynthétiques (Zhelyazkova *et al.*, 2012a). La PEP est néanmoins active tout au long du développement du chloroplaste et transcrit également des

gènes de ménage comme des protéines ribosomales. La PEP est ainsi entièrement sous le contrôle du génome nucléaire : son expression est dépendante des NEP et son activité nécessite les PAP, pTAC et facteurs  $\sigma$  codés par le noyau. Ce contrôle permet la coordination entre les génomes nucléaire et chloroplastique (Gray *et al.*, 2003). L'activation totale de la PEP (le passage de la PEP-B à la PEP-A) est dépendante de la lumière et marque le passage de la skotomorphogénèse à la photomorphogénèse (Liebers *et al.*, 2018). Tout comme pour le génome mitochondrial, l'ensemble de l'ADNcp est transcrit, puis les transcrits sont régulés lors des étapes post-transcriptionnelles (Shi *et al.*, 2016).

La terminaison de la transcription organellaire est encore peu comprise chez les plantes. Aucun homologue de facteur de terminaison bactérien n'a été identifié, et il est considéré que la maturation détermine les extrémités des transcrits (Zhelyazkova *et al.*, 2012a). Les séquences terminatrices habituelles pour les ARN polymérases phagiques n'ont pas été retrouvées dans le génome des organelles. Certains ARN forment néanmoins des tige-boucles à leur extrémité 3', un signal de terminaison de la transcription chez les bactéries. Dans les organelles, ces structures sont plutôt utilisées pour la protection des ARN contre les nucléases (Gagliardi *et al.*, 2004). Récemment, plusieurs protéines jouant un rôle dans la terminaison de la transcription organellaire des plantes ont été identifiées (Ji *et al.*, 2019). RHON1 est un facteur chloroplastique possédant un domaine de liaison à l'ARN similaire aux protéines bactériennes Rho et capable de terminer la transcription du gène *rbcl* (Chi *et al.*, 2014). L'absence du facteur RHON1 conduit à la diminution de la transcription du gène *accD* en aval de *rbcl*, certainement à cause d'un mécanisme d'interférence transcriptionnelle entre la PEP transcrivant *rbcl* et la NEP transcrivant *accD* (Chi *et al.*, 2014). Chez les mammifères, la terminaison de la transcription mitochondriale implique quatre protéines mTERF présentant le motif MTERF (Kruse *et al.*, 1989). De nombreuses protéines possèdent ce motif chez les plantes, dont MOC1 et mTERF6, adressées à la mitochondrie ou aux deux organelles respectivement et dont le rôle dans la terminaison de la transcription a été démontré (Wobbe et Nixon, 2013; Romani *et al.*, 2015).



# Objectifs de ma thèse

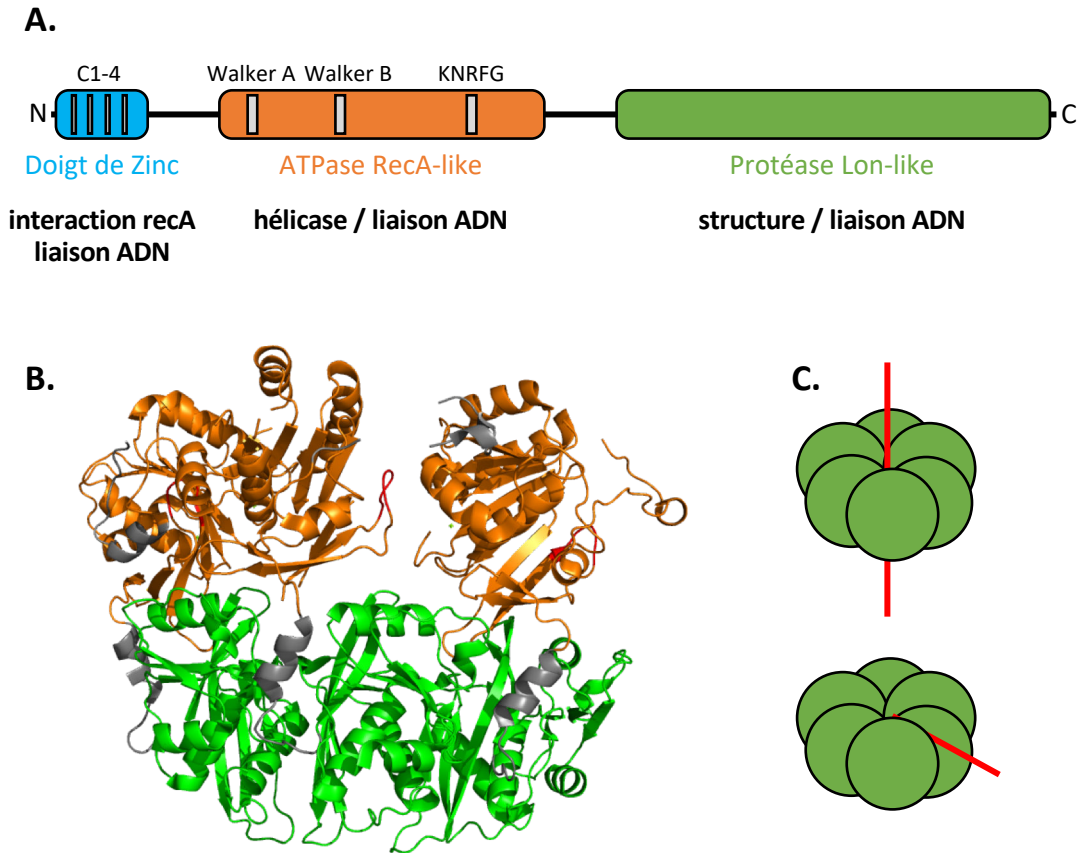
Avant mon arrivée au laboratoire, une étude visant à caractériser des hélicases à ADN impliquées dans la recombinaison des génomes des organelles avait été initiée. Cette étude avait mené à la caractérisation de l'hélicase RECG1, homologue de la protéine de migration de branche RecG bactérienne, responsable de la maturation des intermédiaires de recombinaison et de la suppression de la recombinaison ectopique (Wallet *et al.*, 2015). Alors que la recombinaison organellaire est un processus indispensable au développement des plantes, la mutation *recG1* n'affecte que très peu leur développement malgré l'absence de la seconde voie de migration de branche RuvABC chez les plantes (Miller-Messmer *et al.*, 2012; Wallet *et al.*, 2015). Ces résultats suggèrent donc la présence d'une voie de maturation des intermédiaires de recombinaison alternative et indépendante de RECG1.

Lors de ma thèse, je me suis particulièrement intéressé à deux candidats, pouvant être impliqués dans des fonctions similaires à RECG1 : RADA et TRCF.

Chez les procaryotes, RadA est la troisième voie de migration des branches (avec RuvAB et RecG) permettant la maturation des intermédiaires de recombinaison et la formation d'une HJ (Cooper et Lovett, 2016; Marie *et al.*, 2017). Ces trois voies sont partiellement redondantes et l'absence de l'une d'elles peut être complétée par les deux autres sans affecter la viabilité des bactéries (Cooper *et al.*, 2015). Chez les plantes, une protéine homologue de RadA bactérienne, RADA, pourrait être adressée aux organelles et permettrait ainsi la prise en charge des intermédiaires de recombinaison en absence de RECG1, expliquant l'absence de phénotype développemental observé pour le mutant *recG1*. Pour appréhender le rôle de la protéine RADA, j'ai complété le mutant bactérien *radA* avec RADA d'*Arabidopsis* et testé son activité *in vitro*. Puis, pour comprendre l'importance de RADA chez les plantes, j'ai ensuite caractérisé le mutant *radA*, tant au niveau de son développement que de la structure des génomes de ses organelles.

TRCF, d'abord nommé RECG2, présente le motif hélicase TRG retrouvé chez toutes les protéines de la famille de RecG, présumant d'une activité en lien avec RECG1. Néanmoins, sa séquence générale est plus proche du facteur procaryotique Mfd, possédant lui aussi le motif TRG, impliqué dans le couplage de la transcription avec le NER permettant la réparation préférentielle des régions transcrites (Selby, 2017). Le NER étant décrit comme absent dans les organelles des plantes, je me suis intéressé au rôle de la protéine TRCF d'*Arabidopsis*, dont l'adressage est justement prédit dans les organelles. Pour cela, j'ai voulu identifier les partenaires d'interaction de TRCF par une approche de co-immunoprécipitation afin de voir dans quelles voies la protéine pourrait être impliquée. J'ai également caractérisé le mutant *trcf* afin de comprendre l'importance de la protéine chez les plantes.

# **CHAPITRE 1 : ÉTUDE DE RADA**



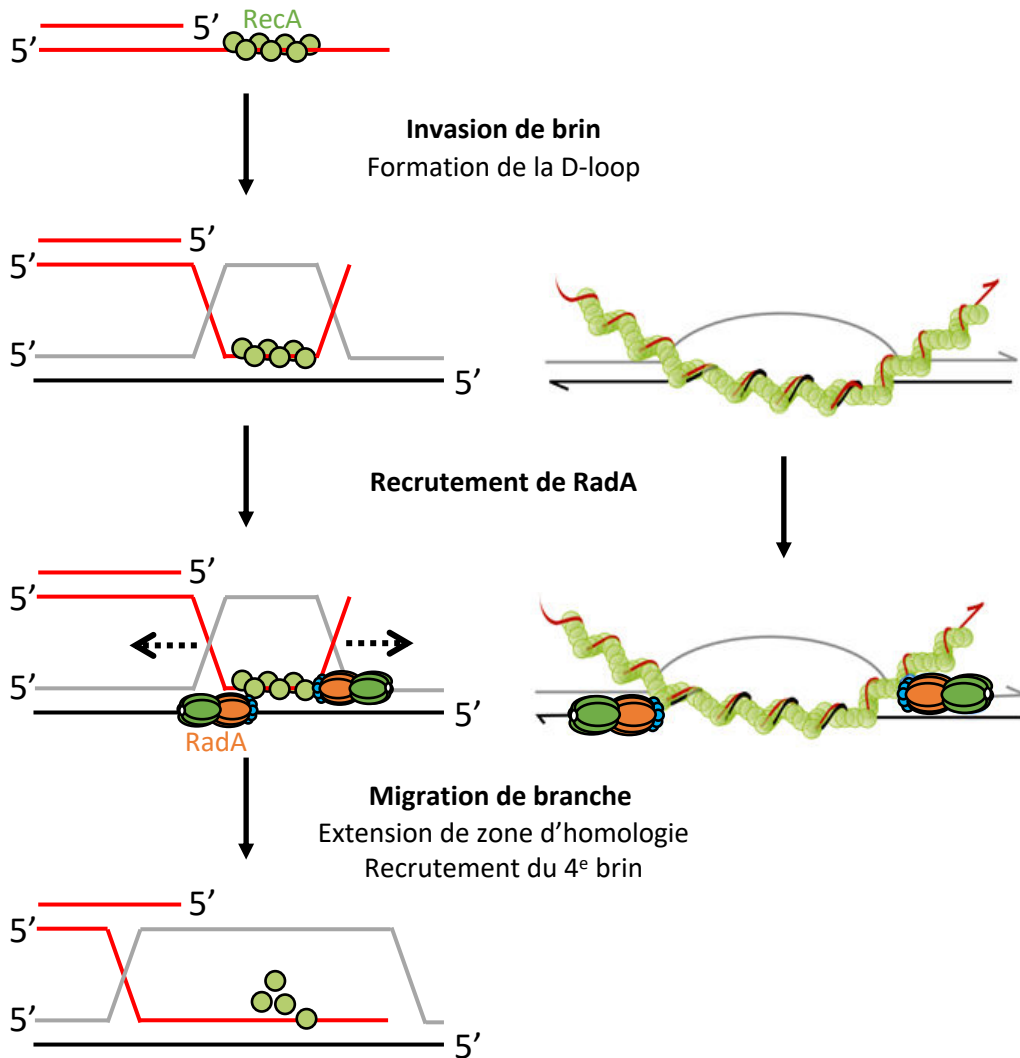
**Figure 23 : Structure de RadA.** **A.** Représentation schématique de la protéine RadA. Les trois domaines fonctionnels sont représentés par les boîtes colorées. La fonction de chacun des domaines est indiquée sous la boîte correspondante. Dans le motif doigt de Zinc, les quatre barres grises représentent les quatre cystéines (C1-4). Dans le domaine ATPasique, les barres grises représentent les deux domaines Walker A et B RecA-like et le motif KNRFG caractéristiques des protéines RadA. **B.** Structure cristallographique d'un demi-hexamère de RadA. Les couleurs de chaque domaine sont conservées par rapport à A. Le domaine doigt de Zinc n'est pas représenté. Le domaine Lon-like permet la structuration en hexamère de la protéine. La seconde moitié de l'hexamère est positionnée derrière celle qui est représentée, créant ainsi un canal au centre du complexe. **C.** Liaison à l'ADN des domaines Lon-like. Les domaines lon-like sont représentés sous forme d'hexamère. L'ADN est représenté en rouge. En haut, l'ADN peut se lier au niveau du canal central de l'hexamère. En bas, l'ADN se lie au niveau de la jonction entre deux monomères de RadA. D'après Marie *et al.*, 2017 ; Inoue *et al.*, 2017

## Introduction

RadA a été initialement identifiée chez *E. coli* lors de la recherche de mutants sensibles aux radiations ionisantes (Diver *et al.*, 1982). Dix années plus tard, RadA est décrit une seconde fois pour la sensibilité de son mutant au MMS (méthyl methanesulfonate) sous le nom de sms (*sensitivity for MMS*) (Neuwald *et al.*, 1992). Les gènes *RadA* et *sms* sont finalement confondus et la dénomination RadA sera préférée, bien que celle de RadA/sms soit couramment utilisée (Song et Sargentini, 1996). La recombinaise RecA-like des archées est également nommée RadA, mais n'est pas lié au gène *RadA/Sms* bactérien. Comme le mutant *radA* est affecté dans les voies de réparation dépendantes de RecA, le rôle de RadA dans la recombinaison homologue est suggéré (Diver *et al.*, 1982; Beam *et al.*, 2002). Ce rôle est confirmé par les effets synergétiques des mutations des facteurs impliqués dans les étapes tardives de la recombinaison (RuvA, RecG) avec celle de RadA (Beam *et al.*, 2002). Son rôle exact dans la recombinaison homologue ne sera finalement identifié que très récemment (Cooper et Lovett, 2016; Marie *et al.*, 2017).

### 1. Structure de RadA

RadA est une hélicase à ADN paralogue de RecA, appartenant à la famille des hélicases SF4 et composée de trois domaines fonctionnels (**figure 23**) (Marie *et al.*, 2017). La partie N-terminale présente quatre cystéines formant un motif doigt de Zinc nécessaire à l'interaction avec la recombinaise RecA. Cette interaction est indispensable aux fonctions de RadA car la substitution d'une cystéine rend la protéine incapable de compléter la déficience de RadA lors de stress génotoxiques (Cooper *et al.*, 2015). A la suite, la protéine RadA présente un domaine ATPase RecA-like composé des motifs Walker A et Walker B et du motif KNRFNG caractéristique des protéines RadA. Les motifs Walker A et KNRFNG sont également indispensables aux fonctions de la protéine et sont impliqués dans la liaison avec l'ADN et l'activité hélicase de RadA lors de la migration de branche (Cooper *et al.*, 2015). La proximité de RadA avec RecA laisse penser que le mécanisme de migration de branche de RadA est similaire à celui de la recombinaise. Là où RecA aurait évolué pour la reconnaissance des séquences homologues et l'invasion de brin, RadA se serait plutôt spécialisé dans les étapes tardives de la recombinaison, améliorant l'activité de migration de branche de RecA. Les mutants des motifs Walker A et KNRFNG sont dominants négatifs, interférant avec les fonctions de la protéine sauvage lorsque les deux formes sont exprimées. Enfin, la partie C-terminale est constituée d'un domaine protéasique de type Lon. Les acides aminés impliqués dans l'activité protéase ne sont cependant pas conservés et le domaine Lon-like de RadA ne possède pas d'activité protéase (Inoue *et al.*, 2017). Ce domaine est impliqué dans la liaison à l'ADN et sert d'échafaudage pour la structure de la protéine. Seuls, les domaines Lon-like s'assemblent en anneaux hexamériques capables d'interagir avec l'ADNsb au niveau du canal central et des



**Figure 24 : Modèle de migration de branche de RadA.** L'invasion de brin est réalisée par le nucléofilament de RecA. Après la formation de la D-loop, RecA recrute un hexamère de RadA sur chacun des brins de la molécule réceptrice. RadA est représenté selon le code couleur de la figure 23 (le motif doigt de Zinc en bleu, le domaine hélicase en orange et le domaine protéase en vert). RadA se déplace sur l'ADNs<sub>b</sub> en ouvrant la molécule réceptrice (brins noir et gris) puis hybride l'un des brins (noir) avec l'ADN donneur (rouge). Le déplacement de RadA permet l'extension de la région d'homologie en 3' de l'ADN envahissant et le recrutement du quatrième brin pour former une jonction de Holliday en 5'. RadA pourrait déstabiliser le nucléofilament de RecA, permettant d'orienter la recombinaison vers les étapes de résolution en recrutant RecG et RuvABC. D'après Cooper et Lovett, 2016 et Marie *et al.*, 2017

jonctions intermoléculaires (**figure 23C**) (Inoue *et al.*, 2017). Ainsi, le domaine Lon-like permet d'oligomériser RadA sous sa forme active hexamérique (Marie *et al.*, 2017).

## 2. Rôle de RadA

RadA est impliquée dans les étapes tardives de la recombinaison homologue et a été identifiée comme la troisième voie de migration de branche avec RecG et RuvAB (Cooper *et al.*, 2015; Cooper et Lovett, 2016). Contrairement à RecG et RuvAB, abondamment étudiées au cours des dernières décennies, le rôle de RadA dans la recombinaison homologue n'a été que très récemment élucidé. Seule, la mutation *radA* n'affecte que très peu la survie de bactérie lors de stress génotoxiques (Beam *et al.*, 2002). La mutation est néanmoins synergétique avec celles d'autres facteurs de migration de branche, indiquant une redondance partielle des trois voies de maturation des intermédiaires de recombinaison (Beam *et al.*, 2002; Cooper *et al.*, 2015). Ces doubles mutants sont plus affectés dans leur survie que le mutant de la recombinase qui supprime complètement la recombinaison. Le double mutant *radA recG* peut être complété par la surexpression de RuvAB ou la mutation de RecF, suggérant que l'accumulation d'intermédiaires de recombinaison non-maturés est létale pour la cellule (Cooper *et al.*, 2015). En absence d'une des voies de migration de branche, les deux autres peuvent compenser la maturation des intermédiaires et permettre l'aboutissement du mécanisme de recombinaison. En absence de deux voies, la quantité ou les différences de types d'intermédiaires sont trop importantes pour être prises en charge par une seule voie, aboutissant à leur accumulation.

Après l'invasion de l'ADN homologue par le nucléofilament de RecA, les extrémités de la D-loop doivent être étendues (**figure 24**). Cela permet d'allonger l'appariement entre les deux molécules d'ADN homologues et de recruter le quatrième brin de l'ADN donneur. RecG et RuvB sont incapables de réaliser la migration de brin dans le contexte du nucléofilament de RecA (WHITBY *ET AL.*, 1993). Dans une expérience d'invasion de brin *in vitro*, l'ajout de RecG ou de RuvAB en présence de RecA entraîne l'inhibition de la recombinaison et le retour vers les formes initiales (WHITBY *ET AL.*, 1993). À l'inverse, il a été montré que RadA était actif dans le contexte du nucléofilament RecA (Cooper et Lovett, 2016; Marie *et al.*, 2017). L'ajout de RadA permet d'accélérer la réaction d'échange de brin réalisée par RecA *in vitro* en présence ou en absence de RecA (Cooper et Lovett, 2016). Il a même été montré que l'interaction de RadA avec RecA était nécessaire pour les fonctions de RadA (Marie *et al.*, 2017). RadA permettrait ainsi de prendre en charge les intermédiaires de recombinaison immédiatement après la formation de la D-loop. L'interaction avec RecA permet de positionner un hexamère de RadA de part et d'autre de la D-loop (**figure 24**). L'activité hélicase à ADNsb permettrait de délier les deux brins d'ADN de la molécule réceptrice puis d'hybrider l'un des brins avec l'ADNsb envahissant, à la manière de RecA (Marie *et al.*, 2017). Le chargement de RadA de chaque côté de la D-

permettrait ainsi la migration de branche dans les deux sens, internalisant les régions homologues sur de longues distances. Enfin, RadA aurait aussi un rôle dans la déstabilisation de RecA, et permettrait ainsi d'orienter le mécanisme de recombinaison vers les étapes de résolution (Cooper et Lovett, 2016).

Chez *Bacillus subtilis*, RadA est impliquée dans le contrôle de l'activité de DisA pour la signalisation des stress réplicatifs (Gandara *et al.*, 2017; Torres *et al.*, 2019). Chez *Deinococcus radiodurans*, RadA est également impliquée dans une voie de réparation indépendante de RecA (Slade *et al.*, 2009). RadA pourrait également y permettre la dissolution des HJ pour favoriser un mécanisme de SDSA. Chez *Oryza sativa* (riz), une protéine RADA a été caractérisée dans le noyau (Ishibashi *et al.*, 2006). Dans cet article, il est décrit que la protéine RADA serait localisée dans le noyau et que l'inactivation du gène par une approche de RNAi affecte peu le développement de la plante, mais augmente la sensibilité à certains stress génotoxiques. Néanmoins, ces résultats diffèrent énormément de ceux obtenus par notre équipe. Nous pensons que c'est à cause de l'utilisation d'un anticorps trop peu spécifique pour la localisation, et la réduction insuffisante de l'expression du gène *RADA* par l'approche de RNAi.

## Résultats

Les résultats obtenus pour la partie RADA sont présentés sous la forme d'une publication scientifique qui sera soumise prochainement. La discussion de l'article est étendue pour couvrir les perspectives à court et long terme de la thèse.

### 1. Résumé de la publication

Les étapes tardives de la recombinaison homologues dans les organelles de plante sont méconnues. Seul RECG1, l'orthologue de RecG a été décrit, mais sa mutation n'affecte que très peu le développement d'*Arabidopsis* alors que le mutant de la recombinase RECA2 n'est pas viable. En absence de RuvAB chez les plantes, nous nous sommes intéressés à la troisième voie de maturation des intermédiaires de recombinaison : RADA.

Nous avons confirmé l'adressage de RADA aux organelles en fusionnant son OTS à la séquence de la GFP. Nos résultats indiquent que RADA est adressé aux mitochondries et aux chloroplastes. Nous avons modélisé la structure de RADA à partir de celle de *S. pneumoniae* cristallographiée. Les structures modélisées sont similaires, suggérant la conservation des activités de la protéine bactérienne chez les plantes. Nous avons montré que RADA pouvait compléter le mutant bactérien *radA* dans sa survie en présence de ciprofloxacine, un antibiotique génotoxique induisant des DSB. Cette complémentarité confirme la conservation des fonctions de RADA entre bactéries et plantes. Nous avons purifié la protéine RADA fusionnée à une étiquette 6-His, ainsi qu'un mutant du domaine ATPase supposé inactif, en

système hétérologue chez *E. coli*. Une analyse par diffusion dynamique de la lumière a montré que RADA est très majoritairement purifiée sous une forme d'hexamères. Nous avons ensuite testé la capacité de liaison à l'ADN et les activités de ces protéines lors de tests *in vitro*. RADA lie préférentiellement les molécules contenant de l'ADNsb. Tout comme la protéine bactérienne, RADA est capable d'accélérer une réaction d'échange de brin *in vitro* lorsqu'elle est initiée par RecA. RADA semble néanmoins inhiber l'initiation de nouvelles réactions, conduisant à une accumulation moindre (mais plus rapide) de produits finaux. La version mutée de RADA inhibe complètement la réaction.

Nous avons sélectionné deux mutants par insertion de T-DNA dans le gène *RADA* d'*Arabidopsis*. Nous avons observé un impact sévère de la mutation sur le développement des plantes. Les mutants présentent un important retard de croissance, une distorsion des feuilles et un raccourcissement des racines. Ils sont également pratiquement stériles et ne produisent que très peu de graines viables. Cette stérilité semble provenir de défauts au niveau des appareils reproducteurs mâles et femelles. L'observation des cellules au microscope optique a montré une diminution de la viabilité des embryons après fécondation. En microscopie électronique à transmission, les mitochondries présentent une densité électronique moindre. La transformation des mutants avec une construction exprimant RADA:HA permet de supprimer les phénotypes développementaux.

Les mutants *radA* montrent une altération importante de la stœchiométrie de l'ADNmt dû à une augmentation de la recombinaison ectopique. Plus les mutants sont affectés au niveau du développement, plus la stœchiométrie de l'ADNmt est altérée. Ces recombinaisons ectopiques conduisent à la formation de sous-génomés se répliquant de manière autonome indépendamment du reste du génome. L'analyse du transcriptome mitochondrial par RT-qPCR n'a pas permis d'identifier les causes des phénotypes développementaux observés. RADA semble avoir un rôle plus important dans la mitochondrie que le chloroplaste car la stichométrie de l'ADNcp n'est pas affectée dans le mutant *radA*. Enfin, nous avons produits plusieurs double mutants en croisant les plantes *radA* avec d'autres mutants de facteurs impliqués dans la recombinaison organellaire. Les mutations *radA* et *recG1* ne sont pas synergétiques, indiquant que RADA est certainement la voie de migration de branche principale dans les organelles. En revanche, le double mutant *radA recA3* n'est pas viable, révélant une voie de recombinaison alternative dépendante de RECA3.

## 2. Article



# Homologous recombination in plant mitochondria requires the main RADA-dependent branch migration pathway

Nicolas Chevigny, Cédric Nadiras, Marc Bichara, Julien Schmitz, André Dietrich and José M. Gualberto

Institut de Biologie Moléculaire des Plantes, CNRS UPR2357, Université de Strasbourg, 12 rue du General Zimmer, 67084 Strasbourg, France

## ABSTRACT

In plant organelles the late stages of homologous recombination are still poorly understood. RECG1, the plant ortholog of bacterial RecG, has little effects in Arabidopsis development, although the loss of the RECA2 recombinase is not viable. Here we present an alternative pathway for the processing of organellar recombination intermediates, the one dependent from RADA. RADA is found in all plants, and our results show that it is targeted to both mitochondria and chloroplasts. RADA is similar in structure to bacterial RadA, and can complement the bacterial *radA* mutant for survival under genotoxic stress, implying similar activities and functions. *In vitro*, RADA preferentially binds ssDNA, and is able to accelerate a strand-exchange reaction initiated by RecA, while at the same time inhibits initiation of the reaction, resulting in less, but faster accumulation of final heteroduplex products.

T-DNA insertion mutants of *RADA* are severely impacted on their development and fertility. The mutants show significant changes in the stoichiometry of mtDNA sequences, due to an increase in ectopic recombination and uncoordinated replication of subgenomes generated by recombination, molecular phenotypes that correlate with the severity of the developmental phenotypes. RADA has a more essential role in mitochondria than in chloroplasts, because no effects were observed in cpDNA stability. The *radA* and *recG1* mutations are not synergistic, indicating that RADA is the main branch migration pathway of plant organelles. In contrast, the double mutation *radA recA3* is lethal at the seedling stage, as the *recA2* mutation, revealing the importance of a RECA3-dependent alternative recombination pathway. Analysis of the mitochondrial transcriptome failed to correlate the developmental phenotypes with defects in mitochondrial gene expression. Rather, we found that activation of genes that repress cell cycle progression might be a cause for *radA* growth defects.

## INTRODUCTION

The mitochondrial genome (mtDNA) of vascular plants is large, ranging from about 200-300 kbs up to 11.3 Mbs in *Silene conica* (Sloan *et al.*, 2012). It is mostly constituted by non-coding sequences and despite its large size it only encodes fewer more genes than the mtDNAs of mammal or yeast. The structure of these large genomes is complex and still not clearly understood, harbouring a heterogeneous population of linear, circular and branched double-stranded (ssDNA) or single-stranded DNA (dsDNA) molecules (Backert *et al.*, 1997). This collection of subgenomic molecules can usually be mapped into a single circular chromosome, called the master circle. But genomes constituted by multiple chromosomes also exist, in particular in the species with the largest plant mitogenomes (Sloan *et al.*, 2012). The complexity of the plant mtDNA comes from frequent recombination events involving repeated sequences. These recombination events are responsible for the rapid evolution of mtDNA structure which leads to important mtDNA rearrangements even within the same specie. Recombination is also responsible for the high conservation of gene nucleotide sequences (Mower *et al.*, 2007; Christensen, 2018).

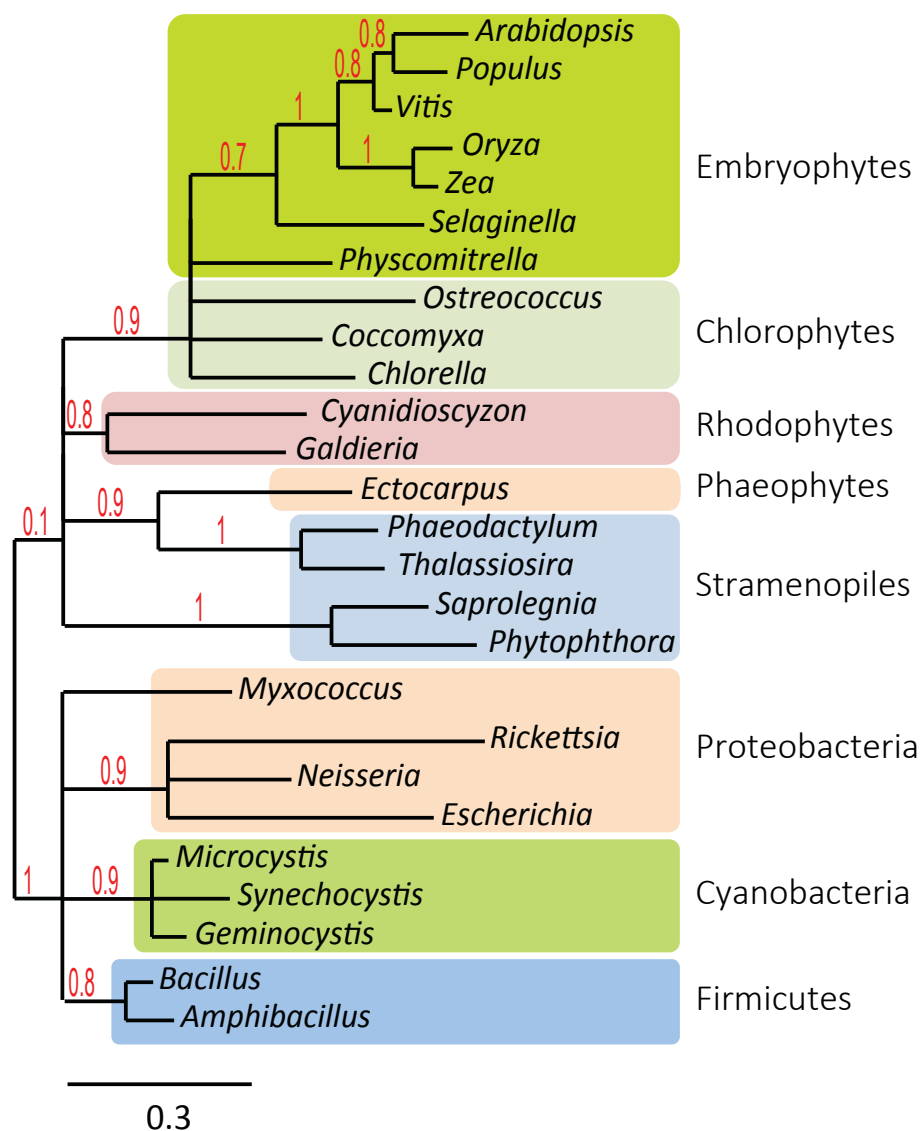
The recombination of the plant mtDNA relies on the high number of repeated sequences that exist in the genome. Large repeated sequences (> 500 pb) are involved in frequent and reversible homologous recombination (HR), while intermediate-size repeats (IRs) (50-500 pb) or microhomologies (< 50 pb) can promote infrequent ectopic or illegitimate recombination, respectively (Woloszynska et Trojanowski, 2009; Maréchal et Brisson, 2010; Janicka *et al.*, 2012). Recombination involving IRs or microhomologies contribute to the heteroplasmic state of mtDNA by creating sub-stoichiometric alternative configurations (mitotypes) that co-exist with the main genome (Small *et al.*, 1989). These variants may present altered gene expression, by the displacement of regulatory sequences or by the disruption of gene sequences (Hanson et Bentolila, 2004). These processes may lead to the creation and expression of chimeric ORFs that can be deleterious for mitochondrial function. Sub-stoichiometric mtDNA variants can become the predominant genome by the yet unclear process of sub-stoichiometric shifting (SSS), which can occur in the time frame of a single plant generation (Small *et al.*, 1989) ; (Chen *et al.*, 2011). SSS is responsible for the occurrence of the cytoplasmic male sterility (CMS) phenotype, which are plant varieties unable to produce viable pollen, in most cases because of the expression of chimeric ORFs (Budar et Pelletier, 2001; Touzet et Meyer, 2014).

Homologous recombination is also the main DNA repair pathway of plant mitochondria, for the repair of double strand breaks (DSB) and the copy-correction of mutations, thus contributing for the very slow evolution of coding sequences. But in order to avoid mtDNA instability and its deleterious outcomes, recombination needs to be tightly regulated. Several factors involved in recombination pathways and in recombination surveillance have been

identified, but their exact roles are still poorly understood. In most cases these factors are homologs of prokaryotic recombination factors that were inherited from the symbiotic ancestors of mitochondria and chloroplasts (Boesch *et al.*, 2011; Gualberto et Newton, 2017). Consequently, recombination mechanisms in plant organelles are considered to be similar to the prokaryotic ones. However, the organellar pathways might be more complex than their bacterial counterpart, as they involve additional factors with apparently partially redundant functions (Edmondson *et al.*, 2005; Zaegel *et al.*, 2006; Cappadocia *et al.*, 2010; Miller-Messmer *et al.*, 2012). Organellar homologous recombination involves several ssDNA-binding proteins whose individual roles still need to be assigned (Edmondson *et al.*, 2005; Zaegel *et al.*, 2006; Cappadocia *et al.*, 2010). In the same way, plant organellar recombination relies on the RecA-like recombinase RECA2, which is addressed to both organelles and whose mutation is lethal, but also involves the chloroplastic RECA1 and mitochondrial RECA3 (Miller-Messmer *et al.*, 2012; Zampini *et al.*, 2015). Defect in factors involved in mitochondrial homologous recombination generally results in increased ectopic recombination involving IRs and microhomologies resulting in genomic rearrangements, because of the enhanced activity of alternative error-prone repair pathways (Miller-Messmer *et al.*, 2012; Wallet *et al.*, 2015).

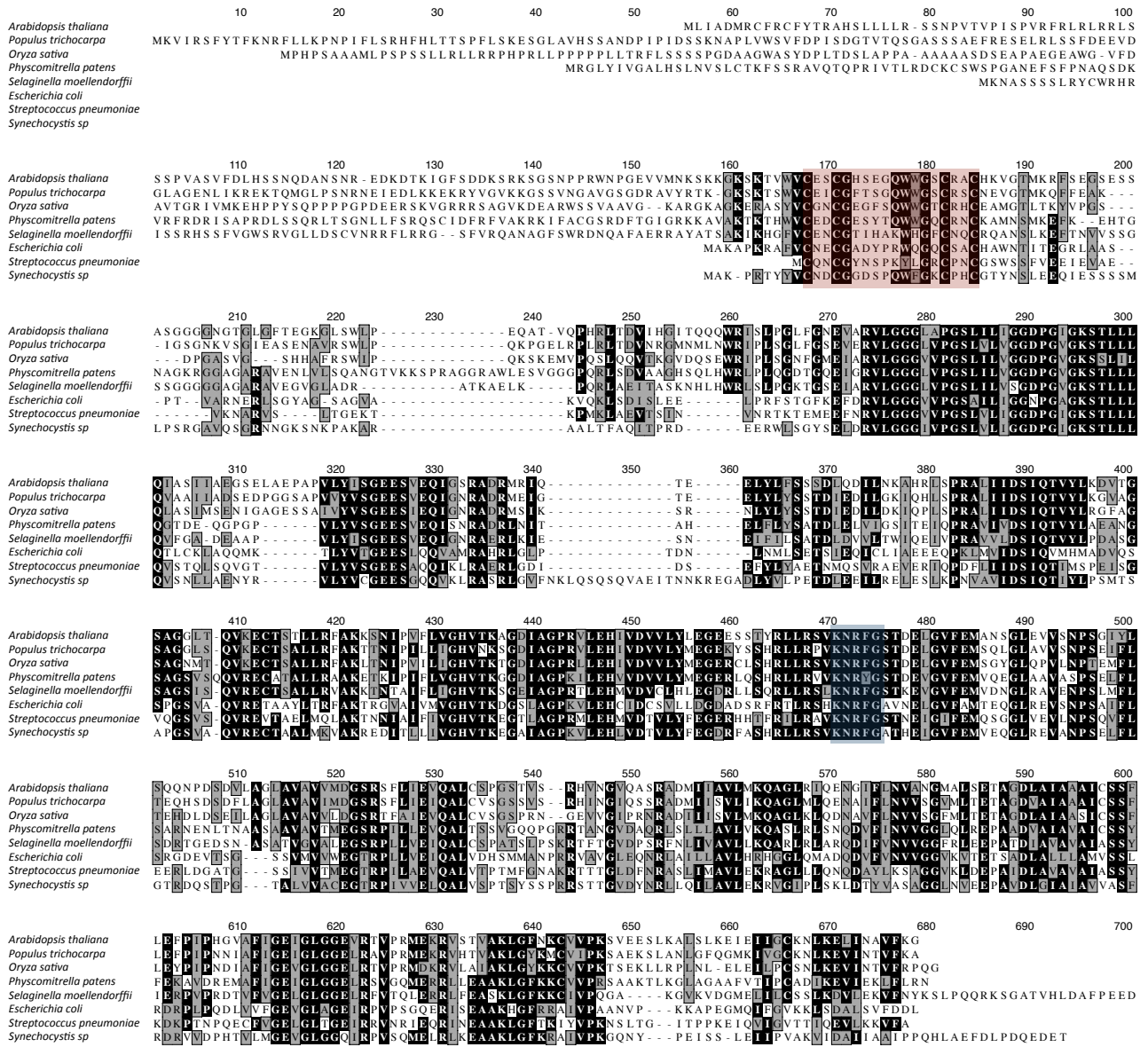
In bacteria, homologous recombination is initiated by the RecBCD-mediated resection of DSB 5' extremities forming 3' ssDNA overhangs (Taylor *et al.*, 2014). Then, the RecBCD complex promotes the formation of the presynaptic filament by loading RecA on ssDNA (Cockram *et al.*, 2015). For single-stranded gaps (SSG), RecA is loaded on ssDNA by the RecFOR mediator complex after the release of SSB (Umezumi et Kolodner, 1994). The RecA nucleofilament is extended and seeks for homologies in the genome by probing multiple heterologous sequences (Forget et Kowalczykowski, 2012; Ragunathan *et al.*, 2012). When the homologous sequence is identified by Watson-Crick pairing, RecA-mediated ATP hydrolysis stabilizes the invading DNA and forms the synaptic complex, known as the D-loop. Finally, post-synaptic steps include branch migration and resolution of the heteroduplex. Branch migration allows the maturation of the recombination intermediates. Branch migration factors use helicase activity to extend homologous region at both side of the D-loop, allowing the recruitment of the fourth DNA strands to form a Holliday Junction (HJ) (Whitby *et al.*, 1994; West, 1997; Cooper et Lovett, 2016). A nuclease resolves HJ and frees both DNA molecules.

Three partially redundant pathways of branch migration have been described in bacteria, respectively involving RuvAB, RecG or RadA (Beam *et al.*, 2002; Cooper et al., 2015). *radA* or *recG* single mutants are mildly affected but *radA recG* double mutant is severely impaired in its survival under genotoxic conditions (Cooper et Lovett, 2016). Deficiency in multiple branch migration pathways is even more deleterious than the complete absence of recombination in the *recA* mutant. Accumulation of toxic unprocessed recombination intermediates is responsible of the enhanced lethality observed in *recG radA* or *recG radA RuvA*



**Figure S1.** Phylogenetic distribution of RADA.

Genes coding for RadA-like proteins are found in all bacteria, in land plants, green, brown and red algae, diatoms and other organisms of the Stramenopile group. *Arabidopsis thaliana*, NP\_199845; *Populus trichocarpa*, EEE84551; *Vitis vinifera*, XP\_002277638; *Oryza sativa*, NP\_001056828; *Zea mays*, NP\_001170708; *Selaginella moellendorffii*, XP\_002976563; *Physcomitrella patens*, XP\_001757578; *Ostreococcus tauri*, XP\_003084463; *Coccomyxa subellipsoidea*, XP\_005643141; *Chlorella variabilis*, EFN56488; *Cyanidioschyzon merolae*, XP\_005536634; *Galdieria sulphuraria*, XP\_005709405; *Ectocarpus siliculosus*, CBJ25917; *Phaeodactylum tricornutum*, XP\_002178713; *Thalassiosira oceanica*, EJK58798; *Saprolegnia diclina*, EQC30001; *Phytophthora infestans*, XP\_002904225; *Myxococcus xanthus*, YP\_629513; *Rickettsia prowazekii*, WP\_014607237; *Neisseria meningitidis*, WP\_002258526; *Escherichia coli*, WP\_001458566; *Microcystis aeruginosa*, WP\_002742386; *Synechocystis* sp, WP\_009633429; *Geminocystis herdmanii*, WP\_017296205; *Bacillus anthracis*, NP\_842650; *Amphibacillus jilinensis*, WP\_017473696.



\*\* Identity Scores (%) \*\*

	<i>Arabidopsis thaliana</i>	<i>Populus trichocarpa</i>	<i>Oryza sativa</i>	<i>Physcomitrella patens</i>	<i>Selaginella moellendorffii</i>	<i>Escherichia coli</i>	<i>Streptococcus pneumoniae</i>	<i>Synechocystis sp</i>
<i>Arabidopsis thaliana</i>	100.0	54.8	48.1	40.4	37.7	28.5	32.4	29.9
<i>Populus trichocarpa</i>	68.3	100.0	54.0	39.3	34.1	25.8	28.7	26.8
<i>Oryza sativa</i>	64.4	67.6	100.0	38.4	33.3	26.1	28.8	28.1
<i>Physcomitrella patens</i>	56.9	52.6	53.2	100.0	44.9	29.2	33.8	31.0
<i>Selaginella moellendorffii</i>	54.5	47.9	48.2	58.3	100.0	30.9	34.8	33.5
<i>Escherichia coli</i>	44.9	41.3	42.3	44.8	46.2	100.0	41.3	42.5
<i>Streptococcus pneumoniae</i>	47.4	44.4	43.9	48.6	49.6	60.5	100.0	43.5
<i>Synechocystis sp</i>	43.9	41.4	42.0	45.0	49.0	61.1	62.1	100.0

\*\* Similarity Scores (%) \*\*

**Figure S2.** High conservation of RADA sequences between plants and bacteria. Sequence alignment (Clustal W) between representative land plant RADA sequences and RadA from proteobacteria and cyanobacteria. The Zinc-finger and KNRF RadA-specific motif are shaded in red and blue respectively. The matrix of sequences identities and similarities is shown below the alignment.

highlighting the crucial role of branch migration factors (Beam *et al.*, 2002; Cooper *et al.*, 2015). RadA was only recently characterized for its ability to accelerate RecA-mediated strand exchange *in vitro* (Cooper et Lovett, 2016). It is an ATP-dependant ssDNA helicase composed of three functional domains: a N-terminal C4 zinc-finger, an ATPase RecA-like domain, and a Lon protease-like domain (Diver *et al.*, 1982; Cooper *et al.*, 2015). The active form of RadA is a hexamer (Marie *et al.*, 2017). In comparison with RecG and RuvAB, RadA interacts with RecA and can function in the RecA filament context (Marie *et al.*, 2017).

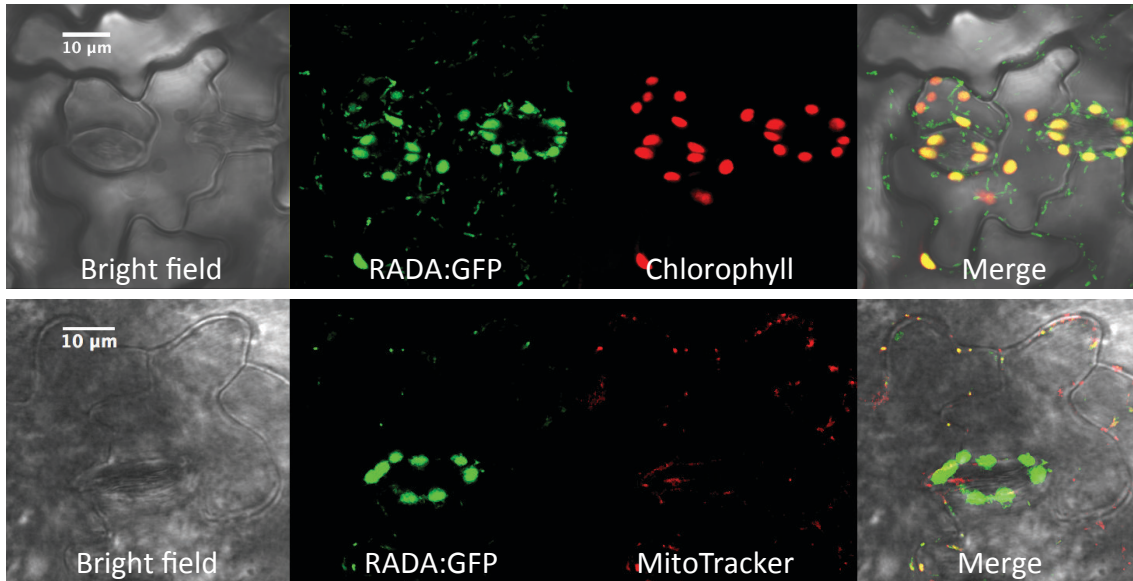
Plant genomes don't encode any homolog of the RuvAB complex, but have a RecG homolog (RECG1) involved in mitochondrial recombination (Wallet *et al.*, 2015). Despite the absence of a RuvAB pathway, plant *recG1* loss of function mutants are only mildly affected. As organellar recombination is essential for plant development we would expect a more deleterious effect of the loss of branch migration activities (Cooper et Lovett, 2016). These results suggest the existence of additional pathways for the maturation of recombination intermediates. We identified a RadA homolog that could be potentially involved in the late steps of organellar recombination pathways. We show that Arabidopsis RADA is indeed involved in mitochondrial recombination and possesses same activities as bacterial RadA. Interestingly, *radA* mutants are severely affected in their development, indicating that RADA has a more important role in plants than in bacteria.

## RESULTS

### **A gene coding for RADA is present in the genomes of all plant species and is targeted to both mitochondria and chloroplasts**

In bacteria, the RecG mutation is strongly synergetic with RadA/Sms for the repair of DNA damages, and the double mutant *recG radA* is as affected as the RecA-deficient *recA* mutant (Beam *et al.*, 2002; Cooper *et al.*, 2015). In Arabidopsis, mitochondrial RECA-dependent recombination is apparently essential (Miller-Messmer *et al.*, 2012), but the loss of the organellar-targeted RECG1 has only mild effects on plant fitness (Wallet *et al.*, 2015). This suggested that additional genes exist coding for redundant or overlapping functions with RECG1 in plant organelles, such as RadA and RecG in bacteria. We have therefore screened the Arabidopsis (*Arabidopsis thaliana*) genome for orthologs of bacterial RadA. We retrieved a gene (At5g50340) coding for a protein 45% similar to bacteria RadA, which we named RADA.

Phylogenetic analysis shows that genes coding for RADA are present in all plants, including land plants, green, brown and red algae, diatoms and also in several organisms of the Stramenopile group that are not photosynthetic, like the water mold *Phytophthora infestans*



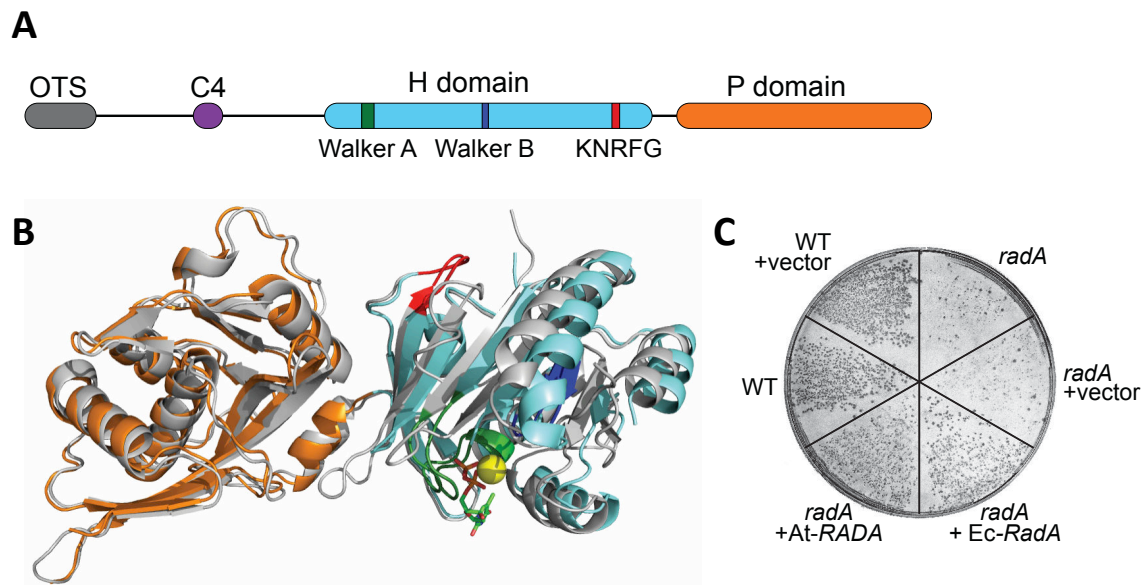
**Figure 1.** Arabidopsis RADA is targeted to both chloroplasts and mitochondria. The RADA:GFP fusion protein was constitutively expressed in transgenic plants and its location observed in leaf epidermal cells. It co-localized with both chloroplasts (autofluorescence of chlorophyll) mitochondria (red fluorescence of MitoTracker).

(Figure S1). While it is probable that plants inherited *RADA* from the prokaryote ancestor of mitochondria or chloroplasts, the strong conservation of the sequences doesn't allowed us to infer if the plant ancestor was a proteobacterial or a cyanobacterial RadA. Sequence alignments (Figure S2) showed that plant RADA proteins have all the important functional motifs that have been described for bacterial RadA (Cooper et Lovett, 2016; Marie *et al.*, 2017). In addition, Plant RADA have non-conserved N-terminal extensions that could correspond to organellar targeting peptides (Figure S2). This is supported by *in silico* predictions that give Arabidopsis RADA as a mitochondrial protein (<http://suba.plantenergy.uwa.edu.au>). This assumption was tested by expression of a construct encoding the N-terminus of RADA including the C4 domain fused to eGFP. The fusion protein was constitutively expressed in transgenic Arabidopsis plants, under control of a double 35S promoter. Observation of the protein fluorescence in leaves showed that it localizes in both chloroplasts and mitochondria, as shown by co-localization with the auto-fluorescence of chlorophyll and the red fluoresce of MitoTracker® (Figure 1). Thus, RADA is apparently a dually targeted organellar protein, like RECG1 and several other factors involved in organellar genomes maintenance (Wallet *et al.*, 2015; Gualberto et Newton, 2017).

### **Structural and functional conservation of RADA as compared to bacterial RadA**

Plant RADA has the same modular structure as bacterial RadA (Figure 2A). A three dimensional model of Arabidopsis RADA could be generated (Figure 2B), on the basis of the known RadA structure of *Streptococcus pneumoniae* (Marie *et al.*, 2017). The RadA structure comprises the N-terminal C4 zinc-finger, which was shown to be required for the interaction with the RecA recombinase, and two main domains: an ATPase RecA-like domain, and a Lon protease-like domain (Diver *et al.*, 1982; Cooper *et al.*, 2015). The interaction between C4 and RecA is apparently indispensable for RadA function, as substitution of one cysteine abolishes complementation of RadA deficiency under genotoxic (Cooper *et al.*, 2015). The RecA-like ATPase domain comprises the Walker A and B motifs, and a KNRFG sequence specific of RadA (Figures S2 and 1A). The Walker A and KNRFG motifs are also indispensable for the branch function of the protein, and are involved in the DNA-binding and helicase activities of RadA (Cooper *et al.*, 2015; Marie *et al.*, 2017). In bacteria, Walker A and KNRFG mutants are dominant negative, interfering with the function of the wild-type protein. The similarity between RadA and RecA suggests that the mechanisms of branch migration of RadA is similar to the one of the recombinases, but while RecA specialized in the recognition of homologous sequences and strand invasion, RadA had rather evolved for functioning in latter steps of recombination, specializing in branch migration. Finally, the C-terminal P-domain of RadA is similar to the Lon protease-like domain of RecA. But it is an inactive domain, since the residues involved in protease activity are not conserved, and no RadA protease activity could be detected (Inoue *et al.*, 2017). Rather, this domain is involved in the binding to dsDNA and works as a scaffold for





**Figure 2.** The Arabidopsis RADA is structurally and functionally homologous to bacterial RadA. **(A)** The modular structure of plant RADA is similar to the one from bacteria, with a N-terminal zinc-finger (C4), an helicase and Lon-protease-like domains (H and P domains respectively). Plant precursor proteins have an N-terminal extension containing an organellar targeting sequence (OTS). **(B)** Model of Arabidopsis RADA superposed on the known structure of *S. pneumoniae* (in gray) (Marie et al., 2017). The color code of relevant domains is as in A. A bound ADP and the Mg<sup>2+</sup> ion (yellow ball) are shown. **(C)** Complementation of a *E. coli radA* mutant for growth in the presence of genotoxic ciprofloxacin. The Arabidopsis RADA gene (*At-RADA*) is as efficient as the bacterial gene (*Ec-RadA*) for complementation

the protein architecture, promoting its oligomerization as hexameric rings (Inoue *et al.*, 2017; Marie *et al.*, 2017).

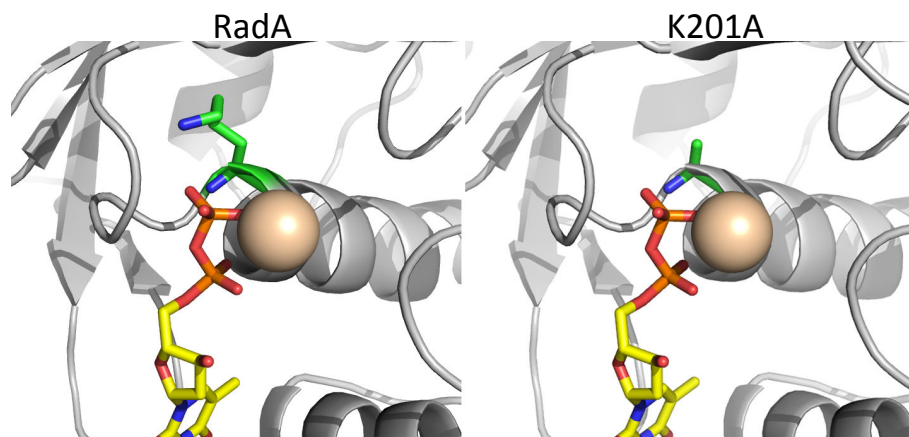
Since all domains of plant RADA perfectly match the bacterial ones, it is expected that it has the same activities and potentially similar functions as the bacterial enzyme. We have therefore tested the ability of plant RADA to complement *Escherichia coli* RadA in the repair of genotoxic stress-induced DNA lesions, as described in Materials and Methods. A *radA785(del)::kan* strain was used for complementation assays, and Arabidopsis RADA was expressed in the low-copy number plasmid pACYC. As described before, we found that as compared to WT the *radA* strain was little affected by UV-C in their growth (Beam *et al.*, 2002; Cooper *et al.*, 2015). In our hands, the best condition to test complementation where we found the highest difference between WT and *radA* was the growth in the presence ciprofloxacin, an antibiotic that induces DNA double-strand breaks. In these conditions we found that expression of Arabidopsis RADA could complement the growth deficiency as well as the homologous bacterial sequence (Figure 2C).

### **RADA is a ssDNA-binding protein that stimulates branch migration in strand-exchange reactions**

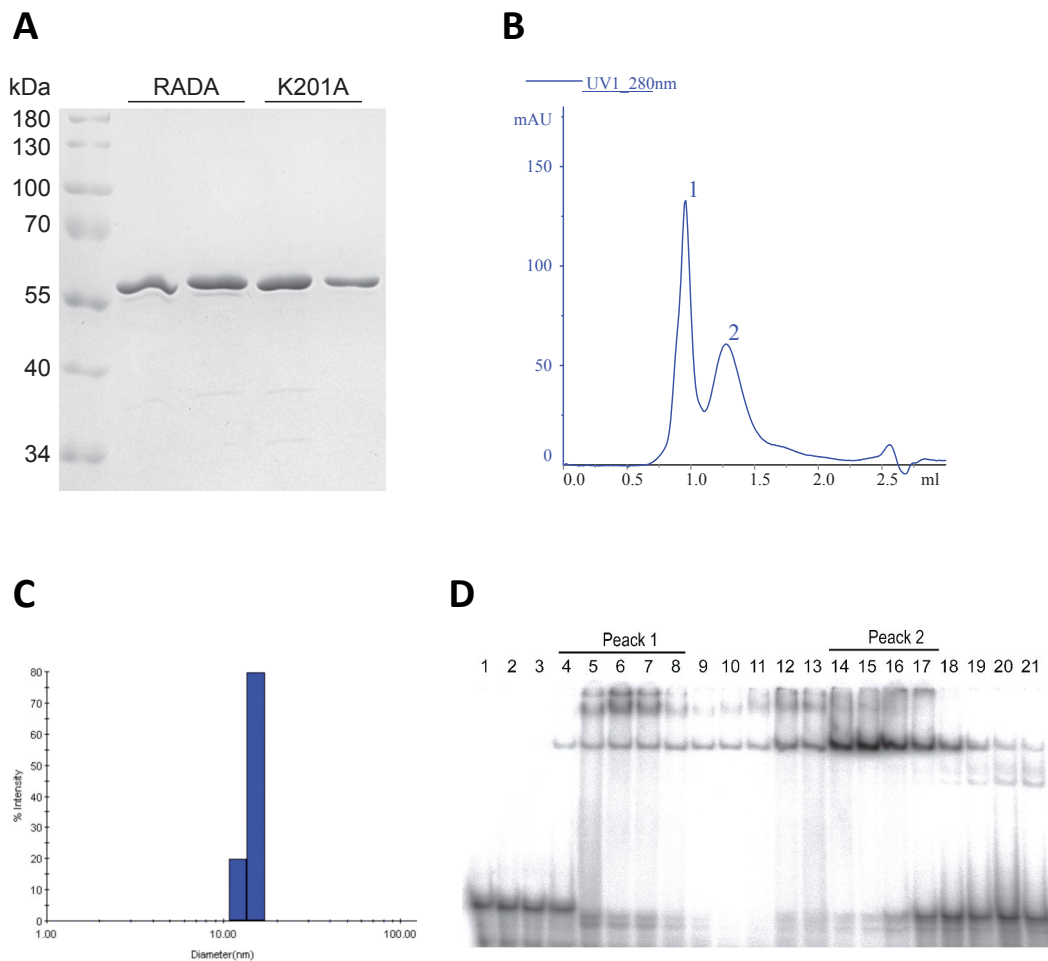
In order to test plant RADA activities we expressed a recombinant protein in bacteria. The Arabidopsis cDNA sequence minus the first 45 codons coding for the putative organellar targeting sequence was cloned in expression vector pET28a, fused to a N-terminal His-tag. A mutant sequence was prepared by point mutagenesis, to express a Walker A deficient protein (K201A). The mutation of the equivalent position in *E. coli* RadA abolished DNA-dependent ATPase activity and it is a dominant negative *radA* allele (Cooper *et al.*, 2015; Cooper et Lovett, 2016). According to the structure of the protein bound to ADP the mutation should not affect ATP or ADP binding, even if the ATPase activity is lost (Figure S3).

Both wild type and K201A could be expressed and purified in nickel columns as soluble proteins (Figure S4A). By gel filtration the purified WT protein resolved as two peaks of high molecular weight (Figure S4B), indicating different degrees of protein oligomerization. Analysis of the smaller peak by dynamic light scattering showed that it was mainly constituted by a monodispersed particle of about 340 kDa, consistent with an hexameric RADA complex (Figure S4C). Preliminary EMSA experiments showed that both peak fractions could bind to a ssDNA oligonucleotide probe giving complexes of same apparent size, with fractions of peak 1 giving predominantly higher molecular weight complexes (Figure S4B).

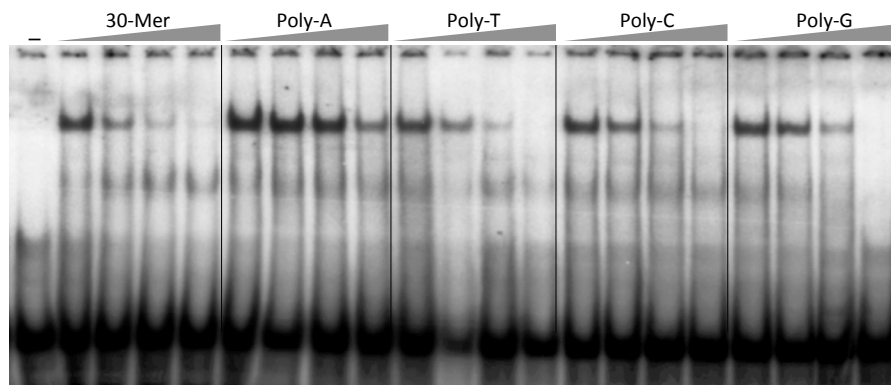
The purified proteins were then used in EMSA experiments to test the DNA-binding specificities of RADA. Different DNA structures were tested, including ssDNA, dsDNA, fork-like structures and double-strand molecules containing 5' or 3' ssDNA overhangs. Using the same



**Figure S3.** Effect of the point mutation K201A (amino acids in green) on the Walker A domain of Arabidopsis RADA in the binding and hydrolysis of ATP. The structure shown is the one from bacterial RadA (Marie et al. 2017), with bound ADP and  $Mg^{2+}$  ion (sphere). The two phosphate groups of ADP are in red.

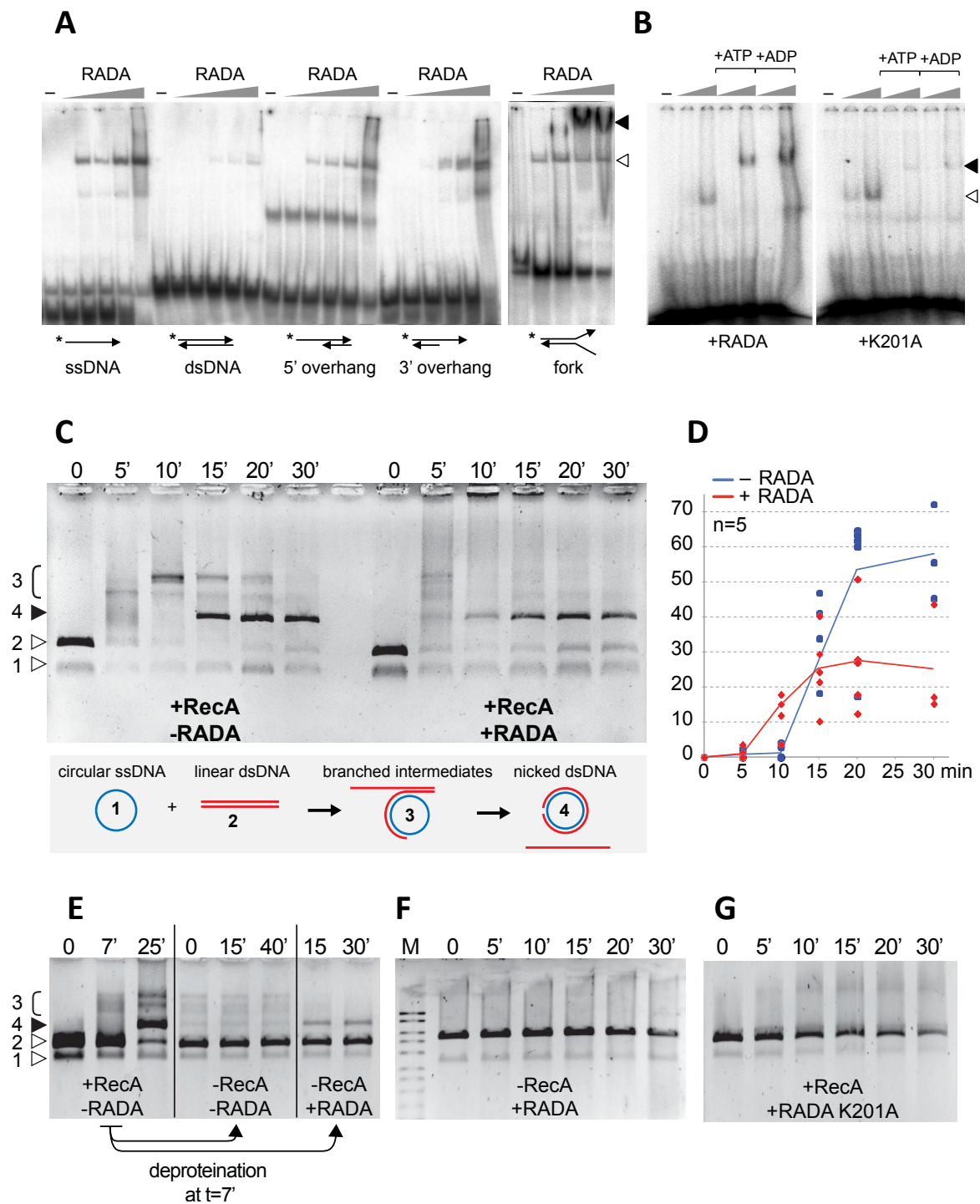


**Figure S4.** Expression, purification and characterization of recombinant RADA. The Arabidopsis RADA sequence minus the first 48 codons corresponding to the OTS was cloned in the expression vector pET28a fused to a N-terminal His-tag. The recombinant RADA and Walker mutant K201A were expressed in the Rosetta(DE3) strain and purified by affinity and gel filtration. **(A)** Coomassie gel staining analysis of the recombinant proteins. **(B)** Gel filtration on Superdex S200 showed that RADA purified as two peaks of high molecular weight. **(C)** Dynamic light scattering of the protein fraction from peak 2 shows that it is monodispersed and corresponding to a size of about 340 kDa, which is consistent with an hexameric RADA molecule. **(D)** EMSA analysis of the binding to a ssDNA oligonucleotide. Fractions corresponding to both peaks give complexes of the same size, although fractions of peak 1 give predominantly higher molecular weight complexes.



**Figure S5.** Sequence specificity of RADA binding to ssDNA.

A 30-mer ssDNA oligonucleotide (7x[AGCT]AG) was used as probe in EMSA experiments with recombinant RADA, and sequence specificity was tested by competition with increasing concentrations of the cold homologous oligonucleotide or with 30-oligomers (poly-A, poly-T, poly-C and poly-G). Competitor/probe ratios tested were 0; 2.5; 10 and 40. Only poly-A showed reduced competition for RADA binding.



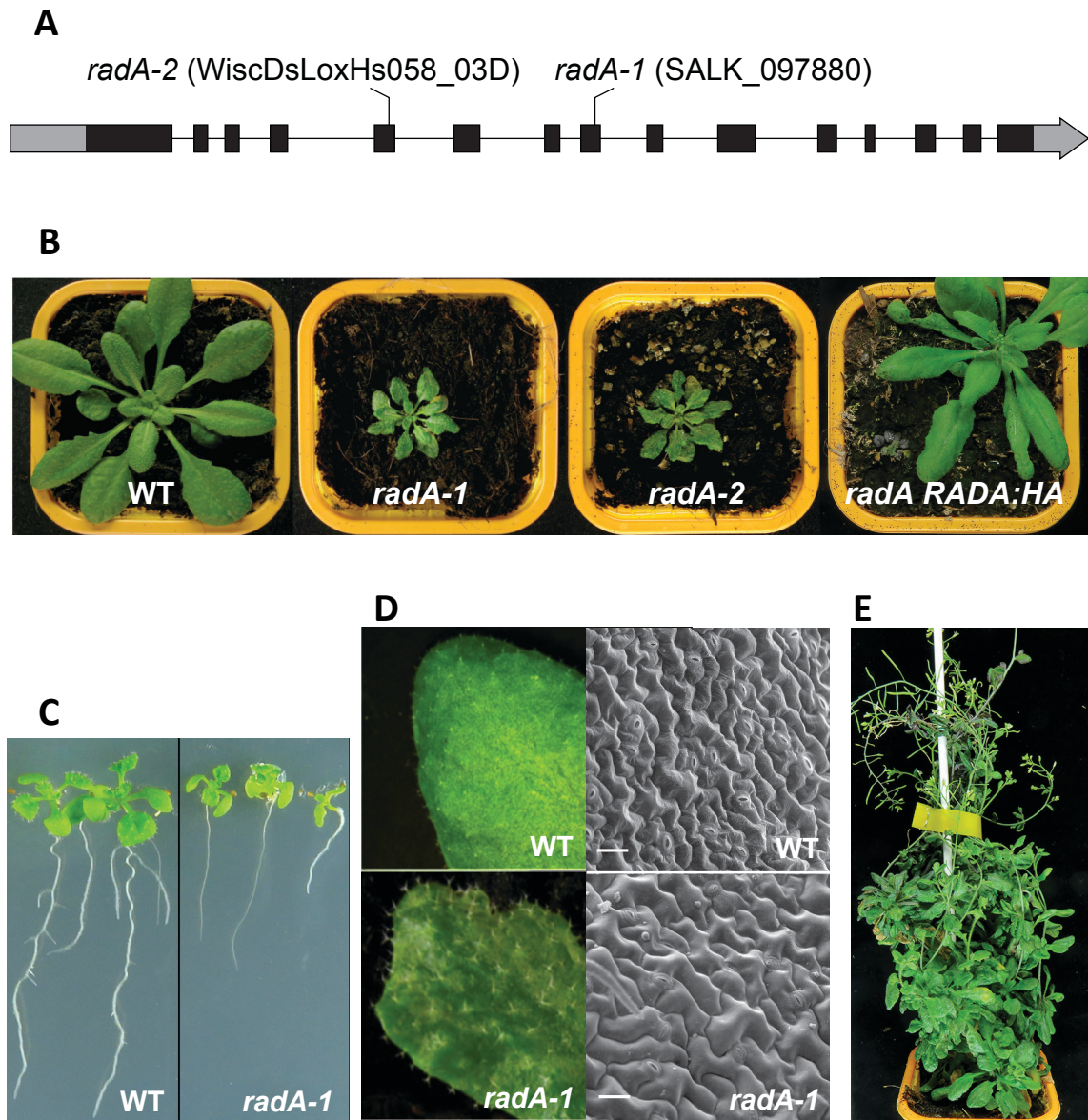
**Figure 3.** DNA-binding and branch-migration activities of RADA. **(A)** EMSA experiments showing that RADA binds any ssDNA-containing DNA structure with higher affinity than dsDNA. Lower and higher molecular weight complexes are indicated by white and black arrowheads respectively **(B)** Analysis on low concentration gel (4.5% as compared to 8% in A) of the formation of a high-molecular weight RADA filament on ssDNA, which is promoted by ATP or ADP (1 mM). The K201A mutant protein binds with equivalent affinity as the WT protein. Increasing concentrations of RADA used in A and B are indicated by the grey triangles. **(C)** In a *in vitro* strand-invasion reaction plant RADA accelerates branch-migration of DNA heteroduplexes initiated by RecA. An explanation of the different substrates and products is shown below the gel. **(D)** Ratio of final product as compared to the initial linear dsDNA substrate in 5 independent experiments, showing that in the presence of RADA there is faster resolution of branched intermediates. However RADA inhibits initiation, as fewer final products are produced. **(E)** RADA can alone finalize branch-migration initiated by RecA: a reaction at T=7 min was arrested by deproteination (left panel) and the DNA purified. Without addition of RecA or RADA there is no spontaneous progression of the reaction (middle panel), but RADA alone can resolve intermediates into the final product (right panel). **(F)** RADA alone cannot initiate strand invasion. **(G)** Mutation of the ATPase Walker domain of RADA (K201A) inhibits the reaction.

probe concentrations and increasing concentrations of recombinant RADA we found that it could bind with equivalent affinities to all structures containing ssDNA regions, and that it could also bind dsDNA, but with much less affinity (Figure 3A). On competition experiments, pol-T, poly-C and poly-G could compete binding as efficiently as the homologous probe, while pol-A was much less efficient as competitor (Figure S5). Thus, Arabidopsis RADA binds preferentially to ssDNA, with little sequence specificity, and apparently without the requirement of ATP or ADP. However, in our experiments we found that higher molecular weight complexes could be formed, but that these were labile during electrophoresis, and could not be well resolved in standard conditions used for lower molecular weight complexes. On 4.5 % polyacrylamide gels and at lower voltage we could resolve these complexes, and found that they were promoted by the presence of ATP or ADP (Figure 3B). These complexes could correspond to polymerization of RADA on the ssDNA, forming nucleofilaments. We did not see significant differences in binding with the K201A mutant protein, which seemed to bind to ssDNA with equivalent affinity as the WT protein.

The branch migration activity of recombinant plant RADA was also tested, in a well-described *in vitro* strand-invasion reaction promoted by bacterial RecA. With this system it has been shown that, after strand invasion catalyzed by RecA, bacterial RadA accelerates strand displacement and the resolution of branched intermediates (Cooper et Lovett, 2016). As described, we could show that RecA in the presence of ATP and bacterial SSB protein can initiate the invasion of dsDNA by homologous ssDNA, and promote branch migration till the final heteroduplex product molecule is completed. In the presence of plant RADA the branched intermediates were resolved faster, leading to an earlier appearance of the final product (Figure 3C). This faster resolution of recombination intermediates was reproducibly observed in five independent experiments (Figure 3D). However, in the presence of RADA there was a decrease in the amount of final product obtained, suggesting that at the same time that promotes branch migration, RADA also inhibits the RecA-dependent strand invasion (Figure 3C-D).

To test if RADA-dependent branch migration requires its interaction with RecA, an experiment was arrested by freezing at a time point when there was already accumulation of branched intermediates, but no visible final heteroduplex product (Figure 3E left panel). The reaction mix was deproteinated by phenol-chloroform extraction and nucleic acids purified by precipitation, before added to a new reaction mix. In the absence of RecA and RADA the branched intermediates could not spontaneously evolve and remained stable (Figure 3E middle panel), but in the presence of RADA they were rapidly converted, showing that RADA is able alone to promote branch migration (Figure 3E right panel). However, in the absence of RecA, RADA alone is not able to initiate the strand-invasion reaction, showing that it is not a recombinase redundant with RecA (Figure 3F). Finally, in the same conditions the K201A mutant protein is unable to promote branch migration, and rather completely inhibited the





**Figure 4.** Arabidopsis *radA* mutants and phenotypes. **(A)** Schematic representation of the Arabidopsis *RADA* gene. Coding sequences are in black and 5'- and 3'-UTRs are in gray. The position of the T-DNA insertions in *radA* mutant lines is shown. **(B)** *radA* plants show severe growth retardation, with distorted leaves presenting chlorotic sectors. These can be complemented by expression of HA-tagged *RADA* (*RADA:HA*) under control of the endogenous promoter. **(C)** root shortening of *radA* plants. **(D)** Detail of leaves phenotype. Scan microscope images show that in *radA* leaves epidermal cells are much larger and there are fewer stomata, half the number for the same leaf surface. Size bar is 20  $\mu\text{m}$  **(E)** Perennial vegetative growth of plants grown on short days (8h light), with development of aerial rosettes (4-month old *radA-1* plant).



reaction (Figure 3G). That could be because of competing with RecA to the binding of DNA, and its incapacity to migrate along the heteroduplexes, an activity dependent on ATP hydrolysis.

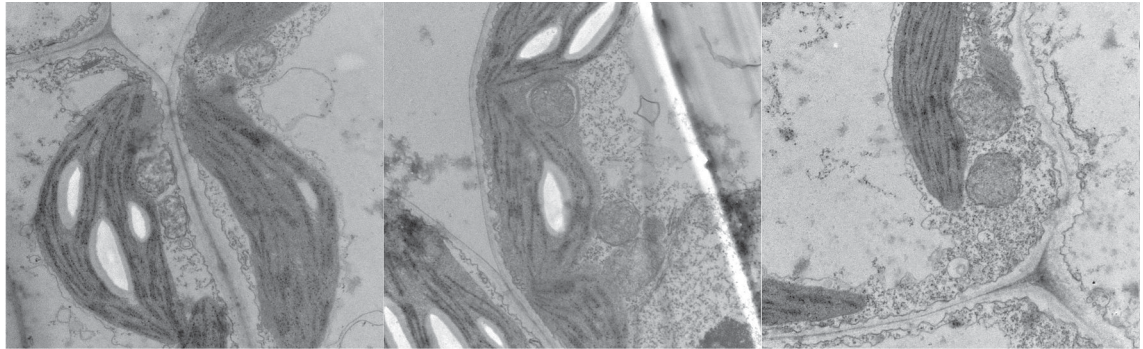
### **Arabidopsis plants deficient in RADA are severely affected in their development**

A search of the available Arabidopsis T-DNA insertion collections allowed us to retrieve several potential *radA* loss of function (KO) lines, from which two (*radA-1* and *radA-2*) could be confirmed by genotyping, with T-DNA insertion in exons 8 and 5 respectively (Figure 4A). Both lines are in the background of accession Col-0. As compared to WT, homozygous plants from both mutant lines displayed severe phenotypes of retarded growth of both leaves and roots, and of distorted variegated leaves showing chlorotic sectors (Figure 4B-D). The phenotype was fully penetrant, with individual plants being more or less affected. Observation of the leaves surface by scanning electron microscopy (SEM) showed that *radA* epidermal leaf cells are much enlarged as compared to cells of WT leaves (Figure 4D). The *radA* leaves also had a significantly reduced number of stomates ( $779.\text{mm}^{-2}$  versus  $384.\text{mm}^{-2}$  in WT and *radA-1* respectively,  $\chi^2$  test= $8.5\text{E-}14$ ). Transmission electron microscope observation showed that *radA* mesophyll cells are almost indistinguishable from WT, with chloroplasts that are morphologically normal. On the other hand, mitochondria looked enlarged in size, and less electron dense as compared to mitochondria from WT leaves, suggesting that it is mitochondria that is predominantly affected in *radA* (Figure S6). The severe dwarf phenotype could be partially relieved by growing *radA* plants under short day photoperiod (8h light/16h dark, as compared to the standard 16h light/8h dark growth conditions). In this conditions plants revealed extended juvenility, with dramatic elongation of lifespan and development of aerial rosettes. Similar phenotypes of perennial growth under short days was described for Arabidopsis *msh1*, deficient in the homolog of bacterial mismatch repair protein MutS, and apparently because of a developmental reprogramming triggered by the chloroplast (Xu *et al.*, 2012).

To fully confirm that these phenotypes are because of a deficiency in RADA, *radA-1* plants were complemented with the WT RADA gene expressed under its own promoter. For that, the sequence was cloned in binary vector pGWB13, fused to a C-terminal 3xHA tag, and used to transform heterozygous *radA-1* plants. Expression of the transgene in the T1 transformants was monitored by western-blot with a HA-specific antibody. In the T2 generation, homozygous *radA-1* plants that also contained the transgene were phenotypically normal (Figure 4B), confirming complementation and the linkage of the growth defects to the *radA* allele.

### **Fertility defects of *radA* plants**

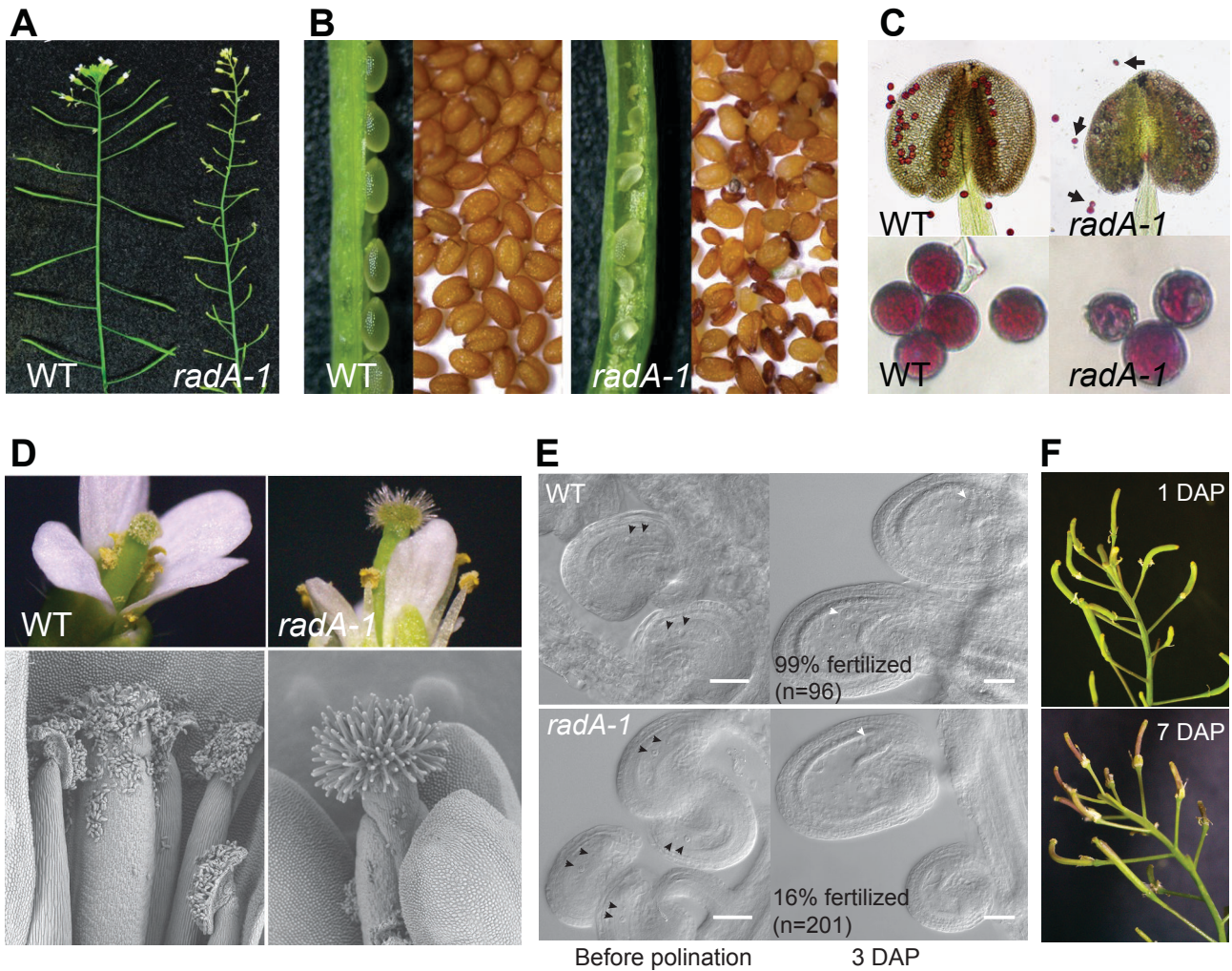
WT Col-0



radA-1



**Figure S6.** Transmission electron microscope (TEM) images of WT col-0 and *radA-1*. TEM images of cells from leaves of same size showed morphologically normal chloroplasts in *radA*. Mitochondria were enlarged and less electron dense as compared to mitochondria from WT cells.



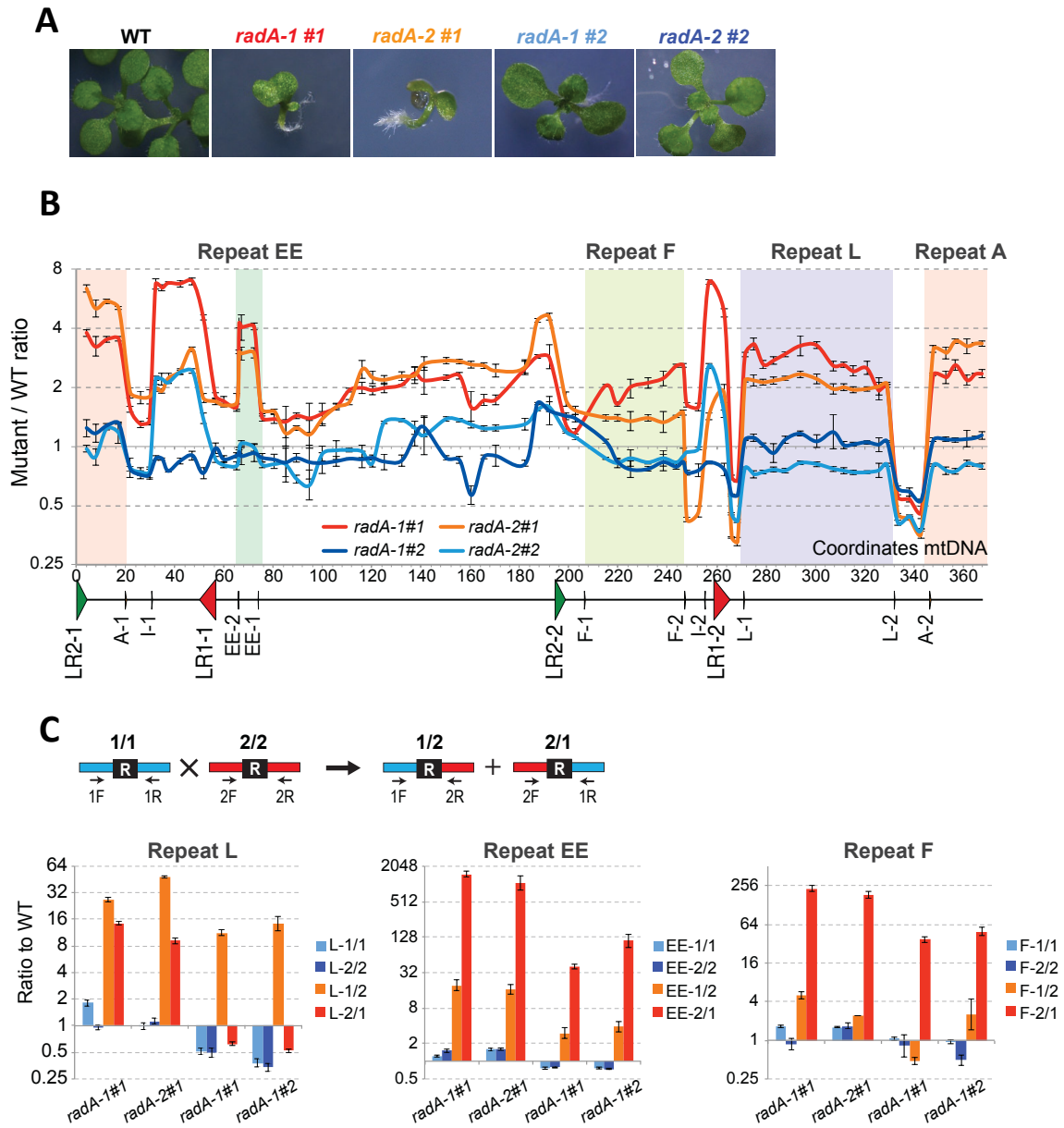
**Figure 5.** Reduced fertility of *radA* plants. **(A)** Comparison of WT and *radA* flower stems showing very small *radA* siliques. **(B)** *radA* siliques have mostly aborted seeds, and the few seeds produced are mostly non-viable. **(C)** Alexander staining of pollen in *radA* anthers as compared to WT, showing little pollen production, and an abundance of small and aberrant pollen grains (indicated by arrows). **(D)** Visible and SEM Images showing that no pollen binds to the papillae in *radA* stigma. **(E)** Differential interference contrast images of ovules on crosses between *radA-1* flowers and WT pollen. Black arrowheads indicate central cell and egg cell nuclei in unfertilized ovules. White arrowheads indicate developing embryos in fertilized ovules, three days after pollination (DAP). Only 16% of *radA* ovules could be fertilized. Size bar is 50  $\mu\text{m}$ . **(F)** After seven DAP the pollinated pistils of *radA* did not develop further.

*Arabidopsis radA* can produce flower stems, but most with very small siliques that mainly contain aborted seeds. The few seeds that are produced are heterogeneous in shape and most do not germinate (Figure 5A and 5B). To better understand the reasons of *radA* reduced fertility we observed both male and female organs. Anthers of *radA* produced pollen that stained positive by Alexander staining, suggesting that it is viable. But microscope inspection of anthers showed that pollen production was reduced compared to WT, and that there was a significant proportion of pollen grains of aberrant size and shape (Figure 5C). Thus, one reason for the partial sterility is low production and low quality of pollen. Regarding female organs, the stigma of *radA* flowers displayed elongated and fully differentiated papillae, but with no attached pollen grains (Figure 5D). Papillae are specialized cells that secrete a polysaccharide-rich extracellular matrix optimal for the germination of pollen. Therefore the absence of attached pollen suggests that either the papillae cells are modified and unable to bind pollen, or that stigma grows faster than pollen matures in the anthers. To test if *radA* female gametes are viable, we backcrossed *radA* plants using WT pollen, and observed ovules before and after pollination. Flowers from both WT and *radA* were emasculated, and 24h later the mature unfertilized ovules were observed on dissected pistils. The *radA* ovules looked morphologically normal and alike WT ovules, with visible and correctly positioned nuclei of the central and egg cells (Figure 5E left panel). Three days after pollination, virtually all WT ovules were fertilized and showed a developing embryo, at the 2 up to 8 cells globular stage. But on *radA* pistils only 1/6 of the ovules had developing embryos (Figure 5E right panel). The remaining unfertilized ones either looked normal or were degenerated, with unidentifiable cell types. This suggested that normal pollen could not fertilize the apparently normal *radA* ovules, apparently because a deficiency in pollen germination in the stigma of *radA* plants. The partial sterility of *radA* would therefore be because of both male and female defects.

### ***radA* mutants are mainly affected in the stability of the mitochondrial genome**

Several *Arabidopsis* mutants affected in recombination functions (ex: *msh1*, *osb1*, *recA2*, *recA3*, *recG1*) showed, in normal growth conditions, increased ectopic recombination of the mtDNA across IRs (Gualberto et Newton, 2017). These can lead to changes in the stoichiometry of the different regions of the genome. The *radA* mutants were thus also tested for such molecular phenotype. The severity of the molecular phenotype of increased ectopic recombination normally correlates with the severity of the developmental phenotypes (Zaegel *et al.*, 2006; Arrieta-Montiel *et al.*, 2009). Plants from both *radA-1* and *radA-2* lines of the second homozygous generation were grown *in vitro* and four of them were further analyzed individually, selected according to the severity of the visible growth defect phenotype (Figure 6A). On these plants the relative copy number of the different mtDNA regions was quantified

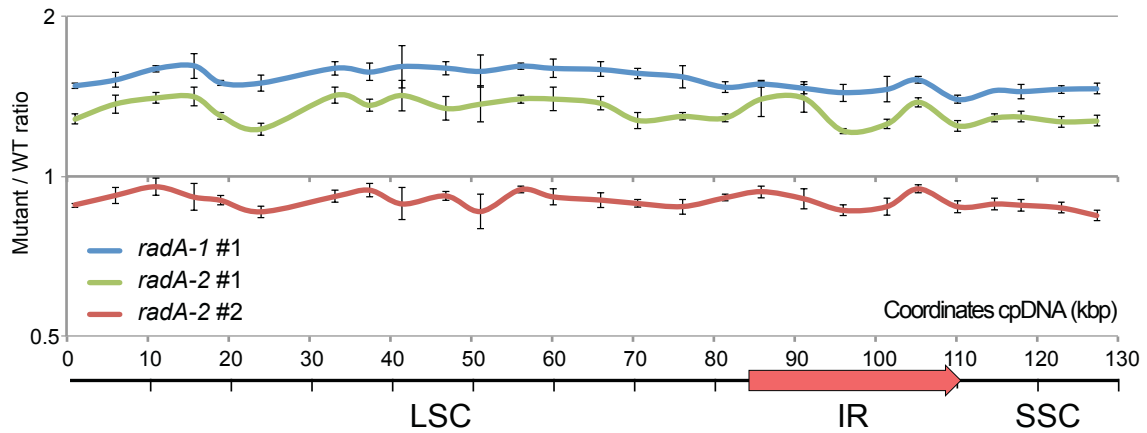




by qPCR, using a set of primer pairs spaced about 5 kb apart across the genome, as described (Wallet *et al.*, 2015). The recently corrected Arabidopsis Col-0 mtDNA sequence was taken as reference (Sloan *et al.*, 2018). Dramatic changes in the stoichiometry of mtDNA sequences were observed in all plants (Figure 6B). These affected similar regions of the genome in the individual plants, but the amplitude of the changes was higher in the severely affected ones (*radA-1#1* and *radA-2#1*) than in the mildly affected ones (*radA-1#2* and *radA-2#2*).

An increase in copy number of large regions could be observed that could be as high as 7 fold. Several of the observed events of stoichiometry variation corresponded to regions flanked by pairs of directly oriented repeats, including the pair of repeats A (556 bp), F (350 bp), L (249 bp), and EE (127 bp) (Figure 6B). This suggested that the process at play is the same described for repeat EE in the *RECG1* KO mutant: the looping out of a subgenome by recombination across directly oriented IRs, followed by their autonomous replication (Wallet *et al.*, 2015). We tested by qPCR the accumulation of corresponding crossover products for repeats F, L and EE, as previously described (Miller-Messmer *et al.*, 2012; Wallet *et al.*, 2015). Analysis of recombination involving repeats A was not possible, because the size of the region to be amplified (larger than 600 bp) was not compatible with qPCR. As expected, a significant increase in crossover products as compared to WT levels was observed for all repeats, in all plants (Figure 6C), with a significant higher accumulation in more affected *radA-1#1* and *radA-2#1* than in the mildly affected *radA-1#2* and *radA-2#2*. Recombination resulted in the asymmetrical accumulation of mainly one of the two products, with the remarkable difference of recombination involving pair of repeats L, which in mildly affected plants resulted in the accumulation of only one the reciprocal crossover products, while both products accumulated in the severely affected plants. Big differences in the relative accumulation of recombination products are seen, depending on the pair of repeats analyzed. But these values are misleading, because they are compared to the basal levels that exist in WT plants. Thus, a 30 fold increase of recombination product L-1/2 is probably equivalent to a 1000 fold increase of product EE-2/1, because the former is already quite abundant in WT Col-0 and easily detected by hybridization, while the latter is virtually absent (Zaegel *et al.*, 2006; Wallet *et al.*, 2015).

Because RADA is apparently targeted to both organelles the cpDNA of *radA* plants was likewise scanned for changes in sequence stoichiometry. No changes were detected between mutant and WT, apart from a slight general increase in cpDNA copy number in several plants, but below two fold and probably not significant (Figure S7). Thus, in contrast to the major effects observed on mtDNA maintenance the loss of RADA apparently does not significantly affect the stoichiometric replication and segregation of the cpDNA in Arabidopsis.



**Figure S7.** No apparent effect of RADA deficiency on cpDNA stability. Scanning of the cpDNA of *radA* mutants for changes in relative copy numbers of the different cpDNA regions. Sequences spaced 5-10 kb apart on the cpDNA were quantified by qPCR, as described for the mtDNA in Figure 6. Coordinates correspond to the ones of the published Col-0 cpDNA sequence. The three major regions of the cpDNA are indicated: LSC, large single copy region; SSC, small single-copy region; IR, inverted repeat.

## The *recG1* and *radA* mutations are not additive

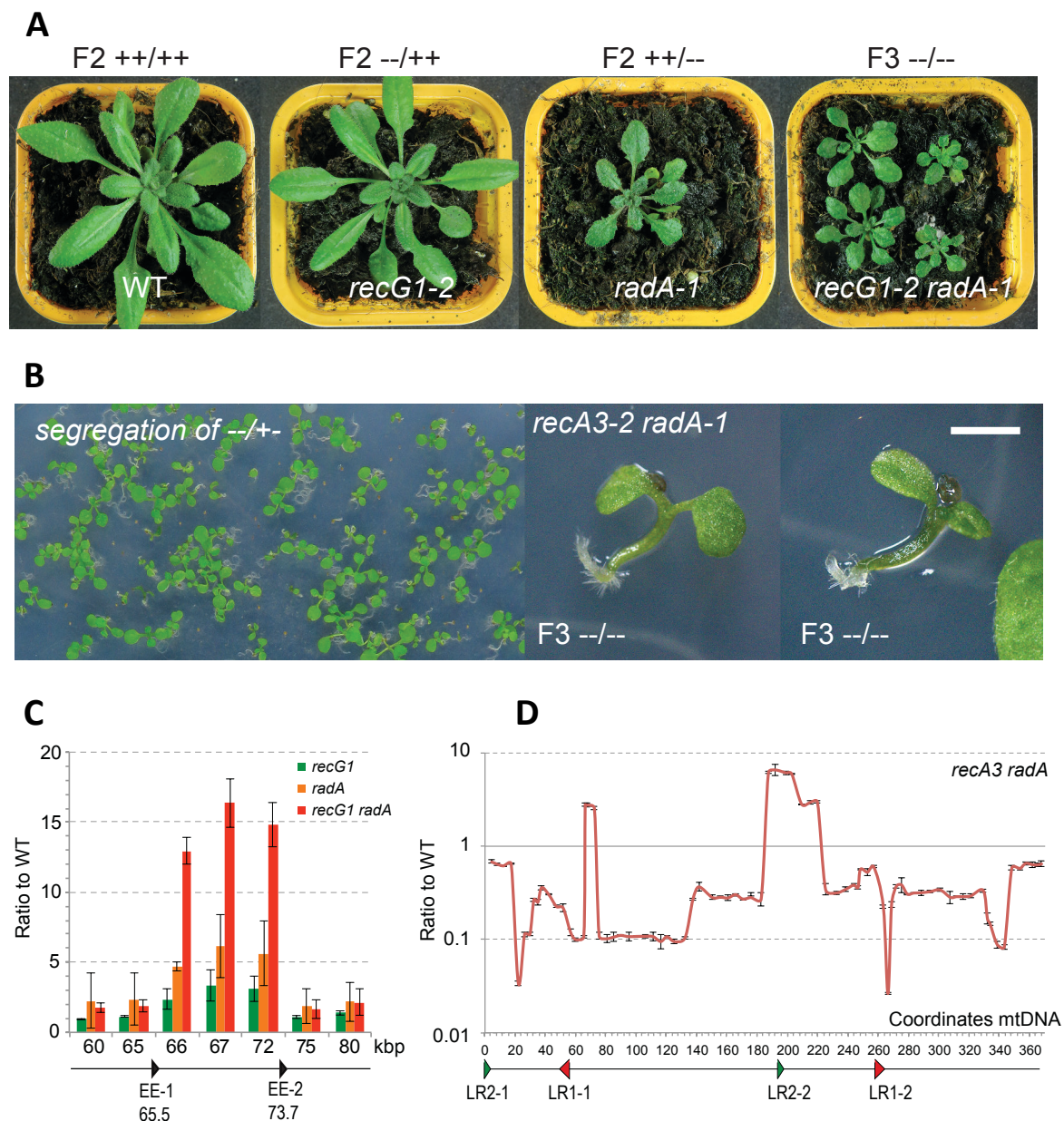
We tested the epistatic relationship between *RADA* and *RECG1*, on a *recG1 radA* double mutant. To compare all mutant plants at the first homozygous generation we followed the following cross protocol: heterozygous *recG1-2* plants (KO for *RECG1*, Wallet et al. 2015), were crossed with *radA-1* heterozygous plants used as pollen donor. Double heterozygous (*RECG1+/-/RADA+/-*) F1 plants were selected, and in the segregating F2 generation we obtained WT, *recG1-2* and *radA-1* single homozygous mutants, and the *recG1-2 radA-1* double homozygous mutants. In this cross plants inherit the mtDNA of accession Ws, which is the genetic background of *recG1-2*. As described before, the *recG1-2* single mutants were similar to WT (Wallet et al., 2015). The *radA-1* single mutants developed the same growth defect phenotypes observed in the Col-0 background. And regarding the *recG1-2 radA-1* double mutants, they were as severely affected in growth as *radA-1* plants (Figure 7A). Thus, the *recG1* and *radA* mutations are not additive, and *RECG1* apparently contributes to a lesser extent to organellar genome maintenance than does *RADA*.

## *RADA* works synergistically with *RECA3* in the maintenance of mtDNA stability

It was previously shown that *recG1* is synergistic with mutants deficient in *RECA3* (Wallet et al., 2015), which codes a RecA ortholog that contrarily to *RECA2* is dispensable (Miller-Messmer et al., 2012). Therefore we also tested the epistatic relationship between *RADA* and *RECA3*. The *recA3-2* and *radA-1* mutants (both in Col-0) were crossed, using heterozygous plants as female and pollen donor, but no double homozygous mutants could be retrieved from F2 plants growing on soil. Seeds from plants heterozygous for *radA-1* and homozygous for *recA3-2* (*RECA3--/RADA+/-*) were then germinated *in vitro* in agar plates with sugar, and seedlings were genotyped. It showed that double homozygous *recA3 radA1* plants could germinate, but were unable to grow roots and to further expand their cotyledons (Figure 7B). This seedling lethal phenotype is similar to the one of *recA2* (Miller-Messmer et al., 2012). The additive effect of the *radA* and *recA3* mutations suggest that the two factors intervene in parallel pathways, and that if both are compromised it leads to a defect equivalent to the loss of the main *RECA2*-dependant recombination pathway.

The double mutants were tested for their cumulative effects on mtDNA stability. In the genetic background of the Ws mtDNA, the effect of the *recG1* KO mutation was mainly in the accumulation of the episome resulting from the recombination between the pair of direct repeats EE (Wallet et al., 2015). We tested its accumulation in *recG1-2*, *radA-1* and in the double mutant *recG1-2 radA-1*. In the first homozygous mutant generation we found a mild effect of *recG1-2* on the accumulation of the EE episome, 3 fold more abundant than the flanking mtDNA





**Figure 7.** Synergistic effects of the *radA* mutation on *recG1* and *recA3*. **(A)** Crosses of *recG1-2* and *radA-1* (pollen donor) show that double homozygote *recG1 radA* plants are as affected as simple homozygote *radA* plants. **(B)** Crosses and segregation of *recA3-2* and *radA-1* (pollen donor) show that double homozygote *recA3 radA* seedlings do not grow roots and do not grow further. The phenotype is similar to the one observed for *recA2* mutants (Miller-Messmer et al. 2012). The size bar is 1 mm. **(C)** qPCR analysis of the copy number of mtDNA sequences around and within the region comprised between pair of repeats EE, previously shown to generate an episome by recombination in *recG1* plants (Wallet et al. 2015). Autonomous replication of the episome is significantly increased in the *recG1 radA* double mutant. Results show the mean and SD error bars from 2 or 3 (for *recG1 radA*) biological replicates. **(D)** Scanning of the copy numbers of the different mtDNA regions in *recA3 radA*, as described in Figure 6, showing severe reduction of several mtDNA regions. Results are in a log scale.

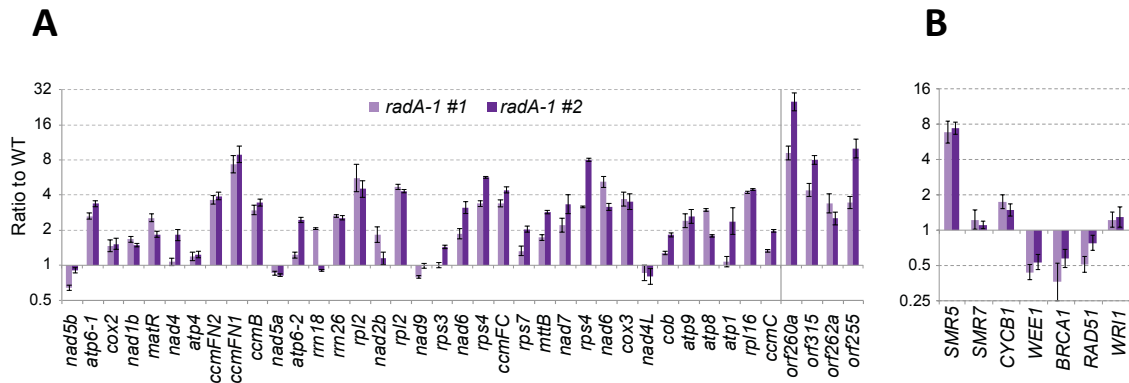
regions. The *radA-1* mutation had an equivalent effect, of an increase of about 3.3 fold as compared to the copy number of the flanking regions. However, in *recG1-2 radA-1* there was a significant increase in episome accumulation, of about 10 fold bigger copy number than flanking regions. Thus, at the molecular level the double mutation has an increased impact in the correct segregation of mtDNA sequences, although without additive effects on plant development.

Regarding *recA3 radA1*, the analysis of the mtDNA of a pool of double mutant seedlings showed that the mtDNA was massively affected in their stability, with an overall reduction in copy number as compared to seedlings of the same size, and with regions more than 20 fold reduced or increased as compared to neighboring sequences. Thus, the seedling lethality can be explained by the problems in the replication and segregation of the mitogenome.

## Transcriptome analysis

In order to understand the why of the *radA* growth defect we quantified by RT-qPCR the relative abundance of almost all mtDNA protein gene transcripts and rRNAs. Surprisingly, we found no apparent defects in gene expression, and for most transcripts rather an increase as compared to WT plants. Because reshuffling of the mtDNA sequences by recombination can change the expression of non-coding orfs and even lead to the expression of new transcription units we tested also the expression of a few orfs whose transcription could be affected by recombination across IRs. These were orf260a that by recombination involving repeat A is part of a chimera with *rps3*, orf315 that by recombination involving repeats G is part of a chimera with *atp9*, orf262a and orf255 that might be transcriptionally activated by recombination involving repeats E, L or H. Indeed we found an overall increase in the transcription of these orfs, in particular of orf260a. But given the overall increase in transcription of most genes it is difficult to conclude that the expression of toxic orfs is the reason of the *radA* growth defects.

The observation that *radA* leaves have enlarged epidermal cells implied of defects in cell cycle progression. It is known that cell cycle checkpoints adjust cellular proliferation to changing growth conditions, arresting it by inhibiting the main cell cycle controllers that include cyclin-dependent kinases (CDKs) (Nowack *et al.*, 2012). The arrest of cell cycle is also one of the effects of the DNA damage response (DDR), controlled by the ATM and ATR signaling cascades (Abraham, 2001). ATM phosphorylation of the transcription factor SOG1 activates the expression of the SIAMESE-RELATED (SMR) group of proteins (Yoshiyama *et al.*, 2013; Yoshiyama, 2016). In plants SMR genes are also transcriptionally induced in response to changing conditions, integrating environmental and metabolic signals with cell cycle control (Yi *et al.*, 2014; Dubois *et al.*, 2018). It has also been shown that a deficiency in chloroplast function



**Figure 8.** Accumulation of mitochondrial and cell cycle transcripts in *radA*. **(A)** Representative plastidial transcripts were quantified by RT-qPCR and normalized against a set of nuclear housekeeping genes. Quantification of transcripts for several orfs whose transcription can be affected by ectopic recombination involving IRs are shown on the right. Two pools of *radA-1* seedlings were analyzed. Results are on a log2 scale **(B)** Quantification of transcripts of selected key cell cycle regulators involved in nuclear DNA damage response (DDR) activated as response to stress and of nuclear DNA repair factors.

could also trigger *SMR5* and *SMR7* dependent arrest of cell growth and concomitant endoreplication of the nuclear genome (Hudik *et al.*, 2014). We have therefore tested the expression of several cell cycle markers (*SMR5* and *SMR7*, the Cyclin-dependent protein kinase *CYCB1* and the *WEE1* kinase homolog) and of several genes involved in DNA repair in the nucleus that are transcriptionally activated as part of the DDR (*BRCA1*, *RAD51* and transcription factor *WRI1*). We found no induction of the nuclear DNA repair factors. But in *radA* there is strong transcription activation of *SMR5*, and a slight increase in *CYCB1* transcripts. Thus, a component in the *radA* phenotype is a retrograde signal that activates cell cycle regulators that inhibit cellular proliferation.

## DISCUSSION

The role of RadA in bacterial homologous recombination has only very recently been described (Cooper et Lovett, 2016; Marie *et al.*, 2017). RadA is a hexameric helicase loaded by RecA on either side of the D-loop, to allow hybridization of the invading ssDNA with the recipient DNA (Marie *et al.*, 2017). Our results indicate similar functions of plant RADA in plant organelles, but a more essential role than the one of RadA in bacteria.

RADA is a RecA paralog, found in all bacteria, algae and land plants. The protein sequence of RADA is well conserved between the various organisms, in particular those of the Walker A, Walker B and KNRF motifs, as well as the N-terminal zinc finger domain. The modeled structure of RADA is very similar to that of RadA of *S. pneumoniae*, suggesting that both proteins have the same activities. This conservation of activity between plant and bacterial proteins is supported by complementation results and by *in vitro* activity tests. Arabidopsis RADA complements the survival of the bacterial *radA* mutant under genotoxic conditions, as well as the *E. coli* protein when brought *in trans*. Similarly, plant RADA preferentially binds ssDNA and accelerates the *in vitro* strand-exchange reaction initiated by RecA, as described for bacterial RadA.

Nevertheless, differences between the activities of plant RADA and bacterial RadA were observed. RADA binds ssDNA-containing molecules without the need for ADP or ATP. But in the presence of ADP or ATP it forms higher molecular weight complexes with ssDNA, suggesting a higher degree of protein oligomerization. However, in bacteria, the presence of ATP activates the translocation of RadA and causes the dissociation of the protein from DNA (Marie *et al.*, 2017 214). Only the mutant of the ATPase domain remains associated with DNA in the presence of ATP, because it is unable to activate translocation (Marie *et al.*, 2017 214). But the Arabidopsis Walker A K201A mutant could also bind to ssDNA and form high molecular weight complexes in the presence of ATP or ADP, just like wild-type RADA. Thus, it is possible that under our test conditions the ATPase activity of RADA is inactive and prevents its translocation

on DNA. Since the same protein could accelerate *in vitro* strand exchange, it should have translocase activity. It is therefore possible that the interaction of RADA with RecA is necessary to activate its ATPase domain. Thus, the conditions of interaction of RadA (or RADA) with DNA remain controversial: Marie et al. (2017) observed the binding of RadA to DNA in the absence of ATP, and translocation in the presence of ATP; Cooper and Lovett (2016) observed RadA binding to DNA only in the presence of ADP; and we observed binding of RADA to DNA, regardless of the presence of ADP or ATP.

The acceleration of *in vitro* strand-exchange by RADA is accompanied by a lesser accumulation of heteroduplex final products. RADA promotes the processing of already formed recombination intermediates, but seems to prevent the formation of new intermediates by RecA. This implies that RADA not only plays a role in branch migration, but might also regulate the recombinase activity of RecA-like organellar proteins. This dual function of RADA would guide homologous recombination towards the late stages of the process. The destabilization of RecA filaments by RadA has already been proposed for the bacterial proteins, but was not confirmed (Cooper, 2016 # 470). Since RadA is the only branch migration factor known that can work in the context of the RecA filament, it seems likely that it acts early in the process of branch migration, before RecG or RuvAB. Also, the destabilization of RecA by RadA makes sense, in order to promote the action of the other branch migration factors, inhibited by RecA (Whitby *et al.*, 1993 511).

Since the absence of RadA is not particularly harmful for the development of bacteria, other mechanisms that destabilize RecA must nevertheless exist. The inhibition of RecA activity by RADA might be due to DNA-binding overcrowding, with RADA competing with RecA for the formation of a nucleofilament. As the reaction progresses, the proportion of RADA to ssDNA would increase, leading to the sequestration of remaining ssDNA molecules. It is also possible that the interaction between RADA and RecA titrates the latter, preventing its activity. Finally, since the RecA protein used in our *in vitro* tests is not the homologous organellar recombinases, it is likely that its activity differs from the ones of RECA2 or RECA3, and may not be fully compatible with RADA. Ideally, the *in vitro* strand invasion experiments should have been done with the organellar RecA-like proteins, but we were unable to express RECA2 or RECA3 in an active form (unpublished results of the laboratory). In addition, it would also be necessary to purify a homologous ssDNA-binding protein required for the reaction (Cooper et Lovett, 2016 470). Plant organelles, however, possess several SSB proteins (SSB1 and SSB2), as well as other proteins with high affinity for ssDNA (OSB1-4, ODB1-2, WHY1-3) whose respective individual roles and partner are still unclear (Gualberto et Newton, 2017 315). It will therefore be difficult to reconstitute *in vitro* the recombination reaction using the homologous organellar factors.

In bacteria, the mutation of a single branch migration factor is only slightly deleterious for cell growth and DNA repair, and that is particular particularly true for RadA. In plants

however, the loss of RADA severely affects plant development and fertility. This result also contrasts with the lack of a notable developmental phenotype observed for the *Arabidopsis recG1* mutants (Wallet *et al.*, 2015 417). It therefore seems that RADA is a more important factor than RECG1 in organellar homologous recombination. Since the RuvAB migration pathway is absent in plants, it is possible that homologous recombination of organelles has evolved by favoring the RADA pathway. Surprisingly, the double mutant *radA recG1* is no more affected in its development than the *radA* single mutant, while in bacteria the mutation of the different branch migration pathways is particularly synergistic (Beam *et al.*, 2002 471; Cooper *et al.*, 2015 469). In bacteria, the double mutant *radA recG* is more severely affected than the RecA single mutant, indicating that the accumulation of unprocessed branched intermediates is more detrimental to the cell than the lack of recombination. In agreement with this hypothesis, the triple mutant *recA radA recG* is less affected than *radA recG*. However, in plant organelles, the lack of homologous recombination in the *recA2* mutant is lethal at the seedling stage, suggesting a much more important role of recombination in organelles than in bacteria (Miller-Messmer *et al.*, 2012 361). It is therefore surprising that the mutation of all known branch migration pathways in plants is no more deleterious than the mutation of the recombinase.

It is possible that the processing of the recombination intermediates is not as detrimental in plant organelles as in bacteria. But it is also possible that an additional way of processing recombination intermediates exists in plant organelles. The synergy of the *radA* and *recA3* mutations recalls that observed in the double mutants of bacterial branch migration factors (Cooper *et al.*, 2015 469). Thus, it is possible that RECA3 is involved in the processing of the recombination intermediates created by the RECA2 recombinase. Like RADA, RECA3 is a RecA paralog, whose evolution would have favored its function of extending the D-loop using a RecA-like mechanism. While prokaryotes have only one RecA paralog (RadA), eukaryotes have several RAD51 paralogs, involved in the regulation of recombinase functions (Adelman et Boulton, 2010 525; Qing *et al.*, 2011 526). As example, RAD51C and XRCC3, two RAD51 paralogs forming the CX3 complex, act in the late stages of nuclear recombination (Chun *et al.*, 2013 527). XRCC3 is necessary for the stabilization of heteroduplexes and controls the extent of gene conversion, roles reminiscent of branch migration factors (Brenneman *et al.*, 2002 528). RADA and RECA3 could thus be partially redundant. Nonetheless, unlike RADA, RECA3 apparently retained strand invasion activity, and could partially complement the bacterial mutant *recA* in the repair of UV-induced lesions (Miller-Messmer *et al.*, 2012 361). RECA3 would therefore also be partially redundant with RECA2. Whereas RECA2 and RADA would have specialized in homologous sequence recognition and invasion and branch migration, respectively, RECA3 would be more versatile and have retained both functions.

It is also conceivable that RECA3 acts in an alternative recombination pathway independent of RECA2 and RADA. In the absence of RADA the RECA2 pathway would be overloaded with unprocessed intermediates and would require activation of the alternative RECA3-dependent recombination pathway. The loss of both RADA and RECA3 would lead to overaccumulation of unresolved recombination intermediates, which would then be lethal for the plant. A co-immunoprecipitation approach to identify the interacting partners of RECA2, RECA3 and RADA might discriminate between these different hypotheses. For that, the *radA* mutant line complemented with the tagged version of RADA has already been produced.

The absence of RADA results in a significant increase in mtDNA ectopic recombination, which leads to drastic changes in the stoichiometry of the genome. These changes are mostly because of the formation of episomes that replicate more rapidly and uncoordinated from the rest of the mtDNA, as it was observed in the *recG1* mutant for the episome resulting from recombination involving repeats EE (Wallet *et al.*, 2015 417). The more the stoichiometry of the mtDNA is modified, the more the development of the plant is affected. However, no region of the mtDNA containing functional genes is lost, and all mitochondrial genes that were tested are expressed (the tRNA genes have not been tested, however). Thus, we could not correlate the developmental phenotypes to the reduced expression of a subunit from a complex of the respiratory chain, or to a factor required for their translation and assembly. Nevertheless, the increase in ectopic recombination can lead to the creation and expression of chimeric genes that could encode toxic proteins (Hanson et Bentolila, 2004 476; Touzet et Meyer, 2014 413). A complete transcriptome analysis of *radA* mutants by RNAseq is required to identify all expressed chimeric ORFs, and any other *radA* effects in the expression of the organellar and nuclear genes.

The absence of RADA does not seem to affect the chloroplast. The stoichiometry of the cpDNA is not modified in *radA*, just like for the *recG1* mutant (Wallet *et al.*, 2015 417). It is possible that RADA does not act in the chloroplast, or that the absence of IRs in the cpDNA of Arabidopsis limits the possibilities of ectopic recombination that could destabilize the genome. RECA2 is also addressed to the chloroplast, but it also does not seem to have any deleterious effects in the maintenance of the cpDNA (Shedge *et al.*, 2007 430; Miller-Messmer *et al.*, 2012 361). Interestingly, the severe developmental phenotypes triggered in *radA* could be because of a nuclear retrograde response, as the increase in epidermal cells size and reduction in stomates suggest. The observed increase in *SMR5* and *CYCB1* indeed suggest effects in the progression and control of the cell cycle.

Alterations of organellar functions can be perceived by the nucleus through retrograde signals leading to a modulation of plant development via the cell cycle (Hudik *et al.*, 2014 541). The cell cycle is an energy demanding process that can be arrested by problems in respiration or photosynthesis (Riou-Khamlichi *et al.*, 2000 545). The increase of *SMR5* and *CYCB* transcripts in

the *radA* mutant indicates that the cell cycle is affected by RADA deficiency. SMR5 (with SMR7) is an inhibitor of the cell cycle induced by oxidative, hydric or light stresses (Hudik *et al.*, 2014 541; Yi *et al.*, 2014 544; Dubois *et al.*, 2018 542). These stresses produce ROS that can be signals for retrograde responses by the nucleus (Mittler *et al.*, 2011 543). Inhibition of the cell cycle can lead to early differentiation of cells causing stunting (Hudik *et al.*, 2014 541). It is thus possible that mitochondrial genome instability in the *radA* mutant provokes sub-optimal function of the respiratory complexes, resulting in productions ROS that trigger genes from DNA-damage response and the arrest of the cell cycle, leading to the developmental phenotypes observed. A nuclear cause for developmental phenotypes has already been proposed for other mutants of other organellar recombinant factors, as MSH1 (Xu, 2012 # 524). This hypothesis can be tested by looking for the presence of ROS in the *radA* mutants and by analyzing the expression of other cell cycle markers such as CYCD, inhibited by SMR5, or of SMR1 that negatively affects CYCB (Dubois *et al.*, 2018 542). Thus, the increase in cell size observed in *radA* could be a consequence of the endoreplication caused by cell cycle arrest. Flow cytometry should be performed to determine the ploidy of *radA* cells.

## METHODS

### Plant Material

Arabidopsis T-DNA insertion mutant lines, all in the Col-0 background, were obtained from the Nottingham Arabidopsis Stock Centre (*radA-1*: SALK\_097880, *radA-2*: WiscDsLoxHs058\_03D). Plant genotypes were determined by PCR using gene and T-DNA specific primers. Seeds were stratified for 3 days at 4°C and plants were grown on soil or on 0.5x MS255 medium (Duchefa) supplemented with 1% (w/v) Sucrose, at 22°C.

DNA was extracted using the cetyltrimethylammonium bromide method. RNA was extracted using TRI Reagent (Molecular Research Centre, Inc.). For RT-qPCR experiments, 5 µg of RNA were depleted from contaminating DNA by treatment with RQ1 RNase-free DNase (Promega) and were reverse-transcribed with Superscript III Reverse Transcriptase (Life Technologies), according to the manufacturer's protocol using random hexamers.

### Bioinformatics analysis

Bacterial and plant sequences were identified in the databases by BLASTP and TBLASTN. Alignments were constructed with ClustalW implemented in the Macvector package using the Gonnet matrix. Phylogenetic trees were built with PhyML v3.1 ([www.phylogeny.fr](http://www.phylogeny.fr)) using the



neighbor-joining method implemented in the BioNJ program. Graphical representations were performed with TreeDyn (v198.3). The Arabidopsis RADA structure was modeled on the structure of RadA from *Streptococcus pneumoniae* (pdb: 5LKM), using Modeller ([http://salilab.org/modeller/about\\_modeller.html](http://salilab.org/modeller/about_modeller.html)).

### **Intracellular Localization**

The cDNA sequence coding the Arabidopsis RADA protein N-terminal domain (first 174 codons) was cloned into the pUCAP-GFP vector, upstream and in frame with eGFP. For constitutive expression in Arabidopsis, the expressing cassette was transferred to the binary vector pBIN+ (van Engelen et al., 1995) and Arabidopsis Col-0 plants were transformed by the floral dip method. Leaves of selected transformants were observed on a Zeiss LSM700 confocal microscope. The fluorescence of GFP and chlorophyll was observed at 505 to 540 nm and beyond 650 nm, respectively. For mitochondrial co-localization, leaves were infiltrated with a 1/1000 dilution of MitoTracker® orange (ThermoFisher) solution. Excitation was at 555 nm and observed at 560 to 615 nm.

### **In vitro strand exchange reaction**

Recombination assays were performed with single-strand linear  $\Phi$ X174 virion DNA and double strand circular  $\Phi$ X174 RFI DNA (New England Biolabs) linearized with PstI in 20 mM Tris-acetate pH 7.4, 12.5 mM phosphocreatine, 10 U/mL creatine kinase, 3 mM ammonium glutamate, 1 mM dithiothreitol, 2 % glycerol and 11 mM magnesium acetate. In our conditions, 20.1  $\mu$ M (in nucleotides) linear single strand DNA (ssDNA), 6.7  $\mu$ M RecA (New England Biolabs), 2  $\mu$ M RADA are incubated with buffer for 8 min at 37 °C. Then, 20.1  $\mu$ M (in nucleotides) linear double-strand DNA (dsDNA) is added and the whole reaction is incubated for 5 min at 37 °C. Finally, strand exchange is initiated by adding 3 mM ATP and 3.1  $\mu$ M SSB (Sigma-Aldrich). Aliquots are stopped at indicated times by addition of 12  $\mu$ M EDTA and 0.8 % SDS. Strand exchange products were analyzed on 0.8 % agarose gels run at 4 °C in Tris-acetate EDTA buffer (TAE) at 50 V and visualized after migration by ethidium bromide staining.

For reaction terminated in the absence of RecA, the RecA-initiated strand exchange reaction was stopped at the indicated time and DNA was deproteinized by phenol-chloroform extraction followed by ethanol precipitation. The DNA pellet was solubilized in reaction buffer, 2  $\mu$ M RADA was added and the reaction was further incubated at 37 °C for the indicated time, before quenching with 12  $\mu$ M EDTA and 0.8 % SDS.

## DNA Binding Assays

For electrophoretic mobility shift assays (EMSA) the purified recombinant protein (50-500 fmol according to the experience) was incubated with oligonucleotide probes (0.01 pmol-10 fmol) that were 5'-radiolabeled with [ $\gamma$ -<sup>32</sup>P]ATP (5000 Ci/mmol; Perkin-Elmer Life Science) and T4 polynucleotide kinase (Fermentas). Different dsDNA structures were prepared by annealing the radiolabeled sense oligonucleotides with a twofold excess of unlabeled complementary oligonucleotide and purified on non-denaturing polyacrylamide gels. The binding reactions were performed in 20 mM Tris-HCl pH 7.5, 50 mM KCl, 5 mM MgCl<sub>2</sub>, 0.5 mM EDTA, 10% glycerol, 1 mM DTT and protease inhibitors (Complete-EDTA; Roche Molecular Biochemicals), incubated at 20°C for 20 min and run on 8 or 4.5% polyacrylamide gels in Tris-Borate-EDTA buffer at 4°C. After drying gels were revealed using a Amersham Typhoon phosphorimager (GE Healthcare Life Sciences). For competition assays, labeled probe and unlabeled competitor were added simultaneously to the reaction mixture.

## Recombinant proteins

Constructs pET28-RADA and pET28A-RADA[K201A] were used to express recombinant proteins in *E. coli* Rosetta 2 (DE3) pLysS (Novagen). Transformed bacteria were grown at 37 °C until OD<sub>600nm</sub> = 0.6. Cultures were then chilled to 4 °C for 30 min before addition of 0.5 mM isopropyl b-D-1-thiogalactopyranoside (IPTG) and overnight incubation at 18 °C. After growth cells were pelleted, resuspended in 50 mM Tris-HCl pH 8.0, 5 % glycerol, 300 mM NaCl, 10 mM imidazole, supplemented with 1 mM PMSF and 1X cComplete protease inhibitors (SigmaAldrich) and lysed with a French press under 1200 PSI. The crude lysate was sonicated for 3 min, clarified by 25 min centrifugation at 17700 g and filtrated through a Filtropur S plus 0.2  $\mu$ m filter (STARDEST). The recombinant RADA and RADA[K201A] proteins were affinity purified in a precalibrated HisTrap FF Crude (GE Healthcare Life Sciences) column run at 0.5 mL/min, washed with 50 mM Tris-HCl pH 8.0, 300 mM NaCl, 5 % glycerol, 50 mM imidazole and eluted with a 50-500 mM imidazole gradient. The recombinant protein fractions was further purified by gel filtration on Superdex S200 columns and aliquots were flash frozen in liquid nitrogen and stored at -80 °C. RADA concentration was determined by spectrophotometry using an extinction coefficient of 42,440 M<sup>-1</sup> cm<sup>-1</sup>.

## Bacterial complementation:

The *E. coli* TOP10 strain was used for routine cloning, whereas the BW25113 *radA+* and JW4352 *radA785(del)::kan* were used for complementation assays. The Arabidopsis RADA

cDNA (sequence coding for amino acids 137 to 627) or the *E. coli* RadA/Sms sequence were cloned between the PstI and BamHI restriction sites of the pACYCLacZ vector (Miller-Messmer et al. 2012) under the control of the *lac* promoter. Both constructs were introduced in the JW4352 strain. The pACYCLacZ empty vector was introduced in the BW25113 and JW4352 strains as control. Bacteria were grown in LB supplemented with 10 µg/mL chloramphenicol till  $OD_{600nm} = 0.4$  before addition of 2.5 mM IPTG. At  $OD_{600nm} = 1.2$  bacteria were diluted  $10^4$  fold and grown on LB agar plates supplemented with 15 nM ciprofloxacin, 2 mM IPTG and 10 µg/mL chloramphenicol.

### qPCR Analysis

qPCR experiments were performed in a LightCycler480 (Roche) in a total volume of 6 µL containing 0.5 mM of each specific primer and 3µl of SYBR Green I Master Mix (Roche Applied Science). The second derivative maximum method was used to determine Cp values and PCR efficiencies were determined using LinRegPCR software (<http://LinRegPCR.nl>). Three technical replicates were performed for each experiment. Results of qPCR and RT-qPCR analysis were standardized as previously described (Wallet et al., 2015). Quantification of mtDNA and cpDNA copy numbers used a set of primer pairs located along the organellar genomes, as described previously (Wallet et al. 2015; Le Ret et al. 2018). Results were normalized against the UBQ10 (At4G05320) and ACT1 (At2G37620) nuclear genes. The accumulation of ectopic recombination in mtDNA was quantified using primers flanking each repeats, as described (Miller-Messmer et al. 2012). The COX2 (AtMG00160) and 18S rRNA (AtMG01390) mitochondrial genes and the 16S rRNA (AtCG00920) chloroplast gene were used for normalization. For RT-qPCR experiments the GAPDH (At1G13440) and ACT2 (At3G18780) transcripts were used as standards.

## FIGURE LEGENDS

**Figure 1.** Arabidopsis RADA is targeted to both chloroplasts and mitochondria. The RADA:GFP fusion protein was constitutively expressed in transgenic plants and its location observed in leaf epidermal cells. It co-localized with both chloroplasts (autofluorescence of chlorophyll) mitochondria (red fluorescence of MitoTracker).

**Figure 2.** The Arabidopsis RADA is structurally and functionally homologous to bacterial RadA. (A) The modular structure of plant RADA is similar to the one from bacteria, with a N-terminal zinc-finger (C4), an helicase and Lon-protease-like domains (H and P domains respectively). Plant precursor proteins have an N-terminal extension containing an organellar

targeting sequence (OTS). (B) Model of Arabidopsis RADA superposed on the known structure of *S. pneumoniae* (in gray) (Marie et al., 2017). The color code of relevant domains is as in A). A bound ADP and the Mg<sup>2+</sup> ion (yellow ball) are shown. (C) Complementation of a *E. coli radA* mutant for growth in the presence of genotoxic ciprofloxacin. The Arabidopsis RADA gene (*At-RADA*) is as efficient as the bacterial gene (*Ec-RadA*) for complementation

**Figure 3.** DNA-binding and branch-migration activities of RADA. (A) EMSA experiments showing that RADA binds any ssDNA-containing DNA structure with higher affinity than dsDNA. Lower and higher molecular weight complexes are indicated by white and black arrowheads respectively (B) Analysis on low concentration gel (4.5% as compared to 8% in A) of the formation of a high-molecular weight RADA filament on ssDNA, which is promoted by ATP or ADP (1 mM). The K201A mutant protein binds with equivalent affinity as the WT protein. Increasing concentrations of RADA used in A and B are indicated by the grey triangles. (C) In a *in vitro* strand-invasion reaction plant RADA accelerates branch-migration of DNA heteroduplexes initiated by RecA. An explanation of the different substrates and products is shown below the gel. (D) Ratio of final product as compared to the initial linear dsDNA substrate in 5 independent experiments, showing that in the presence of RADA there is faster resolution of branched intermediates. However RADA inhibits initiation, as fewer final products are produced. (E) RADA can alone finalize branch-migration initiated by RecA: a reaction at T=7 min was arrested by deproteination (left panel) and the DNA purified. Without addition of RecA or RADA there is no spontaneous progression of the reaction (middle panel), but RADA alone can resolve intermediates into the final product (right panel). (F) RADA alone cannot initiate strand invasion. (G) Mutation of the ATPase Walker domain of RADA (K201A) inhibits the reaction.

**Figure 4.** Arabidopsis *radA* mutants and phenotypes. (A) Schematic representation of the Arabidopsis *RADA* gene. Coding sequences are in black and 5'- and 3'-UTRs are in gray. The position of the T-DNA insertions in *radA* mutant lines is shown. (B) *radA* plants show severe growth retardation, with distorted leaves presenting chlorotic sector. These can be complemented by expression of HA-tagged RADA (RADA:HA) under control of the endogenous promoter. (C) root shortening of *radA* plants. (D) Detail of leaves phenotype. Scan microscope images show that in *radA* leaves epidermal cells are much larger and there are fewer stomata, half the number for the same leaf surface. Size bar is 20 μm (E) Perennial vegetative growth of plants grown on short days (8h light), with development of aerial rosettes (4-month old *radA-1* plant).

**Figure 5.** Reduced fertility of *radA* plants. (A) Comparison of WT and *radA* flower stems showing very small *radA* siliques. (B) *radA* siliques have mostly aborted seeds, and the few seeds produced are mostly non-viable. (C) Alexander staining of pollen in *radA* anthers as compared to WT, showing little pollen production, and an abundance of small and aberrant pollen grains (indicated by arrows). (D) Visible and SEM Images showing that no pollen binds to the papillae in *radA* stigma. (E) Differential interference contrast images of ovules on crosses between *radA-1* flowers and WT pollen. Black arrowheads indicate central cell and egg cell nuclei in unfertilized ovules. White arrowheads indicate developing embryos in fertilized ovules, three days after pollination (DAP). Only 16% of *radA* ovules could be fertilized. Size bar is 50  $\mu\text{m}$ . (F) After seven DAP the pollinated pistils of *radA* did not develop further.

**Figure 6.** Changes in mtDNA sequences stoichiometry in *radA* mutants because of increased ectopic recombination across repeats. (A) Picture of severely affected (*radA-1#1* and *radA-2#1*) and mildly affected (*radA-1#2* and *radA-2#2*) seedlings that were analyzed. (B) Scanning of their mtDNA for changes in relative copy numbers of the different mtDNA regions. Sequences spaced 5-10 kb apart on the mtDNA were quantified by qPCR. Coordinates are those of the Col-0 mtDNA sequence. The position of the mtDNA large repeats LR1 and LR2 and of relevant intermediate size repeats are shown below the graphic. Regions with changed stoichiometries flanked by repeat pairs are shadowed. Error bars are the *SD* values from three technical replicates. (C) Accumulation of crossover products from mtDNA repeats L, F and EE in the *radA* seedlings relative to WT. Results are in a log<sub>2</sub> scale. The scheme shows the qPCR relative quantification of parental sequences 1/1 and 2/2 and of the corresponding crossover products 1/2 and 2/1. Results are the mean of three technical replicates experiments, and error bars correspond to *SD* values.

**Figure 7.** Synergistic effects of the *radA* mutation on *recG1* and *recA3*. (A) Crosses of *recG1-2* and *radA-1* (pollen donor) show that double homozygote *recG1 radA* plants are as affected as simple homozygote *radA* plants. (B) Crosses and segregation of *recA3-2* and *radA-1* (pollen donor) show that double homozygote *recA3 radA* seedlings do not grow roots and do not grow further. The phenotype is similar to the one observed for *recA2* mutants (Miller-Messmer et al. 2012). The size bar is 1 mm. (C) qPCR analysis of the copy number of mtDNA sequences around and within the region comprised between pair of repeats EE, previously shown to generate an episome by recombination in *recG1* plants (Wallet et al. 2015). Autonomous replication of the episome is significantly increased in the *recG1 radA* double mutant. Results show the mean and *SD* error bars from 2 or 3 (for *recG1 radA*) biological replicates. (D) Scanning of the copy numbers of the different mtDNA regions in *recA3 radA*, as

described in Figure 6, showing severe reduction of several mtDNA regions. Results are in a log scale.

**Figure 8.** Accumulation of mitochondrial and cell cycle transcripts in *radA*. (A) Representative plastidial transcripts were quantified by RT-qPCR and normalized against a set of nuclear housekeeping genes. Quantification of transcripts for several orfs whose transcription can be affected by ectopic recombination involving IRs are shown on the right. Two pools of *radA-1* seedlings were analyzed. Results are on a log<sub>2</sub> scale (B) Quantification of transcripts of selected key cell cycle regulators involved in nuclear DNA damage response (DDR) activated as response to stress and of nuclear DNA repair factors.

### Supplemental Figures

**Figure S1.** Phylogenetic distribution of RADA. Genes coding for RadA-like proteins are found in all bacteria, in land plants, green, brown and red algae, diatoms and other organisms of the Stramenopile group. *Arabidopsis thaliana*, NP\_199845; *Populus trichocarpa*, EEE84551; *Vitis vinifera*, XP\_002277638; *Oryza sativa*, NP\_001056828; *Zea mays*, NP\_001170708; *Selaginella moellendorffii*, XP\_002976563; *Physcomitrella patens*, XP\_001757578; *Ostreococcus tauri*, XP\_003084463; *Coccomyxa subellipsoidea*, XP\_005643141; *Chlorella variabilis*, EFN56488; *Cyanidioschyzon merolae*, XP\_005536634; *Galdieria sulphuraria*, XP\_005709405; *Ectocarpus siliculosus*, CBJ25917; *Phaeodactylum tricornutum*, XP\_002178713; *Thalassiosira oceanica*, EJK58798; *Saprolegnia diclina*, EQC30001; *Phytophthora infestans*, XP\_002904225; *Myxococcus xanthus*, YP\_629513; *Rickettsia prowazekii*, WP\_014607237; *Neisseria meningitidis*, WP\_002258526; *Escherichia coli*, WP\_001458566; *Microcystis aeruginosa*, WP\_002742386; *Synechocystis sp*, WP\_009633429; *Geminocystis herdmanii*, WP\_017296205; *Bacillus anthracis*, NP\_842650; *Amphibacillus jilinensis*, WP\_017473696.

**Figure S2.** High conservation of RADA sequences between plants and bacteria. Sequence alignment (Clustal W) between representative land plant RADA sequences and RadA from proteobacteria and cyanobacteria. The Zinc-finger and KNRF G RadA-specific motif are shaded in red and blue respectively. The matrix of sequence identities and similarities is shown below the alignment.

**Figure S3.** Effect of the point mutation K201A (amino acids in green) on the Walker A domain of Arabidopsis RADA in the binding and hydrolysis of ATP. The structure shown is the one from bacterial RadA (Marie et al. 2017), with bound ADP and and Mg<sup>2+</sup> ion (sphere). The two phosphate groups of ADP are in red.

**Figure S4.** Expression, purification and characterization of recombinant RADA. The Arabidopsis RADA sequence minus the first 48 codons corresponding to the OTS was cloned in the expression vector pET28a fused to a N-terminal His-tag. The recombinant RADA and Walker mutant K201A were expressed in the Rosetta(DE3) strain and purified by affinity and gel filtration. **(A)** Coomassie gel staining analysis of the recombinant proteins. **(B)** Gel filtration on Superdex S200 showed that RADA purified as two peaks of high molecular weight. **(C)** Dynamic light scattering of the protein fraction from peak 2 shows that it is monodispersed and corresponding to a size of about 340 kDa, which is consistent with an hexameric RADA molecule. **(D)** EMSA analysis of the binding to a ssDNA oligonucleotide. Fractions corresponding to both peaks give complexes of the same size, although fractions of peak 1 give predominantly higher molecular weight complexes.

**Figure S5.** Sequence specificity of RADA binding to ssDNA. A 30-mer ssDNA oligonucleotide (7x[AGCT]AG) was used as probe in EMSA experiments with recombinant RADA, and sequence specificity was tested by competition with increasing concentrations of the cold homologous oligonucleotide or with 30-oligomers (poly-A, pol-T, poly-C and poly-G). Competitor/probe ratios tested were 0; 2.5; 10 and 40. Only poly-A showed reduced competition for RADA binding.

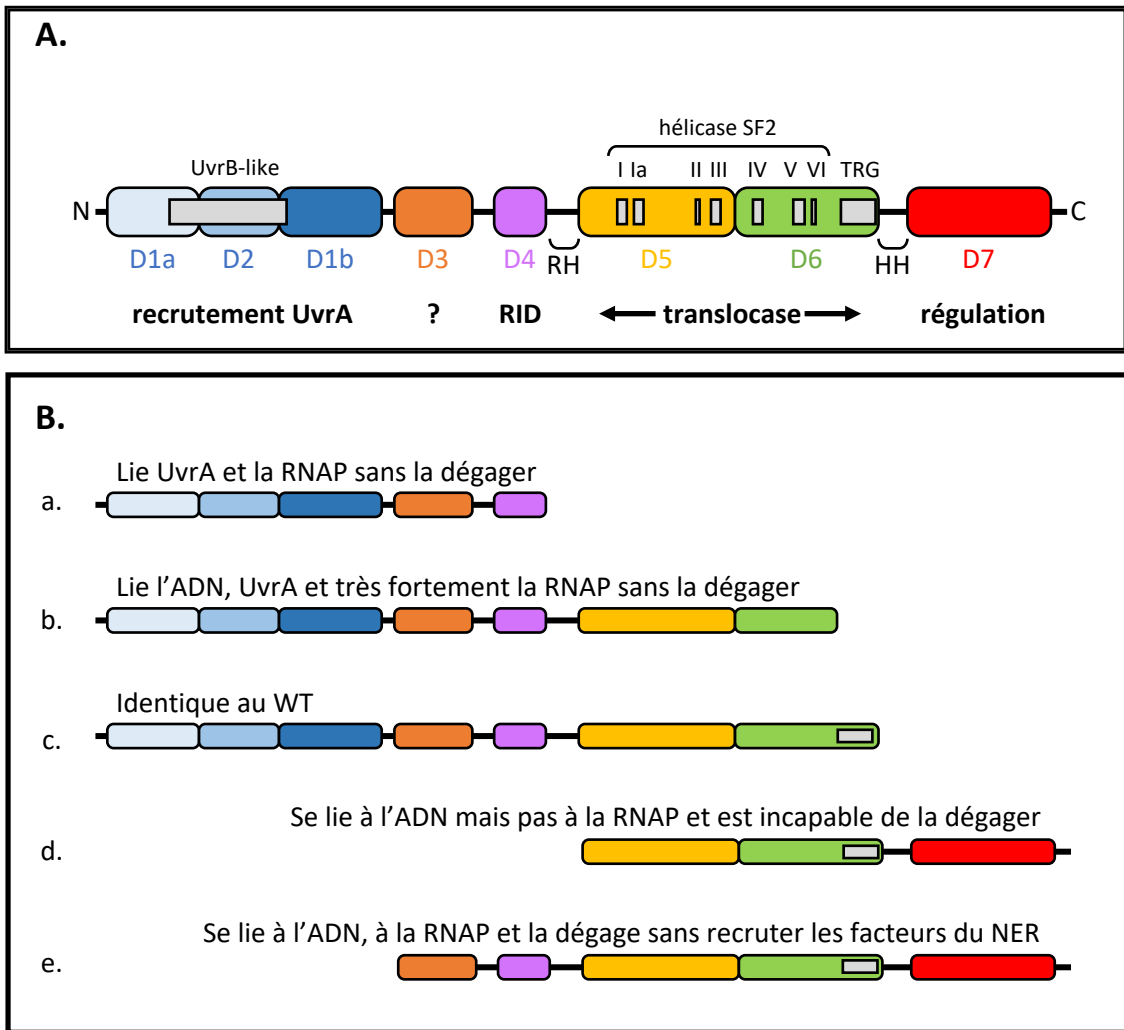
**Figure S6.** Transmission electron microscope (TEM) images of WT col-0 and *radA-1*. TEM images of cells from leaves of same size showed morphologically normal chloroplasts in *radA*. Mitochondria were enlarged and less electron dense as compared to mitochondria from WT cells.

**Figure S7.** No apparent effect of RADA deficiency on cpDNA stability. Scanning of the cpDNA of *radA* mutants for changes in relative copy numbers of the different cpDNA regions. Sequences spaced 5-10 kb apart on the cpDNA were quantified by qPCR, as described for the mtDNA in Figure 6. Coordinates correspond to the ones of the published Col-0 cpDNA

sequence. The three major regions of the cpDNA are indicated: LSC, large single copy region; SSC, small single-copy region; IR, inverted repeat.



## **CHAPITRE 2 : ÉTUDE DE TRCF**



**Figure 25 : Structure et fonctions du TRCF bactérien.** **A.** Les différents domaines fonctionnels de TRCF sont représentés par les boîtes colorées. Les domaines D1a, D2 et D1b permettant le recrutement d'UvrA sont représentés en bleus et contiennent une région similaire en séquence à UvrB représentée en grise. Le domaine D3, en orange n'a pas de fonction connue. D4 permet l'interaction avec les ARN polymérase (RID : *RNA polymerase interacting domain*). Les deux domaines translocase D5 et D6 sont représentés en jaune et en vert. Ils contiennent les sept motifs typiques des hélicases de la famille SF2 (petites barres grises), et le motif TRG retrouvé chez RecG (grande barre grise). D5 et D6 sont impliqués dans la liaison à l'ADN, le déplacement des ARN polymérase bloquées et les changements conformationnels de TRCF. Le domaine D7, en rouge, permet d'inhiber le recrutement des facteurs du NER en dehors du contexte de réparation. Les hélices flexibles RH et HH, potentiellement responsables de la transmission des changements de conformation, sont indiquées par des accolades. **B.** Différents mutants de TRCF délétés d'un ou plusieurs domaines sont représentés. Les domaines D5, D6 et le motif TRG sont indispensables au fonctionnement de TRCF (**a, b et c**). En l'absence de D3, TRCF n'est plus recruté par les ARN polymérase bloquées (**d**). Les domaines D1a, D2, D1b et D7 ne sont pas nécessaires pour le dégagement des ARNP bloquées (**c et e**). (D'après Selby, 2017 et Deanocescu *et al.*, 2006)

## Introduction

TRCF est le produit du gène *Mfd* (*mutation frequency decline*), d'abord identifié par Witkin chez *E. coli* pour sa capacité à réduire l'apparition de mutations suppressives dans un environnement inhibant la synthèse protéique (Witkin, 1956). L'abolition du phénotype de réduction de ces mutations dans les souches mutées pour les facteurs du NER (*Uvr-*) relia ensuite *Mfd* au *dark repair*, la réparation des lésions UV n'impliquant pas la lumière (le NER) (Witkin, 1966). Il sera finalement caractérisé comme le facteur couplant la transcription à la réparation, nécessaire au dégagement de l'ARN polymérase (RNAP) pour recruter les facteurs du NER, et renommé TRCF (Selby *et al.*, 1991).

### 1. Structure de TRCF bactérienne

TRCF est une ADN translocase appartenant à la famille des hélicases SF2 composée de huit domaines fonctionnels reliés par des hélices flexibles (**figure 25**). La partie N-terminale de TRCF est composée des trois domaines D1a, D2, D1b, homologues en séquence aux trois domaines structuraux homonymes de la partie N-terminale de la protéine *UvrB* du NER (Deaconescu *et al.*, 2006). Ce domaine permet à TRCF, comme à *UvrB*, d'interagir avec la protéine *UvrA*. Cette interaction est plus faible avec TRCF qu'avec *UvrB* car la partie N-terminale de TRCF est repliée en dehors du contexte de réparation, rendant inaccessible la région d'homologie à *UvrB* (Deaconescu *et al.*, 2012). L'homologie à *UvrB* ne permet pas à TRCF de réaliser les autres fonctions de la protéine *UvrB*. Le domaine D3 de TRCF est peu conservé, voire absent dans certains TRCF, et n'a pas de rôle connu (Deaconescu *et al.*, 2006). Le cinquième domaine de TRCF (D4) est un domaine d'interaction avec l'ARN polymérase bactérienne (aussi appelé RID : *RNA polymerase interacting domain*). Le domaine adopte une structure de type *Tudor-like* permettant l'interaction spécifique avec la sous-unité  $\beta$  de la RNAP (Rain *et al.*, 2001; Deaconescu *et al.*, 2006). Ce type de structure est retrouvé chez les protéines *CarD* et *CdnL* de *Myxococcus canthus*, des activateurs de la transcription (Bernal-Bernal *et al.*, 2015). La résolution de la structure cristallographique de l'interaction entre TRCF et la RNAP a révélé un changement de conformation de neuf acides aminés au niveau de la sous-unité  $\beta$  de la RNAP (Westblade *et al.*, 2010). Ce changement de conformation pourrait permettre la reconnaissance spécifique des RNAP bloquées ou arrêtées par TRCF (Selby, 2017). Après le domaine RID se trouvent deux domaines hélicase dépendant de l'ATP (D5 et D6) similaires à ceux retrouvés chez l'hélicase *RecG* impliquée dans la HR (Singleton *et al.*, 2001; Deaconescu *et al.*, 2006). Ces domaines permettent la liaison et le déplacement de TRCF sur l'ADNdb, ainsi que le dégagement des RNAP bloquées ou le rétablissement des RNAP ralenties ou arrêtées (Chambers *et al.*, 2003). Le domaine D6 contient le motif TRG (*translocation in RecG*) de 37 acides aminés homologue à *RecG* et nécessaire aux fonctions des deux protéines. Chez TRCF, ce motif TRG serait responsable de l'activité translocase déclenchée après l'interaction avec la

RNAP bloquée, permettant son dégagement en tirant l'ADN à travers la RNAP (Park *et al.*, 2006). Le motif serait également impliqué dans le changement de conformation de la protéine permettant le recrutement de UvrA (Manelyte *et al.*, 2010). Enfin, la partie C-terminale de TRCF comprend le domaine D7 qui recouvre le domaine D2 en dehors du contexte de réparation, empêchant le recrutement d'UvrA (Deaconescu *et al.*, 2006). A cause du domaine D7, la protéine TRCF se retrouve dans un état réprimé jusqu'à l'interaction avec une RNAP qui va induire un premier changement de conformation, activant ses domaines D5 et D6. L'activation des domaines moteurs entraînera ensuite un second changement de conformation permettant le recrutement des facteurs de réparation (Selby, 2017).

En plus des huit domaines, la structure de TRCF présente également des hélices flexibles situées entre les domaines D4 et D5, RH (*relay helix*), et entre D6 et D7, HH (*hook helix*) (Deaconescu *et al.*, 2006). Ces deux hélices permettraient de communiquer les changements de conformation initiés par les domaines hélicase de TRCF aux autres parties de la protéine. Ainsi, la RH pourrait transmettre le mouvement au domaine RID, assistant la dissociation de la RNAP, tandis que HH permettrait le déplacement du domaine D7 (Selby, 2017). Ces hélices pourraient également transmettre le changement de conformation de TRCF lors de l'interaction avec la RNAP, activant ainsi les domaines hélicase.

Plusieurs mutants de TRCF ont été produits pour comprendre le rôle et l'importance de ses différents domaines. Ainsi, les mutants détruisant l'interaction entre D2 et D7 stabilisent la conformation ouverte de TRCF, augmentant sa liaison à l'ADN et ses activités ATPase et translocase (Park *et al.*, 2002; Prabha *et al.*, 2011). Les protéines mutées dans les domaines homologues à UvrB ne sont plus capables de recruter les facteurs du NER (Selby et Sancar, 1995a). Néanmoins, ces mutants sont tout de même capables de se lier à l'ADN pour dégager les RNAP bloquées (Selby et Sancar, 1995a). Les mutants dans les domaines hélicase, et particulièrement dans le motif TRG empêchent l'activité ATPase et le dégagement de la RNAP (Deaconescu *et al.*, 2012). Le mutant tronqué du motif TRG, de HH et de D7 est capable de se lier à l'ADN et s'associe très fortement avec la RNAP, révélant le rôle de TRG dans la régulation de l'interaction avec UvrA (Selby et Sancar, 1995a). De manière similaire, l'expression d'une protéine composée uniquement des domaines homologues à UvrB entraîne la séquestration de UvrA, augmentant la sensibilité aux UV (Prabha *et al.*, 2011).

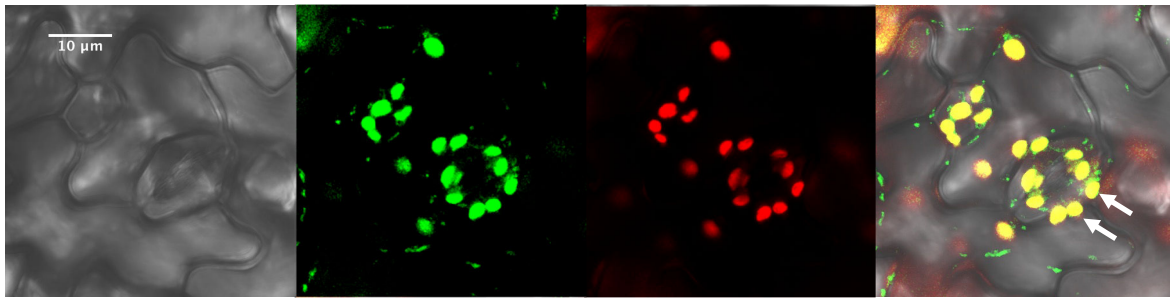
## 2. Rôles de TRCF

TRCF est le facteur permettant de coupler la transcription à la réparation de l'ADN en reconnaissant et dégageant les RNAP bloquées pour recruter les facteurs du NER. En plus de favoriser la réparation des gènes fortement transcrits, et donc nécessaires à la cellule, la voie du TC-NER reconnaît et traite plus rapidement certains dommages relativement mal pris en charge par le GG-NER (Hu *et al.*, 2015). Si le mécanisme du TC-NER est assez bien compris, la

manière dont TRCF recrute et positionne les facteurs UvrA et UvrB au niveau de la lésion reste évasive, notamment car certains mutants déficients dans la reconnaissance des dommages *via* le GG-NER ne sont pas impactés dans le TC-NER (Manelyte *et al.*, 2010). TRCF pourrait donc être capable de reconnaître intrinsèquement les lésions dans le complexe TRCF-UvrA<sub>2</sub>-UvrB. En plus des RNAP bloquées par des dommages distordant l'ADN, TRCF peut également déplacer les RNAP arrêtées par manque de nucléotides à incorporer (Fan *et al.*, 2016). Dans ce cas, après effondrement du complexe d'élongation, la RNAP reste associée à TRCF qui se déplace sur l'ADN pour dissocier d'éventuels blocages en aval (Fan *et al.*, 2016). TRCF peut également jouer le rôle de facteur *anti-backtracking* en poussant les RNAP en pause, permettant de rétablir la transcription et de conserver le transcrit en cours de synthèse (Park *et al.*, 2002).

En dehors du TC-NER, TRCF est impliqué dans la résolution des conflits entre les machineries de transcription et de réplication. Lorsqu'une fourche de réplication rencontre une RNAP bloquée, la collision peut mener à la formation d'une cassure double brin (Sankar *et al.*, 2016). En dégageant les RNAP bloquées TRCF limite ces collisions et l'apparition de DSB (Trautinger *et al.*, 2005). TRCF peut également jouer un rôle dans l'expression des gènes, en dégageant les RNAP bloqués par des répresseurs cataboliques et participant à l'inhibition de la transcription (Belitsky et Sonenshein, 2013). Plus récemment, Martin *et al.* ont montré que TRCF est impliqué dans la réparation de dommages oxydatifs chez *Bacillus subtilis* (Martin *et al.*, 2019). Mfd interagirait avec des composants du BER, telle que la glycosylase MutY, pour réparer les dommages engendrés par des ROS, rappelant l'un des rôles de CSB, l'homologue fonctionnel eucaryotique de TRCF (Menoni *et al.*, 2018).

Malgré ses différents rôles, TRCF n'est pas essentiel pour la bactérie. Le mutant est viable et n'est que faiblement affecté par un traitement génotoxique aux UV (Witkin, 1956, 1966). Le mutant présente néanmoins une augmentation du taux de mutation (Witkin, 1956). L'absence de phénotype sévère pour *mfd* est expliquée par la présence de voies alternatives de réparation couplée à la transcription ne faisant pas intervenir TRCF. Les protéines NusA et UvrD ont été proposées dans la prise en charge des RNAP bloquées par des lésions distordant l'ADN (Cohen *et al.*, 2010; Epshtein *et al.*, 2014). Au lieu de dégager la RNAP, ces protéines la feraient reculer (*backtracking*), démasquant la lésion pour les facteurs de réparation du NER. Contrairement à TRCF, UvrD est régulée par la réponse SOS au même titre que UvrA et UvrB et permettrait de prendre en charge un grand nombre de lésions lors des phases de stress (Epshtein *et al.*, 2014). TRCF aurait plutôt un rôle dans la réparation basale du génome, en absence de stress. Le facteur de terminaison Rho pourrait également participer au dégagement des RNAP bloquées par des lésions, en compétition ou en collaboration avec TRCF (Jain *et al.*, 2019).

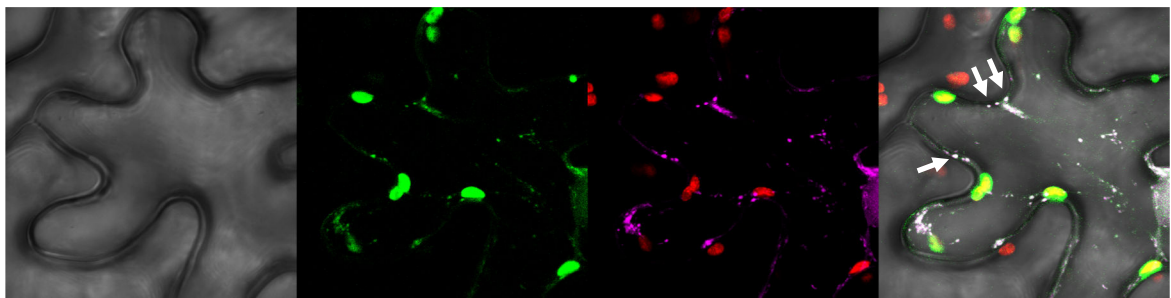


DIC

TRCF:eGFP

Chlorophylle

Superposition



DIC

TRCF:eGFP

CoxIV:dsRED

Superposition

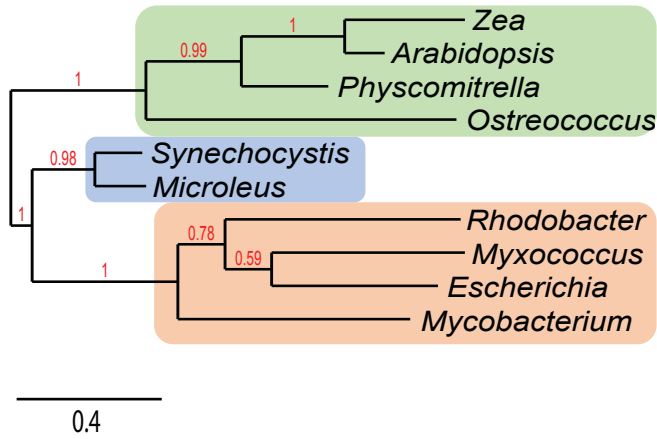
**Figure 26 : Localisation subcellulaire de TRCF.** La séquence d'adressage putative de TRCF est fusionnée à la séquence de l'eGFP en C-terminal. Après transformation stable des plantes (en haut) ou transformation transitoire par biolistique (en bas), les cellules sont observées en microscopie confocale (seconds panneaux, en vert). La fluorescence de la eGFP colocalise avec l'auto-fluorescence de la chlorophylle (troisièmes panneaux, en rouge), et avec la fluorescence d'un marqueur mitochondrial (en bas, troisième panneau, en violet). Les co-localisations sont indiquées par des flèches blanches dans les quatrièmes panneaux.

## Résultats

La protéine TRCF d'*A. thaliana* a été identifiée en même temps que RECG1 avant le début de ma thèse lors de la recherche d'hélicases organellaires pouvant être impliquées dans la HR. TRCF et RECG1 possèdent le motif TRG retrouvé chez la protéine RecG bactérienne. RECG1 a ensuite été caractérisé comme l'homologue fonctionnel de RecG (Wallet *et al.*, 2015) tandis que TRCF s'est révélé phylogénétiquement plus proche de la protéine Mfd bactérienne. Les logiciels de prédiction adressant TRCF dans les organelles, l'équipe s'est intéressée aux rôles que pourraient y avoir TRCF, étant donné qu'aucun autre facteur du NER n'est adressé aux mitochondries ni aux chloroplastes. Si TRCF se retrouve impliqué dans le NER, il s'agirait alors de la première preuve de l'existence de cette voie de réparation de l'ADN dans les organelles. TRCF pourrait également être impliqué dans d'autres mécanismes de réparation sans lien avec le NER. Les facteurs CSA et CSB, couplant la transcription à la réparation dans le TC-NER nucléaire, ont un rôle dans la réparation des dommages oxydatifs dans la mitochondrie des mammifères par exemple, en interagissant avec des facteurs du BER (D'Errico *et al.*, 2013). En plus de ses fonctions dans la réparation de l'ADN, TRCF joue aussi un rôle dans la régulation des RNAP, via son activité *anti-backtracking*, et permet la résolution des conflits entre les machineries transcriptionnelles et répliquatives. Afin de caractériser les fonctions de TRCF dans les organelles d'*Arabidopsis*, nous avons entrepris d'identifier ses partenaires d'interactions afin de savoir dans quelles voies la protéine pourrait être impliquée. Nous avons également étudié l'impact de la mutation de TRCF sur le développement de la plante et sur son génome.

### 1. Localisation subcellulaire de TRCF

Afin de confirmer expérimentalement la prédiction d'adressage aux organelles de TRCF, une construction comprenant les 93 premiers acides aminés de la protéine, correspondant à sa séquence d'adressage putative (OTS : *organellar targeting sequence*), ont été fusionnés à la séquence de l'eGFP (OTS<sub>TRCF</sub>:eGFP). Des plantes transgéniques exprimant cette construction ont été produites, révélant une co-localisation de l'autofluorescence de la chlorophylle avec la fluorescence de OTS<sub>TRCF</sub>:eGFP (**figure 26**). Ces résultats confirment l'adressage chloroplastique de TRCF. Néanmoins, de multiples *loci* ponctués d'eGFP sont également visibles dans le cytoplasme et pourraient correspondre à un adressage mitochondrial. Afin de tester cette hypothèse, des feuilles de tabac ont été co-bombardées avec la construction OTS<sub>TRCF</sub>:eGFP et la construction mRFP (Vermel *et al.*, 2002), correspondant à la protéine DsRed fusionnée à la séquence N-terminale d'adressage vers la mitochondriale de la protéine CoxIV de levure (OTS<sub>CoxIV</sub>:dsRed) (**figure 26**). L'observation en microscopie confocale de cellules transformées montre une co-localisation des *loci* cytoplasmique de GFP avec la fluorescence rouge de la DsRed, indiquant un adressage mitochondrial de TRCF. Ainsi, TRCF est adressé à la fois aux mitochondries et aux chloroplastes.



**Figure 27 : Phylogénie de TRCF.** La comparaison des séquences de la protéine TRCF issues d'organismes pertinents pour l'évolution des organelles indique une acquisition de TRCF par l'ancêtre du chloroplaste chez les plantes. Les valeurs de bootstrap sont indiquées en rouge sur chaque branche.

*Rhodobacter*, *Myxococcus* et *Escherichia* (protéobactéries) et *Mycobacterium* (actinobactérie), *Synechocystis* et *Microcoleus* (cyanobactéries), *Ostreococcus* (algue verte), *Physcomitrella* (bryophyte), *Arabidopsis* (dicotylédone) et *Zea* (monocotylédone).



## 2. Étude des séquences et modélisation de la structure de TRCF

### 2.1. Analyse de l'alignement des séquences protéiques de TRCF

La protéine TRCF est retrouvée au sein des procaryotes, des algues rouges et vertes et des plantes terrestres. L'arbre phylogénétique généré à partir des séquences protéiques de TRCF de différents organismes suggère une acquisition de TRCF chez les plantes terrestres par les cyanobactéries, ancêtres du chloroplaste (**figure 27**).

Si TRCF est conservé chez les procaryotes et les eucaryotes chlorophylliens, l'alignement des séquences protéiques montre un raccourcissement de la protéine chez les organismes de la lignée verte (**figure 28**). La partie N-terminale correspondant aux domaines D1a/D2/D1b et permettant le recrutement d'UvrA est tronquée de ses 200 premiers acides aminés chez les plantes et les algues. Les 150 acides aminés de la partie C-terminale des domaines D1a/D2/D1b ne sont pas du tout conservés entre procaryotes et eucaryotes. De plus, la partie N-terminale de TRCF d'*A. thaliana* doit correspondre à la séquence d'adressage aux organelles, généralement peu conservée entre différentes plantes, impliquant que les domaines D1a/D2/D1b soient raccourcis davantage encore. La séquence homologue à UvrB est d'ailleurs complètement perdue chez le TRCF des plantes. Ces observations suggèrent l'absence complète des domaines fonctionnels de recrutement d'UvrA, ce qui est logique, vu qu'il n'y a aucun gène codant une protéine homologue à UvrA dans le génome d'*Arabidopsis*. A la suite, le domaine D3 n'est également pas conservé. Ce domaine de fonction inconnu est cependant déjà très variable au sein des procaryotes, étant parfois même absent (Deaconescu et Suhanovsky, 2017). Le domaine D4/RID de liaison à la RNAP semble plus conservé entre procaryotes et eucaryotes. Tous les acides aminés impliqués dans l'interaction avec le feuillet  $\beta$  de la RNAP chez *Thermus thermophilus* ne sont néanmoins pas conservés chez les plantes. Seuls les quatre acide aminés Y174 A184, K185 et L186, impliqués dans l'interaction avec RPOB chez *T. thermophilus*, sont conservés chez *Arabidopsis* (Westblade *et al.*, 2010). La modification des autres résidus impliqués peut néanmoins provenir de la modification des séquences des sous-unités de la PEP chloroplastique, potentiel partenaire d'interaction de TRCF. Les domaines D5 et D6 sont bien conservés, notamment au niveau des sept motifs hélicase SF2 et du motif TRG (**figure 28**). Enfin, le domaine C-terminal D7, permettant de masquer les domaines de recrutement d'UvrA, est présent chez les plantes malgré l'apparente absence de domaines D1a/D2/D1b à masquer.

### 2.2. Modélisation de la structure de TRCF d'*Arabidopsis*

Les structures secondaires et tertiaires des protéines étant généralement plus conservées que leurs séquences en acides aminés, nous avons modélisé la structure de TRCF d'*Arabidopsis* à partir de la structure résolue de Mfd d'*E. coli* (**figure 29**) (Deaconescu et

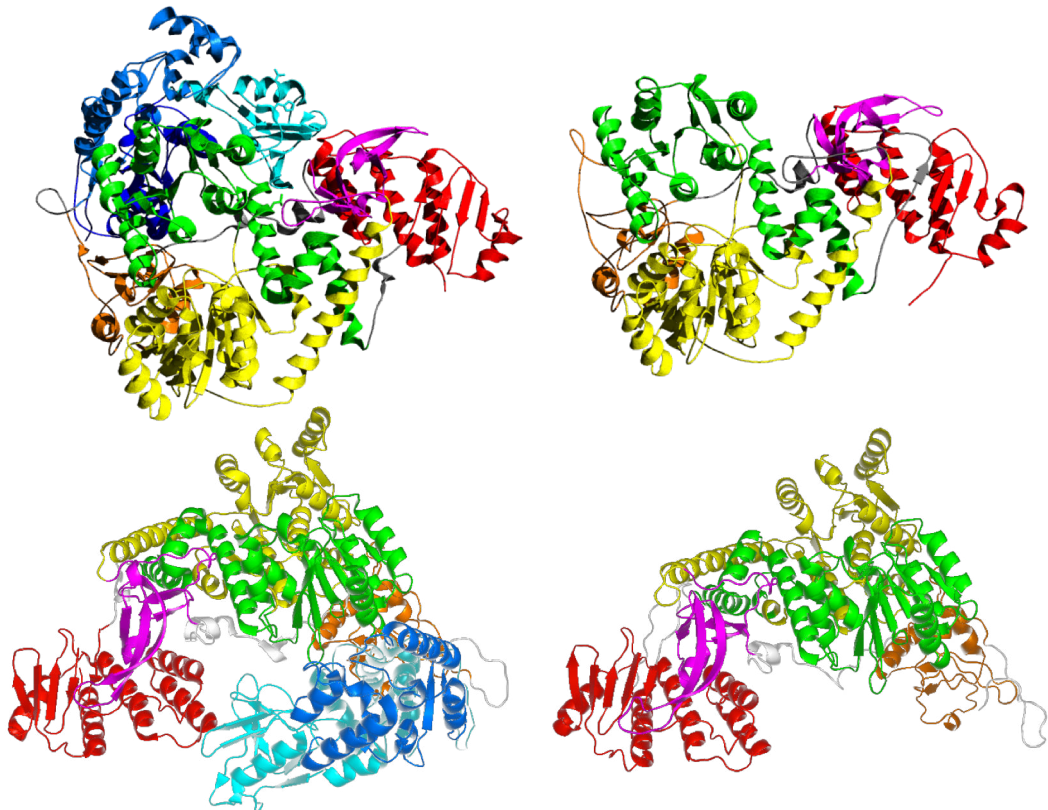


**Figure 28 : Alignements des séquences protéiques de TRCF.** Les séquences de la protéine TRCF issues d'organismes pertinents pour l'évolution des plantes à fleurs sont alignées. Les acides aminés en blanc sur fond noir sont identiques entre les différents organismes, ceux en noir sur fond gris ont des propriétés similaires entre les différents organismes. Les domaines fonctionnels sont représentés dans les mêmes couleurs que la figure R1 (D1a/D2/D1b en bleu, D3 en orange, D4/RID en violet, D5 en jaune, D6 en vert et D7 en rouge). Les domaines D1a, D2, D1b sont absents chez les algues vertes et les plantes terrestres. La région homologue à la protéine UvrB est entourée en rose. Les acides aminés impliqués dans l'interaction entre TRCF et RPOB chez *T. thermophilus* sont entourés en bleu, dans le domaine D4/RID (Westblade et al., 2010). Les acides aminés composants les hélices flexibles des domaines D5 et D6 sont entourés en vert. Les acides aminés des sept motifs hélicase SF2 et du motif TRG sont entourés en rouge.

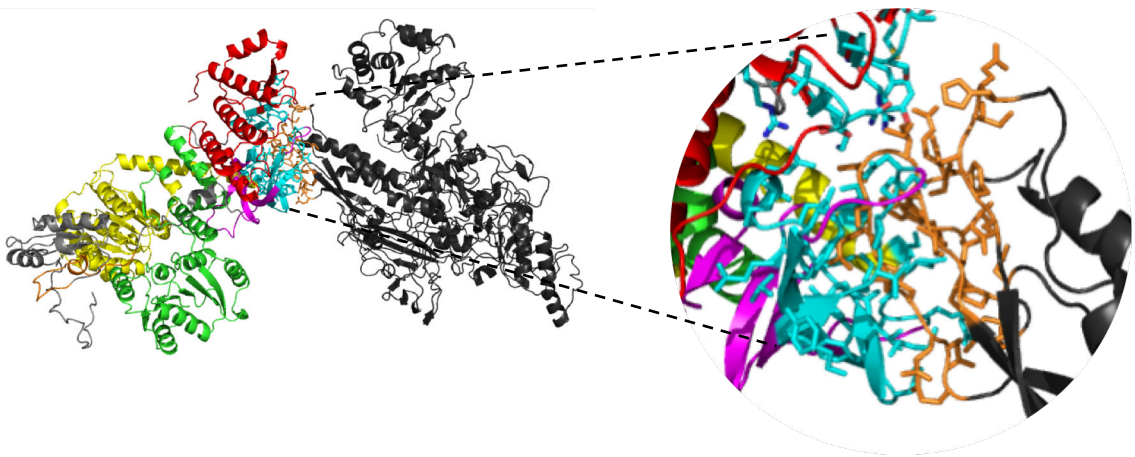
*Escherichia* ( $\gamma$ -protéobactérie) et *Rickettsia* ( $\alpha$ -protéobactérie), *Synechocystis* (cyanobactérie), *Ostreococcus* (algue verte), *Physcomitrella* (bryophyte), *Amborella* (angiosperme primitive), *Oryza* (monocotylédone) et *Arabidopsis* (dicotylédone).

*Escherichia coli*

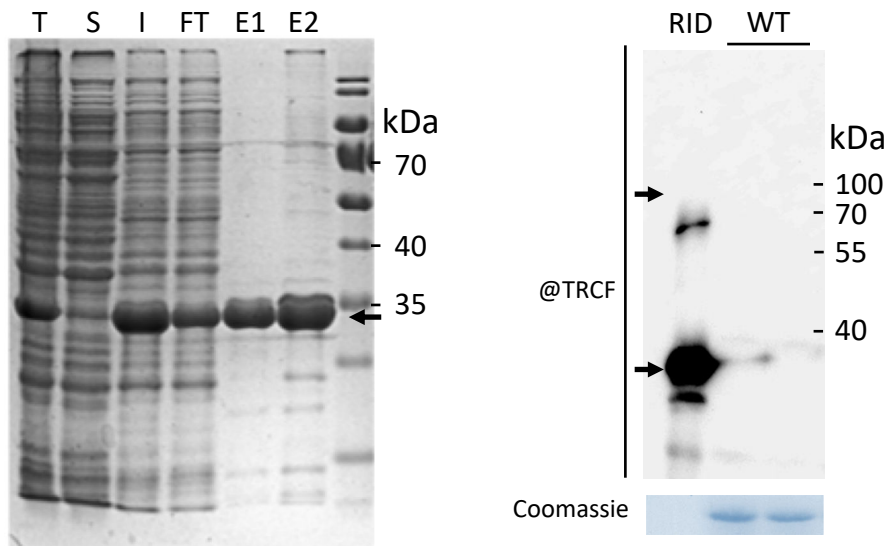
*Arabidopsis thaliana*



**Figure 29 : Modélisation de la structure de la protéine TRCF d'*Arabidopsis thaliana*.** La structure de TRCF d'*Arabidopsis* (à droite) est modélisée à partir de la structure cristallographique résolue de Mfd d'*E. coli* (à gauche) (Deaconescu *et al.*, 2006). Les domaines fonctionnels sont représentés dans les mêmes couleurs que la figure R1 (D1a/D2/D1b en bleu, D3 en orange, D4/RID en violet, D5 en jaune, D6 en vert et D7 en rouge). Les domaines de recrutement de UvrA (D1a, D2, D1b) sont absents chez *Arabidopsis thaliana*. La structure des autres domaines pourrait être conservée et indiquerait un rôle similaire entre les protéines de bactéries et de plantes, sans le recrutement des facteurs du NER pour cette dernière.



**Figure 30 : Modélisation de l'interaction entre les protéines TRCF et RPOB d'*A. thaliana*.** La structure de l'interaction est modélisée à partir de la structure résolue chez la bactérie *T. thermophilus* (Westblade *et al.*, 2010). TRCF est représenté avec le code couleur de la figure R1, et RPOB est représenté en noir. Les acides aminés potentiellement impliqués dans l'interaction sont représentés en bleu ciel pour TRCF et en orange clair pour RPOB. Cette modélisation soutient l'interaction de TRCF avec RPOB selon un modèle similaire à celui retrouvé chez les bactéries.



**Figure 31 : Expression et purification d'une protéine RID recombinante pour produire des anticorps reconnaissant la partie N-terminale de TRCF.** A gauche, la coloration au bleu de Coomassie du gel contenant les différentes fractions de la purification de la protéine recombinante composée d'une étiquette 6-His en N-terminal et des séquences des domaines D4 (RID) et D5 (T : fraction totale ; S : fraction soluble ; I : fractions insoluble ; FT : surnageant ; E1 et E2 : fractions éluées). A droite, l'immunodétection de la protéine cible de 34 kDa (RID) par les anticorps obtenus. Des extraits totaux de plantes (WT) sont également testés mais ne permettent pas de visualiser TRCF (taille attendue : 94 kDa).



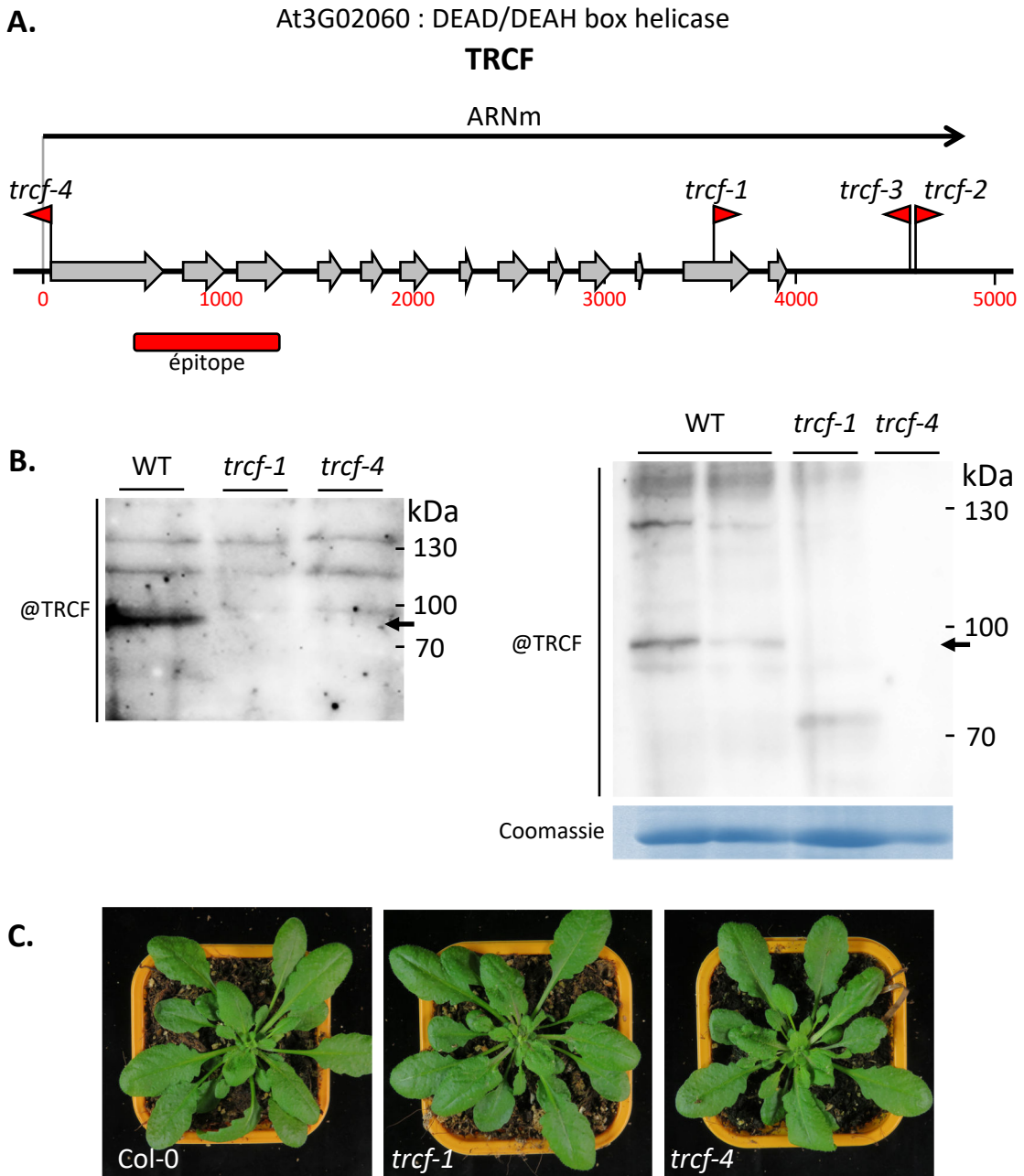
Suhanovsky, 2017). Cette modélisation permet de conserver l'organisation générale du TRCF bactérien pour la protéine d'*Arabidopsis*. Le domaine D4/RID conserve sa structure composée de plusieurs feuillets  $\beta$  antiparallèles malgré la modification de la séquence. Chez les bactéries, ces feuillets  $\beta$  sont impliqués dans l'interaction avec la sous-unité  $\beta$  de la RNAP (Westblade *et al.*, 2010). Les structures des domaines D5 et D6 pourraient être identiques chez *A. thaliana* et *E. coli*, vu l'importante conservation de leurs séquences en acides aminés. Plus étonnant, le domaine D7 pourrait conserver une structure similaire à celle du TRCF bactérien, malgré l'absence des domaines N-terminaux. Actuellement, le domaine D7 n'est connu que pour son rôle inhibiteur, empêchant TRCF de recruter UvrA en dehors du contexte de réparation. Le mutant bactérien tronqué en C-terminal du domaine D7 est fonctionnellement proche de la protéine sauvage (Selby et Sancar, 1995b). La conservation de D7 chez la protéine TRCF d'*Arabidopsis* est donc particulièrement intrigante. La partie N-terminale de la protéine n'a pas pu être modélisée car trop divergente du modèle utilisé.

Nous avons également modélisé l'interaction entre TRCF et RPOB à partir de la structure résolue des interactions chez *T. thermophilus* (**figure 30**) (Westblade *et al.*, 2010). Notre modèle est compatible avec l'interaction des deux protéines. L'orientation des chaînes latérales pourrait permettre un rapprochement suffisant entre TRCF et RPOB pour permettre leur interaction. Cette modélisation ne nous permet néanmoins pas d'affirmer si TRCF interagit réellement avec RPOB.

### 3. Production d'anticorps reconnaissant la partie N-terminale de TRCF

Lors de l'étude de TRCF antérieure au début de ma thèse, un sérum d'anticorps polyclonaux dirigé contre la partie C-terminale de TRCF (@TRCF-C) avait été produit afin de déterminer l'expression de la protéine des mutants (détaillés dans la partie suivante). Néanmoins, ces anticorps présentaient un important bruit de fond sur des extraits totaux d'*Arabidopsis*, et la reconnaissance d'épitopes dans la partie C-terminale de la protéine ne permettait pas la détection d'éventuelles formes tronquées de TRCF.

Afin de palier à ces problèmes, j'ai fait produire un nouvel anticorps dirigé contre la partie N-terminale de TRCF (@TRCF-N), en utilisant une séquence de 296 acides aminés, correspondante au domaine D4/RID et à une partie du domaine D5 (**figure 31** et **figure 32A**). J'ai produit la protéine recombinante en fusion avec une étiquette 6-histidine en système bactérien et l'ai purifiée par chromatographie affinité au nickel en conditions dénaturantes en présence d'urée 8 M. La protéine étant majoritairement insoluble j'ai utilisé un volume de culture bactérienne important (1 L) afin de la purifier en quantité suffisante pour l'immunisation. La protéine recombinante a été éluée avec un gradient jusqu'à 500 mM d'imidazole. Les fractions d'éluion contenant la protéine d'intérêt ont été rassemblées puis concentrées pour permettre l'immunisation de deux lapins (3571 et 3572). L'injection aux



**Figure 32 : Mutants de TRCF par insertion de T-DNA. A.** La carte génétique de TRCF indique les différents exons (flèches grises) et la position de l'insertion du T-DNA dans quatre mutants indépendants (*trcf-1* à *trcf-4*). Les mutants *trcf-2* et *trcf-3* présentent l'insertion dans la région 3' non traduite de l'ARNm de TRCF, très susceptible de ne pas affecter la production de la protéine TRCF. Le mutant *trcf-1* possède l'insertion au niveau de l'avant dernier exons et le mutant *trcf-4* présente l'insertion au niveau du codon initiateur de la protéine. La région reconnue par l'anticorps produit précédemment est indiquée par un rectangle rouge. **B.** L'immunodétection de la protéine TRCF dans des extraits totaux de plantes révèle l'absence d'expression de la protéine pour *trcf-4* et la formation d'une protéine tronquée d'environ 70 kDa pour *trcf-1*. **C.** Les deux mutants analysés, *trcf-1* et *trcf-4*, ne semblent pas montrer de phénotypes développementaux en comparaison avec le Col-0 WT.



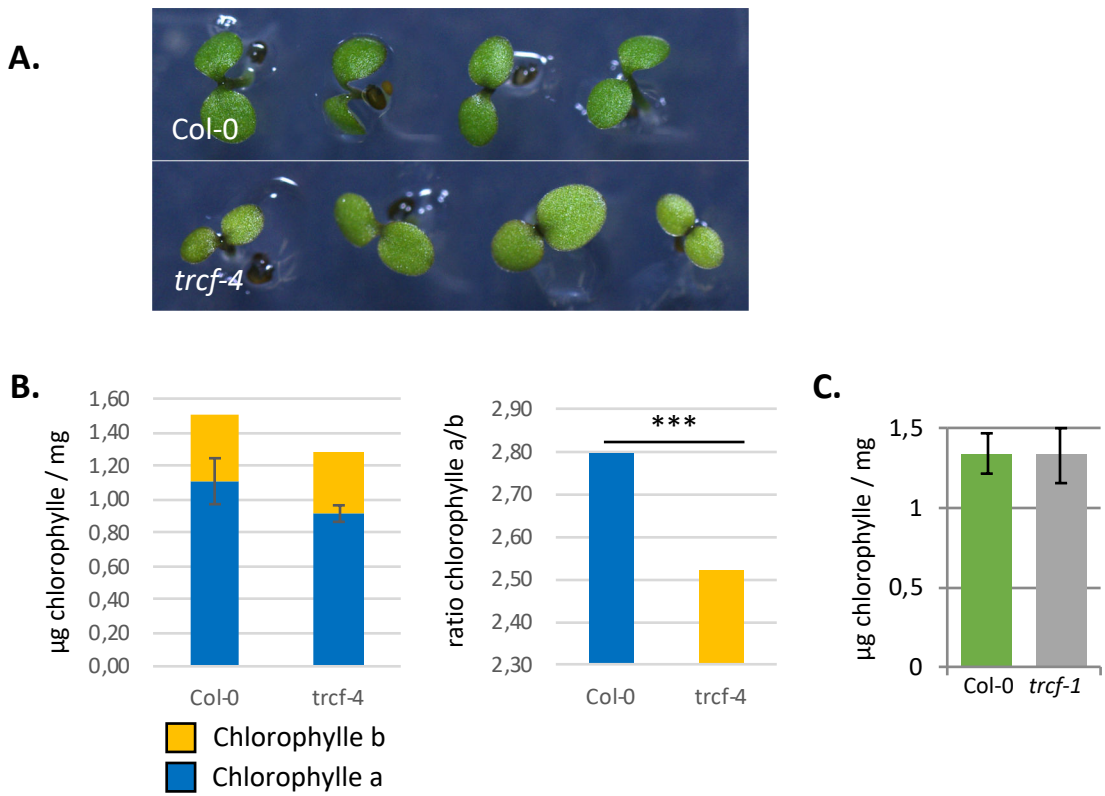
lapins a été réalisée par le service de production d'anticorps de l'institut de génétique et de biologie moléculaire et cellulaire (IGBMC, Illkirch). Les différents séra reçus ont été testés sur des extraits de la protéine recombinante purifiée et sur des extraits totaux de plante (**figure 31**). Si les premiers séra ne permettaient pas d'observer TRCF dans les extraits totaux de plante, les derniers séra du lapin 3572 ont permis de discriminer la protéine, malgré le bruit de fond moins gênant que pour le premier sérum produit (**figure 32B**).

#### 4. Caractérisation de mutants par insertion de T-DNA

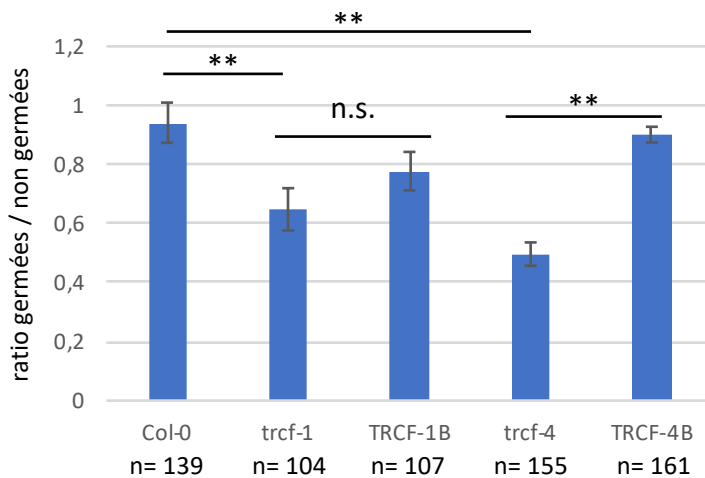
Les mutants *trcf-1*, *trcf-2* et *trcf-3* ont été caractérisés par l'équipe avant le début de ma thèse. L'insertion du T-DNA a été localisée dans la partie C-terminale de TRCF pour les trois mutants : dans l'avant dernier exon pour *trcf-1*, et dans la partie 3' non traduite pour *trcf-2* et *trcf-3* (**Figure 32A**). Aucun des trois mutants ne présentait de phénotype développemental observable (Wallet, 2016). L'immunodétection de TRCF dans les extraits des mutants avec l'anticorps @TRCF-C a montré que *trcf-1* était apparemment un mutant KO (*knock out*), mais que les mutants *trcf-2* et *trcf-3* exprimaient toujours TRCF. Aussi, je ne me suis pas intéressé à *trcf-2* et *trcf-3* lors de ma thèse, cherchant avant tout un mutant n'exprimant plus TRCF. Néanmoins, comme le précédent anticorps ne permettait pas de détecter une forme tronquée en C-terminal de TRCF, il n'était pas certain que *trcf-1* soit réellement un mutant KO. En effet, des expériences de RT-qPCR montraient que les exons en amont de l'insertion T-DNA étaient toujours activement transcrits, même plus que dans la plante sauvage (Wallet, 2016).

J'ai donc caractérisé un quatrième mutant de TRCF par insertion de T-DNA : *trcf-4*. Le séquençage de la bordure gauche (*left border*) du T-DNA a permis de localiser l'insertion au niveau du codon d'initiation de la protéine, présumant une incapacité à produire TRCF (**figure 32A**). Cette impossibilité d'exprimer TRCF a été confirmée par l'immunodétection de la protéine dans les plantes sauvages et mutantes grâce au nouvel anticorps @TRCF-N : TRCF n'est pas détectable dans les extraits de *trcf-4* (**figure 32B**). L'anticorps a aussi permis d'identifier une protéine d'environ 70 kDa dans le mutant *trcf-1*, non présente chez les plantes sauvages, dont la taille correspond à une protéine TRCF tronquée au niveau de l'insertion du T-DNA de *trcf-1*. Le mutant *trcf-1* n'est donc pas un véritable mutant KO, ce qui a pu impacter sur les résultats des expériences de résistance aux stress réalisées sur *trcf-1* avant ma thèse (détaillées dans la partie suivante).

Malgré l'absence d'expression de TRCF, aucun phénotype développemental n'avait été observé chez *trcf-4*, comme pour les trois autres mutants (**figure 32C**), arguant un rôle non fondamental de TRCF dans la plante. Chez les bactéries, le mutant *mfd* n'est également pas affecté dans sa croissance (Witkin, 1956, 1966). Néanmoins, l'observation des plantules au stade cotylédon a révélé une différence dans la pigmentation des feuilles des mutants en comparaison avec les sauvages (**figure 33A**). Les cotylédons des plantules *trcf-4* sont d'une



**Figure 33 : Les mutants *trcf* produisent moins de chlorophylle.** **A.** Au stade cotylédons, les plantules *trcf* montrent une pigmentation légèrement plus claire que les sauvages. **B.** Cette dépigmentation est liée à une quantité de chlorophylles a et b plus faible chez le mutant. La proportion de chlorophylle a par rapport à la chlorophylle b est plus faible chez les plantules *trcf*. **C.** Le mutant *trcf-1* présente la même quantité de chlorophylle totale que le sauvage.



**Figure 34 : Les graines issues des mutants *trcf* ont un taux de germination plus faible.** Le taux de germination sur milieu *in vitro* de graines de plantes sauvages, mutantes et complémentées a été mesuré après stratification. Les graines *trcf-1* et *trcf-4* germent moins fréquemment que les graines sauvages Col-0. La complémentéation de TRCF (TRCF-1B et TRCF-4B) permet d'augmenter le taux de germination de manière significative pour *trcf-4*.

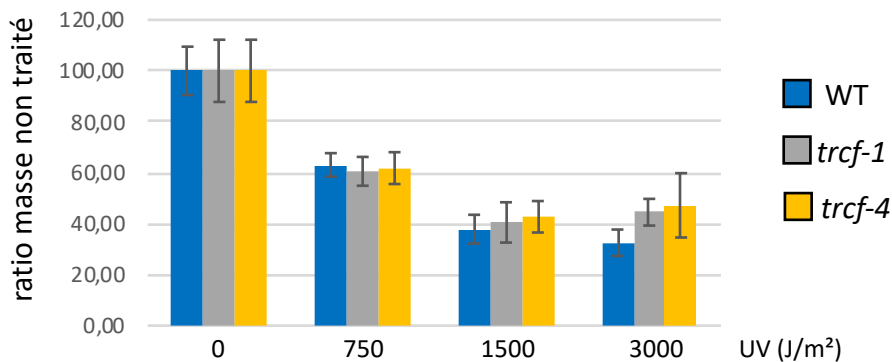
teinte verte légèrement plus claire que ceux des sauvages. Cette légère différence de pigmentation est ensuite observée tout au long du développement de la plante, au niveau des feuilles néoformées. Lorsque la feuille se développe, sa pigmentation devient plus foncée et identique à celle des feuilles d'une plante sauvage. Les mutants *trcf-1* ne présentent pas ce phénotype. Afin de tester cette différence de pigmentation verte, la chlorophylle de plantes sauvages et mutantes a été extraite et quantifiée (**figure 33B**). Comme attendu, aucune différence n'est observée entre le sauvage et le mutant *trcf-1* (**figure 33C**). En revanche, *trcf-4* montre une diminution globale de la quantité de chlorophylle de 15 %. Cette diminution est plus marquée pour la chlorophylle a, qui diminue de 17,5 %, alors que la chlorophylle b ne diminue que de 8,5 %. Cette différence entraîne une modification significative de la proportion de chlorophylle a par rapport à la chlorophylle b. L'absence de TRCF affecte donc la biogénèse des chloroplastes et l'accumulation de chlorophylle, principalement la chlorophylle a, responsable de la pigmentation verte des feuilles. Une réduction du ratio chlorophylle a/b a été décrite pour des plantes soumises à différents régimes de lumière, et aussi dans des conditions de stress osmotique (Brouwer *et al.*, 2014; Wiczarz *et al.*, 2015).

Outre la pigmentation, les mutants *trcf-1* et *trcf-4* présentent une diminution du taux de germination des graines en milieu *in vitro* (**figure 34**). Globalement, 40 % des graines *trcf-1* et 50 % des *trcf-4* ne germent pas. Les mutants complémentés (présentés plus loin) présentent un taux de germination similaire aux plantes sauvages, montrant que le défaut de germination des mutants est bien dû à l'absence de TRCF.

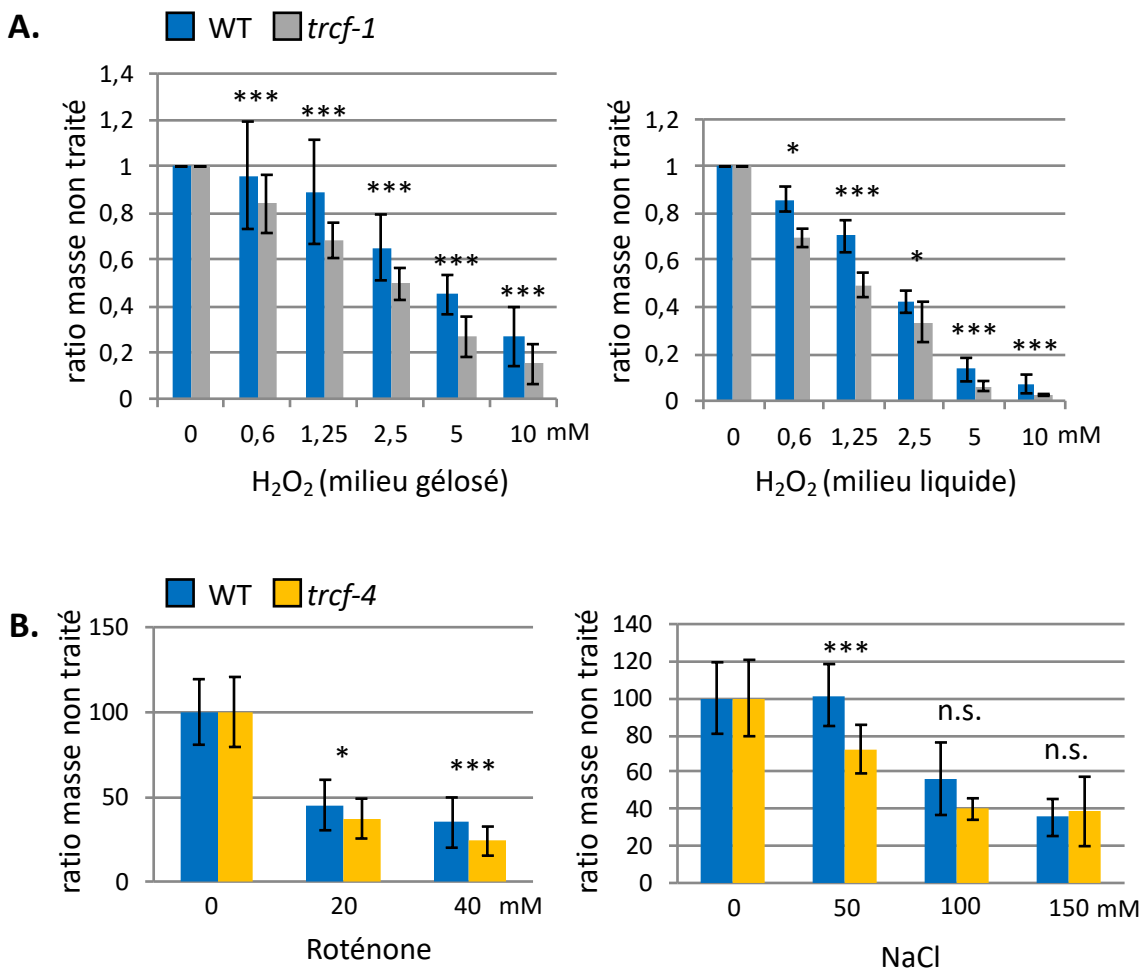
## 5. Sensibilité des mutants à différents stress génotoxiques

Afin de tester l'implication de TRCF dans la réparation de l'ADN, les mutants *trcf* ont été soumis à différents stress génotoxiques. Si TRCF est impliqué dans la réparation de l'ADN comme son homologue bactérien, les plantes mutantes devraient être davantage affectées que les sauvages. Préalablement à ma thèse, des résultats préliminaires suggéraient que le mutant *trcf-1* était plus sensible aux UV-C et au peroxyde d'hydrogène que le témoin, ce qui pointait vers l'implication de TRCF dans la réparation des génomes organellaires (Wallet, 2016). Néanmoins, ces expériences ayant été menées avec des plantes toujours capables de produire une version tronquée de TRCF, elles devaient être reproduites avec le mutant KO *trcf-4*.

Les lésions causées par les UV, déformant l'hélice d'ADN, sont généralement réparées par le NER ou par les photolyases. Pour tester l'éventuelle implication de TRCF dans le NER j'ai soumis des plantules sauvages et *trcf-4* de 14 jours à différentes doses d'UV-C puis j'ai pesé les plantes afin de quantifier leur différence de croissance. Contrairement aux résultats préliminaires trouvés avant ma thèse, les mutants n'ont pas présenté de différence de croissance significative par rapport aux sauvages pour toutes les doses d'UV-C testées (**figure 35**). De plus, je n'ai pas été en mesure de répéter les résultats obtenus pour *trcf-1*. Entre mes



**Figure 35 : Les mutants sont aussi affectés que le sauvage par les traitements aux UV.** Taux de croissance de plantules WT, *trcf-1* et *trcf-4* de 12 jours soumises à des doses croissantes de rayonnement UV. Les masses des plantules traitées sont rapportées à la masse des témoins non traités. Les valeurs pour WT sont représentées en bleu, celles pour *trcf-1* en gris, et celles pour *trcf-4* en jaune.



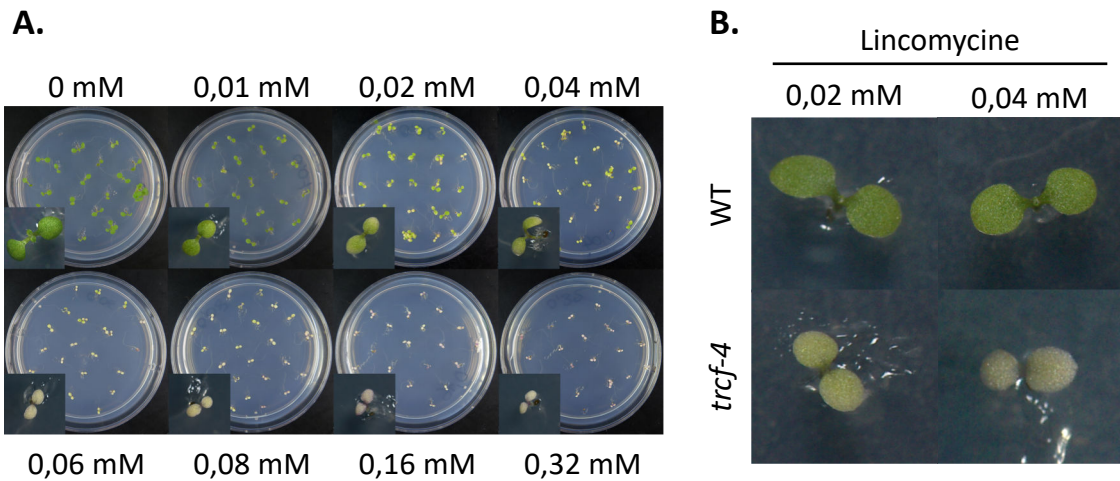
**Figure 36 : Le mutant *trcf-4* n'est pas ou peu affecté par d'autres stress.** **A.** Taux de croissance de plantules WT et *trcf-1* de dix jours exposés au peroxyde d'hydrogène en milieu gélósé (à gauche) ou liquide (à droite). **B.** Taux de croissance de plantules WT et *trcf-4* de 16 jours traitées avec de la roténone (à gauche) ou du NaCl (à droite). Pour chaque expérience, la masse des plantules traitées est rapportée à la masse des témoins non traités. Les valeurs pour WT sont représentées en bleu, celles pour *trcf-1* en gris, et celles pour *trcf-4* en jaune.

mains, ce mutant est aussi sensible aux UV-C que le mutant *trcf-4* ou le sauvage. Les expériences ont été menées en conservant, ou non, les plantes à l'obscurité après le traitement aux UV-C afin de palier à l'éventuelle action de photolyases. Cette différence de protocole n'a entraîné aucun changement significatif des résultats observés. Ces expériences ont été répétées cinq fois indépendamment, sans parvenir à reproduire les résultats observés avant ma thèse. Il est possible que ce soit à cause de différences du matériel biologique, par exemple lié à l'âge et les conditions de conservation des graines. L'absence de sensibilité aux UV des mutants *trcf* laisse penser que TRCF n'est pas impliqué dans la réparation des lésions UV dans les organelles. De façon similaire, le mutant *mfd* bactérien n'est que peu affecté dans sa survie lors de stress aux UV, *Mfd* n'étant pas contrôlé par le système SOS (Witkin, 1956, 1966; Epshtein *et al.*, 2014).

Les lésions oxydatives sont plutôt prises en charge par le BER. L'implication de TRCF dans la réparation de ces lésions avait été testée en cultivant des plantules en présence de peroxyde d'hydrogène, en milieu liquide ou gélosé (**figure 36A**) (Wallet, 2016). Durant ma thèse, je n'ai pas eu l'occasion de répéter ces expériences avec le mutant *trcf-4*. Le lien entre TRCF et les lésions oxydatives a néanmoins été testé avec la roténone, un pesticide bloquant le transfert d'électrons entre le complexe I de la chaîne respiratoire et l'ubiquinone, conduisant à l'accumulation de ROS dans la mitochondrie. Pour cela, les graines ont été directement mises à germer sur un milieu gélosé contenant de la roténone, puis les plantules ont été pesées après 16 jours de croissance. Les résultats indiquent une différence de croissance significative entre les plantes mutants *trcf-4* et les sauvages (**figure 36B**). Comme montré précédemment pour *trcf-1*, la déficience du facteur TRCF augmente la sensibilité de la plante aux dommages oxydatifs. Le facteur TRCF pourrait ainsi être impliqué dans la réparation de ce type de lésions, en lien avec le BER comme CSA et CSB dans la mitochondrie des mammifères (D'Errico *et al.*, 2013).

## 6. Sensibilité du mutant *trcf-4* à d'autres stress abiotiques non génotoxiques

Pour mieux caractériser l'impact de l'absence de TRCF, nous avons également testé la sensibilité du mutant *trcf-4* à deux stress non génotoxiques : le NaCl et la lincomycine. Le NaCl induit un stress osmotique qui est équivalent à un stress hydrique, pouvant affecter de nombreux processus métaboliques (synthèse des lipides et des protéines, photosynthèse). La lincomycine est un antibiotique inhibiteur de la synthèse protéique chloroplastique. Les résultats de Co-IP et de RT-qPCR effectuées (détaillées plus loin dans le manuscrit) nous amenant à penser que TRCF pourrait être impliqué dans la régulation de la transcription chloroplastique, la lincomycine nous a permis de tester si l'absence de TRCF entraînait un stress additionnel au système d'expression des gènes chloroplastiques.



**Figure 37 : Le mutant *trcf-4* est sensible à la lincomycine.** **A.** Des graines de plantes sauvages sont semées sur des boîtes avec une concentration croissante de lincomycine. Les plantules germées de six jours commencent à montrer une dépigmentation des feuilles à partir de 0,04 mM de lincomycine. Les feuilles deviennent majoritairement blanches à partir de 0,06 mM de lincomycine. **B.** Les plantules *trcf-4* de six jours montrent une dépigmentation des feuilles plus importante avec des concentrations plus faibles que pour les plantules sauvages.

Le mutant *trcf-4* présente un défaut de croissance significatif par rapport au sauvage en culture avec 50 mM de NaCl (**figure 36C**), une concentration qui n'affecte quasiment pas la croissance des plantes sauvages. La différence de croissance n'est cependant plus significative pour des concentrations de NaCl plus élevées (100 et 150 mM). Néanmoins, l'aspect général des plantules est plus affecté chez les mutants que chez les sauvages, à toutes les concentrations. La sensibilité au stress osmotique suggère que TRCF est nécessaire au fitness de la plante en conditions sub-optimales de croissance.

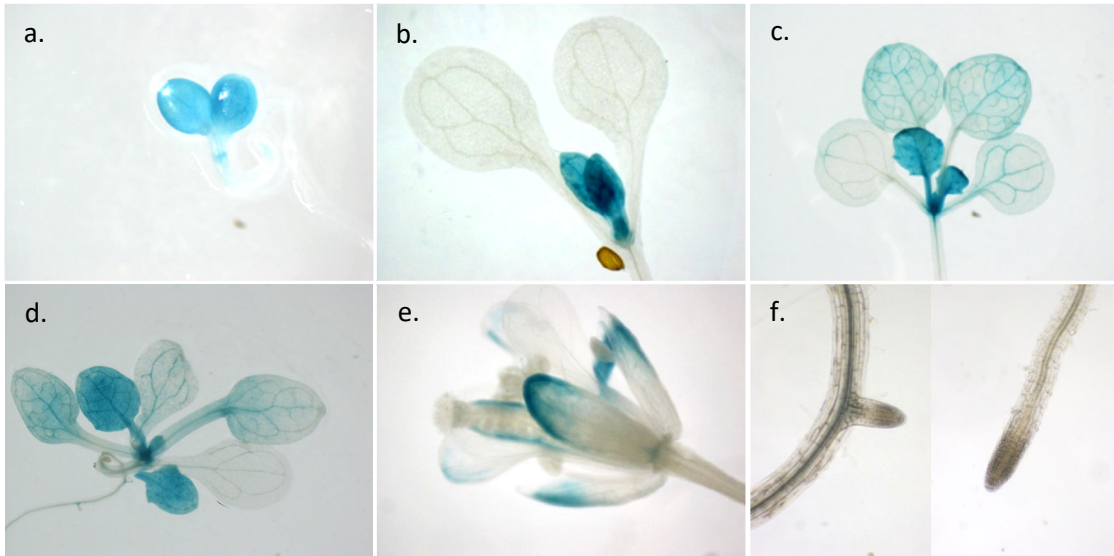
Pour tester la sensibilité du mutant à la lincomycine, j'ai dû définir les concentrations d'antibiotique adéquates (**figure 37A**). Dès 0,08 mM de lincomycine la synthèse protéique chloroplastique devient trop affectée et les plantes présentent un phénotype albinos et un arrêt très précoce de la croissance. Une différence notable de pigmentation est observée sur des plantules cultivées en présence de 0,04 mM de lincomycine : les extrémités des cotylédons blanchissent et la croissance est ralentie. Ces phénotypes s'accroissent avec l'augmentation de la concentration de lincomycine jusqu'à 0,08 mM, où seules quelques rares parties des feuilles de certaines plantules présentent une pigmentation. J'ai ainsi testé les différences de phénotype entre des plantules *trcf-4* et sauvages cultivées *in vitro* sur des milieux gélosés contenant de 0 à 0,08 mM de lincomycine. Les résultats les plus intéressants à 0,02 et 0,04 mM sont présentés dans la **figure 37B**. A ces concentrations les plantules sauvages et mutantes montrent très clairement une différence de pigmentation. Les plantules *trcf-4* sont moins pigmentées à 0,02 mM de lincomycine que ne le sont les sauvages à 0,04 mM. Ainsi, la déficience de TRCF rend plus sensible les plantes soumises à une inhibition de la synthèse protéique plastidiale. L'exacerbation de la dépigmentation des plantules *trcf-4* montre que la synthèse des protéines chloroplastiques est davantage inhibée en absence de TRCF. TRCF pourrait donc jouer un rôle dans la synthèse des protéines ou dans l'expression des gènes, *via* la transcription. Ces expériences de sensibilité à la lincomycine ont été réalisées à la toute fin de la thèse et n'ont donc pas pu être répétées ni être davantage approfondies (en mesurant la quantité de chlorophylle, par exemple).

## 7. Analyse de l'expression de TRCF

### 7.1. Localisation tissulaire de l'expression du promoteur de TRCF

Afin de déterminer les tissus exprimant TRCF au cours du développement d'une plante, le promoteur putatif de TRCF, composé des 381 nucléotides précédents l'ATG initiateur, a été fusionné à la séquence du gène rapporteur *GUS* dans le vecteur binaire pMDC162. Des plantes d'*Arabidopsis* ont ensuite été transformées par la méthode de *flower dip*. Cinq lignées transformées ont été sélectionnées, grâce à leur résistance à la kanamycine, puis testées pour l'expression du gène *GUS* par coloration au X-Gluc. Les résultats observés sont présentés dans la **figure 38** et sont représentatifs de plusieurs lignées de plantes. Sous le contrôle du





**Figure 38 : Profil d'expression tissulaire de TRCF.** Le promoteur de TRCF est fusionné avec la séquence du gène rapporteur *GUS*, conférant une couleur bleue aux tissus où il est exprimé. TRCF est principalement exprimé lors des étapes précoces du développement des tissus chlorophylliens (feuilles (a, b, c et d) et sépales (e)) puis son expression diminue avec le temps. TRCF n'est pas, ou trop faiblement exprimé, dans les tiges et les racines (f). Plantes à 6 (a), 13 (b), 18 (c) et 30 (d) jours ; fleur à 18 jours (e) ; racine et bourgeon racinaire à 18 jours (f).



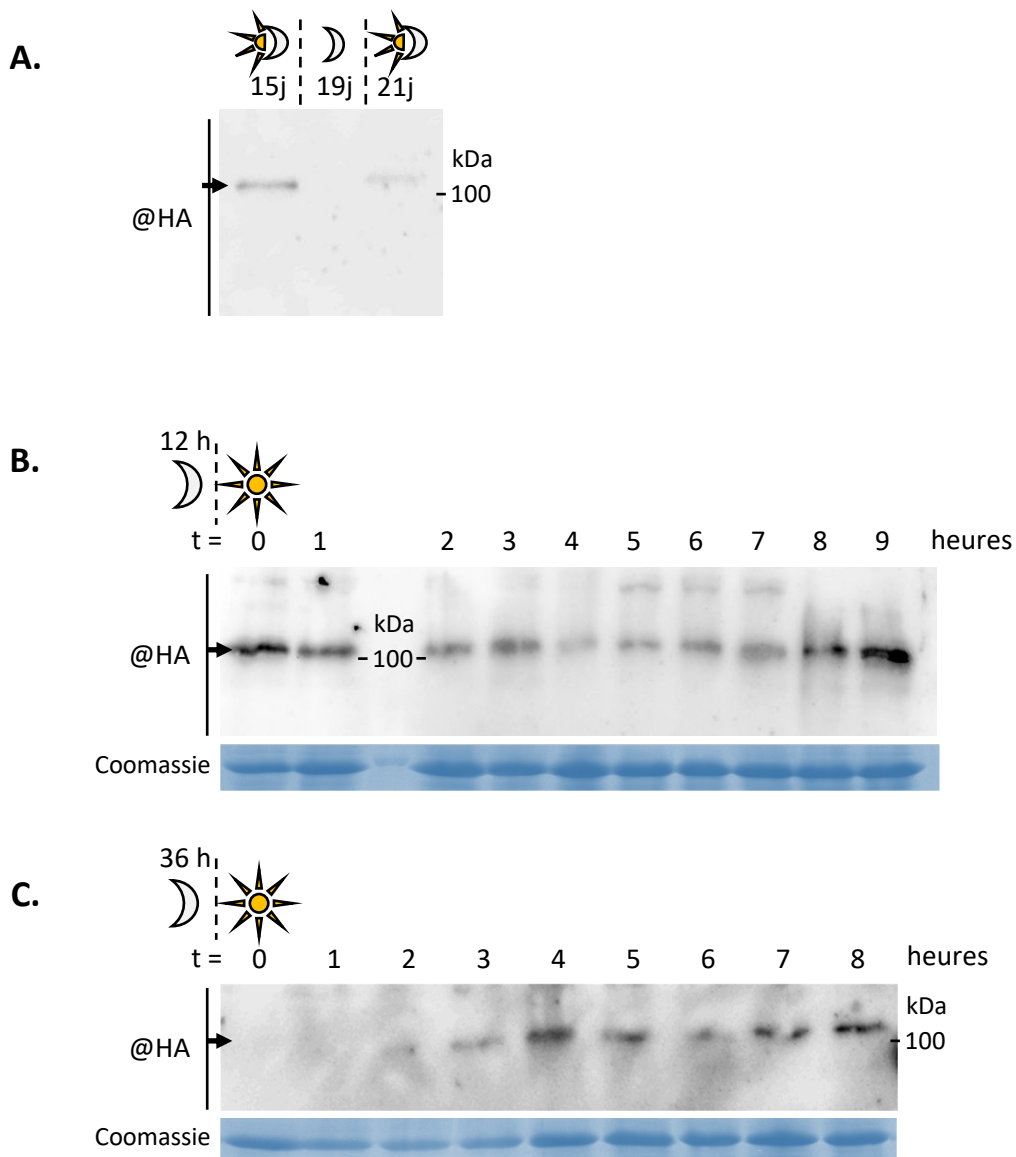
promoteur de TRCF, le gène *GUS* est majoritairement exprimé dans les jeunes tissus chlorophylliens tels que les feuilles, les sépales et les méristèmes apicaux colinéaires. Dans les feuilles, *GUS* est fortement exprimé lors des premiers stades du développement, puis l'expression s'estompe peu à peu dans la feuille vieillissante. En revanche, le promoteur de TRCF ne semble pas actif dans les racines, les fleurs et le système reproducteur, des tissus pourtant riches en mitochondries. Ces résultats laissent penser que TRCF pourrait avoir un rôle dans le développement précoce des tissus chlorophylliens, dont les génomes se répliquent très activement, pouvant entraîner davantage de conflits entre les machineries de transcription et de réplication. L'absence d'expression de *GUS* dans les racines et les fleurs pourrait indiquer un rôle plutôt chloroplastique de TRCF. Il n'est pas rare qu'une protéine adressée aux deux organelles présente une activité prépondérante dans l'un d'eux.

## 7.2. Régulation par la lumière du jour

Lors de l'isolement des chloroplastes, les plantes sont généralement placées à l'obscurité la veille de l'extraction des organelles afin de limiter les granules d'amidon pouvant détruire les fragiles chloroplastes au cours des étapes de purification. En purifiant de la sorte des chloroplastes, j'avais observé une expression excessivement faible de TRCF, aussi bien dans les organelles isolées que dans l'extrait total de plantes utilisées pour l'extraction. Je me suis donc intéressé à l'impact que pouvait avoir le placement à l'obscurité des plantes sur l'expression de TRCF.

Pour cela j'ai utilisé des plantes exprimant la construction TRCF:HA (TRCF fusionné à l'épitope HA en position C-terminale) sous le contrôle du promoteur de TRCF (décrites plus loin). Ces plantes ont été cultivées en conditions normales (photopériode de 16 heures) durant 15 jours après germination, puis à l'obscurité pendant quatre jours, avant de les replacer en conditions normales. J'ai extrait les protéines de plantes à 15 jours avant la mise à l'obscurité, à 19 jours immédiatement à la sortie de la période d'obscurité, et à 21 jours. J'ai ensuite immunodéecté TRCF:HA avec un anticorps @HA dans les différents tissus (**figure 39A**). J'ai ainsi observé une perte d'expression de TRCF:HA à la sortie de la période d'obscurité. L'expression de TRCF:HA s'est ensuite rétablie à la lumière.

Afin d'observer plus précisément l'impact de la lumière sur l'expression de TRCF, j'ai réalisé une cinétique en immunodéectant TRCF:HA toutes les heures à la sortie d'une période d'obscurité de 12 heures (**figure 39B**) ou de 36 heures (**figure 39C**). Lors d'une période d'obscurité courte (12 heures), TRCF:HA est exprimé dès le retour à la lumière, indiquant une expression durant la période d'obscurité. L'expression de TRCF:HA diminue peu à peu durant les quatre premières heures du jour avant d'augmenter à nouveau par la suite. Cette diminution d'expression a été observée lors d'une seconde expérience indépendante. Lors d'une période d'obscurité prolongée (36 heures), TRCF:HA n'est plus exprimé lors du retour



**Figure 39 : L'expression de TRCF est régulée par la lumière.** **A.** L'expression de TRCF:HA est testée dans des plantules de 15 jours cultivées en conditions de culture normales, placées à l'obscurité pendant quatre jours puis replacées en conditions normales. TRCF:HA n'est plus détectée après la mise à l'obscurité des plantes. La protéine est à nouveau détectée lorsque les plantes sont replacées à la lumière. **B.** L'expression de TRCF:HA de plantules de 15 jours est analysée heure après heure à la sortie d'une période de 12 h d'obscurité. TRCF:HA est exprimée immédiatement à la sortie de la période d'obscurité. Son expression diminue légèrement après quatre heures de lumière, puis augmente à nouveau. **C.** Après une période d'obscurité prolongée de 36 h, la protéine TRCF:HA n'est plus exprimée. L'expression de la protéine augmente peu à peu lorsque la plante est mise à la lumière.

des plantes à la lumière, indiquant un arrêt de l'expression de TRCF:HA durant la phase d'obscurité. L'expression de TRCF:HA augmente ensuite peu à peu, devenant à nouveau détectable au bout de deux heures de jour.

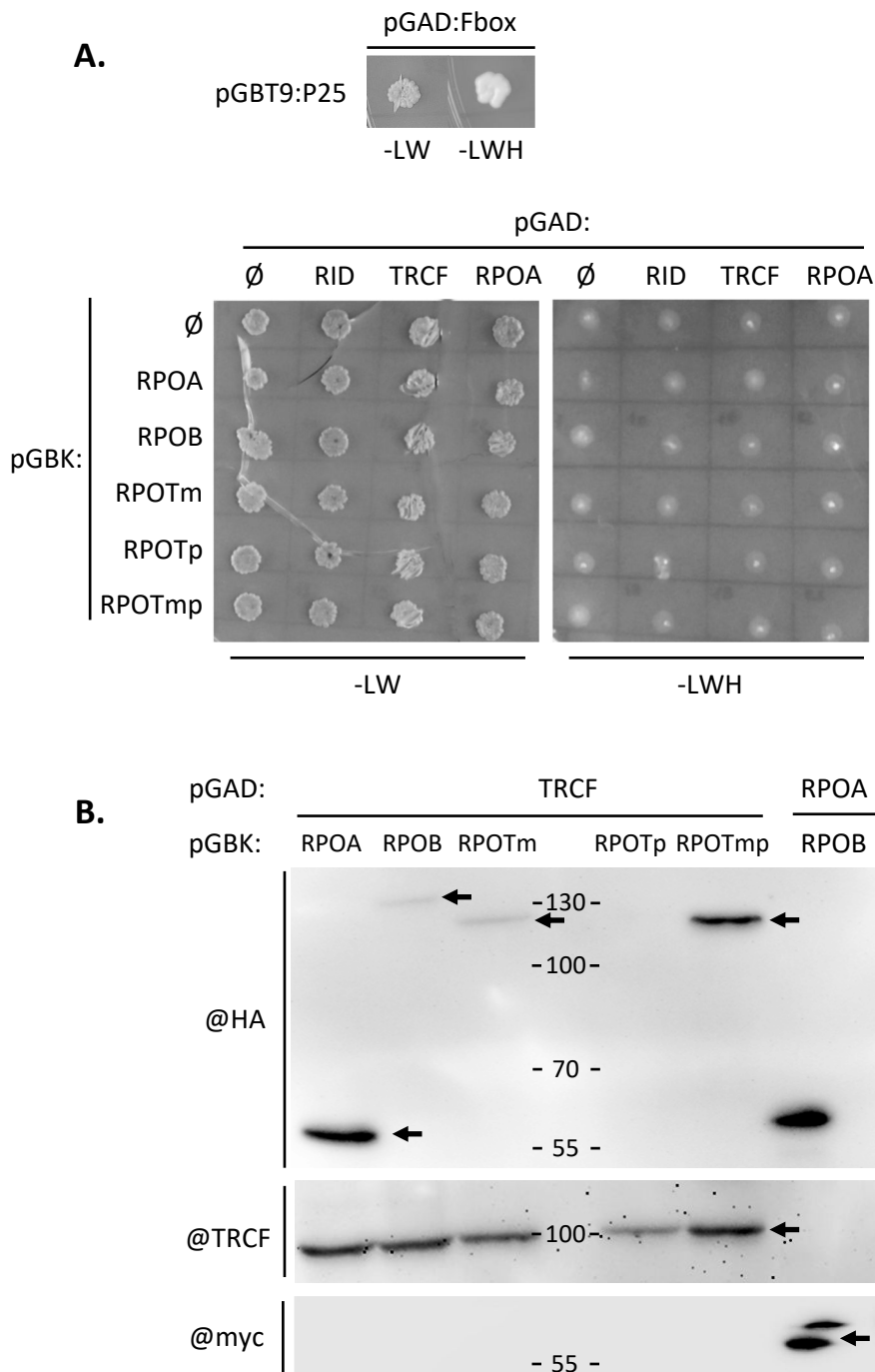
Ces résultats indiquent que l'expression de TRCF est régulée par la lumière du jour, comme de nombreux gènes impliqués dans la biogenèse et l'expression des gènes du chloroplaste. L'activité de la PEP est également modulée par le rythme circadien à travers les facteurs sigma qu'elle utilise. En absence de lumière, la PEP s'associe moins fréquemment avec ses promoteurs, notamment ceux des ARNr et ARNt (Finster *et al.*, 2013). Il est alors possible que TRCF ne soit pas nécessaire en absence de la PEP qu'il pourrait réguler.

## 8. Recherche des partenaires d'interaction de TRCF

Les tests de sensibilités aux stress avec les mutants *trcf* n'ayant pas donné de pistes quant aux rôles de TRCF dans les organelles, j'ai entrepris d'étudier les partenaires d'interaction de TRCF afin de déterminer dans quels mécanismes il pourrait être impliqué. Pour cela, j'ai réalisé une approche par Co-IP suivie d'identification des protéines retenues par spectrométrie de masse, afin d'obtenir une liste d'interactants potentiels de TRCF sans *a priori*. Parallèlement, j'ai effectué un double hybride de levure (Y2H : *yeast two hybrids*) pour tester l'interaction entre TRCF et toutes les ARN polymérases des organelles, partenaires d'interaction les plus probables. Enfin, j'ai tenté de confirmer les interactions déterminées par Co-IP par complémentarité bimoléculaire de la fluorescence (BiFC).

### 8.1. Etude de l'interaction entre TRCF et les ARN polymérases organellaires

Dans l'attente des plantes complémentées permettant l'étude plus poussée des partenaires de TRCF par Co-IP, j'ai entrepris de tester l'interaction de TRCF avec les sous-unités  $\alpha$  et  $\beta$  de la PEP et les trois NEP. Le domaine D4/RID semblant relativement conservé, et l'activité principale de TRCF étant probablement liée aux ARN polymérases, l'interaction entre TRCF et la PEP et/ou les NEP semblait être plutôt probable dans les organelles d'*Arabidopsis*. J'ai ainsi cloné les séquences du gène d'*Arabidopsis* de TRCF ou de son domaine D4/RID seul dans le vecteur pGBK, et les séquences de RPOA, RPOB, RPOTm, RPOTp et RPOTmp dans le vecteur pGAD. J'ai également cloné la séquence du gène RPOA dans pGBK afin d'utiliser l'interaction entre RPOA et RPOB comme témoin positif. Les deux autres sous-unités de la PEP (RPOC1 et RPOC2) n'ont pas été testées pour limiter le nombre de clonages initiaux. De plus, chez la bactérie, le domaine RID de TRCF interagit avec la sous-unité  $\beta$ , correspondant à la protéine RpoB de la PEP. L'interaction avec les autres sous-unités était donc moins probable. Afin d'avoir un témoin positif ayant déjà été testé, j'ai également récupéré les constructions pGAD:Fbox et pGBT9:P25 de l'équipe de David Gilmer à l'IBMP (Thiel *et al.*, 2012).



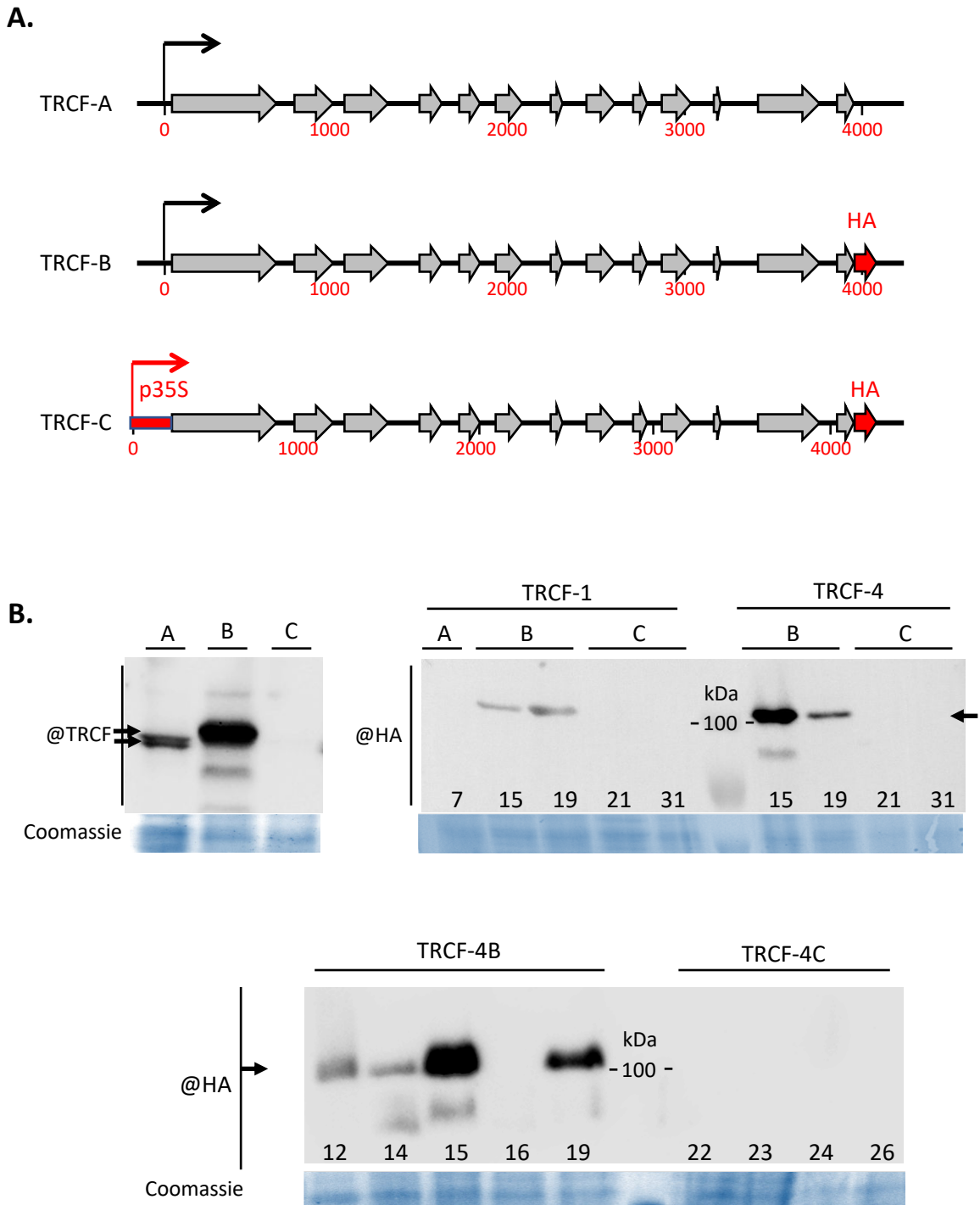
**Figure 40 : Analyse de l'interaction de TRCF avec les ARN polymérases organellaires par double hybride de levure. A.** Les hybrides diploïdes obtenus par croisements entre les souches AH109 transformée avec le plasmide pGAD et Y187 transformée avec le plasmide pGBK sont déposés sur des milieux dépourvus de leucine et tryptophane (-LW) ou dépourvus de leucine, tryptophane et histidine (-LWH). En haut, le témoin positif croît sur le milieu -LWH indiquant l'interaction entre P25-Fbox. En bas, les hybrides croissent tous sur milieu -LW mais aucun sur milieu -LWH, indiquant l'absence d'interaction entre TRCF, ou son domaine RID seul, avec les trois NEP et deux sous-unités de la PEP. Dans ces conditions, il n'y a pas d'interaction entre RPOA et RPOB, deux sous-unités de la PEP. **B.** L'immunodétection des protéines chez les hybrides révèle pourtant l'expression des protéines RPOA:HA (55 kDa), RPOB:HA (135 kDa), RPOTm:HA (110 kDa), RPOTmp:HA (111 kDa), TRCF:myc (100 kDa) et RPOA:myc (55 kDa).

Après de laborieuses étapes de clonage, j'ai pu obtenir tous les hybrides diploïdes de levure capable de croître sur milieu SD-LW (**figure 40**). Cependant, aucun des hybrides obtenus ne s'est révélé capable de croître sur milieu SD-LWH, indiquant l'absence d'interaction entre les protéines testées (**figure 40**). Entre mes mains, ni TRCF ni le domaine D4/RID seul n'interagissent avec RPOA, RPOB ou l'une des trois NEP dans la levure par le système Y2H. Cependant RPOA et RPOB qui devraient être un contrôle positif n'interagissent pas non plus dans les mêmes conditions expérimentales. Seules les deux constructions témoins pGAD:Fbox et pGBT9:P25 validées auparavant par une autre équipe ont permis aux levures de pousser sur milieu SD-LWH. Pour comprendre ces résultats, j'ai vérifié par western-blot l'expression de toutes les protéines de fusion dans les hybrides de levure en utilisant les anticorps spécifiques de TRCF ou spécifiques d'étiquettes fusionnées aux protéines. Les protéines clonées sont fusionnées avec l'étiquette myc dans pGAD ou avec l'étiquette HA dans pGBK. J'ai ainsi pu détecter toutes les protéines aux tailles attendues, excepté RPOTp, indiquant l'expression des différentes protéines de fusion. L'absence de croissance des hybrides sur milieu SD-LWH semble bien due à l'absence d'interaction entre les différentes protéines. Néanmoins, il est possible que les protéines de fusion exprimées ne se replient pas correctement dans la levure, empêchant l'interaction. Il est également possible que certains facteurs d'Arabidopsis soient requis pour permettre l'interaction entre les protéines, ou que l'interaction entre TRCF et les ARN polymérases organellaires soit indirecte. L'interaction entre RPOA et RPOB pourrait nécessiter la présence des autres sous-unités de la PEP, ou de facteurs favorisant l'assemblage absents chez la levure. L'approche préliminaire ne donnant pas les résultats escomptés, la compréhension des fonctions de TRCF reposera donc essentiellement sur la Co-IP de ses interactants.

## 8.2. Recherche des partenaires d'interaction par Co-IP et spectrométrie de masse.

### 8.2.1. Complémentation des mutants *trcf* par différentes versions de TRCF

Afin d'effectuer de Co-IP de bonne qualité des partenaires d'interaction de TRCF, j'ai produit différentes constructions de TRCF fusionné avec une étiquette HA C-terminale. Cette étiquette permet de précipiter de façon très spécifique TRCF:HA à l'aide d'anticorps monoclonaux dirigés contre l'épitope HA, alors que les anticorps polyclonaux contre TRCF produits sont trop peu spécifiques pour des expériences de Co-IP. J'ai ainsi cloné TRCF dans les vecteurs binaires pGWB13 et pGWB14 afin de transformer les mutants homozygotes *trcf-1* et *trcf-4*. J'ai utilisé pGWB13 pour obtenir les constructions TRCF-A, permettant l'expression de TRCF sous son promoteur endogène, et la construction TRCF-B, permettant l'expression d'une protéine TRCF fusionnée en C-terminal avec l'étiquette HA (**figure 41A**). Le vecteur pGWB14 m'a permis d'obtenir la construction TRCF-C, exprimant TRCF:HA sous contrôle du promoteur 35S du CaMV (*Cauliflower mosaic virus*), très fort et actif dans la majorité des tissus (**figure 41A**).



**Figure 41 : Complémentation des mutants *trcf-1* et *trcf-4* par différentes versions de TRCF.** **A.** Les mutants ont été complétés avec trois constructions : TRCF-A avec ses promoteur et terminateur endogènes, TRCF-B avec son promoteur endogène et une fusion C-terminale de l'étiquette HA, et TRCF-C sous contrôle du promoteur 35S et avec une fusion C-terminale de l'étiquette HA. **B.** L'expression de TRCF est analysée avec des anticorps reconnaissant TRCF (@TRCF) ou l'étiquette HA (@HA). La complémentation avec TRCF-C ne permet pas une bonne expression de TRCF, sans doute dû à l'interférence par l'ARN liée à la force du promoteur 35S. La complémentation avec TRCF-B entraîne la reconnaissance d'une protéine de taille supérieure car fusionnée avec l'étiquette HA (101,6 kDa pour TRCF:HA contre 93,7 kDa pour TRCF). La complémentation avec la construction TRCF-A ne peut pas être détectée avec l'anticorps reconnaissant l'étiquette HA.

La fusion de l'étiquette a été préférée en C-terminal pour ne pas modifier l'adressage de TRCF en modifiant sa séquence N-terminale. La construction TRCF-A devait servir de témoin de complémentation en permettant d'observer une éventuelle disparition de phénotypes développementaux liés à la mutation *trcf*. La sensibilité à la lincomycine et la dépigmentation des jeunes feuilles ayant été observées à la fin de la thèse, les plantes TRCF-A n'auront finalement pas servi de témoin car aucun phénotype n'était observé lors de leur obtention.

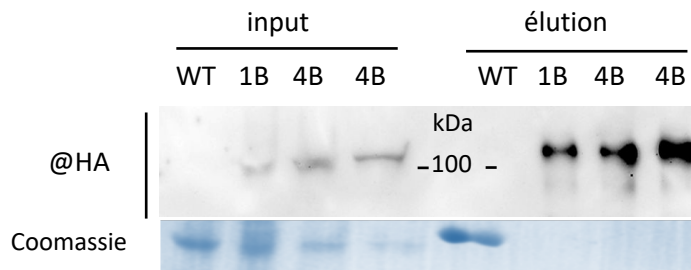
J'ai utilisé ces trois constructions pour transformer les mutants *trcf-1* et *trcf-4*, obtenant ainsi les plantes complémentées TRCF-1A, TRCF-1B et TRCF-1C ou TRCF-4A, TRCF-4B et TRCF-4C, en fonction du mutant transformé et de la construction utilisée. Après génotypage et sélection des plantes transformées, j'ai testé par *western blot* l'expression de TRCF afin de sélectionner les plantes exprimant le mieux TRCF pour les expériences de Co-IP. J'ai utilisé les anticorps polyclonaux @TRCF spécifiques de TRCF, ou des anticorps monoclonaux commerciaux spécifiques de l'étiquette HA (**figure 41B**). Ainsi, j'ai pu observer l'expression de TRCF dans les plantes témoins TRCF-1A et TRCF-4A, à la taille attendue de 94 kDa, et l'expression de la protéine de fusion TRCF:HA de 102 kDa dans les plantes TRCF-1B et TRCF-4B (**figure 41B**). L'expression de TRCF:HA varie d'une plante à l'autre, allant de l'absence d'expression détectable (plante TRCF-4B16) à une très forte expression (plante TRCF-4B15). En revanche, je n'ai jamais détecté une expression satisfaisante dans les plantes complémentées avec la construction TRCF-C, dans toutes les plantes testées (**figure 41B**). Cette absence d'expression provient peut-être d'une expression initiale trop importante conduisant à la dérégulation du transcrit de TRCF:HA par interférence par l'ARN. Ce mécanisme de défense contre les virus est particulièrement développé chez les plantes (Baulcombe, 2004). Son utilisation contre une construction utilisant un promoteur viral nous apparaît donc probable. De la même façon, la plante transformée TRCF-4B15 (**figure 41B**), qui exprimait le plus la protéine de fusion, a fini par ne plus exprimer du tout la protéine de façon détectable par western-blot.

En l'absence de phénotype observable pour les mutants *trcf*, nous ne nous attendions pas à voir une quelconque modification du développement de plantes complémentées. Néanmoins, certaines plantes transformées par la construction TRCF-B ont montré un retard de croissance et le jaunissement de leurs premières feuilles (**figure 47**). L'étude de ce phénotype est traitée plus loin.

### 8.2.2. Co-IP des partenaires d'interaction

- Approches préliminaires

Après l'obtention des mutants complémentés, j'ai utilisé les descendants des lignées TRCF-4B14 et TRCF-1B19 qui présentaient une bonne expression de TRCF:HA. J'ai réalisé la



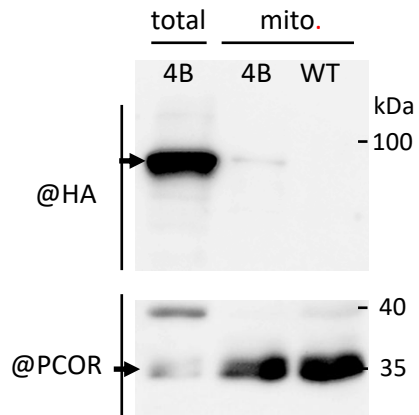
**Figure 42 : Immunodétection de TRCF:HA après co-immunoprécipitation.** Pour analyser les interactants de TRCF, une approche par co-immunoprécipitation des partenaires de TRCF est réalisée sur des extraits totaux de plantes à 14 jours. La présence et l'enrichissement de TRCF:HA dans les fractions d'élution sont testés par *western blot*. La protéine TRCF:HA est largement enrichie après la Co-IP, aussi bien pour des plantes TRCF-1B (1B) que TRCF-4B (4B).



totalité de mes Co-IP avec des plantes hétérozygotes pour l'insertion de TRCF:HA. J'ai commencé par tester mon protocole pour vérifier si je parvenais bien à enrichir TRCF:HA dans mes expériences. TRCF étant majoritairement exprimé dans les jeunes tissus chlorophylliens, j'ai utilisé des plantules de 14 jours exprimant suffisamment TRCF pour permettre des Co-IP de bonne qualité.

J'ai ainsi pu observer la protéine cible TRCF:HA dans les fractions d'élution après Co-IP, attestant du bon fonctionnement de l'expérience (**figure 42**). En effet, les Co-IP étant réalisées en conditions natives afin de conserver les interactions entre protéines, l'étiquette nécessaire à la purification de la protéine cible pourrait se retrouver masquée et ne plus être reconnue par les anticorps. La coloration au bleu de Coomassie de la membrane ne permet pas d'observer de protéine dans les fractions d'élution, étant donné que la grande majorité des protéines n'est pas retenue dans la colonne, conduisant à l'enrichissement de TRCF:HA dans les fractions d'élution.

Les fractions éluées ont ensuite été analysées par spectrométrie de masse afin d'identifier les protéines co-précipitées avec TRCF:HA. Ces résultats sont résumés dans le *volcano plot* décrit plus loin car incluant d'autres expériences également décrites plus loin (**figure 45**). Les premières expériences menées sur des extraits totaux n'ayant pas donné de résultats très satisfaisants, j'ai jugé plus intéressant de ne pas les présenter. Dans mes premières Co-IP, le nombre de spectres de TRCF:HA, correspondant aux nombres de protéines TRCF:HA présentes dans les fractions éluées était bas (variant de 50 à 100 spectres, pour une moyenne de 73 spectres pour 7 expériences de Co-IP). De fait, le nombre de spectres des partenaires co-purifiés étaient également plutôt bas. Parmi les potentiels partenaires uniquement co-précipités avec TRCF:HA, le nombre de spectres le plus élevé était de 16. En moyenne, les protéines uniquement co-purifiées avec TRCF:HA ne présentent que deux spectres par protéine, soit des nombres extrêmement bas et peu significatifs, ne donnant que peu de poids aux résultats obtenus. En analysant statistiquement l'ensemble des protéines éluées (y compris celles également éluées dans les fractions témoins), les partenaires d'interaction les plus probables (déterminés par l'algorithme utilisé) n'ont pas permis d'appréhender précisément les fonctions de TRCF. Ces partenaires d'interaction les plus probables (présentant le plus haut score) appartenaient à des nombreuses catégories fonctionnelles apparemment non liées. Une part importante des protéines identifiées semblait néanmoins avoir un rôle dans le métabolisme des acides nucléiques, ou possédait des domaines de liaison à l'ARN ou à l'ADN. Mais ces protéines étaient localisées dans plusieurs compartiments cellulaires autres que la mitochondrie et le chloroplaste (membrane plasmique, noyau, cytosol, appareil de Golgi, etc.), ne faisant que peu de sens dans le cadre d'une interaction avec TRCF:HA. Ces résultats laissent penser que TRCF n'est pas impliqué dans un complexe, mais interagit plutôt transitoirement avec ses partenaires.



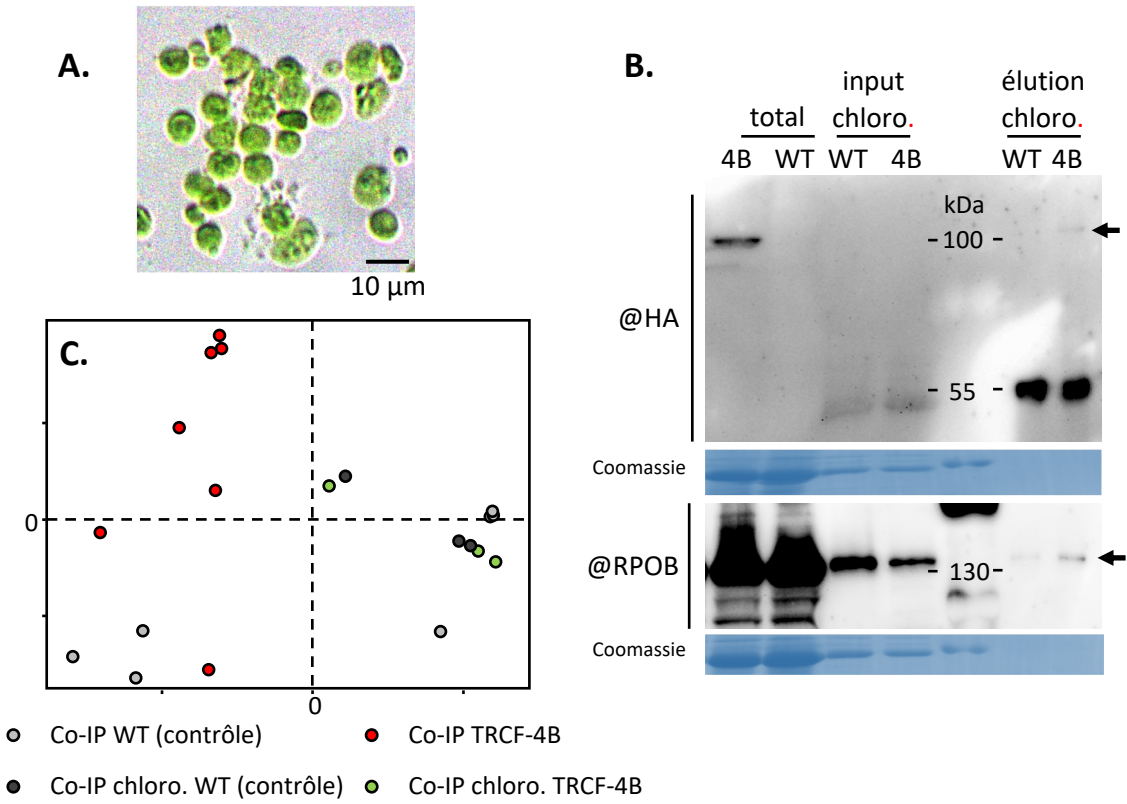
**Figure 43 : Immunodétection de TRCF dans des mitochondries isolées.** L'expression de TRCF:HA est testée dans les purifications de mitochondries réalisées à partir de plantules étiolées de neuf jours. Le témoin (total) correspond à des extraits totaux de plantules de neuf jours cultivées en conditions standard. L'expression de TRCF:HA est particulièrement faible dans la fraction mitochondriale (en haut, seconde colonne) en comparaison avec la fraction totale. De plus, cette expression pourrait provenir d'une contamination chloroplastique, détectée par l'anticorps @PCOR spécifique d'une protéine du chloroplaste.

Les quatre sous-unités de la PEP ainsi que plusieurs PAP sont retrouvées dans les fractions d'élution avec TRCF:HA. La sous-unité  $\alpha$  (RPOA) était la plus enrichie, étant dix fois plus représentée dans les échantillons avec TRCF:HA que dans les IP témoins. L'analyse statistique du nombre de spectre des sous-unités de la PEP et des PAP n'indiquait cependant pas une présence significative de ces protéines avec TRCF:HA, RPOA n'arrivant qu'à la 60<sup>e</sup> position des interactants les plus probables. Aucune NEP n'a été identifiée dans les fractions éluées, confortant plus encore les résultats obtenus par Y2H. Néanmoins, les NEP ne sont présentes qu'en très faible quantité dans chaque organelle, et leur expression diminue avec le vieillissement des tissus. Aussi, il est probable que les plantes utilisées soient déjà trop âgées pour permettre leur co-précipitation, en particulier avec un si faible nombre de spectres pour TRCF:HA.

Au vu de ces résultats préliminaires, il paraissait nécessaire d'optimiser les conditions de Co-IP afin d'augmenter le nombre de spectres de TRCF:HA et de limiter les contaminants provenant de l'extérieur des mitochondries et des chloroplastes. Pour cela, l'isolement des organelles d'intérêt avant Co-IP permettrait d'enrichir davantage TRCF:HA dans la fraction initiale tout en supprimant la grande majorité des contaminants non-organellaires.

- Isolement des mitochondries

Les mitochondries ont été isolées à partir de plantules de neuf jours de la lignée TRCF-4B14, étioilées pour limiter le matériel chloroplastique, selon le protocole indiqué dans les méthodes. L'immunodétection de TRCF:HA dans les mitochondries isolées n'a révélée qu'une quantité extrêmement faible de la protéine (**figure 43**). La quantité de TRCF:HA dans les mitochondries purifiées est nettement inférieure à celle retrouvée chez des extraits totaux de plantes aux mêmes stades de développement, cultivées en conditions normales de culture. Cette quasi-absence de TRCF:HA provient certainement de la dérégulation de l'expression de TRCF à l'obscurité, expliquée précédemment. Néanmoins, au moment de ces expériences la régulation de l'expression de TRCF par la lumière n'avait pas été testée, et ces résultats avaient été interprétés comme preuve du faible adressage de TRCF aux mitochondries. Cette interprétation était alors appuyée par la faible fluorescence mitochondriale obtenue avec les constructions OTS<sub>TRCF</sub>:GFP, par la faible expression des constructions pTRCF:GUS dans certains tissus riches en mitochondries, et l'interaction plus probable de TRCF avec une RNAP de type bactérienne, uniquement présente dans le chloroplaste. Comme contrôle de contamination plastidiale des mitochondries purifiées j'ai utilisé un anticorps dirigé contre PCOR (protochlorophyllide oxidoreductase), une protéine chloroplastique, Le niveau de détection de PCOR dans l'extrait mitochondrial était consistant avec l'idée que le faible niveau de protéine TRCF:HA détecté était du bruit de fond de contamination plastidiale.



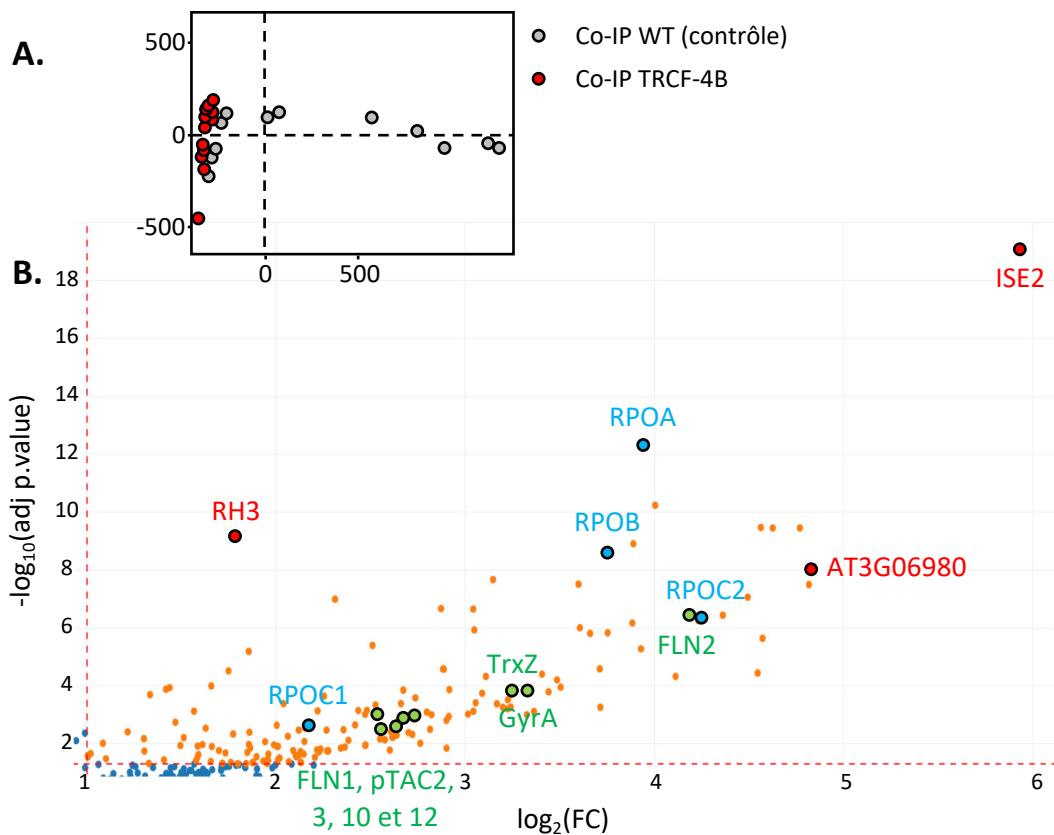
**Figure 44 : La purification des chloroplastes ne permet pas de réaliser d'assez bonnes co-immunoprécipitation des partenaires de TRCF.** **A.** L'intégrité des chloroplastes isolés a été évalué par microscopie. **B.** L'immunodétection de TRCF:HA ne permet pas de détecter la protéine dans la purification de chloroplastes alors qu'elle est clairement détectée dans l'extrait total de plantes de 14 jours ayant servi à isoler ces chloroplastes. La co-immunoprécipitation réalisée avec les chloroplastes isolés permet tout de même de détecter TRCF:HA en faible quantité. La protéine RPOB est enrichie dans la Co-IP avec TRCF:HA, indiquant une potentielle interaction entre les deux protéines. **C.** Le *MDS-plot* réalisé à partir des expériences menées sur des extraits totaux de plantes (points gris et rouges) et sur des chloroplastes isolées (points noirs et verts) révèle d'importantes variations entre les deux conditions.

J'ai également tenté de purifier les mitochondries à partir de fleurs de plantes de la lignée TRCF-4B14, afin d'utiliser des conditions de cultures normales et d'éviter les longues périodes d'obscurité. Mais ces purifications ont échoué et je n'ai pas pu utiliser les extraits mitochondriaux pour effectuer mes Co-IP. A la place, j'ai préféré essayer des extraits de chloroplastes purifiés, davantage susceptibles de contenir TRCF.

- Isolement des chloroplastes

Les premières purifications de chloroplastes, plaçant les plantes à l'obscurité pour limiter les granules d'amidon, ne permettaient pas de détecter TRCF:HA. Après l'observation de la régulation de l'expression de TRCF par la lumière, conduisant à l'absence d'expression lors de placement prolongé des plantes à l'obscurité (**figure 39**), j'ai réalisé de nouvelles purifications des chloroplastes sans placer les plantes à l'obscurité. Afin d'obtenir suffisamment de matériel chlorophyllien, j'ai utilisé des plantules de 18 jours de la lignée TRCF-4B14 cultivées *in vitro* dans les conditions de culture normales. Malgré la présence d'amidon, mes purifications ont permis d'isoler des chloroplastes intacts qui permettaient de détecter TRCF en quantités similaires à un extrait total de plantes (**figure 44A**). Néanmoins, je n'ai pu isoler qu'une quantité faible de chloroplastes intacts, juste suffisante pour réaliser mes Co-IP. Ainsi, après Co-IP, TRCF est détecté bien plus faiblement par *western blot* qu'il ne l'avait été en utilisant des extraits totaux de plantes (**figure 44B**). La diminution de la quantité de TRCF:HA dans les Co-IP a manifestement mené à la diminution du nombre de spectres correspondant à TRCF. Je suis ainsi passé de 50-100 spectres dans les extraits totaux à 5-22 spectres (15 spectres en moyenne) dans les extraits de chloroplastes. En conséquence, le nombre de spectres des protéines co-purifiées était également nettement plus bas, rendant difficile l'exploitation des résultats. En outre, les résultats obtenus variaient significativement d'une expérience à l'autre, comme le montre le MDS-plot (*multidimensional scaling plot*) représenté dans la **figure 44C**. Les protéines identifiées variaient trop peu entre les échantillons témoins (sans TRCF:HA) et les échantillons exprimant TRCF:HA. Ces résultats étaient également aberrants avec ceux obtenus avec les extraits totaux de plantes, se groupant davantage avec les expériences témoins qu'avec celles exprimant TRCF:HA.

Ces expériences réalisées avec les chloroplastes isolées n'ont ainsi pas permis d'appuyer les résultats précédemment obtenus avec les extraits totaux. Ils n'ont permis non plus de mieux appréhender les partenaires d'interaction de TRCF, les interactants obtenus étant systématiquement différents d'une expérience à l'autre. Ces problèmes pourraient provenir de la purification des chloroplastes en elle-même, trop longue, qui entraînerait la perte des interactants associés à TRCF. Si les interactions de TRCF avec ses partenaires sont plutôt transitoires, les manipulations nécessaires à la purification des chloroplastes pourraient facilement les déstabiliser, résultant à des profils d'interaction biaisés et proches des expériences témoins. De plus, la faible quantité de chloroplastes intacts récupérée par un



**C.**

Nom	AGI	Localisation	fonction
AT3G06980	AT3G06980	Cp	hélicase ARN – rôle inconnu
FLN1/PAP6	AT3G54090	Cp	régulation de la transcription
FLN2	AT1G69200	Cp	régulation de la transcription
GyrA	AT3G10690	Cp	réplication
ISE2	AT1G70070	Cp	hélicase ARN – épissage
PTAC10/PAP3	AT3G48500	Cp	métabolisme ADN/ARN
pTAC12/PAP5	AT2G34640	Cp	régulation de la transcription
pTAC2/PAP2	AT1G74850	Cp	métabolisme ADN/ARN
pTAC3/PAP1	AT3G04260	Cp	métabolisme ADN/ARN
RH3	AT5G26742	Cp	hélicase ARN – épissage
RPOA/PEP $\alpha$	ATCG00740	Cp	transcription
RPOB/PEP $\beta$	ATCG00190	Cp	transcription
RPOC1/PEP $\beta'$	ATCG00180	Cp	transcription
RPOC2/PEP $\beta''$	ATCG00170	Cp	transcription
TrxZ/PAP10	AT3G06730	Cp	régulation de la transcription

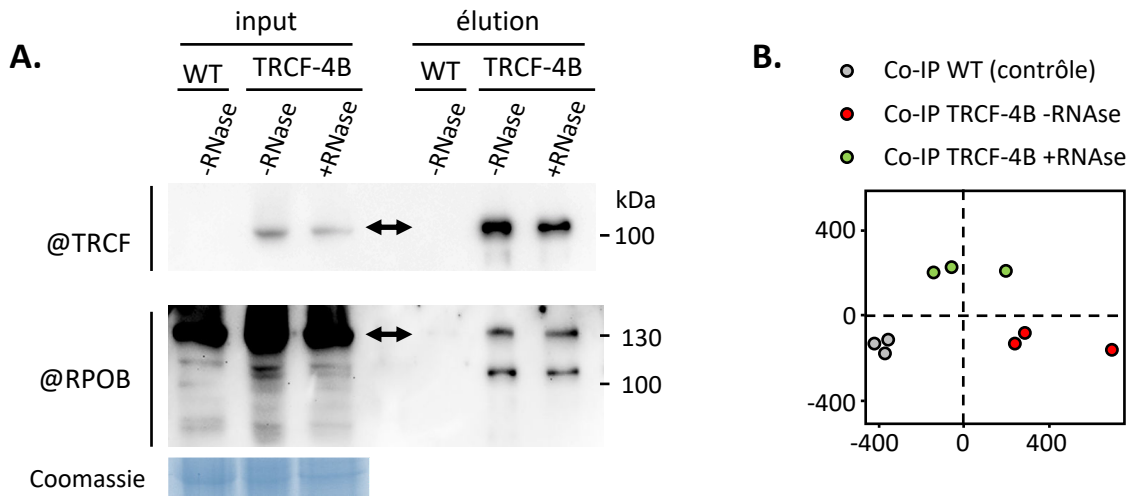
**Figure 45 : TRCF pourrait interagir avec la PEP et les PAP.** **A.** Le *MDS-plot* montre le regroupement des résultats des Co-IP de TRCF:HA réalisés avec des extraits totaux, indiquant une bonne reproductibilité. Les résultats obtenus pour les contrôles sauvages (WT) sont plus variables, en revanche. **B.** Les résultats de 12 Co-IP contrôles et 12 Co-IP TRCF-4B sont représentés sous forme d'un *volcano-plot*. Les interactants les plus intéressants sont indiqués et nommés. **C.** Leur numéro d'accèsion, leur localisation et leur fonction principale sont indiqués dans le tableau. De nombreuses autres protéines dont l'interaction avec TRCF ne fait que peu de sens sont également identifiées. Un grand nombre de ces protéines possède un domaine de liaison avec l'ADN ou l'ARN ou sont liées au métabolisme des acides nucléiques. Ces protéines proviennent de tous les compartiments cellulaires.

protocole ne pouvant se défaire des grains d'amidon limite davantage encore les possibilités d'isoler des interactants transitoires.

J'ai néanmoins profité de l'obtention d'anticorps polyclonaux spécifiques de la protéine RPOB d'*Arabidopsis* pour tester la présence et l'enrichissement de la protéine dans les fractions d'éluion de la Co-IP (**figure 44B**). J'ai ainsi observé la présence plus importante d'une protéine de la taille de RPOB dans l'échantillon exprimant TRCF:HA par rapport au témoin, indiquant qualitativement la possible interaction de TRCF avec RPOB. Malgré l'impossibilité de tirer des conclusions à partir des expériences réalisées avec les chloroplastes isolés, RPOB a néanmoins attiré notre attention en étant la seule protéine systématiquement identifiée dans les échantillons contenant TRCF:HA. Ainsi, malgré le faible nombre de spectres obtenus pour RPOB, la répétabilité de son identification suggère son interaction avec TRCF. En revanche, les trois autres sous-unités de la PEP et les PAP ne sont pas systématiquement observées. Comme pour les expériences précédentes, aucune NEP n'a été identifiée parmi les protéines identifiées.

- Utilisation d'extraits totaux.

Les Co-IP réalisées sur les organelles isolées n'apportant pas les résultats escomptés, j'ai recommencé avec des extraits totaux de plantes. Pour cela j'ai utilisé des plantules de 14 jours des lignées TRCF-1B19 et TRCF-4B14. J'ai ainsi obtenu des expériences plus reproductibles (**figure 45A**) me permettant d'analyser les protéines identifiées par une approche statistique, malgré un nombre de spectres toujours réduit pour TRCF (23 à 60 spectres). Les résultats obtenus sont modélisés sous la forme d'un *volcano-plot*, représentant l'enrichissement des protéines dans la condition TRCF:HA (par rapport au témoin) en fonction de la reproductibilité du résultat (**figure 45B**). Ce *volcano-plot* a été construit en utilisant les résultats de quatre expériences de Co-IP, comprenant chacune trois réplicas techniques pour les conditions témoins et en présence TRCF:HA. Ces résultats montrent que les quatre sous-unités de la PEP ainsi que six PAP et deux pTAC sont significativement enrichies en présence de TRCF:HA, appuyant notre hypothèse initiale d'une interaction entre TRCF et la PEP. Ces résultats sont également appuyés par le fait que RPOB soit la seule protéine qui ait été systématiquement retrouvée dans les Co-IP. Trois hélicases à ARN chloroplastiques, dont deux impliquées dans l'épissage de transcrits (ISE2 et RH3), sont également significativement enrichies. ISE2 est d'ailleurs le partenaire d'interaction présentant le score le plus élevé, selon l'approche statistique, indiquant qu'il est le partenaire d'interaction le plus probable de TRCF. Néanmoins, de nombreuses protéines interagissant avec l'ARN ou l'ADN, dont l'adressage n'est ni mitochondrial, ni chloroplastique sont toujours identifiées et témoignent certainement d'une co-précipitation avec les acides nucléiques. De plus, un grand nombre des protéines co-purifiées identifiées semblent être des contaminants, donc l'interaction avec TRCF ne fait pas de sens, notamment parce qu'il s'agit de protéines extra-organellaires.



**Figure 46 : Traitement RNase des échantillons avant co-immunoprécipitation.** Pour se séparer des éventuels contaminants interagissant avec TRCF *via* l'ARN, les échantillons sont traités avec la RNase A avant co-immunoprécipitation. **A.** RPOB est co-précipité de manière similaire avec ou sans le traitement à la RNase. **B.** Le *MDS-plot* montre le regroupement des valeurs pour chaque condition, indiquant la reproductibilité des résultats. Le traitement à la RNase ne permet néanmoins pas de se débarrasser des protéines dont l'interaction avec TRCF fait peu de sens.



Si TRCF pourrait interagir avec la PEP, ses facteurs associés PAP et certaines protéines impliquées dans le métabolisme de l'ARN, il semblerait en revanche qu'elle soit incapable d'interagir avec les NEP. Bien qu'il soit nécessaire de confirmer ces interactions, les fonctions de TRCF pourraient donc être uniquement liées à la transcription par la PEP. Néanmoins, les plantes utilisées sont très certainement trop développées pour détecter correctement les NEP, très faiblement exprimées aux stades avancés du développement des plantes. C'est pourquoi, dans l'optique d'identifier une possible interaction avec les NEP, j'ai réalisé de nouvelles expériences en utilisant des plantules de trois jours, à peine germées.

L'utilisation de plantules si peu développées a permis l'obtention d'un grand nombre de spectres pour TRCF, bien plus élevé qu'avec des plantes de 14 jours. Le nombre de spectres de TRCF:HA varie de 163 à 167 pour les Co-IP avec la lignée TRCF-4B14, et de 246 à 271 avec la lignée TRCF-1B19. Le nombre de protéines co-précipitées est également plus important. Comme pour les expériences précédentes, malgré l'enrichissement préférentiel de certaines protéines avec TRCF:HA, celles-ci semblent tout de même être des contaminants dont l'interaction avec TRCF ne fait pas de sens. Un nombre important de protéines pouvant lier l'ARN est toujours identifié, notamment des sous-unités d'ARN polymérases nucléaires, appuyant l'idée d'une co-précipitation de contaminants extra-organellaires à travers les acides nucléiques. De plus, les protéines identifiées pour des plantules de trois jours sont très différentes de celles identifiées avec des plantes de 14 jours, ne permettant vraiment pas de comparer les résultats entre les expériences. Néanmoins, les sous-unités de la PEP, ainsi que plusieurs PAP et pTAC sont toujours retrouvés. Il est intéressant de noter, que la NEP RPOTp et le facteur sigma chloroplastique SIG1 ont également été identifiés. Ces deux protéines ne présentent cependant qu'un nombre de spectres excessivement faible (2 spectres maximum), et pourraient aussi être des contaminants co-précipités avec l'ARN ou l'ADN.

Afin d'éliminer la contamination importante semblant provenir des acides nucléiques, et particulièrement des ARN, j'ai réalisé une dernière Co-IP en traitant mes échantillons, extraits de plantules de trois jours, avec de la RNase A. Les échantillons devant rester au froid tout au long de l'expérience pour maintenir les interactions, la RNase doit digérer l'ARN dans des conditions sub-optimales. Ainsi, une quantité importante d'enzyme a dû être utilisée. En conséquence de la digestion des ARN, de nombreuses protéines ribosomales sont retrouvées dans les fractions d'élution après Co-IP, ajoutant du bruit de fond supplémentaire. Néanmoins, l'identité des protéines identifiées avec ou sans traitement à la RNase sont assez similaires. Le partenaire d'interaction putatif le plus intéressant, RPOB, est toujours retrouvé malgré le traitement RNase (**figure 46**).

En résumé, l'ensemble des Co-IP réalisées n'a pas permis d'identifier un complexe protéique stable incluant TRCF. Il semble plutôt que TRCF interagisse transitoirement avec un certain nombre de facteurs impliqués dans la transcription du génome chloroplastique. La

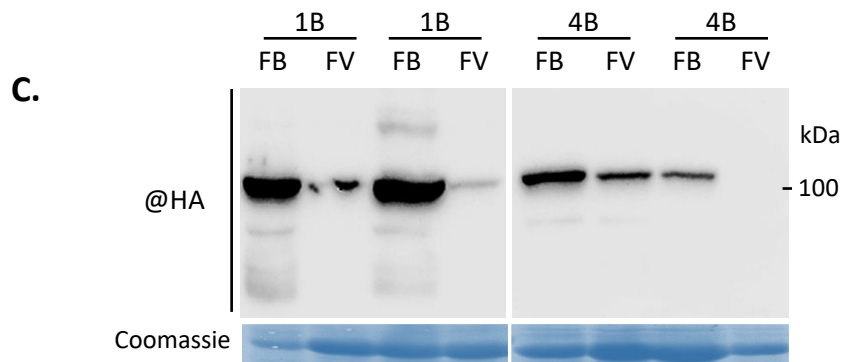
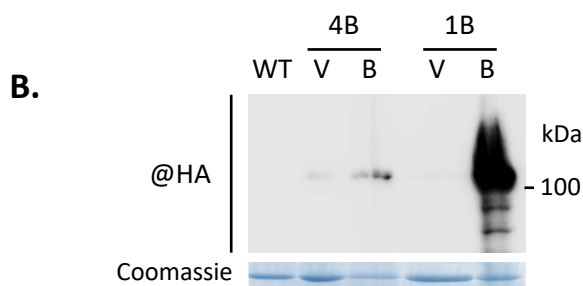
sous-unité  $\beta$  de la PEP est la protéine identifiée le plus fréquemment dans toutes les Co-IP réalisées. Bien qu'avec un nombre de spectres assez restreint, RPOB est identifiée dans 11 des 12 Co-IP. RPOC2 est identifiée dans neuf expériences, tandis que RPOA l'est dans huit. Compte tenu de l'origine et des fonctions de TRCF chez les procaryotes, l'interaction avec la PEP semble probable, bien que je ne sois pas parvenu à le démontrer en système Y2H.

### 8.3. Test des interactions entre TRCF et la PEP par BiFC

Afin de confirmer l'interaction de TRCF avec certaines protéines identifiées à la suite des expériences de Co-IP et n'ayant pas obtenu de résultat probant avec le système Y2H, j'ai entrepris une nouvelle approche par BiFC. J'ai initié ces expériences en fin de thèse, peu de temps avant le début de l'écriture du manuscrit. Ainsi, j'ai décidé de tester uniquement RPOA ou RPOB, dont l'interaction avec TRCF était la plus intéressante à confirmer. Par manque de temps je suis également allé au plus vite et je n'ai volontairement pas cherché à obtenir de témoins positifs ou négatifs, considérant que ces contrôles pourraient être faits par la suite et seulement si des interactions étaient détectées.

J'ai ainsi cloné la séquence de TRCF en fusion avec la partie N-terminale de la protéine YFP dans le vecteur pGW-YN. RPOA et RPOB étant naturellement codés par le génome chloroplastique pour les envoyer dans le chloroplaste j'ai dû les fusionner à une séquence d'adressage plastidial, la séquence d'adressage de la petite sous-unité (SSU : *small subunit*) de la Rubisco. J'ai ensuite cloné les séquences SSU:RPOA et SSU:RPOB dans le vecteur pGW-YC, fusionnées en C-terminal à la partie C-terminale de la YFP. J'ai ensuite transformé des feuilles de tabac avec des combinaisons de deux constructions par agro-infiltration. L'expression des constructions pouvant induire de l'interférence par l'ARN, ces expériences ont été réalisées en ajoutant ou non une construction exprimant la protéine P19, le suppresseur d'interférence par l'ARN du TBSV (*Tomato bushy stunt virus*).

L'observation au microscope confocale de disques foliaires prélevés sur les plantes transformées n'a pas mis en évidence de fluorescence aux longueurs d'onde de la YFP, avec ou sans co-transfection avec P19. Cette absence de fluorescence témoigne de l'impossibilité des parties N- et C-terminale de la YFP de se reconstituer, généralement liée à l'absence d'interaction entre les partenaires testés. Néanmoins, en absence de témoins convenables, il est impossible de conclure quant aux interactions entre TRCF et RPOA ou RPOB dans cette expérience. Il est possible que les protéines RPOA ou RPOB ne soient pas adressées correctement aux chloroplastes, ou que la séquence de l'YFP en C-terminal empêche le repliement correct de la protéine. Il est aussi possible que l'interaction entre TRCF et la PEP nécessite ses quatre sous-unités, mais que l'étiquette utilisée empêche la formation du complexe.



**Figure 47 : La surexpression de TRCF:HA entraîne la dépigmentation des feuilles néoformées.**

**A.** Certaines lignées complémentées par la construction TRCF-B (promoteur endogène, fusion C-terminale à l'étiquette HA) montrent un phénotype de dépigmentation des premières feuilles. Ces lignées présentent une proportion de plantes affectées plus ou moins élevées et une dépigmentation plus ou moins sévère. En vieillissant, les plantes affectées présentent moins de phénotype jusqu'à la disparition complète de la dépigmentation. **B.** TRCF:HA est beaucoup plus exprimé dans les plantes présentant des feuilles blanches (B) que dans les plantes avec uniquement des feuilles vertes (V). **C.** Au sein d'une même plante, TRCF:HA est plus exprimé dans les feuilles blanches (FB) que dans les feuilles redevenues vertes (FV). Chaque lignée exprime (et surexprime) TRCF:HA plus ou moins fortement.

## 9. Caractérisation des phénotypes des mutants complémentés

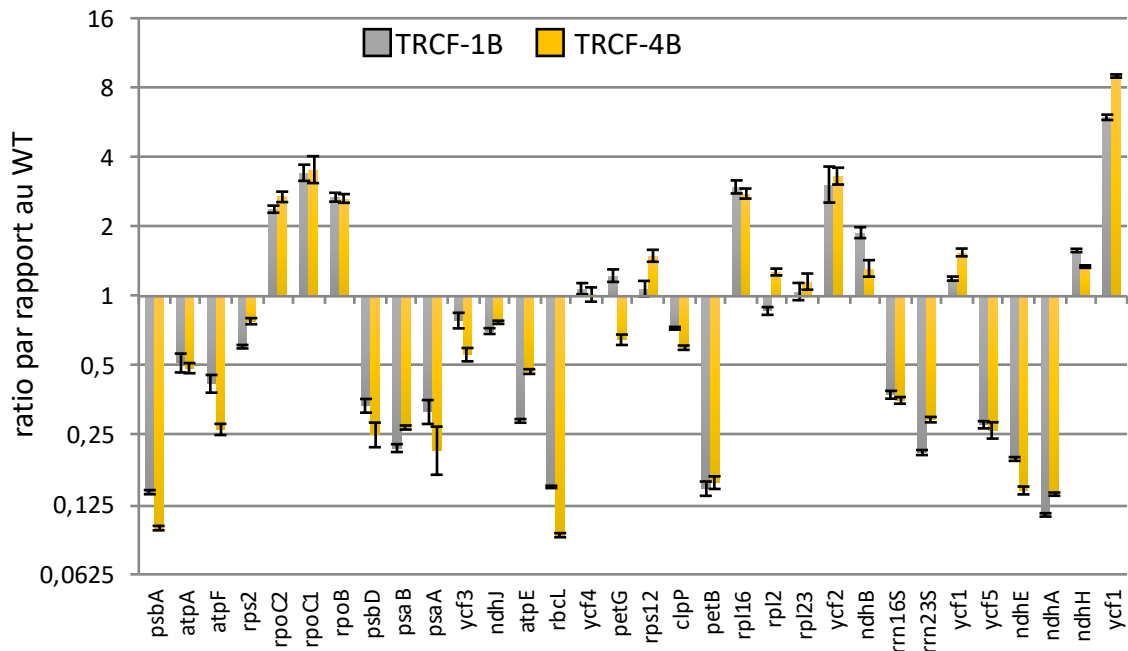
Lors de la sélection des mutants *trcf-1* et *trcf-4* complémentés par différentes versions de TRCF, certaines plantules ont développé un phénotype de jaunissement des feuilles après leur passage d'un milieu *in vitro* contenant du sucre à la terre. Le jaunissement n'affectait que les premières feuilles et entraînait un léger retard de croissance de la plante, puis disparaissait généralement. Ce phénotype n'était observé que pour les plantes complémentées par la construction TRCF-B, avec la séquence du gène TRCF fusionnée à l'étiquette HA en C-terminal et sous le contrôle du promoteur endogène de TRCF, et n'avait touché que quatre plantes sur les 24 sélectionnées, issues des deux lignées mutantes. N'étant pas intéressé par des plantes présentant un phénotype pouvant affecter mes expériences, je les avais laissées de côté. J'interprétais alors ce phénotype comme le résultat de l'insertion du T-DNA TRCF-B dans la séquence d'un autre gène ou à un phénotype développemental lié à au protocole de sélection et de transfert sur terre des plantes.

Par la suite, j'ai observé ce même phénotype chlorotique chez les descendants de lignées indépendantes de la construction TRCF-B (**figure 47A**). Chaque lignée était différentiellement affectée, et certaines n'en développaient pas du tout. Les plantes de la lignée TRCF-1B19 étaient les plus affectées : environ 15-20 % des plantules développant un blanchiment et déformation des premières feuilles (**figure 47A**). Les plantes de la lignée TRCF-4B14 étaient moins affectées, avec seulement 5 % des plantules présentant un jaunissement, parfois léger, des premières feuilles (**figure 47A**). Les plantules des mêmes lignées cultivées sur milieu *in vitro* en présence de sucre n'ont jamais présenté de phénotypes similaires.

Afin de comprendre ce phénotype, j'ai testé par *western blot* l'expression de TRCF:HA dans les plantules présentant oui ou non le phénotype chlorotique. Ainsi, j'ai pu observer une corrélation avec l'expression de TRCF:HA, les plantules chlorotiques exprimant davantage la protéine que les plantules normales (**figure 47B**). De plus, les plantes de la lignée TRCF-1B19, la plus affectée, sur-exprimaient énormément TRCF:HA par rapport aux autres lignées. De manière générale, plus une plante présentait un phénotype sévère, plus TRCF:HA était surexprimé. En comparaison, les plantules sans phénotype n'exprimaient peu, voire pas du tout TRCF:HA.

Les lignées complémentées étant hétérozygotes pour la construction, un quart des plantes résultant de leur ségrégation ne devaient plus contenir la construction TRCF-B, et pas exprimer TRCF:HA. Le génotypage des plantules chlorotiques a montré que toutes contenaient toujours la construction TRCF-B, liant définitivement le phénotype à l'expression de TRCF:HA.

Malgré un léger retard de croissance, les plantes présentant le phénotype parviennent à se développer correctement par la suite, devenant indistinguables des autres plantes. A un stade plus avancé de leur développement j'ai testé l'expression de TRCF:HA dans les nouvelles



**Figure 48 : Effet de la surexpression de TRCF sur l'expression des transcrits chloroplastiques.** Plusieurs transcrits de gènes chloroplastiques de plantes sauvages (WT) ou surexprimant TRCF:HA (TRCF-1B et TRCF-4B) sont quantifiés par RT-qPCR. Les expériences ont été réalisées en triplica et normalisées par rapport aux gènes nucléaires de référence *ACT2* et *GAPDH*. L'expression des gènes des plantes complémentées est rapportée à celle des plantes sauvages.

feuilles chlorotiques et dans les feuilles redevenues vertes au sein d'une même plante. Ces résultats montrent à nouveau que TRCF:HA est surexprimé dans les feuilles présentant un phénotype par rapport aux vertes, où l'expression de TRCF:HA est plus faible, voire indétectable (**figure 47C**).

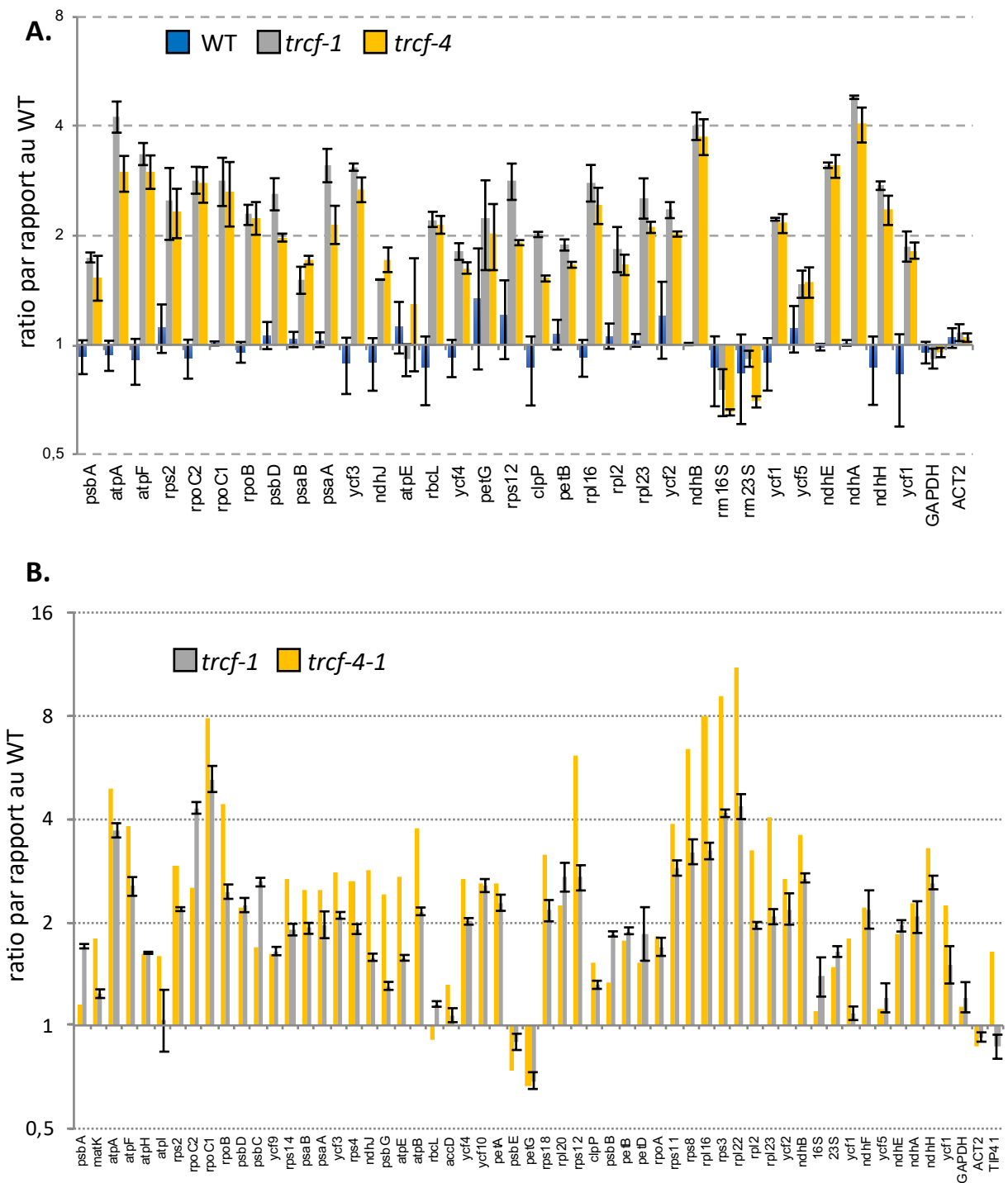
J'ai également recherché des phénotypes similaires chez les plantes des lignées transformées avec les constructions TRCF-A et TRCF-C, mais n'en ai pas trouvé. Les plantes TRCF-1C et TRCF-4C (mutants *trcf-1* ou *trcf-4* transformés avec TRCF-C) n'expriment pas, ou trop peu TRCF:HA, ne permettant donc pas le développement du phénotype. Quant aux plantes TRCF-1A et TRCF-4A (mutants *trcf-1* ou *trcf-4* transformés avec TRCF-A), je n'ai trouvé aucune lignée exprimant TRCF à des niveaux comparables aux expressions détectées dans les plantes TRCF-1B et TRCF-4B. De fait, je ne peux pas exclure que le phénotype observé chez les plantes transformées par la construction TRCF-B soit dû à la présence de l'étiquette HA en fusion avec TRCF. Il est possible que l'étiquette altère sensiblement les fonctions de TRCF:HA causant ainsi des défauts de développement du chloroplaste. Néanmoins, il semble plus probable que la surexpression de TRCF ou de TRCF:HA altère les fonctions d'autres protéines, comme la PEP par exemple, empêchant l'expression correcte du génome chloroplastique.

## 10. Étude des fonctions de TRCF dans la transcription chloroplastique

### 10.1. Impact de la surexpression de TRCF:HA sur la transcription

Le phénotype chlorotique observé chez les plants surexprimant TRCF:HA pouvant provenir d'un défaut de l'expression des gènes chloroplastiques, les transcrits chloroplastiques ont été mesurés chez les mutants complémentés par rapport au sauvage par RT-qPCR (**figure 48**). Nous avons ainsi observé une dérégulation de l'expression de la majorité des gènes testés chez les mutants complémentés. Les transcrits codant des protéines impliquées dans les complexes photosynthétiques sont diminués par rapport au sauvage, à l'exception des transcrits *ndhB*, *ndhH* et *petG* qui sont similaires au sauvage. Les quantités de transcrits codant les protéines ribosomiques ne sont pas significativement modifiées, tandis que celles des deux ARN ribosomiques diminuent de deux à quatre fois par rapport au sauvage. Enfin, la quantité des transcrits de l'opéron *rpoB* (comprenant les gènes *rpoB*, *rpoC1* et *rpoC2*), de la protéine ribosomique *rpl16* et des protéines *ycf1* et *ycf5*, aux fonctions inconnues, est augmentée.

Un tel profil d'expression est caractéristique de mutants et traitements qui inhibent la traduction chloroplastique et est retrouvé dans les mutants qui affectent le fonctionnement de la PEP (Steiner *et al.*, 2011). La transcription des gènes codant les protéines des complexes photosynthétiques, majoritairement transcrits par la PEP, est ici diminuée alors que celle des gènes de ménage (PEP, protéine ribosomique) est plutôt augmentée. Ainsi, lors d'une inhibition de la PEP, l'expression des protéines photosynthétiques est réduite, conduisant à un phénotype



**Figure 49 : Les plantes déficientes pour TRCF présentent une augmentation générale de la transcription des gènes chloroplastiques. A.** Plusieurs transcrits de gènes chloroplastiques de plantes sauvage ou mutées (*trcf-1* et *trcf-4*) sont quantifiés par RT-qPCR. Les barres d'erreur représentent la variation au sein des triplicatas techniques. L'expression des mutants et d'un second triplicata sauvage (WT) est rapportée à celle de plantes sauvages. Les expériences sont normalisées par rapport aux gènes nucléaires de référence *ACT2* et *GAPDH*. **B.** Identique à A avec un nombre de gènes testé plus important.

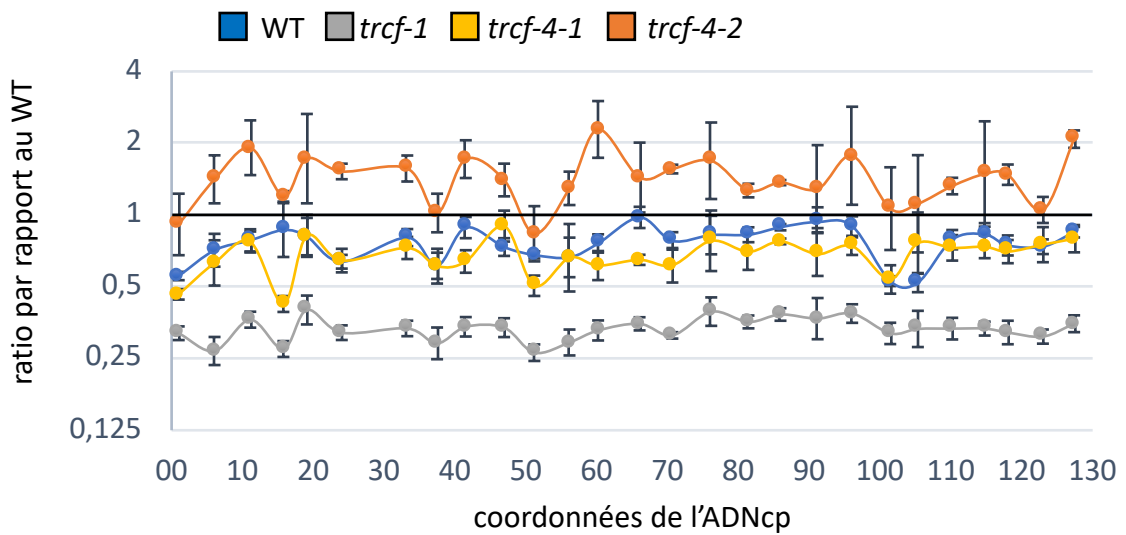
chlorotique, tel que celui qui est observé pour les mutants complémentés TRCF-1B et TRCF-4B. Le défaut de transcription est alors repéré et transmis au noyau par signal rétrograde (Woodson et Chory, 2008). En réponse, le noyau adresse davantage de NEP au chloroplaste afin d'augmenter la transcription des gènes de ménage et de compenser l'inhibition initiale (Emanuel *et al.*, 2006; Zhelyazkova *et al.*, 2012b). A l'inverse des mutants de la PEP ou de ses facteurs associés, le profil observé ici est dû à la surexpression de TRCF:HA, présumant donc d'une activité inhibitrice de la protéine sur la transcription chloroplastique.

## 10.2. Augmentation de la transcription en absence de TRCF

Connaissant l'impact de la surexpression de TRCF:HA sur la transcription chloroplastique, nous avons voulu savoir quel serait l'effet de l'absence de TRCF sur l'expression des gènes chloroplastique. Pour cela, les transcrits de gènes chloroplastique ont été quantifiés dans les mutants *trcf-1* et *trcf-4* par rapport au sauvage par RT-qPCR (**figure 49**). Les résultats montrent que l'absence de TRCF entraîne une augmentation globale des transcrits, à l'exception des ARNt et des transcrits *psbE* et *petG*. Cette augmentation serait due à une augmentation de la transcription, bien qu'à présent on ne peut exclure que ce ne soit pas la stabilité des transcrits qui soit modifiée. Afin de s'assurer que l'augmentation globale des transcrits chloroplastiques n'est pas due à une différence du nombre de copies de l'ADNcp, l'ADNcp a été quantifié par qPCR. Nous avons ainsi observé que le nombre de copies de l'ADNcp ne variait pas significativement entre les mutants et le sauvage (**figure 50**).

Le rôle d'inhibiteur de la transcription de TRCF, suspecté au vu des résultats obtenus avec les mutants complémentés, est ici appuyé par ceux obtenus sur les mutants. L'absence de TRCF permet une transcription augmentée de la très grande majorité des gènes testés, suggérant que TRCF inhibe normalement cette transcription. Cependant, si la surexpression de TRCF impacte surtout la transcription réalisée par la PEP, l'augmentation de la transcription liée à l'absence de TRCF ne semble pas être dépendante de l'ARN polymérase utilisée.





**Figure 50 : Le nombre de copies de l'ADNcp reste stable dans les mutants *trcf*.** Le nombre de copies de chaque régions de l'ADNcp est quantifiée par qPCR. Chaque point indique le nombre de copies d'une région de 5 kb tout au long de l'ADNcp (de 130 kb) représenté en abscisse. Les données obtenues sont rapportées à une plante sauvage (barre noire).

## Discussion et perspectives

Le travail présenté dans cette partie du manuscrit visait à comprendre le rôle du facteur TRCF dans les organelles d'*Arabidopsis*. Après s'être assuré de l'adressage organellaire de TRCF, j'ai étudié sa séquence et modélisé sa structure pour définir mes premières hypothèses de travail. J'ai ensuite recherché les partenaires de TRCF en m'orientant d'abord vers les RNAP des organelles et j'ai étudié l'impact de l'absence ou de la surexpression de TRCF sur la transcription et le développement de la plante. Ces expériences m'ont permis d'appréhender les fonctions des TRCF, mais de nombreuses zones d'ombre persistent et nécessiteraient davantage de travail pour être élucidées.

TRCF est une protéine adressée aux chloroplastes et aux mitochondries. Elle est conservée dans toute la lignée verte, des algues aux plantes à fleurs. L'étude phylogénétique laisse penser que le gène *TRCF* a évolué à partir du gène de la cyanobactérie primitive ancêtre du chloroplaste transféré dans le noyau. Chez les bactéries, TRCF joue un rôle dans la réparation de l'ADN en recrutant les facteurs du NER après le dégagement d'ARN polymérases bloquées par des lésions. Pourtant, chez les plantes, aucun facteur du NER n'a jamais été décrit dans les organelles. TRCF est donc l'unique homologue à un facteur du NER bactérien apparemment présent dans la mitochondrie et le chloroplaste, suggérant ainsi un rôle différent chez les plantes et les bactéries. Cette différence de fonction est appuyée par une modification importante de la partie N-terminale de la protéine. Les trois domaines d'homologie à UvrB, permettant le recrutement d'UvrA chez les bactéries, sont absents chez les plantes et remplacés par la séquence d'adressage aux organelles. De plus, le mutant KO *trcf-4* n'est pas plus affecté que le sauvage dans sa survie lorsqu'il est soumis à différentes doses d'UV, qui induisent des liaisons réparées par le NER dans le noyau. Néanmoins, TRCF pourrait tout de même être impliqué dans la réparation de l'ADN à travers d'autres voies telles que le BER, comme cela a été décrit pour CSA et CSB dans le noyau (D'Errico *et al.*, 2013). Dans ce sens, les mutants *trcf-1* et *trcf-4* présentent une sensibilité plus élevée à de stress oxydatifs (peroxyde d'hydrogène et roténone), comme attendu pour des mutants de facteurs impliqués dans la réparation des lésions oxydatives. Pourtant, lors de la recherche des partenaires d'interaction de TRCF par Co-IP, aucune protéine impliquée dans le BER n'a été identifiée. Chez l'homme, CSA et CSB interagissent avec mtOGG1 et mtSSBP1, deux protéines du BER mitochondrial (Kamenisch *et al.*, 2010). De manière générale, aucune protéine ayant un rôle connu dans la réparation des génomes des organelles n'a été identifiée lors de mes expériences de Co-IP, consolidant l'idée que TRCF n'est pas impliqué dans la réparation de l'ADN.

Si la partie N-terminale du TRCF bactérien a été remplacée chez les plantes, le reste de la protéine est plutôt bien conservé. La modélisation de la structure de TRCF d'*Arabidopsis* montre que les domaines D4 à D7 pourraient conserver leur structure, suggérant le maintien des fonctions de la protéine bactérienne dans les organelles de plantes. Chez *E. coli*, la protéine

TRCF tronquée de sa partie N-terminale, similaire à celle retrouvée dans les organelles, est toujours capable de s'associer avec les RNAP bloquées et de les dissocier (Selby et Sancar, 1995b). Néanmoins, les organelles de plantes présentent plusieurs RNAP, et seule la PEP possède une structure similaire à la RNAP bactérienne. L'interaction avec la PEP était donc initialement jugée la plus probable, bien qu'il ait été montré que TRCF pouvait également dissocier une RNAP de type phagique (Ganesan et Hanawalt, 2010). La modélisation des structures de TRCF et de la sous-unité  $\beta$  de la PEP (RPOB) soutient l'interaction entre les deux protéines, l'orientation de leurs chaînes latérales permettant un rapprochement compatible avec une interaction.

L'interaction entre TRCF et RPOB est appuyée par mes résultats de Co-IP. Bien qu'identifiée très faiblement, RPOB est la seule protéine systématiquement retrouvée dans mes expériences. De manière étonnante, les autres sous-unités de la PEP, particulièrement RPOC1, sont retrouvées dans un nombre plus restreint d'expériences. La PEP étant un complexe stable, ses quatre sous-unités devraient co-précipiter ensemble. La sous-unité  $\alpha$  (RPOA) est identifiée dans 8 des 11 expériences et présente le nombre moyen de spectres le plus élevé, en accord avec la stœchiométrie des protéines de la PEP :  $RPOA_2B_1C_1C_2_1$ . En plus de la PEP, six PAP (PAP1, 2, 3, 5, 6 et 10) et deux pTAC (FLN1 et GyrA) sont également significativement enrichies dans les Co-IP. Les fonctions exactes de ces PAP sont toujours largement inconnues, bien qu'elles soient catégorisées dans des groupes fonctionnels. PAP6, PAP10 et FLN1 seraient impliquées dans la régulation de la transcription par la PEP, tandis que les PAP1, PAP2, PAP3 et PAP5 pourraient se lier à l'ADN et seraient impliquées dans la reconnaissance des promoteurs, l'initiation et l'élongation de la transcription (Pfalz *et al.*, 2006; Yagi *et al.*, 2012; Wang *et al.*, 2016; Chang *et al.*, 2017b; Grübler *et al.*, 2017; Nevarez *et al.*, 2017; He *et al.*, 2018). Étonnamment, PAP8 et PAP11, de fonctions totalement inconnues, et PAP4 et PAP9, impliquées dans la protection contre les ROS, n'ont pas été enrichies dans mes expériences. PAP7 et PAP12 n'ont également pas été enrichies bien qu'elles appartiennent aux mêmes groupes fonctionnels que les six PAP identifiées dans mes expériences. L'absence de la totalité des PAP et la co-précipitation inégale des sous-unités de la PEP doit provenir de la préférence d'interaction de TRCF. Si TRCF interagit avec RPOB, il est probable que les interactants secondaires, voire tertiaires, interagissant avec RPOB soient plus faiblement co-précipités. Il est intéressant de noter que TRCF co-précipite avec les PAP impliquées dans la liaison à l'ADN et la régulation de la transcription, alors que la fonction de TRCF semble justement être de moduler la transcription de la PEP. Néanmoins, là où les PAP favoriseraient la transcription et interagissent stablement avec la PEP, TRCF serait plutôt un inhibiteur interagissant transitoirement avec la PEP.

TRCF étant une translocase à ADN, les partenaires identifiés pourraient avoir co-précipité avec les acides nucléiques auxquels TRCF se lie. Néanmoins, chez les procaryotes, la liaison de

TRCF avec l'ADN nécessite préalablement l'interaction avec l'ARN polymérase bloquée (Deaconescu et Suhanovsky, 2017). La protéine TRCF des plantes ayant conservé son domaine C-terminal D7, celui-ci pourrait toujours servir à maintenir la protéine dans une conformation initiale inhibée, empêchant la liaison avec l'ADN. La co-précipitation de protéines *via* l'ADN n'a cependant pas pu être testée durant ma thèse, faute de temps pour optimiser les conditions. Au vu du grand nombre de protéines de liaison à l'ADN identifiées lors de mes expériences, y compris plusieurs des PAP, de nouvelles Co-IP devront être effectuées en dégradant les ADN. La co-précipitation de protéines *via* l'ARN a pu être testée en utilisant la RNase A, mais n'a pas permis d'améliorer les résultats déjà obtenus. De fait, la grande majorité des protéines identifiées dans mes Co-IP semble être des contaminants dont la co-précipitation avec TRCF:HA ne fait pas beaucoup de sens. La PEP étant une protéine associée aux membranes du chloroplaste, il est également possible que certaines protéines identifiées soient co-précipitées en interagissant avec les membranes (Finster *et al.*, 2013).

La protéine la plus significativement enrichie avec TRCF dans mes expériences est l'hélicase à ARN ISE2. Deux autres hélicases à ARN chloroplastiques sont également significativement enrichies avec TRCF : RH3 et AT3G06980. RH3 et ISE2 sont impliqués dans l'épissage des introns de groupe II chloroplastiques et jouent un rôle dans le développement des plantes. (Gu *et al.*, 2014; Carlotto *et al.*, 2016). Néanmoins, une étude des partenaires d'ISE2 n'a pas identifié TRCF, suggérant l'absence d'interaction directe entre ISE2 et TRCF (Bobik *et al.*, 2019). Cette étude a cependant montré une interaction entre ISE2 et RH3. Chez le maïs, les homologues de ces trois hélicases sont des composants des nucléoïdes chloroplastiques (Majeran *et al.*, 2012). L'interaction entre TRCF et ces ARN hélicases pourrait donc être indirecte, *via* l'interaction de TRCF avec un autre composant des nucléoïdes. Plusieurs des composants importants des nucléoïdes (WHY-1, POL1A) n'ont cependant pas été identifiés lors de mes Co-IP, suggérant que les protéines identifiées ne proviennent pas de la co-précipitation des nucléoïdes en entier. TRCF n'a d'ailleurs pas été identifié comme un composant des nucléoïdes chloroplastiques (Majeran *et al.*, 2012).

Les potentiels partenaires d'interaction identifiés par l'approche de Co-IP nécessiteraient d'être confirmés. La confirmation de l'interaction de TRCF avec les sous-unités de la PEP ou les NEP par Y2H et BiFC n'a cependant pas abouti. Aucune interaction n'a été trouvée entre TRCF et les différentes RNAP organellaires avec les deux approches utilisées. Néanmoins, les expériences de BiFC manquent de témoins, rendant impossible toute analyse de résultats. Ces expériences ont été menées rapidement en fin de thèse pour tenter de confirmer l'interaction la plus intéressante à mes yeux. Mais les vecteurs utilisés, procurés par une autre équipe de l'IBMP, ne sont pas suffisamment caractérisés. Malgré le séquençage de la cassette de clonage et de ses régions environnantes, il n'est pas complètement exclu que ce vecteur ne soit pas un vecteur binaire adapté aux expériences menées. L'adressage des protéines de fusion vers les

chloroplastes n'a pas non plus été confirmé. L'approche par BiFC devra donc être recommencée en incluant les témoins nécessaires et en utilisant potentiellement d'autres vecteurs, caractérisés et testés. Une construction produisant une version mutée de TRCF dans son domaine RID constituerait un bon témoin négatif pour tester l'interaction avec les sous-unités de la PEP. Tandis que tester l'interaction entre RPOA et RPOB permettrait de s'assurer de la fonctionnalité du test et de la conformation adéquate des protéines fusionnées à l'YFP. Enfin, il sera possible de s'assurer de l'expression des protéines de fusion TRCF:YFP, RPOA:YFP et RPOB:YFP par *western-blot*. L'approche par Y2H devrait aussi être recommencée. Bien que l'expression de la majorité des protéines ait été confirmée par *western blot*, il est possible que la fusion d'un des domaines de GAL4 empêche leur repliement correct, altérant leur potentielle interaction. Produire de nouvelles constructions en fusionnant les domaines de GAL4 en N-terminale de la protéine d'intérêt plutôt qu'en C-terminal permettrait de palier à un défaut de repliement. Pour cela, le clonage de TRCF dans le vecteur pGBKCg est actuellement en cours au laboratoire. L'absence d'interaction en système Y2H pourrait également provenir de la présence d'une seule sous-unité de la PEP dans chaque expérience. En absence des autres sous-unités, la conformation de la sous-unité testée pourrait ne pas être identique à celle qu'elle aurait dans le complexe. De même, chez les bactéries, TRCF interagit spécifiquement avec les RNAP bloquées qui présentent une altération de conformation de neuf acides aminés (Westblade *et al.*, 2010). Cette altération pourrait être le signal de recrutement de TRCF (Selby, 2017). Chez la levure, en absence d'une PEP complète, il ne pourrait y avoir cette altération conformationnelle.

La surexpression de TRCF conduit au développement de phénotypes chlorotiques liés à un défaut de transcription par la PEP. Cette surexpression diminue également la quantité d'ARNr, impactant davantage l'expression des gènes. Afin de compenser le défaut de transcription ou de traduction, l'expression des gènes transcrits par la NEP est augmentée. Cette augmentation serait due à un signal rétrograde vers le noyau qui entraînerait une surexpression des NEP (Emanuel *et al.*, 2004). Cette réponse devra être testée en quantifiant l'expression des gènes NEP nucléaires par RT-qPCR. A l'inverse, l'absence de TRCF entraîne une augmentation globale de la transcription chloroplastique, indépendante des promoteurs NEP ou PEP. Ces mutants présentent pourtant un défaut de production de la chlorophylle, particulièrement de la chlorophylle a, ainsi qu'une sensibilité accrue à la lincomycine. Ainsi, malgré l'augmentation de la transcription, l'expression des gènes est tout de même déstabilisée. Le défaut d'expression pourrait alors provenir d'un défaut de maturation des transcrits ou de traduction. Cette inhibition serait alors perçue et entraînerait une réponse rétrograde similaire à celle décrite précédemment. L'augmentation de l'expression nucléaire des NEP augmenterait alors la transcription des gènes possédant un promoteur NEP, conduisant ainsi à l'augmentation de la transcription des gènes avec un promoteur PEP.

L'expression de gènes NEP nucléaire devra aussi être testée dans le mutant. Pour tester ces hypothèses une analyse globale de l'expression et de la maturation des transcrits chloroplastiques est nécessaire, ce qui sera fait par *RNA-Sequencing*.

L'expression du promoteur de TRCF dans les phases précoces du développement des tissus chlorophylliens laisse penser que TRCF a un rôle dans la biogénèse du chloroplaste. Cette hypothèse est appuyée par la régulation de l'expression de TRCF par la lumière du jour. Pourtant, l'absence de TRCF n'entraîne pas de phénotype développemental sévère, à tel point qu'aucun phénotype n'a été observé durant la quasi-totalité de ma thèse. Le rôle de TRCF dans le développement devrait alors plutôt être dans la régulation fine de certains mécanismes. Lors du développement de tissus chlorophylliens, les gènes plastidiaux des éoplastes (graines) ou des proplast (méristèmes) sont d'abord principalement exprimés par la NEP (Liebers *et al.*, 2017). En présence de lumière, la transcription va ensuite être majoritairement réalisée par la PEP, permettant la synthèse massive des protéines photosynthétiques lors du processus de photomorphogénèse (Liebers *et al.*, 2018). TRCF pourrait être impliqué lors de la photomorphogénèse pour réguler l'activité de la PEP et ou de la NEP et permettre un développement optimal. TRCF pourrait ainsi inhiber l'expression de gènes photosynthétiques avant la complète mise en place des machineries de transcription et de traduction. Lors des phases précoces du développement des chloroplastes, l'ADNcp est également activement répliqué et TRCF pourrait limiter les conflits entre les machineries de transcription et de réplication, comme cela a été montré chez les bactéries (Sankar *et al.*, 2016). J'ai voulu tester cette possibilité lors de ma thèse, mais le système utilisé s'est révélé incompatible avec mes vecteurs. Le système reposait sur la possibilité de compléter des mutants bactériens *rep*, *uvrD* ou *rep-uvrD*, déficients dans le dégagement des blocages de la fourche de réplication, par la protéine TRCF d'*Arabidopsis* (Boubakri *et al.*, 2010). L'induction à l'IPTG (isopropyl  $\beta$ -D-1-thiogalactopyranoside) de TRCF cloné dans le vecteur pACYC184LacZ nécessaire à son expression était cependant incompatible avec le système. TRCF devrait être sous-clonée dans un autre vecteur pACYC non dépendant d'une induction à l'IPTG, pour tester l'implication de TRCF dans la résolution des conflits entre réplication et transcription.

En résumant, il semble bien que TRCF ne joue aucun rôle dans le couplage de la transcription à la réparation dans les organelles. Il serait impliqué dans la régulation de l'expression des gènes chloroplastiques, mais son rôle exact reste encore flou. D'autres expériences seront encore nécessaires pour élucider ses fonctions exactes.

# **MATERIEL ET MÉTHODES**

# Matériel

## 1. Matériel bactérien

### 1.1. *Agrobacterium tumefaciens* GV3101 pMP90

Cette souche phytopathogène dérivant de la souche C58C1 a été utilisée pour transformer transitoirement ou stablement les plantes. Elle est résistante à la rifampicine et possède le plasmide Ti désarmé pMP90 conférant une résistance à la gentamycine et portant les gènes de virulence *vir* nécessaires au transfert de l'ADN-T dans la cellule végétale.

### 1.2. *Escherichia coli* BW25113

La souche BW25113 dérive la souche d'*Escherichia coli* K12 et a été utilisée pour réaliser la collection Keio de mutants, de laquelle est issue la souche JW4352 utilisée pour la complémentation de *radA*. La souche BW25113 sert donc de contrôle sauvage.

**BW25113** : F<sup>-</sup>  $\Delta$ (*araD-araB*)567,  $\Delta$ *lacZ*4787(::rrnB-3, LAM<sup>-</sup>, *rph-1*,  $\Delta$  (*rhaD-rhaB*)568, *hsdR*514

### 1.3. *Escherichia coli* DH5 $\alpha$

La souche dérive d'*Escherichia coli* K12. Elle a été utilisée pour le clonage par choc thermique et l'amplification de plasmide. Elle est déficiente pour le gène *recA1* qui réduit le taux de recombinaison et permet le maintien de la construction transfectée. Elle permet l' $\alpha$ -complémentation et la sélection blanc-bleu des clones en présence d'IPTG et de X-gal.

**DH5 $\alpha$**  : F<sup>-</sup> *endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG*  $\Phi$ 80*lacZ* $\Delta$ M15  $\Delta$ (*lacZYA-argF*)U169, *hsdR*17(*r<sub>K</sub><sup>-</sup> m<sub>K</sub><sup>+</sup>*),  $\lambda$ <sup>-</sup>

### 1.4. *Escherichia coli* TOP10

Elle a été utilisée pour le clonage par choc thermique et l'amplification de plasmide. Elle est déficiente pour le gène *recA1*, limitant ainsi les recombinaisons non spécifiques pour assurer la stabilité de la construction transformée. Elle permet l' $\alpha$ -complémentation et la sélection des clones recombinants par la couleur des colonies en présence d'IPTG et de X-gal.

**TOP10** : F<sup>-</sup>*mcrA*  $\Delta$ (*mrr-hsdRMS-mcrBC*)  $\phi$ 80*lacZ* $\Delta$ M15  $\Delta$ *lacX74recA1 araD139*  $\Delta$ (*ara leu*)7697 *galU galK rpsL* (Str<sup>R</sup>) *endA1 nupG*



### 1.5. Escherichia coli JW4352

La souche JW4352 dérive de la souche d'*Escherichia coli* BW25113 et présente une insertion inactivant le gène *radA*. Cette souche a été utilisée dans les expériences de complémentation du mutant bactérien *radA* par le facteur RADA d'*A. thaliana*.

**JW4352** : F<sup>-</sup>  $\Delta(\textit{araD-araB})567, \Delta\textit{lacZ4787}>::\textit{rrnB-3}, \lambda^-, \textit{rph-1}, \Delta(\textit{rhaD-rhaB})568, \textit{hsdR514}, \Delta\textit{radA785}>::\textit{kan}$

### 1.6. Escherichia coli Rosetta 2 (DE3) pLysS

La souche Rosetta 2 (DE3) pLysS (Novagen®) dérivant de la souche BL21 est une souche optimisée pour l'expression de protéines eucaryotes recombinantes. La transcription est réalisée par l'ARN polymérase du phage T7, codée par le chromosome bactérien sous contrôle d'un promoteur inductible à l'IPTG. La souche est aussi déficiente dans la protéase bactérienne OmpT indésirable pour l'expression protéique. Elle possède le plasmide pLysS codant le lysozyme du phage T7 qui supprime l'expression basale de l'ARN polymérase en absence d'induction. Ce plasmide porte également les gènes des ARNt qui décodent des codons rares chez *E. coli*.

**Rosetta(DE3)pLysS** : F<sup>-</sup> *ompT hsdS<sub>B</sub>(R<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>) gal dcm*  $\lambda$  (DE3 [*lacI lacUV5-T7 gene 1 ind1 sam7 nin5*]) pLysSRARE (Cam<sup>R</sup>)

### 1.7. Escherichia coli DB3.1

La souche DB3.1 (Invitrogen®/Thermo Fisher Scientific®, Carlsbad, Californie, États-Unis) a été utilisée pour le clonage par électroporation et l'amplification des vecteurs Gateway® portant un gène suicide. Cette souche contient l'allèle *gyrA462* qui la rend résistante aux effets toxiques du gène *ccdB*, qui code un inhibiteur de la gyrase.

**DB3.1** : *gyrA462 endA1  $\Delta(\textit{sr1-recA}) \textit{mcrB mrr hsdS20 glnV44 (=supE44) ara14 galK2 lacY1 iA2 rpsL20 xyl5 leuB6 mtl1}$*

## 2. Levures

### 2.1. Saccharomyces cerevisiae AH109

Cette souche de signe sexuel *MATa* est utilisée pour détecter les interactions protéine-protéine lors de tests Y2H en système GAL4. Elle contient les gènes rapporteurs HIS3, ADE2, LacZ et MEL1.

**AH109** : *MATa, trp1-901, leu2-3, 112, ura3-52, his3-200, gal4 $\Delta$ , gal80 $\Delta$ , URA3::MEL1<sub>UAS</sub>-MEL1<sub>TATA</sub>-LacZ, LYS2::GAL1<sub>UAS</sub>-GAL1<sub>TATA</sub>-HIS3, GAL2<sub>UAS</sub>-GAL2<sub>TATA</sub>-ADE2*

## 2.2. *Saccharomyces cerevisiae* Y187

Cette souche de signe sexuel *MAT $\alpha$*  est également utilisée en Y2H en partenaire avec la souche AH109 et le gène rapporteur HIS3.

**Y187** : *MAT $\alpha$* , *ura3-52*, *his3-200*, *ade2-101*, *trp1-901*, *leu2-3*, *112*, *gal4 $\Delta$* , *met<sup>-</sup>*, *gal80 $\Delta$* , *URA3::GAL1<sub>UAS</sub>:GAL1<sub>ATA</sub>-lacZ*

## 3. Matériel végétal

### 3.1. *Arabidopsis thaliana* écotype Columbia

La plante modèle *A. thaliana* écotype Columbia (Col-0) a été utilisée pour la majorité des expériences pour étudier le rôle de TRCF et de RADA. Les plantes en terre sont cultivées sous cycle diurne de 16 heures avec des températures de 21 °C le jour et 18 °C la nuit sous lampes néon (*Osram BioluxT5*) ou LED (*Philips LEDtube SA1 Master*). Les cultures *in vitro* sont cultivées sous cycle diurne de 16 heures à 21 °C le jour et 18 °C la nuit sous lampes néon.

#### 3.1.1. Mutant d'insertion TRCF

Les mutants de TRCF (At3g02060) d'*Arabidopsis* par insertion de T-DNA *trcf-1* (SAIL\_682\_C02) et *trcf-4* (SALK\_076350) dérivent de l'écotype Col-0. Ils proviennent des collections 'Syngenta Arabidopsis Insertion Library' (SAIL) et du 'Salk Institute Genome Analysis Laboratory' (SIGnAL) (Alonso *et al.*, 2003) respectivement et sont distribués par le 'Nottingham Arabidopsis Stock Centre' (NASc).

#### 3.1.2. Mutant d'insertion RADA

Les mutants de RADA (At5G50340) d'*Arabidopsis* par insertion de T-DNA *rada-2* (WidcDsLoxHs058\_03D) et *rada-3* (SALK\_097880) dérivent de l'écotype Col-0. Ils proviennent des collections de l'Université de Wisconsin-Madison (Woody *et al.*, 2007) et SIGnAL respectivement et sont distribués par le NASc.

### 3.2. *Nicotiana benthamiana*

Les plants de tabac *N. benthamiana* ont été utilisés pour les expériences de localisation subcellulaire de protéines fusionnées à la eGFP par biolistique, et pour l'étude de l'interaction entre deux protéines par BiFC par agro-infiltration.

**Tableau 2 : Antibiotiques utilisés**

<b>Nom</b>	<b>Concentration</b>	<b>Utilisation</b>
Ampicilline	100 µg/mL	Sélection bactérienne
Carbéniciline	200 µg/mL	Sélection transformant d'Arabidopsis
Céfotaxime	250 µg/mL	Sélection transformant d'Arabidopsis
Chloramphénicol	25 µg/mL	Sélection bactérienne
Ciprofloxacine	9-21 mM	Stress génotoxique bactérien
Hygromycine B	15 µg/mL	Sélection transformant d'Arabidopsis
Kanamycine	50 µg/mL	Sélection bactérienne
Lincomycine	0,01-0,16 mM	Stress génotoxique chloroplastique
Rifampicine	30 µg/mL	Sélection bactérienne
Zéocine	25 µg/mL	Sélection bactérienne

## 4. Milieux de cultures

### 4.1. Milieu LB

Le milieu Luria Bertani, LB (ou *lysogeny broth*) est le plus couramment utilisé pour la culture de bactéries (Bactotryptone 1 % (p/v) ; Extrait de levures 0,5 % (p/v) ; NaCl 0,5 % (p/v) ; pH 7,4). Le milieu est généralement complétement avec les antibiotiques adéquats (**tableau 2**) pour la sélection des bactéries d'intérêt et peut être gélosé lors de l'utilisation de boîtes de Pétri (agar 1,5 %).

### 4.2. Milieu M9

C'est un milieu de culture bactérien minimum composé d'une source de carbone et d'azote et de sels minéraux. Il est préparé en utilisant une solution dix fois concentrée (10 X : Na<sub>2</sub>HPO<sub>4</sub> 33,7 mM ; KH<sub>2</sub>PO<sub>4</sub> 22 mM ; NaCl 8,55 mM ; NH<sub>4</sub>Cl 9,35 mM : pH 7,4) diluée au dixième puis supplémentée de glucose 0,4 % (p/v), d'acide casamino 0,2 % (p/v), de MgSO<sub>4</sub> 1 mM, de CaCl<sub>2</sub> 0,3 mM et de thiamine 0,002 % (p/v). Pour préparer des boîtes de Pétri, 1,5 % d'agar (p/v) est ajouté.

### 4.3. Milieu YPD

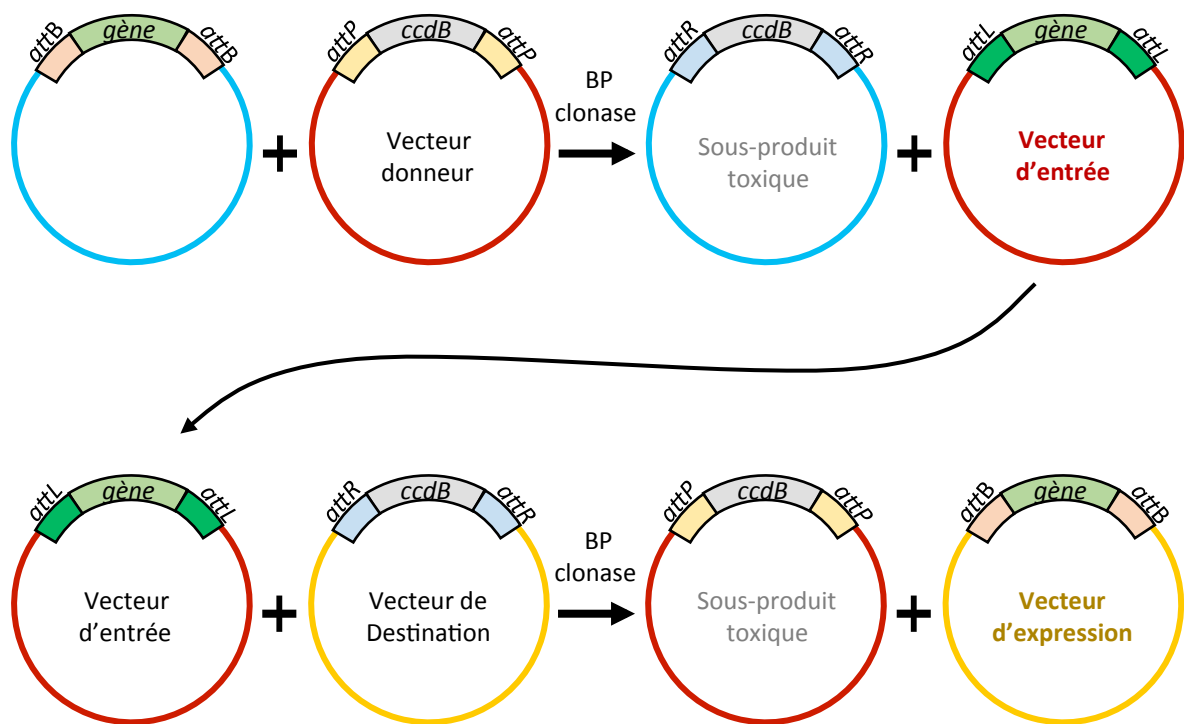
Le milieu *Yeast extract peptone dextrose* (YEPD ou YPD) est un milieu complet permettant la croissance non sélective des levures. Il est composé d'extrait de levure 1 % (p/v), peptone 2 % (p/v), glucose 2 % (p/v) et peut être supplémenté de 1,5 % d'agar.

### 4.4. Milieu SD

Le milieu *synthetic defined* (*Sigma-Aldrich*<sup>®</sup>, Saint-Louis, Missouri, États-Unis) est un milieu minimal permettant la croissance sélective des levures. Ce milieu (SD-Lac ; glucose 2 % ; *yeast nitrogen base without amino acids* 6,7 g/L) est complété avec différents mélanges d'acides aminés (*Drop Out* : 1,62 g/L SD-leucine ; 0,74 g/L SD-tryptophane ; 0,65 g/L SD-leucine-tryptophane ; 0,62 g/L SD-leucine-tryptophane-histidine) dépourvus des acides aminés nécessaires à la sélection des transformants et aux tests d'activation de la transcription lors des tests Y2H. Le milieu peut être supplémenté de 1,5 % d'agar

### 4.5. Milieu MS solide

Le milieu Murashige et Skoog permet la culture de plantes *in vitro*. Il est composé de micro- et macroéléments M0255 (*Duchefa Farma*<sup>®</sup>, Haarlem, Pays-Bas) de saccharose 1 % (p/v) et d'agar 1,5 % (p/v) à un pH ajusté à 5,8. Il peut être supplémenté d'antibiotiques (**tableau 2**) ou d'agent génotoxiques pour les tests de résistance aux stress.



**Figure 51 : Principe du clonage GATEWAY®.** La BP clonase permet la recombinaison entre les sites *attB* et *attP*, transférant la séquence d'intérêt (en vert) dans le vecteur d'entrée, entourée bornée de sites *attL*. La seconde réaction réalisée à l'aide de la LR clonase permet la recombinaison entre les sites *attL* et *attR* pour transférer la séquence d'intérêt dans le vecteur de destination. Chaque réaction forme également un sous-produit exprimant un gène toxique pour les bactéries.

## 5. Vecteurs plasmidiques

### 5.1. pUCAP35S

Ce plasmide dérivé de pUCAP (van Engelen *et al.*, 1995) a été construit au laboratoire. Il est utilisé pour le clonage de gènes (ou partie de gènes) en fusion avec la séquence de la eGFP en C-terminal. La construction est ensuite utilisée pour la localisation intracellulaire de protéine *via* une approche biolistique. Deux copies du promoteur 35S du CaMV en tandem permettent la surexpression de la protéine de fusion dans la plante, conférant une meilleure visualisation de celle-ci. Le plasmide porte également le gène de résistance à l'ampicilline pour la sélection des transformants.

### 5.2. pGEM-T Easy

Le pGEM-T *Easy* (Promega®, Madison, Wisconsin, États-Unis) est utilisé pour le clonage de produits de PCR. Le vecteur est vendu linéarisé avec une thymidine sortante à chaque extrémité 3'OH pour permettre l'insertion de produits de PCR synthétisés avec des adénosines sortantes aux extrémités 3'OH. Cette adénosine supplémentaire peut être ajoutée avec une *Taq* ADN polymérase classique après synthèse du produit de PCR par une polymérase thermostable plus fidèle. Le pGEM-T *Easy* est un vecteur à haut nombre de copies. Il contient les promoteurs des ARN polymérases des phages T7 et SP6 flanquant la cassette de clonage au milieu du gène de la  $\beta$ -galactosidase permettant une sélection blanc-bleu des colonies en présence d'IPTG et de X-gal. Le pGEM-T *Easy* porte le gène de résistance à l'ampicilline.

### 5.3. pBluescript II SK+

Le pBluescript II SK+ est le vecteur dans lequel la construction TRCF (At3g02060) adaptée pour le code génétique d'*E. coli* a été obtenue. Il permet la sélection blanc-bleu des transformants et porte un gène de résistance à l'ampicilline.

### 5.4. pDONR/Zeo

Le pDONR/Zeo est un vecteur donneur possédant les sites de recombinaison *attP1* et *attP2* nécessaires au clonage de type *Gateway*®. Cette méthode est basée sur le système de recombinaison du phage lambda utilisé pour s'intégrer dans le génome d'*E. coli*. Le système requiert les séquences de recombinaison appelées *Gateway att* et un mélange de facteur de recombinaison appelé clonase (**figure 51**). La recombinaison des sites *attB* et *attP* permet de générer un vecteur d'entrée possédant la séquence d'intérêt dans un vecteur donneur. Cette séquence pourra ensuite être facilement transférée dans de multiples vecteurs de destination *via* une nouvelle réaction de recombinaison entre les sites *attL* et *attR* permettant ainsi

d'obtenir la séquence dans différents contextes (expression de protéine dans différents systèmes, avec différentes étiquettes ou fusion, etc.).

Comme tous les plasmides *Gateway*<sup>®</sup>, le pDONR/Zeo possède une cassette de clonage contenant un gène de résistance au chloramphénicol et le gène suicide *ccdB* dont le produit interfère avec l'ADN gyrase bactérienne. Le gène *ccdB* permet ainsi d'éliminer les bactéries non-recombinantes. Le pDONR/Zeo possède également un gène de résistance à la zéocine pour sa sélection.

### 5.5. pGBW13 et pGWB14

Les plasmides de destination type pGWB sont des vecteurs binaires présentant les sites de recombinaison *attR1* et *attR2* nécessaires au clonage *Gateway*<sup>®</sup>, permettant l'insertion d'une séquence dans différents contextes en fonction du vecteur choisi. Le plasmide pGWB14 permet la fusion d'une étiquette HA en C-terminal de la séquence insérée par recombinaison, sous le contrôle du promoteur 35S du CaMV. Le pGWB13 permet également la fusion C-terminale d'une étiquette HA, mais ne présente aucun promoteur. La séquence recombinée devra donc posséder son propre promoteur pour être exprimée. Tous les vecteurs pGWB confèrent une résistance à la kanamycine pour la sélection bactérienne, ainsi qu'une résistance à l'hygromycine B pour la sélection des transformants chez la plante.

### 5.6. pGADT7 et pGADgwT7

Le pGADT7 permet l'expression de protéines en fusion avec le domaine activateur de la transcription de GAL4 (GAL4AD) en N-terminal. Dans la levure, l'expression des protéines de fusion est sous le contrôle du promoteur ADH1 (constitutif) et du terminateur ADH1. Les protéines de fusion sont adressées au noyau grâce au signal de localisation nucléaire (NLS : *nuclear localization signal*) du SV40 (Virus simien 40). Le pGADT7 contient également un promoteur T7, une étiquette HA fusionnée à la séquence GAL4AD, ainsi qu'une cassette de clonage. Cette cassette de clonage est remplacée par une cassette *Gateway*<sup>®</sup> (avec un gène de résistance au chloramphénicol et le gène *ccdB*) dans le cas du pGADgwT7. Les origines de réplication pUC et 2  $\mu$  permettent au vecteur d'être répliqué dans *E. coli* et *S. cerevisiae* respectivement. Un gène de résistance à l'ampicilline permet la sélection dans *E. coli* tandis que le gène *LEU2* permet la sélection chez *S. cerevisiae*.

### 5.7. pGBKT7, pGBKgwT7 et pGCKCg

Le pGBKT7 permet l'expression de protéines de fusion avec le domaine de liaison à l'ADN de GAL4 (GAL4BD) en N-terminal pour les pGBKT7 et pGBKgwT7 ou en C-terminal pour le pGCKCg. Dans la levure, l'expression des protéines de fusion est sous contrôle du promoteur ADH1 et du terminateur ADH1. Les protéines de fusion sont adressées au noyau grâce au signal

de localisation nucléaire NLS du SV40. Le pGADT7 contient également un promoteur T7, une étiquette *c-myc* fusionnée à la séquence GAL4BD, ainsi qu'une cassette de clonage. Cette cassette de clonage est remplacée par une cassette *Gateway*<sup>®</sup> dans le cas des pGBKgwT7 et pGBKcg. Les origines de répllication pUC et 2  $\mu$  permettent au vecteur d'être répliqué dans *E. coli* et *S. cerevisiae* respectivement. Un gène de résistance à la kanamycine permet la sélection dans *E. coli* tandis que le gène *TRP1* permet la sélection chez *S. cerevisiae*.

#### 5.8. pACYC184LacZ

Le pACYC184LacZ a été construit au laboratoire à partir du pACYC184 dans lequel le gène *lacZ* contenant une cassette de clonage multi-sites a été incorporé, permettant ainsi une sélection blanc-bleu des transformants (Chang et Cohen, 1978; Miller-Messmer *et al.*, 2012). C'est un plasmide à faible nombre de copies (origine de répllication p15A) permettant une expression de protéines plus proches des conditions physiologiques des bactéries. C'est pourquoi il a été utilisé dans les expériences de complémentation fonctionnelle des mutants bactériens. La cassette de clonage est sous le contrôle du promoteur *lac* inductible à l'IPTG. Le plasmide contient un gène de résistance au chloramphénicol pour sa sélection.

#### 5.9. pET28A

Le pET28A a été utilisé pour l'expression de protéines en système bactérien. Il permet la fusion d'une étiquette 6-histidine en N-terminal de la protéine d'intérêt pour permettre sa purification sur colonne d'affinité au nickel. L'expression de la protéine de fusion est contrôlée par le promoteur et le terminateur de l'ARN polymérase T7. Le plasmide porte le gène *lacI* codant le répresseur du promoteur *lac* pour réprimer l'expression basale de l'ARN polymérase T7 dans les souches bactériennes dérivées de BL21. Il possède également le gène de résistance à la kanamycine pour sa sélection.

#### 5.10. pGW-YN

Le pGW-YN permet l'expression de protéines avec la partie N-terminale de l'YFP. C'est un vecteur binaire permettant la transformation transitoire de plante par agro-infiltration pour étudier les interactions protéine-protéine par BiFC. Dans la plante, l'expression des protéines de fusion est assurée par le promoteur 35S du CaMV. Le vecteur contient une cassette de clonage *Gateway*<sup>®</sup> en amont de la séquence de la partie N-terminale de l'YFP. Un gène de résistance à la kanamycine permet la sélection dans *E. coli*.



### 5.11. pGW-YC

En tout similaire au pGW-YN, sauf que la fusion est avec la séquence de la partie C-terminale de l'YFP. Le pGW-YC est utilisé pour étudier les interactions protéine-protéine par BiFC en combinaison avec le pGW-YN.

### 5.12. pMDC162

Le pMDC162 permet l'expression de construction associant le promoteur d'un gène d'intérêt avec la séquence du gène de la  $\beta$ -glucuronidase (GUS) (Curtis et Grossniklaus, 2003). C'est un vecteur binaire possédant une cassette de clonage *Gateway*<sup>®</sup> en amont du gène GUS pour cloner le promoteur d'intérêt. Le plasmide confère une résistance à la kanamycine aux bactéries et une résistance à l'hygromycine aux plantes transformées, permettant leur sélection.

## 6. Gène de synthèse

Un gène TRCF de synthèse adapté à la fréquence d'utilisation des codons chez *E. coli* a été construit par *Genecust* et cloné dans pBluescript II SK+.

## 7. Oligonucléotides

La liste des amorces utilisées pour les diverses PCR et qPCR est présentée en annexe.

## 8. Outils informatiques

Les comparaisons et alignements de séquences, les assemblages informatiques des constructions, les analyses de restriction et la phylogénie ont été réalisés avec le logiciel MacVector (Mac Vector, Inc.).

Le logiciel pour créer les arbres phylogénétiques était PhyMLv3 : <http://www.phylogeny.fr/> (Dereeper *et al.*, 2008).

Les modélisations des structures de TRCF et RADA ont été réalisées avec Pymol (<https://pymol.org/2/>), à partir des protéines cristallisées d'*E. coli* et de *S. pneumoniae*, respectivement (Deaconescu *et al.*, 2006; Marie *et al.*, 2017).

L'analyse statistique des résultats de spectrométrie de masse, les *MDS-plot* et les volcano-plot ont été réalisés avec le package R<sup>®</sup> IPinquiry développé à l'institut par Hélène Zuber.

Les quantifications relatives de bandes d'ADN sur gel d'électrophorèse colorées au bromure d'éthidium ont été réalisées avec ImageJ (<https://imagej.nih.gov/ij/>)

Les outils et bases de données de NCBI ont été utilisés pour les recherches bibliographiques et la recherche d'homologies (BLAST) : <https://www.ncbi.nlm.nih.gov>

L'étude *in silico* des gènes d'*A. thaliana* a été réalisée avec les bases de données TAIR (<https://www.arabidopsis.org>) et ePLANT (<https://bar.utoronto.ca/eplant>)

La localisation subcellulaire des différentes protéines ont été prédites grâce à SUBA4 (<http://suba.live>), Predotar (<https://urgi.versailles.inra.fr/Tools/Predotar>) et TargetP (<http://www.cbs.dtu.dk/services/TargetP/>)

## Méthodes

### 1. Préparation de bactéries électro-compétentes

Une culture d'*E. coli* DH5 $\alpha$  ou TOP10 de 50 mL estensemencée avec 1 mL d'une préculture ayant poussée une nuit à 37 °C sous agitation. La culture est incubée à 37 °C sous agitation jusqu'à l'obtention d'une DO<sub>600 nm</sub> = 0,4. Après l'avoir laissée sur glace pendant 15 min, la culture est centrifugée à 2 700g pendant 6 min à 4 °C. Le culot obtenu est lavé deux fois avec 30 mL d'eau stérile puis deux fois avec 10 mL de glycérol 10 % (v/v) avant d'être resuspendu dans 200  $\mu$ L de glycérol 10 %. Enfin, cette préparation est répartie en aliquotes de 50  $\mu$ L congelées dans l'azote liquide puis conservées à -80 °C.

### 2. Préparation de levures chimio-compétentes

Une culture de *S. cerevisiae* AH109 ou Y187 de 50 mL estensemencée avec le volume nécessaire d'une préculture ayant poussée une nuit à 30 °C sous agitation pour obtenir une DO<sub>600 nm</sub> = 0,1-0,2. La culture est incubée à 30 °C sous agitation jusqu'à l'obtention d'une DO<sub>600 nm</sub> = 0,4-0,6 puis centrifugée à 5 000 g pendant 3 min à 4 °C. Le culot est lavé trois fois avec 2 mL d'eau stérile puis deux fois avec 2 mL de Te/LiAc 1 X (Tris-HCl 10 mM pH 7,5 ; EDTA 1 mM ; acétate de lithium 100 mM) avant d'être resuspendu dans 250  $\mu$ L TE/LiAc 1 X.

### 3. Test de complémentation du mutant bactérien *radA* par RADA

#### 3.1. Complémentation du mutant bactérien

Pour les tests de complémentation du mutant bactérien *radA* les souches BW25113 et JW4352 ont été utilisées. L'ADN complémentaire du gène *RADA* d'*A. thaliana* (séquence codant les acides aminés 137 à 627) et la séquence du gène *radA* d'*E. coli* ont été clonés dans pACYC184*LacZ*. Ces constructions ont ensuite servi à transformer la souche JW4352. Comme contrôle les deux souches ont également été transformées avec le vecteur pACYC184*LacZ* vide.

### 3.2. Sensibilité aux UV

Pour chaque construction, une culture bactérienne (LB + chloramphénicol + IPTG) est réalisée à partir d'une préculture incubée sur la nuit. La culture est incubée à 37 °C sous agitation jusqu'à l'obtention d'une  $DO_{600\text{ nm}} = 0,3$ . Les bactéries sont alors placées dans la glace pour stopper leur croissance et sont ensuite diluées 10 000 fois. La dilution est étalée sur des boîtes de LB gélosé supplémenté d'IPTG, incubées à 37 °C à l'obscurité pendant 25 min. Les bactéries sont irradiées à l'obscurité sous lumière UV (254 nm) pour 0, 30, 60, 80 ou 100 J/m<sup>2</sup> sous le contrôle d'un radiomètre. Les irradiations sont réalisées deux fois pour chaque condition. Pour l'irradiation à 100 J/m<sup>2</sup>, une dilution de bactéries au millième est utilisée pour faciliter l'analyse des résultats. Enfin, les bactéries irradiées sont incubées à 37 °C jusqu'au lendemain à l'obscurité pour empêcher l'activité des photolyases.

### 3.3. Sensibilité à la CIP

Pour chaque construction, une culture de bactérie (LB + chloramphénicol) est réalisée à partir d'une préculture incubée sur la nuit. La culture est incubée à 37 °C sous agitation jusqu'à l'obtention d'une  $DO_{600\text{ nm}} = 0,4$  où 2,5 mM d'IPTG sont ajoutés pour induire l'expression du gène porté par le pACYC184LacZ. Les bactéries sont incubées jusqu'à l'obtention d'une  $DO_{600\text{ nm}} = 1,2$ . Les cultures bactériennes sont diluées 10 000 fois puis étalées sur des boîtes de LB gélosé supplémenté de 2 mM d'IPTG, de chloramphénicol et de 9 à 21 mM de ciprofloxacine. Ces boîtes sont incubées à 37 °C sur la nuit.

## 4. Méthodes relatives aux acides nucléiques

### 4.1. Quantification

Les acides nucléiques sont quantifiés par mesure de l'absorbance à 260 nm ( $A_{260}$ ) grâce à un spectrophomètre *NanoDrop 2000* (*Thermo Fisher Scientific*®, Waltham, Massachusetts, États-Unis). Les rapports d'absorbance  $A_{260}/A_{280}$  et  $A_{260}/A_{230}$  permettent de mesurer la contamination par des protéines et des impuretés (composés phénoliques, etc.) respectivement. Une préparation d'acides nucléiques de bonne qualité doit présenter un rapport  $A_{260}/A_{280}$  supérieur à 1,8 et un rapport  $A_{260}/A_{230}$  proche de 2.

### 4.2. Électrophorèse sur gel d'agarose

L'électrophorèse sur gel d'agarose permet de séparer les acides nucléiques à travers un maillage tridimensionnel en fonction de leur taille grâce à un courant électrique. Les analyse d'ADN sont réalisées avec des gels d'agarose de 0.8 % à 2 % (p/v) avec du tampon Tris Acétate EDTA (TAE 0,5 X : Tris-acétate 20mM pH 8,0 ; EDTA 0,5 mM). Les échantillons sont préparés en ajoutant un sixième du volume de tampon de charge (Tris-HCl 10 mM pH 7,6 ; Bleu de

Tableau 3 : Protocole de rétrotranscription

2 µg ARN traités à la DNase I 1 µL oligo-dT 50 µM 1 µL hexamères aléatoires 200 ng/µL 1 µL dNTP 10 mM H <sub>2</sub> O qsp 13 µL
Incubation à 65 °C pendant 5 min
Incubation à 0 °C pendant 5 min
4 µL Tampon <i>Superscript IV 5X (Invitrogen)</i> 1 µL DTT 100 mM 1 µL <i>RNaseOUT 40 U /µL (Invitrogen)</i> 1 µL <i>SuperScript IV 200 U/µL (Invitrogen)</i>
Incubation à 25 °C pendant 5 min
Incubation à 50 °C pendant 60 min
Incubation à 70 °C pendant 15 min

bromophénol 0,03 % p/v ; glycérol 60 % ; EDTA 60 mM). La migration est généralement effectuée à 100 V (50-150 V). Les acides nucléiques sont révélés au bromure d'éthidium (BET) directement incorporé dans le gel d'agarose (BET 0,5 µg/mL). Le BET est un agent intercalant de l'ADN observable sous lumière UV. La taille des fragments d'ADN est comparée avec un marqueur de masse moléculaire *MassRuler DNA Ladder* (Thermo Fisher Scientific®) contenant des fragments d'ADN de taille connue.

## 5. Méthodes relatives à l'ARN

### 5.1. Extraction d'ARN

Les ARN sont extraits à partir de feuilles ou de plantules entières congelées dans l'azote liquide puis broyées à froid avec un broyeur *TissueLyser II* (Qiagen®, Hilden, Allemagne) pendant une minute à 25 Hz. Le broyat est repris dans 1 mL de Trizol (*Life Technologies*®, Carlsbad, Californie, États-Unis) puis mélangé jusqu'à décongélation à température ambiante. 200 µL de chloroforme sont ajoutés puis le mélange est incubé sur glace pendant cinq min après émulsion. Le mélange est ensuite centrifugé à 16 000 g pendant 15 min à 4 °C. La phase supérieure aqueuse contenant les ARN est prélevée puis précipitée avec deux volumes d'éthanol absolu et de l'acétate de sodium 300 mM pH 5 pendant au moins 30 min à -20 °C. Enfin, les ARN sont culotés par centrifugation à 16 000 g pendant 30 min à 4 °C, lavés à l'éthanol 70 %, séchés à l'air libre puis repris dans 60 µL d'eau stérile. Ces ARN sont ensuite analysés par spectrophotométrie et électrophorèse sur gel d'agarose.

### 5.2. Traitement à la DNase

Pour éliminer l'ADN contaminant, les ARN extraits sont soumis à une phase de digestion par la DNase I. Pour cela, 5 µg d'ARN sont incubés dans 100 µL de milieu réactionnel (tampon de DNase RQ1 1 X (Tris-HCl 400 mM pH 8 ; MgSO<sub>4</sub> 100 mM ; CaCl<sub>2</sub> 10 mM) ; RNase-free DNase RQ1 5 U (*Promega*®)) pendant 45 min à 37 °C. Le volume de la réaction est ajusté à 300 µL puis les ARN sont purifiés par une extraction au phénol/chloroforme (1:1 v/v) suivie de précipitation à l'éthanol, comme décrit ci-dessus.

### 5.3. Réaction de transcription inverse

Les ADN complémentaires (ADNc) aux ARN sont synthétisés avec une ADN polymérase ARN-dépendante selon les conditions présentées dans le **tableau 3**. Ces ADNc serviront ensuite de matrice pour les réactions de PCR et de PCR quantitative (qPCR).

Tableau 4 : Conditions de PCR en fonction de l'ADN polymérase thermostable utilisée

Type de cycle	Phire Hot Start II		Phusion High Fidelity		Nombre de cycles
	Température	Temps	Température	Temps	
Dénaturation initiale	Démarrage à chaud ( <i>hot start</i> )		Démarrage à chaud ( <i>hot start</i> )		1
	98 °C	30-90 s	98 °C	30-90 s	
Dénaturation	98 °C	10 s	98 °C	5 s	30-35
Hybridation	52 – 62 °C	10 s	60 – 72 °C	5 s	
Elongation	72 °C	15 s/kb	72 °C	25 s/kb	
Elongation finale	72 °C	5 min	72 °C	5 min	1
	4 °C	pause	4 °C	pause	

## 6. Méthodes relatives à l'ADN

### 6.1. Amplification par PCR

La réaction de polymérisation en chaîne (PCR) permet l'amplification *in vitro* de fragments d'ADN. Deux amorces complémentaires à une séquence spécifique de la région à amplifier, en amont sur un brin et en aval sur l'autre brin, vont permettre la synthèse exponentielle d'ADN bicaténaire grâce à une ADN polymérase thermostable. La réaction est débutée par la dénaturation de l'ADN matrice à haute température. Dans le cas des PCR réalisées directement sur des colonies bactériennes, cette étape permet également de casser la paroi des bactéries. Ensuite, les amorces s'hybrident à l'ADN simple brin à une température dépendant du T<sub>m</sub> de celles-ci (généralement compris entre 52 °C et 72 °C) puis l'ADN polymérase thermostable synthétise le brin complémentaire de l'ADN. Ces trois étapes (dénaturation, hybridation, élongation) sont répétées de 30 à 35 fois pour permettre l'amplification exponentielle de l'ADN d'intérêt. Une phase d'élongation finale plus longue permet de terminer l'élongation de tous les brins.

Différentes ADN polymérases thermostables ont été utilisées en fonction de l'expérience menée. Chacune d'elle possède des conditions (températures et durées des étapes) spécifiques présentées dans le **tableau 4**. L'enzyme Phusion® High Fidelity DNA polymerase (Thermo Fisher Scientific®) a été utilisée pour la synthèse fidèle des séquences à cloner tandis que la Phire Hot Start II DNA polymerase (Thermo Fisher Scientific®), moins onéreuse et moins sensible à la qualité des échantillons a été préférée pour les expériences de géotypage, où la fidélité n'était pas recherchée.

### 6.2. Techniques de clonage

#### 6.2.1. Digestion par des enzymes de restriction

Les enzymes de restriction sont des endonucléases d'origine bactérienne utilisées pour la défense contre les bactériophages. Elles reconnaissent et clivent des séquences d'ADN double brin palindromiques spécifiques, générant des fragments aux extrémités cohésives 5' ou 3' sortantes ou des extrémités franches.

Les digestions ont été réalisées avec les enzymes de restriction de Promega® ou de Thermo Fisher Scientific® (FastDigest™) en utilisant les tampons optimaux fournis et selon les instructions des fabricants. Les produits digérés sont ensuite purifiés par une extraction au phénol-chloroforme.

### 6.2.2. Déphosphorylation d'un vecteur

La déphosphorylation du vecteur favorise l'étape de ligation avec le fragment d'ADN à cloner en empêchant le vecteur de se refermer sur lui-même. Elle est réalisée à l'aide d'une phosphatase qui supprime les groupements phosphates 5'-terminaux du vecteur. Les groupements phosphates nécessaires à la ligation seront ensuite apportés par l'insert.

La réaction est réalisée à la suite de la digestion, dans le même tampon que celle-ci avant l'extraction phénolique, la *FastAP Thermosensitive Alkaline Phosphatase (Thermo Fisher Scientific®)* utilisée ne présentant pas de spécificité étroite vis-à-vis d'un tampon. Après une incubation de 40 min à 37 °C, les ADN digérés et déphosphorylés sont purifiés par une extraction au phénol-chloroforme.

### 6.2.3. Purification des produits de PCR

- Sur colonne de silice

Le clonage d'un fragment d'ADN dans un vecteur nécessite la purification du produit désiré après sa synthèse par PCR. Dans le cas où l'amplification ne produit qu'un seul fragment observable sur gel d'agarose, le produit est purifié avec le kit *PCR and gel clean-up (Macherey-Nagel®, Hœrdt, France)*. L'ADN est fixé sur une membrane de silice en présence d'agents chaotropiques, lavé avec une solution éthanolique pour éliminer les contaminants puis élué avec un tampon à faible force ionique (Tris-HCl 5 mM pH 8,5 ; EDTA 0,1 mM).

- Sur gel d'agarose

Dans le cas où la synthèse du fragment d'ADN désiré présente des produits multiples après électrophorèse sur gel d'agarose, le fragment d'intérêt est excisé du gel sous lumière UV. L'ADN est purifié avec le kit *PCR and gel clean-up (Macherey-Nagel®)*. L'agarose contenant l'ADN est fondue dans un tampon contenant du perchlorate de sodium, puis fixée sur une membrane de silice et purifiée avec les mêmes conditions que ci-dessus.

### 6.2.4. Ligation

- Ligation d'un fragment dans un vecteur linéarisé

La ligation linéarisé est réalisée dans un volume de 10 µL avec approximativement 100 ng de vecteur, le fragment d'ADN à cloner en quantités équimolaires à celle du vecteur, du tampon de ligation rapide (*Promega®*: Tris-HCl 30 mM pH 7,8 ; MgCl<sub>2</sub> 10 mM ; DTT 10 mM ; ATP 1 mM ; PEG-3000 5 % (p/v)), et de 3 U de T4 ADN ligase (*Promega®*). La réaction est incubée à 4 °C toute la nuit et servira directement pour la transformation de bactérie.



- Ajout d'adénosines et ligation dans le pGEM-T Easy

Comme l'ADN polymérase *Phusion High Fidelity* utilisée pour la synthèse des fragments à cloner génère des extrémités franches, l'ajout d'une adénosine à chaque extrémité 3'OH est nécessaire pour permettre le clonage dans le pGEM-T Easy. Cette méthode nommée *A-tailing* est réalisée sur un produit de PCR purifié grâce à une *Taq* ADN polymérase classique.

Le milieu réactionnel de 10 µL est composé de 500 ng d'ADN purifié, du tampon de l'enzyme (Tris-HCl 10 mM pH 8,3 ; KCl 50 mM ; MgCl<sub>2</sub> 1,5 mM), de 1 U de *Taq* ADN polymérase et de 0,2 mM d'ATP. Il est incubé 30 min à 72 °C puis directement utilisé pour la ligation avec *pGEM-T Easy*.

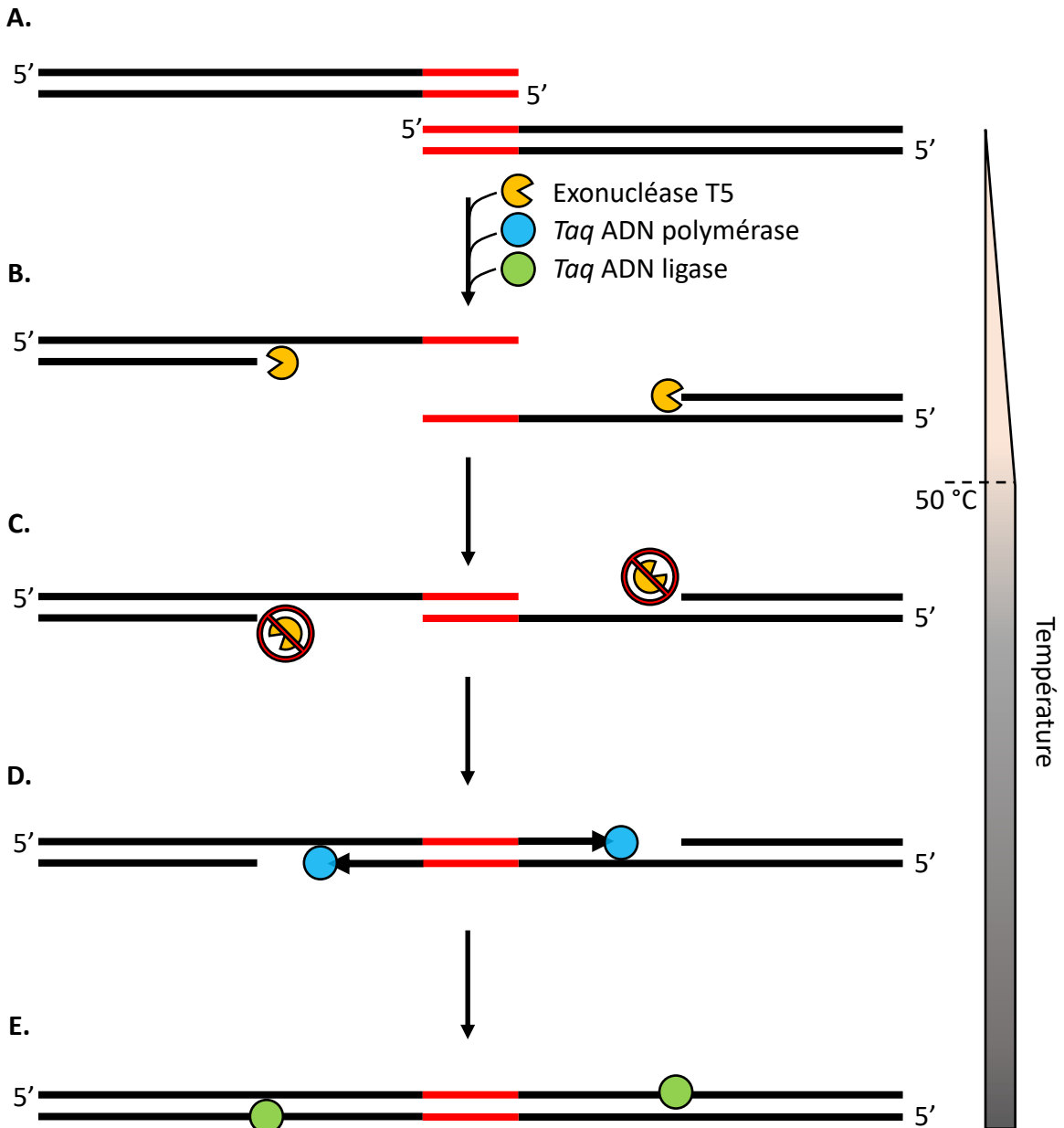
Le milieu réactionnel de 10 µL pour la ligation est composé de 50 ng de *pGEM-T Easy* (*Promega*®), d'approximativement 100 ng d'ADN issu de la réaction de *A-tailing*, du tampon de ligation rapide (*Promega*® : Tris-HCl 30 mM pH 7,8 ; MgCl<sub>2</sub> 10 mM ; DTT 10 mM ; ATP 1 mM ; PEG-3000 5 % (p/v)), et de 3 U de T4 ADN ligase (*Promega*®). Le mélange est incubé sur la nuit à 4 °C et servira directement pour la transformation de bactérie.

#### 6.2.5. Clonage Gateway® : réactions BP et LR

Certains clonages ont été réalisés en utilisant la technologie *Gateway*® comme expliqué dans la **figure 51**. L'insertion de la séquence d'intérêt bordée de séquences *att* dans les différents vecteurs est réalisée par recombinaison.

La première étape du clonage nécessite la recombinaison des séquences *attB* flanquant l'ADN d'intérêt avec les séquences *attP* du vecteur donneur. La réaction est réalisée dans un volume de 5 µL avec 50 fmol de la séquence d'intérêt, 50 fmol de vecteur donneur et 1 µL de mélange enzymatique *BP clonase II* (*Invitrogen*®/ *Thermo Fisher Scientific*®). La réaction se fait à 25 °C sur la nuit puis stoppée par l'ajout de 2 µg de protéinase K (*Thermo Fisher Scientific*®) lors d'une incubation de 10 min à 37 °C. Enfin, 2 µL de milieu réactionnel sont utilisés pour transformer des bactéries TOP10 sensibles au produit du gène *ccdB*.

La seconde étape implique le vecteur d'entrée issue du premier clonage et présentant les séquences *attL* et le vecteur de destination possédant les séquences *attR*. La réaction est réalisée dans un volume de 5 µL avec 50 fmol de vecteur d'entrée, 50 fmol de vecteur de destination et 1 µL de mélange enzymatique *LR clonase II* (*Invitrogen*®/ *Thermo Fisher Scientific*®). La réaction est incubée à 25 °C sur la nuit puis stoppée par l'ajout de 2 µg de protéinase K (*Thermo Fisher Scientific*®) lors d'une incubation de 10 min à 37 °C. Enfin, 2 µL de réaction sont utilisés pour transformer des bactéries TOP10.



**Figure 52 : Principe du clonage *Gibson assembly*®.** La technique Gibson permet le clonage d'un ou plusieurs fragments dans un vecteur sans utiliser de séquences supplémentaires (site de restriction, site de recombinaison, etc.). Vecteur et insert sont d'abord synthétisés par PCR en utilisant des amorces permettant la formation d'extrémités chevauchantes (en rouge) (A). Le mélange réactionnel, comprenant les différentes séquences d'ADN et les trois enzymes, est incubé à 50 °C. L'exonucléase T5 digère rapidement les extrémités 5' des ADN, formant des extrémités 3' sortantes cohésives (B). La température trop élevée inactive l'exonucléase, l'empêchant de dégrader la totalité de l'ADN (C). L'ADN polymérase thermostable synthétise les séquences d'ADN manquantes après hybridation des différents fragments (D). Enfin, une ADN ligase thermostable permet la liaison des extrémités (E).

### 6.2.6. Clonage Gibson®

La technique *Gibson Assembly*® permet l'assemblage rapide et ordonné de multiples fragments d'ADN dans un vecteur, sans ajout de séquences supplémentaires, lors d'une seule réaction. La technique repose sur l'utilisation d'amorces permettant la synthèse de fragments d'ADN possédant des séquences chevauchantes d'au moins 20 nts aux extrémités. Les différents fragments d'ADN (y compris le vecteur) sont d'abord synthétisés par PCR. La réaction permettant la synthèse du vecteur subit ensuite une digestion par l'enzyme *DpnI* afin de supprimer le vecteur original ayant servi de matrice (méthylé). Les fragments sont incubés en présence de l'exonucléase T5, d'une *Taq* ADN polymérase et d'une *Taq* ADN ligase, à 50 °C pendant une heure (**figure 52**). Dans un premier temps, l'exonucléase digère un des brins de l'ADNdb de 5' vers 3', formant ainsi des extrémités 3'OH sortantes capables de s'hybrider entre les différentes molécules. La température de 50 °C inhibe rapidement l'activité de l'exonucléase et empêche la dégradation complète du brin d'ADN, tout en permettant à l'ADN polymérase de synthétiser les parties manquantes de l'ADN. Enfin, les fragments sont reliés grâce à l'activité de l'ADN ligase thermorésistante. Le mélange réactionnel pour une réaction est composé de 0,15 U d'exonucléase T5 et de 80 U de *Taq* ligase, de 0,5 U de *Phusion*® *High Fidelity DNA polymerase* (*Thermo Fisher Scientific*®).

## 6.3. Transformation des bactéries

### 6.3.1. Par choc thermique

Cette technique consiste à placer les bactéries thermo-compétentes en contact avec le plasmide dans un milieu riche en ions calcium sur glace d'abord, pour rigidifier les membranes bactériennes, puis rapidement à 42 °C. Ce choc thermique déstabilise les membranes, créant des pores à travers lesquels l'ADN plasmidique exogène rentre dans la cellule. Puis les bactéries sont replacées sur glace pour permettre aux pores de se refermer.

En pratique, 50 µL de bactéries thermo-compétentes sont décongelées sur glace, puis mises au contact avec 50 à 200 ng d'ADN plasmidique. Le mélange est laissé sur glace au moins 30 min puis placé dans un bain-marie à 42 °C pendant 55 s, avant d'être à nouveau incubé sur glace pendant 2 min. Enfin, 500 µL de LB sont ajoutés et le mélange est incubé à 37 °C pendant une heure pour permettre l'expression des gènes de résistance portés par le plasmide. Les bactéries transformées sont ensuite étalées sur boîte de Pétri supplémentées des antibiotiques adéquats.

### 6.3.2. Par électroporation

Cette méthode de transformation consiste en l'application d'un courant électrique sur les cellules pour augmenter la perméabilité des membranes en créant des pores.

Pour cela, 50  $\mu\text{L}$  de bactéries électro-compétentes sont décongelées sur glace, puis mises au contact de 0,5 à 2 ng d'ADN plasmidique. Le mélange est laissé sur glace pendant au moins 30 min, puis transféré dans une cuve d'électroporation préalablement refroidie. La cuve est ensuite placée dans un électroporateur *Gene Pulser Xcell™* (BioRad®, Hercules, Californie, États-Unis). Les bactéries subissent alors une décharge électrique de 25  $\mu\text{F}$  sous 2,5 kV/cm et 200  $\Omega$ . Après les bactéries sont rapidement additionnées de 700  $\mu\text{L}$  de LB et incubées une heure à 37°C. Enfin, ces bactéries sont étalées sur boîte de Pétri supplémentées des antibiotiques adéquats.

#### 6.4. Transformation de levure par choc thermique

Les levures *S. cerevisiae* AH109 et Y187 sont transformées en utilisant 500 ng de plasmide mélangé à 50  $\mu\text{g}$  d'ADN de sperme de saumon soniqué et bouilli. Ces ADN sont mis en contact avec 50  $\mu\text{L}$  de levures chimio-compétentes fraîchement produites, puis 350  $\mu\text{L}$  de PEG/Te/LiAc (PEG-4000 40 % (p/v) ; Tris-HCl 10 mM pH 7,5 ; EDTA 1 mM ; acétate de lithium 100 mM). Le mélange est incubé à 30 °C pendant 30 min avant de subir un choc thermique à 42 °C pendant 20 min. 700  $\mu\text{L}$  d'eau stérile sont alors ajoutés puis le mélange est centrifugé à 4 000 g pendant une minute. Le culot est repris dans 100  $\mu\text{L}$  d'eau stérile et étalé sur le milieu sélectif approprié : SD-Leu pour AH109 et SD-Trp pour Y187. Enfin, les boîtes sont incubées à 30 °C pendant deux jours.

#### 6.5. Préparation d'ADN plasmidique

L'ADN plasmidique est extrait des bactéries transformées en utilisant les kits *NucleoSpin Plasmid QuickPure* ou *NucleoSpin Plasmid EasyPure* (Macherey-Nagel®). Les colonies transformées sont sélectionnées par PCR puis utilisées pour ensemercer des cultures de 2 à 10 mL de LB supplémentées des antibiotiques adéquats. Après une incubation à 37 °C sous agitation sur la nuit, les bactéries sont sédimentées par centrifugation et lysées dans un tampon alcalin contenant du SDS et de l'hydroxyde de sodium. Un tampon de neutralisation contenant de l'acétate de potassium neutralise le pH du lysat. Les protéines, les débris cellulaires et l'ADN chromosomique attaché aux membranes cellulaires co-sédimentent, tandis que l'ADN plasmidique reste en solution. Cet ADN plasmidique est fixé sur une membrane de silice, lavé avec une solution éthanolique puis élué avec un tampon à faible force ionique (Tris-HCl 10 mM, pH 8 ; EDTA 1 mM). La quantité et la qualité de l'ADN extrait sont ensuite analysées par spectrophotométrie, et l'identité du plasmide confirmée par restriction et par séquençage.

#### 6.6. Séquençage

Les réactions de séquençage sont réalisées par A. Alioua et S. Koechler de la plateforme Analyse de l'expression génique de l'institut. Les ADN sont séquencés selon la méthode de *Sanger* basée sur la synthèse de copies incomplètes du brin d'ADN à séquencer par

Tableau 5 : Conditions des PCR quantitatives

Type de cycle	Température	Durée	Nombre de cycles
Préchauffage	95 °C	10 min	1
Dénaturation	95 °C	15 s	40
Hybridation	58 °C	15 s	
Elongation	72 °C	15 s	
Courbe de fusion	55-95 °C	-	1

incorporation de didésoxyribonucléotides (ddNTP) (Sanger *et al.*, 1977). La méthode a été modifiée pour l'utilisation de la *Taq* ADN polymérase et l'utilisation de ddNTP couplés à des fluorophores (Smith *et al.*, 1986; Murray, 1989). Les séquences sont séparées par électrophorèse capillaire sur un appareil *ABI PRISM® 3100 Genetic Analyzer (Applied Biosystem®*, Foster City, Californie, États-Unis).

### 6.7. Extraction d'ADN total de plante

L'ADN est extrait à partir de feuilles ou de plantules entières congelées dans l'azote liquide, puis broyées à froid avec un broyeur *TissueLyserII (Qiagen®)* à l'aide de billes métalliques pendant une minute à 25 Hz. Le broyat est repris dans 400 µL de tampon CTAB (*cetyl-trimethylammonium bromide* 2 % (p/v) ; NaCl 1,4 M ; EDTA 20 mM ; Tris-HCl 100 mM pH 8) préchauffé à 65 °C et supplémenté de β-mercaptoéthanol 0,2 % (v/v) et de 50 à 100 µg/mL de RNase A. Le mélange est incubé à 65 °C pendant une heure puis supplémenté de 400 µL de chloroforme. Après émulsion le mélange est centrifugé à 16 000 g pendant 10 min. L'ADN contenu dans la phase supérieure aqueuse est précipité à l'éthanol. La qualité et la quantité d'ADN sont ensuite analysées par spectrophotométrie et électrophorèse sur gel d'agarose.

### 6.8. Extraction de l'ADN chromosomique bactérien

L'ADN est extrait à partir d'une culture bactérienne de 2 mL. Les bactéries sont sédimentées par centrifugation à 11 000 g pendant une minute puis repris dans du tampon de lyse (Tris-HCl 10 mM pH 7,4 ; EDTA 1 mM ; SDS 0,7 % (v/v)) supplémenté de 50 µg/mL de RNase A et de 5 µg/mL de protéinase K et incubé à 37 °C pendant 30 min. L'ADN est récupéré après une extraction au phénol/chloroforme (1:1) puis au chloroforme et enfin précipité à l'éthanol.

### 6.9. PCR quantitative

La qPCR permet la quantification d'une séquence d'ADN au sein d'un échantillon grâce à la mesure de la libération de la fluorescence à chaque cycle de PCR. Le nombre de cycles nécessaires pour dépasser le seuil de détection de la fluorescence (Cp : *crossing point*) indique le nombre de molécules initialement présentes dans l'échantillon.

Les réactions de qPCR sont réalisées dans des microplaques optiques de 384 puits sur un appareil *LightCycler 480® Instrument II (Roche®*, Bâle, Suisse) selon les conditions présentées dans le **tableau 5**. Le milieu réactionnel de 6 µL par puits est composé de 0,5 volume de *LightCycler® 480 SYBR Green I Master (Roche®)* de 1,5 µM de chaque amorce et d'approximativement 2 ng d'ADN total par réaction. A la fin des cycles d'amplification, une courbe de fusion est réalisée pour vérifier la spécificité des amorces et la nature unique des

amplicons. Les résultats obtenus sont normalisés par rapports à des gènes de référence endogènes également testés.

## 7. Méthodes relatives aux protéines

### 7.1. Extraction de protéines à partir de matériel végétal

Les protéines sont extraites à partir de feuilles ou de plantules entières congelées dans l'azote liquide puis broyées à froid avec un broyeur *TissueLyserII* (Qiagen®) à l'aide de billes métalliques ( $\varnothing$  3 mm) pendant une minute à 25 Hz. Le broyat est repris dans 50 à 150  $\mu$ L de tampon Laemmli (Tris-HCl 60 mM pH 6,8 ; SDS 2 % (p/v) ;  $\beta$ -mercaptoéthanol 1 % (v/v) ; glycérol 10 % (p/v) ; bleu de bromophénol 0,004 % (p/v)), chauffé à 95 °C pendant 5 min puis centrifugé pour éliminer les débris de broyage.

### 7.2. Extraction de protéines à partir de bactérie

Une culture bactérienne est incubée à 37 °C sous agitation jusqu'à l'obtention d'une  $DO_{600nm} = 2$ , optimalement. Les bactéries sont sédimentées par une centrifugation de 30 s à 11 000 g et le culot est repris dans le volume de tampon de lyse (*BugBuster® Protein Extraction Reagent* 1 X ;  $MgCl_2$  10 mM) nécessaire pour obtenir une  $DO_{600nm} = 20$ . La suspension est supplémentée de 10  $\mu$ g/mL de DNase I, puis incubé à température ambiante pendant 30 minutes sous agitation légère. La réaction est stoppée par l'ajoute d'un volume de Laemmli 2 X (Tris-HCl 120 mM pH 6,8 ; SDS 4 % (p/v) ;  $\beta$ -mercaptoéthanol 2 % (v/v) ; glycérol 20 % (p/v) ; bleu de bromophénol 0,008 % (p/v)) puis chauffé à 95 °C pendant 5 min.

### 7.3. Extraction de protéines à partir de levure

Les levures ou hybrides de levure issus d'une culture liquide sont sédimentés par centrifugation à 11 000 g pendant une minute. Le culot est repris dans 0,2 volume d'hydroxyde de sodium 1 N et incubé 5 min à température ambiante. La suspension est à nouveau centrifugée à 11 000 g pendant une minute puis le broyat est repris dans du tampon Laemmli, chauffé à 95 °C pendant 5 min puis centrifugé pour éliminer les débris de broyage.

### 7.4. Précipitation des protéines

Afin d'éliminer certains contaminants et de concentrer les protéines, celles-ci sont précipitées par la méthode TCA/DOC (acide trichloracétique / deoxycholate de sodium). Un volume de protéine est additionné d'un centième du volume de DOC 0,2 % et incubée à 4 °C pendant 30 minutes. Puis, un dixième du volume de TCA 100 % est ajouté et la réaction est laissée à 4 °C sur la nuit. Les protéines sont ensuite centrifugées à 15 000 g pendant 15 min à 4 °C puis le culot est lavé dans un volume d'acétone 100 % glacial. Après une nouvelle centrifugation à 15 000 g pendant 15 min à 4 °C le culot est séché à l'air puis repris dans du

tampon Laemmli. Si nécessaire (bleu de bromophénol devenant jaune), le pH est ajusté avec du Tris 1 M.

## 7.5. Séparation par électrophorèse sur gel de polyacrylamide en conditions dénaturantes

L'électrophorèse en conditions dénaturantes permet de séparer les protéines à travers un maillage tridimensionnel d'acrylamide/N,N'-méthylènebisacrylamide en fonction de leur taille sous l'effet d'un champ électrique. Le gel est séparé en deux parties : le gel de concentration (Tris-HCl 125 mM pH 6,8 ; acrylamide/N,N'-méthylènebisacrylamide 37,5/1:5 % (p/v) ; SDS 0,1 % ; persulfate d'ammonium (APS) 0,1 % ; tetraméthylethylenediamine (TEMED) 0,05 %) qui concentre les protéines pour qu'elles rentrent dans le gel de séparation de manière uniforme ; et le gel de séparation (Tris-HCl 375 mM pH 8,8 ; acrylamide/N,N'-méthylènebisacrylamide 37,5:1 8-12 % (p/v) ; SDS 0,1 % ; APS 0,1 % ; TEMED 0,05 %). Les échantillons de protéines dans du tampon Laemmli sont dénaturés à 95 °C pendant 5 min avant d'être déposés sur le gel. La migration est effectuée à 25 mA dans du tampon Tris-Glycine-SDS (Tris 25 mM pH 8,3 ; Glycine 192 mM ; SDS 0,1 % (p/v)).

Après migration, les protéines peuvent être transférées sur membrane de PVDF pour réaliser un *western blot*, ou être visualisées par coloration du gel au bleu de Coomassie (acide acétique 10 % (v/v) ; méthanol 40 % (v/v) ; Coomassie brilliant blue R250 0,25 % (p/v)) pendant une heure. Le gel est ensuite décoloré (acide acétique 10 % ; éthanol 20 %) et la taille des protéines est estimée par rapport à l'échelle de protéines *PageRuler™ Prestained Protein Ladder* (Thermo Fisher Scientific®).

## 7.6. Quantification

### 7.6.1. Mesure directe de l'absorbance à 280 nm

La concentration des protéines purifiées est estimée par mesure de l'absorbance à 280 nm ( $A_{280}$ ) grâce à un spectrophomètre *NanoDrop®* (Thermo Fisher Scientific®). Le coefficient d'extinction spécifique de la protéine en solution est alors utilisé.

### 7.6.2. Mesure indirecte par la méthode de Bradford

La quantité de protéines dans un mélange protéique est mesurée grâce à la méthode de Bradford (Bradford, 1976), reposant sur la modification de l'absorbance du bleu de Coomassie de 450 nm à 595 nm en présence de protéines. La quantification est réalisée avec le réactif *Protein Assay Dye Reagent Concentrate* (BioRad®) dilué au sixième dans l'eau. L'absorbance de l'échantillon est comparée à celle d'une gamme étalon réalisée avec des quantités connues de sérum d'albumine bovin (BSA)



### 7.6.3. Utilisation d'une gamme de protéine dans un gel de polyacrylamide

La quantité d'une protéine d'un mélange protéique peut être estimée lors d'une électrophorèse sur gel de polyacrylamide en conditions dénaturantes. Différents volumes de l'échantillon ainsi qu'une gamme de BSA sont déposés sur le gel. Après migration et coloration du gel, l'intensité de la bande protéique d'intérêt est comparée à celles des échantillons de la gamme de BSA dont la quantité déposée sur gel est connue. L'utilisation du logiciel imageJ permet la quantification plus précise de l'intensité des bandes protéiques.

### 7.6.4. Mesure indirecte de la réduction du cuivre en présence de protéine.

Les réactifs de dosage colorimétrique du kit *Pierce™ Rapid Gold BCA Assay* (Thermo Fisher Scientific®) ont été utilisés pour quantifier les protéines d'un mélange protéique en présence des détergents utilisés pour leur extraction, incompatibles avec la méthode de Bradford. Les protéines sont extraites à partir de feuilles, de racines ou de plantules entières congelées dans l'azote liquide puis broyées à froid avec un broyeur *TissueLyserII* (Qiagen®) à l'aide de billes métalliques pendant une minute à 25 Hz. Le broyat est repris dans du tampon SDS-Urée (Tris-HCl 80 mM pH 6,8 ; urée 4 M ; SDS 2 % (p/v)) sans glycérol pour permettre l'utilisation des réactifs de dosage. La quantification est réalisée selon le protocole du fournisseur : 40 µL d'échantillon sont supplémentés de 1 mL de réactifs A et B (50:1). La réaction est incubée à température ambiante pendant 5 min, puis la  $DO_{480\text{ nm}}$  est mesurée. En parallèle, une gamme étalon est réalisée dans les mêmes conditions avec une gamme de BSA de quantité connue.

## **7.7. Surproduction dans E. coli**

### 7.7.1. Production de la protéine RADA en vue de tests in vitro.

Pour réaliser des expériences *in vitro* avec une protéine il est nécessaire de la produire et de la purifier. La fusion d'une étiquette 6-histidines permet la purification d'une protéine par chromatographie d'affinité au nickel.

La souche d'*E. coli* Rosetta 2 (DE3) pLysS (Novagen®) a été transformée avec le vecteur pET28A portant la séquence du gène RADA ou de son mutant RADA[K201A], fusionnée à une étiquette 6-histidines. Les bactéries transformées sont incubées à 37 °C sous agitation jusqu'à l'obtention d'une  $DO_{600\text{ nm}} = 0,6$ , où leur croissance est stoppée sur glace pendant 30 min. Puis 0,5 mM d'IPTG sont ajoutés et la culture est incubée sur la nuit à 18 °C (RADA) pour favoriser l'expression de la protéine recombinante. Les bactéries sont sédimentées par centrifugation à 7 000 g pendant 18 min à 4 °C et le culot est lavé avec de l'eau avant d'être repris dans le volume de tampon de lyse nécessaire pour obtenir un  $DO_{600\text{ nm}} = 20$  (Tris-HCl 50 mM pH 8 ; glycérol 5 % (v/v) ; NaCl 300 mM ; imidazole 10 mM) supplémentés de *Complete™ pretease*

*inhibitor cocktail* (Merck®, Kenilworth, New Jersey, États-Unis) et de 1 mM de PMSF. Les bactéries sont lysées avec une presse de French à 1200 PSI, puis le lysat est soniqué pendant 3 min. Le lysat est ensuite clarifié par une centrifugation à 17 770 g pendant 25 min à 4 °C et le surnageant est filtré à travers un filtre *Filtropur S plus* de 0,2 µm (STARDEST). Les protéines recombinantes sont alors purifiées par affinité sur une colonne *HisTrap FF Crude* (GE Healthcare Life Sciences®) équilibrée avec du tampon de lyse à un débit de 0,5 mL/min. La colonne est lavée avec du tampon de lavage (Tris-HCl 50 mM pH 8 ; glycérol 5 % (v/v) ; NaCl 300 mM ; imidazole 50 mM) puis les protéines sont éluées sur un gradient de 50 à 500 mM d'imidazole. Enfin, les fractions contenant la protéine d'intérêt sont réunies, concentrées, aliquotées et conservées à -80 °C.

#### 7.7.2. Production de TRCF pour immunisation de lapins et production d'anticorps.

La souche d'*E. coli* Rosetta 2 (DE3) pLysS (Novagen®) a été transformée avec le vecteur pET28A portant la séquence du motif CarD (acide aminé 147 à 443) ou du motif SF2 (acide aminé 255 à 656) de TRCF d'*Arabidopsis*. Les bactéries transformées sont incubées à 37 °C sous agitation jusqu'à l'obtention d'une  $DO_{600\text{ nm}} = 0,4$ , puis 0,5 mM d'IPTG sont ajoutés et la culture est incubée à 37 °C pendant deux heures. Les protéines recombinantes ont été par la suite purifiées comme décrit ci-dessus. Afin d'aider la solubilisation des protéines d'intérêt, de l'urée 8 M a été utilisée lors de la lyse des bactéries et la purification. Des anticorps polyclonaux ont été préparés par la plateforme de production d'anticorps de l'IGBMC.

### 7.8. Transfert semi-sec sur membrane de PVDF

Pour détecter ou identifier une protéine dans un gel de polyacrylamide, il est nécessaire de réaliser une immunodétection. Pour cela, l'ensemble des protéines fractionnées sur le gel doivent d'abord être transférées sur une membrane de polyfluorure de vinylidène (PVDF) *Immobilon-P* de 0,45 µm (Merck®). La membrane est mouillée avec de l'éthanol absolu pendant quelques secondes puis équilibrée en même temps que le gel de polyacrylamide dans le tampon de transfert semi-sec (Tris 48 mM pH 8 ; glycine 39 mM) pendant 3 min. Le gel et la membrane sont ensuite assemblés de façon à ce que les protéines soient transférées vers la charge positive, en sandwich entre deux paires de feuilles de papier *Whatman®* 3M. L'électrotransfert est réalisé sous 1 mA/cm<sup>2</sup> de gel pendant 45 à 55 min dans une cuve de transfert semi-sec V10-SDB (*BIOphORETICS™*).

Après le transfert, la membrane peut être colorée avec une solution de coloration (acide acétique 7 % (v/v) ; méthanol 50 % (v/v) ; Coomassie brilliant blue R250 0,1 % (p/v)) pour visualiser les protéines. La membrane est ensuite décolorée avec une solution de décoloration (acide acétique 7 % (v/v) ; méthanol 50 % (v/v)). En général, les membranes sont plutôt colorées après l'immunodétection.

Tableau 6 : Anticorps utilisés

Epitope reconnu	Type	Organisme	Dilution
HA	monoclonal	souris	1/8000
myc	monoclonal	souris	1/10000
6xHis	monoclonal	souris	1/10000
TRCF	polyclonal	lapin	1/8000
RPOB	polyclonal	lapin	1/1000
PCOR	polyclonal	lapin	1/3000
NAD9	polyclonal	lapin	-
H3	polyclonal	lapin	1/20000
IgG de souris	monoclonal	chèvre	1/10000
IgG de lapin	monoclonal	chèvre	1/10000

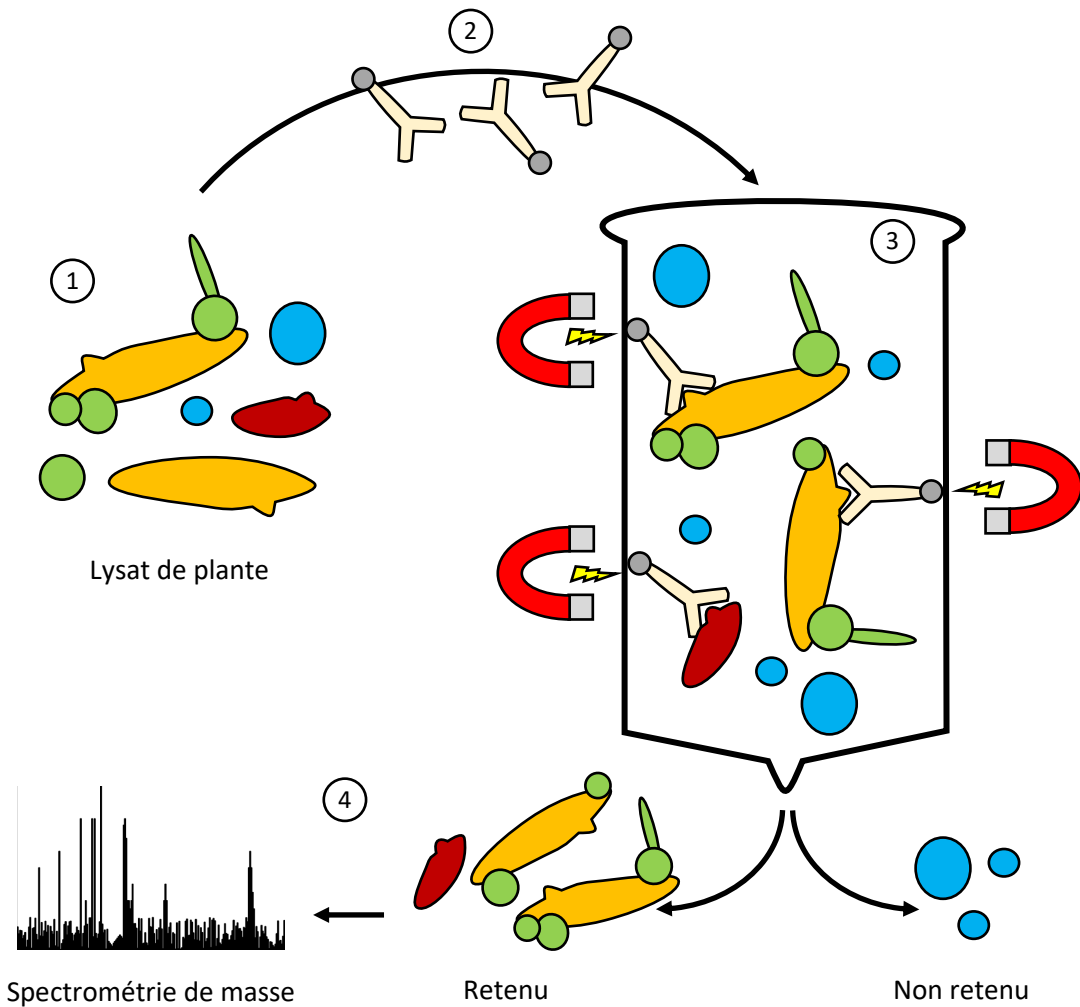
## 7.9. Immunodétection

Après transfert, les protéines d'intérêt sont détectées par des anticorps spécifiques de ces protéines ou d'étiquettes qui leur sont fusionnées. La membrane est préalablement saturée avec des protéines non-spécifiques pour diminuer le bruit de fond, dans du tampon TBS-T (Tris-HCl 20 mM pH 7,4 ; NaCl 150 mM ; Tween®20 0,1 % (v/v)) supplémenté de 5 % de lait écrémé en poudre pendant une heure sous agitation douce. La membrane est ensuite incubée avec l'anticorps primaire (**tableau 6**) dilué dans du TBS-T supplémenté de 2,5 % de lait écrémé à 4 °C sous agitation douce sur la nuit. La membrane est lavée trois fois avec du TBS-T puis incubée avec l'anticorps secondaire anti IgG de lapin ou de souris (**tableau 6**) dilué dans du TBS-T supplémenté de 1 % de lait écrémé pendant une heure sous agitation douce. Après trois nouveaux lavages, l'activité peroxydase des anticorps secondaires est révélée avec l'*Enhanced Chemical Luminescence* (ECL) (Roche®).

## 7.10. Co-immunoprécipitation de partenaires protéiques

La Co-IP permet d'identifier les potentiels partenaires d'interaction d'une protéine d'intérêt. Ces partenaires peuvent ensuite être identifiés par *western blot* ou spectrométrie de masse. La méthode utilisée repose sur l'utilisation d'anticorps reconnaissant la protéine d'intérêt couplés à des billes magnétiques. Après le broyage et la lyse des cellules en conditions natives, la protéine d'intérêt, avec ses interactants, vont être reconnus par les anticorps qui seront retenus dans une colonne grâce aux billes (**figure 53**). Après plusieurs lavages permettant l'élimination des protéines contaminantes la protéine d'intérêt et ses interactants sont élués avec un tampon dénaturant.

Pour une expérience 0,9 à 1 g de plantules de trois ou neuf jours après vernalisation sont broyées à 4 °C avec un mortier et un pilon dans 3 mL de tampon de lyse (Tris-HCl 50 mM pH 8 ; NaCl 100 mM ; Triton X-100 1 % (v/v) ; PMSF 2 mM) supplémentés de *Complete™ protease inhibitor cocktail* (Merck®). Dans le cas de Co-IP réalisée sur des chloroplastes purifiés, 1 mg de protéines chloroplastiques sont repris dans 1 mL de tampon de lyse. Le broyat est clarifié par deux centrifugations à 12 000 g à 4 °C pendant 15 min, puis réparti dans trois tubes traités comme triplicats techniques pour les besoins de la spectrométrie de masse. Le broyat clarifié est ensuite supplémenté de 50 µL de billes magnétiques anti-HA du kit *µMACS HA isolation* (Miltenyi Biotec®, Bergisch Gladbach, Allemagne) puis incubé sur une roue à 5 rpm à 4 °C pendant 30 min. Après la lyse, chaque échantillon est chargé séquentiellement, à raison de 200 µL par 200 µL, dans une colonne M (*Miltenyi Biotec®*) maintenue dans un stand magnétique et pré-équilibrée avec le tampon de lyse. Les colonnes sont ensuite lavées avec cinq à six lavages séquentiels de 200 µL de tampon de lavage (Tris-HCl 50 mM pH 8 ; NaCl 100 mM ; Triton X-100 0,1 % (v/v)). Enfin, l'éluion est réalisée hors du stand avec 130 µL du tampon d'éluion du kit *µMACS HA isolation* (Miltenyi Biotec®).



**Figure 53 : Principe de la co-immunoprécipitation avec anticorps couplés à des billes magnétiques.** La co-immunoprécipitation permet de purifier les partenaires d'interaction d'une protéine d'intérêt pour les identifier. Après le broyage et la lyse non dénaturante à froid des échantillons, le lysat est incubé avec des anticorps dirigés contre une étiquette de la cible (en jaune), couplés à des billes magnétiques. Le lysat est ensuite chargé sur une colonne dans un *stand* magnétique. La rétention des billes permet de retenir la protéine d'intérêt et ses interactants (en vert) alors que les contaminants seront élués (en bleu). Certains contaminants peuvent néanmoins être reconnus à tort par les anticorps et retenus dans la colonne (en rouge). Après plusieurs lavages, toutes les protéines de la colonne sont éluées avec un tampon dénaturant, puis identifiées par spectrométrie de masse.

Les échantillons sont apportés à la plateforme de spectrométrie de masse de l'Institut de Biologie Moléculaire et Cellulaire (IBMC, Strasbourg) où ils sont pris en charge par P. Hammann, L. Kühn et J. Chicher. Les échantillons sont précipités avec cinq volumes de méthanol et 0,1 M d'acétate d'ammonium à -20 °C pendant au moins deux heures. Les échantillons sont ensuite centrifugés à 12 000 g pendant 15 min à 4 °C. Le culot est lavé deux fois avec du méthanol 80 % (v/v) et 0,1 M d'acétate d'ammonium puis à nouveau centrifugé à 12 000 g pendant 15 min à 4 °C. L'échantillon est soniqué pour homogénéiser les billes métalliques dans le méthanol. Le culot est repris dans 50 mM de bicarbonate d'ammonium avec 4,8 mM de DTT puis incubé à 95 °C pendant 10 min. Après une centrifugation rapide, 8,6 mM d'iodoacétamide sont ajoutés et le mélange réactionnel est laissé à l'obscurité durant 45 min. Enfin, 4,2 mM de DTT et une concentration de trypsine 20 fois inférieure à la concentration protéique sont ajoutés. Le mélange est incubé à 37 °C sur la nuit. L'échantillon est ensuite chargé sur une *nanoHPLC Easy-nLC1000* (Thermo Fisher Scientific®) et un spectromètre de masse *QExactive+* (Thermo Fisher Scientific). Les résultats de spectrométrie de masse sont analysés grâce au package R® IPinquiry.

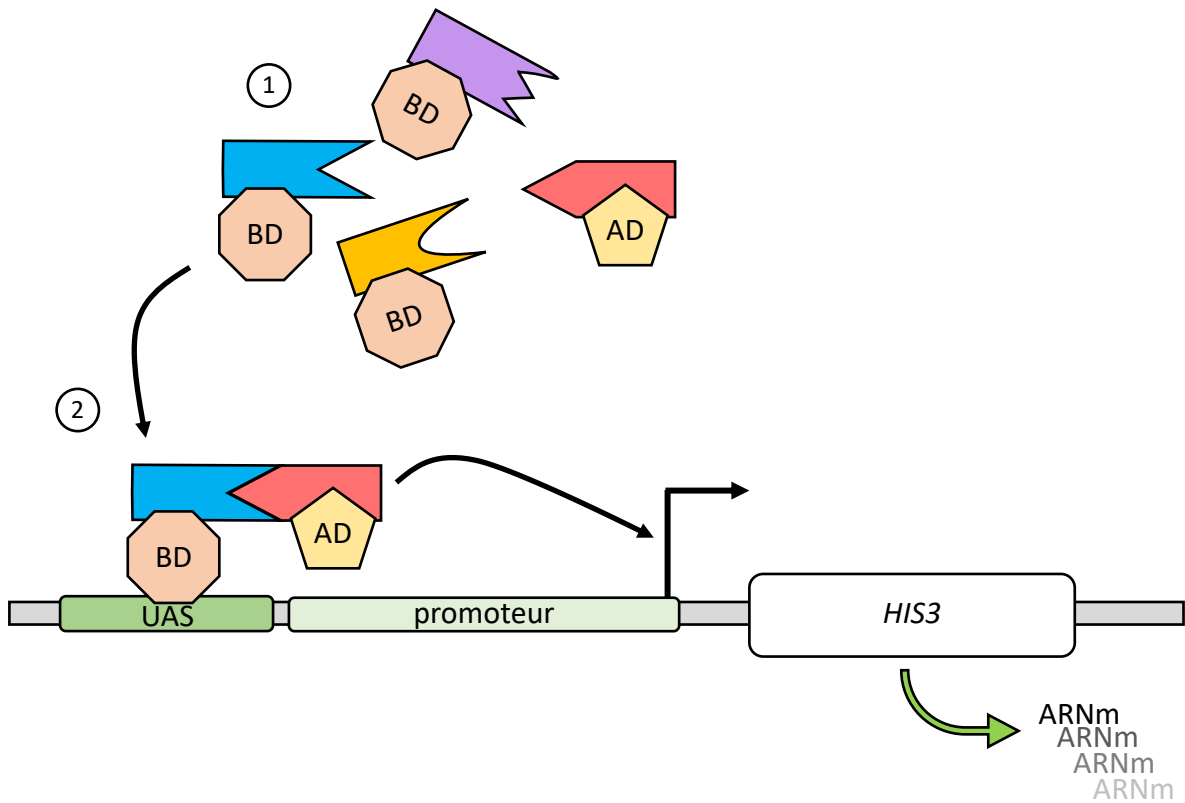
#### 7.11. Etude de l'activation de la transcription par double hybride de levure

La méthode du double hybride de levure Y2H permet d'étudier l'interaction entre deux protéines *in vivo*. Elle repose sur l'expression de protéines fusionnées au domaine de liaison à l'ADN ou au domaine d'activation de la transcription d'un facteur de transcription (GAL4, LexA). L'interaction des protéines fusionnées à chaque domaine permet la reconstitution du facteur de transcription et l'expression *in vivo* d'un gène rapporteur (**figure 54**).

En pratique, deux souches de levures de signe sexuel différents, la souche AH109 *MATa* et la souche Y187 *MATα*, sont transformées avec les plasmides dérivés du pGAD et du pGBK, respectivement. Une culture de levures est réalisée dans le milieu sélectif approprié, puis les cellules sont récupérées par centrifugation à 4 000 g pendant une minute et reprises dans 500 µL d'eau stérile. Les deux souches sont ensuite cultivées dans un même milieu YPD permettant la fusion et la formation d'une cellule diploïde contenant les deux plasmides. Ces hybrides sont sélectionnés sur un milieu SD dépourvu de leucine et de tryptophane (SD-LW) tandis que l'interaction entre les deux protéines est observée dans un milieu dépourvu d'histidine (SD-LWH), révélant l'expression du gène rapporteur *HIS3*.

#### 7.12. Test d'échange de brin *in vitro*

Pour évaluer le rôle de la protéine RADA d'*A. thaliana* dans la formation d'un hétéro-duplex et sa maturation lors de la recombinaison homologue, une expérience *in vitro* d'invasion de brin a été menée avec les protéines RADA purifiées. L'expérience repose sur l'utilisation de deux molécules d'ADN homologues capables de recombiner en présence de la recombinaison RecA et de la protéine de liaison à l'ADNsb SSB bactériennes, et d'un système de



**Figure 54 : Principe du double hybride de levure.** Le double hybride de levure permet d'identifier les partenaires d'interaction d'une protéine d'intérêt. Il repose sur la possibilité de séparer un facteur de transcription en deux domaines. La protéine d'intérêt (la proie) est fusionnée avec la séquence du domaine d'activation de la transcription (AD), tandis que ses interactants potentiels (les appâts) sont fusionnés avec la séquence du domaine de liaison à l'ADN (BD). Lors de l'interaction de la proie avec un appât, le facteur de transcription est reconstitué et permet l'activation d'un promoteur et l'expression d'un gène rapporteur, ici *HIS3*, permettant la croissance des levures. UAS (*upstream activating sequence*) est le site de fixation du facteur de transcription.

régénération de l'ATP (Cooper et Lovett, 2016). Les formes d'ADN initiales, intermédiaires et finales sont discriminables par analyse sur gel d'agarose.

Les réactions de recombinaisons *in vitro* sont réalisées avec l'ADN *single-strand linear*  $\Phi$ X174 virion et *double strand circular*  $\Phi$ X174 RFI (New England Biolabs®, Ipswich, Massachusetts, États-Unis) linéarisé par PstI. En pratique, 20  $\mu$ M d'ADNsb circulaire (en nucléotides), 6,7  $\mu$ M de protéine RecA (New England Biolabs®) et 2  $\mu$ M de protéine RADA purifiée sont incubés dans le tampon (Tris-acétate 20 mM pH 7,4 ; glutamate d'ammonium 3 mM ; acétate d'ammonium 11 mM ; phosphocréatine 12,5 mM ; DTT 1 mM ; créatine kinase 10 U/mL ; glycérol 2 % (p/v) à 37 °C pendant 8 min. Puis, 20  $\mu$ M d'ADNdb linéarisé (en nucléotides) sont ajoutés et la réaction est incubée à 37 °C pendant 5 min. Enfin, la réaction d'invasion de brin est initiée par l'ajout de 3 mM d'ATP et de 3,1  $\mu$ M de protéine SSB. La réaction est incubée à 37 °C et des aliquotes sont prélevées et stoppées par l'addition de 12  $\mu$ M d'EDTA et 0,8 % de SDS (v/v) à différents temps. Les produits de la réaction sont analysés sur un gel d'agarose 0,7 % (p/v) après migration à 5 V/cm et marquage au BET.

Pour les expériences d'échange de brin *in vitro* terminées en l'absence de RecA, la réaction est débutée comme précédemment puis est stoppée au bout de 7 min à 37 °C par une extraction au phénol chloroforme suivie d'une précipitation à l'éthanol. Après lavage, le culot d'ADN est repris dans le tampon de réaction avec 2  $\mu$ M de protéines RADA purifiées. La réaction est incubée à 37 °C puis stoppée par l'ajout de 12  $\mu$ M d'EDTA et 0,8 % de SDS (v/v) à un temps donné. Les résultats sont ensuite analysés comme précédemment.

### 7.13. Test d'interaction avec l'ADN par retard sur gel

Les interactions entre diverses formes d'ADN et les protéines RADA ont été testées par retard sur gel (EMSA : *Electro Mobility Shift Assay*). La sonde est réalisée par marquage de l'extrémité 5'OH de 10 à 50 pmol d'oligonucléotide (annexes) avec 20 pmol de [ $\gamma$ -<sup>32</sup>P]ATP par 10 U de T4 *polynucleotide kinase* (Thermo Fisher Scientific®) à 37 °C pendant 30 min. La réaction est ensuite incubée à 72 °C pendant 5 min pour inactiver l'enzyme puis la sonde est purifiée de la radioactivité libre par centrifugation sur une colonne Sephadex G-50 (Sigma-Aldrich®). Les différentes formes d'ADN utilisés sont générées par hybridation de la sonde avec d'autres oligonucléotides dans du Tris-Acetate 33 mM pH 7.9 ; K-Acetate 66 mM ; Mg-Acetate 10 mM. Elles sont ensuite purifiées sur gel de polyacrylamide non dénaturant (8 % acrylamide/N,N'-méthylènebisacrylamide 29:1) dans du tampon TBE (Tris 90 mM pH 8 ; EDTA 2 mM ; acide borique 90 mM), identifiées par autoradiographie, excisées du gel et éluées dans du tampon TEN (Tris-HCl 10 mM pH 7.5 ; EDTA 1 mM ; NaCl 100 mM) pendant la nuit à 4 °C.



Le retard sur gel est réalisé en incubant des quantités variables de protéine avec 0,1 pmoles sonde radiomarquée (10 000-20 000 cpm) à 20 °C pendant 30 min dans du tampon EMSA (Tris-acétate 20 mM pH 7,5 ; acétate de sodium 50 mM ; acétate de magnésium 5 mM ; glycérol 5 % (v/v) ; EDTA 0,5 mM ; DTT 1 mM). La réaction est analysée sur gel de polyacrylamide 6 % (acrylamide/N,N'-méthylènebisacrylamide 29:1) dans du tampon TBE. La migration est réalisée avec du TBE à 15 V/cm pendant 2 h à 4 °C. Après séchage du gel, la radioactivité est révélée avec un *Phosphorimager* (*Fujifilm FLA7000*<sup>®</sup>, Minato, Tokyo, Japon ; ou *GE Healthcare Typhoon*<sup>®</sup>, GE Healthcare Life Sciences).

## 8. Méthodes relatives aux plantes

### 8.1. Transformation par la méthode biolistique

La méthode biolistique permet la transformation transitoire de matériel végétal *via* l'utilisation de billes d'or ou de tungstène enrobées d'ADN et projetées par un canon à particule. Les bombardements sont effectués sur des feuilles de *N. benthamiana* coupées au pétiole et déposées face adaxiale sur une boîte de Pétri contenant du milieu MS. Le mélange de transformation est composé de 20 à 30 µL d'une suspension de billes d'or (∅ 1 µm) à 30 mg/L dans du glycérol 50 %, 10 µg de plasmide, 3,8 M de CaCl<sub>2</sub> et 60 mM de spermidine. Après une incubation de 20 min à 4 °C sous agitation, les billes recouvertes d'ADN sédimentées sont lavées quatre fois avec de l'éthanol absolu glacial, puis resuspendues dans 35 µL d'éthanol. Les feuilles sont ensuite bombardées avec 15 µL de suspension de billes grâce à un appareil *PDS-1000/He*<sup>TM</sup> (*BioRad*<sup>®</sup>), sous vide (0,1 bar) avec un disque de rupture calibré pour 1 100 psi. Les feuilles transformées sont stockées à l'obscurité sur la nuit, avant d'être découpées pour une observation au microscope confocal.

### 8.2. Transformation transitoire par agro-infiltration

L'agro-infiltration permet de transformer transitoirement des feuilles de *N. benthamiana*. Une culture de 10 mL d'*Agrobacterium tumefaciens* transformée avec le vecteur binaire adéquat est ensemencée à partir d'une préculture, puis incubée à 28 °C sous agitation jusqu'au lendemain. Les bactéries sont sédimentées par une centrifugation à 5 000 g pendant 10 min à température ambiante. Le culot bactérien est resuspendu dans 5 mL de tampon d'agro-infiltration (MgCl<sub>2</sub> 10 mM ; acétosyringone 250 µM). La suspension est ensuite diluée pour obtenir 10 mL de tampon d'agro-infiltration à une DO<sub>600 nm</sub> = 0,8. Les suspensions d'agrobactéries sont incubées 2 h à 28 °C sous agitation. Si nécessaire, différentes souches d'agrobactéries à co-infiltrer sont mélangées. Les bactéries sont infiltrées sur deux feuilles d'une plante de quatre à cinq semaines à l'aide d'une seringue de 1 mL, après une légère incision sur la partie abaxiale des feuilles.

### 8.3. Transformation par floral dip

La méthode *floral dip* permet de transformer stablement les descendants d'une plante en immergeant ses fleurs dans une suspension d'*A. tumefaciens* (Clough et Bent, 1998). Une culture de 500 mL d'agrobactéries transformées avec le vecteur binaire estensemencée à partir d'une préculture, puis incubée à 28 °C sous agitation jusqu'à l'obtention d'une  $DO_{600\text{ nm}} = 0,8$ . Les bactéries sont ensuite centrifugées à 5 000 g pendant 20 min puis resuspendues dans 10 mL de  $MgCl_2$  10 mM. Elles sont alors diluées avec une solution de saccharose 5 % (p/v) et d'agent surfactant Silwet L-77® 0,02 % (v/v) jusqu'à l'obtention d'une  $DO_{600\text{ nm}} = 0,8$ . Les plantes d'*A. thaliana* sont transformées à cinq ou six semaines après germination, lorsque les premières siliques apparaissent. Toutes les siliques déjà formées sont supprimées manuellement pour ne garder que les fleurs. Les hampes florales sont immergées pendant 30 s dans la suspension d'agrobactéries, puis les plantes sont égouttées sur du papier absorbant et placées dans des mini-serres pour maintenir un haut taux d'humidité, dans l'obscurité pendant 24 h. Les conditions d'humidités sont réduites progressivement jusqu'au retour en serre des plantes.

### 8.4. Sélection des transformants obtenus par floral dip

Les graines issues des plantes transformées par *floral dip* sont récoltées en prenant garde de conserver les graines issues des siliques les plus précoces. La transformation de ces graines est ensuite testée sur un milieu MS gélosé contenant l'antibiotique adéquat à la résistance codée par le plasmide binaire, supplémenté de céfotaxime et de carbénicilline (**tableau 2**) pour éliminer les agrobactéries et autres micro-organismes contaminants. Un volume de graines équivalent à 150 µL est stérilisé avec 5 mL d'éthanol 70 % (v/v) pendant deux minutes, puis avec 5 mL de solution chlorée (chlore actif 4 % (v/v) ; Tween-20 4 % (v/v)) pendant 15 min sous agitation lente. Les graines sont lavées quatre fois dans 10 mL d'eau stérile, puis reprises dans 10 mL d'agarose 0,1 % (p/v) stérile fondu et refroidi. Les graines sont étalées uniformément sur le milieu gélosé de sélection grâce à la densité de l'agarose. Les boîtes sont incubées à 4 °C à l'obscurité pendant deux à trois jours pour permettre la stratification des graines.

La sélection des graines par résistance à l'hygromycine est réalisée selon (Harrison *et al.*, 2006). Après stratification, les graines sont exposées à la lumière pendant 6 h pour activer la germination, puis replacées à l'obscurité pendant 48 h. Enfin, les graines germées sont à nouveau placées à la lumière. Les plantes résistantes à l'hygromycine développent un hypocotyle plus allongé facilement observable. Ces plantes sont transférées sur un nouveau milieu gélosé pour leur permettre de croître suffisamment avant d'être transférées en terre.

## 8.5. Croisements

Les plantes d'*A. thaliana* présentent les deux types de gamètes et sont capables d'autofécondation. Il est néanmoins possible de croiser le pollen d'un individu avec le pistil d'un autre, permettant ainsi l'obtention de plantes double mutantes à partir de deux mutants indépendants (homozygotes ou hétérozygotes), pourvue que les mutations ne soient pas liées (très proches dans le même chromosome). Ces croisements permettent également d'étudier l'impact des mutations dans le génome mitochondrial, étant donné que la mitochondrie est maternellement transmise.

Les croisements sont réalisés sur deux à quatre fleurs d'une même inflorescence, sur le point de s'ouvrir. Toutes les autres fleurs et siliques de la hampe florale sont supprimées. Les fleurs restantes sont émasculées pour ne garder que le pistil intact, puis les plantes sont laissées en serre sur la nuit pour qu'elles terminent la maturation des pistils. Ces pistils sont ensuite mis en contact avec une fleur de la plante paternelle pour permettre le dépôt de pollen sur le pistil de la plante maternelle. Les siliques matures sont récupérées avant leur ouverture et conservées dans un tube ouvert pour qu'elles sèchent.

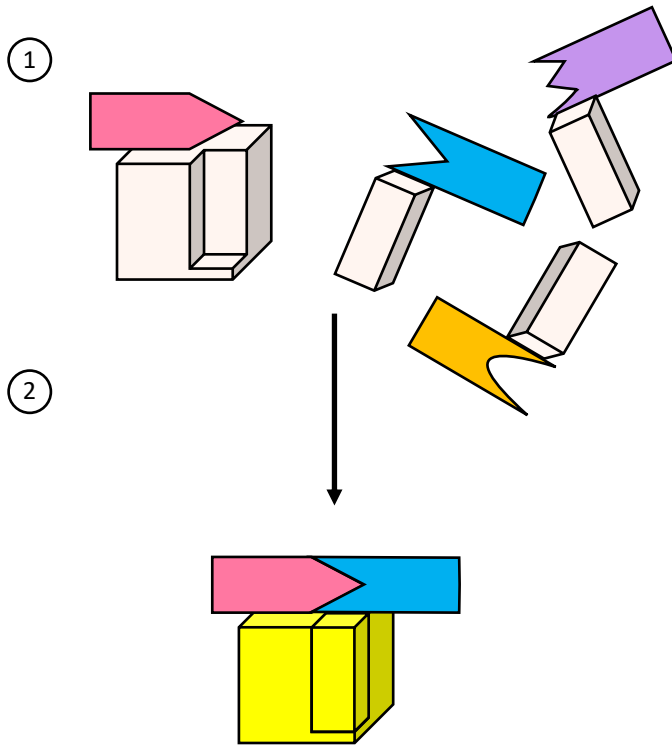
## 8.6. Culture de plantes en milieu *in vitro*

Les graines mises à germer sur un milieu *in vitro* nécessitent d'être stérilisées pour éviter le développement de micro-organismes. Un volume de graines équivalent à 30 µL est stérilisé par une incubation de 10 min sous agitation avec une solution de stérilisation (éthanol 70 % (v/v) ; Triton X-100 0,05 % (v/v)), suivie d'un premier lavage à l'éthanol 70 % (v/v), puis d'un second lavage à l'éthanol absolu. Les graines sont ensuite séchées à l'air libre sous une hotte à flux laminaire stérile.

Les graines sont déposées sur un milieu MS gélosé, supplémenté par des agents de sélection ou non, à l'aide d'un cure-dent stérile (dépôt ordonné) ou d'un morceau de papier *Whatman*<sup>®</sup> 3M stérile (dépôt non ordonné). Les boîtes sont incubées à 4 °C à l'obscurité pendant deux à trois jours pour permettre la stratification des graines.

## 8.7. Fusion d'un promoteur avec le gène rapporteur GUS

Pour visualiser le profil d'expression du promoteur d'un gène dans les différents organes et aux différents stades d'une plante, des plantes d'*A. thaliana* sont transformées avec le gène de la  $\beta$ -glucuronidase (GUS) d'*E. coli* sous contrôle du promoteur d'intérêt. Les tissus de ces plantes sont prélevés et fixés avec de l'acétone glaciale à 90 % (v/v) pendant 15 min sur glace, puis rincés avec du NaPO<sub>4</sub> 50 mM pH 7,5. Les tissus sont ensuite immergés dans le tampon de révélation (NaPO<sub>4</sub> 50 mM pH 7,5 ; EDTA 10 mM ; Triton X-100 0,1 % (v/v) ; K<sub>3</sub>Fe(CN)<sub>6</sub> 1 mM ; K<sub>4</sub>Fe(CN)<sub>6</sub> 1 mM ; 5-bromo-4-chloro-3-indolyl-beta-D-glucuronate (X-Gluc) 0,4 mg/mL (*Sigma* –



**Figure 55 : Principe de la complémentation bimoléculaire de la fluorescence.** La complémentation bimoléculaire de la fluorescence permet d'identifier les partenaires d'interaction d'une protéine d'intérêt. Elle repose sur la possibilité de séparer une protéine fluorescente (ici la YFP) en deux parties non fluorescentes. La protéine d'intérêt (en rose) est fusionnée à la séquence N-terminale de la YFP tandis que ses partenaires hypothétiques sont fusionnés à la séquence C-terminale. En cas d'interaction, la structure de la YFP pourra être reconstituée, rétablissant la fluorescence de la protéine qui pourra être observée par microscopie confocale.

Aldrich®)) à 37 °C sur la nuit à l'obscurité. Enfin, pour une meilleure visualisation de la coloration, la chlorophylle de l'échantillon est supprimée par des bains successifs de concentration croissante en éthanol (15, 30, 50, 70, 85 puis 95 % (v/v)). Les tissus colorés sont conservés dans l'éthanol 50 % (v/v) à 4 °C.

#### **8.8. Test de sensibilité aux UV-C**

Les tests de sensibilité aux UV-C sont réalisés sur des plantules d'*A. thaliana* cultivées *in vitro*. Les graines sont stérilisées, semées de manière non ordonnée, et stratifiées puis placées en condition de lumière pendant six jours. Les plantules germées sont alors repiquées sur une nouvelle boîte de milieu MS de manière ordonnée puis replacées en condition de culture pendant un jour. Les plantules sont irradiées avec un *UV Stratalinker 2400™* (Stratagene®, San Diego, Californie, États-Unis) pour 0, 750, 1 500 ou 3 000 J/m<sup>2</sup> puis placées à l'obscurité durant 6 h pour empêcher l'action des photolyases. Les plantules sont ensuite replacées en conditions de culture standard. Le traitement est répété deux fois, toutes les 48 h. Les plantules sont ensuite photographiées et pesées.

#### **8.9. Test de sensibilité au NaCl, à la roténone et à la lincomycine**

Les tests de sensibilité sont réalisés sur des plantules d'*A. thaliana* cultivées *in vitro*. Les graines sont stérilisées et semées de manière ordonnée sur des boîtes de milieu MS contenant 0 à 150 mM de NaCl ou 0 à 40 mM de roténone ou 0,01 à 0,16 mM de lincomycine. Les graines sont stratifiées pendant 48 h à 4 °C puis placées en condition standard de lumière pour germination. Les plantules sont observées dix jours après germination, et pesées 16 jours après germination.

#### **8.10. Test de sensibilité au H<sub>2</sub>O<sub>2</sub>**

Les tests de sensibilité sont réalisés sur des plantules d'*A. thaliana* cultivées *in vitro*. Les graines sont stérilisées et semées sur des boîtes de milieu MS. Après stratification et germination, les plantules sont cultivées neuf jours avant d'être transférées dans un milieu liquide ou gélosé contenant de 0 à 5 µM de peroxyde d'hydrogène actif. Les plantules sont pesées quatre jours après le début du stress.

#### **8.11. Complémentation bimoléculaire de la fluorescence (BiFC)**

La BiFC est une technique permettant d'étudier l'interaction entre deux protéines. Les séquences des protéines d'intérêt sont fusionnées à la partie N-terminale ou C-terminale de la *yellow fluorescence protein* (YFP). Si les deux protéines interagissent, la séquence protéique de la YFP est reconstituée permettant la fluorescence de la protéine, observable par microscopie confocale (**figure 55**).

Après co-agro-infiltration des vecteurs pGW-YN et pGW-YC exprimant les protéines de fusion d'intérêt, les plantes sont laissées en serre pendant trois jours pour permettre l'expression desdites protéines. Les parties de feuilles agro-infiltrées sont prélevées, placées entre lame et lamelle avec de l'eau puis observées au microscope LSM700 *Laser scanning* (Zeiss®, Oberkochen, Allemagne).

### 8.12. Observation du développement des embryons de plantes fertilisées

Des fleurs émasculées d'*A. thaliana* sauvage et de *radA* sont fixées avec tampon FAA (10:7:2:1 éthanol:eau:acide acétique:formaldéhyde (37 %)) dans la glace pendant 30 min, hydratées avec Na-phosphate 50 mM pH 7,2, disséquées et montées sur lames de microscope dans une solution de 8:2:1 hydrate de chloral:eau:glycérol. Les ovules matures non-fertilisés ont été observés sous microscope photonique *Axiolmager Z1* (Zeiss®) avec optique DIC. D'autres fleurs émasculées sont croisées avec du pollen sauvage exogène. Après trois jours les embryons fertilisés sont également observés sous microscope photonique.

### 8.13. Observation de la forme des cellules et de l'état de pollinisation des fleurs

Des feuilles et des fleurs d'*A. thaliana* sauvages ou *radA* sont prélevées et observées sur microscope électronique à balayage de paillasse *Tabletop SEM, Hitachi TM3030 Plus* (Hitachi High-Tech®, Japon).

### 8.14. Mesure de la chlorophylle

La quantité de chlorophylle est calculée selon la méthode d'Arnon (Arnon, 1949). Des sections de 0,5 cm<sup>2</sup> de feuilles matures d'*A. thaliana* cultivées en condition standard sont broyées dans l'azote avec *TissueLyser II* (Qiagen®) à l'aide d'une bille métallique. Le broyat est repris dans 80 % d'acétone et les débris cellulaire sont sédimentés par centrifugation à 15 000 g pendant 5 min à 4 °C. Les absorbances à 645 nm et 663 nm sont mesurées et la chlorophylle quantifiée selon les formules suivantes :

$$\text{Chlorophylle a (g/L)} = 0.0127 A_{663 \text{ nm}} - 0.00269 A_{645 \text{ nm}}$$

$$\text{Chlorophylle b (g/L)} = 0.0229 A_{645 \text{ nm}} - 0.00468 A_{663 \text{ nm}}$$

$$\text{Chlorophylle Totale (g/L)} = 0.0202 A_{645 \text{ nm}} + 0.00802 A_{663 \text{ nm}}$$

## 9. Méthodes relatives aux organelles

### 9.1. Extraction de mitochondrie à partir de plantules étiolées

L'extraction est réalisée sur des plantules d'*A. thaliana* de neuf jours cultivées à l'obscurité pour empêcher le développement des chloroplastes, et donc limiter la contamination par le matériel chloroplastique, très présent dans la cellule. Toutes les étapes

de la purification sont réalisées au froid pour limiter la dégradation des mitochondries. Les plantules sont récoltées et pesées, puis broyées dans un mortier avec 20 mL de tampon de broyage pour 10 g de plantules (saccharose 0,3 M ; tetrasodium pyrophosphate 5 mM ; EDTA 2 mM ;  $\text{KH}_2\text{PO}_4$  100 mM ; polyvinylpyrrolidone (PVP-40) 1 % (p/v) ; pH 7,5) supplémentés de 1 % BSA (p/v), de 20 mM d'acide ascorbique et de 5 mM de cystéine. Le broyat est filtré sur deux couches de gaze et une couche de *Miracloth* (Merck®) et le matériel récupéré sur les filtres est à nouveau broyer afin de récupérer un maximum de mitochondrie. Le filtrat est ensuite centrifugé une première fois à 2 450 g pendant 5 min à 4 °C pour éliminer les gros débris cellulaires, puis centrifugé une seconde fois dix minutes à 17 400 g à 4 °C pour sédimenter les mitochondries. Le culot est repris dans du tampon de lavage (saccharose 0,3 M ; MOPS 10 mM ; EGTA 1 mM ; pH 7,4) à l'aide d'un pinceau puis à nouveau centrifugé deux fois, comme précédemment, avant d'être repris dans du tampon de lavage. Les mitochondries sont purifiées sur gradient de discontinu de Percoll (1 mL 50 %, 1,5 mL 25 % et 0,5 mL 18 %) dans du tampon de lavage supplémentés de 1 % BSA (p/v). Les gradients sont centrifugés à 25 000 g pendant 25 min à 4 °C sans frein à la fin de la centrifugation. Les mitochondries intactes sont récupérées à l'interface des solutions de Percoll de 25 % et 50 % et reprises dans du tampon de lavage. Le Percoll est éliminé par une série de centrifugation à 18 000 g pendant 15 min en reprenant le culot dans du tampon de lavage. Les mitochondries purifiées sont conservées dans un faible volume final.

## 9.2. Extraction de chloroplaste à partir de plantes et plantules

L'extraction des chloroplastes est réalisée sur des plantules d'*A. thaliana* de 18 jours cultivées sur milieu *in vitro*. Comme pour les mitochondries, toutes les étapes de la purification sont réalisées en chambre froide pour limiter la dégradation des chloroplastes, plus fragiles encore. Les plantules sont récoltées, pesées puis broyées dans un mixeur de laboratoire à faible vitesse pendant deux à trois secondes avec 5 mL/g de tampon d'extraction (Sorbitol 1,65 M ; HEPES 250 mM pH 7,3 ; EDTA 10 mM) supplémenté de 1 mM DTT et de 1 % BSA (p/v) juste avant utilisation. Le broyat est filtré sur deux couches de *Miracloth* (Merck®) puis le filtrat est centrifugé à 2 000 g pendant 5 min à 4 °C sur un rotor à inclinaison variable JS-13.1 (Beckman Coulter®, Brea, Californie, États-Unis) avec une faible force de freinage en fin de centrifugation. Après élimination du surnageant, le culot est repris à l'aide d'un pinceau dans un faible volume de tampon d'extraction puis déposé sur un gradient discontinu de Percoll (5 mL 80 % et 25 mL 40 %) dans du tampon d'extraction supplémenté de BSA 1 % (p/v) et PEG-6 000 3 % (p/v). Le gradient est centrifugé à 7 000 g pendant 20 min à 4 °C sans frein en fin de centrifugation. Les chloroplastes intacts sont récupérés à l'interface des deux solutions de Percoll et repris dans cinq volumes de tampon d'extraction. Le Percoll est éliminé par une centrifugation de 2 000 g pendant 2 min à 4 °C et les chloroplastes sont repris dans un faible volume de tampon d'extraction sans BSA ni DTT. Eventuellement, un deuxième lavage peut être nécessaire.

# **BIBLIOGRAPHIE**



- Abdelnoor RV, Christensen AC, Mohammed S, Munoz-Castillo B, Moriyama H, et Mackenzie SA** (2006). Mitochondrial Genome Dynamics in Plants and Animals: Convergent Gene Fusions of a MutS Homologue. *J Mol Evol* **63**, 165-173.
- Abraham RT** (2001). Cell cycle checkpoint signaling through the ATM and ATR kinases. *Genes Dev* **15**, 2177-2196.
- Adams KL, et Palmer JD** (2003). Evolution of mitochondrial gene content: gene loss and transfer to the nucleus. *Mol Phylogenet Evol* **29**, 380-395.
- Adelman CA, et Boulton SJ** (2010). Metabolism of postsynaptic recombination intermediates. *FEBS Lett* **584**, 3709-3716.
- Albertsson P** (2001). A quantitative model of the domain structure of the photosynthetic membrane. *Trends Plant Sci* **6**, 349-358.
- Albiniak AM, Baglieri J, et Robinson C** (2012). Targeting of luminal proteins across the thylakoid membrane. *J Exp Bot* **63**, 1689-1698.
- Allen JF** (1993). Redox control of transcription: sensors, response regulators, activators and repressors. *FEBS Lett* **332**, 203-207.
- Allen JF, et Raven JA** (1996). Free-radical-induced mutation vs redox regulation: costs and benefits of genes in organelles. *J Mol Evol* **42**, 482-492.
- Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, Gadrinab C, Heller C, Jeske A, Koesema E, Meyers CC, Parker H, Prednis L, Ansari Y, Choy N, Deen H, Geralt M, Hazari N, Hom E, Karnes M, Mulholland C, Ndubaku R, Schmidt I, Guzman P, Aguilar-Henonin L, Schmid M, Weigel D, Carter DE, Marchand T, Risseuw E, Brogden D, Zeko A, Crosby WL, Berry CC, et Ecker JR** (2003). Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* **301**, 653-657.
- Aravind L, Makarova KS, et Koonin EV** (2000). SURVEY AND SUMMARY: holliday junction resolvases and related nucleases: identification of new families, phyletic distribution and evolutionary trajectories. *Nucleic Acids Res* **28**, 3417-3432.
- Arimura S-i, et Tsutsumi N** (2002). A dynamin-like protein (ADL2b), rather than FtsZ, is involved in *Arabidopsis* mitochondrial division. *Proc Natl Acad Sci U S A* **99**, 5727-5731.
- Arnon DI** (1949). Copper Enzymes in Isolated Chloroplasts. Polyphenoloxidase in *Beta Vulgaris*. *Plant Physiol* **24**, 1-15.
- Arrieta-Montiel MP, Shedje V, Davila J, Christensen AC, et Mackenzie SA** (2009). Diversity of the *Arabidopsis* mitochondrial genome occurs via nuclear-controlled recombination activity. *Genetics* **183**, 1261-1268.
- Backert S, Lynn Nielsen B, et Börner T** (1997). The mystery of the rings: structure and replication of mitochondrial genomes from higher plants. *Trends Plant Sci* **2**, 477-483.
- Barkan A, et Small I** (2014). Pentatricopeptide repeat proteins in plants. *Annu Rev Plant Biol* **65**, 415-442.
- Barkan A, Rojas M, Fujii S, Yap A, Chong YS, Bond CS, et Small I** (2012). A Combinatorial Amino Acid Code for RNA Recognition by Pentatricopeptide Repeat Proteins. *PLoS Genet* **8**, e1002910.
- Barrey E, Saint-Auret G, Bonnamy B, Damas D, Boyer O, et Gidrol X** (2011). Pre-microRNA and mature microRNA in human mitochondria. *PLoS ONE* **6**, e20220.
- Barshad G, Marom S, Cohen T, et Mishmar D** (2018). Mitochondrial DNA Transcription and Its Regulation: An Evolutionary Perspective. *Trends Genet* **34**, 682-692.
- Baulcombe D** (2004). RNA silencing in plants. *Nature* **431**, 356-363.
- Beam CE, Saveson CJ, et Lovett ST** (2002). Role for radA/sms in recombination intermediate processing in *Escherichia coli*. *J Bacteriol* **184**, 6836-6844.
- Belitsky BR, et Sonenshein AL** (2013). Genome-wide identification of *Bacillus subtilis* CodY-binding sites at single-nucleotide resolution. *Proc Natl Acad Sci U S A* **110**, 7026-7031.
- Bendich AJ** (2004). Circular chloroplast chromosomes: the grand illusion. *Plant Cell* **16**, 1661-1666.
- Benson FE, Baumann P, et West SC** (1998). Synergistic actions of Rad51 and Rad52 in recombination and DNA repair. *Nature* **391**, 401-404.
- Bernal-Bernal D, Gallego-Garcia A, Garcia-Martinez G, Garcia-Heras F, Jimenez MA, Padmanabhan S, et Elias-Arnanz M** (2015). Structure-function dissection of *Myxococcus xanthus* CarD N-terminal domain, a defining member of the CarD\_CdnL\_TRCF family of RNA polymerase interacting proteins. *PLoS ONE* **10**, e0121322.

- Birky CW, Jr.** (2001). The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms, and models. *Annu Rev Genet* **35**, 125-148.
- Bobik K, Fernandez JC, Hardin SR, Ernest B, Ganusova EE, Staton ME, et Burch-Smith TM** (2019). The essential chloroplast ribosomal protein uL15c interacts with the chloroplast RNA helicase ISE2 and affects intercellular trafficking through plasmodesmata. *New Phytol* **221**, 850-865.
- Boesch P, Ibrahim N, Paulus F, Cosset A, Tarasenko V, et Dietrich A** (2009). Plant mitochondria possess a short-patch base excision DNA repair pathway. *Nucleic Acids Res* **37**, 5690-5700.
- Boesch P, Weber-Lotfi F, Ibrahim N, Tarasenko V, Cosset A, Paulus F, Lightowlers RN, et Dietrich A** (2011). DNA repair in organelles: Pathways, organization, regulation, relevance in disease and aging. *Biochim Biophys Acta* **1813**, 186-200.
- Bohne AV, Teubner M, Liere K, Weihe A, et Borner T** (2016). In vitro promoter recognition by the catalytic subunit of plant phage-type RNA polymerases. *Plant Mol Biol* **92**, 357-369.
- Bonen L, et Gray MW** (1980). Organization and expression of the mitochondrial genome of plants. 1- the genes for wheat mitochondrial ribosomal and transfer RNA: evidence for an unusual arrangement. *Nucleic Acids Res* **8**, 319-335.
- Borner T, Aleynikova AY, Zubo YO, et Kusnetsov VV** (2015). Chloroplast RNA polymerases: Role in chloroplast biogenesis. *Biochim Biophys Acta* **1847**, 761-769.
- Boubakri H, de Septenville AL, Viguera E, et Michel B** (2010). The helicases DinG, Rep and UvrD cooperate to promote replication across transcription units in vivo. *EMBO J* **29**, 145-157.
- Bouchier C, Ma L, Créno S, Dujon B, et Fairhead C** (2009). Complete mitochondrial genome sequences of three *Nakaseomyces* species reveal invasion by palindromic GC clusters and considerable size expansion. *FEMS Yeast Res* **9**, 1283-1292.
- Boussardon C, Avon A, Kindgren P, Bond CS, Challenor M, Lurin C, et Small I** (2014). The cytidine deaminase signature HxE(x)n CxxC of DYW1 binds zinc and is necessary for RNA editing of ndhD-1. *New Phytol* **203**, 1090-1095.
- Bradford MM** (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248-254.
- Brenneman MA, Wagener BM, Miller CA, Allen C, et Nickoloff JA** (2002). XRCC3 controls the fidelity of homologous recombination: roles for XRCC3 in late stages of recombination. *Mol Cell* **10**, 387-395.
- Brettel K, et Byrdin M** (2010). Reaction mechanisms of DNA photolyase. *Curr Opin Struct Biol* **20**, 693-701.
- Brouwer B, Gardestrom P, et Keech O** (2014). In response to partial plant shading, the lack of phytochrome A does not directly induce leaf senescence but alters the fine-tuning of chlorophyll biosynthesis. *J Exp Bot* **65**, 4037-4049.
- Brown GG, Colas des Francs-Small C, et Ostersetzer-Biran O** (2014). Group II intron splicing factors in plant mitochondria. *Front Plant Sci* **5**, 35.
- Brown WM, George M, et Wilson AC** (1979). Rapid evolution of animal mitochondrial DNA. *Proc Natl Acad Sci U S A*, 5.
- Bruce BD** (2000). Chloroplast transit peptides: structure, function and evolution. *Trends Cell Biol* **10**, 440-447.
- Budar F, et Pelletier G** (2001). Male sterility in plants: occurrence, determinism, significance and use. *C R Acad Sci III* **324**, 543-550.
- Cappadocia L, Marechal A, Parent JS, Lepage E, Sygusch J, et Brisson N** (2010). Crystal structures of DNA-Whirly complexes and their role in Arabidopsis organelle genome repair. *Plant Cell* **22**, 1849-1867.
- Cárdenas PP, Carrasco B, Defeu Soufo C, César CE, Herr K, Kaufenstein M, Graumann PL, et Alonso JC** (2012). RecX Facilitates Homologous Recombination by Modulating RecA Activities. *PLoS Genet* **8**, e1003126.
- Carlotto N, Wirth S, Furman N, Ferreyra Solari N, Ariel F, Crespi M, et Kobayashi K** (2016). The chloroplastic DEVH-box RNA helicase INCREASED SIZE EXCLUSION LIMIT 2 involved in plasmodesmata regulation is required for group II intron splicing. *Plant Cell Environ* **39**, 165-173.
- Carrillo C, et Bonen L** (1997). RNA editing status of nad7 intron domains in wheat mitochondria. *Nucleic Acids Res* **25**, 403-409.

- Cerutti H, Johnson AM, Boynton JE, et Gillham NW (1995). Inhibition of chloroplast DNA recombination and repair by dominant negative mutants of Escherichia coli RecA. *Mol Cell Biol* **15**, 3003-3011.
- Chambers AL, Smith AJ, et Savery NJ (2003). A DNA translocation motif in the bacterial transcription--repair coupling factor, Mfd. *Nucleic Acids Res* **31**, 6409-6418.
- Chang AC, et Cohen SN (1978). Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J Bacteriol* **134**, 1141-1156.
- Chang HHY, Pannunzio NR, Adachi N, et Lieber MR (2017a). Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat Rev Mol Cell Biol* **18**, 495-506.
- Chang SH, Lee S, Um TY, Kim J-K, Do Choi Y, et Jang G (2017b). pTAC10, a Key Subunit of Plastid-Encoded RNA Polymerase, Promotes Chloroplast Development. *Plant Physiol* **174**, 435-449.
- Chen JJ, Jiang CZ, et Britt AB (1996). Little or no repair of cyclobutyl pyrimidine dimers is observed in the organellar genomes of the young Arabidopsis seedling. *Plant Physiol* **111**, 19-25.
- Chen SH, Plank JL, Willcox S, Griffith JD, et Hsieh TS (2014). Top3alpha is required during the convergent migration step of double Holliday junction dissolution. *PLoS ONE* **9**, e83582.
- Chi W, He B, Manavski N, Mao J, Ji D, Lu C, Rochaix JD, Meurer J, et Zhang L (2014). RHON1 Mediates a Rho-Like Activity for Transcription Termination in Plastids of *Arabidopsis thaliana*. *Plant Cell* **26**, 4918-4932.
- Cho Y, Qiu YL, Kuhlman P, et Palmer JD (1998). Explosive invasion of plant mitochondria by a group I intron. *Proc Natl Acad Sci U S A* **95**, 14244-14249.
- Christensen AC (2013). Plant mitochondrial genome evolution can be explained by DNA repair mechanisms. *Genome Biol Evol* **5**, 1079-1086.
- Christensen AC (2014). Genes and junk in plant mitochondria-repair mechanisms and selection. *Genome Biol Evol* **6**, 1448-1453.
- Christensen AC (2018). Mitochondrial DNA Repair and Genome Evolution. In Annual Plant Reviews online, pp. 11-32.
- Chrzanowska-Lightowlers Z, Rorbach J, et Minczuk M (2017). Human mitochondrial ribosomes can switch structural tRNAs - but when and why? *RNA Biol* **14**, 1668-1671.
- Chun J, Buechelmaier ES, et Powell SN (2013). Rad51 paralog complexes BCDX2 and CX3 act at different stages in the BRCA1-BRCA2-dependent homologous recombination pathway. *Mol Cell Biol* **33**, 387-395.
- Ciesielski GL, Oliveira MT, et Kaguni LS (2016). Animal Mitochondrial DNA Replication. *Enzymes* **39**, 255-292.
- Clough SJ, et Bent AF (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J* **16**, 735-743.
- Cockram CA, Filatenkova M, Danos V, El Karoui M, et Leach DR (2015). Quantitative genomic analysis of RecA protein binding during DNA double-strand break repair reveals RecBCD action in vivo. *Proc Natl Acad Sci U S A* **112**, E4735-4742.
- Cohen SE, Lewis CA, Mooney RA, Kohanski MA, Collins JJ, Landick R, et Walker GC (2010). Roles for the transcription elongation factor NusA in both DNA repair and damage tolerance pathways in Escherichia coli. *Proc Natl Acad Sci U S A* **107**, 15517-15522.
- Comte J, Maïsterrena B, et Gautheron DC (1976). Lipid composition and protein profiles of outer and inner membranes from pig heart mitochondria. Comparison with microsomes. *Biochim Biophys Acta, Biomembr* **419**, 271-284.
- Cooper DL, et Lovett ST (2016). Recombinational branch migration by the RadA/Sms paralog of RecA in Escherichia coli. *Elife* **5**.
- Cooper DL, Boyle DC, et Lovett ST (2015). Genetic analysis of Escherichia coli RadA: functional motifs and genetic interactions. *Mol Microbiol* **95**, 769-779.
- Covello PS, et Gray MW (1989). RNA editing in plant mitochondria. *Nature* **341**, 662-666.
- Curtis MD, et Grossniklaus U (2003). A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol* **133**, 462-469.
- D'Errico M, Pascucci B, Iorio E, Van Houten B, et Dogliotti E (2013). The role of CSA and CSB protein in the oxidative stress response. *Mech Ageing Dev* **134**, 261-269.

- Dahal S, Dubey S, et Raghavan SC (2018). Homologous recombination-mediated repair of DNA double-strand breaks operates in mammalian mitochondria. *Cell Mol Life Sci* **75**, 1641-1655.
- Daley DO, Clifton R, et Whelan J (2002). Intracellular gene transfer: reduced hydrophobicity facilitates gene transfer for subunit 2 of cytochrome c oxidase. *Proc Natl Acad Sci U S A* **99**, 10510-10515.
- Davidov Y, et Jurkevitch E (2009). Predation between prokaryotes and the origin of eukaryotes. *Bioessays* **31**, 748-757.
- Davila JI, Arrieta-Montiel MP, Wamboldt Y, Cao J, Hagemann J, Shedge V, Xu YZ, Weigel D, et Mackenzie SA (2011). Double-strand break repair processes drive evolution of the mitochondrial genome in Arabidopsis. *BMC Biol* **9**, 64.
- Davis PA, et Hangarter RP (2012). Chloroplast movement provides photoprotection to plants by redistributing PSII damage within leaves. *Photosynth Res* **112**, 153-161.
- de Longevialle AF, Small ID, et Lurin C (2010). Nuclearly encoded splicing factors implicated in RNA splicing in higher plant organelles. *Mol Plant* **3**, 691-705.
- de Paula WBM, Agip A-NA, Missirlis F, Ashworth R, Vizcay-Barrena G, Lucas CH, et Allen JF (2013). Female and Male Gamete Mitochondria Are Distinct and Complementary in Transcription, Structure, and Genome Function. *Genome Biol Evol* **5**, 1969-1977.
- De Santis-Maciossek G, Kofer W, Bock A, Schoch S, Maier RM, Wanner G, Rudiger W, Koop H-U, et Herrmann RG (1999). Targeted disruption of the plastid RNA polymerase genes rpoA, B and C1: molecular biology, biochemistry and ultrastructure. *Plant J* **18**, 477-489.
- Deaconescu AM, et Suhanovsky MM (2017). From Mfd to TRCF and Back Again-A Perspective on Bacterial Transcription-coupled Nucleotide Excision Repair. *Photochem Photobiol* **93**, 268-279.
- Deaconescu AM, Sevostyanova A, Artsimovitch I, et Grigorieff N (2012). Nucleotide excision repair (NER) machinery recruitment by the transcription-repair coupling factor involves unmasking of a conserved intramolecular interface. *Proc Natl Acad Sci U S A* **109**, 3353-3358.
- Deaconescu AM, Chambers AL, Smith AJ, Nickels BE, Hochschild A, Savery NJ, et Darst SA (2006). Structural basis for bacterial transcription-coupled DNA repair. *Cell* **124**, 507-520.
- Del Campo EM, Sabater B, et 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.
- Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, Claverie JM, et Gascuel O (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* **36**, W465-469.
- Dietrich A, Wallet C, Iqbal RK, Gualberto JM, et Lotfi F (2015). Organellar non-coding RNAs: Emerging regulation mechanisms. *Biochimie* **117**, 48-62.
- Diver WP, Sargentini NJ, et Smith KC (1982). A mutation (radA100) in Escherichia coli that selectively sensitizes cells grown in rich medium to x- or u.v.-radiation, or methyl methanesulphonate. *Int J Radiat Biol Relat Stud Phys Chem Med* **42**, 339-346.
- Dodson ML, et Lloyd RS (2002). Mechanistic comparisons among base excision repair glycosylases. *Free Radic Biol Med* **32**, 678-682.
- Drees JC, Lusetti SL, et Cox MM (2004). Inhibition of RecA protein by the Escherichia coli RecX protein: modulation by the RecA C terminus and filament functional state. *J Biol Chem* **279**, 52991-52997.
- Dubois M, Selden K, Bediee A, Rolland G, Baumberger N, Noir S, Bach L, Lamy G, Granier C, et Genschik P (2018). SIAMESE-RELATED1 Is Regulated Posttranslationally and Participates in Repression of Leaf Growth under Moderate Drought. *Plant Physiol* **176**, 2834-2850.
- Dutilleul C, Garmier M, Noctor G, Mathieu C, Chetrit P, Foyer CH, et de Paepe R (2003). Leaf mitochondria modulate whole cell redox homeostasis, set antioxidant capacity, and determine stress resistance through altered signaling and diurnal regulation. *Plant Cell* **15**, 1212-1226.
- Edmondson AC, Song D, Alvarez LA, Wall MK, Almond D, McClellan DA, Maxwell A, et Nielsen BL (2005). Characterization of a mitochondrially targeted single-stranded DNA-binding protein in Arabidopsis thaliana. *Mol Genet Genomics* **273**, 115-122.

- Eggler AL, Lusetti SL, et Cox MM** (2003). The C terminus of the Escherichia coli RecA protein modulates the DNA binding competition with single-stranded DNA-binding protein. *J Biol Chem* **278**, 16389-16396.
- Elliott B, et Jasin M** (2001). Repair of Double-Strand Breaks by Homologous Recombination in Mismatch Repair-Defective Mammalian Cells. *Mol Cell Biol* **21**, 2671-2682.
- Emanuel C, Weihe A, Graner A, Hess WR, et Borner T** (2004). Chloroplast development affects expression of phage-type RNA polymerases in barley leaves. *Plant J* **38**, 460-472.
- Emanuel C, von Groll U, Muller M, Borner T, et Weihe A** (2006). Development- and tissue-specific expression of the RpoT gene family of Arabidopsis encoding mitochondrial and plastid RNA polymerases. *Planta* **223**, 998-1009.
- Eme L, Spang A, Lombard J, Stairs CW, et Ettema TJG** (2017). Archaea and the origin of eukaryotes. *Nat Rev Microbiol* **15**, 711-723.
- Entelis NS, Kolesnikova OA, Dogan S, Martin RP, et Tarassov IA** (2001). 5 S rRNA and tRNA import into human mitochondria. Comparison of in vitro requirements. *J Biol Chem* **276**, 45642-45653.
- Epshtein V, Kamarthapu V, McGary K, Svetlov V, Ueberheide B, Proshkin S, Mironov A, et Nudler E** (2014). UvrD facilitates DNA repair by pulling RNA polymerase backwards. *Nature* **505**, 372-377.
- Falkenberg M, Gaspari M, Rantanen A, Trifunovic A, Larsson N-G, et Gustafsson CM** (2002). Mitochondrial transcription factors B1 and B2 activate transcription of human mtDNA. *Nat Genet* **31**, 289-294.
- Fan J, Leroux-Coyau M, Savery NJ, et Strick TR** (2016). Reconstruction of bacterial transcription-coupled repair at single-molecule resolution. *Nature* **536**, 234-237.
- Ferrando B, Furlanetto A, Gredilla R, Havelund JF, Hebelstrup KH, Moller IM, et Stevnsner T** (2018). DNA repair in plant mitochondria - a complete base excision repair pathway in potato tuber mitochondria. *Physiol Plant*.
- Finster S, Eggert E, Zoschke R, Weihe A, et Schmitz-Linneweber C** (2013). Light-dependent, plastome-wide association of the plastid-encoded RNA polymerase with chloroplast DNA. *Plant J* **76**, 849-860.
- Fisher RP, et Clayton DA** (1988). Purification and characterization of human mitochondrial transcription factor 1. *Mol Cell Biol* **8**, 3496-3509.
- Flores-Perez U, et Jarvis P** (2013). Molecular chaperone involvement in chloroplast protein import. *Biochim Biophys Acta* **1833**, 332-340.
- Forget AL, et Kowalczykowski SC** (2012). Single-molecule imaging of DNA pairing by RecA reveals a three-dimensional homology search. *Nature* **482**, 423-427.
- Foyer CH, et Noctor G** (2005). Redox homeostasis and antioxidant signaling: a metabolic interface between stress perception and physiological responses. *Plant Cell* **17**, 1866-1875.
- Freel KC, Friedrich A, et Schacherer J** (2015). Mitochondrial genome evolution in yeasts: an all-encompassing view. *FEMS Yeast Res* **15**.
- Fukui H, et Moraes CT** (2009). Mechanisms of formation and accumulation of mitochondrial DNA deletions in aging neurons. *Hum Mol Genet* **18**, 1028-1036.
- Gabaldon T** (2018). Relative timing of mitochondrial endosymbiosis and the "pre-mitochondrial symbioses" hypothesis. *IUBMB Life* **70**, 1188-1196.
- Gagliardi D, Stepien PP, Temperley RJ, Lightowlers RN, et Chrzanowska-Lightowlers ZM** (2004). Messenger RNA stability in mitochondria: different means to an end. *Trends Genet* **20**, 260-267.
- Galletto R, Amitani I, Baskin RJ, et Kowalczykowski SC** (2006). Direct observation of individual RecA filaments assembling on single DNA molecules. *Nature* **443**, 875-878.
- Gandara C, de Lucena DKC, Torres R, Serrano E, Altenburger S, Graumann PL, et Alonso JC** (2017). Activity and in vivo dynamics of Bacillus subtilis DisA are affected by RadA/Sms and by Holliday junction-processing proteins. *DNA Repair (Amst)* **55**, 17-30.
- Ganesan AK, et Hanawalt PC** (2010). Transcription-coupled nucleotide excision repair of a gene transcribed by bacteriophage T7 RNA polymerase in Escherichia coli. *DNA Repair (Amst)* **9**, 958-963.
- Garcia-Medel PL, Baruch-Torres N, Peralta-Castro A, Trasvina-Arenas CH, Torres-Larios A, et Briebe LG** (2019). Plant organellar DNA polymerases repair double-stranded breaks by microhomology-mediated end-joining. *Nucleic Acids Res.*

- Giegé P, et Brennicke A (2001). From gene to protein in higher plant mitochondria. *C R Acad Sci II* **324**, 209-217.
- Gilkerson RW, Selker JML, et Capaldi RA (2003). The cristal membrane of mitochondria is the principal site of oxidative phosphorylation. *FEBS Lett* **546**, 355-358.
- Gobert A, Gutmann B, Taschner A, Gößringer M, Holzmann J, Hartmann RK, Rossmannith W, et Giegé P (2010). A single Arabidopsis organellar protein has RNase P activity. *Nat Struct Mol Biol* **17**, 740-744.
- Golczyk H, Greiner S, Wanner G, Weihe A, Bock R, Borner T, et Herrmann RG (2014). Chloroplast DNA in mature and senescing leaves: a reappraisal. *Plant Cell* **26**, 847-854.
- Goldberg AV, Molik S, Tsaousis AD, Neumann K, Kühnke G, Delbac F, Vivares CP, Hirt RP, Lill R, et Embley TM (2008). Localization and functionality of microsporidian iron-sulphur cluster assembly proteins. *Nature* **452**, 624-628.
- Gould SB, Waller RF, et McFadden GI (2008). Plastid evolution. *Annu Rev Plant Biol* **59**, 491-517.
- Gray JC, Sullivan JA, Wang J-H, Jerome CA, et MacLean D (2003). Coordination of plastid and nuclear gene expression. *Philos Trans R Soc Lond B Biol Sci* **358**, 135-145.
- Gray MW (1999). Evolution of organellar genomes. *Curr Opin Genet Dev* **9**, 678-687.
- Grübler B, Merendino L, Twardziok SO, Mininno M, Allorent G, Chevalier F, Liebers M, Blanvillain R, Mayer KFX, Lerbs-Mache S, Ravanel S, et Pfannschmidt T (2017). Light and Plastid Signals Regulate Different Sets of Genes in the Albino Mutant Pap7-1. *Plant Physiol* **175**, 1203-1219.
- Gu L, Xu T, Lee K, Lee KH, et Kang H (2014). A chloroplast-localized DEAD-box RNA helicase AtRH3 is essential for intron splicing and plays an important role in the growth and stress response in Arabidopsis thaliana. *Plant Physiol Biochem* **82**, 309-318.
- Gualberto JM, et Kühn K (2014). DNA-binding proteins in plant mitochondria: implications for transcription. *Mitochondrion* **19 Pt B**, 323-328.
- Gualberto JM, et Newton KJ (2017). Plant Mitochondrial Genomes: Dynamics and Mechanisms of Mutation. *Annu Rev Plant Biol* **68**, 225-252.
- Gualberto JM, Wintz H, Weil JH, et Grienenberger JM (1988). The genes coding for subunit 3 of NADH dehydrogenase and for ribosomal protein S12 are present in the wheat and maize mitochondrial genomes and are co-transcribed. *Mol Gen Genet* **215**, 118-127.
- Gualberto JM, Lamattina L, Bonnard G, Weil JH, et Grienenberger JM (1989). RNA editing in wheat mitochondria results in the conservation of protein sequences. *Nature* **341**, 660-662.
- Gutman BL, et Niyogi KK (2009). Evidence for base excision repair of oxidative DNA damage in chloroplasts of Arabidopsis thaliana. *J Biol Chem* **284**, 17006-17012.
- Gutmann B, Gobert A, et Giege P (2012). PRORP proteins support RNase P activity in both organelles and the nucleus in Arabidopsis. *Genes Dev* **26**, 1022-1027.
- Hada M, Hino K, Buchholz G, Goss J, Wellmann E, et Shin M (2000). Assay of DNA Photolyase Activity in Spinach Leaves in Relation to Cell Compartmentation-Evidence for Lack of DNA Photolyase in Chloroplasts. *Biosci, Biotechnol, Biochem* **64**, 1288-1291.
- Hajdukiewicz PTJ, Allison LA, et 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.
- Handa H (2003). The complete nucleotide sequence and RNA editing content of the mitochondrial genome of rapeseed (*Brassica napus* L.): comparative analysis of the mitochondrial genomes of rapeseed and Arabidopsis thaliana. *Nucleic Acids Res* **31**, 5907-5916.
- Handa N, Amitani I, Gumlaw N, Sandler SJ, et Kowalczykowski SC (2009). Single Molecule Analysis of a Red Fluorescent RecA Protein Reveals a Defect in Nucleoprotein Filament Nucleation That Relates to Its Reduced Biological Functions. *J Biol Chem* **284**, 18664-18673.
- Hanson MR, et Bentolila S (2004). Interactions of mitochondrial and nuclear genes that affect male gametophyte development. *Plant Cell* **16 Suppl**, S154-169.
- Harrison SJ, Mott EK, Parsley K, Aspinnall S, Gray JC, et Cottage A (2006). A rapid and robust method of identifying transformed Arabidopsis thaliana seedlings following floral dip transformation. *Plant Methods* **2**, 19.

- He L, Zhang S, Qiu Z, Zhao J, Nie W, Lin H, Zhu Z, Zeng D, Qian Q, et Zhu L (2018). FRUCTOKINASE-LIKE PROTEIN 1 interacts with TRXz to regulate chloroplast development in rice: FLN1 and TRXz regulate chloroplast development in rice. *J Integr Plant Biol* **60**, 94-111.
- Hedtke B, Borner T, et Weihe A (1997). Mitochondrial and chloroplast phage-type RNA polymerases in Arabidopsis. *Science* **277**, 809-811.
- Hedtke B, Borner T, et Weihe A (2000). One RNA polymerase serving two genomes. *EMBO Rep* **1**, 435-440.
- Herrmann JM, et Riemer J (2010). The intermembrane space of mitochondria. *Antioxid Redox Signal* **13**, 1341-1358.
- Hess WR, et Borner T (1999). Organellar RNA polymerases of higher plants. In *Int Rev Cytol*, Jeon KW, ed (525 B Street, Suite 1900, San Diego, CA 92101-4495, USA: Academic Press Inc), pp. 1-59.
- Hiesel R, Wissinger B, Schuster W, et Brennicke A (1989). RNA editing in plant mitochondria. *Science* **246**, 1632-1634.
- Higgs DC, Kuras R, Kindle KL, Wollman FA, et Stern DB (1998). Inversions in the Chlamydomonas chloroplast genome suppress a petD 5' untranslated region deletion by creating functional chimeric mRNAs. *Plant J* **14**, 663-671.
- Hoch B, Maier RM, Appel K, Igloi GL, et Kossel H (1991). Editing of a chloroplast mRNA by creation of an initiation codon. *Nature* **353**, 178-180.
- Holec S, Lange H, Kühn K, Alioua M, Borner T, et 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.
- Horvath SE, et Daum G (2013). Lipids of mitochondria. *Prog Lipid Res* **52**, 590-614.
- Hricova A, Quesada V, et Micol JL (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.
- Hu J, Adar S, Selby CP, Lieb JD, et Sancar A (2015). Genome-wide analysis of human global and transcription-coupled excision repair of UV damage at single-nucleotide resolution. *Genes Dev* **29**, 948-960.
- Huang CY, Ayliffe MA, et Timmis JN (2003). Direct measurement of the transfer rate of chloroplast DNA into the nucleus. *Nature* **422**, 72-76.
- Hudik E, Yoshioka Y, Domenichini S, Bourge M, Soubigout-Taconnat L, Mazubert C, Yi D, Bujaldon S, Hayashi H, De Veylder L, Bergounioux C, Benhamed M, et Raynaud C (2014). Chloroplast dysfunction causes multiple defects in cell cycle progression in the Arabidopsis crumpled leaf mutant. *Plant Physiol* **166**, 152-167.
- Inoue M, Fukui K, Fujii Y, Nakagawa N, Yano T, Kuramitsu S, et Masui R (2017). The Lon protease-like domain in the bacterial RecA paralog RadA is required for DNA binding and repair. *J Biol Chem* **292**, 9801-9814.
- Ip SCY, Rass U, Blanco MG, Flynn HR, Skehel JM, et West SC (2008). Identification of Holliday junction resolvases from humans and yeast. *Nature* **456**, 357-361.
- Ishibashi T, Isogai M, Kiyohara H, Hosaka M, Chiku H, Koga A, Yamamoto T, Uchiyama Y, Mori Y, Hashimoto J, Ausio J, Kimura S, et Sakaguchi K (2006). Higher plant RecA-like protein is homologous to RadA. *DNA Repair (Amst)* **5**, 80-88.
- Islam MS, Niwa Y, et Takagi S (2009). Light-dependent intracellular positioning of mitochondria in Arabidopsis thaliana mesophyll cells. *Plant Cell Physiol* **50**, 1032-1040.
- Jain S, Gupta R, et Sen R (2019). Rho-dependent transcription termination in bacteria recycles RNA polymerases stalled at DNA lesions. *Nat Commun* **10**, 1207.
- Janicka S, Kühn K, Le Ret M, Bonnard G, Imbault P, Augustyniak H, et Gualberto JM (2012). A RAD52-like single-stranded DNA binding protein affects mitochondrial DNA repair by recombination. *Plant J* **72**, 423-435.
- Ji D, Manavski N, Meurer J, Zhang L, et Chi W (2019). Regulated chloroplast transcription termination. *Biochim Biophys Acta Bioenerg* **1860**, 69-77.
- Johnson MP (2016). Photosynthesis. *Essays Biochem* **60**, 255-273.
- Johzuka K, et Ogawa H (1995). Interaction of Mre11 and Rad50: two proteins required for DNA repair and meiosis-specific double-strand break formation in *Saccharomyces cerevisiae*. *Genetics* **139**, 1521-1532.

- Kamenisch Y, Fousteri M, Knoch J, von Thaler A-K, Fehrenbacher B, Kato H, Becker T, Dollé MET, Kuiper R, Majora M, Schaller M, van der Horst GTJ, van Steeg H, Röcken M, Rapaport D, Krutmann J, Mullenders LH, et Berneburg M (2010). Proteins of nucleotide and base excision repair pathways interact in mitochondria to protect from loss of subcutaneous fat, a hallmark of aging. *J Exp Med* **207**, 379-390.
- Kamenski P, Kolesnikova O, Jubenot V, Entelis N, Krasheninnikov IA, Martin RP, et Tarassov I (2007). Evidence for an adaptation mechanism of mitochondrial translation via tRNA import from the cytosol. *Mol Cell* **26**, 625-637.
- Karnkowska A, Vacek V, Zubacova Z, Treitli SC, Petrzelkova R, Eme L, Novak L, Zarsky V, Barlow LD, Herman EK, Soukal P, Hroudova M, Dolezal P, Stairs CW, Roger AJ, Elias M, Dacks JB, Vlcek C, et Hampl V (2016). A Eukaryote without a Mitochondrial Organelle. *Curr Biol* **26**, 1274-1284.
- Karol MH, et Simpson MV (1968). DNA biosynthesis by isolated mitochondria: a replicative rather than a repair process. *Science* **162**, 470-473.
- Keeling PJ (2013). The number, speed, and impact of plastid endosymbioses in eukaryotic evolution. *Annu Rev Plant Biol* **64**, 583-607.
- Khakhlova O, et Bock R (2006). Elimination of deleterious mutations in plastid genomes by gene conversion. *Plant J* **46**, 85-94.
- Kim KM, Noh JH, Abdelmohsen K, et Gorospe M (2017). Mitochondrial noncoding RNA transport. *BMB Rep* **50**, 164-174.
- Kindgren P, et Strand A (2015). Chloroplast transcription, untangling the Gordian Knot. *New Phytol* **206**, 889-891.
- Kleine T, Lockhart P, et Batschauer A (2003). An *Arabidopsis* protein closely related to *Synechocystis* cryptochrome is targeted to organelles. *Plant J* **35**, 93-103.
- Knoop V (2011). When you can't trust the DNA: RNA editing changes transcript sequences. *Cell Mol Life Sci* **68**, 567-586.
- Komiya T, Sakaguchi M, et Mihara K (1996). Cytoplasmic chaperones determine the targeting pathway of precursor proteins to mitochondria. *EMBO J* **15**, 399-407.
- Kotera E, Tasaka M, et Shikanai T (2005). A pentatricopeptide repeat protein is essential for RNA editing in chloroplasts. *Nature* **433**, 326-330.
- Krejci L, Van Komen S, Li Y, Villemain J, Reddy MS, Klein H, Ellenberger T, et Sung P (2003). DNA helicase Srs2 disrupts the Rad51 presynaptic filament. *Nature* **423**, 305-309.
- Krupinska K, Melonek J, et Krause K (2013). New insights into plastid nucleoid structure and functionality. *Planta* **237**, 653-664.
- Kruse B, Narasimhan N, et Attardi G (1989). Termination of transcription in human mitochondria: Identification and purification of a DNA binding protein factor that promotes termination. *Cell* **58**, 391-397.
- Kubo T, et Newton KJ (2008). Angiosperm mitochondrial genomes and mutations. *Mitochondrion* **8**, 5-14.
- Kühn K, Weihe A, et Borner T (2005). Multiple promoters are a common feature of mitochondrial genes in *Arabidopsis*. *Nucleic Acids Res* **33**, 337-346.
- Kühn K, Bohne AV, Liere K, Weihe A, et Borner T (2007). *Arabidopsis* phage-type RNA polymerases: accurate in vitro transcription of organellar genes. *Plant Cell* **19**, 959-971.
- Kühn K, Richter U, Meyer EH, Delannoy E, de Longevialle AF, O'Toole N, Borner T, Millar AH, Small ID, et Whelan J (2009). Phage-type RNA polymerase RPOTmp performs gene-specific transcription in mitochondria of *Arabidopsis thaliana*. *Plant Cell* **21**, 2762-2779.
- Kukat C, Wurm CA, Spahr H, Falkenberg M, Larsson NG, et Jakobs S (2011). Super-resolution microscopy reveals that mammalian mitochondrial nucleoids have a uniform size and frequently contain a single copy of mtDNA. *Proc Natl Acad Sci U S A* **108**, 13534-13539.
- Kunze M, et Berger J (2015). The similarity between N-terminal targeting signals for protein import into different organelles and its evolutionary relevance. *Front Physiol* **6**, 259.
- Kwon T, Huq E, et Herrin DL (2010). Microhomology-mediated and nonhomologous repair of a double-strand break in the chloroplast genome of *Arabidopsis*. *Proc Natl Acad Sci U S A* **107**, 13954-13959.
- Lang BF, Burger G, O'Kelly CJ, Cedergren R, Golding GB, Lemieux C, Sankoff D, Turmel M, et Gray MW (1997). An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. *Nature* **387**, 493-497.



- Lechner M, Rossmannith W, Hartmann RK, Tholken C, Gutmann B, Giege P, et Gobert A (2015). Distribution of Ribonucleoprotein and Protein-Only RNase P in Eukarya. *Mol Biol Evol* **32**, 3186-3193.
- Lieberman M, Segev O, Gilboa N, Lalazar A, et Levin I (2004). The tomato homolog of the gene encoding UV-damaged DNA binding protein 1 (DDB1) underlined as the gene that causes the high pigment-1 mutant phenotype. *Theor Appl Genet* **108**, 1574-1581.
- Liebers M, Chevalier F, Blanvillain R, et Pfannschmidt T (2018). PAP genes are tissue- and cell-specific markers of chloroplast development. *Planta* **248**, 629-646.
- Liebers M, Grubler B, Chevalier F, Lerbs-Mache S, Merendino L, Blanvillain R, et Pfannschmidt T (2017). Regulatory Shifts in Plastid Transcription Play a Key Role in Morphological Conversions of Plastids during Plant Development. *Front Plant Sci* **8**, 23.
- Liere K, Weihe A, et Börner T (2011). The transcription machineries of plant mitochondria and chloroplasts: Composition, function, and regulation. *J Plant Physiol* **168**, 1345-1360.
- Lilly JW, Havey MJ, Jackson SA, et Jiang J (2001). Cytogenomic analyses reveal the structural plasticity of the chloroplast genome in higher plants. *Plant Cell* **13**, 245-254.
- Liu J, Fang H, Chi Z, Wu Z, Wei D, Mo D, Niu K, Balajee AS, Hei TK, Nie L, et Zhao Y (2015). XPD localizes in mitochondria and protects the mitochondrial genome from oxidative DNA damage. *Nucleic Acids Res* **43**, 5476-5488.
- Logan DC (2006a). The mitochondrial compartment. *J Exp Bot* **57**, 1225-1243.
- Logan DC (2006b). Plant mitochondrial dynamics. *Biochim Biophys Acta, Mol Cell Res* **1763**, 430-441.
- Lovett ST, Hurley RL, Sutera VA, Jr., Aubuchon RH, et Lebedeva MA (2002). Crossing over between regions of limited homology in Escherichia coli. RecA-dependent and RecA-independent pathways. *Genetics* **160**, 851-859.
- Lukeš J, Wheeler R, Jirsová D, David V, et Archibald JM (2018). Massive mitochondrial DNA content in diplomonid and kinetoplastid protists: Massive mitochondrial DNA in protists. *IUBMB Life* **70**, 1267-1274.
- 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 ML, Mireau H, Peeters N, Renou JP, Szurek B, Taconnat L, et Small I (2004). Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell* **16**, 2089-2103.
- Luro S, Germain A, Sharwood RE, et Stern DB (2013). RNase J participates in a pentatricopeptide repeat protein-mediated 5' end maturation of chloroplast mRNAs. *Nucleic Acids Res* **41**, 9141-9151.
- Mackenzie S, et McIntosh L (1999). Higher plant mitochondria. *Plant Cell* **11**, 571-586.
- Macovei A, Balestrazzi A, Confalonieri M, Fae M, et Carbonera D (2011). New insights on the barrel medic MtOGG1 and MtFPG functions in relation to oxidative stress response in planta and during seed imbibition. *Plant Physiol Biochem* **49**, 1040-1050.
- Magalhaes PJ, Andreu AL, et Schon EA (1998). Evidence for the presence of 5S rRNA in mammalian mitochondria. *Mol Biol Cell* **9**, 2375-2382.
- Majeran W, Friso G, Asakura Y, Qu X, Huang M, Ponnala L, Watkins KP, Barkan A, et van Wijk KJ (2012). Nucleoid-enriched proteomes in developing plastids and chloroplasts from maize leaves: a new conceptual framework for nucleoid functions. *Plant Physiol* **158**, 156-189.
- Manelyte L, Kim YI, Smith AJ, Smith RM, et Savery NJ (2010). Regulation and rate enhancement during transcription-coupled DNA repair. *Mol Cell* **40**, 714-724.
- Marechal A, Parent JS, Veronneau-Lafortune F, Joyeux A, Lang BF, et Brisson N (2009). Whirly proteins maintain plastid genome stability in Arabidopsis. *Proc Natl Acad Sci U S A* **106**, 14693-14698.
- Maréchal A, et Brisson N (2010). Recombination and the maintenance of plant organelle genome stability. *New Phytol* **186**, 299-317.
- Marie L, Rapisarda C, Morales V, Bergé M, Perry T, Soulet A-L, Gruget C, Remaut H, Fronzes R, et Polard P (2017). Bacterial RadA is a DnaB-type helicase interacting with RecA to promote bidirectional D-loop extension. *Nat Commun* **8**, 15638.
- Martin HA, Porter KE, Vallin C, Ermi T, Contreras N, Pedraza-Reyes M, et Robleto EA (2019). Mfd protects against oxidative stress in Bacillus subtilis independently of its canonical function in DNA repair. *BMC Microbiol* **19**, 26.

- Massel K, Silke JR, et Bonen L** (2016). Multiple splicing pathways of group II trans-splicing introns in wheat mitochondria. *Mitochondrion* **28**, 23-32.
- Masson JY, Tarsounas MC, Stasiak AZ, Stasiak A, Shah R, McIlwraith MJ, Benson FE, et West SC** (2001). Identification and purification of two distinct complexes containing the five RAD51 paralogs. *Genes Dev* **15**, 3296-3307.
- Massouh A, Schubert J, Yaneva-Roder L, Ulbricht-Jones ES, Zupok A, Johnson MT, Wright SI, Pellizzer T, Sobanski J, Bock R, et Greiner S** (2016). Spontaneous Chloroplast Mutants Mostly Occur by Replication Slippage and Show a Biased Pattern in the Plastome of *Oenothera*. *Plant Cell* **28**, 911-929.
- Matos J, et West SC** (2014). Holliday junction resolution: Regulation in space and time. *DNA Repair* **19**, 176-181.
- Mawer JS, et Leach DR** (2014). Branch migration prevents DNA loss during double-strand break repair. *PLoS Genet* **10**, e1004485.
- Maxwell DP, Wang Y, et McIntosh L** (1999). The alternative oxidase lowers mitochondrial reactive oxygen production in plant cells. *Proc Natl Acad Sci U S A* **96**, 8271-8276.
- McFadden GI** (2001). PRIMARY AND SECONDARY ENDOSYMBIOSIS AND THE ORIGIN. *J Phycol* **37**, 951-959.
- McGlynn P, et Lloyd RG** (2002). Recombinational repair and restart of damaged replication forks. *Nat Rev Mol Cell Biol* **3**, 859-870.
- Meierhoff K, Felder S, Nakamura T, Bechtold N, et 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.
- Melonek J, Oetke S, et Krupinska K** (2016). Multifunctionality of plastid nucleoids as revealed by proteome analyses. *Biochim Biophys Acta, Proteins Proteomics* **1864**, 1016-1038.
- Melonek J, Matros A, Trosch M, Mock H-P, et Krupinska K** (2012). The Core of Chloroplast Nucleoids Contains Architectural SWIB Domain Proteins. *Plant Cell* **24**, 3060-3073.
- Menoni H, Wienholz F, Theil AF, Janssens RC, Lans H, Campalans A, Radicella JP, Marteiijn JA, et Vermeulen W** (2018). The transcription-coupled DNA repair-initiating protein CSB promotes XRCC1 recruitment to oxidative DNA damage. *Nucleic Acids Res* **46**, 7747-7756.
- Miller-Messmer M, Kühn K, Bichara M, Le Ret M, Imbault P, et Gualberto JM** (2012). RecA-dependent DNA repair results in increased heteroplasmy of the Arabidopsis mitochondrial genome. *Plant Physiol* **159**, 211-226.
- Mittler R, Vanderauwera S, Suzuki N, Miller G, Tognetti VB, Vandepoele K, Gollery M, Shulaev V, et Van Breusegem F** (2011). ROS signaling: the new wave? *Trends Plant Sci* **16**, 300-309.
- Molinier J, Lechner E, Dumbliauskas E, et Genschik P** (2008). Regulation and role of Arabidopsis CUL4-DDB1A-DDB2 in maintaining genome integrity upon UV stress. *PLoS Genet* **4**, e1000093.
- Moller IM** (2001). PLANT MITOCHONDRIA AND OXIDATIVE STRESS: Electron Transport, NADPH Turnover, and Metabolism of Reactive Oxygen Species. *Annu Rev Plant Physiol Plant Mol Biol* **52**, 561-591.
- Moreira D, et Lopez-Garcia P** (1998). Symbiosis between methanogenic archaea and delta-proteobacteria as the origin of eukaryotes: the syntrophic hypothesis. *J Mol Evol* **47**, 517-530.
- Morimatsu K, et Kowalczykowski SC** (2014). RecQ helicase and RecJ nuclease provide complementary functions to resect DNA for homologous recombination. *Proc Natl Acad Sci U S A* **111**, E5133-5142.
- Morita R, Nakane S, Shimada A, Inoue M, Iino H, Wakamatsu T, Fukui K, Nakagawa N, Masui R, et Kuramitsu S** (2010). Molecular mechanisms of the whole DNA repair system: a comparison of bacterial and eukaryotic systems. *J Nucleic Acids* **2010**, 179594.
- Mower JP, Touzet P, Gummow JS, Delph LF, et Palmer JD** (2007). Extensive variation in synonymous substitution rates in mitochondrial genes of seed plants. *BMC Evol Biol* **7**, 135.
- Muller M, Mentel M, van Hellemond JJ, Henze K, Woehle C, Gould SB, Yu RY, van der Giezen M, Tielens AG, et Martin WF** (2012). Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. *Microbiol Mol Biol Rev* **76**, 444-495.
- Murray V** (1989). Improved double-stranded DNA sequencing using the linear polymerase chain reaction. *Nucleic Acids Res* **17**, 8889.

- Myouga F, Hosoda C, Umezawa T, Iizumi H, Kuromori T, Motohashi R, Shono Y, Nagata N, Ikeuchi M, et Shinozaki K (2008). A Heterocomplex of Iron Superoxide Dismutases Defends Chloroplast Nucleoids against Oxidative Stress and Is Essential for Chloroplast Development in Arabidopsis. *Plant Cell* **20**, 3148-3162.
- Nakai M (2018). New Perspectives on Chloroplast Protein Import. *Plant Cell Physiol* **59**, 1111-1119.
- Nakayama T, et Ishida K (2009). Another acquisition of a primary photosynthetic organelle is underway in *Paulinella chromatophora*. *Curr Biol* **19**, R284-285.
- Nassif N, Penney J, Pal S, Engels WR, et Gloor GB (1994). Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. *Mol Cell Biol* **14**, 1613-1625.
- Neale MJ, Pan J, et Keeney S (2005). Endonucleolytic processing of covalent protein-linked DNA double-strand breaks. *Nature* **436**, 1053-1057.
- Neuwald AF, Berg DE, et Stauffer GV (1992). Mutational analysis of the *Escherichia coli* serB promoter region reveals transcriptional linkage to a downstream gene. *Gene* **120**, 1-9.
- Nevarez PA, Qiu Y, Inoue H, Yoo CY, Benfey PN, Schnell DJ, et Chen M (2017). Mechanism of Dual Targeting of the Phytochrome Signaling Component HEMERA/pTAC12 to Plastids and the Nucleus. *Plant Physiol* **173**, 1953-1966.
- Nimonkar AV, Sica RA, et Kowalczykowski SC (2009). Rad52 promotes second-end DNA capture in double-stranded break repair to form complement-stabilized joint molecules. *Proc Natl Acad Sci U S A* **106**, 3077-3082.
- Noh JH, Kim KM, Abdelmohsen K, Yoon JH, Panda AC, Munk R, Kim J, Curtis J, Moad CA, Wohler CM, Indig FE, de Paula W, Dudekula DB, De S, Piao Y, Yang X, Martindale JL, de Cabo R, et Gorospe M (2016). HuR and GRSF1 modulate the nuclear export and mitochondrial localization of the lncRNA RMRP. *Genes Dev* **30**, 1224-1239.
- Nowack EC, Vogel H, Groth M, Grossman AR, Melkonian M, et Glockner G (2011). Endosymbiotic gene transfer and transcriptional regulation of transferred genes in *Paulinella chromatophora*. *Mol Biol Evol* **28**, 407-422.
- Nowack MK, Harashima H, Dissmeyer N, Zhao X, Bouyer D, Weimer AK, De Winter F, Yang F, et Schnittger A (2012). Genetic framework of cyclin-dependent kinase function in Arabidopsis. *Dev Cell* **22**, 1030-1040.
- O'Rourke B (2007). Mitochondrial ion channels. *Annu Rev Physiol* **69**, 19-49.
- Obara K, Sumi K, et Fukuda H (2002). The use of multiple transcription starts causes the dual targeting of Arabidopsis putative monodehydroascorbate reductase to both mitochondria and chloroplasts. *Plant Cell Physiol* **43**, 697-705.
- Odahara M, et Sekine Y (2018). RECX interacts with mitochondrial RECA to maintain mitochondrial genome stability. *Plant Physiol*, pp.00218.02018.
- Odahara M, Inouye T, Nishimura Y, et Sekine Y (2015). RECA plays a dual role in the maintenance of chloroplast genome stability in *Physcomitrella patens*. *Plant J* **84**, 516-526.
- Oldenkott B, Yang Y, Lesch E, Knoop V, et Schallenberg-Rudinger M (2019). Plant-type pentatricopeptide repeat proteins with a DYW domain drive C-to-U RNA editing in *Escherichia coli*. *Commun Biol* **2**, 85.
- Orthwein A, Noordermeer SM, Wilson MD, Landry S, Enchev RI, Sherker A, Munro M, Pinder J, Salsman J, Dellaire G, Xia B, Peter M, et Durocher D (2015). A mechanism for the suppression of homologous recombination in G1 cells. *Nature* **528**, 422-426.
- Palmer JD, et Herbon LA (1988). Plant mitochondrial DNA evolves rapidly in structure but slowly in sequence. *J Mol Evol* **28**, 87-97.
- Parent JS, Lepage E, et Brisson N (2011). Divergent roles for the two PolII-like organelle DNA polymerases of Arabidopsis. *Plant Physiol* **156**, 254-262.
- Park JG, Kim DW, Hong CW, Nam BH, Shin YK, Hong SH, Kim IJ, Lim SB, Aronson M, Bisgaard ML, Brown GJ, Burn J, Chow E, Conrad P, Douglas F, Dunlop M, Ford J, Greenblatt MS, Heikki J, Heinimann K, Lynch EL, Macrae F, McKinnon WC, Moeslein G, Rossi BM, Rozen P, Schofield L, Vaccaro C, Vasen H, Velthuisen M, Viel A, et Wijnen J (2006). Germ line mutations of mismatch repair genes in hereditary nonpolyposis colorectal cancer patients with small bowel cancer: International Society for Gastrointestinal Hereditary Tumours Collaborative Study. *Clin Cancer Res* **12**, 3389-3393.

- Park JS, Marr MT, et Roberts JW (2002). E. coli Transcription repair coupling factor (Mfd protein) rescues arrested complexes by promoting forward translocation. *Cell* **109**, 757-767.
- Pereira de Souza A, Jubier MF, Delcher E, Lancelin D, et Lejeune B (1991). A trans-splicing model for the expression of the tripartite nad5 gene in wheat and maize mitochondria. *Plant Cell* **3**, 1363-1378.
- Perrin R, Meyer EH, Zaepfel M, Kim YJ, Mache R, Grienenberger JM, Gualberto JM, et Gagliardi D (2004). Two exoribonucleases act sequentially to process mature 3'-ends of atp9 mRNAs in Arabidopsis mitochondria. *J Biol Chem* **279**, 25440-25446.
- Petrova V, Chen SH, Molzberger ET, Tomko E, Chitteni-Pattu S, Jia H, Ordabayev Y, Lohman TM, et Cox MM (2015). Active displacement of RecA filaments by UvrD translocase activity. *Nucleic Acids Res* **43**, 4133-4149.
- Pfalz J, et Pfannschmidt T (2013). Essential nucleoid proteins in early chloroplast development. *Trends Plant Sci* **18**, 186-194.
- Pfalz J, Liere K, Kandlbinder A, Dietz KJ, et 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.
- Pittis AA, et Gabaldon T (2016). Late acquisition of mitochondria by a host with chimaeric prokaryotic ancestry. *Nature* **531**, 101-104.
- Pogson BJ, Ganguly D, et Albrecht-Borth V (2015). Insights into chloroplast biogenesis and development. *Biochim Biophys Acta* **1847**, 1017-1024.
- Poole AM, et Gribaldo S (2014). Eukaryotic origins: How and when was the mitochondrion acquired? *Cold Spring Harb Perspect Biol* **6**, a015990.
- Prabha S, Rao DN, et Nagaraja V (2011). Distinct properties of hexameric but functionally conserved Mycobacterium tuberculosis transcription-repair coupling factor. *PLoS ONE* **6**, e19131.
- Pramateftaki PV, Kouvelis VN, Lanaridis P, et Typas MA (2006). The mitochondrial genome of the wine yeast *Hanseniaspora uvarum* : a unique genome organization among yeast/fungal counterparts. *FEMS Yeast Res* **6**, 77-90.
- Preuten T, Cincu E, Fuchs J, Zoschke R, Liere K, et Borner T (2010). Fewer genes than organelles: extremely low and variable gene copy numbers in mitochondria of somatic plant cells. *Plant J* **64**, 948-959.
- Puchta H (1999). Double-strand break-induced recombination between ectopic homologous sequences in somatic plant cells. *Genetics* **152**, 1173-1181.
- Pyke KA (1999). Plastid division and development. *Plant Cell* **11**, 549-556.
- Qing Y, Yamazoe M, Hirota K, Dejsuphong D, Sakai W, Yamamoto KN, Bishop DK, Wu X, et Takeda S (2011). The epistatic relationship between BRCA2 and the other RAD51 mediators in homologous recombination. *PLoS Genet* **7**, e1002148.
- Ragunathan K, Liu C, et Ha T (2012). RecA filament sliding on DNA facilitates homology search. *Elife* **1**, e00067.
- Rain JC, Selig L, De Reuse H, Battaglia V, Reverdy C, Simon S, Lenzen G, Petel F, Wojcik J, Schachter V, Chemama Y, Labigne A, et Legrain P (2001). The protein-protein interaction map of Helicobacter pylori. *Nature* **409**, 211-215.
- Ranjha L, Howard SM, et Cejka P (2018). Main steps in DNA double-strand break repair: an introduction to homologous recombination and related processes. *Chromosoma* **127**, 187-214.
- Richter U, Kühn K, Okada S, Brennicke A, Weihe A, et Borner T (2010). A mitochondrial rRNA dimethyladenosine methyltransferase in Arabidopsis. *Plant J* **61**, 558-569.
- Riou-Khamlichi C, Menges M, Healy JM, et Murray JA (2000). Sugar control of the plant cell cycle: differential regulation of Arabidopsis D-type cyclin gene expression. *Mol Cell Biol* **20**, 4513-4521.
- Rödiger A, Baudisch B, Langner U, et Klosgen RB (2011). Dual targeting of a mitochondrial protein: the case study of cytochrome c1. *Mol Plant* **4**, 679-687.
- Romani I, Manavski N, Morosetti A, Tadini L, Maier S, Kühn K, Ruwe H, Schmitz-Linneweber C, Wanner G, Leister D, et Kleine T (2015). A Member of the Arabidopsis Mitochondrial Transcription Termination Factor Family Is Required for Maturation of Chloroplast Transfer RNA<sup>lle</sup> (GAU). *Plant Physiol* **169**, 627-646.

- Roy R, Kozlov AG, Lohman TM, et Ha T (2009). SSB protein diffusion on single-stranded DNA stimulates RecA filament formation. *Nature* **461**, 1092-1097.
- Ruthel G, et Hollenbeck PJ (2003). Response of mitochondrial traffic to axon determination and differential branch growth. *J Neurosci* **23**, 8618-8624.
- Sagan L (1967). On the origin of mitosing cells. *J Theor Biol* **14**, 255-274.
- Sakofsky CJ, et Malkova A (2017). Break induced replication in eukaryotes: mechanisms, functions, and consequences. *Crit Rev Biochem Mol Biol* **52**, 395-413.
- Salinas T, Duchene AM, Delage L, Nilsson S, Glaser E, Zaepfel M, et Marechal-Drouard L (2006). The voltage-dependent anion channel, a major component of the tRNA import machinery in plant mitochondria. *Proc Natl Acad Sci U S A* **103**, 18362-18367.
- Salinas-Giege T, Giege R, et Giege P (2015). tRNA biology in mitochondria. *Int J Mol Sci* **16**, 4518-4559.
- Samach A, Melamed-Bessudo C, Avivi-Ragolski N, Pietrokovski S, et Levy AA (2011). Identification of plant RAD52 homologs and characterization of the Arabidopsis thaliana RAD52-like genes. *Plant Cell* **23**, 4266-4279.
- Sancar A, et Sancar GB (1988). DNA repair enzymes. *Annu Rev Biochem* **57**, 29-67.
- Sandler SJ, et Clark AJ (1994). RecOR suppression of recF mutant phenotypes in Escherichia coli K-12. *J Bacteriol* **176**, 3661-3672.
- Sanger F, Nicklen S, et Coulson AR (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74**, 5463-5467.
- Sankar TS, Wastuwidyaningtyas BD, Dong Y, Lewis SA, et Wang JD (2016). The nature of mutations induced by replication-transcription collisions. *Nature* **535**, 178-181.
- Sato S, Nakamura Y, Kaneko T, Asamizu E, et Tabata S (1999). Complete structure of the chloroplast genome of Arabidopsis thaliana. *DNA Res* **6**, 283-290.
- Schuermann D, Molinier J, Fritsch O, et Hohn B (2005). The dual nature of homologous recombination in plants. *Trends Genet* **21**, 172-181.
- Scott I, et Logan DC (2008). Mitochondrial morphology transition is an early indicator of subsequent cell death in Arabidopsis. *New Phytol* **177**, 90-101.
- Selby CP (2017). Mfd Protein and Transcription-Repair Coupling in *Escherichia coli*. *Photochem Photobiol* **93**, 280-295.
- Selby CP, et Sancar A (1995a). Structure and function of transcription-repair coupling factor. I. Structural domains and binding properties. *J Biol Chem* **270**, 4882-4889.
- Selby CP, et Sancar A (1995b). Structure and function of transcription-repair coupling factor. II. Catalytic properties. *J Biol Chem* **270**, 4890-4895.
- Selby CP, Witkin EM, et Sancar A (1991). Escherichia coli mfd mutant deficient in "mutation frequency decline" lacks strand-specific repair: in vitro complementation with purified coupling factor. *Proc Natl Acad Sci U S A* **88**, 11574-11578.
- Severinov K, Mustaev A, Kukarin A, Muzzin O, Bass I, Darst SA, et Goldfarb A (1996). Structural modules of the large subunits of RNA polymerase. Introducing archaebacterial and chloroplast split sites in the beta and beta' subunits of Escherichia coli RNA polymerase. *J Biol Chem* **271**, 27969-27974.
- Shedge V, Arrieta-Montiel M, Christensen AC, et Mackenzie SA (2007). Plant mitochondrial recombination surveillance requires unusual RecA and MutS homologs. *Plant Cell* **19**, 1251-1264.
- Shi C, Wang S, Xia E-H, Jiang J-J, Zeng F-C, et Gao L-Z (2016). Full transcription of the chloroplast genome in photosynthetic eukaryotes. *Sci Rep* **6**.
- Shibata A (2017). Regulation of repair pathway choice at two-ended DNA double-strand breaks. *Mutat Res* **803-805**, 51-55.
- Silva-Filho MC (2003). One ticket for multiple destinations: dual targeting of proteins to distinct subcellular locations. *Current Opin Plant Biol* **6**, 589-595.
- Singleton MR, Scaife S, et Wigley DB (2001). Structural analysis of DNA replication fork reversal by RecG. *Cell* **107**, 79-89.
- Slade D, Lindner AB, Paul G, et Radman M (2009). Recombination and replication in DNA repair of heavily irradiated Deinococcus radiodurans. *Cell* **136**, 1044-1055.
- Sloan DB, Wu Z, et Sharbrough J (2018). Correction of Persistent Errors in Arabidopsis Reference Mitochondrial Genomes. *Plant Cell* **30**, 525-527.

- Sloan DB, Oxelman B, Rautenberg A, et Taylor DR (2009). Phylogenetic analysis of mitochondrial substitution rate variation in the angiosperm tribe Sileneae. *BMC Evol Biol* **9**, 260.
- Sloan DB, Alverson AJ, Chuckalovcak JP, Wu M, McCauley DE, Palmer JD, et Taylor DR (2012). Rapid Evolution of Enormous, Multichromosomal Genomes in Flowering Plant Mitochondria with Exceptionally High Mutation Rates. *PLoS Biol* **10**, e1001241.
- Small GD (1987). Repair systems for nuclear and chloroplast DNA in *Chlamydomonas reinhardtii*. *Mutat Res* **181**, 31-35.
- Small I, Suffolk R, et Leaver CJ (1989). Evolution of plant mitochondrial genomes via substoichiometric intermediates. *Cell* **58**, 69-76.
- Small ID, et Peeters N (2000). The PPR motif - a TPR-related motif prevalent in plant organellar proteins. *Trends Biochem Sci* **25**, 46-47.
- Small ID, Isaac PG, et Leaver CJ (1987). Stoichiometric differences in DNA molecules containing the atpA gene suggest mechanisms for the generation of mitochondrial genome diversity in maize. *EMBO J* **6**, 865-869.
- Small ID, Rackham O, et Filipovska A (2013). Organelle transcriptomes: products of a deconstructed genome. *Curr Opin Microbiol* **16**, 652-658.
- Smith LM, Sanders JZ, Kaiser RJ, Hughes P, Dodd C, Connell CR, Heiner C, Kent SB, et Hood LE (1986). Fluorescence detection in automated DNA sequence analysis. *Nature* **321**, 674-679.
- Solinger JA, Kiiantsa K, et Heyer W-D (2002). Rad54, a Swi2/Snf2-like Recombinational Repair Protein, Disassembles Rad51:dsDNA Filaments. *Mol Cell* **10**, 1175-1188.
- Song S, Pursell ZF, Copeland WC, Longley MJ, Kunkel TA, et Mathews CK (2005). DNA precursor asymmetries in mammalian tissue mitochondria and possible contribution to mutagenesis through reduced replication fidelity. *Proc Natl Acad Sci U S A* **102**, 4990-4995.
- Song Y, et Sargentini NJ (1996). Escherichia coli DNA repair genes radA and sms are the same gene. *J Bacteriol* **178**, 5045-5048.
- Sousa JS, D'Imprima E, et Vonck J (2018). Mitochondrial Respiratory Chain Complexes. *Subcell Biochem* **87**, 167-227.
- Stein A, et Sia EA (2017). Mitochondrial DNA repair and damage tolerance. *Front Biosci (Landmark Ed)* **22**, 920-943.
- Steiner S, Schroter Y, Pfalz J, et Pfannschmidt T (2011). Identification of Essential Subunits in the Plastid-Encoded RNA Polymerase Complex Reveals Building Blocks for Proper Plastid Development. *Plant Physiol* **157**, 1043-1055.
- Steinmetz A, Gubbins EJ, et Bogorad L (1982). The anticodon of the maize chloroplast gene for tRNA Leu UAA is split by a large intron. *Nucleic Acids Res* **10**, 3027-3037.
- Stevnsner T, Nyaga S, de Souza-Pinto NC, van der Horst GTJ, Gorgels TGMF, Hogue BA, Thorslund T, et Bohr VA (2002). Mitochondrial repair of 8-oxoguanine is deficient in Cockayne syndrome group B. *Oncogene* **21**, 8675-8682.
- Stock D, Gibbons C, Arechaga I, Leslie AG, et Walker JE (2000). The rotary mechanism of ATP synthase. *Curr Opin Struct Biol* **10**, 672-679.
- Stupar RM, Lilly JW, Town CD, Cheng Z, Kaul S, Buell CR, et Jiang J (2001). Complex mtDNA constitutes an approximate 620-kb insertion on Arabidopsis thaliana chromosome 2: implication of potential sequencing errors caused by large-unit repeats. *Proc Natl Acad Sci U S A* **98**, 5099-5103.
- Tadi SK, Sebastian R, Dahal S, Babu RK, Choudhary B, et Raghavan SC (2016). Microhomology-mediated end joining is the principal mediator of double-strand break repair during mitochondrial DNA lesions. *Mol Biol Cell* **27**, 223-235.
- Takahashi M, Teranishi M, Ishida H, Kawasaki J, Takeuchi A, Yamaya T, Watanabe M, Makino A, et Hidema J (2011). Cyclobutane pyrimidine dimer (CPD) photolyase repairs ultraviolet-B-induced CPDs in rice chloroplast and mitochondrial DNA: CPD photolyase in rice chloroplasts and mitochondria. *Plant J* **66**, 433-442.
- Takenaka M, Zehrmann A, Verbitskiy D, Hartel B, et Brennicke A (2013). RNA editing in plants and its evolution. *Annu Rev Genet* **47**, 335-352.
- Taylor AF, Amundsen SK, Guttman M, Lee KK, Luo J, Ranish J, et Smith GR (2014). Control of RecBCD enzyme activity by DNA binding- and Chi hotspot-dependent conformational changes. *J Mol Biol* **426**, 3479-3499.

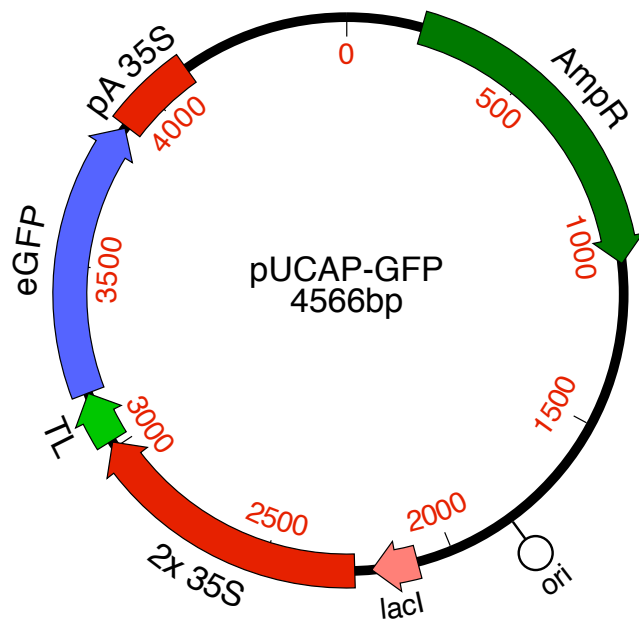
- Teardo E, Carraretto L, De Bortoli S, Costa A, Behera S, Wagner R, Lo Schiavo F, Formentin E, et Szabo I (2015). Alternative splicing-mediated targeting of the Arabidopsis GLUTAMATE RECEPTOR3.5 to mitochondria affects organelle morphology. *Plant Physiol* **167**, 216-227.
- Thiel H, Hleibieh K, Gilmer D, et Varrelmann M (2012). The P25 pathogenicity factor of Beet necrotic yellow vein virus targets the sugar beet 26S proteasome involved in the induction of a hypersensitive resistance response via interaction with an F-box protein. *Mol Plant Microbe Int* **25**, 1058-1072.
- Timmis JN, Ayliffe MA, Huang CY, et Martin W (2004). Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nat Rev Genet* **5**, 123-135.
- Torres R, Serrano E, Tramm K, et Alonso JC (2019). Bacillus subtilis RadA/Sms contributes to chromosomal transformation and DNA repair in concert with RecA and circumvents replicative stress in concert with DisA. *DNA Repair (Amst)* **77**, 45-57.
- Touzet P, et Meyer EH (2014). Cytoplasmic male sterility and mitochondrial metabolism in plants. *Mitochondrion* **19 Pt B**, 166-171.
- Trautinger BW, Jaktaji RP, Rusakova E, et Lloyd RG (2005). RNA polymerase modulators and DNA repair activities resolve conflicts between DNA replication and transcription. *Mol Cell* **19**, 247-258.
- Umez K, et Kolodner RD (1994). Protein interactions in genetic recombination in Escherichia coli. Interactions involving RecO and RecR overcome the inhibition of RecA by single-stranded DNA-binding protein. *J Biol Chem* **269**, 30005-30013.
- Unsel M, Marienfeld JR, Brandt P, et Brennicke A (1997). The mitochondrial genome of Arabidopsis thaliana contains 57 genes in 366,924 nucleotides. *Nat Genet* **15**, 57-61.
- van Engelen FA, Molthoff JW, Conner AJ, Nap JP, Pereira A, et Stiekema WJ (1995). pBINPLUS: an improved plant transformation vector based on pBIN19. *Transgenic Res* **4**, 288-290.
- Vermel M, Guermann B, Delage L, Grienemberger JM, Marechal-Drouard L, et Gualberto JM (2002). A family of RRM-type RNA-binding proteins specific to plant mitochondria. *Proc Natl Acad Sci U S A* **99**, 5866-5871.
- von Heijne G, Steppuhn J, et Herrmann RG (1989). Domain structure of mitochondrial and chloroplast targeting peptides. *Eur J Biochem* **180**, 535-545.
- Wallet C (2016). L'hélicase RECG1, un facteur-clé dans le maintien et la ségrégation de l'ADN mitochondrial d'Arabidopsis thaliana (Université de Strasbourg).
- Wallet C, Le Ret M, Bergdoll M, Bichara M, Dietrich A, et Gualberto JM (2015). The RECG1 DNA Translocase Is a Key Factor in Recombination Surveillance, Repair, and Segregation of the Mitochondrial DNA in Arabidopsis. *Plant Cell* **27**, 2907-2925.
- Wamboldt Y, Mohammed S, Elowsky C, Wittgren C, de Paula WB, et Mackenzie SA (2009). Participation of leaky ribosome scanning in protein dual targeting by alternative translation initiation in higher plants. *Plant Cell* **21**, 157-167.
- Wang D, Liu H, Zhai G, Wang L, Shao J, et Tao Y (2016). OspTAC2 encodes a pentatricopeptide repeat protein and regulates rice chloroplast development. *J Genet Genomics* **43**, 601-608.
- Wardrope L, Okely E, et Leach D (2009). Resolution of joint molecules by RuvABC and RecG following cleavage of the Escherichia coli chromosome by EcoKI. *PLoS ONE* **4**, e6542.
- West SC (1997). Processing of recombination intermediates by the RuvABC proteins. *Annu Rev Genet* **31**, 213-244.
- Westblade LF, Campbell EA, Pukhrambam C, Padovan JC, Nickels BE, Lamour V, et Darst SA (2010). Structural basis for the bacterial transcription-repair coupling factor/RNA polymerase interaction. *Nucleic Acids Res* **38**, 8357-8369.
- Whitby MC, Ryder L, et Lloyd RG (1993). Reverse branch migration of Holliday junctions by RecG protein: a new mechanism for resolution of intermediates in recombination and DNA repair. *Cell* **75**, 341-350.
- Whitby MC, Vincent SD, et Lloyd RG (1994). Branch migration of Holliday junctions: identification of RecG protein as a junction specific DNA helicase. *EMBO J* **13**, 5220-5228.
- Wiczarz M, Gubernator B, Kruk J, et Niewiadomska E (2015). Enhanced chloroplastic generation of H<sub>2</sub>O<sub>2</sub> in stress-resistant Thellungiella salsuginea in comparison to Arabidopsis thaliana. *Physiol Plant* **153**, 467-476.
- Wiedemann N, et Pfanner N (2017). Mitochondrial Machineries for Protein Import and Assembly. *Annu Rev Biochem* **86**, 685-714.

- Witkin EM (1956). Time, temperature, and protein synthesis: a study of ultraviolet-induced mutation in bacteria. *Cold Spring Harb Symp Quant Biol* **21**, 123-140.
- Witkin EM (1966). Radiation-induced mutations and their repair. *Science* **152**, 1345-1353.
- Wobbe L, et Nixon PJ (2013). The mTERF protein MOC1 terminates mitochondrial DNA transcription in the unicellular green alga *Chlamydomonas reinhardtii*. *Nucleic Acids Res* **41**, 6553-6567.
- Wolfe KH, Li WH, et Sharp PM (1987). Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc Natl Acad Sci U S A* **84**, 9054-9058.
- Woloszynska M (2010). Heteroplasmy and stoichiometric complexity of plant mitochondrial genomes--though this be madness, yet there's method in't. *J Exp Bot* **61**, 657-671.
- Woloszynska M, et Trojanowski D (2009). Counting mtDNA molecules in *Phaseolus vulgaris*: sublimons are constantly produced by recombination via short repeats and undergo rigorous selection during substoichiometric shifting. *Plant Mol Biol* **70**, 511-521.
- Woodson JD, et Chory J (2008). Coordination of gene expression between organellar and nuclear genomes. *Nat Rev Genet* **9**, 383-395.
- Woody ST, Austin-Phillips S, Amasino RM, et Krysan PJ (2007). The WiscDsLox T-DNA collection: an arabidopsis community resource generated by using an improved high-throughput T-DNA sequencing pipeline. *J Plant Res* **120**, 157-165.
- Wu Z, Cuthbert JM, Taylor DR, et Sloan DB (2015). The massive mitochondrial genome of the angiosperm *Silene noctiflora* is evolving by gain or loss of entire chromosomes. *Proc Natl Acad Sci U S A* **112**, 10185-10191.
- Xu Y-Z, Arrieta-Montiel MP, Viridi KS, de Paula WBM, Widhalm JR, Basset GJ, Davila JI, Elthon TE, Elowsky CG, Sato SJ, Clemente TE, et Mackenzie SA (2011). MutS HOMOLOG1 Is a Nucleoid Protein That Alters Mitochondrial and Plastid Properties and Plant Response to High Light. *Plant Cell* **23**, 3428-3441.
- Xu YZ, Santamaria Rde L, Viridi KS, Arrieta-Montiel MP, Razvi F, Li S, Ren G, Yu B, Alexander D, Guo L, Feng X, Dweikat IM, Clemente TE, et Mackenzie SA (2012). The chloroplast triggers developmental reprogramming when mutS HOMOLOG1 is suppressed in plants. *Plant Physiol* **159**, 710-720.
- Yagi Y, et Shiina T (2014). Recent advances in the study of chloroplast gene expression and its evolution. *Front Plant Sci* **5**.
- Yagi Y, Ishizaki Y, Nakahira Y, Tozawa Y, et Shiina T (2012). Eukaryotic-type plastid nucleoid protein pTAC3 is essential for transcription by the bacterial-type plastid RNA polymerase. *Proc Natl Acad Sci U S A* **109**, 7541-7546.
- Yan J, Yao Y, Hong S, Yang Y, Shen C, Zhang Q, Zhang D, Zou T, et Yin P (2019). Delineation of pentatricopeptide repeat codes for target RNA prediction. *Nucleic Acids Res* **47**, 3728-3738.
- Yang H, Li Q, Fan J, Holloman WK, et Pavletich NP (2005). The BRCA2 homologue Brh2 nucleates RAD51 filament formation at a dsDNA--ssDNA junction. *Nature* **433**, 653-657.
- Yasui A, Yajima H, Kobayashi T, Eker AP, et Oikawa A (1992). Mitochondrial DNA repair by photolyase. *Mutat Res* **273**, 231-236.
- Yi D, Alvim Kamei CL, Cools T, Vanderauwera S, Takahashi N, Okushima Y, Eekhout T, Yoshiyama KO, Larkin J, Van den Daele H, Conklin P, Britt A, Umeda M, et De Veylder L (2014). The Arabidopsis SIAMESE-RELATED cyclin-dependent kinase inhibitors SMR5 and SMR7 regulate the DNA damage checkpoint in response to reactive oxygen species. *Plant Cell* **26**, 296-309.
- Yin P, Li Q, Yan C, Liu Y, Liu J, Yu F, Wang Z, Long J, He J, Wang HW, Wang J, Zhu JK, Shi Y, et Yan N (2013). Structural basis for the modular recognition of single-stranded RNA by PPR proteins. *Nature* **504**, 168-171.
- Yoshiyama KO (2016). SOG1: a master regulator of the DNA damage response in plants. *Genes Genet Syst* **90**, 209-216.
- Yoshiyama KO, Kobayashi J, Ogita N, Ueda M, Kimura S, Maki H, et Umeda M (2013). ATM-mediated phosphorylation of SOG1 is essential for the DNA damage response in Arabidopsis. *EMBO Rep* **14**, 817-822.
- Yu Q-B, Huang C, et Yang Z-N (2014). Nuclear-encoded factors associated with the chloroplast transcription machinery of higher plants. *Front Plant Sci* **5**.

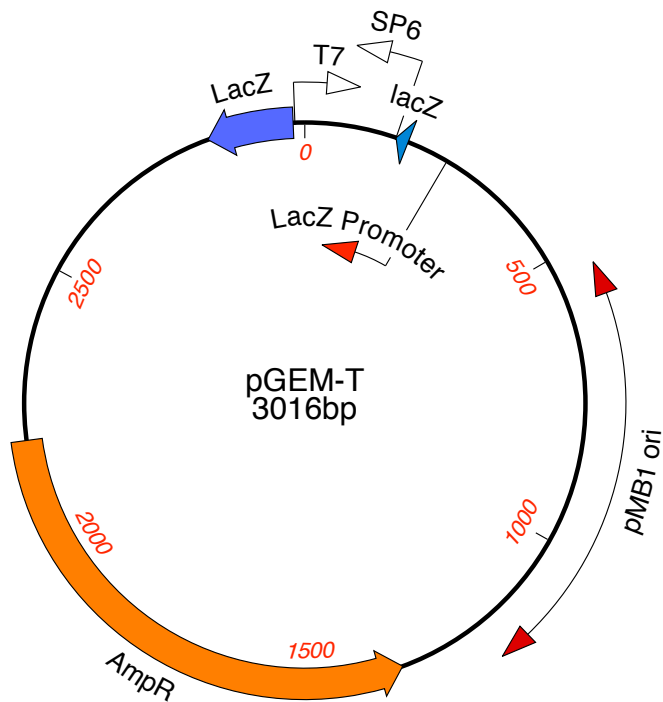


- Zaegel V, Guermann B, Le Ret M, Andres C, Meyer D, Erhardt M, Canaday J, Gualberto JM, et Imbault P (2006). The plant-specific ssDNA binding protein OSB1 is involved in the stoichiometric transmission of mitochondrial DNA in Arabidopsis. *Plant Cell* **18**, 3548-3563.
- Zampini E, Lepage E, Tremblay-Belzile S, Truche S, et Brisson N (2015). Organelle DNA rearrangement mapping reveals U-turn-like inversions as a major source of genomic instability in Arabidopsis and humans. *Genome Res* **25**, 645-654.
- Zhao W, Steinfeld JB, Liang F, Chen X, Maranon DG, Jian Ma C, Kwon Y, Rao T, Wang W, Sheng C, Song X, Deng Y, Jimenez-Sainz J, Lu L, Jensen RB, Xiong Y, Kupfer GM, Wiese C, Greene EC, et Sung P (2017). BRCA1-BARD1 promotes RAD51-mediated homologous DNA pairing. *Nature* **550**, 360-365.
- Zhelyazkova P, Sharma CM, Forstner KU, Liere K, Vogel J, et Borner T (2012a). The Primary Transcriptome of Barley Chloroplasts: Numerous Noncoding RNAs and the Dominating Role of the Plastid-Encoded RNA Polymerase. *Plant Cell* **24**, 123-136.
- Zhelyazkova P, Sharma CM, Forstner KU, Liere K, Vogel J, et Borner T (2012b). The primary transcriptome of barley chloroplasts: numerous noncoding RNAs and the dominating role of the plastid-encoded RNA polymerase. *Plant Cell* **24**, 123-136.

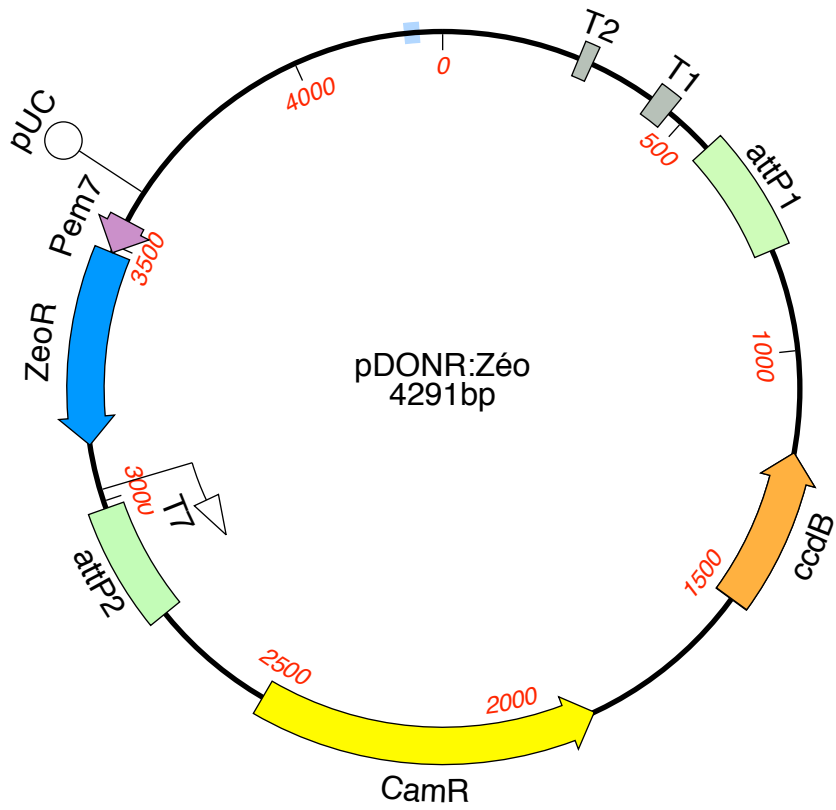
# Annexes



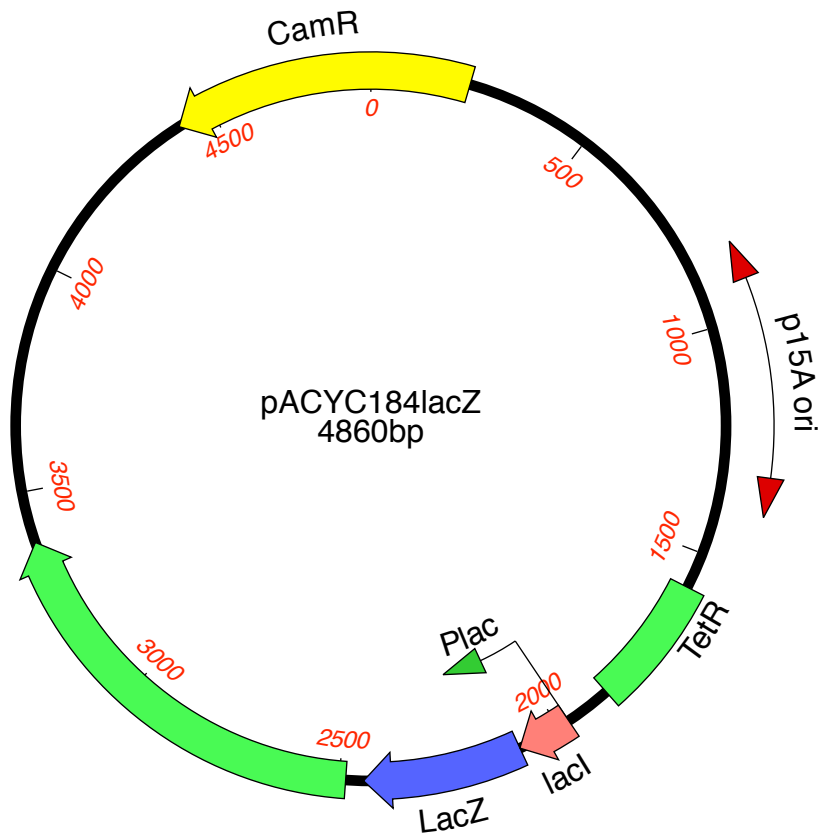
Annexe 1 : Carte plasmidique du pUCAP-GFP.



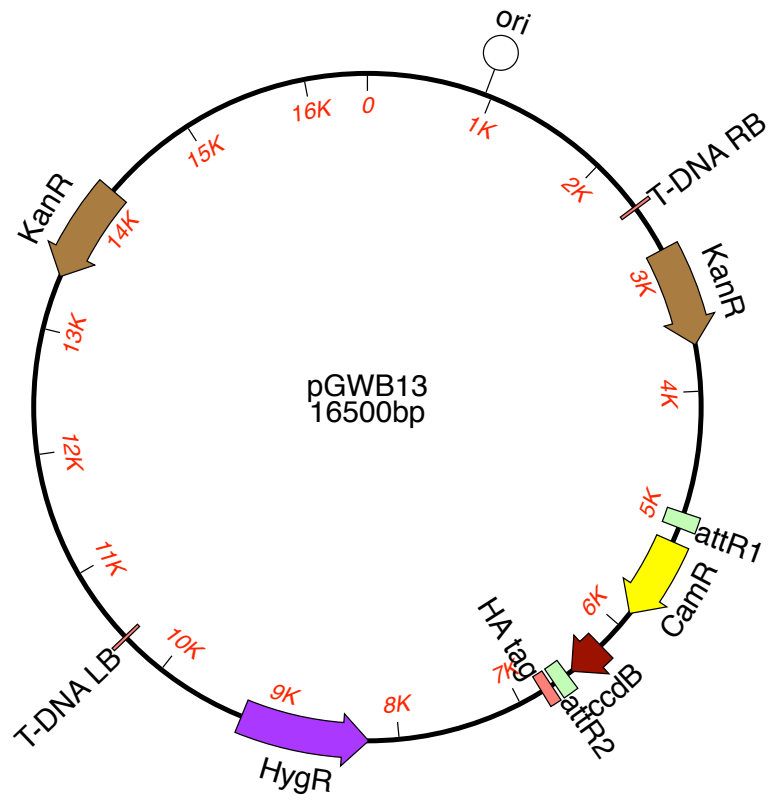
Annexe 2 : Carte plasmidique du pGEM-T easy.



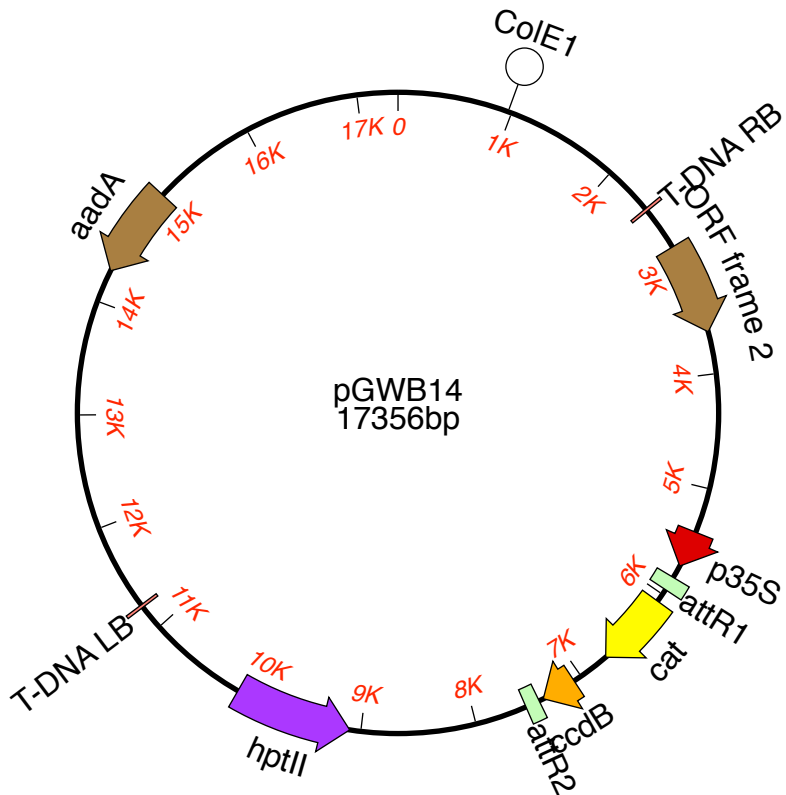
Annexe 3: Carte plasmidique du pDONR/Zeo.

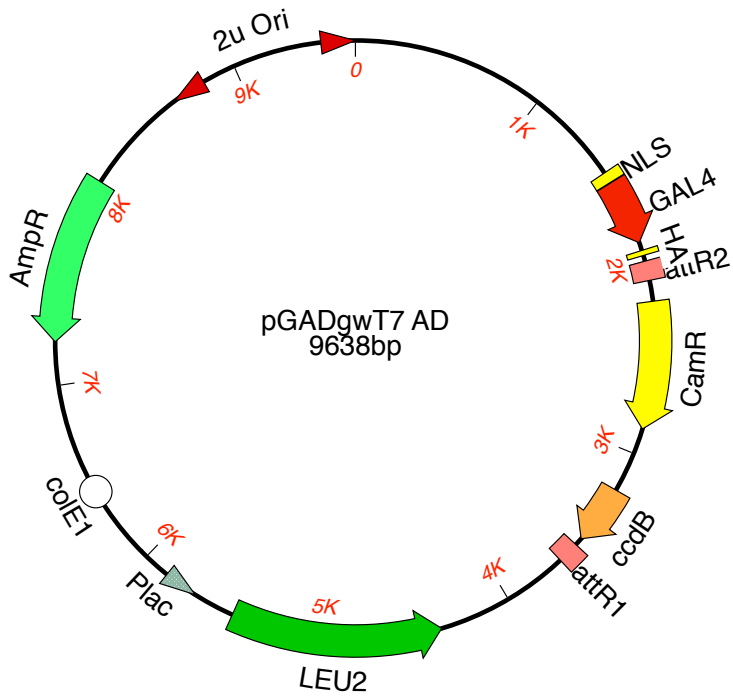


Annexe 7 : Carte plasmidique du pACYC184LacZ.

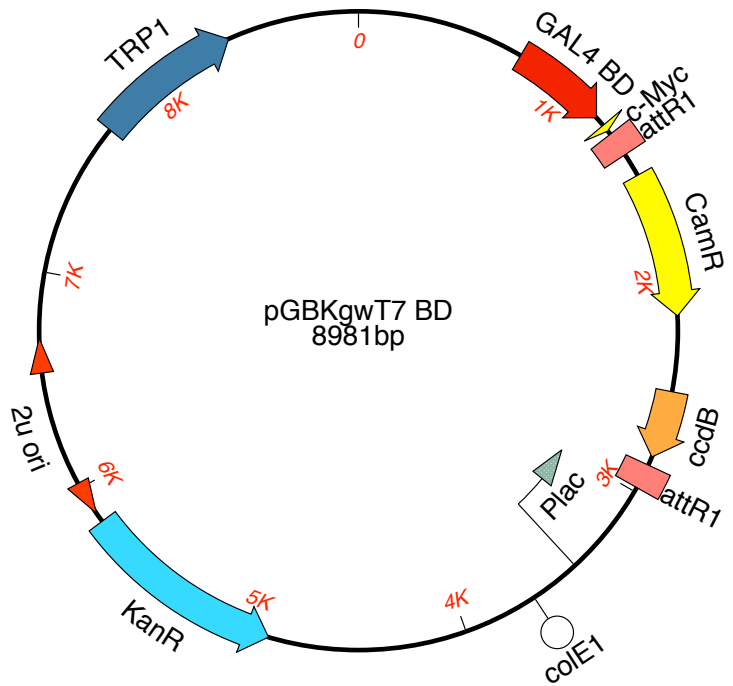


Annexe 4 : Cartes plasmidiques des pGWB13 et pGWB14.

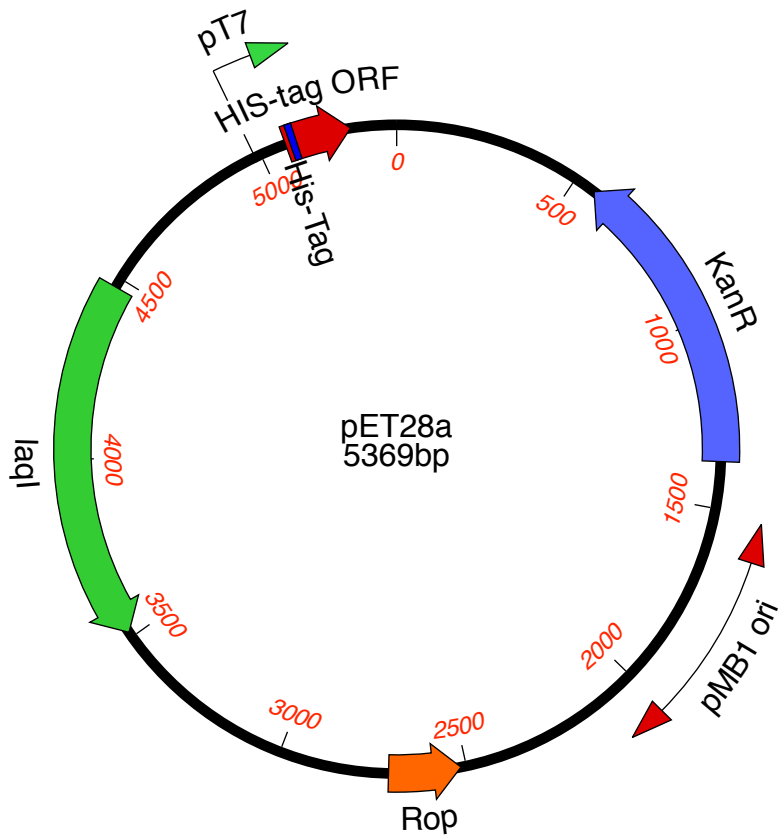




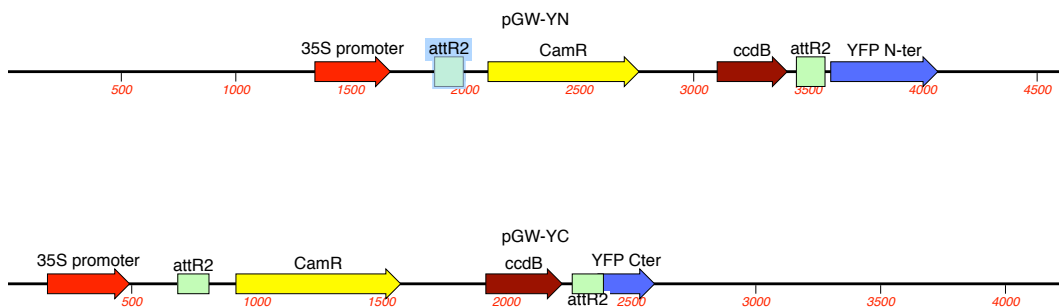
Annexe 5 : Carte plasmidique du pGADT7.



Annexe 6 : Carte plasmidique du PGBKT7.



Annexe 9 : Carte plasmidique du pET28A.



Annexe 10 : Cartes des régions séquencées des pGW-YN et pGW-YC.

## Annexe 11 : Liste des amorces utilisées pour le PCR

Complémentation bimoléculaire de la fluorescence		
Nom	Séquence	Fonction
TRCFgWF	GGGGACAAGTTTGTACAAAAAAGCAGGCTCAAACATGATGTCTTTGCTCCCAAATCC	Clonage de TRCF
TRCFgWR	GGGGACCACTTTGTACAAGAAAGCTGGGTCAATTTGATAAGAGCAGGAAG	Clonage de TRCF
CpTSgWF	GGGGACAAGTTTGTACAAAAAAGCAGGCTCAAACATGGCTTCCTCTATGCTC	Clonage de SSU
RPOAgWR	GGGGACCACTTTGTACAAGAAAGCTGGGTCTTTTTTTCTAGAATGTCTAAT	Clonage de SSU:RPOA
CpTS-RPOAF	TGAGACTCTCTTACCTTCTGACCTTACCGATTCCGAAGTTCGAGAGAAAGTCAAAGTATC	Clonage de SSU:RPOA
RPOA-CpTSR	ACTGTAGTGTCCGAGTAGATACTTTGACTTTCTCTCGAAGTTCGGAATCGGTAAGGTCAG	Clonage de SSU:RPOA
RPOBgWR	GGGGACCACTTTGTACAAGAAAGCTGGGTCAACTTCCTTCTATTAATCTGG	Clonage de SSU:RPOB
CpTS-RPOBF	TGAGACTCTCTTACCTTCTGACCTTACCGATTCCGAAGTTCGGGATGAAAAAGAGGG	Clonage de SSU:RPOB
RPOB-CpTSR	ATCCGGGTATTGCAGATGTTCCCTCTTTTTCATCCCAAGTTCGGAATCGGTAAGGTCAG	Clonage de SSU:RPOB

Double hybride de levure		
Nom	Séquence	Fonction
TRCF R	TGTGGATCCGATTAAGTTGGGTAACGCCAGG	Clonage de TRCF
EcoCarDF	TGTCATATGCTTGGGGATGAAAAAGAGG	Clonage de TRCF et de RID
Eco-CarDR	TGTGGATCCCAGACAGACGAGACAGAGTACG	Clonage de RID
NdeRpoBF	TGTCATATGCTTGGGGATGAAAAAGAGG	Clonage de RPOB
XhoRpoBR	TGTCTCGAGCTTCTATTAATCTGGAAATTC	Clonage de RPOB
XmaRpoTpF	ATCCCGGGACAAGACCCACCATGGATCTCTG	Clonage de RPOTp
XhoRpoTpR	TCTCTCGAGTCTTCAGATGGTTCCTGCTG	Clonage de RPOTp
RPOAAttBF	GGGGACAAGTTTGTACAAAAAAGCAGGCTCTGTTTCGAGAGAAAGTCAAAGTA	Clonage gateway de RPOA
RPOAAttBR	GGGGACCACTTTGTACAAGAAAGCTGGGTCCGACTGGAAAATAGAAATTACG	Clonage gateway de RPOA
RPOtmGWF	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGCAGCTATACGCCAAAGAGA	Clonage gateway de RPOTm
RPOtmGWR	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCAGTTGAAGAAGTAAGTCG	Clonage gateway de RPOTm
RPOtmAttBF	GGGGACAAGTTTGTACAAAAAAGCAGGCTCTATCCCGGCCAATTTGGAATGGGTCGTAC	Clonage gateway de RPOTmp
RPOtmAttBR	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCAGTTGAAGAAATAAGGTGA	Clonage gateway de RPOTmp
RPOtmAttBF2	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCCAATTTGGAATGGGTCGTAC	Clonage gateway de RPOTmp
XhoRpoTmpR	ACACCTCGAGACCAACTTCGGTACTCCCAAAC	Clonage GIBSON de RPOTmp
RpoTmpExF1	ACGACGTACCAGATTACGCTTCCAGTGCTCAAACCCCACTG	Clonage GIBSON de RPOTmp
RpoTmpExR1	GGCGCATAATATCAAGAACCCTGGTAGCTATTCTGAATAAAC	Clonage GIBSON de RPOTmp
RpoTmpExF2	TATTCAGGAATAGCTACCAGGGTCTTGATATTATGCGCC	Clonage GIBSON de RPOTmp
RpoTmpExR2	CTGGCCTCCATGGCCATATGTCAGTTGAAGAAATAAGGTGA	Clonage GIBSON de RPOTmp
pGADGIBR	AGCGTAATCTGGTACGTCGTATG	Clonage GIBSON de RPOTmp
pGADGIBF	CATATGGCCATGGAGGCCAGTG	Clonage GIBSON de RPOTmp



Production d'anticorps contre la partie N-terminale de TRCF		
Nom	Séquence	Fonction
CarDF	AGAGGATCCGTTGTCCATAAGAAGGTCGGTATC	Clonage TRCF
CarDR	GTGAAGCTTTAGTCATTCTTTCTGTTTACACCGAAACG	Clonage TRCF

Complémentation des mutants radA		
Nom	Séquence	Fonction
AttB1-RADA35S	GGGGACAAGTTTGTACAAAAAAGCAGGCTCAGGTTCAAATCCTCCACGATG	Complémentation de rada
AttB-RADA-Cter R	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCCCTTAACACAGCATTGATTA	Complémentation de rada

Complémentation des mutants trcf		
Nom	Séquence	Fonction
p2613-AttB2R	GGGGACCACTTTGTACAAGAAAGCTGGGTTACCCAATAGATTCGCAAACAG	Clonage de TRCF-A
TRCF35SF	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCCCTACACTTCTCTTCTA	Clonage de TRCF-C
AttB RecG2 d1	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCATTGATTGTTGTGTTCTGCC	Clonage de TRCF-B et TRCF-C
AttBRG2 Cter R	GGGGACCACTTTGTACAAGAAAGCTGGGTTCATATTTGATAAGAGCAGGAAGT	Clonage de TRCF-A et TRCF-B

## Résumé

La mitochondrie et le chloroplaste possèdent un génome dont l'expression et la maintenance dépendent de facteurs codés par le génome nucléaire. L'ADN mitochondrial des plantes (ADNmt) est caractérisé par d'importantes activités de recombinaison qui modulent sa structure et contribuent à son évolution. Au cours de ma thèse, j'ai caractérisé l'hélicase RADA qui est impliquée dans les étapes tardives de la recombinaison homologue organellaire. RADA possède des fonctions similaires à son homologue bactérien (RadA), mais semble jouer un rôle plus important dans les organelles de plantes.

L'expression du génome chloroplastique (ADNcp) repose sur deux types d'ARN polymérases : la NEP, codés par le génome nucléaire et la PEP, codés par l'ADNcp. Durant ma thèse, j'ai également étudié le rôle de TRCF, homologue du facteur Mfd et adressé aux chloroplastes. Chez les bactéries, Mfd est impliqué dans le dégagement des ARN polymérases bloquées par des lésions et le recrutement de facteurs de réparation. Chez Arabidopsis, TRCF n'est plus impliqué dans la réparation, mais semble jouer un rôle dans la régulation de l'expression des gènes chloroplastiques.

**Mots clés :** ADNmt, recombinaison, migration de branche, RADA, ADNcp régulation de l'expression, PEP, TRCF

## Résumé en anglais

The mitochondria and chloroplasts have their own genomes whose expression and maintenance depend on factors encoded by the nuclear genome. The plant mitochondrial DNA (mtDNA) is characterized by important recombination activities that modulate its structure and contribute to its evolution. During my thesis, I characterized the RADA helicase that is involved in the late stages of organelle homologous recombination. RADA has similar functions to its bacterial homolog (RadA), but seems to play a more important role in plant organelles.

The expression of the chloroplast genome (cpDNA) relies on two types of RNA polymerases: NEP encoded by the nuclear genome and PEP encoded by the cpDNA. During my thesis, I also studied the role of TRCF, an homolog of the Mfd factor which is addressed to chloroplasts. In bacteria, Mfd is involved in the release of stalled RNA polymerases by lesions and the recruitment of repair factors. In Arabidopsis, TRCF is no longer involved in repair but appears to play a role in the regulation of chloroplast gene expression.

**Key words:** mtDNA, recombination, branch migration, RADA, cpDNA expression regulation, PEP, TRCF