L'hélicase RECG1, un facteur-clé dans le maintien et la ségrégation de l'ADN mitochondrial d'Arabidopsis thaliana

Clementine Wallet

▶ **To cite this version:**

Clementine Wallet. L'hélicase RECG1, un facteur-clé dans le maintien et la ségrégation de l'ADN mitochondrial d'Arabidopsis thaliana. Biologie végétale. Université de Strasbourg, 2016. Français. NNT : 2016STRAJ016 . tel-01402317

HAL Id: tel-01402317

https://theses.hal.science/tel-01402317

Submitted on 24 Nov 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTE

Institut de Biologie Moléculaire des Plantes – CNRS – UPR2357

THÈSE DE DOCTORAT

présentée par :

Clémentine WALLET

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

Discipline/ Spécialité : Sciences du vivant/ Aspects Moléculaires et Cellulaires de la Biologie

L'hélicase RECG1, un facteur-clé dans le maintien et la ségrégation de l'ADN mitochondrial d'*Arabidopsis thaliana*

Thèse soutenue le 25 avril 2016 devant le jury composé de:

Pr. FUCHS Robert	(Université d'Aix-Marseille)	Rapporteur externe
Pr. TOUZET Pascal	(Université de Lille)	Rapporteur externe
Dr. SCHREIBER Valérie	(CNRS - Université de Strasbourg)	Examinatrice interne
Pr. BRENNICKE Axel	(Université d'Ulm)	Examinateur externe
Dr. DIETRICH André	(CNRS - Université de Strasbourg)	Directeur de thèse
Dr. GUALBERTO José Manuel	(CNRS - Université de Strasbourg)	Directeur de thèse

Remerciements

Les personnes que je souhaite remercier sont nombreuses et je n'aurais pas la place de toutes les énumérer ici.

Tout d'abord je souhaite remercier les Docteurs Robert Fuchs et Valérie Schreiber ainsi que les Professeurs Pascal Touzet, et Axel Brennicke pour avoir accepté d'évaluer mon travail de thèse.

Je remercie mes directeurs de thèse José Gualberto et André Dietrich, pour leur patience, et leur confiance. J'ai beaucoup appris à leurs côtés. Merci à José de m'avoir appris les ficelles du métier de chercheur.

Je remercie tous les anciens et les actuels membres du labo 312. Et tout particulièrement Monique Le Ret qui m'a beaucoup aidée pour les manips et Pierre Mercier qui a pris le relais. Merci à Pierre, Nicolas et Khalid pour les nombreuses discussions, scientifiques ou non et pour la bonne ambiance du box. Les bonnes ondes sont indispensables dans un si petit espace. Merci à Déborah Schatz pour le travail fournit durant son stage de Master.

Merci aux membres du département Génex, pour leurs conseils, leurs coups de pouce, leur disponibilité et les moments de détentes passés ensemble. Je n'oublie pas les membres des plateformes de l'IBMP en particulier Malek Alioua de la plateforme de séquençage, Marc Bergdoll de la plateforme de biologie structurale, Nicolas Baumberger de la plateforme de production et purification des protéines ainsi que toute l'équipe de production des plantes de l'IBMP. Je remercie également tous les membres de l'IBMP qui ont pu m'apporter leur aide durant ces années. Je tiens aussi à remercier notre collaborateur Marc Bichara, qui m'a ouvert le monde des bactéries.

Je remercie mes amis, proches ou éloignés et ma famille, pour leur précieux soutien. Merci à mes parents pour leur soutien moral indéfectible et leurs précieux conseils. Merci à ma mère pour sa relecture. Enfin je souhaite remercier Denis pour sa patience et son affection, et pour son grand soutien moral et logistique tout au long de cette thèse.

8-oxoG	8-oxoguanine
ADN	Acide désoxyribonucléique ou DNA
ADP	Adénosine diphosphate
AOX	"Alternative oxydase"
ARN	Acide ribonucléique ou RNA
ATP, dATP	Adénosine triphosphate, désoxyadénosine triphosphate
A_x	Absorbance à x nm
BER	"Base excision repair"
BET	Bromure d'ethidium
BIR	"Break-induced replication"
BSA	"Bovine serum albumine"
CaMV	"Cauliflower mosaic virus"
cDNA	"Complementary DNA" ou ADN complémentaire
Ci, µCi	Curie, microcurie
CMS	"Cytoplasmic male sterility"
CPD	Cyclobutylpyromidine
cpDNA	"Chloroplastic DNA" ou ADN chloroplastique
dCTP	Désoxycytidine triphosphate
dGTP	Désoxyguanine triphosphate
DR	"Direct repair"
DSB	"Double strand break"
DSBR	"Double strand break repair"
dsDNA	"Double-stranded DNA" ou ADN double brin
DTT	Dithiothreitol
dTTT	Désoxythymidine triphosphate
EDTA	Ethylène diamine tétra-acétate de sodium
eGFP	"Enhanced green fluorescent protein"
EGTA	Ethylène glycol tétra-acétate de sodium
g	Accélération centrifuge
g, mg, µg, ng	Gramme, milligramme, microgramme, nanogramme
GGR	"Global genomic repair"
h	heure
His	Histidine
HJ	"Holliday junction"
HR	"Homologous recombination"
IPTG	Isopropyl β-D-1thiogalactopyranoside
IR	"Intermediate-size repeat"
kDa	kilodalton
L, mL, µL	Litre, millilitre, microlitre
LB	Milieu de Luria Bertani
M, mM, µM	Molaire, millimolaire, micromolaire
mA	milliampère
MCS	"Multiple cloning site"
MCS	"Multiple cloning site"

min	minute
MMBIR	"Microhomology mediated break-induced replication"
MMEJ	"Microhomology-mediated end-joining"
MMR	"Mismatch repair"
mtDNA	"Mitochondrial DNA" ADN mitochondrial
NER	"Nucleotide excision repair"
NGS	"Next generation sequencing"
NHEJ	"Non-homologous end joining"
nm	nanomètre
NOS	Nopaline synthase
nt	Nucléotide
NTP, dNTP	Nucléotide triphosphate, desoxynucléotide triphosphate
ORF	"Open reading frame"
OSB	"Organellar single-stranded DNA-binding"
p/v	Poids/volume
pb, kb	Paires de bases, kilobases
PCR, qPCR,	"Polymerase chain reaction", "quantitative PCR",
PFGE	"Pulse-field gel electrophoresis"
Photoproduit 6-4	Dimère pyrimidine pyrimidone (6-4')
RDR	"Recombinaison-dependent replication"
RFP	"Red fluorescent protein"
RID	"RNAP interacting domain"
RMR	"recombination-mediated replication"
RNAP	"RNA polymérase"
ROS	"Reactive oxygenous specie"
RT	"Reverse transcription"
SDSA	"Synthesis-dependent strand annealing"
SSA	"Single strand annealing"
SSB	"Single strand binding"
ssDNA	"Single-stranded DNA" ou ADN simple brin
SSG	"Single-stranded gap"
SSS	"Sub-stoichiometric shifting"
T-DNA	"Transferred DNA" ou ADN de transfert
TCR	"Transcription coupled repair"
TD	"Translocation domain"
TLS	"Trans-lesion synthesis"
tRNA	"transfer RNA" ou ARN de transfert
U	Unité enzymatique
UV	Ultra-violet
V	volts
v/v	Volume/volume

Table des matières

Introduction

1. La réparation de l'ADN 5
1.1. Principaux dommages causés à l'ADN 6
1.1.1. Lésions oxydatives 7
1.1.2. Lésions induites par les UV 8
1.1.3. Alkylations 9
1.1.4. Hydrolyses 9
1.1.5. Erreurs de réplication 9
1.2. Mécanismes de réparation 10
1.2.1. Réversion directe des lésions (DR) 10
1.2.2. Excision de la lésion 11
1.2.2.1. Réparation par excision de base (BER) 11
1.2.2.1.1. Les glycosylases 12
1.2.2.1.2. Les endonucléases AP 14
1.2.2.1.3. Autres facteurs 14
1.2.2.2. Réparation par excision de nucléotide (NER) 15
1.2.2.2.1. Réparation globale du génome (GGR-NER) 16
1.2.2.2.2. Réparation couplée à la transcription (TCR-NER) 17
1.2.2.3. Réparation des mésappariements (MMR) 19
1.2.3. Réparation des cassures double-brin 20
1.2.3.1. Réparation par jonction d'extrémités non-homologues (NHEJ) 21
1.2.3.2. Réparation par recombinaison homologue (HR) 22
1.2.3.2.1. Réparation de cassures double-brin (DSB) ou de lacunes simple-brin (SSG) 22
1.2.3.2.2. Sauvetage des fourches de réplication 25
1.2.3.2.2.1. Ré-initiation de la réplication et régression des fourches. 25
1.2.3.2.2.2. La synthèse trans-lésion (TLS) 27

2. Les hélicases 28
2.1. Les grandes familles d'hélicases 29
2.1.1. SPF1 et SPF2 29
2.1.2. SPF3, SPF4, SPF5 et SPF6 30
2.2. RecG 31
2.2.1. Structure de RecG 31
2.2.2. Rôles de RecG dans *E. coli* 32
2.3. TRCF 34
2.3.1. Sa structure 34
2.3.2. Son rôle dans *E. coli* 34

3. L'ADN mitochondrial de plantes 35
3.1. Structure de l'ADN mitochondrial 36
3.2. Évolution et ségrégation du mtDNA 38
3.3. Les facteurs impliqués dans la recombinaison 40
3.3.1. Recombinases RecA-like 41
3.3.2. Protéines liant le ssDNA 41
3.3.2.1. SSB de type bactérien 42
3.3.2.2. Protéines liant le ssDNA spécifiques aux organelles (OSB) 42
3.3.2.3. Protéines de type Whirly (WHY) 42
3.3.2.4. Homologue à Rad52 (ODB) 43
3.3.3. Homologue à MutS (MSH1) 44
3.3.4. ADN polymérases de type Pol I 44
3.3.5. ADN Hélicases 44
3.3.6. Ligase 45
3.3.7. Topoisomérases 45

3.3.8. Autres protéines 45
3.4. Réplication du mtDNA des plantes 46
3.5. Réparation du mtDNA des plantes 48
3.5.1. Réparation par la voie courte du BER 48
3.5.2. Mécanisme du MMR 49
3.5.3. Réparation par Recombinaison Homologue 49

4. Objectifs de thèse 51

Résultats

1. Import d'ADN et recombinaison *in organello* 54
1.1. Problématique 54
1.2. Prérequis 55
1.3. Tests des différentes sortes de mitochondries 55
1.4. Construction à importer 56
1.5. Test d'import 56
1.6. Analyse de la recombinaison *in organello* par PCR directe ou inverse 57

2. Etude d'une protéine homologue à RecG 59
2.1. Préparation d'anticorps spécifiques à RECG1 60
2.2. Etude de RECG1 (publication Wallet *et al.* 2015, The Plant Cell) 61
2.3. Complémentations du mutant *recG1-2* 62
2.4. Etude du phénotype des plantes bicolores, issues de rétrocroisement entre *recG1-2* et WT 64
2.5. Etude du double mutant *pol1B-2 recG1-1* 65
2.5.1. Phénotype de développement 66
2.5.2. Etude du mtDNA de *pol1B recG1* 67
2.5.3. Etude des fonctions chloroplastiques 67

3. Etude de TRCF 69
3.1. Fabrication d'anticorps 69
3.2. Phylogénie de TRCF 70
3.3. Localisation tissulaire et intracellulaire de TRCF 70
3.4. Caractérisation des mutants 71
3.4.1. Localisation de l'insertion T-DNA 71
3.4.2. Quantification du transcrit de *TRCF* dans les mutants 71
3.5. Etude des rôles de TRCF 72
3.5.1. Sensibilité au stress oxydative: peroxyde d'hydrogène 72
3.5.2. Sensibilité aux UV-C 72
3.5.3. Nombre de copies du mtDNA et du cpDNA dans les plantes stressées 73
3.5.4. Mesure de la recombinaison ectopique 73

Discussion et Perspectives

1. Recombinaison *in organello* 76

2. RECG1 76
2.1. RECG1 et son rôle dans les mitochondries 77
2.1.1. RECG1 a un rôle dans la réparation par HR, mais des pièces sont manquantes 77
2.1.2. RECG1 et son rôle dans la réplication 79
2.1.2.1. RECG1 et la régression de fourches de réplication 79
2.1.2.2. RECG1 et l'inhibition de la réplication 80
2.1.3. RECG1 et son rôle dans la ségrégation 81
2.2. RECG1 et son rôle dans les chloroplastes 83

3. Le rôle de TRCF dans les organelles 84
3.1. TRCF et le NER dans les organelles 84
3.2. Protéines partenaires 85
3.3. Rôle de TRCF dans les chloroplastes 85

Matériel et Méthodes

1. Matériel **89**
1.1. Matériel bactérien 89
1.2. Matériel végétal 90
1.3. Vecteurs plasmidiques 91
1.4. Gènes de synthèse 95
1.5. Oligonucléotides 95
1.6. Outils informatiques 95

2. Méthodes **96**
2.1. Bactéries 96
2.1.1. Préparation de bactéries compétentes 96
2.1.1.1. Electro-compétentes 96
2.1.1.2. Chimio-compétentes 97
2.1.2. Tests de complémentation de RECG1 dans *E. coli* 97
2.1.2.1. Tests de sensibilité aux UV 97
2.1.2.2. Tests de réplication 98
2.2. Acides nucléiques 98
2.2.1. Quantification 98
2.2.2. Analyse par électrophorèse 99
2.3. Méthodes relatives à l'ARN 99
2.3.1. Extraction d'ARN total 99
2.3.2. Traitement DNase 100
2.3.3. Réaction de transcription inverse (RT) 100
2.4. Méthodes relatives à l'ADN 101
2.4.1. Amplification par PCR 101
2.4.2. Techniques de clonage 101
2.4.2.1. Digestion 101
2.4.2.2. Ligation 101
2.4.2.2.1. Ligation d'un fragment dans un vecteur linéarisé 101
2.4.2.2.2. A-tailing et ligation avec pGEM-T easy 102
2.4.2.3. Purification de l'ADN 102
2.4.2.3.1. Sur colonne de silice 102
2.4.2.3.2. Sur gel d'agarose 103
2.4.2.4. Réaction BP et LR 103
2.4.3. Transformation de bactéries 104
2.4.3.1. Transformation par électroporation 104
2.4.3.2. Transformation par choc thermique 104
2.4.4. Préparation d'ADN plasmidique 104
2.4.5. Séquençage 105
2.4.6. Extraction d'ADN total 105
2.4.7. Réaction de PCR quantitative 106
2.4.8. Détection d'ADN par hybridation avec une sonde radioactive (Southern blot) 106
2.4.8.1. Synthèse de la sonde radioactive 106
2.4.8.2. Transfert d'ADN sur membrane et hybridation 107
2.5. Méthodes relatives aux protéines 108
2.5.1. Extraction de protéines à partir de matériel végétal 108
2.5.2. Précipitation (purification et concentration) 108
2.5.3. Séparation des protéines en conditions dénaturantes : SDS-PAGE 109
2.5.4. Quantification 110
2.5.5. Surproduction dans *E. coli* 110
2.5.6. Transfert semi-sec sur membrane de PVDF 111
2.5.7. Immunodétection 111
2.6. Méthodes relatives aux Plantes 112
2.6.1. Transformation d'Arabidopsis par *Agrobacterium tumefaciens* 112
2.6.1.1. Transformation par « floral-dip » 112

2.6.1.2. Sélection des transformants sur boîte 113
2.6.2. Croisements entre deux plantes 114
2.6.3. Transfection par méthode biolistique 114
2.6.4. Coloration GUS 115
2.6.5. Test de respiration 115
2.6.6. Culture de plantules *in vitro* en milieu solide 116
2.6.7. Tests de sensibilité aux stress génotoxiques 116
2.6.7.1. Exposition au peroxyde d'hydrogène 116
2.6.7.2. Exposition aux UV-C 116
2.7. Méthodes relatives aux organelles 117
2.7.1. Extraction de mitochondries 117
2.7.1.1. A partir de fleurs, de rosettes, ou de plantules *in vitro* d'Arabidopsis 117
2.7.1.2. A partir de protoplastes de cultures cellulaires d'Arabidopsis 118
2.7.2. Tests de contrôle respiratoire des mitochondries 118
2.7.3. Tests d'import d'ADN dans les mitochondries isolées 119
2.7.3.1. Marquage de l'ADN à importer 119
2.7.3.2. Import et recombinaison *in organello* 119

Références **121**

Annexe **137**

Introduction

L'introduction qui va suivre présente les mécanismes de réparation de l'ADN chez les êtres vivants, l'implication des hélicases dans ces processus, et les connaissances sur ces mécanismes dans les mitochondries de plantes. La première partie va donc s'attacher à décrire les différents mécanismes de réparation de l'ADN, ainsi que les facteurs impliqués chez les bactéries, dont les mitochondries sont évolutivement proches, et chez les eucaryotes. La deuxième partie aborde le fonctionnement plus particulier de la famille des hélicases, protéines très importantes auxquelles je me suis intéressée dans mon travail de recherche. Enfin la troisième partie traite de l'ADN des organelles qui est l'objet plus particulier de mon travail. Cette synthèse est suivie de l'exposé des objectifs de mon travail de doctorat.

1. La réparation de l'ADN

L'ADN, molécule universelle qui porte l'information génétique dans les cellules des organismes vivants a été découvert il y a plus de 60 ans. Il est indispensable au fonctionnement des cellules et constitue l'identité de chaque être vivant. L'ADN est notamment à la base de la transmission et de l'évolution des caractères. La connaissance des mécanismes cellulaires touchant son fonctionnement est fondamentale, non seulement pour comprendre le vivant, mais également pour soigner les maladies humaines, ou améliorer les végétaux.

L'ADN, molécule complexe, constituée de nombreux composants agencés en double-hélice peut subir des dommages plus ou moins importants tout au long de la vie de la cellule qui le contient. Or, les changements qui en découlent peuvent avoir des conséquences très graves pour la cellule, voire pour l'organisme entier. En effet, des altérations chimiques, des mutations ou des cassures peuvent entraîner une perte ou une modification de l'information génétique et avoir des effets dramatiques aussi bien à court terme (mort cellulaire) qu'à plus long terme (maladies, dégradation de l'intégrité de l'organisme et/ou de sa descendance, processus vitaux altérés, etc…).

Plusieurs causes d'altération de l'ADN sont connues : l'une concerne les erreurs de réplication faites par la machinerie moléculaire, erreurs qui peuvent entraîner des mutations dans la molécule d'ADN, l'autre est relative aux stress. En effet, les organismes vivants sont sans cesse exposés à des stress physiques ou chimiques qui endommagent les composants

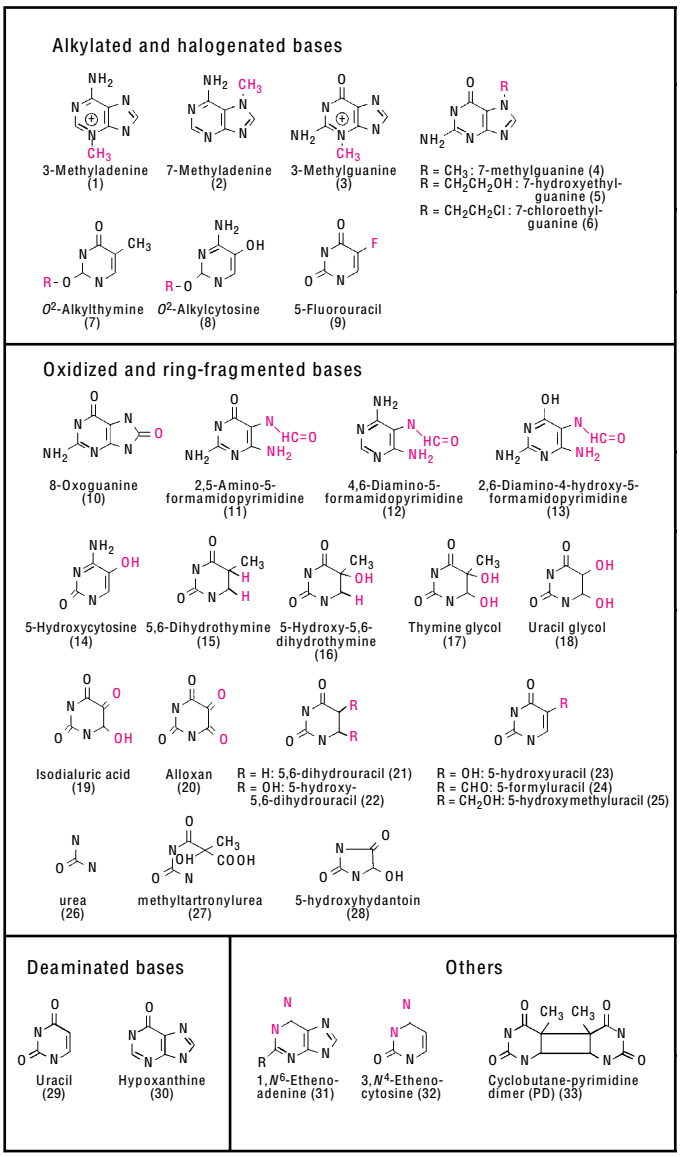


Figure 1. Représentation des différents dommages des bases.
D'après Krokan *et al*., 1997

cellulaires. Les stress peuvent être endogènes, ils proviennent alors de la cellule qui, par son métabolisme et son fonctionnement, produit des molécules pouvant réagir avec les composants cellulaires et les endommager. Par exemple, l'activité mitochondriale, par l'action de la chaîne respiratoire, est une source de radicaux libres qui peuvent altérer les molécules biologiques. Ces stress peuvent aussi être exogènes et provenir de l'environnement extérieur. C'est le cas de nombreuses molécules dont certaines, comme l'eau qui est impliquée dans les réactions d'hydrolyse, sont paradoxalement indispensables à la vie. Le cas des plantes est particulier car elles sont fixées et elles ne peuvent échapper au stress par la fuite. Elles sont donc constamment exposées aux rayonnements ultra-violets (UV) émis par le soleil et qui sont à la fois une source importante de stress et une source d'énergie indispensable pour la croissance. En effet, la lumière est utilisée par la réaction de photosynthèse, qui transforme l'énergie lumineuse en ATP, molécule énergétique.

Au cours de l'évolution, chez tous les êtres vivants, des mécanismes efficaces de réparation de l'ADN ont été sélectionnés, afin de préserver une information génétique fonctionnelle. Cette machinerie moléculaire est complexe, et ceci est particulièrement vrai chez les plantes, organismes eucaryotes dont les cellules renferment trois génomes. En effet, en plus de celui du noyau, les végétaux ont deux génomes organellaires, celui des mitochondries et celui des chloroplastes. Ces deux génomes dérivent de ceux des endosymbiotes procaryotiques qui sont les ancêtres des organelles.

Dans cette première partie, j'expose les lésions pouvant être observées dans l'ADN, ainsi que les principaux mécanismes de réparation et leurs facteurs impliqués chez *E. coli* et chez les eucaryotes.

1.1. Principaux dommages causés à l'ADN

Les dommages causés à l'ADN ont des origines diverses. Les UV sont une origine majeure de lésions, mais les agents oxydants et les agents génotoxiques sont aussi à la source de nombreuses modifications de l'ADN (figure 1). Les principaux dommages peuvent être classés en trois catégories et conditionnent parfois la voie de réparation utilisée :

1. Les bases chimiquement modifiées créent des distorsions dans le duplex d'ADN et bloquent généralement la progression des complexes de réplication et de transcription.

2. Les mésappariements de bases peuvent entraîner des distorsions mais aussi et surtout des mutations.
3. Enfin les cassures double-brin sont une dégradation grave car elles interrompent la continuité de l'information génétique.

1.1.1. Lésions oxydatives

Les lésions oxydatives sont causées par des espèces réactives de l'oxygène (ROS pour "reactive oxygen species"). Dans les cellules eucaryotes, les principales sources de ROS sont les mitochondries, où se déroule la respiration (Gill et Tuteja, 2010; Wallace, 1992), et les peroxysomes, où de nombreuses réactions enzymatiques ont lieu (del Rio *et al.*, 2006). Chez les plantes, les chloroplastes sont une source supplémentaire de ROS car la photosynthèse est réalisée dans la membrane des thylakoïdes par transfert d'électrons (Foyer et Shigeoka, 2011). Certains électrons des chaînes de transfert réagissent parfois avec l'oxygène pour former des radicaux libres (anions superoxydes, $O_2^{\bullet -}$). Par dégradation et diverses réactions, ceux-ci peuvent former du peroxyde d'hydrogène (H_2O_2) et des radicaux hydroxyles (•OH). Les peroxysomes et les chloroplastes produisent de l'oxyde nitrique (NO• qui fait aussi partie des ROS (Demple et Harrison, 1994; Gill et Tuteja, 2010; Temple *et al.*, 2005). Les ROS sont également générés suite à une exposition à des stress environnementaux comme les UV-B, les radiations ionisantes, l'ozone, les métaux lourds, ou encore pour les plantes la sécheresse et la chaleur (Britt, 1996; Gill et Tuteja, 2010).

Les ROS peuvent entraîner toute une variété de modifications dans l'ADN. L'oxydation des oses provoque des cassures simple-brin, qui peuvent faciliter l'hydrolyse des bases à proximité de la cassure. Un des dommages les plus courants est l'oxydation de la thymine, donnant du thymine glycol ainsi que de l'urée issue de sa dégradation. L'on trouve aussi l'un des produits d'oxydation ou d'alkylation des purines : la formamidopyrimidine (FAPy). Tous ces produits bloquent le complexe réplicatif (Demple et Harrison, 1994). L'oxydation peut aussi conduire à des produits mutagènes. En effet, par oxydation ou hydrolyse, les cytosines peuvent être désaminées et former des uraciles qui seront mutagènes. Mais le produit le plus nocif est la 8-hydroxyguanine (appelée 8-oxoG). Ce produit est très mutagène car il s'apparie aussi bien avec un A qu'avec un C et il peut tout à fait servir de substrat pour la synthèse d'ADN (Maki et Sekiguchi, 1992).

Divers mécanismes sont en place dans les organelles pour limiter l'impact de ces stress oxydatifs. Chez *E. coli*, l'enzyme MutT dégrade les formes mutagènes du dGTP (8-oxoG) en

nucléotides monophosphate, évitant ainsi leur incorporation dans l'ADN (Maki et Sekiguchi, 1992). Un homologue de MutT existe chez Arabidopsis (Britt, 2002). Dans les plastes, des antioxydants limitent, en temps normal, le rejet de ROS. Dans les mitochondries des plantes, l'oxydase alternative (AOX pour "alternative oxydase") catalyse la réduction de l'O_2 par l'ubiquinone et permet ainsi de diminuer l'accumulation de ROS (Cvetkovska et Vanlerberghe, 2012). De plus, dans les plantes, des enzymes catalysant des réactions de transformation des ROS permettent de contrôler leur concentration. La superoxide dismutase (SOD) catalyse la transformation du superoxyde $O_2^{\bullet -}$ en H_2O_2, et l'ascorbate peroxydase (APX), la glutathion peroxydase (GPX) et une catalase (CAT) catalysent la dégradation de l'H_2O_2 (Apel et Hirt, 2004). Cependant, lors de stress importants, ces mécanismes ne suffisent pas toujours à évacuer les ROS et l'ADN des cellules se trouve exposé à des stress oxydatifs. L'ADN des organelles est d'autant plus exposé qu'il est proche de la source des ROS.

1.1.2. Lésions induites par les UV

Si la lumière émise par le soleil est indispensable à la photosynthèse, les UV qu'elle contient créent des dommages dans les macromolécules des êtres vivants. Les plantes sont d'autant plus exposées à ce stress qu'elles ont besoin de lumière. Les UV-A, B et C ont des longueurs d'onde respectives comprises entre 400 nm et 320 nm, 320 nm et 280 nm, et inférieures à 280 nm. Les UV-B sont décrits comme particulièrement nocifs (Landry *et al.*, 1997) du fait de leur présence élevée dans les rayonnements solaires qui franchissent l'atmosphère. Mais les UV-A et C sont aussi délétères et peuvent provoquer les mêmes dommages (Britt, 1996). Les UV affectent la photosynthèse et en particulier inactivent le photosystème II (PSII). Ils provoquent également une réduction de l'activité de la ribulose bis-phosphate carboxylase (rubisco) et inhibent les réactions de la photosynthèse en général. En réponse à ces agressions, les plantes accumulent des métabolites secondaires de la famille des flavonoïdes, notamment les anthocyanes, qui font office d'écran solaire (Jansen *et al.*, 1998). Chez l'homme, la nocivité des rayons solaires se manifeste par des cancers de la peau, car l'exposition aux UV peut conduire à des modifications importantes de l'ADN dont les deux principales sont la formation de dimères cyclobutylpyrimidine (CPD) et de dimères pyrimidine pyrimidone (6-4') appelés photoproduits 6-4 (Landry *et al.*, 1997). Ces dimères représentent respectivement 75% et 25% des lésions induites par les UV (Mitchell et Nairn, 1989). Ils peuvent être extrêmement délétères car de tels dimères empêchent la transcription et la réplication. De plus, ils forment des distorsions dans la double hélice, ce qui peut

entraîner des cassures double-brin de l'ADN (Britt, 1999). Les UV peuvent aussi provoquer des dommages oxydatifs, ou des pontages ADN-protéine.

1.1.3. Alkylations

Les 4 bases qui entrent dans la composition de l'ADN peuvent être méthylées en présence d'agents alkylants. Certains de ces agents sont utilisés en recherche, en mutagenèse, ainsi que pour étudier les mécanismes de réparation (Wyatt et Pittman, 2006). Les bactéries ont développé plusieurs mécanismes de défense et de réparation de ces lésions, suggérant que d'autres de ces agents sont naturellement présents dans l'environnement et peuvent réagir avec l'ADN des organismes (Vaughan *et al.*, 1991). Ce sont parfois des produits méthylés qui ont été oxydés en radicaux méthyl. Les plus délétères sont la 3-méthyladénine qui bloque la réplication, et la O^6-méthylguanine qui s'apparie avec la thymine et a un fort pouvoir mutagène (Sedgwick et Vaughan, 1991).

1.1.4. Hydrolyses

Le composé majeur de la vie est l'eau. Cependant, elle peut réagir avec l'ADN et entraîner des réactions d'hydrolyse. Si la liaison entre l'ose et la base de l'acide nucléique est hydrolysée, cela donne des sites abasiques qui sont hautement mutagènes. Ces sites abasiques sont également générés lors du mécanisme de réparation par "Base Excision Repair" (BER) et sont donc pris en charge par ce mécanisme de réparation décrit plus bas (I.1.2.1). La cytosine peut subir une désamination par hydrolyse et entraîner la formation d'uracile qui s'apparie avec l'adénine (Lindahl, 1993).

1.1.5. Erreurs de réplication

Une source endogène de dommages à l'ADN est la polymérase (Pol) réplicative qui peut commettre des erreurs. Chez *E. coli*, la Pol III peut répliquer les deux brins d'ADN avec une activité correctrice exonucléase 3'-5' (activité "proof-reading"). Ce qui signifie que, lorsqu'il y a un mésappariement des bases dû à une erreur de polymérisation, la polymérase change de conformation et exerce alors son activité exonucléase pour corriger l'erreur. L'on estime cependant la fréquence d'erreur à 10^{-7} par paire de bases et par réplication (Iyer *et al.*, 2006; Kunkel, 2004). Cette fréquence est à peu près la même chez les eucaryotes, bien que la réplication soit assurée par 3 polymérases : Polα, Polδ et Polε. Polα joue le rôle de primase et ne possède pas d'activité proof-reading, mais Polδ corrige ses erreurs car Polδ et Polε possèdent toutes deux une activité correctrice exonucléase 3'-5' (Johansson et Dixon, 2013).

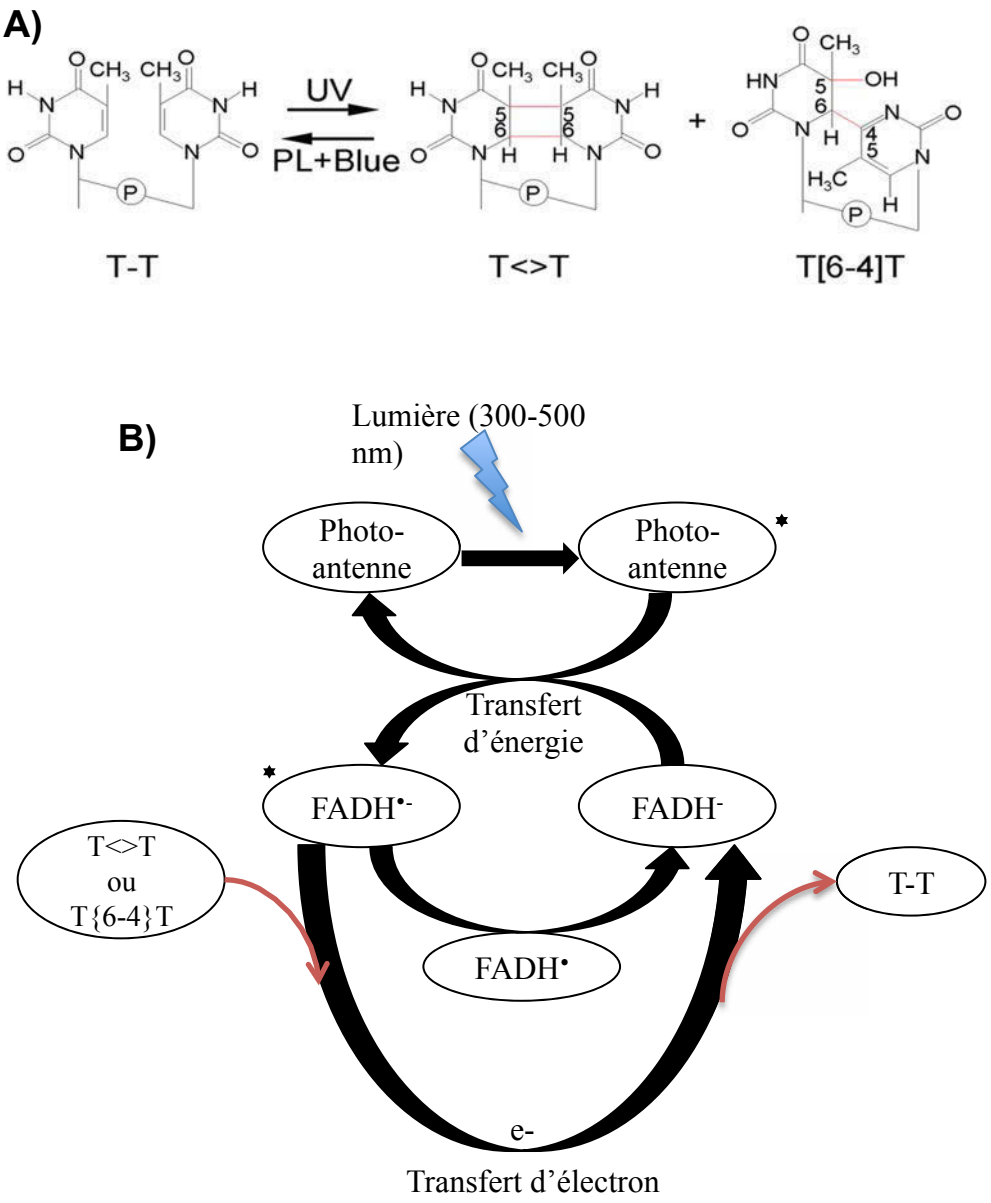


Figure 2. Photoproduits UV et exemple de mécanisme d'une photolyase. A. structure des deux photproduits UV les plus courants: le dimère de thymine T<>T et le photoproduit 6-4 T(6-4)T. PL: photolyase. **B.** Mécanismes de réaction d'une photolyase. Les astérices représentent les éléctrons en excitation. D'après Sancar, 2008

Les erreurs de réplication peuvent entraîner des mutations irréversibles si elles ne sont pas corrigées rapidement (Kunkel et Erie, 2015).

1.2. Mécanismes de réparation

En regard de la diversité et de l'abondance des dommages qui peuvent être causés à l'ADN, des mécanismes de réparation tout aussi diversifiés existent chez les êtres vivants. Dans cette partie, je présente l'ensemble des mécanismes moléculaires impliqués dans la réparation de l'ADN, et plus particulièrement chez *E. coli*. En effet, par leur origine procaryotique, les mitochondries peuvent être rapprochées des bactéries.

1.2.1. Réversion directe des lésions (DR)

Il a été observé depuis longtemps que des bactéries irradiées aux UV peuvent se rétablir si elles sont exposées selon le cas à la lumière bleue ou aux UV-A (320-400 nm) après irradiation (Kelner, 1949). Ce phénomène est appelé photoréactivation et il est dû à l'activité de photolyases, qui rétablissent les bases sans excision de nucléotide par conversion d'énergie lumineuse en énergie chimique (Sancar, 1996). Les photolyases sont composées de deux cofacteurs chromophores photo-transmetteurs. Le premier est un cofacteur catalytique flavine (FAD) qui se lie spécifiquement au site endommagé de l'ADN. Il transmet ses électrons à l'ADN, ce qui a pour effet d'enlever le pontage entre les bases et de transformer les dimères de pyrimidines en monomères. Le deuxième cofacteur porte une photo-antenne, qui est excitée le plus souvent par la lumière bleue ou par les UV-A et qui transmet ensuite ses deux électrons au cofacteur flavine (figure 2) (Brettel et Byrdin, 2010; Sancar, 2003). Chaque photolyase répare spécifiquement un des deux types de lésion UV (CPD ou photoproduit 6-4). Dans les bactéries, seules des photolyases réparant spécifiquement les CPD ont été identifiées (Sancar, 1996).

Chez l'homme, aucune photolyase impliquée dans la réparation des dommages induits par les UV n'a été identifiée (Britt, 1999; Vonarx *et al.*, 1998). Cependant, deux enzymes ayant une homologie avec les récepteurs à la lumière bleue ont été trouvées. Mais elles ne présentent pas d'activité photolyase et semblent être impliquées dans la mise en place du rythme circadien (Sutherland et Sutherland, 1975; Chaves *et al.,* 2011). Ces enzymes n'ayant pas de rôle dans la réparation de l'ADN ont été appelées cryptochromes. La répartition de chacune des photolyases et des cryptochromes est très variée d'une espèce à l'autre.

Les plantes, qui sont plus exposées aux UV que les mammifères, ont gardé ou développé des stratégies de réparation par réversion directe en utilisant la lumière. Chez Arabidopsis, il y a au moins deux types différents de photolyase. L'une (PHR1 ou UVR2) répare spécifiquement les CPD (Ahmad *et al.*, 1997; Landry *et al.*, 1997). L'autre (UVR3) répare les photoproduits 6-4, et son mutant est hypersensible aux UV (Jiang *et al.*, 1997; Nakajima *et al.*, 1998).

D'autres enzymes permettant la réversion d'une lésion par une simple réaction catalytique existent. Dans les bactéries, la protéine OGT (pour "O^6 guanine transferase"), enlève le groupement méthyle du produit O^6-méthylguanine et le transfère sur l'un de ses résidus cystéine. L'équivalent eucaryote (MGMT) a été trouvé chez les plantes (Olsson et Lindahl, 1980). Une autre protéine de *E. coli* appelée AlkB enlève les produits 1-méthyladénine et 3-méthylcytosine par déméthylation oxidative. Il y a deux homologues de cette protéine chez l'homme (ABH2 et ABH3) ayant la même fonction, et l'on trouve un homologue chez Arabidopsis, bien que l'activité ALKB n'ait pas encore été montrée chez les plantes (Sedgwick *et al.*, 2006).

1.2.2. Excision de la lésion

Les lésions que subit l'ADN sont variées et ne se limitent pas aux dimères de pyrimidines. De plus, certaines espèces ne possèdent pas de photolyases. Quand la lésion ne peut être directement réparée, le ou les nucléotides endommagés doivent être remplacés. Ces mécanismes sont essentiels chez tous les êtres vivants.

1.2.2.1. Réparation par excision de base (BER)

Le mécanisme de réparation appelé BER (pour "Base Excision Repair") est initié par l'excision de la base endommagée et non du nucléotide en entier. C'est un mécanisme efficace que l'on trouve au sein de toutes les espèces. Il reconnaît et répare les lésions distordant modérément la double hélice de l'ADN. Son efficacité est basée sur la diversité des lésions que peuvent reconnaître les enzymes principales impliquées dans ce processus : les ADN glycosylases. Il en existe en effet un grand nombre dans le monde du vivant, et chacune d'entre elles reconnaît spécifiquement un type de lésion (Krokan *et al.*, 1997). Ce sont toutes des petites protéines monomériques, et elles peuvent se diviser en deux catégories fonctionnelles selon qu'elles possèdent ou non une activité supplémentaire lyase.

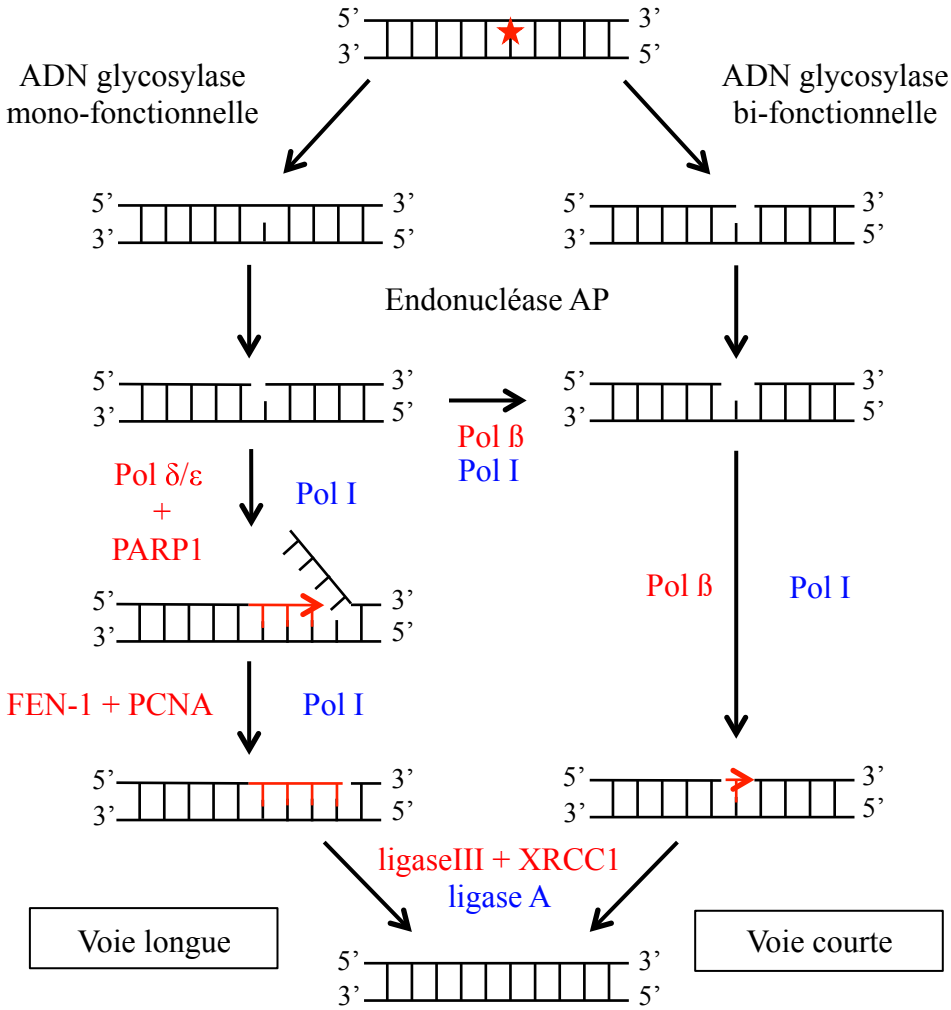


Figure 3. Etapes de la voie BER. Les glycosylases agissent sur une lésion et laissent un site abasique. Les glycosylases bifonctionnelles créent une coupure en 3’. Une endonucléase AP clive ensuite la liaison en 5’ de la lésion. Dans la voie longue à gauche, la polymérase synthétise plusieurs nucléotides et enlève les anciens. Dans la voie courte, à droite, la polymérase synthétise uniquement le nucléotide manquant. En rouge les protéines spécifiques aux eucaryotes, en bleu, aux bactéries, et en noir, les protéines communes aux deux. Adapté de Fromme et Verdine, 2004.!

Le mécanisme est le suivant : une glycosylase excise la base endommagée et laisse un site apurinique/apyrimidique (AP). Deux types de glycosylase existent. Les glycosylases monofonctionnelles se contentent de cliver la liaison entre la base et l'ose, en laissant un site AP, alors que les glycosylases bifonctionnelles ont aussi une activité lyase. Elles transfèrent la base lésée sur une amine, et par une cascade de réactions entraînent le clivage du désoxyribose abasique en 3' (Dodson et Lloyd, 2002). Un site abasique est lui-même potentiellement mutagène, il est donc pris en charge par la suite par une AP endonucléase qui clive le désoxyribose abasique en 5'. Si la base a été excisée par une glycosylase monofonctionelle, le clivage par l'endonucléase donnera un produit 5'-desoxyribose phosphate (5'-dRP). Si au contraire elle a été traitée par une glycosylase bifonctionnelle, le clivage en 5' va permettre l'excision du nucléotide en entier et laisser un vide d'un seul nucléotide (figure 3) (Fromme et Verdine, 2004).

A partir de là, deux voies peuvent se mettre en place. Ces mécanismes ont été découverts chez l'homme et les facteurs humains sont donc utilisés pour décrire le fonctionnement du BER. Par la voie courte ("short patch"), un seul nucléotide est remplacé. C'est l'ADN polymérase impliquée dans la réparation (Pol I pour *E. coli* et Pol ß pour les mammifères) qui, grâce à son activité lyase, enlève le 5'-dRP restant et qui polymérise le nucléotide manquant (Allinson *et al.*, 2001). Le nucléotide est ensuite soudé grâce à une ADN ligase, la ligase III et XRCC1 chez les mammifères (Cappelli *et al.*, 1997). Par la voie longue ("long-patch"), les deux Pol δ/ε initient la polymérisation de 5' en 3', et se déplacent le long du brin complémentaire sur 2 à 15 nucléotides (Frosina *et al.*, 1996). La partie simple brin libre contenant le 5'-dRP (appelée "5'-flap") (figure 3) est ensuite excisée grâce à l'exonucléase FEN1 chez les mammifères (Klungland et Lindahl, 1997). Ce mécanisme requiert également l'aide de la protéine PCNA (pour "proliferating cell nuclear antigen"). Ce facteur fait partie de la famille des "clamps" glissants dont le rôle est d'augmenter la processivité des ADN polymérases. La voie longue est une voie également utilisée lors de la réparation associée à la réplication car il arrive que celle-ci mène à des lacunes simple brin (SSG pour "single-strand gap"). La voie courte est privilégiée chez les mammifères, mais les deux voies existent chez les eucaryotes et les bactéries.

1.2.2.1.1. ***Les glycosylases***

Dans le règne du vivant il existe une grande diversité de glycosylases, reflétant la diversité des lésions pouvant affecter l'ADN.

La famille de glycosylases la plus répandue et qui a été largement étudiée est responsable de la détection et de l'excision des bases uraciles insérées dans l'ADN. Elles sont simplement appelées uracile-ADN glycosylases (UDG pour "Uracyl-DNA glycosylase") (Krokan *et al.*, 1997). Il en existe 5 familles différentes selon leur spécificité. Les glycosylases appelées UNG, humaines et bactériennes, détectent les uraciles issus de la désamination d'une cytosine (Slupphaug *et al.*, 1995; Varshney *et al.*, 1988). Chez les plantes, la présence du BER ainsi que celle d'une uracile-ADN glycosylase ont d'abord été montrées chez la carotte (Talpaert-Borle et Liuzzi, 1982), mais il a été confirmé qu'Arabidopsis possède une UNG pouvant détecter les uraciles.

Chez les bactéries et les mammifères, l'on trouve des glycosylases détectant les mésappariements U/G, T/G, mais aussi l'appariement de thymines avec le produit d'alkylation O^6-methylguanine. Elles sont appelées MUG (*E. coli*) ou TDG (animaux), mais aucun équivalent n'a encore été trouvé chez les plantes (Schormann *et al.*, 2014).

Chez *E. coli*, de nombreuses glycosylases réparent les lésions issues de stress oxydatif, et le BER semble être le premier mécanisme responsable de la réparation de tels dommages. EndoIII (endonucléase III) excise les bases pyrimidiques oxydées. Codée par le gène *nth*, c'est une ADN glycosylase-lyase qui excise les thymines glycol et d'autres bases oxydées (Demple et Linn, 1980). Un homologue de ce gène existe et a été caractérisé chez Arabidopsis (Roldan-Arjona *et al.*, 2000), ainsi que chez les mammifères, où les lésions FAPy sont également reconnues par cette glycosylase. Deux protéines en partie redondantes se chargent de l'excision de 8-oxoG. La protéine MutY détecte les mésappariements dus à la présence de 8-oxoG qui s'apparie avec l'adénine. Elle est aussi décrite dans la détection des mésappariements A/G ou A/C (Nghiem *et al.*, 1988). Les mammifères et les plantes ont un homologue de la protéine MutY, mais pas la levure (Morita *et al.*, 2010). La protéine MutM (ou Fpg) d'*E. coli* répare les bases puriques oxydées. Elle reconnaît les FAPy et les 8-oxoG (Tchou *et al.*, 1991). Chez la levure et les mammifères, l'on ne trouve pas d'homologue de MutM, mais une protéine différente nommée OGG1 a la même fonction (Gao et Murphy, 2001). Chez les plantes, l'on trouve les homologues de MutM et de OGG1 (Dany et Tissier, 2001; Garcia-Ortiz *et al.*, 2001; Macovei *et al.*, 2011). La présence de deux protéines redondantes chez les plantes pourrait suggérer que l'une d'entre elles est adressée aux organelles.

Chez *E. coli*, deux glycosylases sont impliquées dans la réparation des bases alkylées. Tag prend en charge les 3-methyladénines (3-meA) et les 3-méthylguanines (3-meG) (Steinum et Seeberg, 1986). AlkA a une grande diversité de substrat, à commencer par les purines et les pyrimidines endommagées. Elle reconnaît également un grand nombre de produits d'alkylation ou d'oxydation (Nakabeppu *et al.*, 1984). Chez la levure, l'on trouve seulement un homologue d'AlkA, alors que chez les mammifères, une autre glycosylase, AAG, se charge de l'excision de toute une série de substrats issus de l'alkylation de bases. Chez Arabidopsis, l'on trouve 6 homologues du gène *tag*, et il se pourrait que certaines des protéines correspondantes soient adressées aux organelles. L'on trouve également deux protéines homologues à AlkA (Britt, 2002), ainsi qu'un homologue de la protéine AAG des mammifères (AMAG) (Santerre et Britt, 1994).

1.2.2.1.2. ***Les endonucléases AP***

Chez *E. coli*, il en existe deux : ExoIII (pour exonucléase III) qui a une activité de proof-reading 3'-5', en plus de son activité AP endonucléase, et EndoIV (pour endonucléase IV) dont la synthèse est stimulée par les stress oxydatifs.

Chez les mammifères, l'on trouve deux endonucléases. La première, appelée APE1, est homologue à ExoIII, et la deuxième, APE2 n'est présente ni dans la levure, ni chez les procaryotes. Chez Arabidopsis, l'on trouve deux endonucléases AP. L'une est similaire à l'APE1 des mammifères, et la deuxième est plutôt proche de l'APE2 (Britt, 2002; Morita *et al.*, 2010).

1.2.2.1.3. ***Autres facteurs***

La Pol I d'*E. coli* possède également une activité de déplacement de brin, ainsi qu'une activité 5'-3' exonucléase qui peut dégrader un 5'-flap. La Pol I intervient donc surement dans l'élongation dans la voie longue du BER. Pol III quant à elle pourrait jouer le rôle de clamp glissant, comme PCNA, dans la voie longue (Morita *et al.*, 2010).

Bien que les mécanismes du BER soient conservés dans les bactéries et les eucaryotes, des facteurs supplémentaires sont présents chez les eucaryotes. C'est le cas des protéines XRCC1 et PARP1. XRCC1 possède un domaine d'interaction N-terminal avec Pol β. Il interagit également avec la ligase III et PARP1, et participe certainement à la régulation des mécanismes de réparation (Cuneo et London, 2010). PARP1 est une poly(ADP-ribose) polymérase qui régule l'activité de Pol β dans la voie longue. Plus PARP1 est elle-même

poly(ADP-ribosyl)ée, plus l'activité polymérase de la voie longue est favorisée (Sukhanova *et al.*, 2010).

De nombreux facteurs impliqués dans le BER ont été identifiés chez Arabidopsis. Les plantes possèdent des glycosylases, des endonucléases AP, mais également un homologue de l'exonucléase FEN1. L'on trouve également XRCC1 ainsi que PARP1. L'on ne trouve pas d'homologue de la polymérase β. Il est suggéré que la polymérase λ qui appartient à la même famille pourrait jouer ce rôle (Kimura et Sakaguchi, 2006). Les plantes ont aussi deux ADN polymérases de type PolI, qui sont celles impliquées dans la réplication des génomes des organelles. Mais ces protéines ne contiennent pas le domaine 5'-3' exonucléase qui pourrait agir sur les 5'-flaps (Cupp et Nielsen, 2013). Cette activité est probablement assurée par une autre protéine.

1.2.2.2. Réparation par excision de nucléotide (NER)

Parmi les autres mécanismes d'excision qui existent, il y a le mécanisme appelé réparation par excision de nucléotide (NER pour "Nucleotide Excision Repair"). Le complexe de réparation est composé de protéines pouvant reconnaître une grande diversité de lésions dans l'ADN. Tandis que le BER prend en charge des lésions ayant un impact léger sur la distorsion de la molécule d'ADN, le NER répare les lésions plus encombrantes comme par exemple les dimères de thymines, mais aussi les pontages ADN-protéine et ADN-ADN (Britt, 1996). Le complexe de réparation des eucaryotes n'est évolutivement pas proche de celui des procaryotes, mais les mécanismes sont les mêmes. La lésion est reconnue par une protéine qui forme un complexe stable ADN-protéine. Puis une réaction ATP-dépendante le convertit en complexe de pré-incision, et le duplex est ouvert pour permettre une meilleure conformation de la nucléase et avoir accès à l'ADN. Enfin, une incision est effectuée sur le brin abimé, d'abord en 3' puis en 5'. L'action couplée d'une polymérase et d'une ligase permet la synthèse du morceau d'ADN manquant (Sancar et Reardon, 2004). Deux voies existent dans le NER. Le "Global Genomic Repair" (GGR) qui scanne et répare les lésions de tout le génome, et le "Transcription-Coupled Repair" (TCR) qui associe la réparation à la transcription et permet une réparation préférentielle des séquences transcrites. Chez les mammifères, où les photolyases n'ont pas de rôle dans la réparation, ce mécanisme semble être primordial pour la réparation de l'ADN. Toutefois, ce mécanisme reste très important chez tous les êtres vivants où il est présent. Bien qu'elles possèdent des photolyases pouvant réparer les lésions UV, les plantes possèdent aussi les enzymes impliquées dans le NER (Kimura et Sakaguchi, 2006).

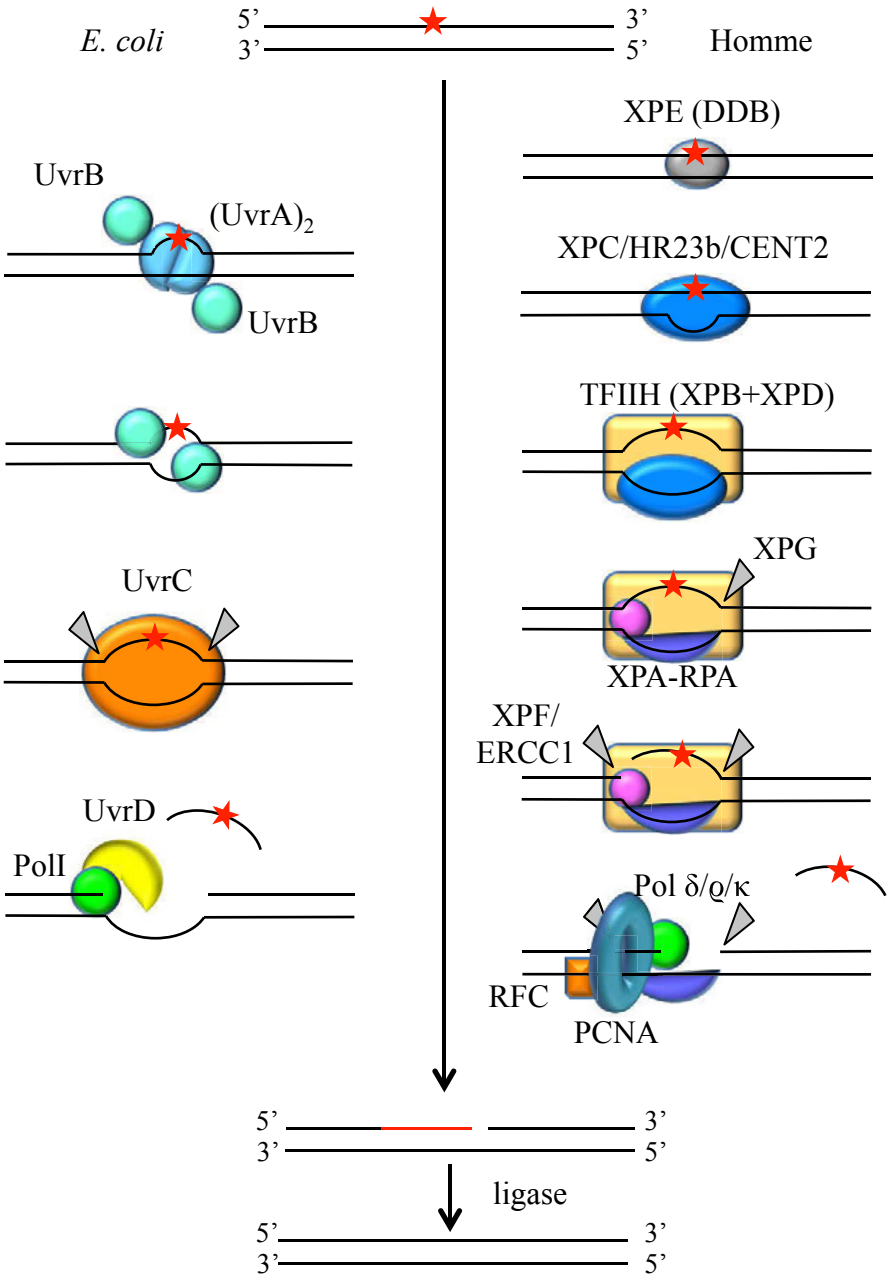


Figure 4. Etapes de la voie NER GGR chez *E. coli* (gauche) et chez l'homme (droite). La lésion est représentée par une étoile rouge. Chez *E. coli*, UvrA repère la lésion et recrute UvrB. UvrB ouvre le duplexe ADN. UvrC clive l'ADN en deux endroits. Puis UvrD enlève le fragment excisé et PolI synthétise la portion manquante. Une ligase soude le nouveau fragment. Chez l'homme, XPE/DDB repère la lésion et recrute XPC/HR23b/CENT2 qui recrute TFIIH. TFIIH ouvre le duplexe pendant que XPC se fixe sur le brin opposé à la lésion. Puis il est remplacé par XPA-RPA qui est recrutés en même temps que XPG. Enfin, XPA recrute XPF/ERCC1 qui clive en premier en 5' de la lésion. XPG clive à son tour et le fragment est détaché en même temps que TFIIH. Enfin, la polymérase δ/ϱ/κ accompagnée de RFC et PCNA synthétise la portion manquante avant qu'elle ne soit religuée. Adapté de Spivak, 2015.!

1.2.2.2.1. ***Réparation globale du génome (GGR-NER)***

Chez les bactéries, ce sont les 3 enzymes UvrA, UvrB, UvrC qui détectent la lésion et prennent en charge l'excision (figure 4). Un dimère d'UvrA et UvrB forme un complexe A_2B_1, qui détecte les lésions de l'ADN (Orren et Sancar, 1989). UvrB, associée à UvrA, possède une activité hélicase qui ouvre la double hélice d'ADN sur 5 pb autour de la lésion et crée une bulle permettant un meilleur accès à l'ADN. Cette réaction, ATP-dépendante, entraîne un changement de conformation d'UvrB, qui se lie de façon stable à l'ADN, et un détachement d'UvrA du complexe (Zou et Van Houten, 1999). UvrC s'associe ensuite avec le complexe et catalyse les incisions permettant l'excision du nucléotide endommagé. UvrC possède un domaine N-terminal similaire à celui de l'endonucléase I-*Tev*I, responsable de la coupure en 3' de la lésion (Verhoeven *et al.*, 2000). Celle-ci précède la coupure en 5' effectuée par le domaine C-terminal EndoV (Lin et Sancar, 1992). Chez les bactéries, le nombre de nucléotides excisés est de 12 ou 13. Une fois l'excision effectuée, UvrC se détache du complexe, tandis que la translocase UvrD (ou helicase II) désolidarise l'oligomère de la molécule d'ADN principale. UvrB est ensuite déplacé par l'ADN polymérase Pol I qui synthétise la portion complémentaire manquante (Orren *et al.*, 1992). Une ligase agit finalement pour souder l'ADN nouvellement synthétisé.

Chez l'homme, de nombreux facteurs impliqués dans le NER ont été identifiés. Plusieurs protéines sont impliquées dans la reconnaissance des lésions. Dans le cas de lésions distordant fortement la double hélice, c'est le facteur XPC, en complexe avec HR23B et CENTRIN2, qui reconnaît la lésion (Araki, 2001). Dans le cas de lésions de type CPD, qui ne déforment pas beaucoup la molécule d'ADN, c'est le facteur XPE (aussi appelé DDB2) en complexe avec DDB1 qui se lie à la lésion et recrute le complexe XPC (Fitch *et al.*, 2003). Le complexe XPC se fixe sur le brin opposé à la lésion, ouvre la double hélice et recrute TFIIH (Lee *et al.*, 2014). Ce facteur est un complexe contenant 10 protéines. Parmi elles, les deux hélicases et ATPases XPB et XPD déroulent l'ADN sur 30 pb et le complexe de pré-incision est alors mis en place. RPA est recruté et se fixe sur l'ADN simple brin (ssDNA) en face de la lésion, ce qui le stabilise et le protège. XPA et XPG sont recrutées et se positionnent respectivement en 5' et en 3' de la lésion. Puis XPA recrute ERCC1-XPF, l'endonucléase qui coupe l'ADN en 5'. XPG coupe enfin l'ADN en 3' de la lésion. Dans les cellules à fort taux de réplication, ce sont ensuite les ADN polymérases Pol δ/ϱ/κ en complexe avec PCNA, RFC et RPA ou l'ADN polymérase Pol ε qui ajoutent les nucléotides manquants, et la ligase I qui soude l'oligomère (figure 4). Dans les cellules quiescentes, ce sont l'ADN polymérase Pol ε avec

PCNA, RFC et RPA, et la ligase IIIα avec XRCC1 qui sont mobilisées pour la synthèse de la portion manquante (Sancar et Reardon, 2004).

Quant aux plantes, leur système de NER semble être proche de celui des mammifères et de la levure. Des homologues des protéines du NER de ces organismes ont été identifiés dans Arabidopsis, ainsi que dans le riz (Kimura et Sakaguchi, 2006) et d'autres plantes. Tous les homologues ne semblent pas impliqués dans le NER ou en tout cas pas dans les mêmes mécanismes que chez les mammifères. Mais il a été montré que cette voie est importante pour la réparation de l'ADN des plantes. Ainsi, l'on trouve un homologue de XPC (ou RAD4 chez la levure) et 4 homologues de HR23B (ou RAD23 chez la levure). Mais aucun des mutants correspondants n'a montré de déficience dans la réparation, signifiant que peut-être un autre processus reconnaît les lésions (Hays, 2002). Un homologue de la protéine CENTRIN2 (CEN2) qui agit en complexe avec XPC et HR23B a été caractérisé chez Arabidopsis. Cette protéine serait impliquée dans le NER mais aussi dans la réparation par recombinaison (Molinier *et al.*, 2004a). Parmi les composants de TFIIH, l'on trouve un homologue de l'hélicase XPD (*AtRad3*) dont la mutation est responsable de la sensibilité aux UV chez le mutant *uvh6* (Liu *et al.*, 2003). Deux homologues de l'hélicase XPB (*AtRad25*) ont été caractérisés dans Arabidopsis. Les deux mutants montrent une sensibilité aux UV mais aussi à l'agent alkylant méthyl-méthanesulfonate (MMS). Ces deux protéines redondantes agissent probablement dans la même voie (Costa *et al.*, 2001; Morgante *et al.*, 2005). Des homologues des endonucléases responsables de l'excision, XPF (*AtRad1* ou *UVH1*), ERCC1 (*At-ERCC1*) et XPG (*AtRad2* ou *UVH3*) ont également été caractérisés dans Arabidopsis (Gallego *et al.*, 2000; Hefner *et al.*, 2003; Liu *et al.*, 2001; Liu *et al.*, 2000).

1.2.2.2.2. ***Réparation couplée à la transcription (TCR-NER)***

Il a été montré depuis longtemps que la transcription stimule la réparation du brin transcrit, et qu'il y a donc un mécanisme couplant la réparation à la transcription (Mellon et Hanawalt, 1989; Mellon *et al.*, 1987). Ce mécanisme ("Transcription-Coupled Repair" ou TCR) est bien compris chez *E. coli*, mais il est plus complexe chez les eucaryotes. Le fait que l'ADN des eucaryotes soit enroulé grâce aux histones dans la chromatine complique son accès et de fait, le ssDNA est beaucoup moins stable que chez la bactérie. C'est sûrement pourquoi l'on observe la présence de protéines supplémentaires impliquées dans la stabilisation du ssDNA.

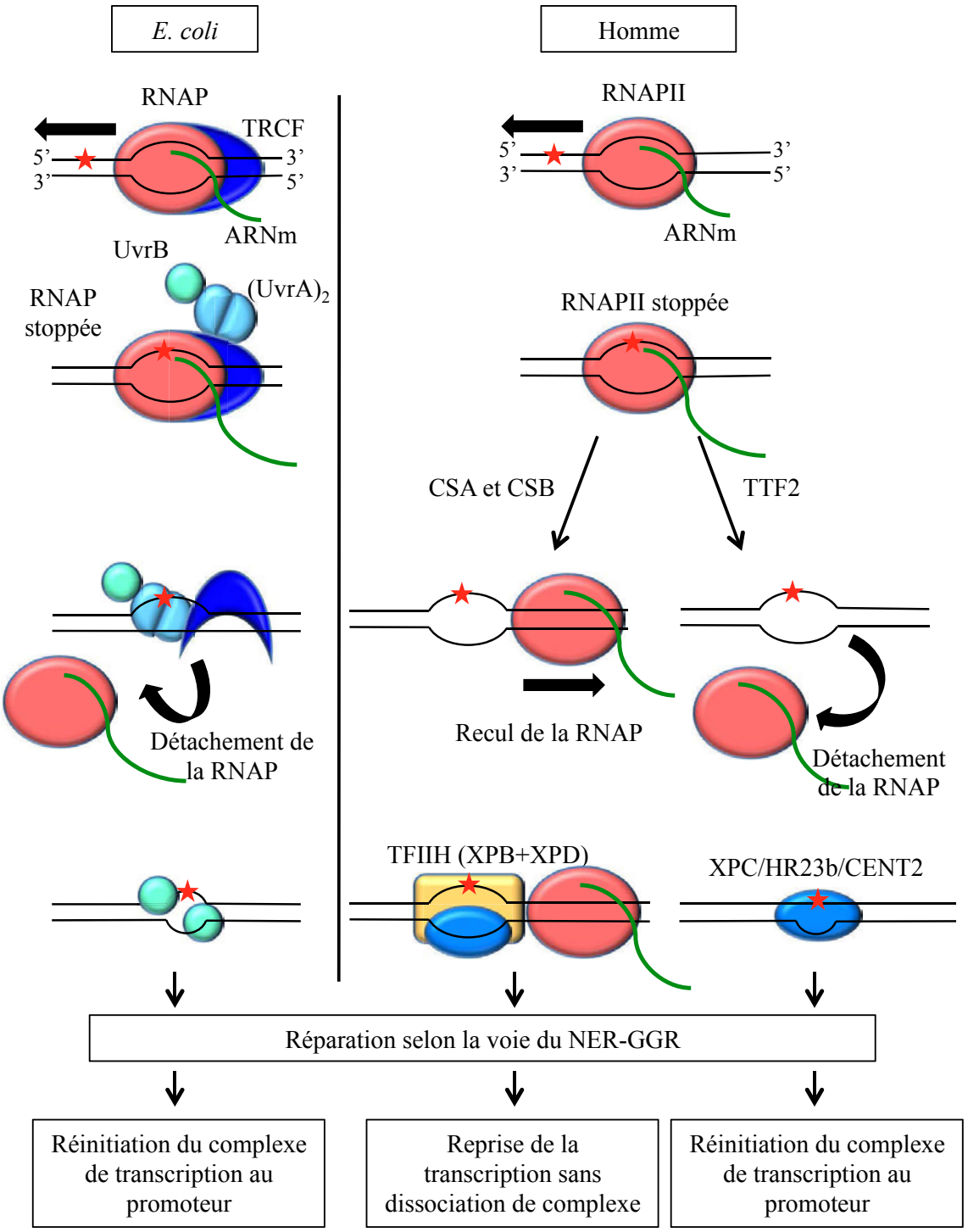


Figure 5. Etapes de la voie TCR-NER chez *E. coli* et chez l'homme. Chez *E. coli*, TRCF/mfd en interaction avec RNAP change de conformation quand celle-ci est bloquée par une lésion (représentée par une étoile rouge) et recrute UvrA. RNAP est décrochée du duplexe et UvrA recrute UvrB qui se positionne sur le duplexe. La réparation peut ensuite s'opérer selon le NER, mais le complexe de transcription doit être ré-initié. Chez l'homme, CSA et CSB permettent le recul de la RNAP et le recrutement direct de TFIIH. TTF2 permet le détachement de la RNAP. La lésion sera prise en charge par la voie du NER GGR qui commence par la fixation de XPC sur la lésion. Dans le cas du recul de RNAP, la transcription pourra continuer sans ré-initiation du complexe.

Chez *E. coli*, le principal acteur de cette voie est le facteur couplant la réparation à la transcription (TRCF pour "transcription-repair coupling factor") codé par le gène *mfd* (pour "mutation frequency decrease") (Selby *et al.*, 1991). C'est une translocase qui possède un domaine de liaison à l'ARN polymérase (RNAP pour "RNA-polymerase") bactérienne. Elle contient un motif doigt de zinc, ainsi qu'un domaine homologue à UvrB ayant une affinité pour UvrA dans le complexe A_2B_1. TRCF agit comme une translocase pour RNAP, et favorise le décrochement de la RNAP de la molécule d'ADN quand l'élongation est stoppée par une lésion importante (Park *et al.*, 2002) (figure 5). TRCF recrute ensuite les enzymes de réparation UvrA et UvrB, puis se détache de la lésion. UvrABC procède à l'excision de la lésion (Selby et Sancar, 1995). C'est donc TRCF qui est responsable de la détection de la lésion, et sa forte affinité pour la RNAP permet une surveillance accrue des gènes hautement transcrits dont la mutation pourrait s'avérer dramatique, même à court-terme. Récemment, une étude de Epshtein (Epshtein *et al.*, 2014) a montré que l'hélicase UvrD aurait aussi un rôle dans la voie TCR. Une voie de réparation alternative est proposée où TRCF n'interviendrait pas. UvrD promouvrait le recul de RNAP, laissant un accès à la lésion au complexe de réparation. A l'inverse de TRCF, UvrD ne détacherait pas RNAP, ce qui permettrait de continuer la transcription après réparation et serait plus efficace d'un point de vue de la transcription (Epshtein, 2015). Le facteur d'élongation de la transcription NusA serait également un autre acteur de la voie TCR bactérienne. Il permettrait le recul de la RNAP au même titre que UvrD. Cependant ces observations doivent encore être confirmées *in vivo* (Cohen et Walker, 2010).

Chez les eucaryotes, le système de réparation associé à la transcription est moins compris. La protéine CSB (pour "cokayne syndrome B") est structuralement proche de TRCF, car elle possède un motif hélicase, elle interagit avec l'ADN et la RNAP et elle hydrolyse l'ATP. Chez l'homme, la réparation couplée à la transcription a été montrée seulement pour les gènes transcrits par l'ARN polymérase II (RNApolII). Chez la levure, ce mécanisme existe aussi en association avec l'ARN polymérase I. Chez l'homme, c'est donc CSB qui a une affinité pour la RNApolII bloquée (Beerens *et al.*, 2005). Cette liaison entraîne un changement de conformation du complexe et un recrutement de CSA, ce qui permet à TFIIH de procéder à la réparation comme décrit précédemment. Dans ce scénario, la RNApolII n'est pas détachée de l'ADN, et CSA favoriserait le recul de la polymérase afin d'accéder à la lésion. Ce mécanisme permettrait de ne pas recommencer la transcription du début. Une autre protéine,

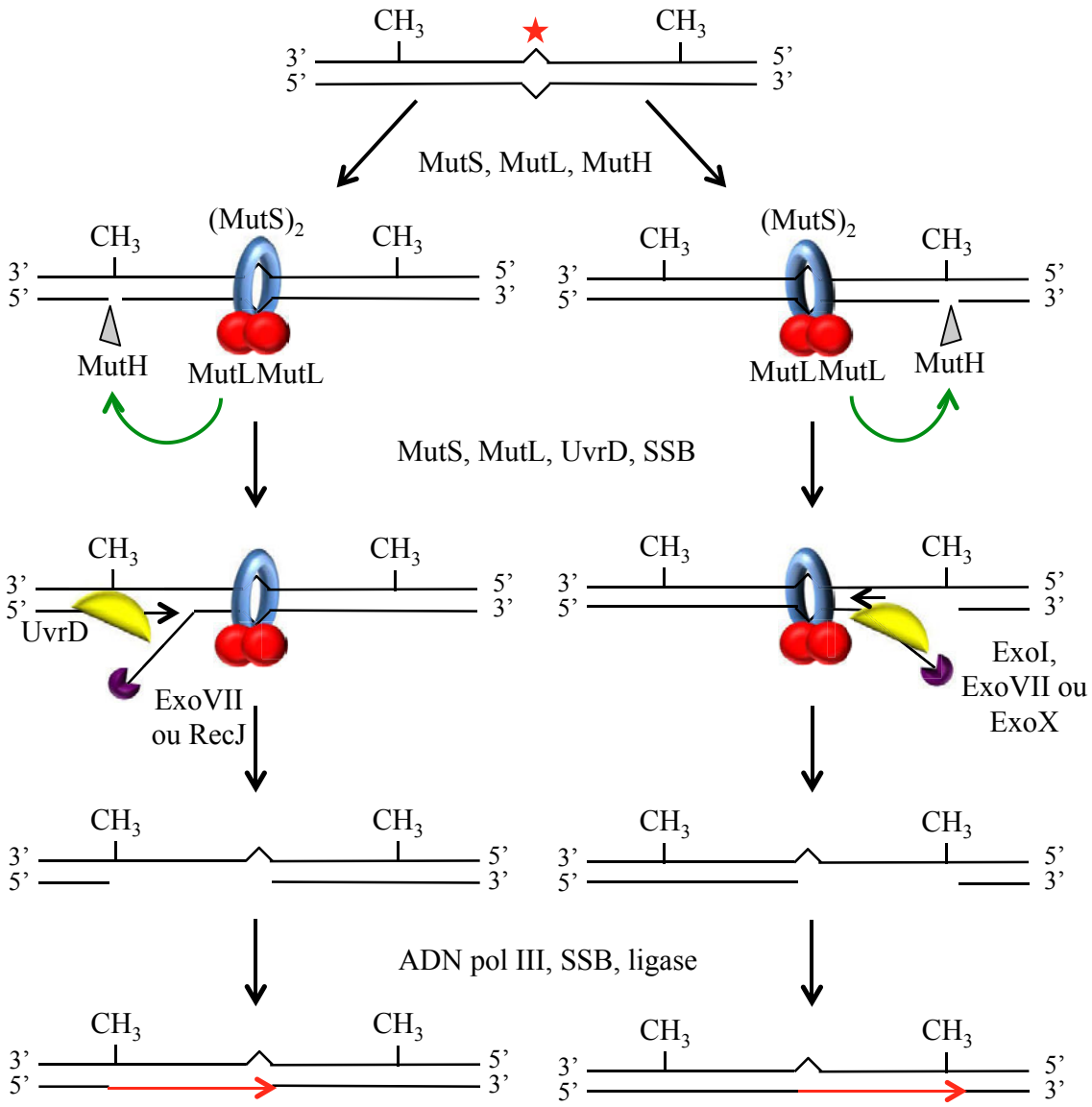


Figure 6. Etapes de la voie MMR chez *E. coli*. Le mésappariement est représentée par une étoile rouge. MutS reconnaît la lésion et recrute MutL. MutH clive alors le brin non-méthylé dans la séquence GATC la plus proche. Si le clivage est effectué en 5' de la lésion, les exonucléases ExoVII ou RecJ dégradent le brin jusqu'au complexe MutS/MutL avec l'aide d'UvrD. Si le clivage est en 3' de la lésion, les exonucléases ExoI, ExoVII ou ExoX prennent en charge la résection. L'ADN polIII synthétise ensuite la portion manquante à l'aide de SSB et le tout est lié par une ligase. Adapté de Iyer *et al*. 2006. !

TTF2, a la capacité de détacher les RNApolI et III (Hara *et al.*, 1999), et pourrait agir comme la TRCF bactérienne en libérant le complexe de transcription (figure 5).

En ce qui concerne la voie TCR chez les plantes, les deux protéines CSA et CSB ont des homologues dans Arabidopsis. Il a été montré que le mutant d'un des homologues de CSB (At5g63950, appelé ERCC6 ou At-Rad26) est sensible aux UV (Shaked *et al.*, 2006). Chez Arabidopsis il n'y a pas d'UvrABC, mais l'on trouve un homologue de TRCF. Cette protéine a une prédiction d'adressage aux organelles et a été l'un des objets d'étude de ma thèse. J'y reviendrai donc par la suite.

1.2.2.3. Réparation des mésappariements (MMR)

Lors d'un mésappariement dû à une erreur de réplication, ou à une oxydation de base, le nucléotide n'est pas reconnu comme une lésion. La voie du "Mismatch Repair" (MMR) permet la reconnaissance et l'excision de mésappariements. Dès lors, le mécanisme de réparation doit identifier laquelle des deux bases du mésappariement est la base mutée. Chez les bactéries, l'ADN nouvellement synthétisé n'est pas tout de suite méthylé. Le système de réparation identifie donc la base sur le brin non-méthylé comme incorrecte et l'excise. C'est la protéine MutS qui reconnaît le mésappariement et se lie d'abord à l'ADN. MutL se fixe ensuite sur ce complexe MutS-ADN et le stabilise afin de faciliter l'action de MutH. Cette dernière reconnaît le brin non-méthylé. Dans *E. coli*, toutes les adénines dans la séquence GATC sont méthylées par la méthylase dam, sauf après le passage de la polymérase pendant la réplication. MutH produit une coupure simple brin du côté non-méthylé (Au *et al.*, 1992). Cela permet à UvrD et à la protéine "single strand-binding" (SSB) qui stabilise le ssDNA de se fixer au complexe. L'excision peut ensuite s'effectuer dans l'un ou l'autre sens selon que MutH ait coupé en 5' ou en 3' de l'erreur (Cooper *et al.*, 1993). Si la coupure est en 5' du mésappariement, MutL recrute les endonucléases RecJ ou ExoVII. A l'inverse, si la coupure est en 3', MutL recrute les endonucléases ExoI, ExoVII ou ExoX (Burdett *et al.*, 2001). UvrD en tant qu'hélicase, déroule le brin d'ADN (figure 6).

Chez les eucaryotes, le système de reconnaissance du brin néoformé est différent. Il a été montré *in vitro* que le système de réparation MMR est dirigé par les lacunes simple-brin (SSG). En effet, le brin retard est préférentiellement réparé *in vitro* (Pavlov *et al.*, 2003), montrant que les interruptions entre les fragments d'Okasaki sont reconnues par le système de réparation. C'est aussi le cas de l'extrémité 3' du brin néoformé. Chez les mammifères, l'on trouve deux complexes agissant en hétérodimère avec une activité semblable à MutS.

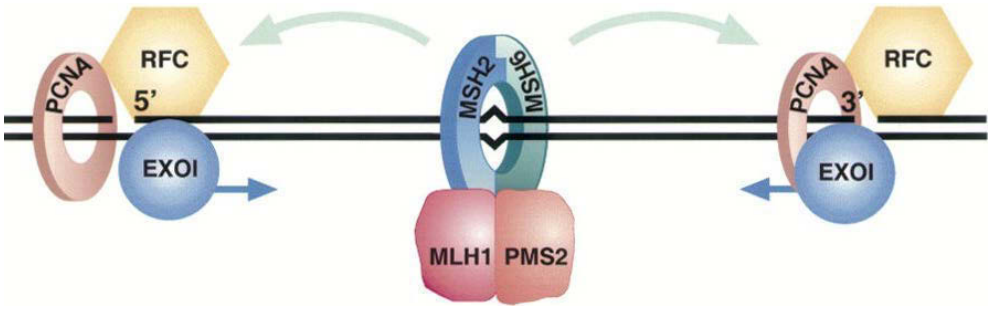


Figure 7. Clivage directionnel dans la voie du MMR chez l'homme. Le déplacement du complexe MutSα/MutLα le long de l'hélice ADN permettrait de signaler le sens de déplacement du complexe d'excision (PCNA+RFC+EXOI) par rapport à la cassure simple brin. PCNA est chargé avec une orientation définie sur l'extrémité 3' de la cassure simple brin. Le complexe mobile MutSα/MutLα rencontre donc différentes orientations de PCNA selon que la cassure est située en 3' ou en 5' du mésappariement. Cette interaction résulte en l'activation différentielle de l'activité hydrolytique 5'-3' ou 3'-5' d'EXOI. D'après Dzantiev *et al*. 2004.

L'hétérodimère composé de MSH2/MSH6, MutSα reconnaît préférentiellement les mésappariements, les insertions et délétions de 1 à 2 nucléotides, alors que MutSβ (MSH2/MSH3) reconnaît des mésappariements comprenant 2 à 10 nucléotides. L'on trouve également 3 hétérodimères homologues à MutL. MutLα (MLH1/PMS2) a la même activité que MutL dans la réparation et agit avec MutSα et MutSβ. Le rôle de MutLβ (MLH1/PMS1) n'a pas été démontré dans la réparation. MutLγ (MLH1/MLH3) semble avoir un rôle partiel dans la réparation des complexes initiés par MutSα (Modrich, 2006). Il n'y a pas d'homologue de MutH chez les mammifères. L'endonucléase impliquée ensuite dans l'excision des nucléotides est ExoI. Il a été montré que cette endonucléase est impliquée dans la réparation des lésions situées en 3' de la coupure mais aussi en 5', alors qu'elle a uniquement une activité exonucléase 5'-3'. Il s'est finalement avéré que ExoI pourrait avoir une activité exonucléase 3'-5' lorsqu'elle est associée en hétérodimère avec d'autres composants du système de réparation. Les protéines PCNA et RFC sont impliquées dans le contrôle de la polarité de l'excision des nucléotides à partir de la coupure (figure 7) (Constantin *et al.*, 2005; Dzantiev *et al.*, 2004; Kadyrov *et al.*, 2006).

Chez Arabidopsis, l'on trouve également MutSα et MutSβ, ainsi que deux homologues de MutL, un homologue de PCNA et de RFC. Il a été montré que ces protéines sont impliquées dans la réparation et un mécanisme MMR proche de celui des mammifères pourrait donc être conservé chez les plantes (Morita *et al.*, 2010). Les plantes possèdent également un homologue de MutS (MSH1) adressé aux mitochondries et aux chloroplastes (Arrieta-Montiel *et al.*, 2009).

1.2.3. Réparation des cassures double-brin

Les cassures double-brin (DSB pour "double-strand break") sont fréquentes et proviennent de lésions non-réparées, qui créent des distorsions dans la double-hélice. Elles sont produites par des agents ionisants, des ROS ou des UV. Une DSB peut aussi apparaître suite à une erreur de fonctionnement de nucléases, mais aussi lorsqu'une fourche de réplication s'arrête. Les SSG peuvent apparaître de manière spontanée, ou peuvent être le résultat d'une lésion dont le processus de réparation (BER) n'a pas été terminé. Les DSB peuvent avoir de lourdes conséquences sur le génome comme des délétions ou des réarrangements. Ce type de lésion entraîne l'activation de mécanismes de réparation par recombinaison homologue (HR pour "homologous recombination") ou par jonction d'extrémités non-homologues (NHEJ pour "non-homologous end joining"). Ces deux voies de

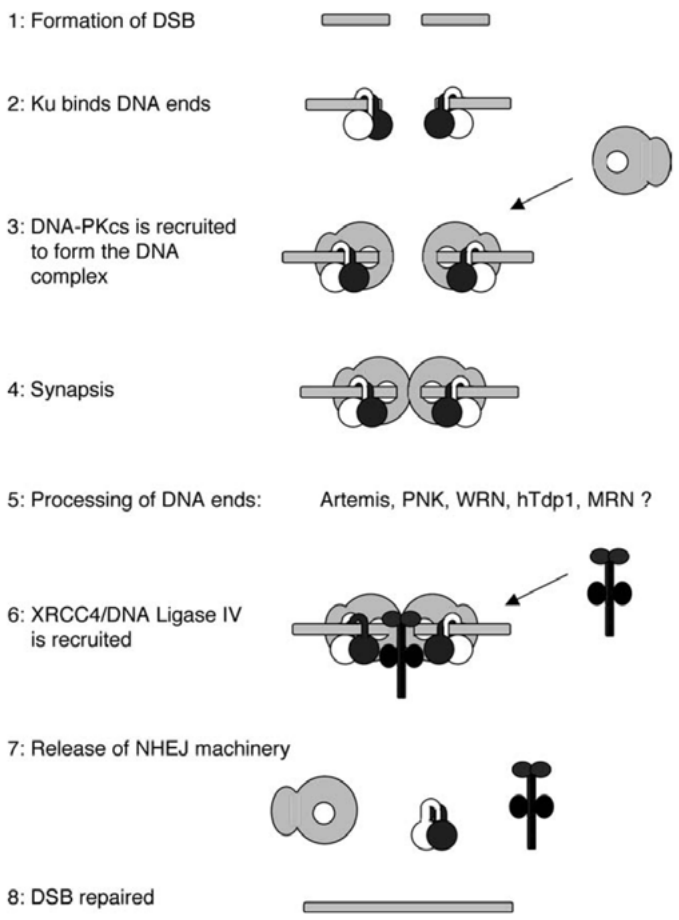


Figure 8. Modèle simplifié pour le NHEJ chez les eucaryotes. 1. Formation de DSB. **2.** L'hétérodimère Ku70/80 reconnait la DSB et se lie aux extrémités de l'ADN. **3.** DNA-PKc est recrutée par Ku liée à l'ADN. **4.** La synapse se met en place quand deux extrémités d'ADN sont rassemblées par les deux molécules DNA-PK. **5.** Les extrémités de l'ADN sont traitées car les DSB produisent rarement des extrémités franches. **6**. Le complexe XRCC4/ligase IV est recruté au niveau du complexe DNA-PK. **7.** Les protéines de la voie du NHEJ se détachent. **8.** La DSB est réparée bien que des nucléotides soient souvent supprimés, donnant des réparation non-fidèles. D'après Lees-Miller et Meek, 2003.

réparation sont une source de diversité, voire de mutations, de délétions ou d'insertions. En effet, dans le cas de la voie NHEJ, deux extrémités coupées et physiquement proches sont reliées ensemble. Mais dans le cas de multiples cassures, il y a des chances pour que les extrémités reliées ne soient pas correspondantes, ce qui donne des insertions ou des délétions. Dans la voie HR, la recombinaison avec une région répétée ou avec une région très proche en séquence présente ailleurs dans le génome peut également entraîner des réarrangements. De plus, les mécanismes de recombinaison passent par des structures d'ADN complexes dont la résolution peut mener à différents produits.

1.2.3.1. Réparation par jonction d'extrémités non-homologues (NHEJ)

Chez les bactéries et les levures, c'est le mécanisme HR qui est privilégié par rapport au NHEJ. Ce dernier est surtout utilisé chez les mammifères. Chez l'homme, un hétérodimère Ku70-Ku80 se lie à la cassure, la protégeant ainsi de toute résection. Le dimère Ku a une faible affinité pour l'ADN simple brin. Puis une protéine kinase ADN-dépendante (DNA-PK pour "DNA-dependent protein kinase") se lie aux extrémités de la cassure. DNA-PK est structuralement proche des deux kinases ATM et ATR. Elles sont toutes trois impliquées dans la réparation des DSB. L'activité kinase de DNA-PK, est déclenchée uniquement lorsqu'elle est assemblée aux DSB. En effet, la phosphorylation de Ku70-Ku80 est indispensable à la réparation par NHEJ et catalyse le rapprochement et la synapse par appariement de 1 à 5 nt des deux fragments d'ADN. XRCC4 en complexe avec la ligase IV joint ensuite les deux bouts et achève la réparation (figure 8). XRCC4 pourrait aussi être phosphorylé par DNA-PK, ce qui aurait un rôle dans la régulation de la réparation (Lees-Miller et Meek, 2003). ATM est également recrutée dans la réparation des DSB mais par le complexe MRN (Mre11/Rad50/Nbs1). ATR est recrutée quand des DSB apparaissent au niveau des fourches de réplication (Caron *et al.*, 2015; Falck *et al.*, 2005). Les voies ATM et ATR ne sont pas formellement considérées comme faisant partie du NHEJ. Ces deux kinases, qui définissent les voies de réparation dites ATM- ou ATR-dépendantes, sont des facteurs-clés des mécanismes de contrôle de l'intégrité de l'ADN pendant le cycle cellulaire. En cas de reconnaissance de lésions, elles sont recrutées et elles phosphorylent des facteurs en aval du processus de réparation. Cela peut entraîner un arrêt du cycle cellulaire et de la division de la cellule, pour donner du temps à la réparation, ou éventuellement déclencher l'apoptose de la cellule.

Chez Arabidopsis, des homologues de Ku70, Ku80, XRCC4 et ligase IV ont été identifiés et la voie NHEJ est effectivement active dans les cellules. Il existe aussi des gènes

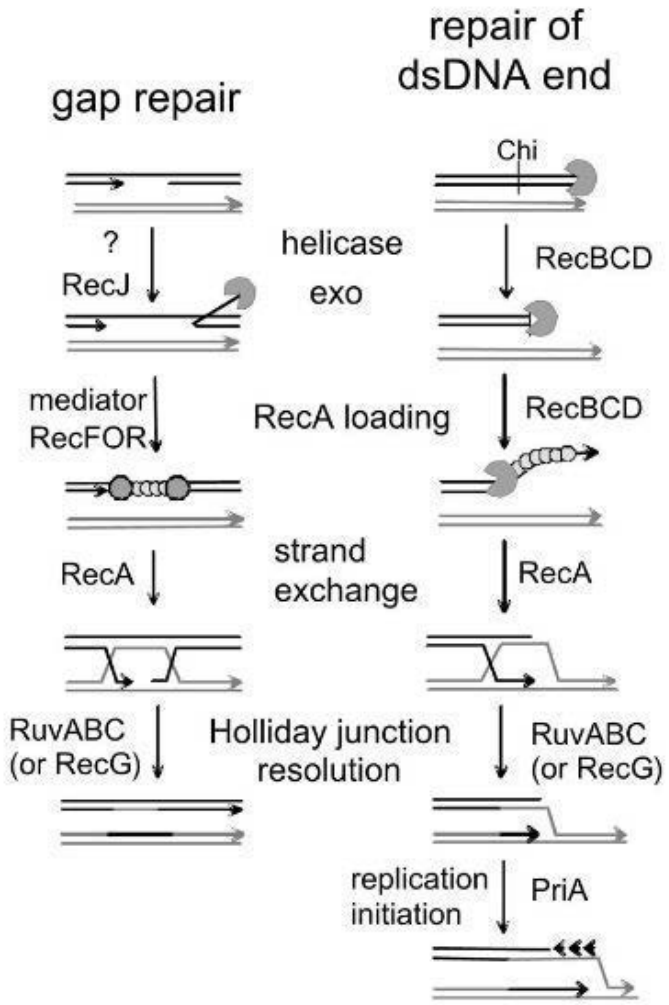


Figure 9. Réparation par HR de cassures simple brin (gap repair) ou double brin (repair of dsDNA end) chez *E. coli*. Les enzymes représentées ont toutes un rôle biochimique connu. Au niveau des cassures simple brin, l'exonucléase RecJ agrandit la région ssDNA, puis RecF, RecO et RecR permettent le chargement de RecA sur le ssDNA stabilisé par SSB. Aux DSB, RecBCD dégrade l'ADN jusqu'à rencontrer un site Chi (χ). Son activité hélicase-nucléase est alors modifiée pour laisser une extrémité 3' et charger RecA. La phase synaptique est catalysée par RecA et produit une HJ. La phase post-synaptique consiste en la migration (synthèse) et la résolution des HJ, catalysé par RuvABC et RecG. D'après *Rocha et al.*, 2005.!

codant pour ATM et ATR, qui sont ubiquitaires chez les eucaryotes. Cependant, il n'y a pas d'orthologue de DNA-PK. Le NHEJ des plantes pourrait donc fonctionner sans kinase, ou avec une kinase différente (Lees-Miller et Meek, 2003). Il a été montré que l'ADN nucléaire des plantes est moins sensible aux DSB que celui des autres organismes. Le NHEJ serait aussi le mécanisme favorisé chez les plantes (Gorbunova et Levy, 1999) et la prédominance de cette voie, qui est considérée comme source d'erreurs ("error-prone"), pourrait avoir une influence sur l'évolution génomique et les capacités d'adaptation des plantes.

1.2.3.2. Réparation par recombinaison homologue (HR)

1.2.3.2.1. ***Réparation de cassures double-brin (DSB) ou de lacunes simple-brin (SSG)***

Cette voie de réparation par recombinaison est prédominante chez les bactéries. Elle comprend 3 phases. Dans la phase pré-synaptique, les deux extrémités des DSB sont prises en charge pour former des extrémités simple-brin sur lesquelles se fixe RecA ou une recombinase équivalente. Pendant la phase synaptique, le filament nucléoprotéique cherche une homologie dans l'ADN et catalyse un échange de brin avec une molécule d'ADN donneur. La formation d'une jonction de Holliday (HJ) (Holliday, 2007), structure en X formée de 4 brins d'ADN associés en 4 duplex, permet la copie d'informations à partir du brin donneur. La phase post-synaptique comprend la résolution des HJ et la génération des produits de réparation (Yeeles et Dillingham, 2010).

Dans le cas de DSB, chaque extrémité de la cassure est excisée de 5' en 3', laissant de chaque côté une extrémité simple brin 3' sortante. RecBCD est un complexe permettant cette résection de brin. RecB contient un domaine hélicase 3'-5' associé à un domaine nucléase. RecC contient un domaine de reconnaissance des séquences Chi (pour "Crossover hotspot instigator"), et RecD contient un domaine hélicase 5'-3'. Chez *E. coli*, la séquence Chi est une séquence de 8 nucléotides où la recombinaison s'initie de manière préférentielle. RecB et RecD progressent sur chacun des brins, tout en dégradant l'ADN, jusqu'à ce que RecC reconnaisse une séquence Chi. Une telle séquence entraîne un changement de conformation du complexe et empêche la dégradation du brin de 3' en 5' par RecB. Cela laisse donc une extrémité 3' sortante sur laquelle est chargée RecA (figure 9) (Wigley, 2013). La présence de nombreuses séquences Chi dans le génome expliquerait pourquoi cette séquence stimule la recombinaison (Kowalczykowski, 2000). RecA est une recombinase essentielle pour la réparation par recombinaison homologue. C'est elle qui permet l'invasion de ssDNA dans une région homologue d'un duplex et son appariement avec le brin complémentaire. Cette

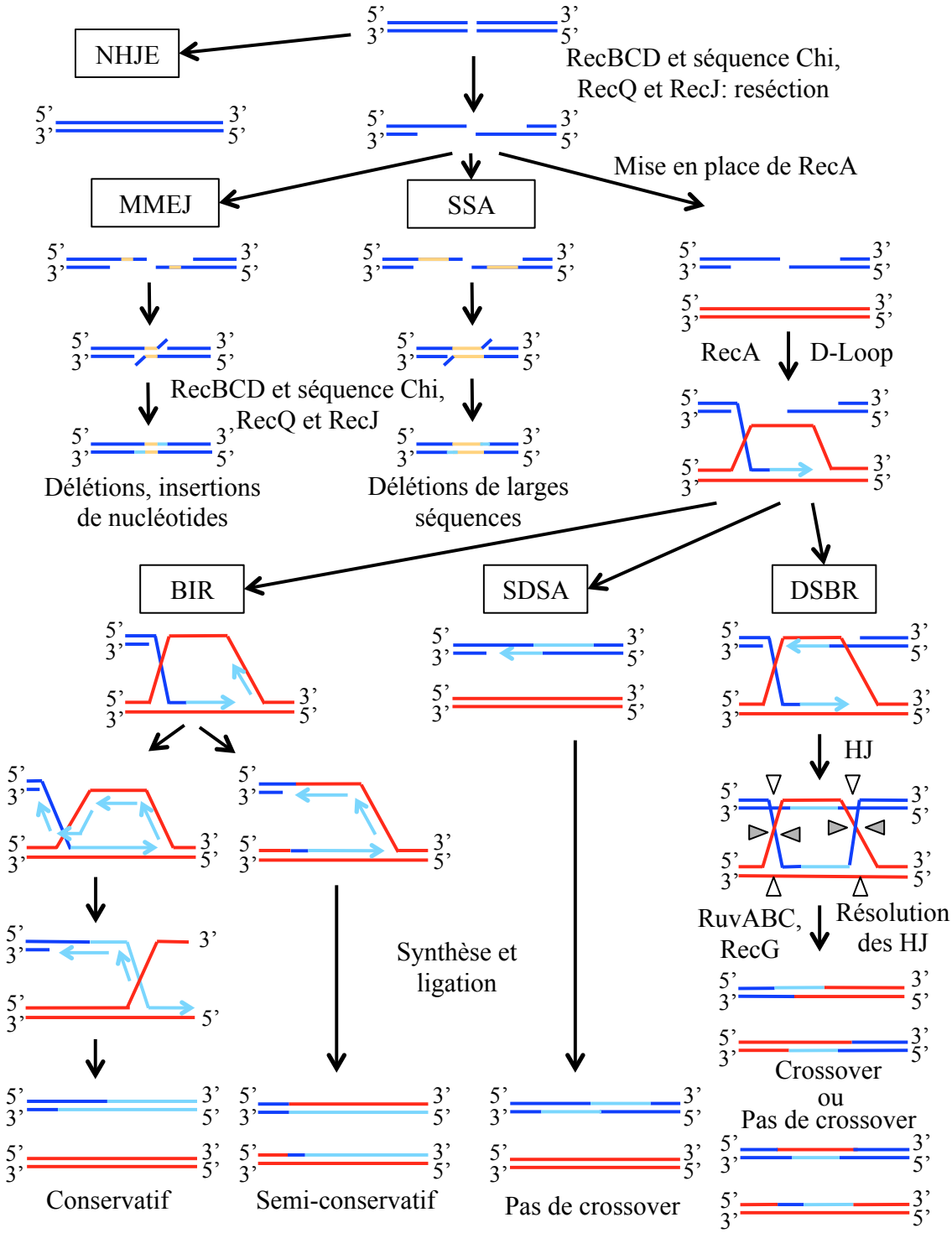


Figure 10. Etapes des voies NHEJ et HR. Les extrémités de la cassure double brin peuvent être directement reliées par la voie du NHEJ. Dans les autres voies, il y a résection des brins, donnant des extrémités 3' sortantes. S'il y a deux séquences complémentaires sur ces extrémités, elles peuvent s'apparier sans l'aide de RecA, entrainant des délétions. Dans la voie des MMEJ, les séquences homologues sont en dessous de 30 pb, contrairement à la voie SSA, où elles sont plus grandes. Quand RecA se met en place sur l'extrémité 3' elle catalyse la recherche de séquences homologues dans le génome et la formation d'une D-loop. Dans le cas où il y a recrutement du deuxième brin (DSBR) les deux HJ mises en places peuvent être résolues dans des sens différents, produisant un crossover (clivage gris), ou non (clivage blanc). Dans le cas du SDSA, le brin néoformé se détache pour permettre la synthèse du complémentaire. Dans le cas du BIR, une fourche de réplication est mise en place et l'ADN est synthétisé jusqu'à la fin de la molécule.

structure d'ADN 3 brins est appelée D-loop et initie souvent des processus de recombinaison ou de réplication.

Le complexe RecFOR peut intervenir dans une voie alternative pour la résection de l'ADN au niveau des cassures et charger RecA sur le ssDNA. Il prend en charge les DSB, mais intervient surtout pour la réparation des SSG. En effet, il permet l'accès de RecA à l'ADN malgré la présence du facteur SSB qui se lie fortement au ssDNA afin de le stabiliser (Morimatsu et Kowalczykowski, 2003). Dans cette voie, l'hélicase RecQ déroule le brin d'ADN pour permettre son accès à RecJ, une 5'-3' exonucléase, et à SSB qui stabilise la partie simple brin sortante. C'est ensuite le complexe RecFOR qui charge RecA sur le ssDNA. RecO enlève SSB et RecR charge RecA (Morimatsu et Kowalczykowski, 2014). Le complexe RecFOR, l'ADN PolI et l'ADN PolII synthétisent ensuite la partie manquante (figure 9).

Les hélicases RecG et RuvAB permettent l'ouverture des deux brins-matrice, et l'appariement de chacun de ces brins avec les brins du duplex endommagé. Cela donne une structure d'ADN de type HJ. La résolvase RuvC permet de résoudre les HJ et de revenir à deux duplex d'ADN distincts (Michel et Leach, 2012). Il a été montré que RusA peut également cliver de façon symétrique les HJ (Mahdi *et al.*, 1996). Cependant cette protéine est codée par un gène cryptique qui n'est pas exprimé en temps normal.

A partir de l'établissement d'une D-loop, différentes voies peuvent se mettre en place et générer des produits finaux plus ou moins réorganisés (figure 10). Dans le cas où il n'y a pas de deuxième extrémité simple brin, la D-loop peut se transformer en fourche de réplication et la voie de "break induced replication" (BIR) se met en place. Soit une fourche de réplication est mise en place et l'ADN est synthétisé jusqu'à la fin du chromosome, donnant un brin néoformé, soit la synthèse du brin retard s'effectue en prenant pour modèle le brin invasif. Dans le cas où une deuxième extrémité simple brin est disponible, deux voies peuvent être empruntées. Dans le "synthesis-dependent strand annealing" (SDSA), la D-loop est dissociée une fois que le brin a été allongé. Le brin néoformé se ré-apparie alors avec la deuxième extrémité simple brin, afin de servir de matrice à sa réparation. Cette voie évite les "crossover" et minimise les réarrangements, contrairement à la voie du "double-strand break repair" (DSBR). Dans cette voie, les brins 3' de chaque côté de la cassure forment chacun une D-loop avec le brin homologue d'un autre duplex d'ADN. Il y a donc formation de deux HJ qui peuvent être résolues de différentes manières et produire ou non un "crossover" selon le sens de clivage des HJ (Heyer *et al.*, 2010).

Enfin, deux mécanismes impliquant la recombinaison homologue qui ne font pas intervenir RecA peuvent exister. En effet, dans le cas ou il y a deux séquences homologues de plus de 5 nt sur les brins 3'-sortants, elles peuvent s'apparier sans l'aide de RecA. Cette configuration peut se rencontrer dans le cas ou deux séquences homologues flanquent une cassure double brin. Dans la voie appelée "single-strand annealing" (SSA), la taille des séquences homologues est supérieure à 30 nt. Elles s'apparient ensemble puis les extrémités 3' sortantes non-homologues sont clivées et une polymérase et une ligase terminent le processus. L'information génétique entre ces deux séquences homologues est généralement perdue (Huertas, 2010). Chez les eucaryotes, une autre voie similaire appelée "microhomology-mediated end joining" (MMEJ) peut être empruntée quand une séquence homologue de 5 à 30 nt se situe sur les extrémités 3'-sortantes (figure 10). Cette voie entraîne souvent des délétions et des insertions de nucléotides (McVey et Lee, 2008).

Chez les eucaryotes, le mécanisme de résection de l'ADN se rapproche de celui de RecFOR. Chez l'Homme et la levure, c'est le complexe MRN/X (Mre11-Rad50-NSB1/Xrs2) qui initie le clivage de l'extrémité 5' à l'aide de CtlP/Sae2. La petite extrémité sortante 3' sert de point de départ pour une résection plus importante. Celle-ci est réalisée par l'exonucléase 5'-3' ExoI ou par le complexe BLM/Sgs1-Top3-Rmi1 associé avec l'endonucléase Dna2 (Symington, 2014). Dna2 agit comme la nucléase bactérienne RecJ. BLM est un des homologues de l'hélicase bactérienne RecQ et agit comme tel, même s'il a été montré qu'elle exerce également des fonctions translocase proches de celles de RecG (Killen *et al.*, 2012). Une fois l'extrémité 3'-sortante générée, RPA (homologue fonctionnel de SSB) se fixe au ssDNA pour le stabiliser. C'est ensuite RAD51 (homologue fonctionnel de RecA) qui se lie au ssDNA afin de promouvoir l'invasion de brin dans une région homologue. RAD52 et RPA permettent l'invasion du second brin dans le cas du DSBR ou du SDSA. RecQ1 et FANCM sont des hélicases impliquées dans la migration des HJ (Kowalczykowski, 2015). Les nucléases GEN1 ou SLX1-SLX4 contrôlées par ERCC4-ERCC1 permettent le clivage et la résolution des HJ.

Les plantes possèdent un homologue au gène RAD51, mais aussi RAD50 et Mre11, nécessaire à l'invasion de brin. Une autre protéine INO80 a un rôle dans la recombinaison de l'ADN des plantes (Fritsch *et al.*, 2004). Plusieurs protéines du NER (XPF, ERCC1 et CENTRIN2) sont impliqué dans le HR, montrant un lien entre le NER et le HR (Molinier *et al.*, 2008; Molinier *et al.*, 2004a). Chez les plantes, un mécanisme de réparation, dépendant d'une protéine SNM a été montré dans la réparation des lésion oxydative par recombinaison

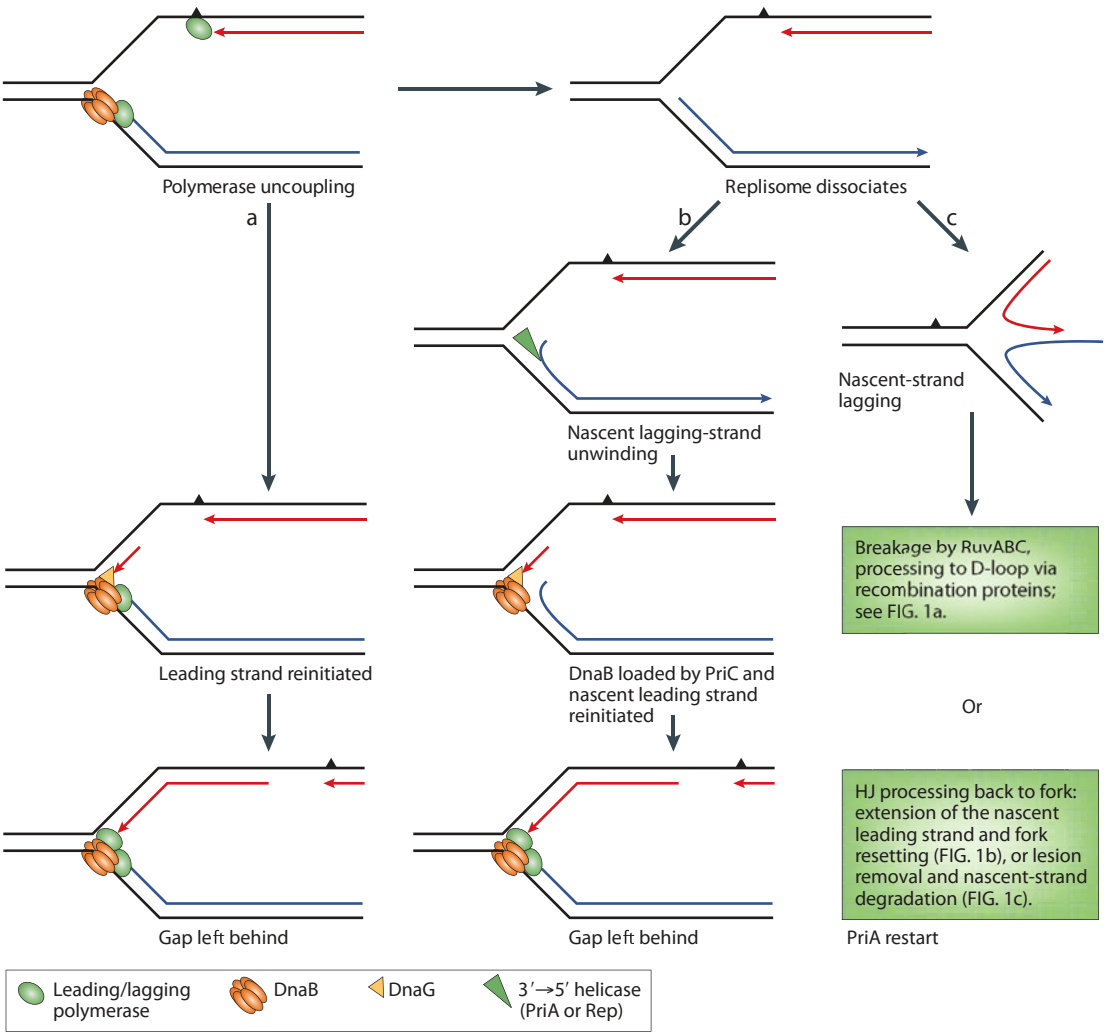


Figure 11. Réactivation de la fourche de réplication suite à une lésion sur le brin précoce chez *E. col.* a. Quand le complexe de réplication rencontre une lésion (triangle noir) sur le brin précoce les polymérases se découplent. Il est possible que l'hélicase DnaB continue de dérouler les brin parentaux après la lésion et que la polymérase du brin retard reste en place. Dans ce cas, la primase DnaG synthétyse une amorce en aval de la lésion, et le réplisome se remet en place, laissant un SSG. **b.** Dans le cas où le réplisome se dissocie, une hélicase tel que PriA ou Rep déroule le brin retard de 3' en 5' et expose une portion de ssDNA pour que le système de ré-initiation dépendant de PriC charge DnaB. La fourche de réplication se remet en place et laisse un SSG. **c.** Pour les modèles qui impliquent une régression de fourche, le brin précoce et retard s'apparient ensemble et la fourche se réverse en formant une HJ. A partir de là, il peut y avoir la mise en place de mécanismes de réplication recombinaison-dépendante ou de redémarrage direct, dont le détail est représentées dans la Figure 12. Ces processus génèrent une structure qui est reconnue par PriA pour la ré-initiation. Les brins noirs représentent les brins parentaux, le brin rouge représente le brin précoce et le brin bleu représente le brin retard. D'après Heller et Marians, 2006.

(Molinier *et al.*, 2004b). De nombreux homologues de RecQ existent et RECQ4 a un grand rôle dans le contrôle de la recombinaison et de la réparation par recombinaison (Schropfer *et al.*, 2014). Un homologue de FANCM est aussi impliqué dans ce contrôle (Knoll *et al.*, 2012), et deux paralogues de la résolvase GEN1 sont aussi présents (AtGEN1 and AtSEND1) (Bauknecht et Kobbe, 2014).

1.2.3.2.2. ***Sauvetage des fourches de réplication***

Lorsque les cellules sont en phase de prolifération l'ADN est répliqué. Or il arrive que les lésions dues à l'environnement extérieur ou à des facteurs endogènes ne soient pas réparées avant la réplication. Dans ce cas des cassures ou des lésions peuvent arrêter le complexe de réplication et ainsi engendrer des versions du génome tronquées ce qui est délétère pour l'organisme.

1.2.3.2.2.1. Ré-initiation de la réplication et régression des fourches.

Chez *E. coli* il a été montré depuis longtemps que la réplication s'amorce au site OriC, dépendant du facteur d'initiation DnaA. Or, les cas où le complexe de réplication se dissocie à cause d'une lésion sont fréquents, et la réplication doit pouvoir se ré-initier indépendamment de OriC, afin de maintenir une bonne intégrité génétique. Les mécanismes permettant de contourner ou de réparer ces lésions diffèrent selon la position de la lésion sur le brin précoce ou retardé ou selon le type de lésion. Des mécanismes de ré-initiation directe du réplisome existent dans le cas où le réplisome se dissocie. Il s'est avéré que la réparation par recombinaison est intimement liée à la réplication. Des mécanismes de recombinaisons permettent de réparer ces lésions, mais ils semblent aussi importants pour la ré-initiation du complexe de réplication. Cette ré-initiation de la réplication est recombinaison-dépendante (RDR pour "recombination-dependent replication").

Si la lésion est située sur la matrice du brin retardé, le réplisome ne se désassemble pas et seule la synthèse du fragment d'Okasaki correspondant est affectée. La polymérase se déplace jusqu'à la prochaine amorce et laisse une SSG qui sera réparée par la suite. Le brin précoce ne sera pas affecté (Yeeles *et al.*, 2013). Si la lésion est située sur la matrice du brin précoce, la polymérase qui le réplique se bloque et la synthèse du brin retard continue sur quelques nucléotides (Pages et Fuchs, 2003). Il arrive que le réplisome ne soit pas dissociée et qu'il puisse passer outre la lésion. Une amorce est synthétisée au niveau du réplisome qui a progressé le long du brin retard, et la polymérase du brin précoce se réamorce, laissant un trou (figure 11) (Yeeles et Marians, 2011). Cependant, ce mécanisme n'est pas toujours efficace,

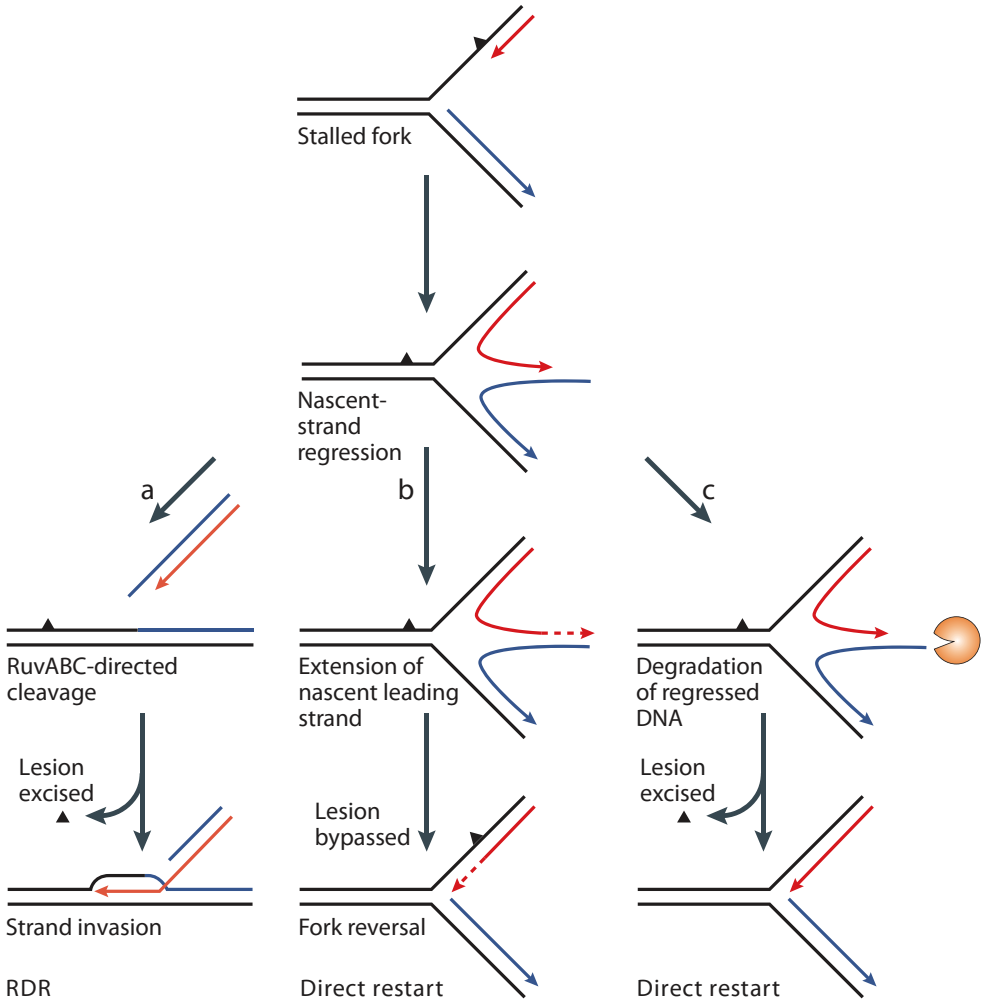


Figure 12. Réplication recombinaison-dépendante et mécanismes de redémarrage directe dans le cas de régression de fourche. a. Réplication recombinaison-dépendante. La HJ est résolue par RuvABC. Après l'excision de la lésion, l'extrémité d'ADN libre peut envahir le duplexe intacte et former une D-loop. **b.** Dans le cas d'un changement de matrice, le brin retard est étendu en prenant le brin précoce pour matrice. Après rétablissement de la fourche, la lésion est contournée. **c.** Alternativement, la lésion est excisée, et la jonction 4 brins est réorganisée en fourche de réplication par l'action d'une exonucléase représentée en orange. Ces trois processus génèrent une structure qui est reconnue par PriA pour la ré-initiation. Les brins noirs représentent les brins parentaux, le brin rouge le brin précoce et le brin bleu le brin retard. D'après Heller et Marians, 2006.

notamment dans le cas de nombreuses lésions. Si le réplisome est dissocié, il peut y avoir une ré-initiation directe de la réplication par PriC (figure 11), ou une régression de la fourche. PriC est une hélicase permettant la ré-initiation directe du complexe de réplication. Cependant elle ne peut agir que lorsqu'il y a un SSG suffisamment important sur le brin précoce et lorsque la fourche de réplication n'est pas régressée (Heller et Marians, 2006). Ce mécanisme de ré-initiation est moins important que celui de régression de la fourche de réplication. Lorsque le réplisome est dissocié, RecG peut alors catalyser la régression de la fourche de réplication en appariant les 2 brins néosynthétisés et former une HJ (figure 12) (McGlynn et Lloyd, 2002a). Si la résolvase RuvABC intervient sur la HJ, la molécule parentale est dissociée de la molécule néoformée. La réparation peut alors s'opérer via le mécanisme de RDR. RecA catalyse une invasion de brin ce qui permet la mise en place d'une D-Loop. La réinitiation du complexe de réplication a alors lieu et l'élongation se déroule selon le BIR. Cette réaction requiert PriA qui réinitie la réplication, mais aussi RecBCD et les sites Chi. Il a également été montré que la réplication peut être réinitié par une R-Loop, en faisant intervenir RecA et PriA (Kogoma, 1997). Une R-loop est une structure où un brin d'ARN est apparié à l'un des brin d'un duplex ADN. L'ARN sert donc d'amorce pour la synthèse d'ADN. L'enzyme PriA est une hélicase reconnaissant les D-Loop, ainsi que les fourche de réplication arrêtées. PriB stabilise le complexe PriA-ADN et permet le recrutement de l'hélicase réplicative DnaB, puis du complexe entier. Ce mécanisme est important dans la réparation des lésions UV bloquant la fourche de réplication (Yeeles *et al.*, 2013).

D'autres mécanismes permettent de contourner la lésion qui est réparée postérieurement, ce qui permet de poursuivre la réplication de façon efficace (Heller et Marians, 2006). L'action translocase de RecG sur la HJ mise en place après la régression de brin permet l'allongement du brin précoce en prenant le brin retard, plus long, pour matrice (figure 12). RecG peut ensuite prendre en charge le rappariement de chacun des brins néo-formés avec leur matrice complémentaire. La lésion n'est pas réparée mais la réplication est achevé et la lésion pourra être prise en charge par un autre mécanisme de réparation (McGlynn et Lloyd, 2002b). Alternativement, si la lésion est rapidement excisée, une fourche de réplication peut être re-établie par l'action d'une exonucléase et de PriA ou PriC (figure 12).

La réparation des SSGs laissées lors du contournement des lésions du brin précoce ou du brin retard s'opère par le mécanisme impliquant RecFOR, en prenant pour modèle l'autre molécule d'ADN issue de la réplication (figure 13). La proximité de celle-ci facilite l'invasion de brin par RecA (Heller et Marians, 2006).

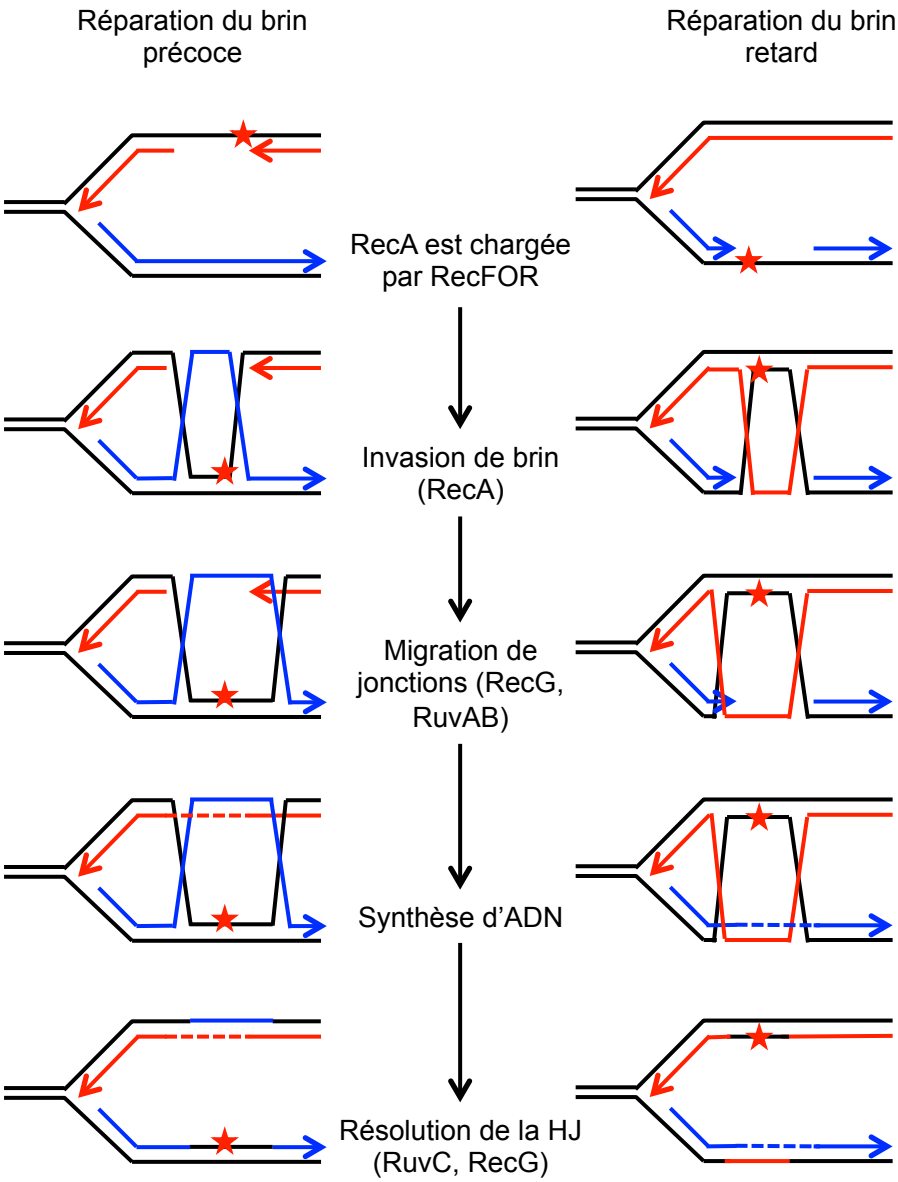


Figure 14. Réparation des lacunes simple brin (SSG) dans *E. coli*. Dans les bactéries, les SSG peuvent se situer sur le brin retard, lorsqu'une lésion est contournée, ou sur le brin précoce lors du redémarrage de la fourche par PriA. A l'aide de RecFOR, RecA est chargée sur le brin néoformé. Puis RecA catalyse l'invasion de brin et la réparation se déroule comme dans le mécanisme HR. Brin en noir : ADN parental ; pointillés : ADN synthétisé lors de la réparation ; brin bleu : brin retard ; brin rouge : brin précoce.

Chez l'homme et la levure, il existe un système de signalisation complexe faisant intervenir Mec1/ATR pour protéger les fourches de réplication bloquées. Dans certains cas, le réplisome peut être déplacé sans être dissocié pour permettre l'accès à la lésion par les mécanismes de réparation. Mais ces mécanismes sont complexes et loin d'être entièrement compris. Mre11 et Brca2, des facteurs de recombinaison, ont également un rôle dans la protection et la réparation des fourches de réplication. La première élargit les SSG pour permettre la mise en place de mécanismes de réparations post-réplicatifs, et la deuxième charge Rad51 sur le ssDNA pour éviter sa dégradation (Costanzo, 2011). La régression de fourche a également été montrée chez les animaux et la levure. PARP1 a été identifiée comme un facteur important dans ce mécanisme (Ray Chaudhuri *et al.*, 2012). L'activité des facteurs HTLF (ou Rad5 chez la levure) et de l'hélicase FANCM (Fml1 chez la levure) dans la régression de fourche a été montrée *in vitro*. La réplication est ensuite réinitiée après une dégradation des brins régressés, ou des mécanismes faisant intervenir la HR. L'extrémité 3' du brin précoce peut envahir le duplex parental en aval de la lésion, en créant ainsi une HJ qui est résolue par l'hélicase BLM (homologue fonctionnel à RecQ) et la topo-isomérase Top3, et permet ainsi le contournement de la lésion. Il a été suggéré récemment que SMACARL1 agirait comme un homologue fonctionnel de RecG dans la régression de fourche (Betous *et al.*, 2013). De nombreux facteurs sont impliqués mais les mécanismes sont encore méconnus (Yeeles *et al.*, 2013). SrS2 a été identifiée dans la levure comme un homologue de UvrD et inhibe la formation d'intermédiaires de recombinaison non-productifs (Heller et Marians, 2006).

L'ADN génomique des plantes a une structure proche de celui des vertébrés et on peut donc s'attendre à ce que les mécanismes de réparation des fourches soient proches de celui des vertébrés et que des facteurs communs aux animaux soient présents. Le facteur de recombinaison BRCA2 a un rôle dans le contrôle d'invasion de ssDNA lors de HR (Seeliger *et al.*, 2012). Les plantes possèdent également trois PARP qui sont impliquées dans la réparation des DSB. PARP3 a un rôle dans la réparation de l'ADN des plantes au stade graine, jouant un rôle dans leur viabilité au court du temps. PARP1 et PARP2 sont impliqués dans la réparation des DSB chez les plantes, bien que PARP2 semble la plus importante dans ces mécanismes (Rissel *et al.*, 2014; Song *et al.*, 2015).

1.2.3.2.2.2. La synthèse trans-lésion (TLS)

Alternativement, le mécanisme du Translesion Synthesis (TLS) permet de passer par-dessus une lésion bloquant la réplication, en échangeant de polymérase. Ainsi, chez *E. coli*, il

y a recrutement de l'ADN PolV par β-clamp, pour remplacer l'ADN PolIII. β-clamp est une protéine dimèrique qui encercle les amorces et attache PolV à l'ADN pour permettre une réplication processive. Une fois la lésion passée la PolIII reprend sa place (Friedberg, 2005).

Chez les eucaryotes, le contournement des lésions est favorisé et le TLS est donc important. Quand une lésion n'est pas réparée avant la réplication, la Pol δ/ε réplicative s'arrête et est échangée par une des Pol ζ, η, ι, κ avec RevI. Ces polymérases permettent la synthèse d'ADN malgré la présence de la lésion, avec une haute fidélité pour la plupart. Ce mécanisme permet la réplication d'ADN comportant des dimères de thymine et dépend de RevI (Kimura et Sakaguchi, 2006; Yoon *et al.*, 2015).

Dans les plantes, plusieurs facteurs du TLS ont été trouvés. Les ADN Pol ζ, η et κ et Rev1 ont été identifiées dans Arabidopsis et dans le riz, et la sensibilité aux UV des mutant Arabidopsis pour Polζ ont montré qu'elle avait bien un rôle dans le TLS (Kimura et Sakaguchi, 2006).

De nombreux facteurs différents sont impliqués dans la réparation. Parmi eux, l'une des catégories importantes sont les hélicases. En effet elles ont un rôle non seulement dans le déroulement des brins d'ADN, mais aussi dans la résolution et la migration de structures plus complexes faisant intervenir plusieurs brins d'ADN et/ou d'ARN. Je détaille dans le chapitre suivant les différentes catégories d'hélicases répertoriées, ainsi que deux hélicases particulières, RecG et TRCF, sur lesquelles j'ai travaillé pendant ma thèse.

2. Les hélicases

Les hélicases sont une classe de protéines essentielles à tous les organismes. Ce sont des ATPases capables de dérouler des duplexes ADN-ADN, ARN-ADN ou ARN-ARN. De fait, elles interviennent dans les mécanismes importants de la vie impliquant des acides nucléiques. Si la fonction historique des hélicases est l'ouverture de la double hélice d'ADN pour la réplication, la réparation et la recombinaison, leur rôle a depuis été montré dans la transcription, la traduction, la synthèse des ribosomes, la maturation et l'épissage des ARN et le remodelage de la chromatine (Lohman *et al.*, 2008).

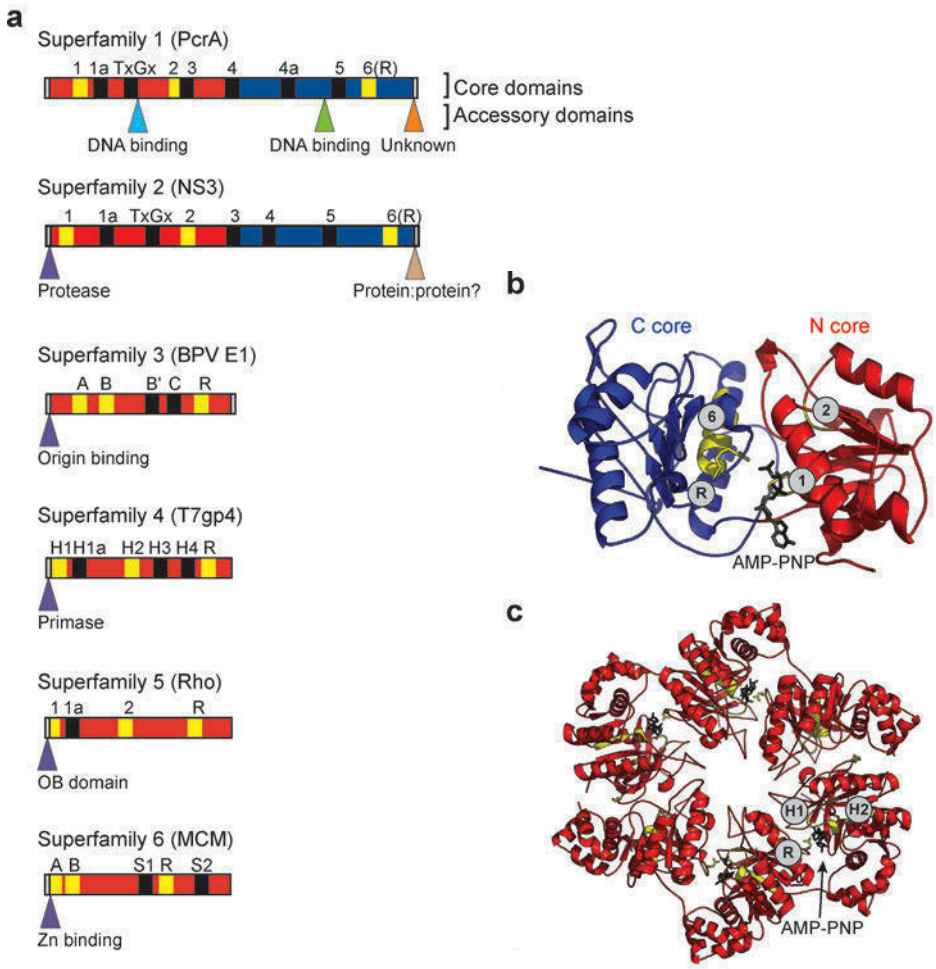


Figure 14. Classification des helicases et translocases. a. Le nom de la protéine rprésentée comme exemple de chaque famille est inscrit entre parenthèses. Le domaine noyau ("core-domain") et les positions des motifs spécifiques sont représentés pour chaque type d'hélicase. Les motifs en jaune représentent les éléments structuraux universels présents dans toutes les hélicases. La position et la fonction des domaines supplémentaires sont aussi représentés et sont spécifiques à chaque hélicase. Les domaines en rouge correspondent au domaine N du noyau. Les domaines en bleu correspondent au domaine C du noyau. **b et c.** Structure du domaine noyau. Les éléments structuraux responsables de la liaison et l'hydrolyse de NTP sont montrés en jaune. En bleu, le domaine C, et en rouge le domaine N **b.** Les enzymes SF1 et SF2 contiennent un noyau monomère formé par le tandem de deux feuillets RecA-like identiques. Les domaines N et C-terminal RecA-like sont représentés avec un analogue au NTP lié à l'interface des deux domaines. Les motifs 1 et 2 sont les motifs Walker A et B. Le motif 6 contient un doigt arginine. Ce domaine noyau constitue le moteur de translocation minimal. **c.** Exemple d'une enzyme de la famille SF3 (T7gp4) constituée de six domaines RecA-like individuels (en rouge) arrangés en anneau. D'après Singleton *et al.*, 2007.

Pendant ma thèse, je me suis intéressée à deux protéines d'Arabidopsis homologues aux translocases bactériennes RecG et TRCF (Mfd). Après avoir présenté les différents types d'hélicases, je présenterai plus en détail les rôles qui ont été identifiés pour ces deux protéines

2.1. Les grandes familles d'hélicases

Si les ADN hélicases se définissent par leur capacité à séparer un dsDNA (pour "double-stranded DNA") en deux brins séparés grâce à l'hydrolyse de NTP, les translocases sont des enzymes capables de se déplacer dans un sens donné le long d'un acide nucléique en hydrolysant du NTP. Les hélicases sont donc des translocases. On trouve une grande diversité d'hélicases, reflétant la diversité de leur fonction. Les hélicases sont définies par leurs domaines dont deux sont communs à toutes. Elles possèdent un domaine ATPase RecA-like composé des boites Walker A et B (figure 14b), impliquées dans la liaison et l'hydrolyse du NTP. On trouve aussi un "doigt d'arginine", qui a un rôle dans le couplage de l'énergie (Ye *et al.*, 2004). Ces hélicases ont été classées en six superfamilles (SPF). Avec les données structurales récentes, on sait maintenant que certaines de ces hélicases ne sont que des translocases et ne possèdent même parfois pas d'activité ATPase. Les hélicases se différencient par leur fonctionnement en monomère ou en multimère. Les hélicases SPF1 et SPF2 sont considérées comme agissant en monomères, bien qu'en fait leur activité hélicase soit parfois stimulée par leur dimérisation. Les hélicases des familles SPF3 à SPF6 fonctionnent toutes en hexamères ou en double hexamères et forment des structures toriques ("donut structure") (figures 14c et 15). Les hélicases progressent dans un sens sur un brin donné, et on distingue les hélicases en sous-groupes A et B selon que leur processivité soit de 3' en 5' ou de 5' en 3', respectivement. C'est pourquoi les processus de reconnaissance de brin sont indispensables pour déterminer la direction de l'hélicase. L'addition de domaines protéiques supplémentaires en N ou C-terminal peut changer l'activité ou la spécificité d'une hélicase, donnant ainsi un grand panel de possibilités (Lohman *et al.*, 2008; Singleton *et al.*, 2007).

2.1.1. SPF1 et SPF2

Les hélicases des familles SPF1 et SPF2 sont structuralement différentes des quatre autres familles. En effets chaque unité protéique comprend deux motifs RecA-like avec un domaine de liaison à l'ATP entre les deux, alors que les familles 3 à 6 forment des hexamères de 6 ou 12 motifs RecA-like individuels. C'est peut-être la raison pour laquelle elles peuvent agir en monomère ou dimère (figure 14a). L'activité hélicase des protéines des familles SPF1

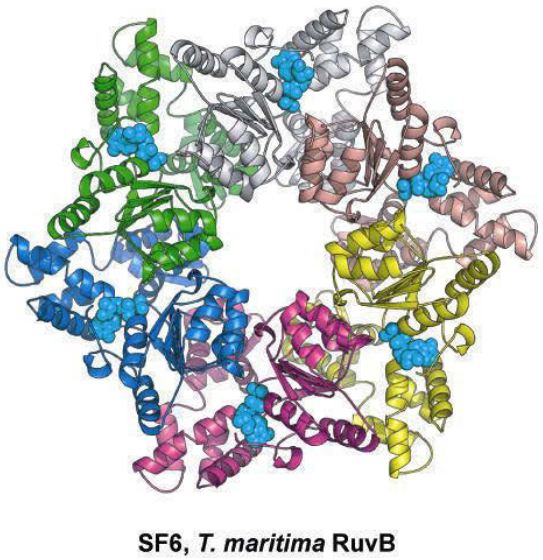


Figure 15: Représentation en ruban de l'hélicase héxamérique RuvB. D'après Singleton *et al.*, 2007

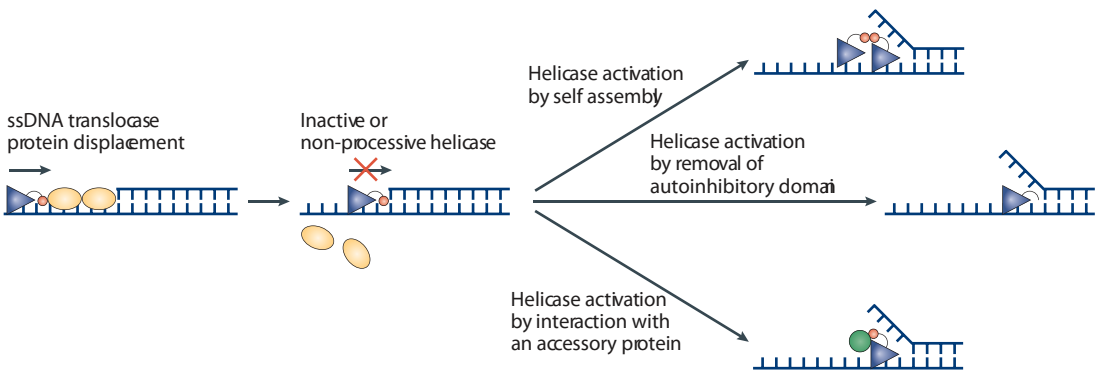


Figure 16. Activation de l'activité hélicase d'une translocase monomérique. Une translocase de ssDNA monomérique de la superfamille 1 (triangle bleu) peut déplacer des protéines (ovales jaunes) qui sont liées au ssDNA. La capacité de l'enzyme à dérouler le dsDNA peut être auto-inhibée par l'un de ces sous-domaines (orange). Cependant l'activité hélicase d'une enzyme SF1 peut être activée ou stimulée par auto-assemblage, par suppression du domaine auto-inhibiteur, ou par interaction avec une autre protéine accessoire (cercle vert). D'après Lohman et *al.*, 2008

et SPF2 est souvent déclenchée par leur assemblement en dimère, la suppression d'un domaine inhibiteur, ou l'association avec une autre protéine (figure 16) (Lohman *et al.*, 2008).

La famille SF1 est celle qui a été le mieux étudié au niveau mécanistique. Ces hélicases sont toutes de type α, c'est à dire qu'elles ont toutes du ssDNA comme substrat. Les enzymes de la famille SF1A ont une activité de 3' en 5'. Parmi elles on trouve deux des hélicases impliquées dans le NER ou le HR : Rep et UvrD. Il a été montré *in vitro* que UvrD a une bonne activité de translocase sur du ssDNA mais une activité hélicase limitée. Elle est activée quand la protéine est associée en dimère (Maluf *et al.*, 2003). Parmi les hélicase SF1B on trouve RecD qui, associée à RecB et RecC est impliqué dans la résection des brins d'ADN pour permettre la fixation de RecA.

La famille SF2 est la plus importante et comprend des hélicases impliquées dans de nombreux processus différents. Elle comprend les ARN-hélicases à motif DEAD-box, la famille des RecQ-like et des Snf2-like ADN-hélicases. La plupart sont de type A, excepté la famille des XPD qui se déplace de 5' en 3'. La particularité de cette famille est dans les divers ajouts de domaines protéiques supplémentaires en N ou C-terminal qui confère à l'hélicase une activité supplémentaire. Par exemple la protéine RecG (famille SF2Aβ) possède un domaine "wedge" qui lui donne la capacité de se lier à des structures de type HJ ou en Y. UvrB, qui est impliqué dans le NER, possède un motif en ruban β (constitué de deux feuillets β en épingle à cheveux) qui se glisse entre les deux brins d'ADN et lui donne sa spécificité pour l'ADN endommagé. TRCF quant à elle, possède un domaine supplémentaire de liaison à la RNAP et lui permet de détecter les complexes de réplication bloqués (Lohman *et al.*, 2008; Singleton *et al.*, 2007).

2.1.2. SPF3, SPF4, SPF5 et SPF6

Les hélicases de ces familles sont composées d'hexamères de protéines assemblées en formes toriques. Les domaines de liaison aux nucléotides sont situés à l'interface entre les monomères. Cependant cette structure en forme d'anneau est fermée et pose le problème du chargement de l'hélicase sur la molécule d'ADN. Elles sont donc souvent aidées de façon ATP-dépendante par des protéines partenaires. Les protéines de ces trois familles sont fortement similaires. La première hélicase bactérienne réplicative identifiée fait partie de la famille SPF4. La protéine Rho qui détache les hybrides DNA/RNA fait partie de la famille SPF5. Elle a la particularité d'exister sous forme ouverte ou fermée. Cette conformation permet le passage du ssRNA dans le canal central. L'hélicase RuvB fait partie de la famille

SPF6. Les hélicases sont donc une famille très diversifiée (Lohman *et al.*, 2008; Singleton *et al.*, 2007).

2.2. RecG

RecG est une hélicase qui a été identifiée dans *E. coli*. Ses fonctions dans la réparation, la réplication et de manière générale dans le maintien du génome ont été définies. Elle peut d'un côté promouvoir la recombinaison par la formation de HJ, mais déstabilise les D-loop ou R-loop, étape préliminaire à la recombinaison. Ses activités peuvent donc paraître contradictoires et dépendent des mécanismes dans lesquels elle agit.

2.2.1. Structure de RecG

RecG est une transférase hélicase de la famille SF2Aβ. Son rôle dans la migration et la formation des jonctions de Holliday (HJ) a été montré il y a longtemps (Lloyd et Buckman, 1991; Whitby *et al.*, 1993). RuvAB est la voie alternative permettant la migration de HJ. Cependant ces deux protéines ont une structure bien différente. RuvB est une hélicase héxamérique (figure 15) (Putnam *et al.*, 2001; Yamada *et al.*, 2001) et elle est recrutée par le tétramère RuvA (Rafferty *et al.*, 1996) quand il est en complexe avec une HJ. Les quatre unités se positionnent entre les duplexes d'ADN et le motif central en ruban β permet la séparation de brin (Hargreaves *et al.*, 1998; Roe *et al.*, 1998). L'hélicase RuvB sert de moteur et permet la migration de HJ en poussant l'un des duplex d'ADN vers RuvA. RecG est structurellement bien différente car elle est active en monomère, et c'est sa structure particulière qui lui permet d'agir aussi efficacement qu'un hexamère et un tétramère ensemble.

Singleton et ses collaborateurs ont réussi à cristalliser la protéine RecG de *Thermotoga maritima*, associée à son substrat (Singleton *et al.*, 2001). Cette structure a permis de mieux comprendre comment RecG se lie à son substrat et comment elle active la migration des HJ. Le substrat utilisé pour la cristallisation mime une fourche de réplication dont le brin précoce néoformé est manquant. En quelque sorte, elle mime une jonction en 3 brins, le substrat favori de RecG *in vitro* (McGlynn et Lloyd, 2001). Le domaine 1 en N-terminal est appelé domaine wedge (ce qui signifie cale). Il est très conservé dans toutes les bactéries. La structure a montré que l'interaction avec l'ADN se fait surtout au niveau du domaine wedge, mais aussi d'un motif en ruban β. Ces deux structures forment un motif qui permet la séparation de deux duplexes d'ADN. Les deux brins du duplex parental sont positionnés de chaque côté du domaine wedge dans des rainures trop étroites pour accueillir un duplex. Les deux brins

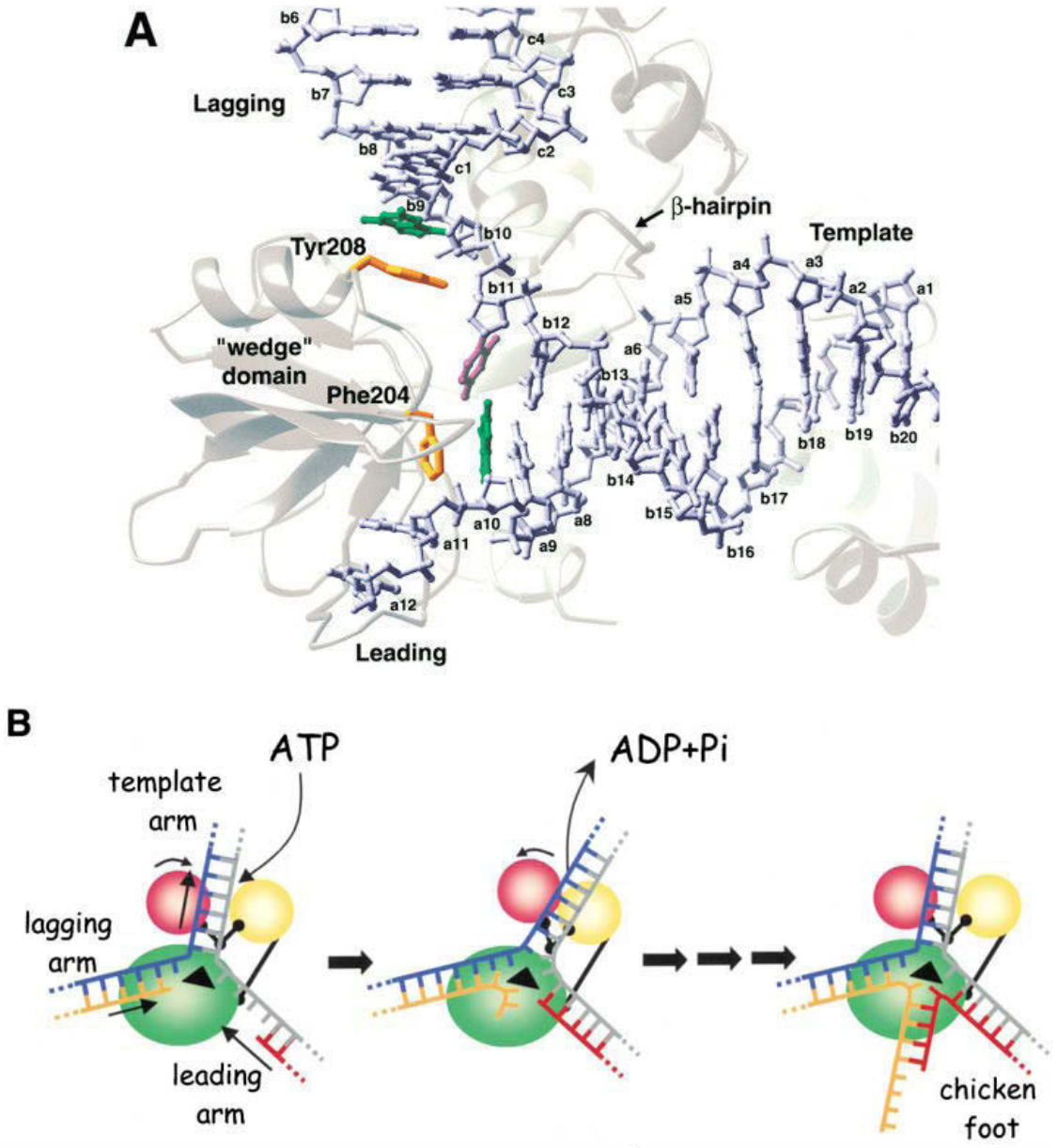


Figure 17. Représentation structural du domaine "wedge" de RecG. A. Interactions entre la protéine (en gris) et l'ADN (bleu pâle). Les interactions aromatiques (Phe204 et Tyr208 en jaune) stabilisent les bases orphelines (a10 et b10 en vert) à la jonction. La base b11 représentée en magenta est tournée vers l'intérieur pour une meilleure protection. La structure ADN représentée est une structure en Y, avec un simple brin précoce et un double brin retard. **B.** Domaine 1 en vert, 2 en jaune et 3 en rose. Les brins d'ADN parentaux sont représentés en gris et bleu. Le brin retard est représenté en orange et le précoce en rouge. Le domaine "wedge" est représenté par un triangle noir. Le brin précoce de la jonction bloquée est d'abord lié à la protéine avant hydrolyse de l'ATP. Quand l'ATP se lie, la fente entre le domaine 2 et 3 se ferme, ce qui induit un changement de conformation et tire sur les brin parentaux. Ce mouvement permet la séparation du duplexe formé avec le brin retard via le domaine "wedge". Quand l'ATP est hydrolysé, les domaines 2 et 3 s'écartent et glissent le long des brins parentaux qui est maintenu par des jonction étroite avec le domaine 1. D'après Singleton *et al.*, 2001.

parentaux sont alors rassemblés base après base. La base du brin parental provisoirement orpheline est protégée par la structure de la protéine (figure 17A). Un rapprochement entre le domaine 1 et le domaine 3 laisse à penser que les domaines 2 et 3 servent de moteur. En effet les domaines 2 et 3 sont deux domaines RecA-like, dit motifs hélicase. Le domaine DEAD-box contient le domaine de liaison à l'ATP, et le domaine hélicase contient les deux motifs Walker. Singleton propose donc que ces deux domaines se lient activement à l'un des duplexes, et grâce à l'hydrolyse d'ATP "poussent" la molécule d'ADN à travers le domaine wedge (figure 17B) (Singleton *et al.*, 2001).

2.2.2. Rôles de RecG dans *E. coli*

Les premières études sur la protéine RecG ont montré qu'elle catalyse la migration des HJ (Lloyd et Sharples, 1993). Les HJ sont des structures caractéristiques des intermédiaires de recombinaison et sont formées de 4 duplexes de DNA associés entre eux. Cette fonction est aussi réalisée par la resolvase RuvABC, qui agit dans une autre voie. Ceci a été confirmé par le fait que le double mutant *recG ruv* montre une très grande sensibilité aux stresses génotoxiques, quand le simple mutant *recG* a une sensibilité modérée. Cependant RecG n'a pas d'activité nucléase et ne clive donc pas les HJ. Il a été montré *in vitro* que RuvC peut cliver les HJ catalysées par RecG (Gupta *et al.*, 2014). Mais en l'absence de RuvABC, RusA pourrait résoudre les HJ formées par RecG (Chan *et al.*, 1998).

De par sa fonction, la protéine RecG est un facteur de recombinaison intervenant à plusieurs niveaux. Elle permet la migration des HJ mais catalyse aussi leur mise en place (McGlynn et Lloyd, 2001). Plusieurs situations font intervenir les HJ en tant qu'intermédiaire de recombinaison. Lors de la réparation par recombinaison homologue, comme décrit en I.1, une fois que le filament nucléoprotéique RecA a envahi une région homologue une ou plusieurs HJ sont mises en place pour permettre la correction de la lésion ou combler les SSG. RecG a donc un rôle important dans la réparation des lésions. D'autre part, la régression des fourches de réplication est une étape essentielle lorsque celles-ci sont bloquées par une lésion. Le désappariement des brins néoformés des brins parentaux et la formation d'une HJ comme intermédiaire de recombinaison permettant la réparation est catalysée par RecG (McGlynn et Lloyd, 2002a). La redondance de la fonction de RecG et RuvAB pourrait être un moyen d'assurer la réparation par deux voies différentes. Cependant, RecG possède des activités supplémentaires et différentes de RuvAB (Zhang *et al.*, 2010).

Des tests ont montré que la fréquence de recombinaison entre un plasmide et l'ADN génomique diminue fortement chez le mutant *ruv*. A l'inverse, et contre toute attente, chez le mutant *recG* cette fréquence augmente, signifiant que RecG a un rôle d'inhibiteur sur la recombinaison (Grove *et al.*, 2008). De plus, il a été montré *in vitro* que RecG dissocie les structures D-loop et R-loop (McGlynn *et al.*, 1997; Vincent *et al.*, 1996). Ceci signifie que RecG a un rôle dans le maintien du génome, en inhibant les réactions de recombinaison surnuméraire pouvant mener à une perte dans l'information génétique. Ce rôle de "gardien" est aussi appliqué à la réplication. En effet, les R-loop sont des intermédiaires de recombinaison permettant l'initiation de la réplication chez certains plasmides (origine de réplication ColE1). RecG contrôle donc la réplication de ces plasmides en déstabilisant les R-loop. Mais elle a aussi un rôle dans la réplication de l'ADN génomique. En effet il a été montré que les mutants *recG* ont un nombre accru de fourches de réplication DnaA-indépendantes. En d'autres termes, le nombre de sites d'initiation de la réplication recombinaison-dépendants est très élevé et mène à des graves problèmes de ségrégation des chromosomes. Ce phénotype est diminué quand l'hélicase PriA est mutée (Grove *et al.*, 2008). Cette dernière permet la ré-initiation du complexe de réplication par recrutement de DnaB.

La réplication du chromosome circulaire bactérien commence à l'origine de réplication OriC. Deux fourches de réplications se mettent en place et progressent dans le sens inverse le long du chromosome. Dans des cellules en phase de multiplication, un autre cycle de réplication peut commencer avant que le premier ne soit achevé. Lorsque les deux fourches de réplication se rencontrent dans les zones de terminaison, il peut y avoir une collision des fourches et déplacement du brin précoce de l'une d'elles sur l'autre fourche, laissant une extrémité 3' sortante. PriA a une affinité pour les extrémité 3' sortantes, qui servent de substrat pour ré-initier la réplication. Or cette réplication pathologique peut mener à la formation de chromosomes inachevés. Dans cette situation, le bon déroulement de la ségrégation et la stabilité chromosomique peuvent être menacés, et la cycle de division retardé. RecG diminue le risque de ré-initiation en convertissant les extrémités 3' sortantes en 5' sortantes. Une exonucléase simple brin 5'-3' peut ensuite dégrader le brin flottant (Rudolph *et al.*, 2009; Rudolph *et al.*, 2013).

Les rôles de RecG sont donc importants et multiples dans le maintien de l'ADN dans *E. coli*, car ils concernent la recombinaison, la réparation, la réplication et même la ségrégation de l'ADN. RecG agit comme facteur de recombinaison en la stimulant ou l'inhibant selon les cas de figure.

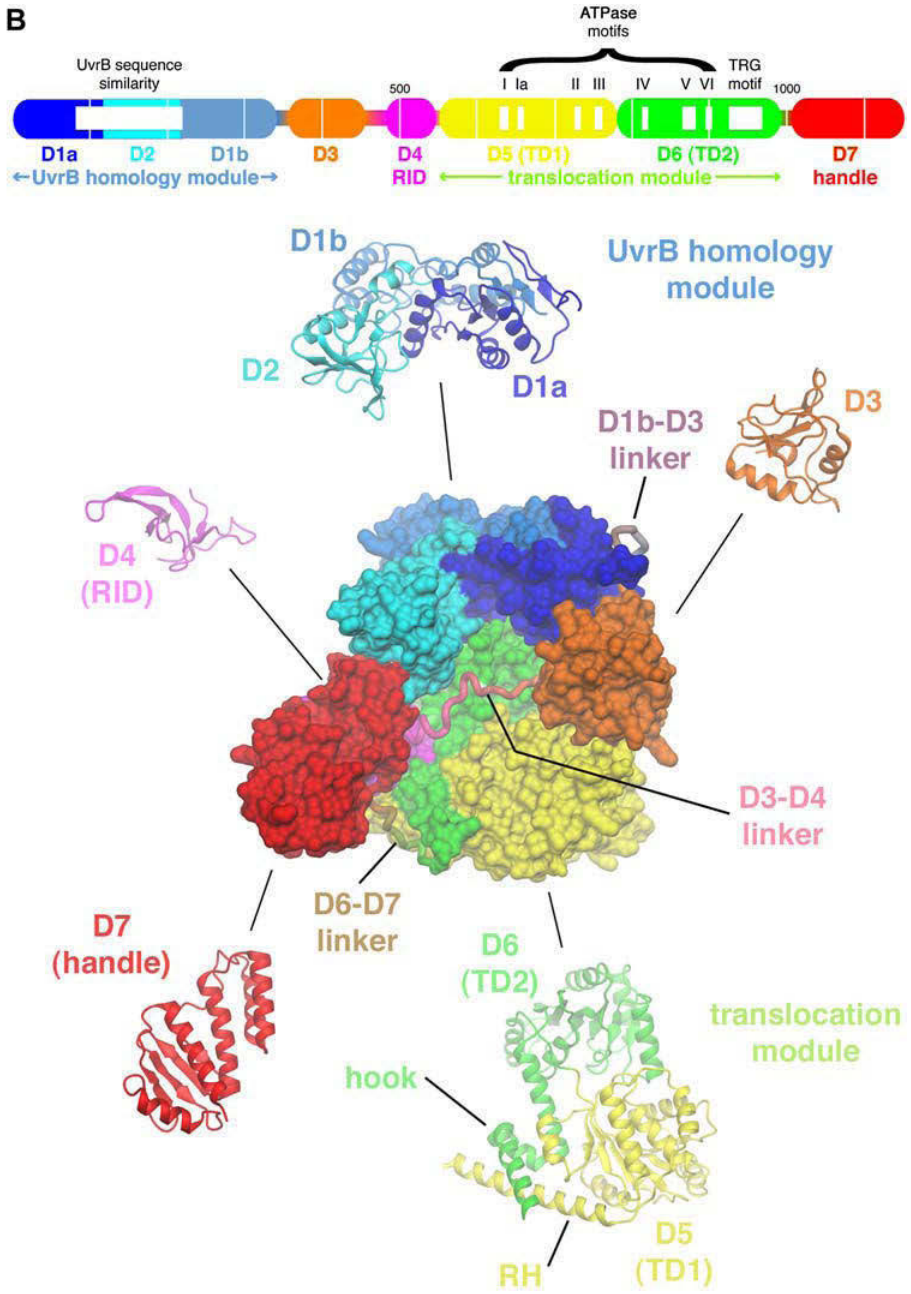


Figure 18. Structure de TRCF

En haut: représentation schématique des domaines de TRCF. D1a/D2/D1b contient une similarité de séquence avec UvrB. D4 est le domaine d'interaction avec la RNAP (domaine RID). D5 et D6 sont aussi appelés TD1 et TD2 pour "translocation domain" et contiennent les motifs SF2 helicase. Ils forment le même motif TRG que RecG, et constituent le module de translocation. En bas, les mêmes couleurs ont été conservées pour la représentation structurale de la protéine, le détail des différents domaine est représenté en ruban. D'après Deaconescu *et al.*, 2006.

2.3. TRCF

TRCF, aussi appelée protéine Mfd, est une translocase qui agit dans la voie de la réparation couplée à la transcription du NER. Elle détecte la RNAP bloquée par des lésions sur l'ADN matrice et recrute les facteurs du NER.

2.3.1. Sa structure

La structure de cette protéine a été étudiée par Deaconescu et ses collaborateurs (Deaconescu *et al.*, 2006). Elle comprend 8 domaines structuraux (figure 18). Les 3 premiers : D1a, D2 et D1b, correspondent à un domaine similaire à UvrB. D4 est un domaine d'interaction avec la RNAP (RID pour "RNAP interacting domain"). D5 et D6 sont les domaines comprenant les modules de translocation. Ils sont appelés respectivement TD1 et TD2 pour "Translocation Domain". Ces deux domaines confèrent l'activité translocase et sont fortement homologues aux domaines "moteurs" de RecG. TD1 et TD2, avec l'aide de l'hydrolyse d'ATP semblent être responsables du détachement de RNAP. Dans le NER, UvrB reconnait l'ADN endommagé grâce à un ruban β situé entre les deux domaines D1a et D1b. TRCF en est démunie et sert donc uniquement de marqueur protéine-protéine pour signaler l'arrêt de la RNAP et recruter UvrA. Dans la structure 3D, la proximité de D7 et D2 permet de supposer qu'en temps normal D7 cache le domaine de liaison à UvrA situé en D2, pour éviter les liaisons improductives et la diminution de l'efficacité du NER. La reconnaissance par TRCF de RNAP bloquée changerait la conformation de la protéine et dévoilerait le domaine D2. Le domaine RID de TRCF est en interaction avec le domaine β de la RNAP, permettant de détecter son arrêt. TRCF suit de près la RNAP sur le duplex, et il a été montré qu'environ 25 nucléotides en amont de la polymérase sont nécessaires à la fixation de TRCF. Le domaine D1b est aussi en interaction avec ce domaine de la RNAP, et son changement de conformation par liaison à UvrA pourrait jouer sur la RNAP. Le détail de l'action de TRCF sur la RNAP n'est pas encore complétement compris (Deaconescu *et al.*, 2006; Deaconescu *et al.*, 2007).

2.3.2. Son rôle dans *E. coli*

Comme décrit précédemment, TRCF est liée à la RNAP et détecte quand celle-ci est bloquée. Grace à son domaine homologue à UvrB elle recrute UvrA et déclenche la voie de réparation du NER. D'autre part elle permet le détachement de la RNAP pour permettre l'accès aux enzymes impliqués dans la réparation à la lésion. Un autre facteur impliqué dans

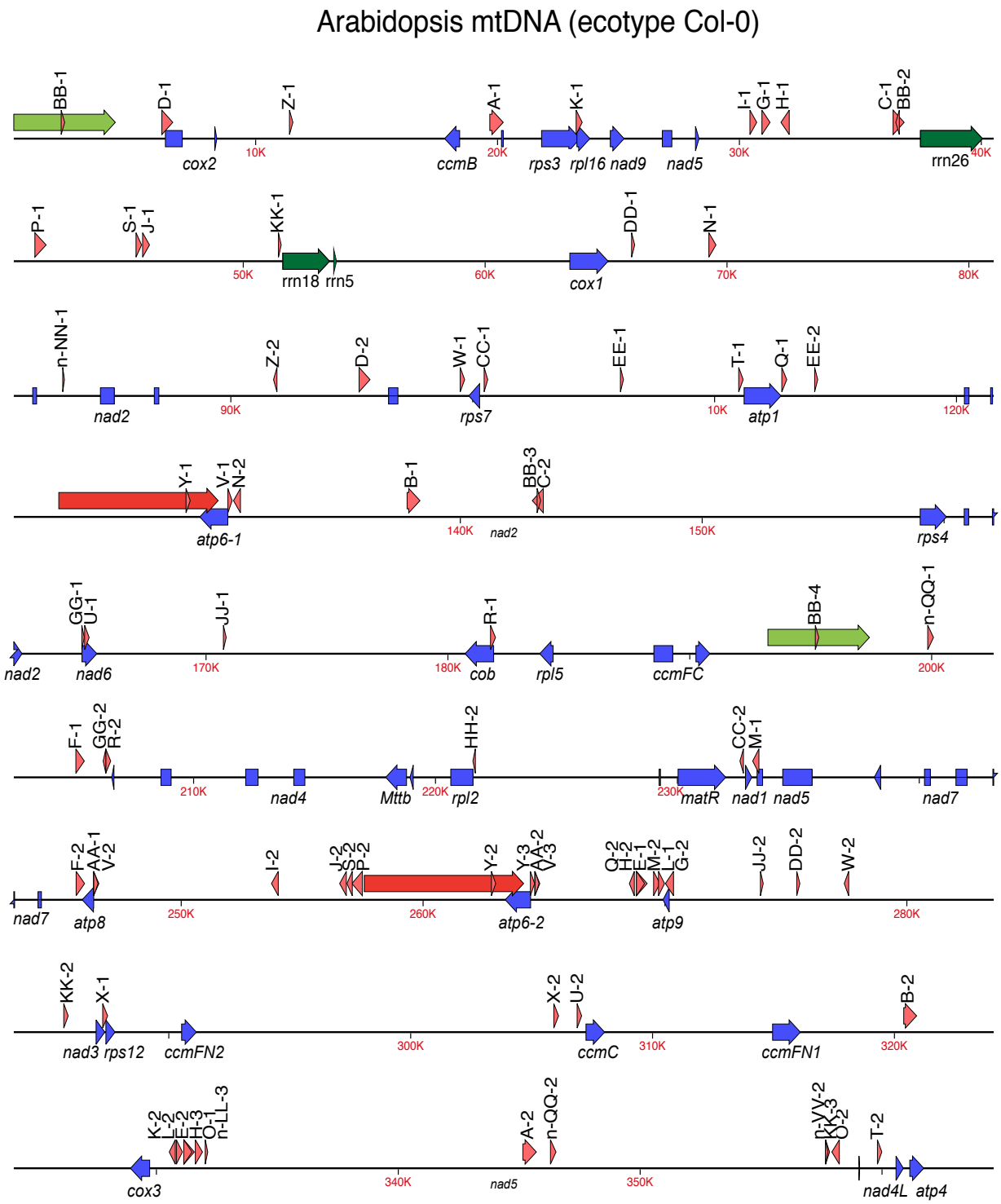


Figure 19. Carte du mtDNA d'Arabidopsis, ecotype Col-0. Les gènes sont représentés par des flèches bleues et vertes (rRNA) localisées sur la séquence. Les deux paire de grandes séquences répétées sont représentées au dessus de la séquence, en rouge et en vert. Les flèches roses représentent les paires de petites séquences répétées (IRs) de taille supérieur à 100 bb.

le "Transcription coupled repair" a été identifié récemment chez *E. coli* : le modulateur de transcription NusA (Deaconescu, 2013). Ce dernier a un rôle dans le recrutement d'enzymes du TC-TLS ("transcription-coupled translesion synthesis") impliquées dans la tolérance aux dommages ADN qui permettent de contourner la lésion (Cohen *et al.*, 2009). Elle recruterait également UvrA et mobiliserait ainsi les enzymes du NER, mais ne possèderait pas d'activité de dissociation de la RNAP. Cette voie de réparation interviendrait pour des lésions pouvant être facilement contournées pendant la réplication mais étant délétère pour la transcription. NusA permettrait le recul du complexe de réplication sans dissociation (Cohen et Walker, 2010). TRCF serait aussi l'un des facteurs permettant de moduler la RNAP dans la résolution des conflits entre réplication et transcription. Cette fonction est importante pour les bactéries où le génome est à la fois activement répliqué et transcrit, et où les deux processus se superposent (Trautinger *et al.*, 2005). TRCF est donc un facteur essentiel pour la réparation notamment des gènes très exprimés.

Les hélicases ont un rôle primordial dans de nombreux mécanismes et notamment la recombinaison. Certains facteurs mitochondriaux ont été identifiés ces dernières années, et parmi eux, de nombreux homologues à des protéines bactériennes. Cependant, avant le démarrage de ma thèse, aucune hélicase impliquée dans la réparation de l'ADN des organelles des plantes n'avait été identifiée.

3. L'ADN mitochondrial de plantes

Chez les eucaryotes photosynthétiques, les mitochondries et les chloroplastes sont essentiels pour la production d'énergie sous forme d'ATP qui alimente de nombreuses réactions du vivant. Selon la théorie endosymbiotique, les ancêtres de ces organelles seraient des bactéries endocytées par l'ancêtre de la cellule eucaryote. Au cours de l'évolution, de nombreux gènes mitochondriaux et chloroplastiques ont migré vers le noyau. Il ne reste maintenant dans les génomes de ces organelles que quelques gènes codant pour certaines des protéines de la chaine respiratoire ou photosynthétique, ainsi que pour des facteurs nécessaires à l'expression de ces gènes (Adams et Palmer, 2003). Cette centralisation des gènes dans le noyau représente un gain énergétique pour la cellule, lui permettant d'exercer une meilleure surveillance sur l'ensemble de l'ADN. Elle lui permet également d'éloigner l'ADN de potentiel ROS générés par les organelles. Mais cela implique aussi des mécanismes

Name	Start	End	Size	Orientation	Distance (kb)
Large-1-1	123400	129989	6589	+	127,6
Large-1-2	257567	264156	6589	+	
Large-2-1	1	4197	4196	+	189,0
Large-2-2	193231	197427	4196	+	
A-1	19680	20235	555	+	324,9
A-2	345153	345708	555	+	
B-1	137800	138332	532	+	182,1
B-2	320385	320917	532	+	
P-1	41386	41850	464	+	215,2
P-2	257058	257487	429	—	
C-1	36354	36810	456	+	106,2
C-2	142976	143431	455	—	
D-1	6122	6573	451	+	88,7
D-2	95303	95754	451	+	
E-1	268822	269262	440	+	61,9
E-2	331137	331566	429	+	
F-1	205132	205481	349	+	40,2
F-2	245656	246005	349	+	
H-1	31718	32059	341	—	236,8
H-2	268822	269157	335	+	62,0
H-3	331131	331472	341	+	
G-1	30931	31265	334	+	238,8
G-2	270019	270353	334	—	
O-1	331610	331905	295	+	26,0
O-2	357925	358235	310	—	
N-1	69251	69543	292	+	61,1
N-2	130624	130909	285	—	
J-1	45840	46124	284	+	210,4
J-2	256550	256826	276	—	
I-1	30435	30715	280	+	223,0
I-2	253732	254012	280	—	
M-1	233113	233371	258	—	36,2
M-2	269531	269788	257	+	
K-1	23250	23500	250	+	307,0
K-2	330531	330781	250	—	
L-1	269733	269981	248	+	60,8
L-2	330827	331075	248	+	
n-QQ-1	199838	200076	238	+	146,2
n-QQ-2	346292	346520	228	+	
S-1	45567	45796	229	+	211,1
S-2	256852	257072	220	—	

Name	Start	End	Size	Orientation	Distance (kb)
AA-1	246374	246596	222	+	18,0
AA-2	264617	264825	208	+	
R-1	181762	181967	205	+	24,4
R-2	206366	206571	205	+	
Q-1	112779	112983	204	+	155,5
Q-2	268528	268733	205	—	
X-1	287268	287471	203	+	18,4
X-2	305919	306114	195	+	
U-1	164981	165167	186	+	141,7
U-2	306882	307068	186	+	
T-1	110997	111181	184	+	248,6
T-2	359811	359991	180	+	
W-1	99483	99666	183	+	177,7
W-2	277413	277595	182	—	
n-VV-1	285658	285840	182	+	71,8
n-VV-2	357669	357830	161	+	
V-1	130391	130566	175	+	115,8
V-2	246384	246563	179	+	18,1
V-3	264627	264796	169	+	
Y-1	128659	128833	174	+	134,0
Y-2	262826	263000	174	+	1,4
Y-3	264436	264610	174	+	
BB-1	1960	2104	144	+	34,4
BB-2	36465	36611	146	—	106,6
BB-3	143175	143321	146	+	51,9
BB-4	195190	195334	144	+	
CC-1	100473	100620	147	+	132,0
CC-2	232586	232730	144	—	
Z-1	11398	11539	141	+	80,2
Z-2	91765	91906	141	—	
DD-1	66056	66183	127	+	209,3
DD-2	275450	275577	127	+	
EE-1	106106	106232	126	+	7,9
EE-2	114137	114263	126	+	
JJ-1	170719	170833	114	+	103,1
JJ-2	273948	274062	114	+	
KK-1	51456	51568	112	+	234,1
KK-2	285658	285761	103	+	
GG-1	164865	164965	100	+	41,3
GG-2	206262	206369	107	—	

Figure 20. Liste des IRs au dessus de 100 pb d'Arabidopsis, ecotype Col-0. Les IRs sont représentés selon leur nom, position dans le génome, taille, orientation et distance les uns des autres.

efficaces d'adressage des protéines synthétisées dans le cytosol et de leur import au sein des organelles.

De nombreuses protéines mitochondriales et chloroplastiques sont donc des homologues de protéines bactériennes. Cela est particulièrement vrai dans le cas de protéines impliquées dans le maintien et l'expression des génomes des organelles. A l'inverse, des facteurs d'origine eucaryote indispensables au maintien du génome, ont aussi été recrutés par les organelles, en parallèle de la perte probable des facteurs procaryotiques assurant les mêmes fonctions initialement.

Chez l'homme, des mutations de l'ADN mitochondrial (mtDNA) peuvent avoir de graves conséquences et donner des maladies héréditaires graves (myopaties, dégénérescences, etc). Alors que l'ADN des mammifères évolue très rapidement, le mtDNA des plantes est différent car il est plus stable en termes de séquences. Cependant sa plasticité en termes de structure est unique. En effet le mtDNA végétal comporte de nombreuses séquences répétées, susceptibles d'être impliquées dans des processus de recombinaison. Chez les plantes, le chloroplaste contribue lui aussi à la production d'énergie par photosynthèse. Cependant, les mitochondries de plantes ont une influence sur certains traits agronomiques comme la fertilité, la compatibilité de croisements et la vigueur de la plante. Ces génomes doivent donc être entretenus efficacement, c'est-à-dire répliqués et réparés correctement, pour pouvoir permettre une expression correcte des gènes. Au cours de ma thèse je me suis intéressée à la recombinaison et à la ségrégation du mtDNA d'Arabidopsis. Ce paragraphe sera donc consacré aux particularités de l'ADN mitochondrial des plantes, aux connaissances accumulées sur ce génome et aux protéines déjà identifiées nécessaires à son maintien.

3.1. Structure de l'ADN mitochondrial

Le mtDNA présente des spécificités que l'on ne trouve que dans le règne végétal. Chez les animaux, le mtDNA est d'une taille constante (entre 14 kb et 20 kb), alors qu'il est compris entre 200 et plus de 2500 kb chez les angiospermes (Kubo et Newton, 2008). Il atteint même une taille record de 11,3 Mb chez *Silene conica*.(Sloan *et al.*, 2012). Malgré ces différences de taille, le mtDNA des plantes ne contient pas beaucoup plus de gènes que le mtDNA des animaux. En effet on trouve en moyenne 50 gènes dans le mtDNA des plantes, et en moyenne 30 gènes dans le mtDNA des animaux. La différence de taille entre ces génomes s'explique donc par la présence de nombreuses séquences non codantes chez les plantes. Le

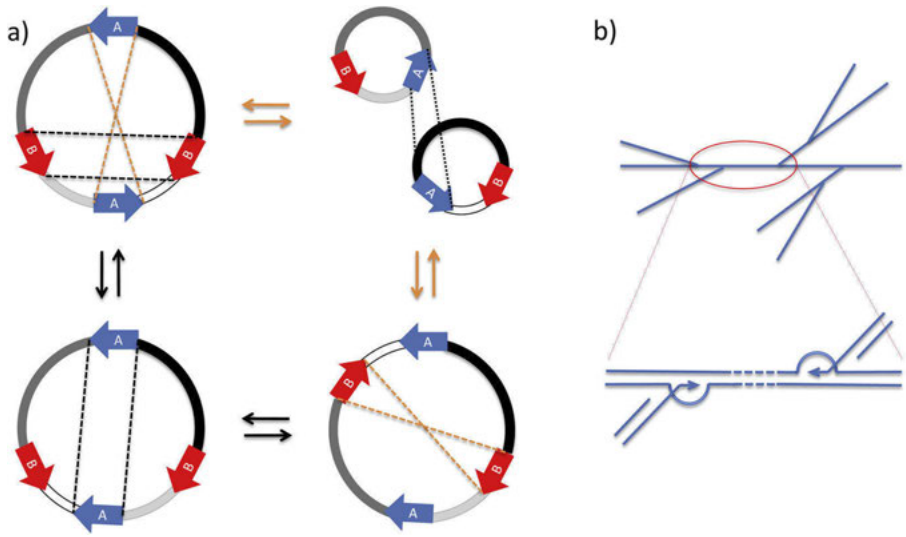


Figure 21. Modèle circulaire et linéaire du mtDNA.
a. Modèle structural du génome circulaire multipartite dans l'écotype C24, le grand cercle représente le "master circle", les différentes recombinaisons intramoléculaire qui peuvent avoir lieu (tirets), et celles qui forment les sous-génomes (orange). Les deux paire de grandes séquences répétées recombinant fréquemment (A et B) sont représentées par les flèches bleues et rouges. Les possible recombinaisons intermoléculaire sont représentées en pointillés.
b. Modèle linéaire branché du mtDNA. Les molécules branchées pourraient être des intermédiaires de réplication recombinaison-dépendante, dues à l'invasion de brin simple brin et à la formation d'une D-loop. D'après Gualberto *et al*., 2014.

premier mtDNA de plantes entièrement séquencé a été celui d'Arabidopsis (Unseld *et al.*, 1997). Il est composé de 57 gènes codant pour 32 protéines, 3 ARN ribosomiques (rRNA) et 22 ARN de transfert (tRNA) (figure 19). Mais la spécificité du génome mitochondrial des plantes vient aussi de la présence de nombreuses séquences répétées. Ces séquences répétées stimulent la recombinaison et font du mtDNA un génome très dynamique d'un point de vue structural.

Chez Arabidopsis accession C24, on trouve deux paires de grandes séquences répétées (6,5 et 4,2 kb) en sens direct et inverse (Klein *et al.*, 1994). La présence de ces deux paires de grandes séquences répétées est conservée dans d'autres accessions déjà séquencées, mais pas forcément leur orientation (Davila *et al.*, 2011). Le mtDNA d'Arabidopsis contient aussi de multiples séquences répétées de taille intermédiaire (en dessous de 1 kb) appelées IR (pour "Intermediate-size repeat") (figures 19 et 20). Chaque séquence peut non seulement recombiner avec son homologue à une fréquence plus ou moins élevée, selon la taille de la séquence répétée, mais aussi échanger ainsi des portions d'ADN. Les grandes séquences répétées se recombinent souvent et facilement, ce qui donne des sous-génomes résultants de ces réarrangements. D'une manière générale, le mtDNA a longtemps été considéré comme étant présent sous la forme d'un chromosome circulaire double brin appelé "master circle" et d'un ensemble de formes sous-génomiques, également circulaires issues de la recombinaison entre les grandes séquences répétées (figure 21) (Fauron *et al.*, 1990; Lonsdale *et al.*, 1984; Palmer et Shields, 1984). Le mtDNA des plantes est donc multipartite et peut exister sous la forme d'une population de structures alternatives. Cependant des études plus poussées en microscopie électronique, ou par électrophorèse en champ pulsé (PFGE pour "pulse-field gel electrophoresis") ont permis de préciser la structure du mtDNA (Oldenburg et Bendich, 1998). Le mtDNA serait présent sous forme d'un ensemble de molécules linéaires, circulaires et même branchées (figure 21), et sous différentes tailles (Backert et Borner, 2000; Oldenburg et Bendich, 1996). Ces structures complexes seraient des intermédiaires de recombinaison et leur existence pourrait suggérer des mécanismes de réplication du mtDNA recombinaison-dépendants (Backert et Borner, 2000; Bendich, 1993; Oldenburg et Bendich, 2001). La présence d'une proportion importante de ssDNA suggère également l'existence de mécanismes de recombinaison (Backert *et al.*, 1997).

Ce sont les données de cartographie par restriction et de séquençage qui ont servi à établir des cartes circulaires du mtDNA. Cependant, il est généralement admis que ces données peuvent aussi s'expliquer par la présence de concatémères, c'est à dire d'un assemblage de

séquences du mtDNA en tête à queue. Mais aussi sous forme de sous-génomes linéaires circulairement permutables, comme dans le cas de certains bactériophages. L'essor des technologies de séquençage a permis une approche plus précise des structures et confirmeraient l'existence de structures complexes, dont la forme et la composition varient selon les espèces (Mower *et al.*, 2012; Sloan, 2013).

Le séquençage du mtDNA de nombreuses plantes a également révélé que l'organisation structurale du mtDNA peut changer d'une espèce à l'autre. Chez certaines plantes, le génome est divisé en plusieurs chromosomes, qui ne sont pas des sous-génomes car ils ne présentent pas de séquences en commun entre eux, excepté quelques séquences répétées (Sloan, 2013). Par exemple, dans le concombre (*Cucumis sativus*) on trouve un chromosome de 1,6 Mb et deux autres chromosomes de 45 et 36 kb dont les séquences sont distinctes (Alverson *et al.*, 2011). Chez les silènes, on en trouve de 1 à 128 selon l'espèce (Sloan *et al.*, 2012). Cette organisation se distingue donc de la structure multipartite conventionnellement observée et montre encore la complexité des structures de mtDNA de plantes. Arabidopsis est considérée comme faisant partie des plantes dont l'organisation est plutôt classique et multipartite.

La présence de séquences répétées et la dynamique structurale du mtDNA augmentent la difficulté de compréhension des mécanismes nécessaires à son maintien. Cette dynamique participe aussi à son évolution.

3.2. Évolution et ségrégation du mtDNA

La plus grande particularité du mtDNA des angiospermes est le nombre de séquences répétées pouvant être impliquées dans des processus de recombinaison. Comme décrit en I.3.1, les séquences homologues de plus de 1 kb sont engagées dans des recombinaisons fréquentes et réciproques, participant ainsi à la structure multipartite du génome. Mais on trouve également de nombreuses séquences répétées de tailles intermédiaires (IR). Chez Arabidopsis on en compte environ 30 paires. Les IRs sont d'autant de possibilités de recombinaison et de remaniements de l'ADN. Toutefois, ces événements sont rares et conduisent à la création de versions alternatives du génome présentes en très peu de copies, souvent moins d'une par cellule (Andre *et al.*, 1992; Arrieta-Montiel *et al.*, 2009). Ces évènements de recombinaison rares sont quand même détectable dans les plantes sauvages par hybridation et par longue exposition des Southern (Small *et al.*, 1987), ou par PCR. Ils contribuent à l'hétéroplasmie du

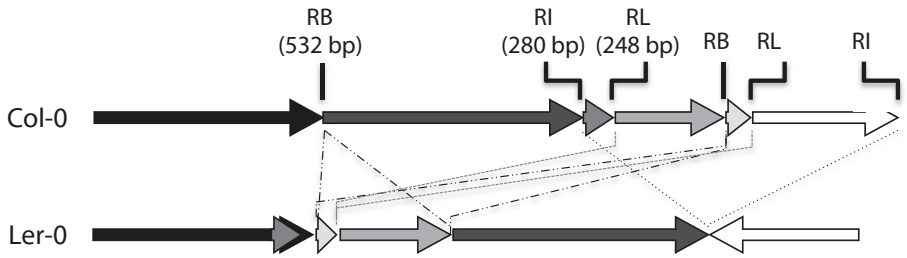


Figure 22. Evolution rapide du mtDNA entre différents écotypes d'Arabidopsis. Le mtDNA d'Arabidopsis ecotypes Col-0 et *Landsberg erecta* (Ler-0) (accession JF729201 et JF729202) sont représentés sous forme linéaire. Les interruptions de séquences coïncident la plupart du temps avec la présence d'IR (séquences RB, RL et RI), indiquant que l'évolution est conduite par la recombinaison entre IRs. D'après Kühn et Gualberto, 2012.

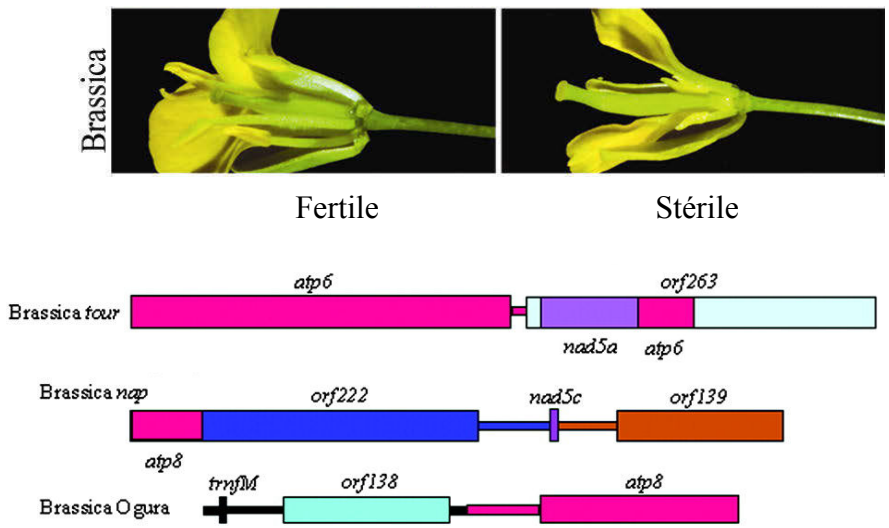


Figure 23. Exemple de CMS sur Brassica. En haut, comparaison entre une fleur male fertile avec anthères et pollen, et une fleur male stérile. En bas représentation des gènes chimériques responsable de CMS dans différentes espèces (*B. tournefortii, B. napus* et *B. ogura*). D'après Hanson et Bentolila, 2004.

mtDNA et les versions alternatives du génome peuvent participer à l'évolution du mtDNA (Small *et al.*, 1987).

Le séquençage du mtDNA de plusieurs variétés d'Arabidopsis a montré que certaines régions du mtDNA avaient été réorganisées via la recombinaison entre ces séquences (Arrieta-Montiel *et al.*, 2009; Davila *et al.*, 2011). Par exemple entre deux accessions d'Arabidopsis, Col-0 et *Landsberg erecta* (Ler-0) une région située entre deux IRs a été supprimée, et le génome a été réorganisé via des recombinaisons impliquant d'autres IRs (figure 22). Les processus de recombinaison sont aussi à l'origine de variétés CMS. Comme son nom l'indique, les plantes stériles mâles cytoplasmiques (CMS pour "cytoplasmic male sterility") ont un système reproductif mâle stérile (Budar et Pelletier, 2001). Cette stérilité peut venir d'une malformation ou absence de pollen, ou d'une immaturité des anthères par exemple. Ce phénotype CMS est très étudié car il apporte un avantage pour la sélection de semences en agronomie. De nombreux cas de CMS résultent de séquences partiellement dupliquées et réarrangées entre elles. Ces séquences forment des ORF fonctionnelles, et les protéines qu'elles codent ont un effet délétère sur les mitochondries et la cellule qui est seulement traduit dans un phénotype visible pendant la floraison (Hanson et Bentolila, 2004) (figure 23). Il a été observé dans la nature que certaines plantes deviennent CMS, ce qui confère un gain énergétique à la plante et favorise la reproduction hétérogame (Touzet et Meyer, 2014). Le séquençage du mtDNA de 5 écotypes de maïs dont 3 sont CMS a montré que les régions intergéniques évoluaient rapidement et que la différence de tailles entre ces génomes provenait de large délétions ou insertions (Allen *et al.*, 2007). Un scénario a également été proposé à partir de l'étude de variétés de maïs et de téosintes. Ces lignées auraient évoluées à partir d'un ancêtre commun par duplications en tandem suivis d'inversions (Darracq *et al.*, 2010).

Dans les tissus somatiques, des versions alternatives du mtDNA peuvent être continuellement générées par recombinaisons entre deux IRs d'une même paire. Elles sont toutefois maintenues dans des quantités bien moins élevées que leur version parentale, de 20 à plus de 100 fois dans une étude (Woloszynska et Trojanowski, 2009), et même plus de 1000 fois moins abondantes, selon la taille des IRs et leur positions dans le génome (résultats du laboratoire). Une de ces versions alternative sous-génomique peut parfois ségréger, être amplifiée et devenir alors la version majoritaire. Ce processus est appelé SSS pour "sub-stoichiometric shifting" (Small *et al.*, 1989). Ce changement peut se produire très rapidement d'un point de vue évolutif, voir en seulement une génération. Le SSS pourrait être une

conséquence d'une levée de contrôle sur la recombinaison entre IRs (recombinaison ectopique), et augmenterait donc le nombre de produits de recombinaison formant le génome alternatif. En effet plusieurs mutants des gènes mitochondriaux régulant la recombinaison ont montré de tels SSS (Abdelnoor *et al.*, 2003; Kuzmin *et al.*, 2005; Zaegel *et al.*, 2006). Les cassures double-brin de l'ADN (DSB) influenceraient directement l'évolution (Christensen, 2013; Davila *et al.*, 2011). En effet, la ligation directe de l'ADN au niveau des DSB par la voie du NHEJ et du MHMR peut mener à des insertions, des délétions et des réorganisations donnant lieu par exemple à des gènes chimériques comme ceux responsables de la CMS. En addition, la réplication du mtDNA est probablement recombinaison-dépendante, et les cassures double brin pourraient favoriser la réplication de séquences réparées par recombinaison mais par conséquence réarrangées. La réplication préférentielle des formes alternatives pourrait également entraîner un SSS. Mais les mécanismes et les facteurs qui les contrôlent sont loin d'être compris.

En effet les processus de réplication et de ségrégation du mtDNA et donc de sa transmission à la descendance sont encore loin d'être entièrement identifiés. Dans Arabidopsis on ne trouve qu'entre 40 à 60 copies du mtDNA par cellule, et en quantités comparables entre les feuilles et la cellule œuf qui transmet le génome mitochondrial (Preuten *et al.*, 2010; Wang *et al.*, 2010). C'est-à-dire moins d'une copie par mitochondrie. Il est donc difficile de modéliser comment un tel génome peut être transmis à la descendance en toute intégrité. Il est difficile d'étudier ces processus, car il n'y a pas encore de méthodes pour transformer le mtDNA des plantes pour tester les hypothèses. Pourtant, il serait intéressant de pouvoir comprendre et maitriser cette ségrégation, pour pouvoir comprendre les processus d'évolution, dans une optique de sélection pour une amélioration du végétal.

La recombinaison est donc au cœur des processus évolutif du mtDNA. C'est pourquoi des facteurs nucléaires qui sont impliqués dans le contrôle de la recombinaison et le maintien du génome, existent dans les mitochondries.

3.3. Les facteurs impliqués dans la recombinaison

Des facteurs régulant la recombinaison homologue ont été identifiés et certains sont homologues à des facteurs bactériens.

3.3.1. Recombinases RecA-like

Dans les mitochondries de plantes supérieures, on trouve deux protéines homologues à RecA, un facteur de recombinaison qui permet l'invasion de brin dans une région homologue chez les bactéries. Ces protéines, RECA2 et RECA3, sont adressées aux mitochondries et RECA2 est également adressée aux chloroplastes. RECA1 est une protéine seulement chloroplastique. Les mutants *recA3* et *recA2* montrent une augmentation de recombinaison entre les IRs, aggravé en particulier dans les plantes *recA2*. La mutation *recA2* est létale à un stade précoce du développement, car les plantes ne se développent pas au-delà du cotylédon, suggérant qu'en absence de RECA2 la structure du mtDNA est trop instable pour permettre la survie de la plante. La fonction exacte de ces deux RecA-like est encore inconnue mais elles semblent bel et bien impliquées dans la surveillance de la recombinaison. RecA de *E. coli* peut apparier des séquences homologues à partir de 25 pb. Cependant l'efficacité d'appariement augmente quand la taille des séquences homologues est supérieure à 200 bp (Lovett *et al.*, 2002). Cela est en accord avec les observations réalisées sur la fréquence de recombinaisons impliquant des IRs dans le mtDNA des plantes (Arrieta-Montiel *et al.*, 2009; Sloan *et al.*, 2012). Cependant la réplication du mtDNA ne semble pas compromise dans les mutants individuels *recA2* ou *recA3* (Miller-Messmer *et al.*, 2012; Shedge *et al.*, 2007). Mais comme le double mutant n'a pu être obtenu, il est possible qu'une fonction RecA soit nécessaire pour la réplication, et que RECA2 et RECA3 soient redondantes pour cette fonction.

Il a été montré dans Arabidopsis que les mutants *recA3* sont hypersensibles aux stress génotoxiques créant des DSB dans le mtDNA. Ces stress provoquent une augmentation de la recombinaison des IRs dans les plantes sauvage. Cet effet est perdu dans les mutants, suggérant que RECA3 est impliqué dans la réparation des DSB par recombinaison, comme son homologue bactérien (Miller-Messmer *et al.*, 2012).

3.3.2. Protéines liant le ssDNA

Le ssDNA est le substrat de nombreux mécanismes de recombinaison et de réparation. Le ssDNA étant vulnérable à la dégradation par des nucléases, les protéines liant le ssDNA sont donc indispensable pour sa protection, mais aussi pour promouvoir la recombinaison et la réplication ou détecter des dommages (Shereda *et al.*, 2008).

3.3.2.1. SSB de type bactérien

Plusieurs types de protéines se liant au ssDNA ont été identifiés dans les mitochondries végétales. Ainsi on trouve deux homologues à la protéine bactérienne SSB qui est responsable de la stabilisation du ssDNA, lors par exemple de la résection de brin avant que RecA ne se fixe sur l'extrémité 3' sortante. Appelées SSB1 et SSB2, elles sont toutes deux adressées aux mitochondries. La première a des fonctions identiques à la protéine bactérienne car il a été montré qu'elle a une grande affinité pour le ssDNA et stimule l'invasion de brin RecA-dépendante (Edmondson *et al.*, 2005). La deuxième semble essentielle car sa mutation est létale. Dans les bactéries, SSB est détachée du ssDNA à l'aide d'un autre facteur comme par exemple une translocase, ou alors par des médiateurs de recombinaison comme RecO ou Rad52. Cette fonction est proposée pour les protéines WHY2 et/ou ODB1 (Janicka *et al.*, 2012).

3.3.2.2. Protéines liant le ssDNA spécifiques aux organelles (OSB)

Un autre groupe de protéines spécifiques aux organelles possède une grande affinité avec le ssDNA (Vermel *et al.*, 2002). Les OSB (pour "organellar single-stranded DNA-binding") sont au nombre de quatre. OSB1, et 4 sont adressées aux mitochondries. OSB2 aux chloroplastes et OSB3 est adressée aux deux organelles (Zaegel *et al.*, 2006). Ces protéines possèdent un, deux ou trois domaines PDF, spécifique aux plantes, qui sont impliqué dans l'interaction avec l'ADN. Chez les bactéries, SSB peut inhiber la fixation de RecA sur le ssDNA. Une protéine partenaire, liant le ssDNA est souvent requise pour détacher SSB et fixer RecA. OSB pourrait remplir cette fonction, mais aucune observation dans ce sens n'a été rapportée. Les mutants *osb1* et *osb4* peuvent présenter un phénotype de développement avec des feuilles bigarrées, déformées, et une stérilité partielle, les mêmes phénotypes observés dans certaines plantes *recA3*. Ce phénotype serait causé par l'accumulation de produits de recombinaison entre les IRs. Il s'aggrave dans les générations suivantes, en corrélation avec un SSS par lequel certaine de ces formes recombinées deviennent prédominantes, suggérant un rôle de OSB1/OSB4 dans le contrôle de la stœchiométrie des formes alternatives du mtDNA (Kühn et Gualberto, 2012; Zaegel *et al.*, 2006).

3.3.2.3. Protéines de type Whirly (WHY)

Une protéine de type WHY (WHY2) est présente dans les mitochondries. Deux autres, WHY1 et WHY3, sont adressées aux chloroplastes. Les protéines WHY se lient préférentiellement au ssDNA, et sont impliquées dans le métabolisme de l'ADN des

organelles, ainsi que la régulation de la transcription dans le noyau (Desveaux *et al.*, 2005; Marechal et Brisson, 2010). Leur implication dans la surveillance de la recombinaison non-homologue a été montrée autant dans les plastes que dans la mitochondrie (Marechal *et al.*, 2008; Marechal *et al.*, 2009). Le mutant *why2* montre une augmentation de la recombinaison entre micro-homologies quand les plantes sont traitées à la ciprofloxacine. La ciprofloxacine bloque la gyrase, qui est adressée aux organelles, et provoque ainsi des DSB dans leurs génomes. Ainsi, il a été proposé que les protéines WHY agissent en se liant au ssDNA pour promouvoir la réparation de cassure double brin par la voie BIR. Sans la protéine ce serait la voie MMBIR ("micro-homology mediated break-induced replication") qui serait favorisée (Cappadocia *et al.*, 2010; Marechal *et al.*, 2009).

3.3.2.4. Homologue à Rad52 (ODB)

Chez Arabidopsis on trouve deux protéines contenant des similarités significatives avec Rad52, ODB1 qui est adressée à la mitochondrie et ODB2 aux chloroplastes (Janicka *et al.*, 2012; Samach *et al.*, 2011). Chez la levure, Rad52 lie les séquences de ssDNA qui interviennent dans la voie du SSA ou du MMEJ (McVey et Lee, 2008). ODB1 a été copurifié avec la protéine WHY2, et comme WHY2 elle est affine pour le ssDNA, mais se lie aussi au dsDNA. Les mutants *odb1* traités à la ciprofloxacine montrent un taux de recombinaison diminué par rapport au sauvage, suggérant un rôle d'ODB1 dans la réparation recombinaison-dépendante (Janicka *et al.*, 2012) au même titre que WHY2 (Cappadocia *et al.*, 2010). ODB1 présente également des similarités avec Mgm101 qui est impliquée dans la recombinaison du mtDNA de levure comme médiateur de la recombinaison (Mbantenkhu *et al.*, 2011). Les similarités d'ODB1 avec Rad52 et Mgm101 suggèrent qu'ODB1 a la même fonction chez les plantes dans la mitochondrie. La co-purification de WHY2 et ODB1 et leur similarité fonctionnelle suggèrent également qu'elles pourraient être redondantes, ce que pourrait expliquer l'absence de phénotypes visibles dans les mutants respectifs. Il n'a pas été possible d'obtenir de double mutant *why2 odb1* car chez Arabidopsis, les deux gènes ne sont qu'à 8 kb de distance sur le même chromosome. ODB1 est aussi un exemple de protéine bi-fonctionnelle, car elle affecte en plus l'épissage de deux introns de la mitochondrie (Gualberto *et al.*, 2015).

3.3.3. Homologue à MutS (MSH1)

Un autre facteur qui semble très important dans la surveillance de la recombinaison est MSH1, une protéine MutS-like homologue de Msh1 de la levure. MSH1 est adressé aux mitochondries et aux chloroplastes (Shedge *et al.*, 2007). Les mutants de ce gène présentent un mélange de deux phénotypes, d'une part de bigarrure, et plus rarement de déformation des feuilles et retard de croissance similaire aux mutants *recA3* et *osb1*. Mais le phénotype moléculaire correspond à un important réarrangement du mtDNA (Sakamoto *et al.*, 1996). Dans les bactéries, MutS est une protéine reconnaissant les mésappariements qui recrute les enzymes de la voie du MMR. MutS a également un rôle d'inhibiteur de la recombinaison (Schofield et Hsieh, 2003). Il n'y pas d'homologue à MutH chez les plantes, mais MSH1 contient un domaine GIY-YIG qui laisse supposer qu'elle assure le rôle de détection mais aussi d'endonucléase. Les mutants d'Arabidopsis *msh1* montrent une accumulation des produits de recombinaisons entre IRs (Arrieta-Montiel *et al.*, 2009; Shedge *et al.*, 2007). Bien souvent, cette recombinaison est asymétrique ce qui mène à une différence de concentration entre les deux formes recombinées. De plus, ce phénotype moléculaire s'aggrave au fil des générations, et les mutants homozygotes de générations avancées montrent une accumulation importante des produits de recombinaison (Arrieta-Montiel *et al.*, 2009). MSH1 assurerait un rôle de surveillance de la recombinaison en dissociant les hétéroduplexes lorsque leur homologie n'est pas assez importante (Abdelnoor *et al.*, 2006; Abdelnoor *et al.*, 2003).

3.3.4. ADN polymérases de type Pol I

Deux homologues à l'ADN Pol I bactérienne ont été identifiés. POL1A et POL1B. Les deux sont doublement adressées aux mitochondries et aux chloroplastes. La mutation d'une d'entre elle n'affecte pas les plantes, mais la double mutation est létale. Cela montre qu'elles ont surement un rôle redondant dans la réplication du mtDNA (Cupp et Nielsen, 2013; Parent *et al.*, 2011). L'insensibilité de *pol1A,* et la sensibilité du mutant *pol1B* à la ciprofloxacine, désignent POL1B comme la polymérase intervenant dans les réactions de réparation. Ni Pol1A ni Pol1B ne contiennent le domaine 5'-3' exonucléase qui existe chez la Pol I bactérienne. Cette fonction est probablement assurée par une autre protéine agissant en *trans.*

3.3.5. ADN Hélicases

Une hélicase homologue à l'hélicase Twinkle de la mitochondrie animale a été identifiée. Chez les animaux, Twinkle est l'hélicase réplicative qui a un rôle pour le déroulement des brins dans la fourche de réplication. Cette hélicase a comme ancêtre la protéine gp4 du phage

T7, qui combine la fonction d'hélicase et de primase. Contrairement aux animaux, les résidus responsable de l'activité primase sont conservés dans la TWINKLE des plantes (Shutt et Gray, 2006), et effectivement une activité primase a été mise en évidence *in vitro* (Diray-Arce *et al.*, 2013). Twinkle pourrait ainsi dérouler le brin et synthétiser les oligonucléotides amorces pour la réplication du brin retard. La TWINKLE d'Arabidopsis est adressée aux chloroplastes et aux mitochondries, et elle est essentielle, car le mutant est létal au stade embryonnaire (Carrie *et al.*, 2009).

3.3.6. Ligase

La ligase I (LIG1) est adressée au noyau et aux mitochondries dans Arabidopsis. Le gène donne deux protéines différentes selon le codon AUG utilisé pour démarrer la transcription, la plus longue contenant un signal d'adressage mitochondrial (Sunderland *et al.*, 2006).. Mais on la trouve également dans le protéome de nucléoides plastidiaux de maïs, suggérant également un adressage aux chloroplastes (Majeran *et al.*, 2012). L'ADN ligase I est nécessaire pour la réparation et la réplication de l'ADN nucléaire (Lohman *et al.*, 2011). Pour l'instant son rôle dans les mitochondries n'a pas encore étudié, mais se rapproche probablement de son rôle dans le noyau. Le mutant *lig1* est létal au stade embryonnaire, comme attendu pour une enzyme nécessaire au maintien des trois génomes de la cellule végétale

3.3.7. Topoisomérases

Les topoisomérases sont nécessaire à la réduction des forces de torsions dans l'ADN. Elles catalysent le clivage et la religation de l'ADN afin de faciliter son déroulement. Les topoisomérases de type I clivent un seul brin, alors que les topoisomérases de type II, appelée gyrase chez les procaryotes, génèrent des DSB (Champoux, 2001). Une de chaque type a été identifié dans les mitochondries de plantes. La gyrase est divisée en deux sous-unités, GYRA et GYRB. Un unique GyrA est doublement adressée aux mitochondries et aux chloroplastes. Trois GyrB existent dans Arabidopsis, et deux possèdent une séquence d'adressage, l'une aux mitochondries et l'autre aux chloroplastes (Wall *et al.*, 2004). Comme pour les autres gènes nécessaires à la réplication des génomes des organelles, les mutants de la gyrase ne sont pas viables.

3.3.8. Autres protéines

Une homologue à la RadA (Sms) bactérienne existe dans les mitochondries d'Arabidopsis. RadA est une protéine dont les fonctions sont mal connues. Sa mutation dans les bactéries n'a

pas beaucoup d'effet sur sa sensibilité aux stress génotoxiques, mais est synergique avec celle de RecG (Beam *et al.*, 2002; Cooper *et al.*, 2015). De plus, la perte de RadA combinée avec celle de RuvABC et RecG diminue la fréquence de HR dans *E. coli* (Beam *et al.*, 2002). Elle pourrait donc avoir un rôle dans la migration de fourche et agir dans une voie parallèle à celle de RecA. En effet, récemment il a été montré que RadA ne permet pas l'invasion de brin, mais peut promouvoir l'extension des hétéroduplxes pendant la recombinaison, même en absence de RecA (Cooper et Lovett, 2016). Des données non-publiées du laboratoire ont montré qu'un homologue à RadA existe dans tout le règne végétal. La protéine est adressée aux organelles et a un rôle dans la recombinaison et/ou ségrégation du mtDNA d'Arabidopsis.

3.4. Réplication du mtDNA des plantes

La réplication du mtDNA est loin d'être entièrement comprise. En effet la complexité structurale participe à la difficulté d'étudier et comprendre les mécanismes impliqués. Plusieurs facteurs nécessaires à la réplication ont été identifiés dans les mitochondries de plantes (voir I.3.3). En effet POL1A, POL1B, TWINKLE, TopoI, GyrA/B, LIG1 ainsi que les nombreuses protéines liant le ssDNA, constituent un ensemble de protéines pouvant intervenir dans la progression de la fourche de réplication.

Deux possibles origines de réplication ont été découvertes dans le mtDNA du pétunia (de Haas *et al.*, 1991), capables de promouvoir de la réplication *in vitro* à partir d'extraits mitochondriaux. Les auteurs ont proposé à l'époque un modèle de réplication similaire à celui du mtDNA des mammifères : chacune des origines de réplication est située sur un brin et ces derniers sont répliqués indépendamment, chacun dans un sens différent. Cependant cette forme de réplication n'est pas applicable à la structure multipartite du génome mitochondrial. Des analyses par PFGE, permettant de visionner des grandes molécules d'ADN, ont permis de détecter des formes de mtDNA linéaires et circulaires en concatémères. C'est-à-dire des formes ou le génome est dupliqué plusieurs fois à la suite (Oldenburg et Bendich, 2001). Ces concatémères suggèrent une réplication de type "rolling-circle" ou bien une réplication recombinaison-dépendante (RMR pour "recombinaison-mediated replication") similaire à celle du phage T4.

La réplication de type "rolling-circle" est initiée soit par la création d'une coupure simple brin sur une molécule d'ADN circulaire soit par la formation d'une D-Loop. La synthèse d'ADN s'initie donc à cet endroit en partant de l'extrémité 3' et sur un seul brin (figure 24).

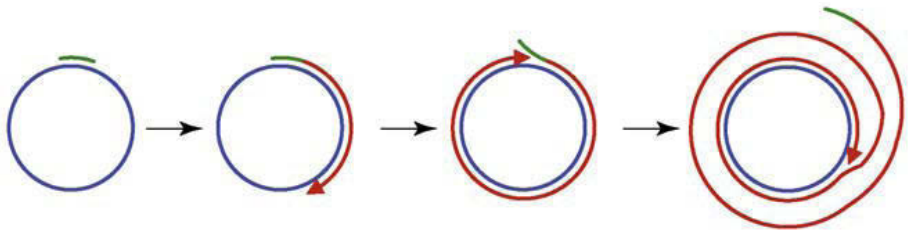


Figure 24. Principe de l'amplification par rolling-circle d'une molécule simple brin. La polymérase synthétise un brin complémentaire (en rouge). Après un tour, l'amorce (en vert) et le nouveau brin sont déplacés et la synthèse est continuée sur plusieurs tours. La conséquence est la synthèse d'une molécule concatémèrique. D'après Johne *et al.*, 2009.

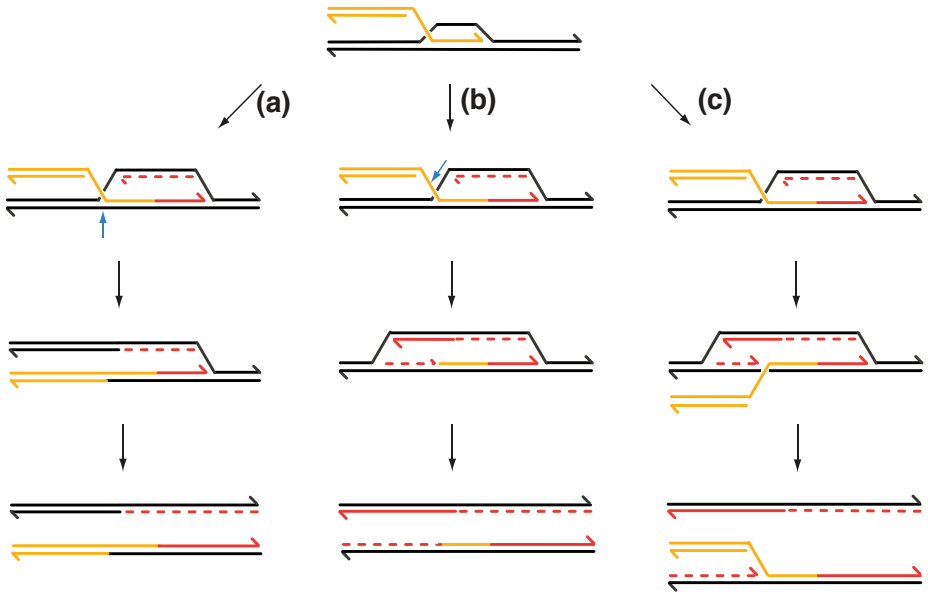


Figure 25. Modèles uni-, bi- et tri-directionnels de la réplication recombinaison-dépendante du phage T4. Le clivage de l'un des brins à l'extrémité de la D-loop peut générer une réplication unidirectionnelle **(a)** ou bidirectionnelle **(b)**.
(a) Réplication unidirectionnelle, la structure D-loop est résolue par le clivage du brin du bas (flèche bleue). Alternativement, une HJ peut se former et translater le long du brin.
(b) Réplication bidirectionnelle, le brin invasif est clivé à l'extrémité de la D-loop, et devient une amorce pour initier deux fourches de réplication en sens opposés.
(c) Réplication tri-directionnelle. Aucune nucléase n'intervient dans cette voie, et deux fourches de réplications sont initialement mises en place comme dans (b). Une structure en Y est donc laissée en place sur l'un des brins. Une nouvelle fourche peut alors être mise en place à cet endroit. Le brin invasif est représenté en jaune, les duplexes homologues en noir et l'ADN nouvellement synthétisé en rouge. Les brins précoce sont représentés en ligne continue et les brin retards en pointillés. La pointe de flèche représente l'extrémité 3'. D'après Kreuzer, 2000.

Pendant le déplacement du complexe le long de l'ADN parental, le brin complémentaire au brin matrice se déroule pour laisser place au brin néoformé. Un complexe de réplication vient ensuite se mettre en place sur ce ssDNA pour permettre la synthèse de fragments d'Okazaki qui formeront le brin complémentaire (Johne *et al.*, 2009). L'ADN néoformé est donc un concatémère des copies de l'ADN. Par microscopie électronique des structures en forme de sigma (σ) ont été observées (Backert *et al.*, 1996). Ces structures sont caractéristiques de la réplication de type rolling-circle car elles sont un intermédiaire de la réplication. Ces observations étayent donc la théorie du rolling-circle.

La réplication de type phage T4 est une réplication de type BIR où une D-loop se met en place, et où la synthèse s'effectue en prenant pour modèle chacun des brins (figure 25). Plusieurs scénarios sont possibles pour la suite selon le sens de clivage de la HJ. L'action de nucléase sur la HJ peut entraîner un échange de brin par religation. Elle peut aussi entraîner le clivage de la partie qui n'est pas engagée dans la D-Loop. La partie homologue du duplex invasif sert alors de primase. S'il n'y a pas d'action de nucléases on peut observer des structures d'ADN branché (Kreuzer, 2000). Formes qui ont justement été observées dans le génome mitochondrial par microscopie électronique. De plus aucune nucléase pouvant cliver les HJ n'a encore été identifiée dans les mitochondries de plantes. La réplication du génome mitochondrial de plante dépend probablement des caractéristiques du mtDNA de l'espèce concernée. Le mtDNA linéaire est probablement pris en charge par la RDR, alors que les sous-génomes circulaires générés par recombinaison sont probablement répliqués par le mécanisme du rolling-circle (Backert *et al.*, 1996). Des analyses dans des cultures de cellules ont confirmé la présence de réplication de type RDR (Backert et Borner, 2000). Shedge et ses collaborateurs (Shedge *et al.*, 2007) ont proposé un modèle où la réplication s'initie dans les grandes séquences répétées, à l'instar de la réplication du mtDNA de la levure *Candida albicans* (Gerhold *et al.*, 2010). Cependant ces mécanismes n'expliquent pas comment un génome si complexe peut être transmis à la descendance. La réplication des génomes complexes eucaryotes comporte des points de contrôle pour permettre la réplication du génome en entier. Chez les bactéries, les zones d'initiation et de terminaisons contrôlent la réplication stœchiométrique du génome. Finalement, chez des bactériophages type T4 c'est le processus d'empaquetage qui permet la ségrégation de génomes entiers. Pour l'instant, aucun point de contrôle ou zone de terminaison n'ont été identifiés dans le mtDNA de plantes. Ces mécanismes de réplication et de son contrôle sont encore inconnus dans les mitochondries de

Mécanismes	Function moléculaire	Catégorie	*E. coli*	*A. thaliana*	Adressage
Réplication	Synthèse d'ADN	ADN polymérase	PolIII	POL1A POL1B	Mt + Cp Mt + Cp
	Déroulement de brin	Hélicase	DnaB	TWINKLE	Mt + Cp
	Synthèse amorces	Primase	DnaG		Mt + Cp
	ligation d'ADN	ADN ligase	LigA	LIG1	nu + Mt + Cp
	Réduction des torsions	Gyrase Topoisomérase I	GyrA GyrB TopoI	GYRIIA GYRIIB2 TOPOI	Mt + Cp Mt Mt + Cp
	stabilisation ssDNA	ssDNA-binding	SSB	SSB1 SSB2	Mt** Mt**
Réparation					
Réversion directe du dommage (DDR)	Conversion des CPD	Photolyase	PhrA	CPD	nu + Mt + Cp
Par excision de base (BER)	Excision 3-meA Excision de U Excision de U ou U,T,C apparié Excision de A apparié à 8-oxoG Excision de 8-oxG apparié à C	ADN glycosylase	AlkA Ung Mug MutY Fpg	– UNG – – FPG OGG1	– Mt – – nu* nu + Cp*
	Prise en charge des sites AP	AP endonucléase	EndoIV EndoIII	– APE1L APE2	– Cp nu* + Mt*
	Synthèse d'ADN	ADN polymérase Ligase	PolI LigA	POL1B LIG1	Mt + Cp Mt + noyau
Par excision de nucléotide (NER)	Détection de l'ADN endommagé Catalyse la mise en place du Nucléase 3' et 5' Détachement de la portion Facteur de réparation associé à	DNA-binding Hélicase Endonucléase Hélicase Translocase	UvrA UvrB UvrC UvrD TRCF	– – – – TRCF	– – – – Mt + Cp
Mésappariements (MMR)	Reconnaissance du mésappariement	DNA-binding	MutS	MSH1	Mt + Cp
	Stabilisation du complexe MutS-ADN	ATPase	MutL	–	–
	Excision de brin et synthèse d'ADN	Endonucléase	MutH	MSH1 ?	Mt + Cp
	Exonucléase	Exonucléase 5'-3' Exonucléase 3'-5' Exonucléase 5'-3' Exonucléase 3'-5'	RecJ ExoI ExoVII ExoX	–	–
	stabilisation ssDNA	ssDNA binding	SSB	SSB1 SSB2	Mt** Mt**
	Déroulement de brin	Hélicase	UvrD		
Recombinaison Homologue	Résection de brin	Hélicase 3'-5'/nucléase Hélicase/nucléase (Chi) Hélicase 5'-3'/nucléase Nucléase 5'-3' ATPase Hélicase	RecB RecC RecD RecJ RecF RecQ	– Plusieurs	– nu*
	Formation du filament RecA-ssDNA	ssDNA-binding (médiateur de recombinaison)	RecO	ODB1(Rad52) WHY2 OSB1 OSB3 OSB4	Mt + noyau? Mt Mt Mt + Cp Mt**
		DNA-binding	RecR	–	–
		ssDNA-binding	SSB	SSB1 SSB2	Mt** Mt**
	Formation de D-Loop	Recombinase	RecA	RECA2 RECA3	Mt + Cp Mt
	Migration de jonctions de Holliday	Complexe avec RuvB Hélicase Hélicase ATPase RecA-like Hélicase	RuvA RuvB RecG RadA/Sms RecQ	– – RECG1 RADA Plusieurs	– – Mt + Cp Mt + Cp** nu*
	Résolution de jonctions de Holliday	Résolvase	RuvC RusA	– –	– –
	Supresseur de la recombinaison	Hélicase	UvrD/Srs2	RECG1	Mt + Cp
	Sauvetage des fourches de réplication	Complexe avec RuvB Hélicase Hélicase Recombinase ATPase RecA-like	RuvA RuvB RecG RecA RadA/Sms	– – RECG1 RECA2 RECA3 RADA	– – Mt + Cp Mt + Cp Mt Mt + Cp**
	Synthèse d'ADN ligation d'ADN	ADN polymérase Ligase	PolI LigA	POL1B LIG1	Mt + Cp nu + Mt + Cp

Figure 26. Tableau récapitulatif de différents facteurs bactériens impliqués dans les mécanismes de réplication/réparation et leurs équivalents dans les mitochondries d'Arabidopsis. Mt: mitochondries, Cp: chloroplastes, nu: noyau. *localisation prédite par SUB3. **résultats expérimentaux non-publiés du laboratoire.

plantes. Les mécanismes de réplication du mtDNA pourraient aussi dépendre du stade de développement de la plante.

3.5. Réparation du mtDNA des plantes

L'ADN des organelles est plus exposé aux dommages oxydatifs car physiquement proche des ROS produits par leur métabolisme. Pourtant, le mtDNA des plantes est malgré tout très stable en termes de mutation par substitution (Palmer et Herbon, 1988), sauf quelques exceptions bien caractérisées (Cho *et al.*, 2004; Zhu *et al.*, 2014). Les mécanismes de réparation du mtDNA sont donc efficaces et probablement liés aux activités de recombinaison qui permettent la correction de copie. En addition, il a été montré que la voie courte du BER est active dans les mitochondries de plantes. Et plusieurs facteurs ont été identifiés (figure 26). A ce jour, aucune preuve de l'existence du NER n'a été mise à jour dans les organelles de plantes, et aucun facteur n'a encore été identifié.

3.5.1. Réparation par la voie courte du BER

Le BER est l'une des voies de réparation majeure des lésions causées par le stress oxydatifs. Dans les mitochondries de plantes, plusieurs facteurs impliqués dans le BER ont été identifiés. Des glycosylases ont été trouvées dans les organelles de plantes. Une UDG, clivant les uraciles, a été détectée dans le maïs, et son activité a été démontrée *in organello* dans la pomme de terre et Arabidopsis (Boesch *et al.*, 2009; Cordoba-Canero *et al.*, 2010). Cette enzyme ne présente pas d'activité lyase et semble agir uniquement dans la voie du short-patch BER, car seul un nucléotide est remplacé. Dans les plantes, deux glycosylases peuvent prendre en charge les lésions 8-oxoG, OGG1 et FPG. OGG1 est ubiquitaire chez les eucaryotes, et FPG (ou MutM) ne se trouve que chez les bactéries et les plantes. Par fusion à la GFP il a été montré qu'OGG1 est située dans le noyau (Chen *et al.*, 2012). Mais elle pourrait potentiellement être aussi envoyée dans les mitochondries et les chloroplastes (Macovei *et al.*, 2011). L'un ou plusieurs des 7 isoformes de FPG pourrait également être adressé aux mitochondries, car une double mutation d'OGG1 et FPG entraîne une accumulation de mutations dans l'ADN nucléaire, mais aussi dans le mtDNA (Cordoba-Canero *et al.*, 2014).

La deuxième étape du BER fait intervenir une AP endonucléase. Dans Arabidopsis, on en trouve 3 différentes. Des expériences de fusion à la GFP ont montré qu'ARP est adressée aux chloroplastes (Gutman et Niyogi, 2009). APE1L a été trouvé dans le protéome du

chloroplaste d'Arabidopsis, mais elle est aussi prédite comme étant située dans le noyau (Kleffmann *et al.*, 2004). Enfin APE2 possède deux isoformes provenant d'un épissage alternatif, et le plus long a une prédiction d'adressage aux mitochondries. Les polymérases POL1A ou POL1B pourraient ensuite intervenir pour synthétiser le nucléotide manquant, mais cela n'a pas été montré. Pour la dernière étape, une ligase est nécessaire pour fermer la molécule d'ADN. Dans le noyau d'Arabidopsis c'est la ligase I (LIG1) qui intervient dans le short-patch et le long-patch du BER. Certaines pièces sont encore manquantes mais un bon nombre de facteurs du BER semblent être présents dans les mitochondries. Il a été montré chez la levure que les facteurs du BER sont envoyés dans les compartiments selon les stress et les besoins en réparation (Griffiths *et al.*, 2009; Swartzlander *et al.*, 2010). Il est possible que ce même système soit aussi utilisé dans les plantes, ce qui permettrait une optimisation de l'utilisation des facteurs de réparation selon les besoins de la cellule. Mais cela reste à être démontré.

3.5.2. Mécanisme du MMR

La présence de MSH1, un homologue de MutS, pose la question de la présence du MMR dans les mitochondries de plantes. Pour l'instant cette protéine est la seule identifiée dans cette voie, mais elle possède un domaine qui pourrait lui conférer une activité d'endonucléase. Cependant, son rôle a plutôt été caractérisé dans la déstabilisation d'intermédiaires de recombinaison lorsque l'homologie n'est pas assez importante (Abdelnoor *et al.*, 2006; Abdelnoor *et al.*, 2003; Davila *et al.*, 2011; Xu *et al.*, 2011). L'existence de MMR reste donc incertaine.

3.5.3. Réparation par Recombinaison Homologue

Les séquences répétées sont d'autant de possibilités de recombinaison lors de la réparation des DSB par HR. Cependant les recombinaisons ectopiques entraînent une modification du mtDNA et sont finement contrôlées. La mutation de certains des facteurs qui interviennent dans la recombinaison mitochondriale augmente la fréquence des recombinaisons ectopiques. C'est le cas des OSB1 et 4, des RECA2 et 3 et de MSH1 (Arrieta-Montiel *et al.*, 2009; Miller-Messmer *et al.*, 2012; Zaegel *et al.*, 2006). Certains facteurs contrôlent également la réparation par recombinaison. La ciprofloxacine est souvent utilisée pour étudier les mécanismes de réparation dans les organelles. En effet cette substance bloque les gyrases d'origine bactérienne dans une conformation qui entraîne des DSB. Dans les plantes la gyrase ne se trouve que dans les organelles et son utilisation cible uniquement ces

génomes, sans affecter l'ADN nucléaire. Dans les plantes sauvages, la réponse à ce traitement se manifeste par l'augmentation de recombinaisons entre les IRs dans le mtDNA. Cela montre que la réparation induite pour réparer les DSB fait intervenir la recombinaison homologue. Dans le contexte du mtDNA riche en IRs, il en résulte l'augmentation de la recombinaison ectopique impliquant ces IRs. Ces expériences montrent aussi qu'en conditions de croissance normales ces voies de recombinaison sont inhibées. Dans les mutants de RECA3, ODB1, WHY2 et POL1B, la réponse aux DSB par HR est diminuée, suggérant que ces facteurs ont un rôle dans la recombinaison associée à la réparation (Cappadocia *et al.*, 2010; Cupp et Nielsen, 2013; Janicka *et al.*, 2012; Miller-Messmer *et al.*, 2012). Dans les mutants *odb1* et *why2*, la voie de réparation par recombinaison via des micro-homologies (MHMR) est favorisée quand ils sont exposés à des stress génotoxiques. Ces observations suggèrent qu'un mécanisme alternatif de réparation par MHMR existe dans les mitochondries de plante est qu'il est défavorisé si les processus de recombinaison homologue sont actifs (Cappadocia *et al.*, 2010; Janicka *et al.*, 2012). Le cas de RECA3 semble contradictoire. En effet cette protéine inhibe la recombinaison ectopique d'un côté, mais de l'autre favorise la recombinaison pour la réparation (Miller-Messmer *et al.*, 2012). Une voie de réparation RecA-dépendante existe donc dans les plantes. Mais l'augmentation de produits de recombinaison dans des conditions normales de culture chez *recA3* suggère qu'il existe aussi une voie de recombinaison alternative RecA-indépendante, comme la voie du SSA ou du MMEJ.

Ainsi un modèle des voies de réparation, basé sur ces observations a été proposé (Gualberto *et al.*, 2014). Les DSB sont prises en charge par une voie de réparation RecA-dépendante favorisée par l'invasion de brin catalysée par les RECA2, 3, ODB1 et WHY2. Ces facteurs inhibent la recombinaison entre micro-homologies, susceptible d'insérer des erreurs. Cette invasion de brin permet une réparation par la voie du BIR. La voie du SDSA est probablement active pour limiter les crossover. D'un autre côté, les protéines RECA2, 3, OSB1, 4 et MSH1 inhibent la recombinaison RecA-indépendante. Cela a pour effet d'inhiber la voie de réparation par SSA qui introduit des larges délétions, ce qui est source d'erreurs. Mais il est aussi possible que les fonctions de RecA dans la réparation concernent surtout les mécanismes de réparation couplées à la réplication, comme discuté dans le chapitre 1.2.3.2.2. A ce jour, aucune hélicase permettant la régression des fourches de réplication n'a encore été identifiée. On a donc peu d'informations sur les mécanismes de sauvetage des fourches de réplication.

4. Objectifs de thèse

Contrairement aux mitochondries animales, le mtDNA des plantes, se caractérise par les nombreuses activités de recombinaison qui modulent sa structure. Ces activités de recombinaison sont impliquées à la fois dans la réparation et la réplication, et mais aussi dans l'évolution rapide de la structure du mtDNA. Le "shifting" sous-stœchiométrique qui contribue à l'évolution des génomes est favorisé par le taux d'hétéroplasmie du mtDNA. Plus il y a des versions différentes du mtDNA, plus l'une d'entre elles a des chances d'être ségrégée et de devenir majoritaire. Les activités de recombinaison ectopique ou illégitime, comme celles entre IRs, contribuent significativement à l'hétéroplasmie du mtDNA et sont donc finement contrôlées. Plusieurs types de protéines comme des recombinases, des polymérases ou des protéines liant le ssDNA ont été décrites dans la recombinaison du mtDNA. Mais encore plusieurs facteurs sont manquants pour reconstituer les mécanismes.

Nous avons suivi la piste des hélicases dont le rôle important dans les mécanismes de recombinaison était connu. En effet les hélicases UvrD, Rep, RecQ et RecG chez *E. coli*, Srs2 chez la levure et BLM chez l'homme sont essentielles au contrôle de la recombinaison et peuvent même agir comme des anti-recombinases. Dans les mitochondries des plantes, à l'exception de l'hélicase réplicative TWINKLE, aucune autre ADN hélicase n'avait encore été identifiée. TWINKLE pourrait éventuellement avoir une double fonction, dans la réplication et recombinaison, comme récemment proposé pour la TWINKLE humaine (Sen *et al.*, 2016). Mais une telle fonction de la TWINKLE des plantes n'était pas soupçonnée au moment où j'ai commencé ma thèse.

Dans ce contexte, les objectifs de ma thèse étaient:

- De mieux comprendre les mécanismes de recombinaison intervenant dans les mitochondries de plantes pour pouvoir manipuler le génome mitochondrial
- D'identifier les facteurs manquants, et notamment les hélicases pour mieux comprendre les mécanismes de recombinaison, de réparation et d'évolution du mtDNA de plantes.

Pour atteindre ces objectifs, j'ai utilisé deux approches. La première est basée sur un système de recombinaison *in organello*. Ce système devrait permettre de déterminer les facteurs qui influencent l'efficacité de recombinaison. Cette approche est également intéressante dans l'optique de manipuler le mtDNA et d'obtenir un système de transformation

des mitochondries. La deuxième approche consiste en l'identification de protéines mitochondriales et en l'étude de mutants pour comprendre les mécanismes dans lesquelles elles sont impliquées. Je me suis intéressée aux hélicases mitochondriales qui pourraient avoir des rôles essentiels dans le contrôle de la recombinaison et la ségrégation du mtDNA.

Résultats

1. Import d'ADN et recombinaison *in organello*

1.1. Problématique

Dans le contexte actuel, la transformation des mitochondries semblerait une approche évidente pour étudier le fonctionnement génétique des organelles et maîtriser la CMS. Les premiers essais de manipulations génétiques des mitochondries ont eu lieu dans les années 80. Clark et Shay ont réussi à importer des mitochondries purifiées portant un gène de résistance à un antibiotique dans des cellules de mammifère (Clark et Shay, 1982). Des essais de transformation conventionnelle des mitochondries dans les cellules entières, comme cela se fait pour les chloroplastes, ont été effectués par la méthode biolistique. Cette approche consiste à bombarder les cellules avec des particules d'or ou de tungstène recouvertes d'ADN. Elle permet l'expression de ces constructions dans les cellules transfectées. Mais à une fréquence réduite, ces constructions peuvent être insérées dans le génome par recombinaison et transmises à la descendance, si un protocole de sélection adéquat est disponible. L'approche n'a été couronnée de succès que dans le cas de *Saccharomyces cerevisiae*, *Chlamydomonas reinhardtii* et *Candida glabrata* (Bonnefoy *et al.*, 2007; Zhou *et al.*, 2010).

Il a été montré que les mitochondries isolées étaient capables d'importer de l'ADN (Koulintchenko *et al.*, 2003; Koulintchenko *et al.*, 2006). Des études portent donc sur l'apport d'ADN à l'intérieur de la cellule, jusqu'à la mitochondrie. Des vésicules mitochondriotropiques comme les DQasomes ou les liposomes, capables de relarguer de l'ADN au contact des organelles dans des cellules de mammifères sont à l'étude (Lyrawati *et al.*, 2011; Wagle *et al.*, 2011). Les MITO-Porter quant à eux sont des vésicules entourées de deux membranes, l'une pouvant fusionner avec la membrane plasmique, puis l'autre pouvant fusionner avec celle des mitochondries (Yamada et Harashima, 2013). Seul le transport de protéines dans des cellules animales a été testé pour l'instant par cette approche. Si les recherches sur le transfert d'ADN jusque dans la mitochondrie sont en cours, les modalités de recombinaison et les mécanismes au niveau du mtDNA sont encore mal connus. C'est pourquoi des études sur la recombinaison après import dans des mitochondries isolées permettraient d'en savoir plus sur le devenir de l'ADN importé et des approches favorisant son insertion dans le mtDNA par recombinaison.

Le but était de définir des constructions d'ADN qui pourraient être maintenues dans les mitochondries par recombinaison ou réplication et remplir les fonctions attendues. Nous avons profité pour ces études de la possibilité d'importer l'ADN dans les mitochondries

isolées. L'objectif à long terme serait d'intégrer un gène responsable de la CMS, pour directement obtenir des lignées mâle-stériles. Pour cela il est indispensable de pouvoir exercer une pression de sélection, comme la résistance à un antibiotique, afin de promouvoir une transformation mitochondriale homoplasmique. Le but est également d'obtenir un transcrit fonctionnel, et pour cela d'avoir une unité fonctionnelle contenant un promoteur fort d'un gène mitochondrial. Il a été montré que l'import d'ADN dans les mitochondries dépend de la séquence pour les gros fragments (>7 kb) (Ibrahim *et al.*, 2011).

1.2. Prérequis

L'étude de ce système de transfection *in organello* dans la mitochondrie a montré que l'intégration d'une séquence d'ADN exogène était possible par cette méthode, dans des mitochondries de pomme de terre (Mileshina *et al.*, 2011). Mais la séquence du génome mitochondrial de la pomme de terre n'a pas encore été assemblée et il est donc difficile d'aller plus loin dans l'étude de cette technique. Nous avons donc trouvé intéressant de passer au modèle Arabidopsis dans lequel nous avons de nombreux outils génétiques à notre disposition. Le but de notre projet n'était pas de redémontrer le travail déjà publié par Mileshina, mais aussi de tester l'effet de différents facteurs mitochondriaux de la recombinaison, déjà identifiés, dont RECG1, pour lesquelles des mutants Arabidopsis sont disponibles. Le test de ce système dans les mitochondries de mutants pourrait nos donner un aperçu des mécanismes impliqués dans la recombinaison.

1.3. Tests des différentes sortes de mitochondries

Pour les tests d'import et de recombinaison *in organello*, il est nécessaire d'avoir des mitochondries fonctionnelles, pures et en grandes quantités. Les tubercules de pommes de terre sont un matériel adapté car il est riche en mitochondries et permet de les purifier en quantités importantes et de très bonne qualité fonctionnelle. Au contraire, il est plus difficile d'obtenir des mitochondries fonctionnelles à partir d'Arabidopsis. J'ai donc voulu tester quel est le matériel végétal d'Arabidopsis permettant d'obtenir le maximum de rendement. Pour cela, j'ai extrait des mitochondries à partir de différents types de tissus d'Arabidopsis, et j'ai mesuré le contrôle respiratoire de chacun des extraits. J'ai testé des rosettes entières cultivées en terre, des plantules de 3 semaines cultivées *in vitro*, et de cellules vertes d'une suspension cultivée *in vitro*. Les mesures à l'électrode d'oxygène ont montré que les mitochondries de chacune des préparations consomment de l'oxygène, ce qui signifie qu'elles respirent. Cependant, les mitochondries issues à partir de plantes entières n'ont pas montré de contrôle

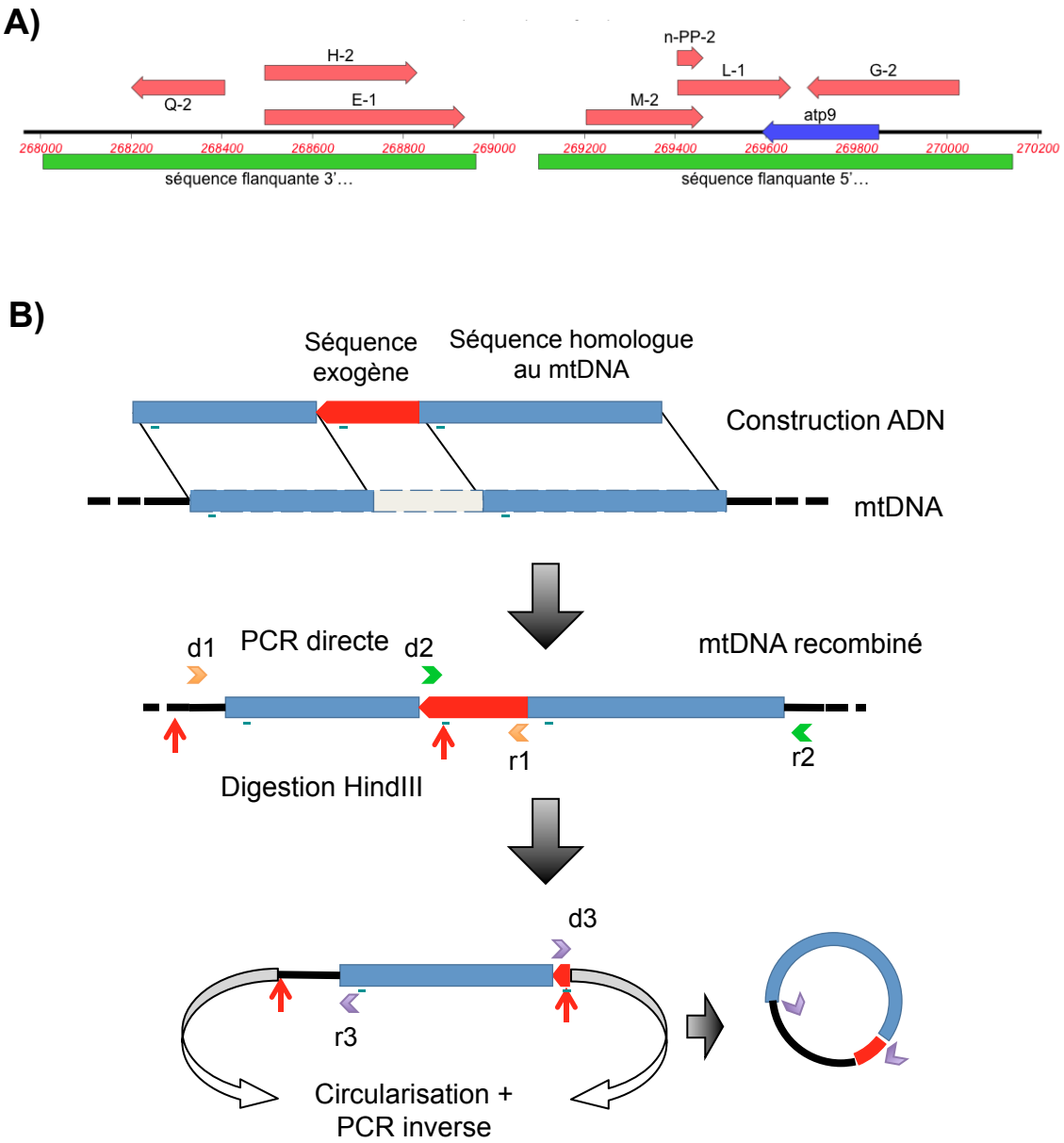


Figure 27: Recombinaison *in organello*.
A. Carte de la région du mtDNA cible. En rouge, les séquences répétées. En bleu, le gène *ATP9*. En vert, les séquences flanquantes utilisées.
B. Schéma de la recombinaison entre la construction et le mtDNA, et stratégie d'analyse par PCR directe et inverse. d1, r1, d2, r2 , d3 et r3 représentent les amorces utilisées pour la PCR directe 5' (d1 + r1), directe 3' (d2 + r2) et inverse (d3 + r3). Les flèches rouges montrent les sites HindIII.

respiratoire. Les mitochondries présentant le meilleur contrôle respiratoire se sont avéré être celles purifiées à partir de culture cellulaires en suspension (contrôle respiratoire = 1,3). Nous avons donc choisit de continuer les expériences avec ce matériel.

1.4. Construction à importer

En se reposant sur les travaux de Mileshina (Mileshina *et al.*, 2011) j'ai élaboré une construction pour les expériences d'import. Elle comprend des séquences flanquantes homologues à une région du mtDNA, qui seraient un substrat pour la recombinaison homologue, et une séquence témoin exogène à insérer dans la région du mtDNA ciblé. Pour favoriser la recombinaison de la séquence témoin j'ai choisi comme cible dans le mtDNA une zone riche en séquences répétées, montré auparavant comme fréquemment impliquée dans des événements de recombinaison (Miller-Messmer *et al.*, 2012; Zaegel *et al.*, 2006). La séquence flanquante en 5', en antisens, va de 269426 à 270471 (coordonnées du génome mitochondrial de Col-0, accession JF729201) et contient le gène *ATP9* ainsi que les séquences répétées M-2 ; L-1 et G-2. La séquence flanquante en 3' va de 268333 à 269286, en antisens, et contient les séquences répétées Q-2, H-2 et E-1 (figure 27A). La séquence exogène à insérer par recombinaison est le "multiple cloning site" (MCS, 150 nt) du plasmide pUC18, que nous avons choisi dans l'optique d'expériences ultérieures. Pour réaliser la construction, les différentes régions ont été amplifiées et assemblées par PCR, puis clonées dans le vecteur pGEM-T easy. La construction entière (2,1 kb) a ensuite été amplifiée par PCR et purifiée (kit PCR clean-up, Macherey-Nagel). Pour obtenir des fragments plus courts, et tester ainsi l'effet de la taille des séquences flanquantes sur la recombinaison, la construction a été amplifiée en utilisant d'autres amorces pour donner des fragments de 1,7 et 0,7 kb.

1.5. Test d'import

Nous avons analysé à chaque fois trois échantillons. Un contrôle négatif en absence d'ADN (NT), les témoins T_0 de contamination à l'ADN, arrêtés à Temps=0 où il ne devrait pas y avoir d'import, et les expériences d'import T_{130} (40 min d'import + 90 min post-incubation après traitement à la DNase I). Les expérimentations ont montré que l'ADN est bel et bien importé dans les mitochondries comme le montre la figure 28A On distingue souvent la présence d'ADN marqué dans le témoin de contamination d'ADN T_0, bien qu'en très faible quantité par rapport à T_{130}.

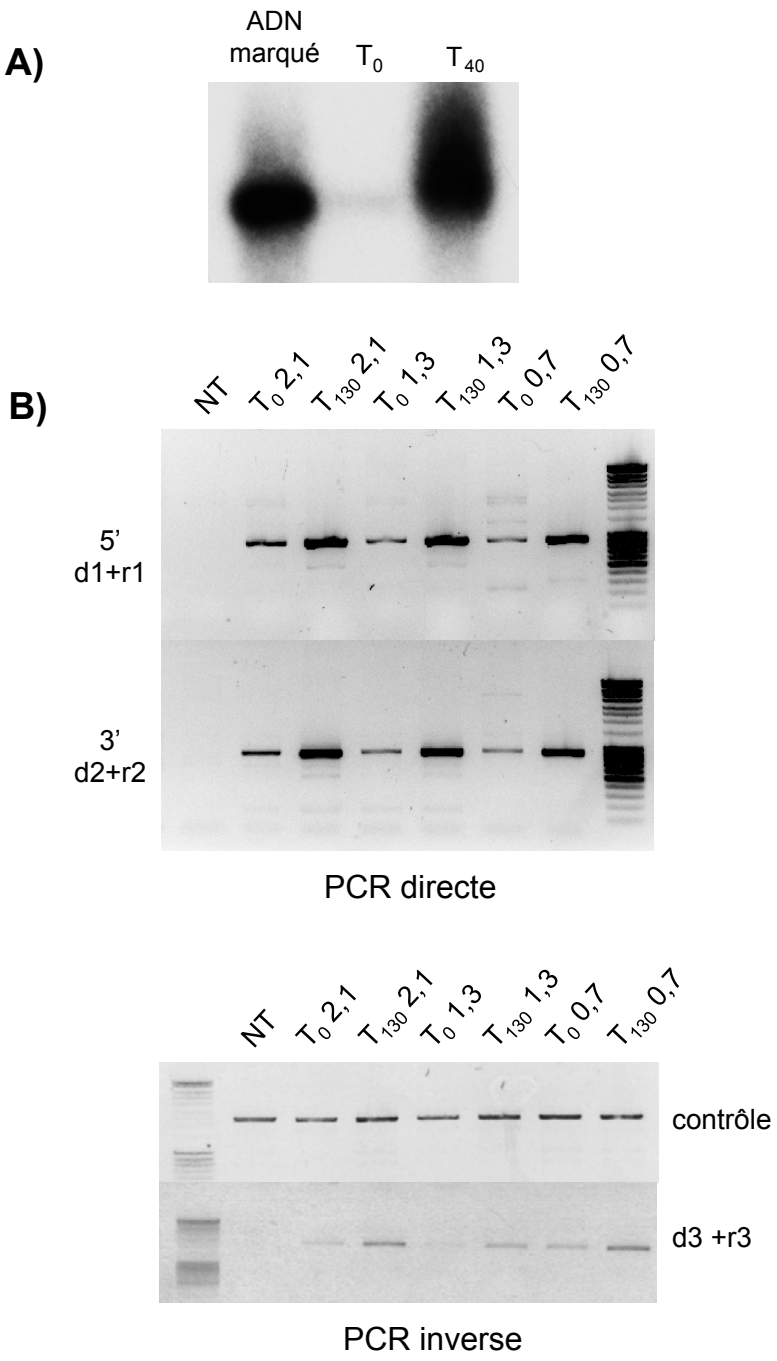


Figure 28. Résultats de la recombinaison *in organello*.
A. Contrôle de l'import d'ADN marqué radioactivement, par autoradiographie. Pour T_0 et T_{40}, l'ADN mitochondrial a été extrait après la mise en contact avec la construction, le traitement à la DNase et le lavage (T_0) ou après la phase d'import, le traitement à la DNase et le lavage (T_{40}).
B. En haut, PCR directe en 5' ou 3' du mtDNA. En bas, contrôle de la circularisation, et PCR inverse. Des constructions de taille variable ont été testées (2,1 kb, 1,3 kb et 0,7 kb). NT: mitochondries non traitées. T_{130}: ADN extrait après la phase d'import, le traitement à la DNase, et la phase de post-incubation.

1.6. Analyse de la recombinaison *in organello* par PCR directe ou inverse

Pour analyser l'ADN importé qui est intégré dans le mtDNA par recombinaison homologue j'ai utilisé les approches de PCR directe et inverse utilisées dans (Mileshina *et al.*, 2011) et représentées dans la figure 27B. Après isolement des mitochondries, import du fragment et post-incubation, les acides nucléiques ont été extraits. Une première série d'analyses a reposé sur de la PCR directe. Pour chaque échantillon, deux couples d'amorces ont été utilisés, un premier amplifiant la région en 3' de la séquence MCS et un deuxième couple d'amorces amplifiant la région en 5'. Cela permet de déterminer si le fragment a bien été inséré des deux côtés dans la région homologue du mtDNA. Les amorces utilisées étaient :

- Une amorce directe dans le mtDNA (d1) située en amont de la région mise en jeu dans l'expérience et une amorce réverse (r1), située dans la séquence MCS
- Une amorce directe dans la séquence MCS (d2) et une amorce réverse dans le mtDNA (r2) située en aval de la région mise en jeu dans l'expérience.

On s'attendait donc à obtenir des produits de PCR à partir de l'ADN extrait de mitochondries où il y a eu recombinaison, à T_{130}, mais aucun produit dans les échantillons qui n'avait pas été en contact avec le substrat d'import, à T_0.

La figure 28B montre l'analyse par PCR après une expérience d'import et de post-incubation avec les trois constructions de différentes tailles (2,1 ; 1,7 ; 0,7 kb). Il n'y a pas d'amplification dans la piste du témoin négatif sans ADN ajouté (NT). Cependant, on observe un produit dans les Témoin T_0, ayant été en contact avec l'ADN d'import mais n'ayant pas subit les phases d'import et de post-incubation. Cela signifie que la polymérase amplifie les résidus d'ADN présents à l'extérieur des mitochondries qui n'ont pas été complètement dégradés par le traitement à la DNase. De plus, pour la PCR en 5', la taille du produit obtenu ne correspond pas à celle attendue, qui devrait être de 1425 nt. Après plusieurs expériences avec des résultats semblables nous avons conclu que l'analyse par PCR directe donne trop d'artéfacts, probablement par changement de matrice pendant l'amplification, qui pourraient être dus aux caractéristiques des polymérases utilisées pour la PCR.

Pour éliminer la possibilité de formation d'artéfacts, un protocole de PCR inverse a été mis en place. La région du gène *ATP9* d'Arabidopsis présente plusieurs sites de digestion HindIII. La

séquence MCS, lorsqu'elle est insérée, ajoute un site de digestion. Le principe est donc de faire une digestion avec l'enzyme de restriction HindIII. On obtient alors divers fragments de différentes tailles, dont l'un contenant la jonction entre l'ADN endogène et le MCS. On effectue ensuite une religation des fragments sur eux-mêmes pour obtenir des petites molécules d'ADN circulaire. Seuls les échantillons ayant subi une recombinaison pourront être religués selon la configuration montrée dans la figure 27B. En utilisant des amorces divergentes dont l'une directe, spécifique de la séquence MCS, et l'autre réverse, située dans le mtDNA en amont de la région ciblée, l'amplification d'un produit ne sera effective qu'en cas d'insertion par recombinaison. On peut noter qu'il peut avoir également d'autres produits plus longs synthétisés, lorsque d'autres fragments HindIII générés de la digestion du mtDNA s'insèrent au moment de la religation.

Pour ces expériences j'ai dû optimiser le protocole pour obtenir une circularisation efficace. Le résultat de la figure 28B montre un contrôle de la digestion du mtDNA et de religation, par PCR utilisant des amorces divergentes. Mais dans l'analyse des expériences d'import et recombinaison par PCR inverse j'ai aussi rencontré des problèmes d'artéfacts, comme le montre la ure 28.B. En effet, un produit d'amplification a été obtenu dans toutes les pistes, y compris celles des contrôles T_0.

De plus, ces approches de PCR ne sont pas quantitatives et il n'est donc pas possible de mesurer l'efficacité de l'insertion. J'ai donc tenté de quantifier l'insertion par qPCR absolue, en utilisant comme référence les données d'amplification d'une construction synthétique de concentration connue. Mais là encore je me suis heurtée à des problèmes de reproductibilité, et je ne suis pas arrivée à une conclusion franche quant à la réalité et à l'efficacité de l'insertion par recombinaison. J'en ai conclu que par rapport aux résultats publiés le changement de système (pomme de terre à Arabidopsis) a changé les paramètres de l'expérience. En effet, il y a moins de mitochondries dans les cellules vertes d'Arabidopsis que dans la pomme de terre et les mitochondries obtenues sont de moins bonne qualité. L'expérience demanderait donc une optimisation plus poussée et de nombreuses conditions devraient être testées. De plus, le fait d'utiliser des séquences homologues à une région du mtDNA contenant de nombreuses séquences répétées peut possiblement favoriser la formation d'artéfacts de PCR. Après nombreuses expériences de mise au point j'ai décidé de mettre de côté ce projet et de me consacrer entièrement à l'étude des facteurs de recombinaisons des mitochondries d'Arabidopsis.

2. Etude d'une protéine homologue à RecG

Quand j'ai démarré mon travail de recherche, aucune hélicase mitochondriale impliquée dans la recombinaison n'avait encore été identifiée chez les plantes. Pourtant, les mécanismes de réparation et d'évolution du mtDNA, décrits dans l'introduction, impliquent des mécanismes de recombinaison. Or des structures complexes d'ADN interviennent forcément dans ces processus, et une ou plusieurs ADN hélicases permettant de résoudre des structures complexes comme les HJ sont indispensables. La seule hélicase mitochondriale déjà identifiée, Twinkle-like, est celle associées au complexe de réplication, responsable du déroulement de l'ADN pendant la réplication, et n'est apparemment pas impliquée dans la recombinaison (Diray-Arce *et al.*, 2013). De plus, il a été montré que la recombinaison mitochondriale est sous le contrôle de systèmes de surveillance qui défavorisent la recombinaison impliquant des séquences répétées de courte taille (Miller-Messmer *et al.*, 2012; Shedge *et al.*, 2007; Zaegel *et al.*, 2006). Une façon d'expliquer un tel mécanisme de surveillance est de postuler qu'une hélicase dissocie les intermédiaires de recombinaison indésirables, comme décrit pour d'autres hélicases, par exemple FANCM (Knoll *et al.*, 2012) ou l'hélicase Srs2 de la levure (Krejci *et al.*, 2003). Nous avons effectué une recherche bio-informatique de gènes d'Arabidopsis codant pour des protéines orthologues à des hélicases bactérienne, car le système de recombinaison du mtDNA est évolutivement proche de celui des bactéries. De nombreux homologues de l'hélicase RecQ bactérienne sont présents dans les noyaux de plantes et sont impliquées dans des processus de réparation, recombinaison et dans la régression des fourches de réplication (Schropfer *et al.*, 2014). Cependant, selon les programmes de prédiction, aucune d'entre-elles ne semble adressée aux mitochondries ou aux chloroplastes. De même, un orthologue de SRS2/UvrD qui a été caractérisé n'a pas été décrit comme adressé aux organelles (Blanck *et al.*, 2009). D'autre part, l'hélicase RuvAB n'a pas d'homologues dans Arabidopsis. Cependant, un homologue de l'hélicase RecG bactérienne est codé par le génome d'Arabidopsis, et les différents logiciels de prédiction la situent dans les organelles. Nous avons donc voulu étudier cette protéine d'Arabidopsis, que nous avons appelée RECG1, et son rôle dans les plantes et les organelles. La plupart de mon travail sur cette protéine a été publié en 2015 dans The Plant Cell (Wallet *et al.*, 2015). Mais quelques expériences et études n'ont pas été publiées ou ont été faites à la suite. Elles sont répertoriées ici.

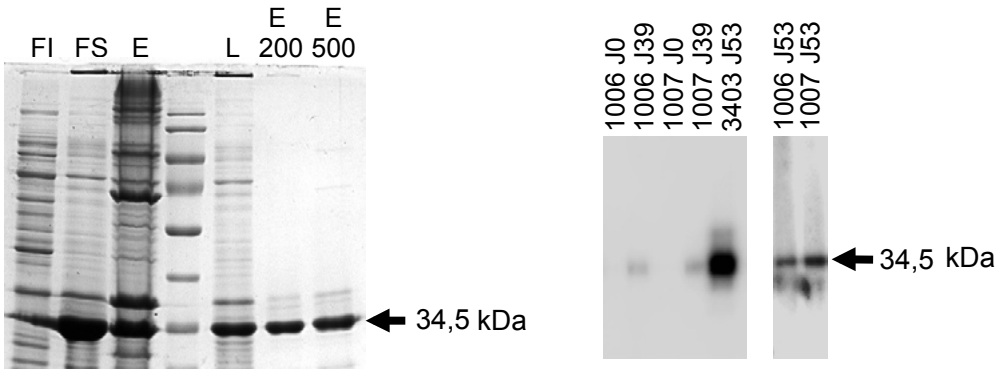


Figure 29. Expression et purification de la protéine recombinante RECG1.
A gauche, gel coloré au bleu de Coomassie avec les différentes fractions de la purification de RECG1 recombinante. A droite, immunodéctection de RECG1 recombinante à l'aide des sérums immuns des différents lapins, récoltés aux jours J0, J39 et J53 après début de l'immunisation. FI: fraction insoluble; FS: fraction soluble; E: éluat; L: lavage à 30 mM imidazole; E200: élution à 200 mM imidazole; E500: élution à 500 mM imidazole. 1006, 1007, 3403: Code numéro des lapins. Les flèches montrent les bandes correspondant aux tailles attendues pour la protéine recombinante.

2.1. Préparation d'anticorps spécifiques à RECG1

Les anticorps sont toujours utiles dans les études de protéines. En effet ils permettent un suivi de la diminution de la protéine dans les mutants, mais aussi le contrôle de son expression dans des bactéries ou dans des plantes, lors d'expériences de complémentation. L'obtention d'anticorps bien spécifiques permet également d'envisager des approches de détection *in situ*, ainsi que des expériences d'immuno-précipitation afin d'identifier des protéines partenaires. J'ai donc fait produire des anticorps contre la protéine RECG1. J'ai exprimé une partie de la protéine dans *E. coli* et je l'ai purifiée avant de l'envoyer à un service de production d'anticorps qui l'a injecté à des lapins. Pour cela, j'ai cloné la séquence correspondant aux 260 derniers acides aminés de la protéine RECG1 d'Arabidopsis dans le vecteur p0GWA. L'induction et la purification de la protéine sont décrites en II.5.5. La protéine exprimée s'est avérée être insoluble (figure 29). Nous avons donc travaillé en conditions dénaturantes, avec la protéine solubilisée dans une solution contenant de l'urée. Pendant la purification par affinité au nickel une grande partie de la protéine a été éluée de la résine Ni-NTA avec très peu d'imidazole, dans les fractions de lavage, et le reste a été récupéré dans les fractions d'élution avec 200 mM et 500 mM d'imidazole. Les fractions contenant la protéine ont été groupées, puis concentrées par ultrafiltration (Ultrafree, Millipore). A la fin la protéine a été diluée jusqu'à une concentration optimale en protéine, mais réduite en urée et imidazole, compatible avec l'injection dans les lapins. Deux lapins (codes 1006 et 1007) ont été immunisés par l'entreprise Covalab et un lapin (code 3403) par le service de production d'anticorps de l'Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC, Illkirch). Après avoir vérifié que le sérum pré-immun des lapins ne contenait pas d'anticorps pouvant interférer avec la détection de notre protéine, les animaux ont subi plusieurs injections réparties sur 2 mois. A plusieurs intervalles de temps (jours 39, 53 et 70 après injection), du sérum de l'animal a été prélevé et nous a été envoyé. J'ai testé les différents anticorps sur la protéine recombinante purifiée (figure 29). Comme le montre la figure 29, l'anticorps du lapin 3403 est le plus affin pour la protéine, tout en étant spécifique. J'ai donc choisi de l'utiliser par la suite. Cet anticorps est efficace car il a permis de détecter la protéine RECG1 dans un extrait total d'Arabidopsis, et a pu montrer une baisse importante, voire une disparition de la protéine chez les plantes mutantes ((Wallet *et al.*, 2015), figure 3C).

2.2. Etude de RECG1 (publication Wallet *et al.* 2015, The Plant Cell)

This article is a *Plant Cell* Advance Online Publication. The date of its first appearance online is the official date of publication. The article has been edited and the authors have corrected proofs, but minor changes could be made before the final version is published. Posting this version online reduces the time to publication by several weeks.

The RECG1 DNA Translocase Is a Key Factor in Recombination Surveillance, Repair, and Segregation of the Mitochondrial DNA in Arabidopsis[OPEN]

Clémentine Wallet,[a] Monique Le Ret,[a] Marc Bergdoll,[a] Marc Bichara,[b] André Dietrich,[a] and José M. Gualberto[a,1]

[a] Institut de Biologie Moléculaire des Plantes, CNRS UPR2357, Université de Strasbourg, 67084 Strasbourg, France
[b] CNRS UMR7242, IREBS, Université de Strasbourg, 67412 Illkirch, France

ORCID IDs: 0000-0002-2388-4510 (C.W.); 0000-0002-7296-2618 (J.M.G.)

The mitochondria of flowering plants have considerably larger and more complex genomes than the mitochondria of animals or fungi, mostly due to recombination activities that modulate their genomic structures. These activities most probably participate in the repair of mitochondrial DNA (mtDNA) lesions by recombination-dependent processes. Rare ectopic recombination across short repeats generates new genomic configurations that contribute to mtDNA heteroplasmy, which drives rapid evolution of the sequence organization of plant mtDNAs. We found that *Arabidopsis thaliana* RECG1, an ortholog of the bacterial RecG translocase, is an organellar protein with multiple roles in mtDNA maintenance. RECG1 targets to mitochondria and plastids and can complement a bacterial *recG* mutant that shows defects in repair and replication control. Characterization of Arabidopsis *recG1* mutants showed that RECG1 is required for recombination-dependent repair and for suppression of ectopic recombination in mitochondria, most likely because of its role in recovery of stalled replication forks. The analysis of alternative mitotypes present in a *recG1* line and of their segregation following backcross allowed us to build a model to explain how a new stable mtDNA configuration, compatible with normal plant development, can be generated by stoichiometric shift.

INTRODUCTION

Vascular plants have large mitochondrial genomes, varying from a few hundred kilobases to several megabases in size (Sloan et al., 2012) and are mostly constituted by noncoding sequences. The actual structure of these large genomes remains unclear. Although mapping and sequencing data generally produce large circular chromosomes comprising all the genetic information of mitochondria, physical observation of the mitochondrial DNA (mtDNA) using electrophoretic and microscopic approaches revealed a complex and fragmented distribution into heterogeneous populations of circular, linear, and multibranched double- and single-stranded molecules that apparently exist in dynamic equilibrium (Oldenburg and Bendich, 1996; Backert et al., 1997; Backert and Börner, 2000). Multichromosomal mitochondrial genomes have also been found in several species of flowering plants (Alverson et al., 2011; Sloan et al., 2012). Moreover, the structure of plant mitochondrial genomes evolves rapidly, driven by rearrangements that result from high rates of recombination. This rapid evolution of the mtDNA is manifest not just interspecies, but also intraspecies, and significantly different mtDNA structures can be found in different plant accessions (Allen et al., 2007; Arrieta-Montiel et al., 2009). On the contrary, gene sequences evolve very slowly in most species, apart from a few significant exceptions (Mower et al., 2007).

Our understanding of the processes that control the maintenance, replication, inheritance, and evolution of such complex and fragmented genomes remains limited. It is possible that replication of the plant mtDNA initiates by recombination-mediated processes, as in the yeast *Candida albicans* (Gerhold et al., 2010). Branched, linear, and sigma-like forms of the plant mtDNA that have been observed would be intermediates of rolling circle and/or recombination-dependent replication mechanisms, referring to the model of phage T4 (Backert and Börner, 2000). However, recombination-dependent replication does not explain how stoichiometric replication and inheritance of the fragmented mitochondrial genome is achieved. Whether there are replication checkpoints has not been documented, and it is not known whether replication of the two mtDNA strands is coupled or independent. The former is possible because the organellar replicative helicase has intrinsic DNA primase activity (Diray-Arce et al., 2013).

As to the shuffling of mtDNA sequences, it is promoted by the numerous repeated sequences present in plant mitochondrial genomes and involves (1) frequent homologous recombination (HR) via large repeated sequences, (2) ectopic recombination involving intermediate-size repeats (IRs), or (3) illegitimate recombination involving sequence microhomologies (Maréchal and Brisson, 2010; Gualberto et al., 2014). These recombination activities contribute to the intrinsic heteroplasmic state of plant mitochondrial genomes, in which alternative mitotypes coexist with the predominant mtDNA at substoichiometric levels (Small et al., 1989; Kmiec et al., 2006). In some conditions, one of the minor mitotypes can become predominant, by a stoichiometry shift of the relative copy numbers. This process has been described in many different species (Small et al.,

[1] Address correspondence to jose.gualberto@ibmp-cnrs.unistra.fr.
The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: José M. Gualberto (jose.gualberto@ibmp-cnrs.unistra.fr).
[OPEN] Articles can be viewed online without a subscription.
www.plantcell.org/cgi/doi/10.1105/tpc.15.00680

 © 2015 American Society of Plant Biologists. All rights reserved.

1989; Kanazawa et al., 1994; Chen et al., 2011) and cybrids (Sakai and Imamura, 1992; Bellaoui et al., 1998) and can occur in the time frame of just a few generations. It is also at the origin of cytoplasmic male sterile varieties of agronomic interest and of corresponding fertile revertants (Janska et al., 1998; Allen et al., 2007; Feng et al., 2009; Matera et al., 2011). Thus, rapid evolution of the mtDNA involves recombination processes that promote mtDNA heteroplasmy and sorting of the alternative mitotypes by stoichiometric shifting.

The mechanisms of mitochondrial recombination that promote mtDNA variability have only started to be unraveled. The "raison d'être" of these recombination pathways apparently is the repair of mtDNA lesions (Davila et al., 2011; Christensen, 2013). The response to genotoxic stress that comes out with increased HR between IRs is controlled by genes involved in repair, several of which have been identified (*RECA HOMOLOG3* [*RECA3*], *WHIRLY2* [*WHY2*], *ORGANELLAR DNA BINDING PROTEIN1* [*ODB1*], and *DNA POLYMERASE 1B* [*POL1B*]). In the corresponding mutants, there is reduced repair by homologous recombination and there might be increased illegitimate recombination involving sequence microhomologies (Cappadocia et al., 2010; Parent et al., 2011; Janicka et al., 2012; Miller-Messmer et al., 2012). These illegitimate recombination events are possibly responsible for U-turn genomic inversions that accumulate in plant organelles and also in human mitochondria (Zampini et al., 2015).

Because of the deleterious nature of the rearrangements that can be created by recombination processes, these are under tight control, and undesired recombination activities are normally suppressed. Restriction of ectopic and illegitimate mtDNA recombination might occur during replication, while rescuing stalled replication forks, or at the postreplicative level, through dissociation of D-loops formed by strand invasion at ectopic DNA locations. The recombination surveillance pathways are disrupted in the mutants of several genes, such as the mutants of *MUTS HOMOLOG1* (*MSH1*), *ORGANELLAR SINGLE-STRANDED DNA BINDING PROTEIN1* (*OSB1*), *RECA2*, and *RECA3* (Zaegel et al., 2006; Arrieta-Montiel et al., 2009; Miller-Messmer et al., 2012). In these mutants, there are increased rates of ectopic recombination involving IRs, which in later mutant generations can result in stoichiometric shifting (Zaegel et al., 2006; Davila et al., 2011). The study of these mutants advanced our understanding of the factors that can promote mtDNA evolution by stoichiometric shifting, but the examples that have been described involved multiple events of recombination, difficult to follow individually, and the dynamics of the shifting process leading from one predominant mtDNA structure to another could not be understood.

For a deeper understanding of these processes, we searched for specific DNA helicases that could play essential roles in mtDNA maintenance, in particular in the orientation toward error-free recombination and replication processes. We identified the *Arabidopsis thaliana* RECG1 helicase, an ortholog of bacterial RecG. We show that RECG1 has multiple roles in mtDNA repair by recombination and in the suppression of ectopic recombination events. Remarkably, whereas mutations of so far identified mitochondrial recombination factors triggered multiple and complex events of recombination, a *recG1* mutation generated only two alternative configurations of the mitochondrial genome differing in a single locus. This convenient system allowed us to study the dynamics of their segregation in the following generations, a process influenced by RECG1, and to build a model for the transition to a new stable mtDNA structure compatible with normal plant development by stoichiometric shift.

RESULTS

We screened the *Arabidopsis* genome for genes coding for candidate DNA helicases potentially involved in mtDNA recombination. Because the recombination machineries of plant organelles are partially inherited from their symbiotic ancestors, we have in particular searched for orthologs of bacterial DNA helicases with known functions in recombination processes. Our scrutiny mainly retrieved a homolog of bacterial RecG. Other prokaryotic-type DNA helicases either are not encoded in plant genomes (i.e., RuvAB), have no putative organellar targeting sequences, or have been studied already, with no hint of possible functions in organellar processes (i.e., RecQ-like proteins)(Knoll and Puchta, 2011). In bacteria, RecG is a DNA translocase that can perform multiple functions, including Holliday junction (HJ) resolution, rescuing of stalled replication forks, modulation of RecA-promoted strand exchange, and suppression of pathological DNA replication (Vincent et al., 1996; McGlynn and Lloyd, 2002; Rudolph et al., 2010, 2013). The RecG homolog is encoded in all sequenced plant genomes, including vascular plants, mosses, and algae (Supplemental Figure 1). We named the Arabidopsis RecG RECG1.

RECG1 Is Conserved in All Plant Species and Is Targeted to Both Mitochondria and Chloroplasts

Phylogenetic analysis shows that plant RECG1s branch together with cyanobacterial RecG proteins, suggesting that the plant gene was inherited from the symbiotic ancestor of the chloroplast (Figures 1A; Supplemental Figure 1). A three-dimensional model of the Arabidopsis RECG1 could be generated (Figure 1B) on the basis of the known RecG structure of *Thermatoga maritima* that has been solved in complex with a replication fork analog (Singleton et al., 2001). The RecG structure comprises three domains, two of which form the core helicase motor. The large third domain contains a Greek key fold that acts as a "wedge," directing the template strand toward the core motor and steering the second strands of the side arms to facilitate their annealing. This mechanism can explain HJ branch migration, regression of replication forks into HJ, and unwinding of D-loops and R-loops, i.e., all functions that have been proposed for RecG (McGlynn and Lloyd, 2002; Briggs et al., 2004) and for its phage T4 functional homolog UvsW (Carles-Kinch et al., 1997; Dudas and Kreuzer, 2001; Manosas et al., 2013). All major domains of the plant RECG1 model perfectly match the bacterial structure (Figure 1B), supporting the assumption that RECG1 has the same activities as the bacterial enzyme and potentially has similar functions. Notably, all plant proteins share three additional short extensions that are not found in bacterial sequences (Figure 1A; Supplemental Figure 2) and that could not be accommodated into the model. These extensions (of 20, 15, and 19 amino acids, respectively, in Arabidopsis RECG1) might confer specific characteristics to plant RECG1s.

All plant sequences present N-terminal extensions that could correspond to organellar targeting peptides (Supplemental Figure 2), and the Arabidopsis RECG1 is predicted to be targeted to

organelles (http://suba.plantenergy.uwa.edu.au). This assumption was tested by expression of a construct encoding the N terminus of RECG1 fused to eGFP. Subcellular localization of the fusion protein was analyzed both in transfected *Nicotiana benthamiana* leaf cells and in stable Arabidopsis transformants, and in both systems the protein was found in chloroplasts and in mitochondria (Figure 1C), implying that RECG1 is an additional example of a dually targeted organellar protein. Recently, the *Physcomitrella patens* ortholog of Arabidopsis RECG1 was described and shown as well to be targeted to both mitochondria and chloroplasts (Odahara et al., 2015).

Expression of *RECG1* during plant development was studied in Arabidopsis lines stably transformed with the β-glucuronidase gene under the control of the *RECG1* promoter. The consensus observation from several independent transformants was that the *RECG1* promoter was active at all developmental stages and in most tissues, including shoots of young plants, root tips, and secondary root primordia. High expression was observed in the vascular tissues of leaves and roots (Figure 1D). In mature flowers, β-glucuronidase was present in sepals and in stamen filaments, but not in pollen, ovules, or developing seeds.

RECG1 Partially Complements Bacterial RecG

The ability of plant RECG1 to complement *Escherichia coli* RecG in repair of UV-C-induced DNA lesions was tested, as described in Methods. UV-C promotes formation of pyrimidine dimers that block the replication fork and induce repair mechanisms in which RecG is implicated (Meddows et al., 2004). After UV-C irradiation, the Δ*recG* strain presented a 30-fold decrease in survival, as described before (Lloyd, 1991). Expression of the plant RECG1 significantly increased the survival rate of Δ*recG* upon UV exposure (Figure 2A). When bacteria were irradiated with up to 20 J/m^2 of UV-C, the survival of complemented mutant cells increased by a factor of 15-fold compared with the noncomplemented mutant, reaching half of the survival rate of wild-type bacteria. We also tested

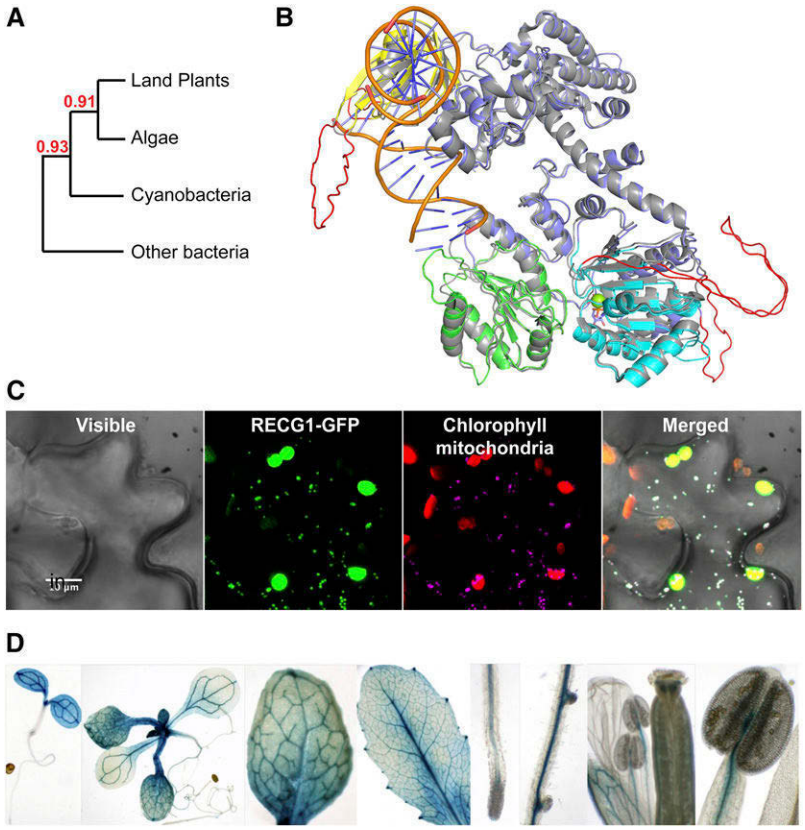


Figure 1. A RecG Protein Is Present in the Organelles of Plants.

(A) Detail of the RecG phylogenetic tree supporting that plant RecGs derive from a cyanobacterial ancestor. Bootstrap values are indicated.
(B) Model of the Arabidopsis RECG1 protein superposed on the known structure of *T. maritima* RecG (in gray) bound to a DNA substrate mimicking a stalled replication fork (Singleton et al., 2001). The large N-terminal domain is in purple, with the wedge domain in yellow. The two helicase domains are in green and light blue, with complexed ADP and Mg^{2+} (green ball). The three sequences specific to plant RecGs (see alignment in Supplemental Figure 2) could not be modeled and are shown in red.
(C) Transient expression of the RECG1-GFP construction in *N. benthamiana* leaf epidermal cells showing dual targeting of RECG1. Chloroplasts and mitochondria are identified by the fluorescence of the chlorophyll and by coexpression of a CoxIV-dsRED construct, respectively.
(D) Widespread expression of *RECG1* in Arabidopsis, according to promoter-GUS fusion at different developmental stages. From left to right, 3-d-old and 7-d-old seedlings, rosette and cauline leaves, roots, flowers, and anthers.

whether the Arabidopsis RECG1 could complement a deficiency in *RuvAB* because both RecG and RuvAB are translocases that drive branch migration and have been described as either partially redundant or complementary in the resolution of HJ (Zhang et al., 2010; Mawer and Leach, 2014). In contrast to Δ*recG*, the Δ*ruvA* mutation could not be rescued by RECG1 (Figure 2A). Thus, despite the absence of RuvAB in plant organelles, RECG1 has not acquired RuvAB-like functions during evolution and apparently has similar specificity in DNA repair as bacterial RecG.

In addition to its function in DNA repair, bacterial RecG has recently been assigned as a guardian of the bacterial genome that prevents pathological re-initiation of DNA replication (Rudolph et al., 2013). In *E. coli*, genome replication initiates bidirectionally at the *OriC* region and proceeds around the circular genome in both directions. Therefore, in wild-type exponentially growing bacteria, a maximum in copy number is found for the sequences around the *OriC* region and a minimum for the region between the termination sequences *TerA* and *TerB* (Figure 2B). In the absence of RecG, replication can reinitiate where forks collide by a process that depends on recombination functions (Rudolph et al., 2013). We tested the ability of RECG1 to prevent replication reinitiation by quantifying the relative copy number of the *YdcR* gene located equidistantly between *TerA* and *TerB*, as described in Methods. It was found that in Δ*recG* there was indeed a significant increase in *YdcR* copy number compared with the wild type (Figure 2B), indicating replication restart in the termination zone and confirming published results (Rudolph et al., 2013). When RECG1 was expressed in the Δ*recG* mutant, this effect was reduced, although not abolished, showing that RECG1 can partially replace bacterial *recG* in the suppression of uncontrolled replication reinitiation. This result also suggests that RECG1 might be necessary for stoichiometric replication of the plant organellar genomes.

Finally, we tested whether RECG1, like RecG, is able to dissociate R-loops, and as a consequence inhibit the replication of plasmids containing ColE1 replicons (Vincent et al., 1996). Replication of these plasmids is primed by a short RNA expressed from the replication origin (Dasgupta et al., 1987). The wild-type, Δ*recG*, and Δ*recG RECG1*+ strains were transformed with ColE1-based plasmids (pBAD/Thio or pUC18) whose relative copy numbers in exponentially growing cultures were compared with the copy numbers in saturated precultures. In nonselective growth conditions, a 3-fold lower copy number of the plasmids was retained in the complemented strain compared with the Δ*recG* strain, comparable to the relative copy numbers found in the wild-type strain. This result indicated that expression of RECG1 inhibited the replication of ColE1-based replicons (Figure 2C).

Arabidopsis Plants Deficient in RECG1 Have Reduced Recombination-Dependent mtDNA Repair

From a search of the available T-DNA insertion mutant collections, only two confirmed *recG1* lines were obtained (Figure 3A), one in accession Col-0 containing an insertion in the 5′-untranslated region (UTR) of *RECG1* (*recG1-1*) and a second one in accession Wassilewskija (Ws) containing a T-DNA insertion in the last but one exon (*recG1-2*). The insertion in *recG1-2* results in the loss of the last 126 codons, part of which encode a highly conserved helical hairpin motif essential for helicase activity (Mahdi et al., 2003). Quantification by RT-qPCR showed that both lines have a significant decrease in *RECG1* transcript, of about 5-fold (Figure 3B). Immunoblot analysis with a specific antibody against RECG1 showed that *recG1-1* is a hypomorphic mutant that has a significant reduction in RECG1 protein, while *recG1-2* is a knockout in which the protein is no longer detected (Figure 3C). Under normal growth conditions, both *recG1-1* and *recG1-2* plants were indistinguishable from wild-type plants.

In bacteria, RecG is required for recombination-dependent DNA repair (Lloyd and Buckman, 1991), and the above assays suggest that RECG1 can complement this function. To test whether RECG1 plays a similar role in plant mitochondria, the growth of *recG1-1* and *recG1-2* seedlings was examined under genotoxic stress, more specifically under exposure to ciprofloxacin (CIP). CIP is an inhibitor of gyrase, an enzyme that in plants is active in both mitochondria and chloroplasts but not in the nucleus (Wall et al., 2004). Gyrase inhibition by CIP promotes double-strand breaks (DSBs), and mutants for genes involved in organellar HR-dependent repair are more sensitive to CIP than the wild type (Cappadocia et al., 2010; Rowan et al., 2010; Parent et al., 2011; Janicka et al., 2012; Miller-Messmer et al., 2012; Lepage et al., 2013). Seedlings of *recG1-1*, *recG1-2*, and of the corresponding wild-type lines were grown in the presence of CIP, and the ability to develop normal first true leaves was scored, as previously described (Parent et al., 2011; Miller-Messmer et al., 2012). The Col-0 and Ws wild-type accessions were differentially sensitive to CIP, with Ws slightly more sensitive than Col-0: At 0.75 μM CIP, ~20% of the Col-0 plants still developed first leaves while virtually none of the Ws seedlings did (Figure 4A). Regarding the mutant lines, both *recG1-1* and *recG1-2* were slightly more sensitive than the corresponding wild-type accessions, with fewer mutant seedlings developing true first leaves (Figures 4A and 4B).

Repair by recombination of mtDNA DSBs is also correlated to the molecular phenotype of increased ectopic recombination across IR sequences. This response is reduced in several mutants impaired in HR-dependent repair (Janicka et al., 2012; Miller-Messmer et al., 2012). Therefore, the levels of ectopic recombination across mtDNA IRs were also compared between CIP-treated and untreated seedlings, for wild-type and *recG1* lines (Figures 4C and 4D). The sequence and organization of the mtDNA are known for Col-0 (accession JF729201) but not for Ws. To compare the two *recG1* mutant alleles, which have different genetic backgrounds, we first tested the mtDNA of the Ws accession for the presence of IRs that are also contained in the Col-0 mitochondrial genome. We found that few IRs previously described as involved in ectopic recombination in Col-0 mitochondria are conserved in the Ws mtDNA (Supplemental Figure 3). We analyzed the ectopic recombination involving repeats D, L, X, and EE, according to the nomenclature of the published mtDNA of Col-0. Sequence L could not be tested in *recG1-2* because it is not repeated in the mtDNA of Ws. In the two wild-type accessions, upon CIP exposure there was a significant increase in ectopic recombination involving all repeats tested, as previously described (Miller-Messmer et al., 2012). The recombination was in all cases asymmetric, with one of the crossover products preferentially increased compared with the reciprocal one (Figure 4D). The magnitude of the increase varied significantly depending on the repeat and the Arabidopsis wild-type accession: In the case of repeat X, there was only a slight increase in Ws plants of crossover product X-2/1, while the increase of the same sequence was 10-fold higher in

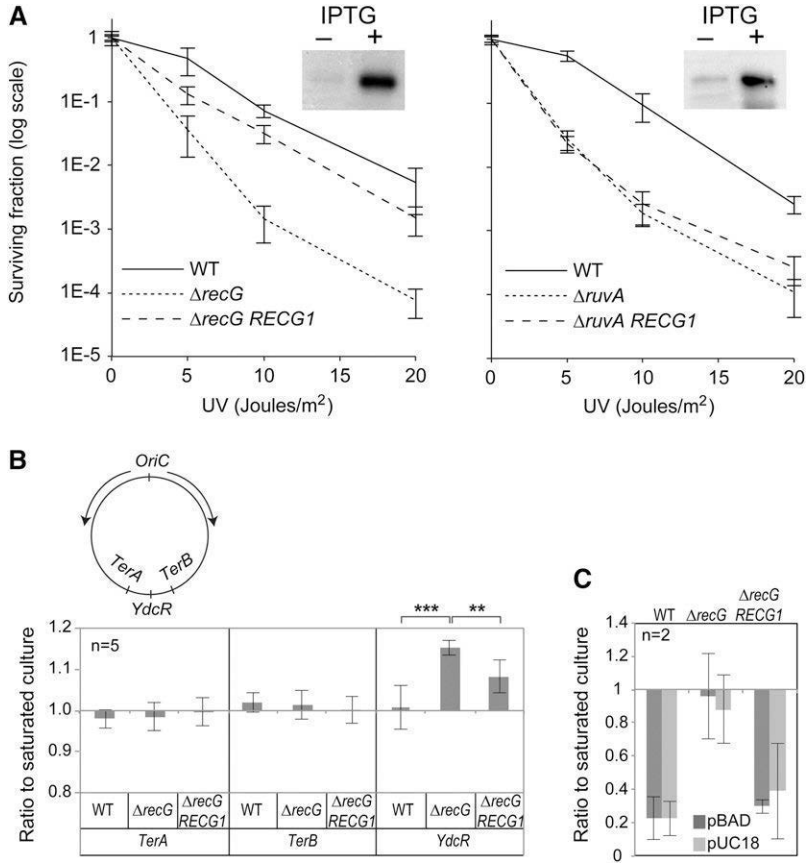


Figure 2. Arabidopsis RECG1 Partially Complements RecG Deficiency in Bacteria.

(A) Partial complementation of *E. coli* Δ*recG* but not of Δ*ruvA* in the repair of UV-C induced damage. JM103 (WT), JM103 *recG::kan* (Δ*recG*), and JM103 *ruvA::Tc* (Δ*ruvA*) transformed with the *pACYClacz* empty vector were used as control strains and compared with the ones transformed with *pACYClacz:RECG1*. Recombinant RECG1 expression was induced with IPTG prior to treatment with UV-C, as described in Methods, and controlled by immunodetection with the RECG1 antibody (on the upper right of the graphics, +/− IPTG). Results are means from two independent experiments, with three technical replicates each, and are represented on a logarithmic scale.

(B) Complementation of Δ*recG* by RECG1 in the suppression of pathologic replication reinitiation. The sequences *TerA*, *TerB*, and *YdcR* were quantified in growing cultures and the levels were compared with those in the corresponding saturated cultures. The relative positions in the *E. coli* chromosome and the divergent replication from *OriC* are shown in the scheme. Results correspond to five independent experiments ± SD. Asterisks show the statistical significance (unpaired *t* test; **P < 0.01, ***P < 0.001).

(C) Effect of RECG1 expression on the replication of plasmids (pBAD/Thio and pUC18) containing a ColE1 replication origin. Results were normalized against the bacterial chromosome (*TerA* sequence).

Col-0 plants. In the case of crossover product EE-2/1, the increase was more than 1000-fold in CIP-treated Col-0 plants, but only ~50-fold in Ws. These differences do not correlate with repeat size, but are probably related to the genomic context. But in all cases, the CIP-induced increase in crossover products was significantly lower in both mutant *recG1* lines. These results imply that reduction or loss of RECG1 expression impairs recombination-dependent repair of DSBs in the mtDNA.

We also tested whether RECG1 takes part in chloroplastic recombination-dependent repair. In contrast to the mtDNA, the chloroplast DNA (cpDNA) of Arabidopsis is not rich in IRs whose recombination can be used to monitor HR-dependent repair. We identified a single imperfect repeat (76 bp, 92% identity) amenable to qPCR analysis (a repeat that we named R2, coordinates 38705:38780 and 40929:41004 on the published cpDNA of Arabidopsis). We found that CIP-treated wild-type plants accumulated high levels of a crossover product from repeat R2, confirming that recombination involving this repeat is a good sensor of HR-dependent DSB repair in the Arabidopsis chloroplast (Figure 4E). However, *recG1* mutants displayed a similar response with no or little quantitative difference compared with the wild type, suggesting that RECG1 is not essential for chloroplastic repair of DSBs.

RECG1 Participates in the Surveillance of mtDNA Recombination

Several Arabidopsis mutants described as affected in recombination surveillance (*msh1*, *osb1*, *recA2*, and *recA3*) showed, in normal growth conditions, increased ectopic recombination of

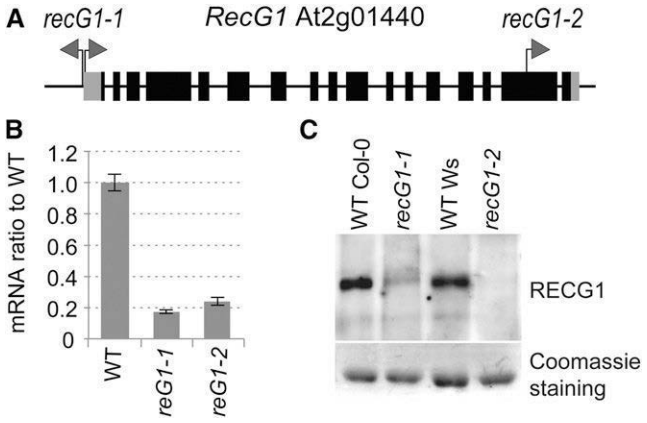


Figure 3. Arabidopsis *recG1* Lines.

(A) Schematic representation of the Arabidopsis *RECG1* gene. Coding sequences are in black and 5′- and 3′-UTRs are in gray. The position of the T-DNA insertions in the mutant lines is shown, with arrows indicating the orientation of the left border.

(B) RT-qPCR relative quantification of the *RECG1* transcript in the mutant lines.

(C) Immunodetection of RECG1 in total inflorescence extracts. Coomassie blue staining of the membrane is shown as a loading control. A protein of ~110 kD is detected in extracts from wild-type plants, in agreement with the predicted size of 108 kD. The protein is reduced in *recG1-1* and absent in *recG1-2*.

the mtDNA across IRs (Zaegel et al., 2006; Shedge et al., 2007; Miller-Messmer et al., 2012). In these mutants, without any genotoxic treatment there is already a constitutive increase in error-prone recombination activities. The *recG1* mutants were thus tested for such a molecular phenotype. The IRs D, F, X, and EE that are conserved in the mtDNA of both Col-0 and Ws (Supplemental Figure 3) were selected for qPCR analysis of the accumulation of crossover products, as previously described (Miller-Messmer et al., 2012). A significant increase in crossover products compared with wild-type levels was observed in both mutants (Figure 5A). The most important effects were in *recG1-2* KO plants, which for certain repeats showed an increase in crossover products one order of magnitude higher than in the hypomorphic *recG1-1* mutant. Recombination in *recG1* plants resulted either in equivalent accumulation of both reciprocal products (as for repeats D, F, and EE) or in asymmetrical accumulation of mainly one of the two products (as for repeat X). As the ectopic recombination induced by CIP treatment, depending on the repeats there were remarkable quantitative differences (Figure 5A). In particular, crossover products from the recombination involving repeats EE increased ~1000-fold in *recG1-2* compared with wild-type plants. To test whether loss of RECG1 also affects surveillance of recombination in the chloroplast, we monitored the ectopic recombination involving repeat R2. We found that RECG1 had little or no effect on the recombination involving this chloroplast repeat (Figure 5B). This suggests that RECG1 has no significant roles in the surveillance of recombination in the chloroplast, contrarily to its function in mitochondria.

It has been described that several mutants affected in organellar genome recombination surveillance or in repair by recombination can display visible phenotypes of reduced growth, distorted leaves, variegation, and reduced fertility, but with low penetrance or only in late mutant generations (Zaegel et al., 2006; Maréchal and Brisson, 2010; Davila et al., 2011). However, no such phenotypes were observed in homozygous *recG1* plants of generations up to T5. We screened the development of more than 100 individual T5 mutant plants from the *recG1-2* knockout line, but none showed an altered visible phenotype.

Increased Recombination in *RECG1* Knockout Plants Causes Segregation of a Small Crossover Product That Replicates Autonomously

To test the effect of increased recombination in *recG1* mutants on the relative copy number and segregation of the different mtDNA regions, the complete mtDNA was scanned by qPCR using a set of primer pairs spaced 5 to 10 kb apart (Figure 6A). The published Arabidopsis Col-0 mtDNA sequence was taken as a reference. No significant changes in the relative copy number of the different regions were detected in the *recG1-1* hypomorphic line, consistent with the partial reduction in RECG1 and the mild effect on ectopic recombination. However, in the *recG1-2* knockout, the specific amplification of an mtDNA discrete region was observed. Amplification was by a factor ranging from 4 to 8, depending on the pool of plants analyzed, and it was seen in all independent first generation homozygote plants tested. A detailed analysis showed that this stoichiometry change corresponded to the 8.0-kb region flanked by the 127-bp-long EE repeats (Figure 6B). This region contains the *ATP1* gene, coding for subunit alpha of the mitochondrial ATP synthase. The increase in copy number of this sequence resulted from the massive increase in crossover products EE-1/2 and EE-2/1 compared with the wild type (Figure 5) and was accompanied by a decrease of about 2-fold in the parental sequences EE-1/1 and EE-2/2 (Figures 5 and 6B). From this result, it was inferred that in *recG1-2*, recombination involving repeats EE produced an alternative form of the mitochondrial genome that coexists with the wild-type mtDNA. It consists in the deleted mtDNA sequence without the region between the EE repeats (EE-1/2 configuration; Figure 6D) plus an episome comprising *ATP1* (EE-2/1 configuration). Moreover, the EE-2/1 form was present at a significantly higher copy number, suggesting that once generated through recombination it is replicated autonomously and faster than the main genome.

Upon DNA gel blot analysis, an extra band that could correspond to the EE-2/1 circular episome was detected in nondigested DNA of *recG1-2*, in addition to the bulk of high molecular weight mtDNA (Figure 6C). As expected, this form was not detected with a probe hybridizing upstream of EE-1/1. Surprisingly, according to the signal intensities, the putative circular episome seemed to exist at a lower copy number than the high molecular weight mtDNA. Conversely, after restriction, the 2.7-kb *Bam*HI fragment specific of EE-2/1 was at least 4-fold more abundant than the wild-type-specific fragment of 7.7 kb according to phosphor imager quantification (Figure 6C), which was in agreement with the above qPCR results. This suggests that amplification of the recombination-generated EE-2/1 circular form produces high molecular weight DNA. It could be catenated forms of the replicating episome or tandem-repeated concatemer

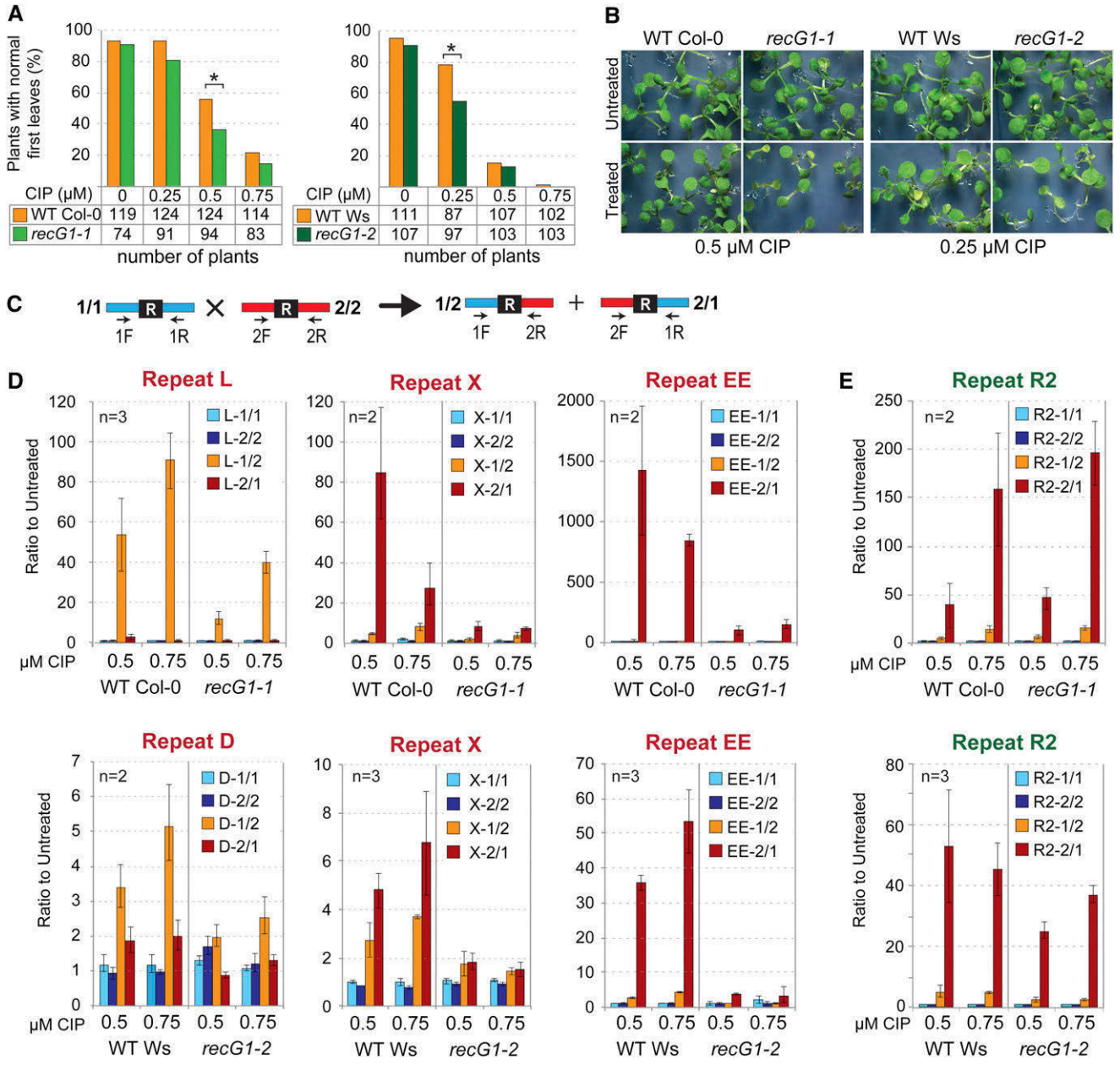


Figure 4. The *recG1* Mutants Are More Sensitive to Genotoxic Treatment and Are Deficient in HR-Dependent Repair of DSBs Induced by CIP.

(A) Percentage of plants developing normal first true leaves at the indicated concentrations of CIP. The number of plants analyzed is indicated under the graphs. Asterisks show statistical significance (unpaired *t* test; $P < 0.05$).

(B) Phenotypes of plants grown on agar plates containing the indicated concentrations of CIP. Most *recG1* plants did not develop normal first true leaves at CIP concentrations that had a mild effect on the wild type.

(C) Scheme of the qPCR relative quantification of parental sequences 1/1 and 2/2 that comprise a same repeated sequence (box R) and of the corresponding crossover products 1/2 and 2/1 using combinations of forward (F) and reverse (R) primers.

(D) Analysis of the accumulation of crossover products in CIP-treated wild-type Col-0, *recG1-1*, wild-type Ws, and *recG1-2* plants. qPCR results for repeats were normalized against the mitochondrial *COX2* and *RRN18* gene sequences. Results are the mean of two or three independent experiments, as indicated, and error bars correspond to SD values. The repair of DSBs induced by CIP in the mtDNA results in increased mobilization of IRs and concomitant accumulation of crossover products. This effect is strongly impaired in *recG1* plants.

(E) Accumulation of crossover products from chloroplast repeat R2, showing no significant effect of RECG1 in the repair of DSBs induced by CIP in the cpDNA. Results were normalized against the chloroplastic *RRN16* gene sequence.

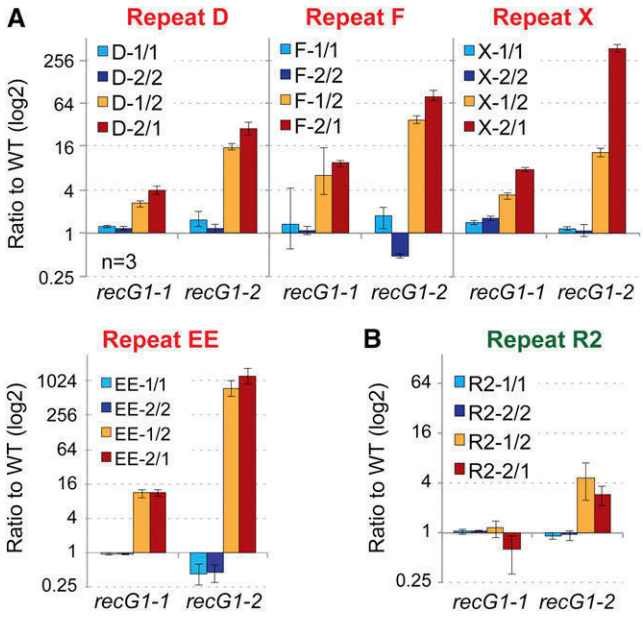


Figure 5. *recG1* Mutants Have Increased Ectopic Recombination across Mitochondrial Intermediate Size Repeats.

(A) As described in Figure 4C, quantification of the accumulation of crossover products from mitochondrial repeats D, F, X, and EE in *recG1-1* and *recG1-2* relative to the corresponding Col-0 and Ws wild-type accessions. Results are in a $\log_2$ scale.
(B) As in **(A)**, for chloroplastic repeat R2.

sequences generated by "rolling circle" replication, a process that has been proposed to occur in plant mitochondria (Backert et al., 1996). To confirm that the putative episome is indeed a circular molecule, it was digested with *Pac*I, an enzyme that cuts only once in the sequence and resolved on low percentage agarose gels (Figure 6C). *Pac*I digestion converted the molecular form migrating at an apparent size of 10 kb, which could correspond to a circular nicked molecule, into the expected linear fragment of 8 kb. Because the EE-2/1 recombination product contains *ATP1*, we checked whether its amplification results in increased expression of this gene. Indeed, a 3-fold increase in *ATP1* transcript was found in *recG1-2*, while no such increase was observed for other mtDNA gene transcripts (Figure 6E).

The cpDNA of *recG1-1* and *recG1-2* was likewise scanned for changes in sequence stoichiometry. No changes were detected between the mutants and corresponding wild-type accessions, apart from a slight general increase in cpDNA copy number (Supplemental Figure 4). Thus, loss of RECG1 does not seem to significantly affect the stoichiometric replication and segregation of the cpDNA in Arabidopsis.

RECG1 Works Synergistically with RECA3 in the Maintenance of mtDNA Stability

A comparison between the molecular effects observed in mutants of factors involved in mtDNA recombination suggests that RECG1 acts in the same pathways as the mitochondrial RecA-like protein RECA3. Both *recG1* (this work) and *recA3* (Miller-Messmer et al., 2012) mutants show increased ectopic recombination involving IRs in nonselective conditions, as well as reduced HR-dependent repair of CIP-induced mtDNA breaks. In contrast, other mutants described as involved in recombination surveillance and/or repair in plant mitochondria either are only affected in repair of induced DNA lesions (i.e., *odb1*, *why2*, and *pol1B*; Cappadocia et al., 2010; Parent et al., 2011; Janicka et al., 2012) or show only increased ectopic recombination in normal growth conditions (i.e., *osb1*; Zaegel et al., 2006). We tested the effect of the double mutation *recG1-1 recA3-2* (both mutations in the Col-0 background) on the maintenance of the Arabidopsis mtDNA. While the parental plants used for reciprocal crossing had no visible developmental phenotypes, the homozygous F2 double mutants were all affected, to different extents, by slow growth, deformed and variegated leaves, and partial or complete sterility (Figure 7A). We scanned individual plants for changes in the stoichiometry of the different mtDNA regions, as described above. All *recG1-1 recA3-2* plants tested showed substantial changes in the stoichiometry of mtDNA sequences compared with wild-type plants (Figure 7B). An increase in copy number of large regions was mainly observed that could be as high as 12-fold in some cases (e.g., plant #6). The mtDNA sequences affected and the magnitude of the changes varied from plant to plant, however, with an involvement of common specific regions. Several of the observed events of stoichiometric shift could be explained by the same process as the one involving the EE repeats in *recG1-2*: a looping out of subgenomic regions by recombination across directly oriented IRs, followed by their autonomous replication. The increase of the EE repeat regions was also observed in several plants, although it had not been seen in the hypomorphic *recG1-1* single mutant. Thus, reduction in RECG1 has a synergistic effect with the *recA3* mutation, leading to the generation by recombination of subgenomes whose replication is apparently not coordinated in these mutants.

Backcrossing Elicits Segregation of the Alternative mtDNA Forms

We have found that in *recG1-2* there is coexistence of two alternative mtDNA configurations interconvertible through a single recombination event involving the EE repeat. These are the wild-type mtDNA containing the EE-1/1-EE-2/2 region and the alternative mitotype comprising the mtDNA deprived of the EE region (EE-1/2) plus the EE-2/1 episome (Figure 6D). The simplicity of this genomic configuration makes it a convenient model to study the mechanisms and dynamics of mtDNA evolution by stoichiometric shifting and to determine how RECG1 can affect these processes. For such studies, we first analyzed the stability of the different genomic environments and their possible segregation in the wild-type and *recG1-2* populations. We found that the ratios between the main genome sequences and crossover products EE-1/2 and EE-2/1 were very well conserved among individual wild-type plants (Figure 8A), confirming the tight control that exists on mtDNA recombination.

Next, we tested whether the coexistence of the alternative mitotypes was stable in the *recG1-2* line. All first homozygous mutant plants that were tested had the molecular phenotype of increased copy number of the EE-2/1 sequence. In the progeny of one of these plants, we tested the possible segregation of the different EE configurations. More than 100 plants from generation

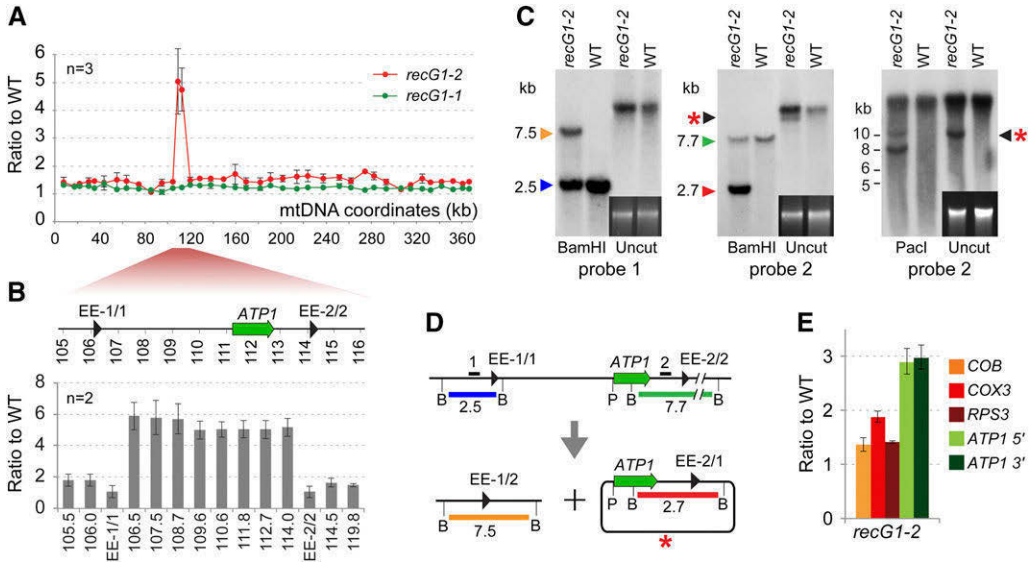


Figure 6. Recombination-Mediated Segregation of a Small Replisome Containing the *ATP1* Gene in the *recG1-2* Knockout.

(A) Scanning of the mtDNA for changes in relative copy numbers of the different mtDNA regions. Sequences spaced 5 to 10 kb apart on the mtDNA were quantified by qPCR on *recG1* mutants and corresponding wild-type accessions. Results are the mean and SD values from three biological replicates. Coordinates are those of the published Col-0 mtDNA sequence.

(B) Detailed analysis of the region amplified in *recG1-2*, using probes spaced 1 kb apart and primers that specifically amplify repeated sequences EE-1/1 and EE-2/2.

(C) DNA gel blot hybridization of wild-type and mutant DNA from inflorescences with probes external (1) and internal (2) to the amplified region. The ethidium bromide-stained uncut DNA is shown as loading control.

(D) Schematic interpretation of the results of **(C)**. *Bam*HI and *Pac*I sites are indicated and the relevant restriction fragments are color-coded. The thick lines labeled 1 and 2 indicate the sequences used as probes. Recombination involving direct repeats EE produces an alternative genome deleted of the region comprising *ATP1* and a circular sequence containing *ATP1* that is autonomously replicated and amplified in the *recG1-2* line. A red asterisk indicates the signal corresponding to the EE-2/1 episome in uncut DNA. *Pac*I digestion showed that it is a circular molecule.

(E) Increased accumulation of *ATP1* transcripts correlates with the increased gene copy number. The transcripts of several mtDNA-encoded genes including *ATP1* were quantified by RT-qPCR. Two different *ATP1* sequences from the 5′ and 3′ parts of the gene were tested.

T5 were tested by competitive PCR, but in none of these plants we found evidence for differential segregation of the alternative mitotypes. From these plants, 42 were further analyzed by qPCR, which confirmed that the ratios between the alternative EE sequences were relatively stable in the *recG1-2* background (Figure 8A). These results contrast with the data obtained for other Arabidopsis mutants affected in mtDNA recombination surveillance, including *osb1* and *msh1*, for which shifts in mtDNA structure could be observed between individuals in the mutant population, as well as segregation of plants developing specific phenotypes aggravated in late mutant generations (Zaegel et al., 2006; Shedge et al., 2007; Davila et al., 2011).

We next verified whether reintroduction of the *RECG1* allele could promote reversion to the wild-type mtDNA stoichiometry. The *recG1-2* mutant was backcrossed with wild-type Ws as pollen donor and F1 plants were analyzed. Surprisingly, among the F1 plants, ~20% displayed characteristic phenotypes of leaf distortion, variegation, slow growth, and reduced fertility (Figure 8B). These phenotypes were similar to the ones observed in the *recG1-1 recA3-3* plants and in other mutants affected in mtDNA recombination surveillance. Quantification of the EE parental genomic sequences and crossover products in the individual F1 plants revealed that all had undertaken sorting of the alternative genomes (Figure 8C). These results were reproduced in two independent backcross experiments. Among 76 F1 plants analyzed, one-third (26/76) had drastically reduced copy numbers of the recombined genomic sequences (deleted mtDNA sequence EE-1/2 and episome EE-2/1), while maintaining wild-type levels of the parental sequence EE-1/1-EE-2/2. An additional third of the plants (22/76) had intermediate levels of mtDNA reversion, with normal EE-1/1-EE-2/2 copy numbers and significant reduction of EE-1/2 and/or EE-2/1. The recombined genomic sequences showed a differential segregation in many plants, which confirmed that they were not genetically linked. The loss of EE-2/1 in revertants was thus not necessarily through recombination with EE-1/2, which should result in stoichiometric reduction of both sequences, but rather occurred by suppression of its replication or transmission.

The analysis of the F1 plants displaying variegated phenotypes showed that most had asymmetrically segregated the alternative EE sequences (Figure 8C), with simultaneous reduction of the two configurations carrying the *ATP1* gene (EE-1/1-EE-2/2 and episome EE-2/1). This suggested that the observed phenotype resulted from a reduction in ATP1 expression. Indeed, the reduction in copy number of the *ATP1* gene was confirmed in all variegated plants tested (Figure 9A). Moreover, quantification of *ATP1* in different leaves of a chimeric plant containing both normal-looking

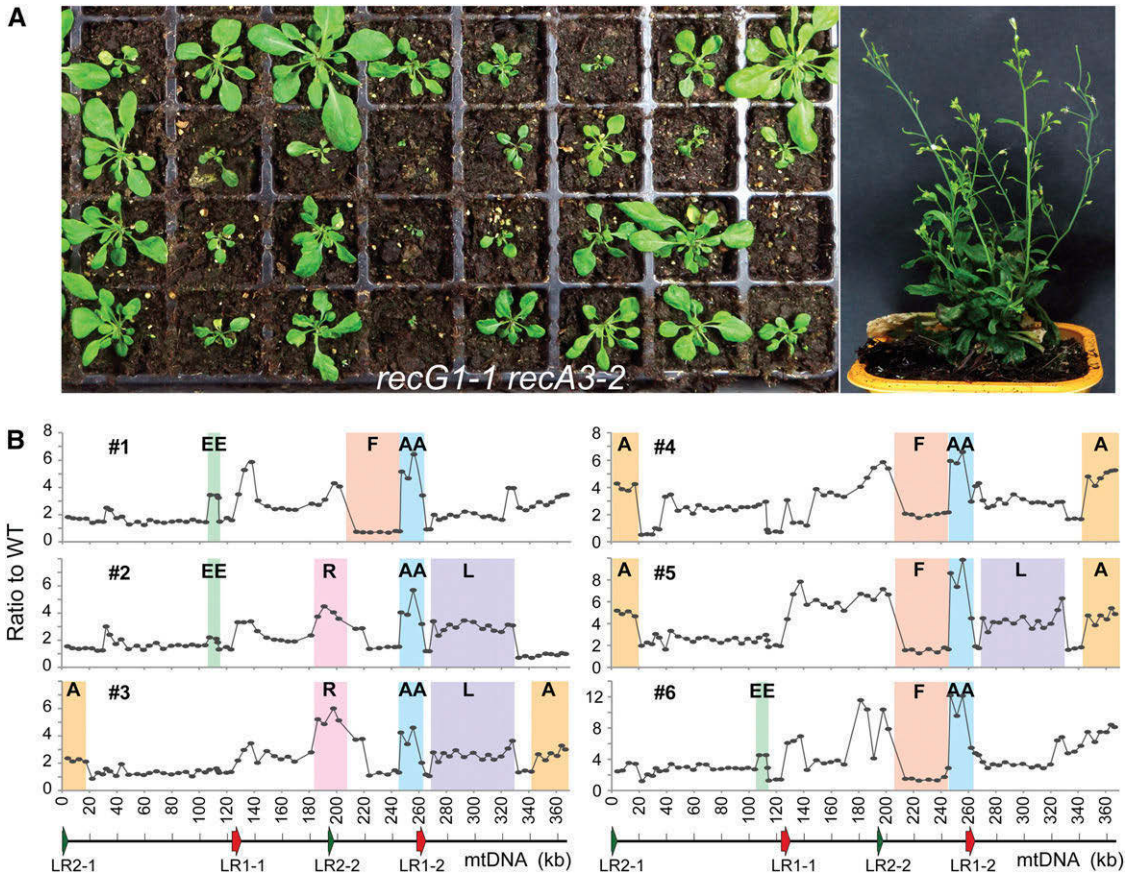


Figure 7. Synergistic Effect of *recG1* on *recA3*.

Knockdown of *RECG1* potentiates the *recA3* mutation, which results in plants that segregate different mitotypes resulting from recombination.
(A) Variable severity of the phenotypes developed by *recG1-1 recA3-2* double mutants. The plant on the right illustrates a severe phenotype of dwarfism, leaf distortion, variegation, and sterility.
(B) Changes in the copy number and stoichiometry of mtDNA sequences in individual *recG1-1 recA3-2* plants. The mtDNA was scanned for differences in copy number as described in Figure 6A. The linear representation of the Col-0 mtDNA is displayed below the graphics, with the coordinates in kilobases and the two pairs of large repeated sequences LR1 and LR2. Colored regions indicate sequences whose increase or decrease can be explained by the loop-out of subgenomes by recombination involving the indicated direct repeats, followed by their autonomous replication and/or segregation.

and variegated leaves (plant F1-51; Figure 8B) confirmed that the severity of the phenotype correlated with the reduction in *ATP1* (Figure 9A). The decrease in gene copy number also correlated with a significant reduction of the *ATP1* transcript (Figure 9B). Detection of the ATP1 protein in immunoblots of total leaf extracts showed that the *recG1*-2 mutant had normal or slightly higher amounts of ATP1 but that the protein was significantly reduced, by a factor of 2-fold, in the variegated plant F1-55 (Figure 9C). The limited reduction in protein compared with the dramatic reduction in gene transcript is possibly the result of the differential turnover rates. Therefore, the variegated phenotype apparently resulted from clonal asymmetrical segregation of the alternative mitotypes, with a compromised oxidative phosphorylation in tissues inheriting below-threshold levels of *ATP1*. To confirm this hypothesis, we measured the respiration capacity of normal and severely affected variegated leaves, of equivalent size and weight. We did indeed find that the affected tissues had much reduced respiration rates compared with normal leaves (Figure 9D).

Remarkably, among the F1 plants, a significant proportion (14/76) displayed completely shifted mtDNA genomes, having almost completely lost the wild-type genome (EE-1/1-EE-2/2) and retained the recombined forms EE-1/2 and EE-2/1. These plants had a new near-homoplasmic mtDNA and were phenotypically indistinguishable from wild-type plants. Hybridization to undigested DNA showed that the fast-migrating mtDNA form (the EE-2/1 episome) was lost in the F1 shifted plants (Figure 10A). The high molecular weight EE-2/1 configuration that was retained either corresponded to a head-to-tail concatemeric autonomous replicon or resulted from a reintegration of the EE-2/1 sequence into a new locus of the main genome. The latter hypothesis was confirmed by DNA restriction analysis, as well as by qPCR (Figure 10). According to these experiments, the EE-2/1 sequence reintegrated the genome through recombination involving the repeated sequence T, which is present just upstream of *ATP1* and is repeated elsewhere in the mtDNA (sequences T-1/1 and T2/2 respectively; Figure 10B). The product of this secondary

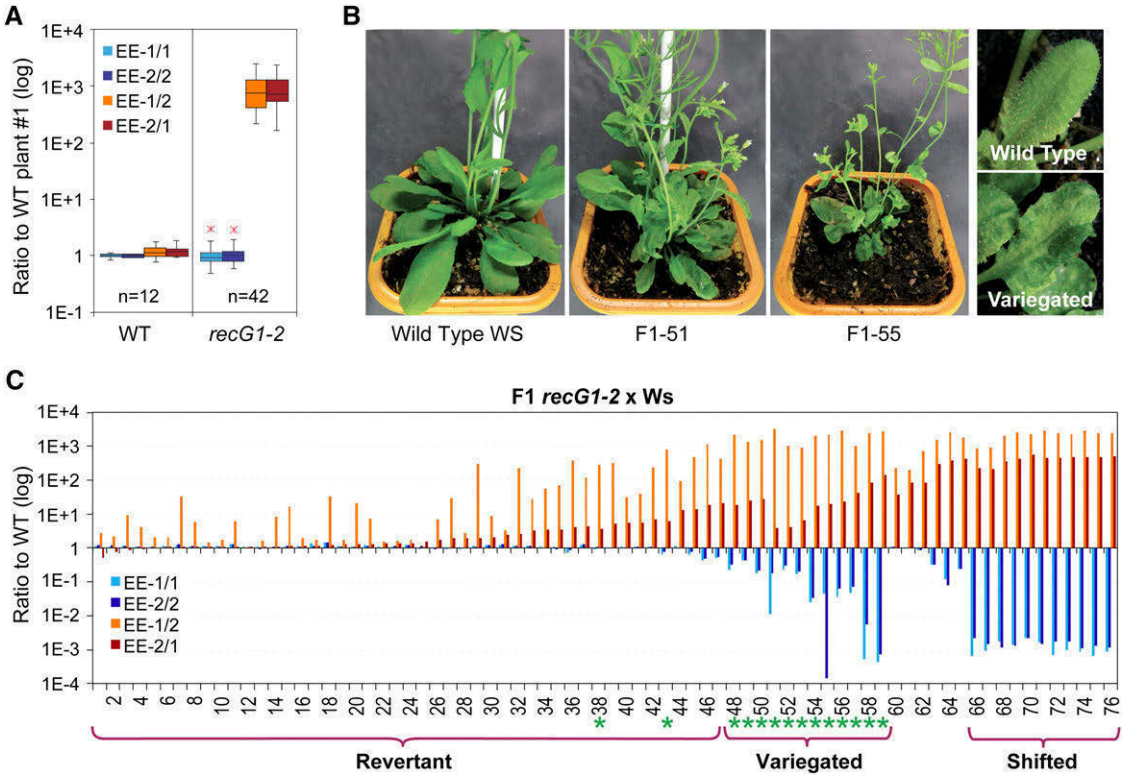


Figure 8. Backcross Triggers Genomic Shift Involving EE Repeats.

The stable mtDNA structure found in *recG1-2* knockout plants (Figure 6D) is destabilized by reintroduction of the RECG1 allele by backcross, resulting in segregation of plants with different mitotypes.

(A) Stable ratios of the sequences derived from repeats EE in wild-type and *recG1-2* plants. The values were compared with a same individual wild-type plant.

(B) Examples of F1 plants from backcross presenting phenotypes of leaf variegation and distortion.

(C) Analysis of the EE configurations, as in Figure 5, in 76 individual F1 plants. Green asterisks indicate plants displaying the variegation phenotypes shown in **(B)**. Revertants are plants that have significant reduction in the EE-2/1 subgenome containing sequence. Shifted plants are the ones that retain the EE-1/2 and EE-2/1 subgenomes and have nearly lost the EE-1/1 and EE-2/2 parental sequence.

recombination event is already present in homozygous *recG1-2* plants, giving a 5.4-kb diagnostic *Spe*I restriction fragment hybridizing with the *ATP1* internal probe (Figure 10A). This modified genome was preferentially segregated in the shifted plants, to the detriment of the wild-type mtDNA, which was virtually lost. The qPCR analysis supports this interpretation, showing a 9-fold increase of the T-1/2 and T2/1 recombined forms in shifted plant compared with homozygous *recG1-2* plants, with a corresponding loss of the T-1/1 and T-2/2 wild-type sequences (Figure 10C). Thus, our analysis shows that in *recG1-2* F1 plants, there was a differential segregation of three alternative mitotypes that were already present in the homozygous mutant: (1) the wild-type mtDNA, (2) the mitotype subdividing into the EE-1/2 form and the EE-2/1 episome, and (3) the mtDNA resulting from the reintegration of EE-2/1 into the main mtDNA by a secondary recombination event involving repeats T-1/1 and T-2/2.

Finally, analysis of the F2 generation of the shifted plants showed that the mtDNA remained stable in all *RECG1:recG1-2* and *RECG1:RECG1* plants, with no further changes in the ratios of the EE-containing configurations (Figure 10D). However, in all F2 *recG1:recG1* homozygous plants (13 plants analyzed), there was a remarkable reincrease of the parental EE-1/1-EE-2/2 sequence, by a factor of ~100-fold. Thus, in these plants, a new cycle of mtDNA changes by recombination was again elicited by the loss of RECG1.

DISCUSSION

Multiple Activities of RecG-Like Proteins

Our results show that RECG1 acts at multiple stages of mtDNA maintenance, both in recombination-dependent repair and in the surveillance of recombination activities that compromise the integrity of the organellar genomes. In addition, characterizing the segregation of the mitotypes coexisting in *recG1* plants allowed us to address the mechanisms underlying mtDNA stoichiometric shifting and to better understand how these can lead to stable novel mtDNA configurations.

According to our computer modeling, the plant RECG1 appears to be structurally homologous to bacterial RecG, which suggests that it shares similar activities. This is supported by the ability of the

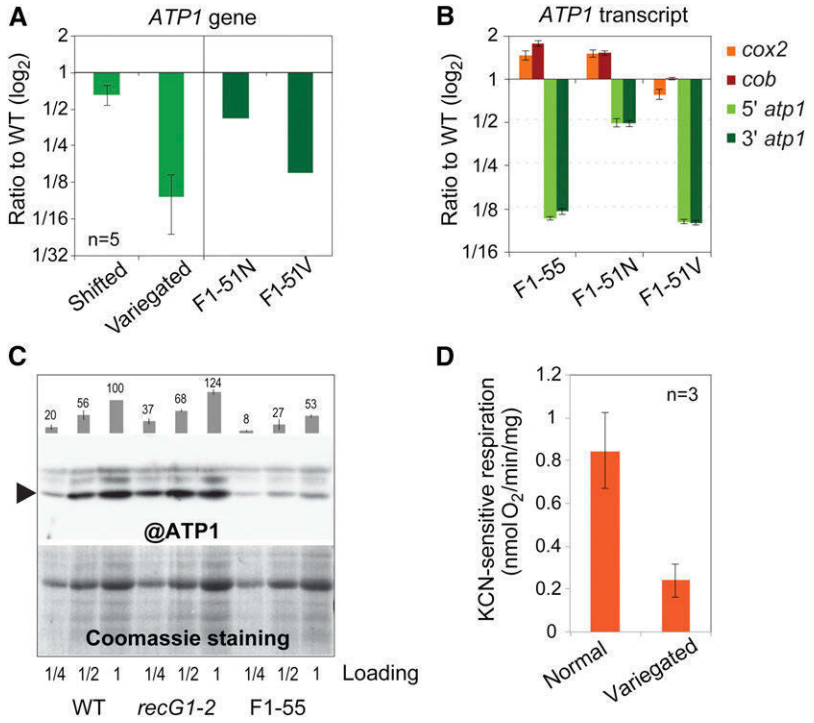


Figure 9. The Variegation Phenotypes Result from Reduced Transmission and Expression of the *ATP1* Gene.

(A) Correlation between variegation phenotypes and reduction in *ATP1* copy number in F1 plants. The results are the means and SD from five plants analyzed. The results obtained for a variegated leaf (V) and a normal leaf (N) from plant F1-51 are shown.

(B) As in **(A)**, but for the abundance of the *ATP1* transcript, compared with two other mtDNA-encoded genes. Two different *ATP1* sequences from the 5′ and 3′ parts of the gene were tested.

(C) Immunodetection and relative quantification of the ATP1 protein (@ATP1), showing a correlation between the affected phenotype of plant F1-55 and a reduced steady state level of ATP1 versus the wild type and *recG1-2*. The Coomassie blue staining of the membrane is shown as loading control. The arrowhead indicates the protein migrating at the expected size of 55 kD.

(D) Measure of the KCN-sensitive oxygen consumption in 3-week-old normal or variegated leaves.

plant sequence to replace the bacterial protein in all functional tests that we performed. One can speculate that the few extra sequences that are present in RECG1 might modulate plant-specific interactions and probably do not change the overall modular organization of the protein and its activity. Thus, RECG1 possibly has similar functions in plant organelles as RecG in bacteria.

Bacterial RecG has multiple roles in DNA metabolism and can target a plethora of different nucleic acid structures, including replication forks, HJs, D-loops, and R-loops (Supplemental Figure 5). RecG is associated with later stages of repair by recombination because of its ability to catalyze branch migration of HJs, as does RuvAB. But RecG is not redundant with RuvAB, since it has no associated nuclease activity that could resolve HJs, whereas RuvAB recruits the RuvC nuclease (Zhang et al., 2010). Analysis of recombination intermediates also revealed independent effects of *ruvAB* and *recG* mutations in the resolution of HJs (Mawer and Leach, 2014). As RecG in bacteria, in plant organelles, RECG1 should be able to translocate HJs and have major roles in their resolution. But in plants there are no orthologs of RuvABC, which conversely suggests that plant RECG1 can take over RuvAB functions. However, we found that RECG1 cannot complement *ruvAB* in bacteria, and functional homologs that act in the nucleus such as the RAD54 translocase and the GEN1 and SEND1 resolvases are apparently not targeted to organelles (Osakabe et al., 2006; Bauknecht and Kobbe, 2014). RECG1 should then perform in cooperation with a not yet identified resolvase. Alternatively, RECG1 alone might be able to eliminate HJs, simply by driving branch migration till they merge with replication forks, as has been suggested for bacterial RecG (Wardrope et al., 2009).

An additional major function of bacterial RecG is the suppression of replication initiation, associated with its ability to dissociate R-loops and D-loops. This allows RecG to repress replication of plasmids that require an RNA primer for replication initiation (Vincent et al., 1996; Fukuoh et al., 1997) and to suppress potentially pathological replication reinitiation, such as when converging replication forks meet (Rudolph et al., 2010, 2013). Our complementation results demonstrated that plant RECG1 can ensure such functions in bacteria, which suggests that it can dissociate R-loops and D-loops. As discussed below, such an activity could be responsible for the loss of the EE episome, when the *RECG1* allele is reintroduced by backcross. Finally, bacterial RecG can convert three-way to four-way DNA junctions, an ability that has mainly been associated with the rescue of stalled replication forks and the bypass of blocking

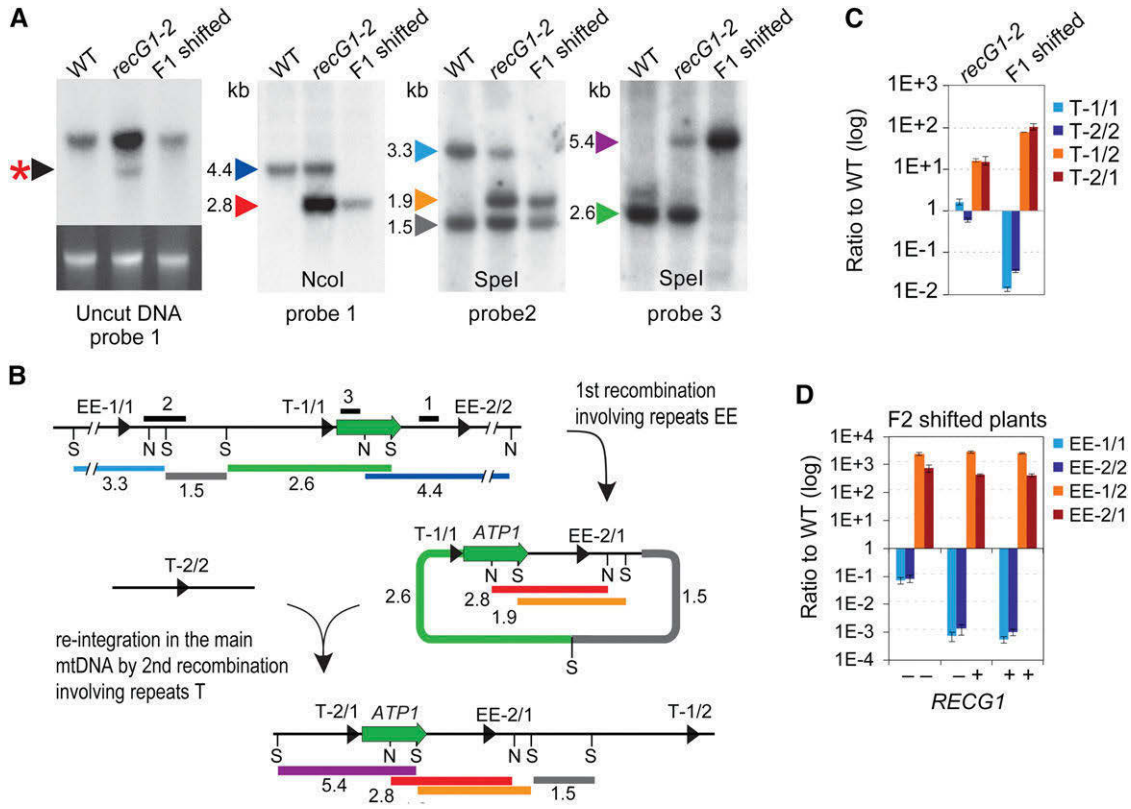


Figure 10. Genomic Shift Results in a New mtDNA Genome.

(A) and **(B)** DNA gel blot hybridizations **(A)** supporting the model **(B)** proposed to explain how F1-shifted plants inherited a new mtDNA configuration by reintegration of the EE-2/1 subgenome into a new genomic locus of the main mtDNA, via recombination involving repeats T-1/1 and T-2/2. The asterisk indicates the EE-2/1 replisome visible in uncut DNA of homozygous *recG1-2* and lost in shifted F1 plants. The arrowheads indicate hybridizing restriction fragments with a color code corresponding to the fragments represented in the explaining scheme. *Spe*I (S) and *Nco*I (N) sites are indicated. The thick black lines labeled 1, 2, and 3 indicate the sequences used as probes.

(C) Relative qPCR quantification of the parental configurations (T-1/1 and T2/2) of mitochondrial repeat T and of the corresponding crossover products (T-1/2 and T-2/1) in *recG1-2* and F1 shifted plants, as outlined in Figure 4C.

(D) Relative qPCR analysis of the EE sequences in the F2 progeny of F1-shifted plants. The genotype regarding the *recG1-2* mutation is indicated.

lesions (McGlynn and Lloyd, 2000; McGlynn et al., 2001). Whether RECG1 is involved in the rescue of stalled forks during replication of the organellar genomes remains to be established but would have implications discussed below.

Roles of RECG1 in Controlling Proper Homologous Recombination in Plant Organelles

We observed that Arabidopsis *recG1* mutants display several distinct molecular phenotypes. First, under unrestricted growth conditions, they accumulate mtDNA sequences generated by ectopic HR, in which IRs are used as loci of strand invasion (Figure 5). In wild-type plants, these ectopic recombination events are tightly suppressed. This same molecular phenotype has been observed in several other mutants, affected in the expression of unrelated proteins, including the RecA-like mitochondrial recombinases RECA2 and RECA3, the single-stranded DNA binding protein OSB1, and the MutS-like protein MSH1 (Zaegel et al., 2006; Arrieta-Montiel et al., 2009; Shedge et al., 2010; Miller-Messmer et al., 2012). This implies that these proteins participate in similar pathways of recombination and that when these pathways are compromised, there is activation of error-prone recombination activities that are normally repressed. These are most likely break-induced replication-related activities that result in increased ectopic recombination due to the repeat-rich mitochondrial genome of plants (Christensen, 2013). The pathways of recombination that are compromised in mutants of these proteins are RecA-dependent, since loss of either RECA2 or RECA3 also leads to the same molecular phenotype (Miller-Messmer et al., 2012).

Second, the *recG1* mutants are more sensitive to CIP treatment, which induces DSBs in the organellar genomes (Figure 4). Induction of HR-dependent mtDNA repair in wild-type plants also results in increased mobilization of IRs as possible sites of strand invasion. This effect is reduced in *recG1* plants, as it is in other mutants impaired in mtDNA repair by recombination (Janicka et al., 2012; Miller-Messmer et al., 2012). Therefore, it seems that the induction of DSBs by CIP treatment in wild-type lines

phenocopies the effect of mutations that induce ectopic recombination in regular growth conditions. It suggests that the effect of such mutations is an overall increase in mtDNA breaks that is mimicked by the genotoxic treatments. The increase could result from the inability to repair DNA lesions by break-avoiding pathways.

These pathways could be involved in the recovery of blocked replication forks. In organelles, the main threats to DNA integrity are modified bases and cross-links that are frequently generated in the oxidative environment of mitochondria and chloroplasts. Such lesions can block DNA polymerase progression, challenging the duplication of the organellar genomes. If the stalled forks cannot be rescued, by repair or bypass of the blocking lesion, the replisome might dissociate, resulting in fork collapse. Further DNA synthesis will then require recombination to restart replication (Supplemental Figure 6). It is probable that RECG1, like bacterial RecG, promotes fork regression to bypass unrepaired replication blocks. Other factors whose mutations result in increased ectopic recombination might also be required for efficient replication repair and restart of stalled replication forks, avoiding fork collapse. Conversely, failure to recover stalled forks in *recG1* or other mutants would result in DNA breaks. The broken DNA ends would then be recovered for replication restart by recombination through the break-induced replication pathway, which is error-prone and can give rise to gross rearrangements, copy number variations, and loss of sequence synteny if it occurs in the context of repeated sequences (Saini et al., 2013). This scenario can explain both the increase in ectopic recombination in *recG1* and other mutants and the phenocopy of mutation effects by drugs that induce DSBs, as discussed above. RecA-like proteins also promote replication fork reversal, in particular in the presence of single-strand gaps that result from the blocking of lagging strand progression (Robu et al., 2004; Cox, 2007). Contrarily, RecG should preferentially target forks blocked on the leading strand (Supplemental Figure 6). In the context of plant organelles, complementary and partially redundant functions of RecA proteins and RECG1 in the rescue of stalled replication could explain why (1) the single *recG1* mutation has a mild effect on organellar genome maintenance and does not impair normal plant growth, and (2) the double mutation *recG1-1 recA3-2* has a synergistic effect on the stoichiometric replication of the Arabidopsis mtDNA.

Finally, our results also suggest that RECG1 is able to displace D-loops and R-loops, an activity that allows bacterial RecG to selectively support recombination at locations with extensive DNA sequence homology (Vincent et al., 1996; Fukuoh et al., 1997). Indeed, it was shown in bacteria that recombination between correctly aligned sister sequences involves an unstable intermediate that is stabilized by RecG-dependent branch migration, which converts three-way to four-way DNA junctions (Mawer and Leach, 2014). It can thus be speculated that branch migration of D-loops catalyzed by RECG1 might help to discriminate between regions of limited or extensive sequence identity for strand invasion in plant mitochondria, preventing undesirable genomic rearrangements. The loss of RECG1 would in turn result in increased ectopic recombination.

No Significant Effect of RECG1 in the Maintenance of the cpDNA in Arabidopsis

RECG1 is apparently dually targeted to mitochondria and chloroplasts, as shown through both transient and constitutive expression of a GFP fusion protein in plants. The RECG1 promoter-GUS fusion results also showed important gene activity in photosynthetic tissues. It was therefore expected that RECG1 would exert the same functions in both organelles and affect the maintenance of both the mtDNA and the cpDNA. However, in *recG1* knockdown and knockout plants, we have not seen any significant effects on the copy number of the cpDNA, on the stoichiometry of the cpDNA regions, or on the surveillance of recombination mobilizing repeated sequences. These observations contrast with recently published results on the *P. patens* ortholog of RECG1, which is also dually targeted to both organelles and whose deletion caused an increase in recombination mobilizing mtDNA IRs of 47 to 79 bp, but also cpDNA IRs of 48 to 63 bp (Odahara et al., 2015). There are several possible explanations for these functional differences. First, it cannot be excluded that overexpression under the strong 35S promoter results in mistargeting of the GFP fusion protein into chloroplasts. Second, it has already been seen with dually targeted proteins that one location can be favored versus the other (Carrie and Whelan, 2013). That seems to be the case for Arabidopsis RECA2, which is dually targeted but appeared to affect only the maintenance of the mtDNA (Shedge et al., 2007; Miller-Messmer et al., 2012). Third, the organellar recombination machineries of *P. patens* and Arabidopsis show further differences, and not all intervening factors are shared between the two species. For instance, RECA3 and WHY proteins, which are important for organellar HR-dependent repair in Arabidopsis, are not found in *P. patens*. Finally, it is possible that the effects of RECG1 on the maintenance of the Arabidopsis cpDNA are not visible: Long IRs of more than 100 bp of perfect identity, which are fairly abundant in the Arabidopsis mtDNA and are the major sites of ectopic HR, are not present in the cpDNA. The chloroplast repeated sequence R2 that we monitored is a 76-bp imperfect repeat with five mismatches, whose two copies only share a stretch of 32 bp of perfect identity. It is therefore unlikely to constitute an efficient substrate for HR. In Arabidopsis mitochondria, efficient mobilization of such small repeats for recombination has not been observed (Davila et al., 2011). On the contrary, mitochondrial and chloroplastic repeated sequences that are active in ectopic HR in *P. patens* are essentially small (Odahara et al., 2009, 2015), definitely reflecting differences between the organellar recombination machineries of the two species. In this context, it might be noted that recombinational activities that reshuffle mtDNA gene arrangements only came with the onset of spermatophytes (Knoop, 2013). Finally, in CIP-treated wild-type plants we have seen that recombination involving R2 is significantly increased, indicating that this repeat can be efficiently mobilized during HR-dependent repair of cpDNA DSBs. However, the absence of RECG1 in mutant plants does not significantly affect this mobilization, suggesting that it is mediated by other factors.

Multiple Mitotypes in *recG1* Plants

In this study, we observed that in *recG1-2* plants the increased ectopic recombination produced an episome that apparently could replicate autonomously, resulting in the generation of two alternative genomes differing in a single locus. In the *recG1-2* background, there is a stable coexistence of these alternative

mitotypes, at least for two plant generations. It is not clear why additional autonomous subgenomes were not created in the mutant. One possible reason is that EE repeats are the IRs closest to each other in the mtDNA and that recombination involving repeats far apart is much less efficient. The small size of the EE-2/1 episome should also support a faster replication rate. How replication of the EE-2/1 episome initiates is not known, but the presence of the very actively transcribed *ATP1* gene might help, either by locally melting the DNA and facilitating strand invasion for recombination-dependent replication or by providing RNA primers for replication initiation. In *recG1-1 recA3-2* plants, we observed additional changes in the stoichiometry of several mtDNA regions, not occurring in single mutants and explainable by similar processes of recombination involving directly oriented IRs and independent replication of the resulting products. This synergistic effect of the *recG1-1* hypomorphic mutation and the *recA3* knockout mutation implies that the two genes act in partially redundant pathways that control the correct stoichiometric transmission of the mtDNA. Independent and preferential replication explains how these sequences that are generated in somatic tissues at relatively low frequencies can become predominant. While qPCR data show that virtually all pairs of IRs are targets of increased ectopic recombination, only a few of them promote significant changes in the copy number of the sequences comprised between the repeats. These might be the few recombination products that are able to replicate and segregate autonomously.

A Model for mtDNA Segregation and Stoichiometric Shifting

The alternative mitotypes discussed above segregated in the F1 plants of the *recG1-2* backcross, after reintroduction of the *RECG1* allele. The analysis of these events allowed us to understand characteristics of stoichiometric shifting and generation of a new stable mtDNA structure compatible with normal plant development. We could also directly link phenotypes of leaf distortion and variegation to a reduced expression of a gene essential for oxidative phosphorylation. The subpopulation of F1 plants that we classified as variegated developed specific developmental phenotypes that also occur in several other Arabidopsis mutants impaired in mitochondrial recombination functions (Sakamoto et al., 1996; Zaegel et al., 2006; Arrieta-Montiel et al., 2009). As we observed, these phenotypes are also widespread in *recG1-1 recA3-2* plants. In mutant and double mutant plants that have been previously described, the penetrance and severity of phenotypes correlated with increased mtDNA rearrangements by recombination. However, given the variability and extension of the mtDNA changes in those plants, it has not been possible to directly link phenotypes to specific changes in organellar gene expression. On the contrary, in the variegated F1 plants, the phenotypes clearly correlated with the loss of the *ATP1* gene and the reduction of the corresponding transcript. This should result in a significant decrease in ATP1 protein and oxidative phosphorylation. Thus, we show here that the likely explanation for these phenotypes is the loss of mtDNA regions coding for genes essential to OXPHOS functions. The clonal tissue distribution, with regions that are more or less affected depending on the differential inheritance of the deleted mtDNA, recalls what has been observed in NCS (nonchromosomal stripe) maize (*Zea mays*) mutants (Gu et al., 1993).

In the *recG1-2* F1 plants, there was a segregation of mtDNA mitotypes that was not observed in homozygous *recG1-2*, at least up to generation T5. Therefore, it seems that sorting of the alternative mitotypes coexisting in *recG1-2* was not possible in the absence of the *RECG1* allele. In almost all segregating F1 plants, the sorting of the alternative mitotypes correlated with a highly reduced copy number of the EE-2/1 episome, either because of its loss in revertants or because of its integration into the mtDNA by secondary recombination in the shifted plants. This suggests that sorting was contingent on the ability of mitochondria to lose the EE-2/1 episome, in a RECG1-dependent manner. That could be because of multiple reasons. The ability of RECG1 to dissociate D-loops or R-loops could explain how its reintroduction by backcross suppresses the replication of the episome. As we have shown, expression of RECG1 in bacteria is able to repress the copy number of ColE1-dependent plasmids, which require an RNA primer for replication initiation. Indeed, one of the possible functions of RECG1 in mitochondria could be to suppress replication of aberrant mtDNA sequences, if the process involves an R- or D-loop. Another possible explanation is that the EE episome might be catenated with the main mtDNA molecules, from which it cannot dissociate during replication in the absence of RECG1. As we have seen, a major part of the EE-2/1 sequence in *recG1-2* comigrates with high molecular weight DNA. It can be hypothesized that these sequences would correspond to duplications in the main mtDNA genome. However, in that case, we would not expect to see a reduction of the EE-1/1 and EE-2/2 sequences compared with the neighboring mtDNA regions, as it was observed (Figures 5A and 6B). A possible catenation of the EE-2/1 episome better explains these results. Therefore, it is possible that RECG1 cooperates in the decatenation of the DNA, as it has already been described for other helicases (Ramamoorthy et al., 2012).

Finally, our backcross experiments generated shifted plants with a stable rearranged mtDNA, differing from the original mtDNA by the transposition of a specific sequence to a different locus. This study gave us a comprehensive scenario of how the mtDNA can rapidly evolve by stoichiometric shifting, in two steps (Figure 9B). First, recombination involving directly oriented IRs engenders subgenomes whose copy number significantly increases by autonomous replication. In a second step, the high copy number reached by these subgenomes favors additional recombination events that permit their reintegration into a different locus of the main genome, changing the structural organization of the mtDNA sequences. Factors such as RECG1 modulate such processes, by controlling the recombination frequency and the replication of subgenomes.

METHODS

Plant Material and Growth Conditions

Arabidopsis thaliana T-DNA insertion lines *recG1-1* (WiscDsLox485-488E5) and *recG1-2* (FLAG_630A09) were obtained from the Nottingham Arabidopsis Stock Centre and from the Versailles Arabidopsis Stock Centre, respectively. Plants were grown on soil or in vitro at 23°C under

a 16-h light photoperiod. For in vitro cultures, surface-sterilized seeds were sown on agar plates containing MS255 (Duchefa) supplemented with 1% (w/v) sucrose and stratified during 3 d at 4°C.

Phylogenetic Analysis

Bacterial and plant sequences were identified in the databases using the *Escherichia coli* RecG sequence as query. An alignment was constructed with ClustalW implemented in the MacVector package using the GonneT matrix. Phylogenetic trees were built with the PhyML v3.1 online software (www.phylogeny.fr; Dereeper et al., 2008) using the neighbor-joining method implemented in the BioNJ program and 100 bootstrap support. Graphical representations were performed with TreeDyn (v198.3).

E. coli Complementation Assays

The *E. coli* DH5α strain was used for routine cloning, whereas the JM103 (*recG*+ and *ruvA*+), JM103 *recG::kan* (Δ*recG*), and JM103 *ruvA::Tc* (Δ*ruvA*) strains were used for complementation assays. The *RECG1* cDNA without targeting sequence (codons 58 to 957) was synthesized (http://www.genecust.com) and transferred into the *Pac*I and *Asc*I restriction sites of *pACYClacz* under the control of the isopropyl β-D-1-thiogalactopyranoside (IPTG)-inducible *lac* promoter. (Miller-Messmer et al., 2012). Codons were adapted to tRNA frequency in bacteria. The resulting construct was introduced in Δ*recG* and Δ*ruvA* strains, and complementation results were compared with wild-type and mutant strains transformed with the empty *pACYClacz* vector. Induced expression of the recombinant protein was monitored with a polyclonal antibody directed against RECG1. To raise the antibody, the sequence corresponding to the last 260 C-terminal amino acids of Arabidopsis RECG1 was cloned into the p0GWA expression vector (Busso et al., 2005). The recombinant protein containing a C-terminal His-tag was expressed in the Rosetta(DE3)pLysS *E. coli* strain (Novagen), purified under denaturing conditions on Ni-NTA agarose (Qiagen), and injected into rabbits.

For DNA repair assays, 5 mL of Lysogeny broth supplemented with 1% (w/v) glucose and 10 μg/mL chloramphenicol was inoculated with 25 μL of overnight cultures and bacteria were grown at 37°C. Expression of RECG1 was induced at A_{650} = 0.2 by the addition of 0.25 mM IPTG and bacteria further grown till A_{650} = 0.4. Appropriate dilutions were spread on Lysogeny broth plates in triplicate and exposed to different UV-C doses (checked with an UV radiometer) before incubation at 37°C in the dark. The colony-forming units were counted and survival was calculated relative to the minus UV control.

For the overreplication complementation assays, cultures were grown to A_{650} = 1.2, diluted 400-fold with chloramphenicol-containing medium plus 0.25 mM IPTG, and further grown at 37°C. Samples at exponential growth were taken at A_{650} = 0.4 and A_{650} = 1.0 (~3 and 4 h after inoculation). Samples at the stationary phase were taken several hours after the A_{650} was stable. DNA was extracted with the Gene Elute Bacteria kit (Sigma-Aldrich), and the relative copy numbers of the sequences corresponding to *OriC*, *TerA*, *TerB*, and to the *YdcR* gene (located equidistantly between *TerA* and *TerB*) were determined by qPCR. Results were normalized against the mean of *TerA* and *TerB*. When analyzing the effect of RECG1 expression on the replication of the pBAD/Thio (Invitrogen) and pUC19 (NEB) plasmids, the relative plasmid copy numbers between exponentially growing cultures and saturated precultures were compared, and results were normalized against the *TerA* sequence.

Intracellular Localization and RECG1 Promoter-GUS Fusion

For in vivo intracellular localization, the sequence encoding the N terminus of RECG1 up to the region conserved in all RecG sequences (the first 208 codons) was cloned into the pCK-GFP3 vector, upstream and in frame with the GFP coding sequence, under the control of the CaMV 35S promoter, and young leaves of *Nicotiana benthamiana* were transfected by bombardment as described before (Vermel et al., 2002). The pCKmRFP plasmid (Vermel et al., 2002) was cotransfected as an internal mitochondrial marker. After 24 h, the fluorescence of GFP and chlorophyll was observed at 505 to 540 nm and beyond 650 nm, respectively, after excitation at 488 nm on a Zeiss LSM700 confocal microscope. The DsRED fluorescence was excited at 555 nm and observed at 560 to 615 nm. For constitutive expression in Arabidopsis, the expressing cassette was transferred from pCK-GFP3 into the *Asc*I and *Pac*I restriction sites of pBINPlus (van Engelen et al., 1995), and Arabidopsis Col-0 plants were transformed by the floral dip method. Leaves of selected transformants were observed at the confocal microscope at different stages of development. For promoter-GUS histochemical analysis, the 5′ upstream region of *RECG1* (595 nucleotides, starting just after the upstream gene and comprising 161 nucleotides of *RECG1* 5′-UTR) was cloned upstream of the *GUS* gene in the binary vector pMDC162 (Curtis and Grossniklaus, 2003), and the construct was used to generate stable Arabidopsis transformants. Tissues from six independent lines were stained with 5-bromo-4-chloro-3-indolyl-β-glucuronic acid (X-Gluc; Biosynth) as described (Zaegel et al., 2006) and observed by stereomicroscopy and transmission microscopy.

DNA Gel Blots, Genotoxic Stress Treatments, and Immunoblots

Genomic DNA was extracted using the cetyltrimethylammonium bromide method (Murray and Thompson, 1980). Plant genotypes were determined by PCR using gene- and T-DNA-specific primers. Total RNA was extracted with TRIzol (Life Technologies). For genotoxic stress assays, seeds were surface sterilized and spread on MS255 1% sucrose plates supplemented with different concentrations of ciprofloxacin (Sigma-Aldrich). Pools of 10 plants were harvested at 10 d, and total DNA was extracted for further analysis. For immunoblots, total leaf proteins were extracted in the presence of 2% (w/v) SDS and 4 M urea and precipitated with methanol-chloroform before SDS-PAGE fractionation and transfer to Immobilon-P membranes. The ATP1 protein was detected with an antiserum raised against the α-subunit of the mitochondrial ATP synthase from yeast. Chemiluminescence was quantified with a Fusion-FX7 camera system (Vilber Lourmat). Respiration capacity was measured in an Oxygraph oxygen electrode (Hansatech Instruments) in young 3-week-old leaves. The rates of oxygen consumption were measured after tissue addition and subtracted from the rates after addition of KCN.

qPCR Analysis

qPCR assays were performed in a LightCycler480 (Roche), in 6 μL of reaction containing 1× LightCycler480 SYBR Green I Master Mix (Roche) and 0.5 μM of each primer. The thermocycling program was as follows: a 7-min denaturing step at 95°C, then 40 cycles of 10 s at 95°C, 15 s at 58°C, and 15 s at 72°C. The second derivative maximum method was used to determine Cp values, and PCR efficiencies were determined from DNA serial dilution curves or using LinRegPCR software (http://LinRegPCR.nl). Three technical replicates were performed for each experiment. The numbers of biological replicates are indicated in the figures. To measure the accumulation of ectopic recombination in mtDNA and cpDNA, primers flanking each repeat were used, as described in Figure 4C. The *COX2* (AtMG00160) and 18S rRNA (AtMG01390) mitochondrial genes and the 16S rRNA (AtCG00920) chloroplast gene were used for normalization. For quantification of mtDNA and cpDNA copy numbers, a set of primers located along the organellar genomes was used, and results were normalized against the *UBQ10* (At4G05320) and *ACT1* (At2G37620) nuclear genes. For RT-qPCR experiments, the *GAPDH* (At1G13440) and *ACT2* (At3G18780) transcripts were used as standards.

Modeling of the RECG1 Structure

Plant RECG1 sequences were aligned using ClustalX (Thompson et al., 1997). The Arabidopsis RECG1 sequence was modeled on the structure of *Thermotoga maritima* RecG bound to a three-way DNA junction

(pdb1GM5), using Modeler (Sali and Blundell, 1993; http://salilab.org/modeller/about_modeller.html). The score and E-value between the two sequences is 370 and 7e-114, respectively. Three insertions found in all plant RECG1 sequences have no equivalents in bacterial sequences and could not be modeled accurately. No constraints were imposed and these additional sequences were displayed as loops and hand-edited to adjust to figure size.

Accession Numbers

Sequence data from this article can be found in the Arabidopsis Genome Initiative or GenBank/EMBL databases under the following accession numbers: *RECG1*, At2g01440; *RECA3*, At3g10140; and *ATP1*, AtMg01190. Accession numbers of the sequences used for phylogenetic analysis are given in Supplemental Data Set 1.

Supplemental Data

Supplemental Figure 1. Phylogenetic tree of RecG proteins.

Supplemental Figure 2. Alignment of plant RECG1 protein sequences.

Supplemental Figure 3. Analysis of representative intermediate size repeats in accessions Col-0 and Ws.

Supplemental Figure 4. Scanning of the cpDNA of *recG1* mutants for changes in relative copy numbers of the different cpDNA regions.

Supplemental Figure 5. Activities catalyzed by bacterial RecG in genome maintenance.

Supplemental Figure 6. Recovery of blocked replication forks by RecG and RecA complementary functions.

Supplemental Table 1. Oligonucleotide primers.

Supplemental Data Set 1. Text file of alignment used for phylogenetic analysis.

ACKNOWLEDGMENTS

We thank Géraldine Bonnard for providing ATP1 antibodies and Jean Molinier and Marie-Edith Chaboute for helpful discussions. This work has been published under the framework of the LABEX [ANR-11-LABX-0057_MITOCROSS] and benefits from a funding from the state managed by the French National Research Agency as part of the "Investments for the Future" program.

AUTHOR CONTRIBUTIONS

C.W., M.L.R., and J.M.G. performed research. M. Bergdoll constructed the protein model. C.W., M. Bichara, A.D., and J.M.G. designed the research and analyzed data. C.W., A.D., and J.M.G. wrote the article.

Received July 31, 2015; revised September 14, 2015; accepted September 25, 2015; published October 13, 2015.

REFERENCES

Allen, J.O., et al. (2007). Comparisons among two fertile and three male-sterile mitochondrial genomes of maize. Genetics **177:** 1173–1192.

Alverson, A.J., Rice, D.W., Dickinson, S., Barry, K., and Palmer, J.D. (2011). Origins and recombination of the bacterial-sized multichromosomal mitochondrial genome of cucumber. Plant Cell **23:** 2499–2513.

Arrieta-Montiel, M.P., Shedge, V., Davila, J., Christensen, A.C., and Mackenzie, S.A. (2009). Diversity of the Arabidopsis mitochondrial genome occurs via nuclear-controlled recombination activity. Genetics **183:** 1261–1268.

Backert, S., and Börner, T. (2000). Phage T4-like intermediates of DNA replication and recombination in the mitochondria of the higher plant *Chenopodium album* (L.). Curr. Genet. **37:** 304–314.

Backert, S., Dörfel, P., Lurz, R., and Börner, T. (1996). Rolling-circle replication of mitochondrial DNA in the higher plant *Chenopodium album* (L.). Mol. Cell. Biol. **16:** 6285–6294.

Backert, S., Nielsen, B.L., and Borner, T. (1997). The mystery of the rings: structure and replication of mitochondrial genomes from higher plants. Trends Plant Sci. **2:** 477–483.

Bauknecht, M., and Kobbe, D. (2014). AtGEN1 and AtSEND1, two paralogs in Arabidopsis, possess holliday junction resolvase activity. Plant Physiol. **166:** 202–216.

Bellaoui, M., Martin-Canadell, A., Pelletier, G., and Budar, F. (1998). Low-copy-number molecules are produced by recombination, actively maintained and can be amplified in the mitochondrial genome of Brassicaceae: relationship to reversion of the male sterile phenotype in some cybrids. Mol. Gen. Genet. **257:** 177–185.

Briggs, G.S., Mahdi, A.A., Weller, G.R., Wen, Q., and Lloyd, R.G. (2004). Interplay between DNA replication, recombination and repair based on the structure of RecG helicase. Philos. Trans. R. Soc. Lond. B Biol. Sci. **359:** 49–59.

Busso, D., Delagoutte-Busso, B., and Moras, D. (2005). Construction of a set Gateway-based destination vectors for high-throughput cloning and expression screening in *Escherichia coli*. Anal. Biochem. **343:** 313–321.

Cappadocia, L., Maréchal, A., Parent, J.S., Lepage, E., Sygusch, J., and Brisson, N. (2010). Crystal structures of DNA-Whirly complexes and their role in Arabidopsis organelle genome repair. Plant Cell **22:** 1849–1867.

Carles-Kinch, K., George, J.W., and Kreuzer, K.N. (1997). Bacteriophage T4 UvsW protein is a helicase involved in recombination, repair and the regulation of DNA replication origins. EMBO J. **16:** 4142–4151.

Carrie, C., and Whelan, J. (2013). Widespread dual targeting of proteins in land plants: when, where, how and why. Plant Signal. Behav. **8:** e25034.

Chen, J., Guan, R., Chang, S., Du, T., Zhang, H., and Xing, H. (2011). Substoichiometrically different mitotypes coexist in mitochondrial genomes of *Brassica napus* L. PLoS One **6:** e17662.

Christensen, A.C. (2013). Plant mitochondrial genome evolution can be explained by DNA repair mechanisms. Genome Biol. Evol. **5:** 1079–1086.

Cox, M.M. (2007). Motoring along with the bacterial RecA protein. Nat. Rev. Mol. Cell Biol. **8:** 127–138.

Curtis, M.D., and Grossniklaus, U. (2003). A gateway cloning vector set for high-throughput functional analysis of genes in planta. Plant Physiol. **133:** 462–469.

Dasgupta, S., Masukata, H., and Tomizawa, J. (1987). Multiple mechanisms for initiation of ColE1 DNA replication: DNA synthesis in the presence and absence of ribonuclease H. Cell **51:** 1113–1122.

Davila, J.I., Arrieta-Montiel, M.P., Wamboldt, Y., Cao, J., Hagmann, J., Shedge, V., Xu, Y.Z., Weigel, D., and Mackenzie, S.A. (2011). Double-strand break repair processes drive evolution of the mitochondrial genome in Arabidopsis. BMC Biol. **9:** 64.

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J.F., Guindon, S., Lefort, V., Lescot,

M., Claverie, J.M., and Gascuel, O. (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. **36:** W465–W469.

Diray-Arce, J., Liu, B., Cupp, J.D., Hunt, T., and Nielsen, B.L. (2013). The Arabidopsis At1g30680 gene encodes a homologue to the phage T7 gp4 protein that has both DNA primase and DNA helicase activities. BMC Plant Biol. **13:** 36.

Dudas, K.C., and Kreuzer, K.N. (2001). UvsW protein regulates bacteriophage T4 origin-dependent replication by unwinding R-loops. Mol. Cell. Biol. **21:** 2706–2715.

Feng, X., Kaur, A.P., Mackenzie, S.A., and Dweikat, I.M. (2009). Substoichiometric shifting in the fertility reversion of cytoplasmic male sterile pearl millet. Theor. Appl. Genet. **118:** 1361–1370.

Fukuoh, A., Iwasaki, H., Ishioka, K., and Shinagawa, H. (1997). ATP-dependent resolution of R-loops at the ColE1 replication origin by *Escherichia coli* RecG protein, a Holliday junction-specific helicase. EMBO J. **16:** 203–209.

Gerhold, J.M., Aun, A., Sedman, T., Jõers, P., and Sedman, J. (2010). Strand invasion structures in the inverted repeat of *Candida albicans* mitochondrial DNA reveal a role for homologous recombination in replication. Mol. Cell **39:** 851–861.

Gu, J., Miles, D., and Newton, K.J. (1993). Analysis of leaf sectors in the NCS6 mitochondrial mutant of maize. Plant Cell **5:** 963–971.

Gualberto, J.M., Mileshina, D., Wallet, C., Niazi, A.K., Weber-Lotfi, F., and Dietrich, A. (2014). The plant mitochondrial genome: dynamics and maintenance. Biochimie **100:** 107–120.

Janicka, S., Kühn, K., Le Ret, M., Bonnard, G., Imbault, P., Augustyniak, H., and Gualberto, J.M. (2012). A RAD52-like single-stranded DNA binding protein affects mitochondrial DNA repair by recombination. Plant J. **72:** 423–435.

Janska, H., Sarria, R., Woloszynska, M., Arrieta-Montiel, M., and Mackenzie, S.A. (1998). Stoichiometric shifts in the common bean mitochondrial genome leading to male sterility and spontaneous reversion to fertility. Plant Cell **10:** 1163–1180.

Kanazawa, A., Tsutsumi, N., and Hirai, A. (1994). Reversible changes in the composition of the population of mtDNAs during dedifferentiation and regeneration in tobacco. Genetics **138:** 865–870.

Kmiec, B., Woloszynska, M., and Janska, H. (2006). Heteroplasmy as a common state of mitochondrial genetic information in plants and animals. Curr. Genet. **50:** 149–159.

Knoll, A., and Puchta, H. (2011). The role of DNA helicases and their interaction partners in genome stability and meiotic recombination in plants. J. Exp. Bot. **62:** 1565–1579.

Knoop, V. (2013). Plant mitochondrial genome peculiarities evolving in the earliest vascular plant lineages. J. Syst. Evol. **51:** 1–12.

Lepage, É., Zampini, É., and Brisson, N. (2013). Plastid genome instability leads to reactive oxygen species production and plastid-to-nucleus retrograde signaling in Arabidopsis. Plant Physiol. **163:** 867–881.

Lloyd, R.G. (1991). Conjugational recombination in resolvase-deficient ruvC mutants of *Escherichia coli* K-12 depends on recG. J. Bacteriol. **173:** 5414–5418.

Lloyd, R.G., and Buckman, C. (1991). Genetic analysis of the recG locus of Escherichia coli K-12 and of its role in recombination and DNA repair. J. Bacteriol. **173:** 1004–1011.

Mahdi, A.A., Briggs, G.S., Sharples, G.J., Wen, Q., and Lloyd, R.G. (2003). A model for dsDNA translocation revealed by a structural motif common to RecG and Mfd proteins. EMBO J. **22:** 724–734.

Manosas, M., Perumal, S.K., Bianco, P.R., Ritort, F., Benkovic, S.J., and Croquette, V. (2013). RecG and UvsW catalyse robust DNA rewinding critical for stalled DNA replication fork rescue. Nat. Commun. **4:** 2368. Erratum. Nat. Commun. **5:** 4210.

Maréchal, A., and Brisson, N. (2010). Recombination and the maintenance of plant organelle genome stability. New Phytol. **186:** 299–317.

Matera, J.T., Monroe, J., Smelser, W., Gabay-Laughnan, S., and Newton, K.J. (2011). Unique changes in mitochondrial genomes associated with reversions of S-type cytoplasmic male sterility in maizemar. PLoS One **6:** e23405.

Mawer, J.S., and Leach, D.R. (2014). Branch migration prevents DNA loss during double-strand break repair. PLoS Genet. **10:** e1004485.

McGlynn, P., and Lloyd, R.G. (2000). Modulation of RNA polymerase by (p)ppGpp reveals a RecG-dependent mechanism for replication fork progression. Cell **101:** 35–45.

McGlynn, P., and Lloyd, R.G. (2002). Genome stability and the processing of damaged replication forks by RecG. Trends Genet. **18:** 413–419.

McGlynn, P., Lloyd, R.G., and Marians, K.J. (2001). Formation of Holliday junctions by regression of nascent DNA in intermediates containing stalled replication forks: RecG stimulates regression even when the DNA is negatively supercoiled. Proc. Natl. Acad. Sci. USA **98:** 8235–8240.

Meddows, T.R., Savory, A.P., and Lloyd, R.G. (2004). RecG helicase promotes DNA double-strand break repair. Mol. Microbiol. **52:** 119–132.

Miller-Messmer, M., Kühn, K., Bichara, M., Le Ret, M., Imbault, P., and Gualberto, J.M. (2012). RecA-dependent DNA repair results in increased heteroplasmy of the Arabidopsis mitochondrial genome. Plant Physiol. **159:** 211–226.

Mower, J.P., Touzet, P., Gummow, J.S., Delph, L.F., and Palmer, J.D. (2007). Extensive variation in synonymous substitution rates in mitochondrial genes of seed plants. BMC Evol. Biol. **7:** 135.

Murray, M.G., and Thompson, W.F. (1980). Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res. **8:** 4321–4325.

Odahara, M., Kuroiwa, H., Kuroiwa, T., and Sekine, Y. (2009). Suppression of repeat-mediated gross mitochondrial genome rearrangements by RecA in the moss *Physcomitrella patens*. Plant Cell **21:** 1182–1194.

Odahara, M., Masuda, Y., Sato, M., Wakazaki, M., Harada, C., Toyooka, K., and Sekine, Y. (2015). RECG maintains plastid and mitochondrial genome stability by suppressing extensive recombination between short dispersed repeats. PLoS Genet. **11:** e1005080.

Oldenburg, D.J., and Bendich, A.J. (1996). Size and structure of replicating mitochondrial DNA in cultured tobacco cells. Plant Cell **8:** 447–461.

Osakabe, K., Abe, K., Yoshioka, T., Osakabe, Y., Todoriki, S., Ichikawa, H., Hohn, B., and Toki, S. (2006). Isolation and characterization of the RAD54 gene from *Arabidopsis thaliana*. Plant J. **48:** 827–842.

Parent, J.S., Lepage, E., and Brisson, N. (2011). Divergent roles for the two PolI-like organelle DNA polymerases of Arabidopsis. Plant Physiol. **156:** 254–262.

Ramamoorthy, M., Tadokoro, T., Rybanska, I., Ghosh, A.K., Wersto, R., May, A., Kulikowicz, T., Sykora, P., Croteau, D.L., and Bohr, V.A. (2012). RECQL5 cooperates with Topoisomerase II alpha in DNA decatenation and cell cycle progression. Nucleic Acids Res. **40:** 1621–1635.

Robu, M.E., Inman, R.B., and Cox, M.M. (2004). Situational repair of replication forks: roles of RecG and RecA proteins. J. Biol. Chem. **279:** 10973–10981.

Rowan, B.A., Oldenburg, D.J., and Bendich, A.J. (2010). RecA maintains the integrity of chloroplast DNA molecules in Arabidopsis. J. Exp. Bot. **61:** 2575–2588.

Rudolph, C.J., Upton, A.L., Briggs, G.S., and Lloyd, R.G. (2010). Is RecG a general guardian of the bacterial genome? DNA Repair (Amst.) **9:** 210–223.

Rudolph, C.J., Upton, A.L., Stockum, A., Nieduszynski, C.A., and Lloyd, R.G. (2013). Avoiding chromosome pathology when replication forks collide. Nature **500:** 608–611.

Saini, N., Ramakrishnan, S., Elango, R., Ayyar, S., Zhang, Y., Deem, A., Ira, G., Haber, J.E., Lobachev, K.S., and Malkova, A. (2013). Migrating bubble during break-induced replication drives conservative DNA synthesis. Nature **502:** 389–392.

Sakai, T., and Imamura, J. (1992). Alteration of mitochondrial genomes containing atpA genes in the sexual progeny of cybrids between *Raphanus sativus* cms line and *Brassica napus* cv. Westar. Theor. Appl. Genet. **84:** 923–929.

Sakamoto, W., Kondo, H., Murata, M., and Motoyoshi, F. (1996). Altered mitochondrial gene expression in a maternal distorted leaf mutant of Arabidopsis induced by chloroplast mutator. Plant Cell **8:** 1377–1390.

Sali, A., and Blundell, T.L. (1993). Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. **234:** 779–815.

Shedge, V., Arrieta-Montiel, M., Christensen, A.C., and Mackenzie, S.A. (2007). Plant mitochondrial recombination surveillance requires unusual RecA and MutS homologs. Plant Cell **19:** 1251–1264.

Shedge, V., Davila, J., Arrieta-Montiel, M.P., Mohammed, S., and Mackenzie, S.A. (2010). Extensive rearrangement of the Arabidopsis mitochondrial genome elicits cellular conditions for thermotolerance. Plant Physiol. **152:** 1960–1970.

Singleton, M.R., Scaife, S., and Wigley, D.B. (2001). Structural analysis of DNA replication fork reversal by RecG. Cell **107:** 79–89.

Sloan, D.B., Alverson, A.J., Chuckalovcak, J.P., Wu, M., McCauley, D.E., Palmer, J.D., and Taylor, D.R. (2012). Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. PLoS Biol. **10:** e1001241.

Small, I., Suffolk, R., and Leaver, C.J. (1989). Evolution of plant mitochondrial genomes via substoichiometric intermediates. Cell **58:** 69–76.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and Higgins, D.G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **25:** 4876–4882.

van Engelen, F.A., Molthoff, J.W., Conner, A.J., Nap, J.P., Pereira, A., and Stiekema, W.J. (1995). pBINPLUS: an improved plant transformation vector based on pBIN19. Transgenic Res. **4:** 288–290.

Vermel, M., Guermann, B., Delage, L., Grienenberger, J.M., Maréchal-Drouard, L., and Gualberto, J.M. (2002). A family of RRM-type RNA-binding proteins specific to plant mitochondria. Proc. Natl. Acad. Sci. USA **99:** 5866–5871.

Vincent, S.D., Mahdi, A.A., and Lloyd, R.G. (1996). The RecG branch migration protein of *Escherichia coli* dissociates R-loops. J. Mol. Biol. **264:** 713–721.

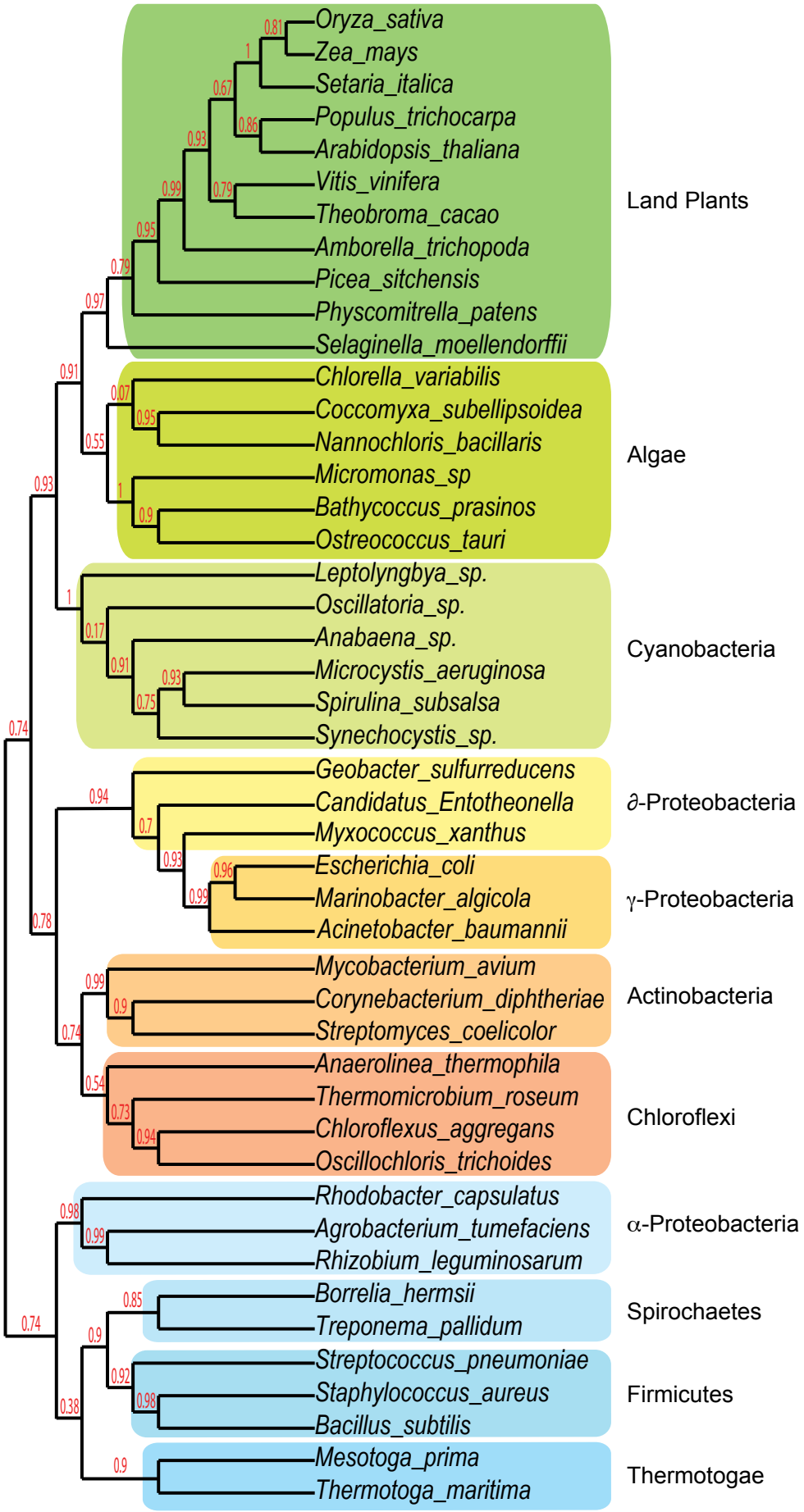
Wall, M.K., Mitchenall, L.A., and Maxwell, A. (2004). Arabidopsis thaliana DNA gyrase is targeted to chloroplasts and mitochondria. Proc. Natl. Acad. Sci. USA **101:** 7821–7826.

Wardrope, L., Okely, E., and Leach, D. (2009). Resolution of joint molecules by RuvABC and RecG following cleavage of the *Escherichia coli* chromosome by EcoKI. PLoS One **4:** e6542.

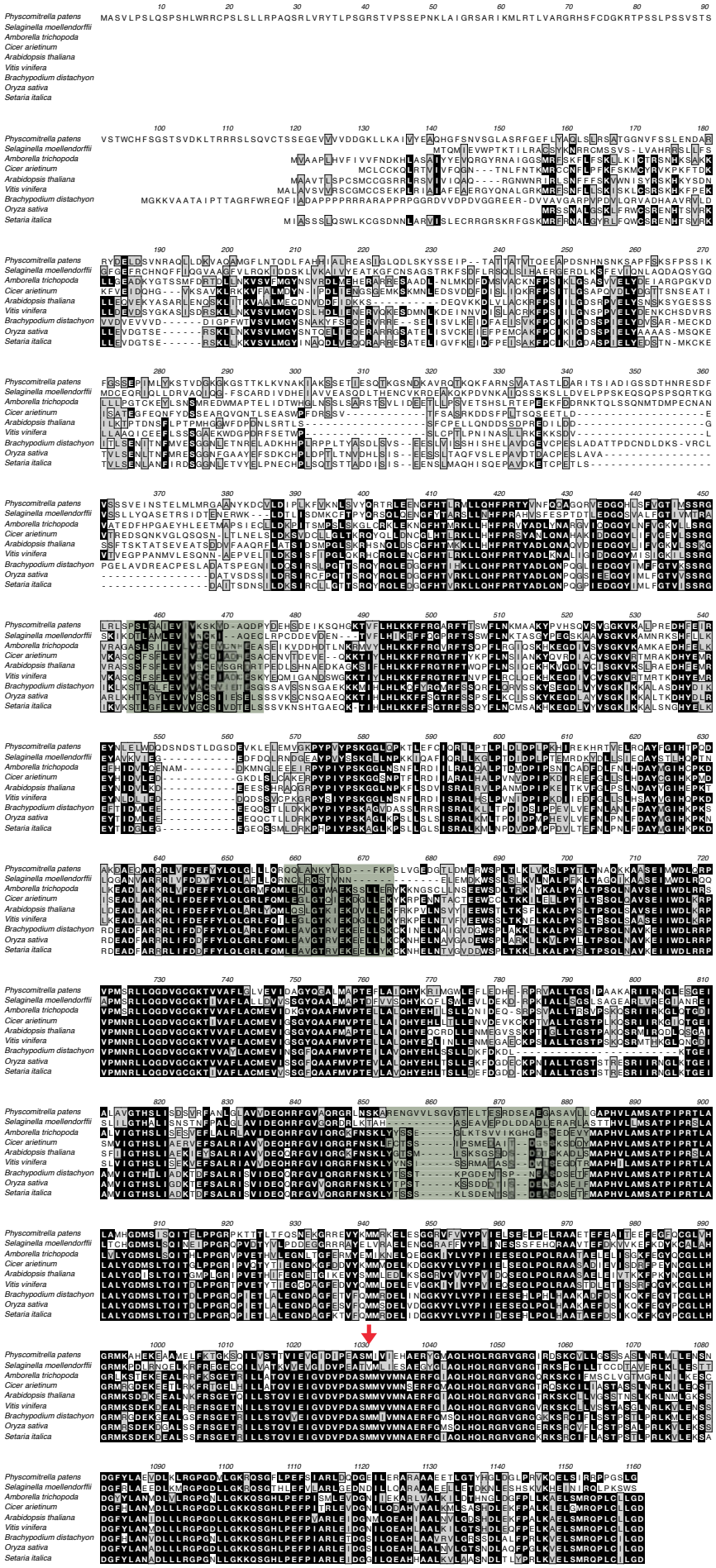
Zaegel, V., Guermann, B., Le Ret, M., Andrés, C., Meyer, D., Erhardt, M., Canaday, J., Gualberto, J.M., and Imbault, P. (2006). The plant-specific ssDNA binding protein OSB1 is involved in the stoichiometric transmission of mitochondrial DNA in Arabidopsis. Plant Cell **18:** 3548–3563.

Zampini, É., Lepage, É., Tremblay-Belzile, S., Truche, S., and 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.

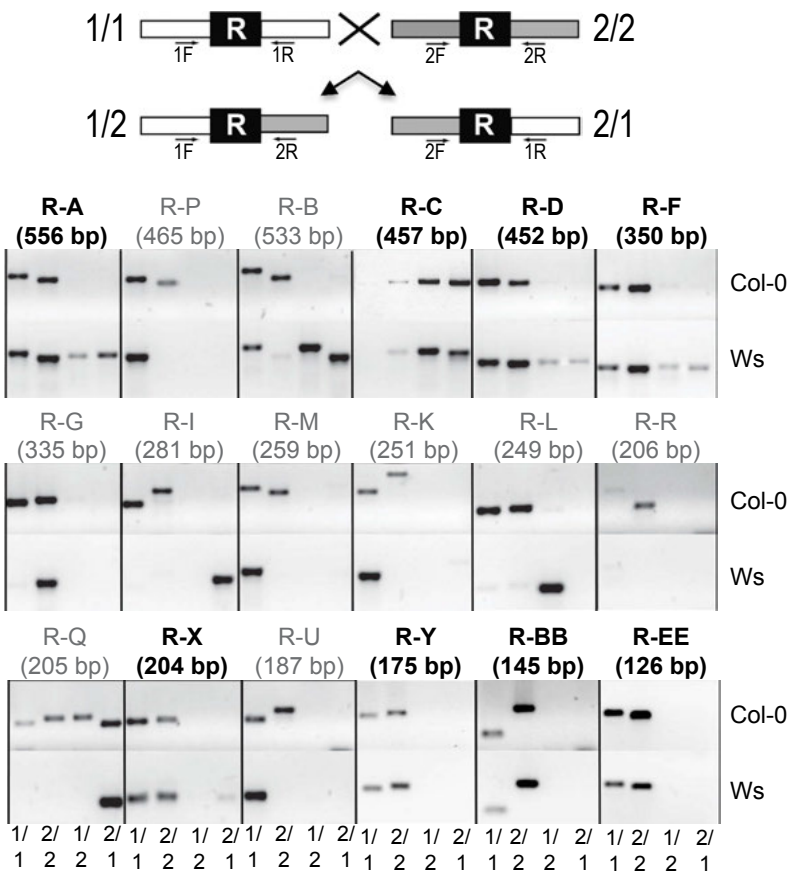
Zhang, J., Mahdi, A.A., Briggs, G.S., and Lloyd, R.G. (2010). Promoting and avoiding recombination: contrasting activities of the *Escherichia coli* RuvABC Holliday junction resolvase and RecG DNA translocase. Genetics **185:** 23–37.



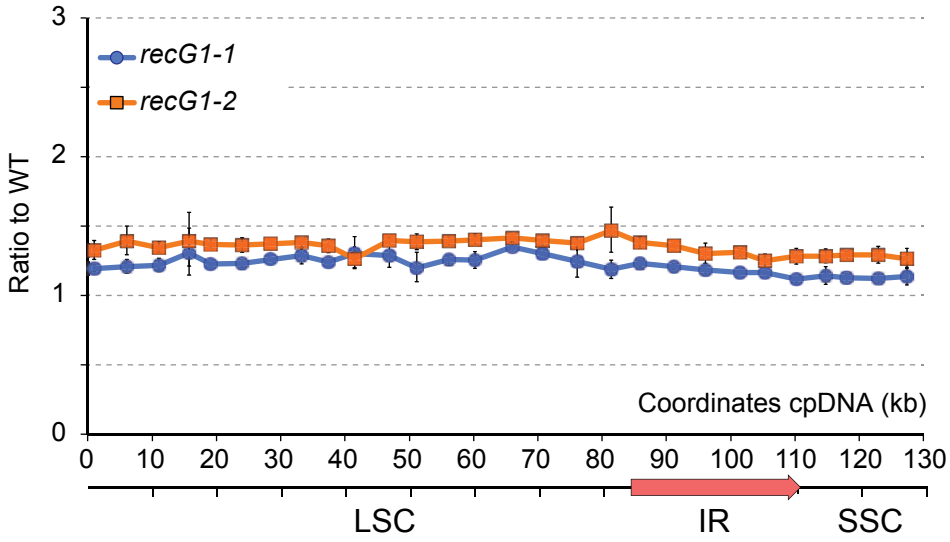
Supplemental Figure 1. Unrooted phylogenetic tree of RecG proteins in bacteria and plants. Bootstrap values are indicated.



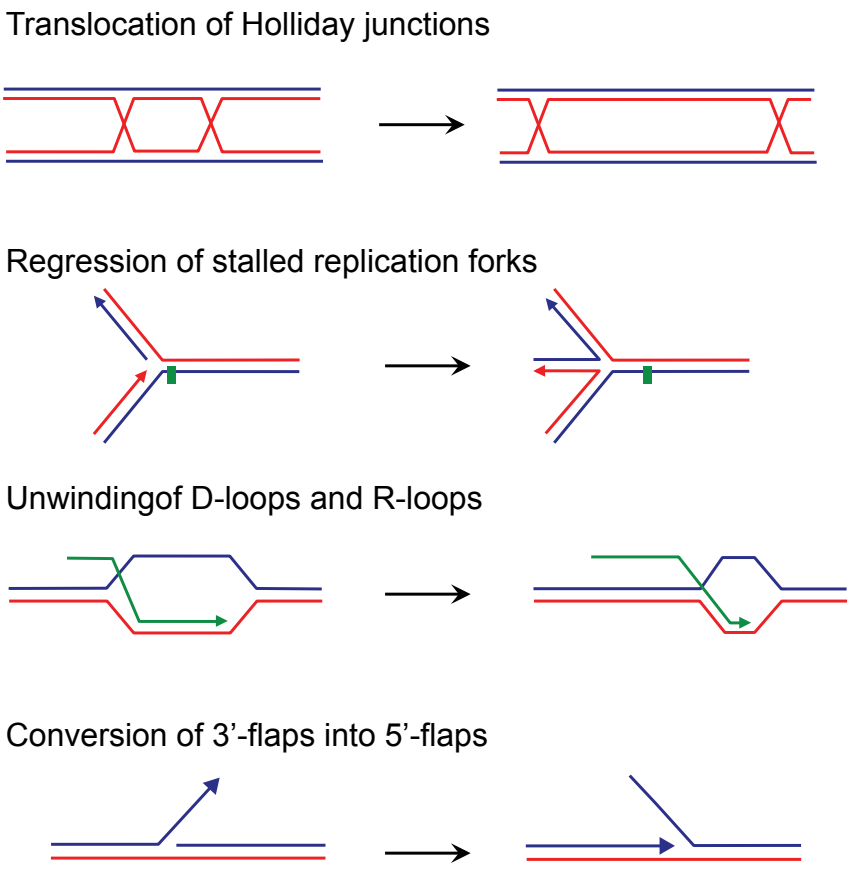
Supplemental Figure 2. Alignment of representative plant RECG1 protein sequences. The three short additional sequences that are absent in bacterial RecG proteins are shaded in green. The red arrowhead indicates the site corresponding to the T-DNA insertion in the gene of mutant *recG1-2.*



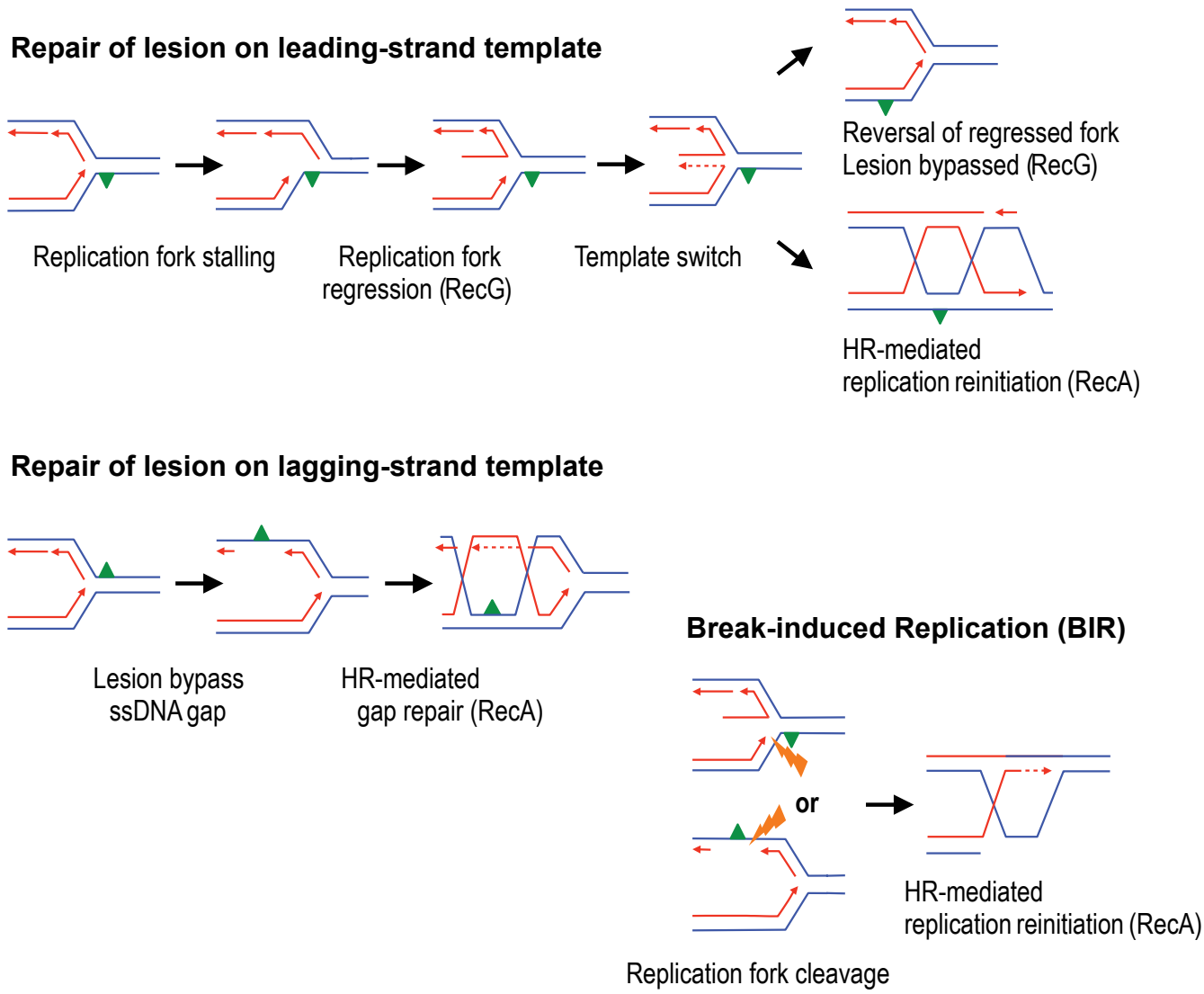
Supplemental Figure 3. PCR analysis for the presence of representative intermediate size repeated sequences (IRs) in the Arabidopsis accessions Col-0 and Ws. The scheme in the upper part explains the PCR strategy. Results showed that only a few IRs previously described as involved in ectopic recombination in Col-0 mitochondria are conserved in the Ws mtDNA. These are labeled in black bold characters. Repeats Y and BB were not retained for further experiments because the sequences are present in more than two copies in the mtDNA of Col-0.



Supplemental Figure 4. Scanning of the cpDNA of *recG1* 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 6A. 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.



Supplemental Figure 5. Different activities catalyzed by bacterial RecG in genome maintenance. Brach-migration of Holliday junctions is consistent with roles in the resolution of such structures during homologous recombination. Regression of replication forks stalled because of a blocking lesion (green rectangle) permits template shifting and bypassing of the lesion. Unwinding of R-loops and D-loops controls unwanted replication initiation by and recombination at sites sharing limited sequence identity. The conversion of 3'-flaps into 5'-flaps permits their processing by ssDNA-exonucleases and avoids re-initiation by PriA that recognizes 3'-flaps.



Supplemental Figure 6. Recovery of blocked replication forks by the complementary and partially redundant functions of RecG and RecA. If the stalled forks are not recovered and the replisome dissociates, the action of nucleases will convert it into a DSB that will be recovered by the BIR pathway for replication re-initiation through homologous recombination.

Supplemental Table 1. Oligonucleotide Primers

	Forward	Reverse
Genotyping *recG1-1*	GGAATCTCGTGTGTAATGGTTTC	TCTACAGACACGTCAATCGAGC
Genotyping *recG1-2*	ACTGTTACACGGGAGGATGAAG	TGTTTCTTCCCAAGCAAATCAC
GFP fusion	TAGGTACCTCTTCCAATGGCGGCAGTGAC	ACGGATCCACGGGACGAGAATCACCAAG
Fusion protein for antibodies	ATGGCTCCTCATGTTCTTGCCAT	TTAGTCGCCCAGAAGACAAAGT
Promoter-GUS fusion	GAAATCCTCGTCTTGGTTGAAG	AGAGAAGTGTAGGGGATAAT
RT-qPCR RECG1	GGCATCAAAGAAGTCTACTCAATG	CAACGTATACTCTGCCTCCAGAC
qPCR OriC	GTATAAGCTGGGATCAGAATG	CGATCTACTGTGGATAACTCTG
qPCR TerA	GTCAACCATTAAACCGATTCGC	GTTGCCGGAGCACAATGGAAC
qPCR TerB	ATTTACCTCTGCCTGACACTAC	GTAACGCGCCATAATATTAACC
qPCR YdcR	CTGATGACTAATAGCCAGAATC	CGTTACGTTGTACTGATTGAG
qPCR pBAD/pUC18	AGCTGAATGAAGCCATACCAAACG	GAAGTGGTCCTGCAACTTTATCCG
qPCR pACYC	CAGACGGCATGATGAACCTGAATC	CCAATATGGACAACTTCTTCGCCC
qPCR act	CATCTTGGCCTCCCTCAGTA	GAGTAAACAAGTGATGGGACTGTG
qPCR ubq	TCACCGGAAAGACCATCACT	CGGTGGGATACCCTCTTTG
RT-qPCR EXP	GAGCTGAAGTGGCTTCAATGAC	GGTCCGACATACCCATGATCC
RT-qPCR GAPDH	TTGGTGACAACAGGTCAAGCA	AAACTTGTCGCTCAATGCAATC
qPCR COX2	AATAAACGTGATTGACCCAATTCT	TCCGATGAGCAGTCACTCAC
qPCR mt18S	CCTTGAGCTAGGAGCCTCTTT	CATGCAAGTCGAACGTTGTT
qPCR ATP6	TGGCATTACTATAGTGGGATTTCA	TCCTGCGGGTAATAAAAAGC
RT-qPCR COB	GAACACATTATGAGAGATGTTGAAGG	AACATACTTGCCCCATTAGCA
RT-qPCR COX3	CCTTGGGAAATCCCTTTTCT	TTTCTTTCCCCGCGAGTAT
RT-qPCR RPS3	CTCGACCAGCGAGAAAAAGT	CGAATGAAGTGGGTCAACCT
RT-qPCR COX2	TCCGATGAGCAGTCACTCAC	AATAAACGTGATTGACCCAATTCT
RT-qPCR ATP1 3'	GCCTTACCCGTCATTGAAAC	CACATTGGTGGGAATATAGGC
RT-qPCR ATP1 5'	GTTTATGGATTGAACGAGATTC	GACATTCTCATTCTCAAGATTC
Repeat EE-1	CAGCATGGTGAAAAGTAAGC	TTTAATCAAAGCTAGCTACAGGG
Repeat EE-2	AGGCTTCAGGGTAGCTTCTG	CACAAGGAGCAGTTGAGACAG
Repeat X-1	TTATGCTGCAAAAGACTGGGATG	TTTGACAGCTCAGCTTTGTCTC
Repeat X-2	TCTTACGTGGAGCGGAATTAG	CTTTCTTTCCTTGGGCAGTACCTC
Repeat F-1	ATCCATCTTGCTGCGACAAAGC	CGAAAAATATGAAGGTAAACCAC
Repeat F-2	CAAATTTCATTGAGAGGAATCAGC	CATCAAACCAGAGACCAAAGCAG
Repeat D-1	ATAAGAGAGCAAATGATGGTATAGC	GCTGCTTTTAGGAGAGTGATCTG
Repeat D-2	CAAAGCAAAGGATGGTTCAAC	CGATTGTATCCGACTATTGAGTG
Repeat L-1	GTTATGCCATTTTGGGCTTTGCTC	CGCTCGACCGAAAGAAATGAGTAAC
Repeat L-2	CCGGTTGAAAGCTAAGCCGGAG	CCCTCACTGAACCGACTTGAATCTG
Repeat T-1	GAATTAGTAAGGCAACAGGAG	TTAAGAGGCAAGGAGGAATACC
Repeat T-2	CCGCATCAGTCGAAGTGGTG	GGGTATTTTCTCCTTCCGGTTTG

PrimersmtDNA

Coordinates on Col-0 mtDNA		Forward	Reverse
4234	4329	TCCTCCGGTCGAGTCATCTTTC	CGGTTTCGGACAGCTCATGATG
7816	7912	CACCTTTCTTTGTTTCAGGCTC	AGGAATCGCAAGAATTGAGAAG
12410	12515	GACCTTGAATGCGCTCTTGGAG	AGAGGACGAAGGCTACTTTCG
17273	17382	AACTTTATGCAATCCGCCATAC	GGGCTTGGAGTTAATCATCTTG
22005	22096	AACTAGACCCGGAAAGGAGAAG	CTTCTGTATCGTCGCTTGAATG
26152	26244	TGCTCGCAGTCGTCTCATAGAG	GACCTTAGCAACCCGAAACGTG
29683	29774	TACTATCCAGAGGCGAAGAAGC	AGGACCCTAAAGAAGGATCGTC
32274	32383	AGTCAGTGGGAAATCGGTAGC	AATGGGCCAGCTGCATAATCAC
34904	35000	CATTGAGGAGAGAGGAGAGAG	CGCCCATACTCTTAACTTTTCG
39428	39536	TAAGCCTGCCTGACTGTGAGAC	GCGTACCTTTGATCCGTTGAGC
43283	43378	CTTGAATGGGAGTTTCTTGGTC	TAGCTTTCGCTTCTTCAGATCC
48567	48661	ATTACGTCTGCCTCCCAGGATG	AGGTGATGCGGGAACTAAGTCC
54799	54917	CCTTGGCGGATAACTAGAAGTG	CCCGAATGTCTTTAATCGAGAG
59534	59642	GGCTCTTCCCTTGGGTCATAGG	ACCCAGCTCAGGAACCTAAACC
63647	63760	ATTCTTGGTGGGAATCATCAAC	AACAAACCAATTACCAAATCCAC
68782	68892	TCACTTCGACGACGGAAGACTC	ATGACCCGGATCGAGTACACTG
73964	74053	AAGATGACAAGCGGAGTAAAGC	GGACGATAGTAGTTTGCCCTTG
79716	79828	AGACCTCCTGAGCTGTTGCAAG	TTCCCAGACTATGATGCCTGG
84857	84974	TATTCGTACTGGTTTCTCATGTG	TTGCCGTAATGCTGAAACTATG
90211	90297	GAGCTTGACCTGCACACACTTG	TCACAGCAGCCAGACAGAAGAC
94946	94954	TTTAAGGAGTTAGCTGCATTTCC	TCCAATCTTGTTGCTAATGTGC
99214	99312	CGACCCTGCCCTTCTTACTCAG	AGGTAGCTAAGCCTGCCGAATC
104322	104436	GAGATGGGGTGACTAAACCTTG	AGTTTTGACTTGCTTTGGCTTC
107511	107614	CCAAGAATAAGGATTCTAAAGG	GCTTTGAGTGCTCCTTAATAG
112496	112595	AGGCTGACAGAAGTACCGAAAC	TTCGATCACAGAATCCATTGAC
113547	113645	GTCTTTCTCGTCAGGTCTATG	TAAGCTAGAGGAGACTTGTTTG
114471	114564	AGCATAGCATAGTAGTGCAAG	CGCTACTTTGTTATCTAAGGTC
119849	119943	GGAGTTTGCTTTACCCAACAAC	GGGCACTCTCATTCTTCATTC
123165	123271	TGCCTCTCTAGCTGGCTTGATC	CGGAGACATCAGAGGAAGGAGC
128104	128210	CATCGTCTGATGTAGCCTTTTG	TGGAATTGATTGTCGTGAACTC
132432	132549	CGCAGGCATGTGATTGTAAGG	ACTTACTTGACCGCTAGCCAGG
137615	137698	ACAGGATTCGAACCTATGGC	GCAATTACAGTGAGACGCGAGG
142333	142437	AAGGTCTGGCGAAGTCTTGAGG	CATCTTAGCAGCAGGGCTTTCG
149136	149238	CCCTTATGAATGGATCTTGCTC	AGTTCACCCCGCTGTTAGATAG
154724	154823	GAAGAGCTGGTAGGACGGACTG	AGGCCTTCTGTTCCAGTGACAG
159570	159689	AACCGCTAAGAATGTGGAGAAG	CTTGCATAGAAGAACGCAACTG
164577	164690	TCTCGCTACCATACCACAAGG	TTCGATCTTGTACCCACGGAGC
169363	169479	ATTGGTGACCTGTTCGATTACC	TAAACCCTGGTGAATCTTTTGG
175463	175551	AGGATTTCTGGTCTCCTGCAGG	GAGTGGTTCTAGTCTCCCTGG
181437	181541	TACTGTAGCTCCCCAAAAGCTC	TTTGTTTGGTGTCTTGGAGTTG
186252	186347	TCCGTTTGGAGGGACAGATGAG	AGGACTTCGCCTTGACCTTCTG
190999	191099	AGGACTCTCTTCTGTCCGTTTG	CTTTCATAGATGTCGTGCGAAG
197765	197855	ATGGCTACAAGTGGGCTACGAG	TCGCAATGAAGTCTTCCCGTTG
201598	201690	AAGCCTGACACGAAGAATAAGC	CTTCTCTCGACCTCTCTCCTTG
208193	208300	TCTCCTCGTGCTCTGCCATAAG	AGCCAATAACTCGAACGCGTTC
213948	214056	CGGTTCTTCTTAGTTGGATTGG	CGTCCGTCGTCGTAAAGTAGTC
218611	218726	AGGCAGGGTAAGAAAGGGTGAC	AAGTTCGAATCCGTTCCGTTCG
223926	224015	AAATTCCACATCCTGACATTCC	TATGTTCTTACCCACCCGAAC
231026	231115	TCTGGCCTGAACCTTTGGGTAG	TTCCCGAATGACCGTACCAAGG
237257	237348	GTTTCCGTCCCAGTACCTGTAG	TGTCACTAAGTTGCCAATCGAC
242471	242579	AGGCGCTGTCATCTATCTCGAC	ACCCTTCCTCCTCTGCTTTGAG
245308	245427	GACATGTTCTGCCGCATACCAC	CATAGCCTTCAACCCGCAAAGG
246842	246960	AAGACAGAGAGCTTGGCAGAG	GGGCTGTTTCCCTCTCCTAGAG
251558	251649	ACTTCGATTCCAAAGAAAGCAG	GATTACCATTAGGCGAAAGCAG
255737	255846	GACGGTTGGCTAAAGACATGG	TGGCAATGTCGGAGATCCTAC
262023	262125	ATGGAAGACCCCTCTCCTTATC	GAATCTCGGAAAACGACAAATC
265093	265190	TGATTGATCCAACACGCCAAGC	TCCTGACTAAAGCACCTCTCC
267468	267574	AGCAGCACCAACCTAGCTCTAC	GCTTATTTCGATACCGCTTCTG
270478	270578	CGAGAGCACCAGATACACCATG	ACTAGCTCGCCTTCTTTCCTCG
274140	274247	ATCCAATCTTCCTTGAGAGCTG	TGATTTCTTGATGGTTGACTGC
277941	278045	CGACAAGTCGTGCCTGAAGTTG	GAATTCACTGCAAGGTGGCGAG
282177	282273	TTTGAAACATTTGTGAGATGCC	TTCCTACCAATTCCATTTCGAG
286792	286911	TGGAAGGACCTCCAAGAAACC	AAAGGCTTGCAGACCTTATGC
292828	292948	CTTTCCCCACCTCACATAGAAG	GAGTTGACATGAAATGGATCAC
299589	299691	GTTCAGTTGTTGCCTCCGGTTG	AGCAGAGAACCACAGGGAAGAC
306260	306369	TGGCAACTGATTTGGTTTTATG	GGGTAGAACTGAGAAATGGCAC
310532	310649	TTAGCTCGTCCAGGATCGCAAG	TGCCCTTGTGTTCTAGACCTCG
314566	314674	ATCTCCAAGCCCCTTTACTAGG	ACAAGGACCAGCTTACTTTTCG
319855	319968	TCCGGATCCTTCAGACTTTCG	GTAGGCTCATGCATTCGCTGAC
324834	324942	TCCCTTTACTGTTGACATGGTG	CTTGACACAGGTTGTTCTTTCG
328085	328165	GTAGATGGCGCTCAACTTCCTG	GAGGTCATGTGCGATGCAAGTG
332522	332640	GTCTGAGCAGCAAACGTACAAC	GATGTTGAGGAGAAAGCAGGAG
337297	337401	TCGTCGCAAAGGTTCAAGCAAG	ATCGGACGCTAACCTGGAACTG
341942	342038	GTGGACTTCCCAAGACTTTCAG	TCTGGAACATGGAAATTGATTG
346953	347035	TAGTCCAACGCTCTCGCTATGG	TAAGCACACAAGCAGGGATGC
352185	352280	GACATAAGCCGTTCCTTGTTC	AAGCAGCAATACCTCAAGCAC
356004	356123	GATGTTGTTCTCGACTCCGTG	CGGATTGGAGGTCTTGGGAATG
360635	360742	GGGGAATCCTCCTTAATAGACG	CATCCGAAGAAACGGAAAATAC
363962	364069	GCGACAAGAAGATCCTGCATCG	TATCGGGAGGGTCGCAGATTTG
366597	366694	GTTGATCCATTGTGATGTTTGG	AAGTCCCATCTTGCCTATGTTG

Primers cpDNA

Coordinates on Col-0 cpDNA		Forward	Reverse
1002	1120	GAATATGCAACAGCAATCCAAG	GGCGGTCCTTATGAACTAATTG
6058	6168	CGAGCCAAGAGCACCTATATTC	ACGATGTGGTAGAAAGCAACG
11104	11219	CTCGACCATCAATAGGGTTAGC	GATGATCCAAGAAGGAAGTTCAG
15777	15894	CGATGGAACCTTCTCTTGTACC	TTGCTGTAGTCATTTGCTTTCG
19090	19206	ATTGACCCGCGATAATACCTAC	GAACTCCCTTTACTTGTCGGAG
24016	24141	AAGTCGATACCTGAATGCCAAG	TAGGATCGTGGGTTGAAACAG
33265	33362	TTTCCAAGGGTTTCATAATTGG	TAGCACCATGAATAGCGCATAG
37469	37575	TGCACGATTGAAAGAGCTACTG	GGTTATTGGCAGGAATTGATTG
41520	41636	TGAGCACTAGGTCCAATGTGAG	TTTCGGCCAACTCTCTATCATC
46929	47054	AATTGGATTGAGCCTTGGTATG	TTGTAAACTCCGGTTTGTTCG
51198	51327	TTAGGGCTATACGGACTCGAAC	AGTTAATGAAAGAGCCCAATGC
56158	56270	CAATTCGGTGGAGGAACTTTAG	GCAAGATCACGTCCCTCATTAC
60199	60303	ATATGGAAATTCGAGGTCAAGG	GCAAGAAGTAAGCCAATTCAGC
66049	66176	TTGGCTGGATTATTCGTAACTG	CTGAATTAGACCTCCCGCATAG
70683	70808	TGCTCCTCCGACTAGGATAAAG	CCTGGTGGATGGGTAATATCTG
76165	76289	GTGTTGGACAATCCACTTTGAC	TAAGGGACCAGAAATACCTTGC
81419	81535	GTCGTACATTTCGTGTCATTGC	CTTCTCGTGGGAATCGTATTTG
85937	86047	TGTATGTCCCAGAATAGGTCC	ACAGAAATAAAGCATTGGGTCG
91180	91290	TGAACAACCGGGAACAATTTAC	GCAAGGAATATCCGTCTTTCTG
96056	96181	ACGGATGCTCCTATTACACTCG	ATTACGGGTAGTTCCTGCAAAG
101430	101550	AAGAAGCAATGACGGTATCTGG	AAGCCACCTACAGACGCTTTAC
105328	105429	TAGCGAAAGCGAGTCTTCATAG	CTCCACTTAGTTTCACCCAAGC
110181	110290	TGGGACTAAACAGGACCAAGAG	CTCGGATTTCTTCCATTTCATC
114762	114881	TAAACACAATAACTGCGCCAAG	TCATTAACCACTGGGACTGGAG
118021	118122	GCCCTAACCATATTTCGACTTG	AAGGAGTTGCTGAATGATACTCG
122976	123095	AATAATCCCATCGCGTTACATAC	ATGGTGAGGATGTTGTTGACTG
127385	127499	GCCTTCCAATTCTTCCAACTC	AACTGATAGCAACGATCAAACG

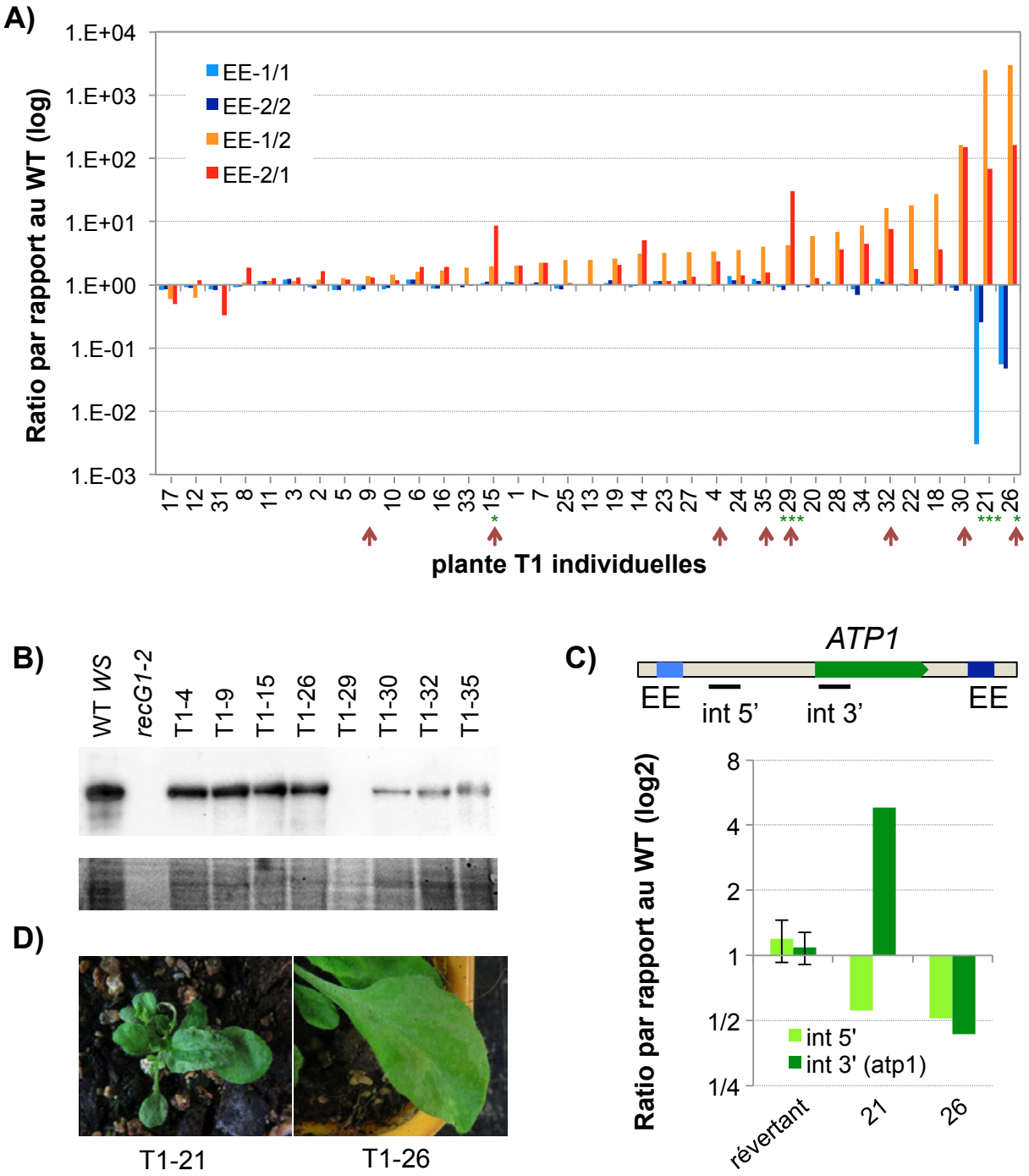


Figure 30. Analyse de plantes *recG1-2* complémentées avec *RECG1*.
A. abondance des formes, parentales (1/1 et 2/2) ou recombinées (1/2 et 2/1) de la séquence répétée EE, mesurées par qPCR dans chacune des plantes T1, par rapport au sauvage (WT). Les étoiles vertes marquent les plantes affectées. Les flèches rouges pointent les plantes dont la complémentation a été vérifiée. **B.** haut: Immunodéctection de RECG1 avec l'anticorps 3403 dans les plantes T1. Bas: Contrôle de dépôt par coloration au bleu de Coomassie. **C.** nombre de copies de la région comprise entre les séquences répétées EE mesurée par qPCR. Résultats normalisés par rapport aux gènes mitochondriaux *COX2* et *ATP6.* Les données individuelles pour les plantes T1-21 et T1-26 sont représentées. Les valeurs représentés sur "révertants" correspondent à la moyenne des valeurs obtenus pour tous les révertants, ± écart type. **D.** feuilles bigarrées et déformées des plantes T-21 et T-26.

2.3. Complémentations du mutant *recG1-2*

Dans notre publication (Wallet *et al.*, 2015) nous avons montré la présence dans *recG1-2* d'une molécule d'ADN circulaire contenant le gène *ATP1* et se répliquant de façon indépendante. Nous avons également montré que la version d'origine de cette séquence est toujours présente chez le mutant, ainsi qu'une version alternative, où le gène *ATP1* a été réintégré dans le mtDNA à un autre endroit, par un deuxième événement de recombinaison. La réintroduction de *RECG1* par rétrocroisement entraîne une ségrégation des différentes versions de cette séquence. Le rétrocroisement est une manière simple de réintroduire le gène *RECG1* sauvage dans un contexte de cytoplasme mutant. En effet la plante mutante utilisée comme maternelle a transmis son cytoplasme à la génération F1. Cependant, on ne peut exclure que des évènements secondaires comme d'autres mutations ou modifications épigénétiques interviennent lors du rétrocroisement, favorisant ainsi l'apparition de plantes affectées dans la descendance et interférant avec les effets du gène *RECG1*. Nous avons donc voulu confirmer le rôle de RECG1 dans la ségrégation par complémentation avec le gène sauvage réintroduit par agrotransformation. Pour cela j'ai cloné *RECG1*, introns et promoteur compris (jusqu'au gène en amont, soit 434 nt avant le premier AUG de *RECG1*, 5987 pb au total), flanqué des extrémités *att*B dans le vecteur pGEM-T easy. J'ai ensuite introduit la construction dans le vecteur d'entrée pDONRZeo par recombinaison Gateway, puis dans le vecteur binaire pGWB40. Ce vecteur a été transféré dans *Agrobacterium tumefaciens* pour pouvoir transformer par "floral-dip" des plantes mutantes *recG1-2*. J'ai également transformé des plantes sauvages (WT) pour contrôle. Après croissance sur milieu MS contenant de l'hygromycine j'ai sélectionné 35 plantes *recG1-2* et 25 plantes WT transformées avec la construction de complémentation (génération T1).

La réintroduction du gène *RECG1* sauvage dans le mutant *recG1-2* homozygote a été vérifiée par western blot avec l'anticorps anti-RECG1 sur les protéines totales de fleurs. Sur un échantillon des plantes complémentées j'ai détecté la protéine RECG1 dans toutes les plantes, sauf dans la plante T1-29 (figure 30B). Je n'ai pas pu faire ce test sur la plante T1-21 car elle était trop affectée et n'a pas survécu jusqu'au prélèvement. Dans les mutants complémentés, sur 35 plantes observés, 3 présentaient un phénotype de feuilles déformées et bigarrées, ressemblant à celui des plantes F1 appelées "variegated" dans la publication, quand aucun phénotype de la sorte n'a pas été observé dans mon groupe contrôle WT complémenté (figure 30D).

Comme dans le cas des plantes F1 décrites dans la publication, l'abondance des formes recombinées et sauvages contenant la séquence répétée EE a été mesurée par qPCR (figure 30A). Pour cette séquence, la majorité des plantes complémentées montrent un phénotype moléculaire de type révertant, c'est à dire avec des quantités relatives des séquences EE équivalentes à celles des plantes sauvages. Chez la plante moyennement affectée T1-26, on voit que la forme recombinée EE-1/2 dépourvue d'*ATP1* est 1000 fois plus abondante que dans le sauvage. Les formes parentales (EE-1/1 et EE-2/2) sont quant à elles diminuées de 20 fois. Ce qui n'est pas autant que chez les plantes "shifted" décrites dans la publication, où ces séquences sont diminuées de plus de 1000 fois par rapport au sauvage, mais dans la même ordre de grandeur que dans certaines plantes F1 classées comme "variegated" (figure 30C). Dans l'ensemble il y a bien une diminution du nombre de copies du gène *ATP1* dans T1-26, et le phénotype serait dû à une mauvaise ségrégation des séquences contenant le gène *ATP1*.

La plante T1-21 était la plus affectée au niveau du phénotype visible, et présentait le même phénotype de déformation et de faible croissance observé dans de nombreuses plantes F1. Cependant le phénotype moléculaire est difficile à expliquer. L'allure de la plante est cohérente avec l'abondance relative des formes parentales ou recombinées de la séquence EE (figure 30A). Il y a en effet une ségrégation asymétrique de ces molécules, car la forme parentale EE-1/1 est réduite 100 fois par rapport à la forme EE-2/2. Cependant, alors que toutes les plantes F1 "variegated" issues de rétrocroisements avaient un nombre réduit de copies du gène *ATP1*, ce n'est pas le cas pour T1-21, car le nombre de copies d'*ATP1* est augmenté de 4 fois par rapport aux révertants. Cela pourrait être la conséquence d'un évènement de recombinaison secondaire, qui produit un gène tronqué ou qui affecte sa transcription. Malheureusement, l'étude de ces phénomènes est limitée du fait de la mort de la plante. L'étude d'un nombre supplémentaire de plantes complémentées de la génération T1 (premiers transformants) permettrait peut-être de trouver d'autres plantes similaires.

Dans cette étude nous n'avons pas trouvé de plantes similaires aux plantes "shifted" décrites dans l'article, c'est-à-dire avec une disparition des formes parentales EE-1/1 et EE-2/2 et une rétention des deux formes EE-1/2 et EE-2/1. Cela peut être dû au fait que nous n'avons pas étudié assez de plantes. La présence de plantes "shifted" peut aussi dépendre de la plante mère utilisée pour les rétrocroisements ou les complémentations. Mais même s'ils étaient présents dans des proportions différentes, nous avons obtenu deux des trois types de phénotypes moléculaire que nous avons observé par rétrocroisement, et notamment les formes avec un phénotype de feuilles déformées ("variegated") qui sont dues à une réduction du

nombre de copies du gène *ATP1*. Ces résultats de complémentation confirment donc ce que nous avions observé par rétrocroisement, et éliminent la possibilité d'effets épigénétiques ou d'autres mutations non-liées à *RECG1* pouvant jouer dans la ségrégation du mtDNA. L'étude de plantes supplémentaire permettrait de confirmer ces données.

2.4. Etude du phénotype des plantes bicolores, issues de rétrocroisement entre *recG1-2* et WT

Après avoir analysé un grand nombre de plantes F1 issues de rétrocroisements, nous avons observé dans 2 plantes, F1-81 et F1-83, un autre type de phénotype, différent de ceux décrits dans l'article et aussi observé dans quelques-unes des plantes complémentées. Ces plantes présentaient quelques feuilles avec des grandes parties décolorées, à moitié albinos (feuilles bigarrées), ainsi que quelques rares hampes et siliques bicolores (figure 31A et figure 33A). Certaines siliques complétement albinos ne se sont pas développées. Ce phénotype développemental est révélateur d'un dysfonctionnement chloroplastique. Il faut rappeler que nous avons montré le double adressage de RECG1 dans les mitochondries et dans les chloroplastes, mais que jusqu'à présent nous n'avons pu détecter des effets sur le génome chloroplastique dans aucun des mutants *recG1*. C'est pourquoi nous nous sommes intéressés à ces plantes. Les deux plantes F1-81 et F1-83 présentant ce phénotype "semi-albinos" sont représentés dans la figure 31A. Nous avons extrait séparément l'ADN de la partie verte et de la partie blanche des feuilles atteintes afin d'analyser les différences au sein du mtDNA. Une feuille de la plante F1-81 a pu être divisée en 3 parties: blanche, pâle et verte (figure 33A). Comme on a pu le voir par qPCR, ces plantes possèdent les versions sauvages de la séquence répétée EE (1/1 et 2/2). Nous avons également mesuré le nombre de copies du gène *ATP1* dans les deux plantes bigarrées. Mais le nombre de copies d'*ATP1* est comparable à celui du WT pour F1-81, et légèrement réduit (moins de la moitié) pour F1-81. En comparaison, les plantes déformées F1-77 et F1-84 issues de la même expérience de rétrocroisement présentent une réduction de 8 fois du nombre de copies d'*ATP1* (figure 31C). Et pour ce qui concerne la recombinaison impliquant la séquence répétée EE, on n'observe pas de grandes différences entre les parties blanches et vertes des feuilles, excepté pour la partie pâle de la plante F1-81, ce qui exclut une corrélation entre le niveau d'*ATP1* et le phénotype. Ces observations suggèrent que ces phénotypes résultent de changements au niveau du cpDNA, ou du moins au niveau du chloroplaste. Pour détecter un éventuel effet sur le cpDNA, j'ai mesuré par qPCR le nombre de copie de chaque région du cpDNA à l'aide de couples d'amorces localisés tous les

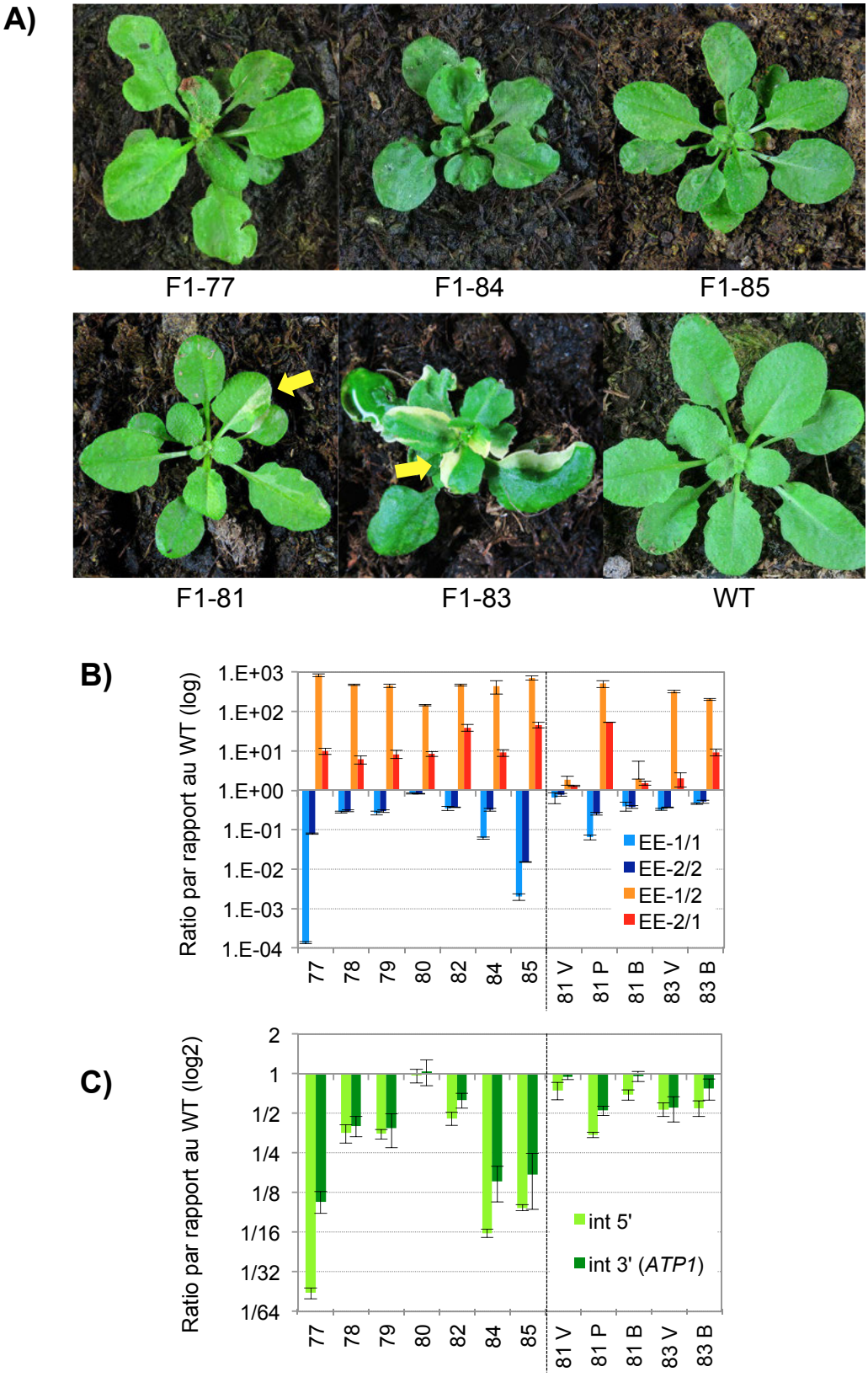


Figure 31. Analyse de plantes *recG1-2* rétro-croisées avec WT.
A. Plantes affectées, bigarrées-déformées, ou simplement bigarrées, présentant des secteurs albinos, indiqués par les flèches. **B.** abondance des formes parentales (1/1 et 2/2) ou recombinées (1/2 et 2/1) de la séquence répétée EE dans chacune des plantes F1, par rapport au sauvage (WT). **C.** nombre de copies de la région comprise entre les séquences répétées EE, correspondant à la séquence en 5' et interne au gène *ATP1*. V: partie verte d'une feuille bicolore; P: partie pale d'une feuille bicolore, B: partie blanche d'une feuille bicolore.

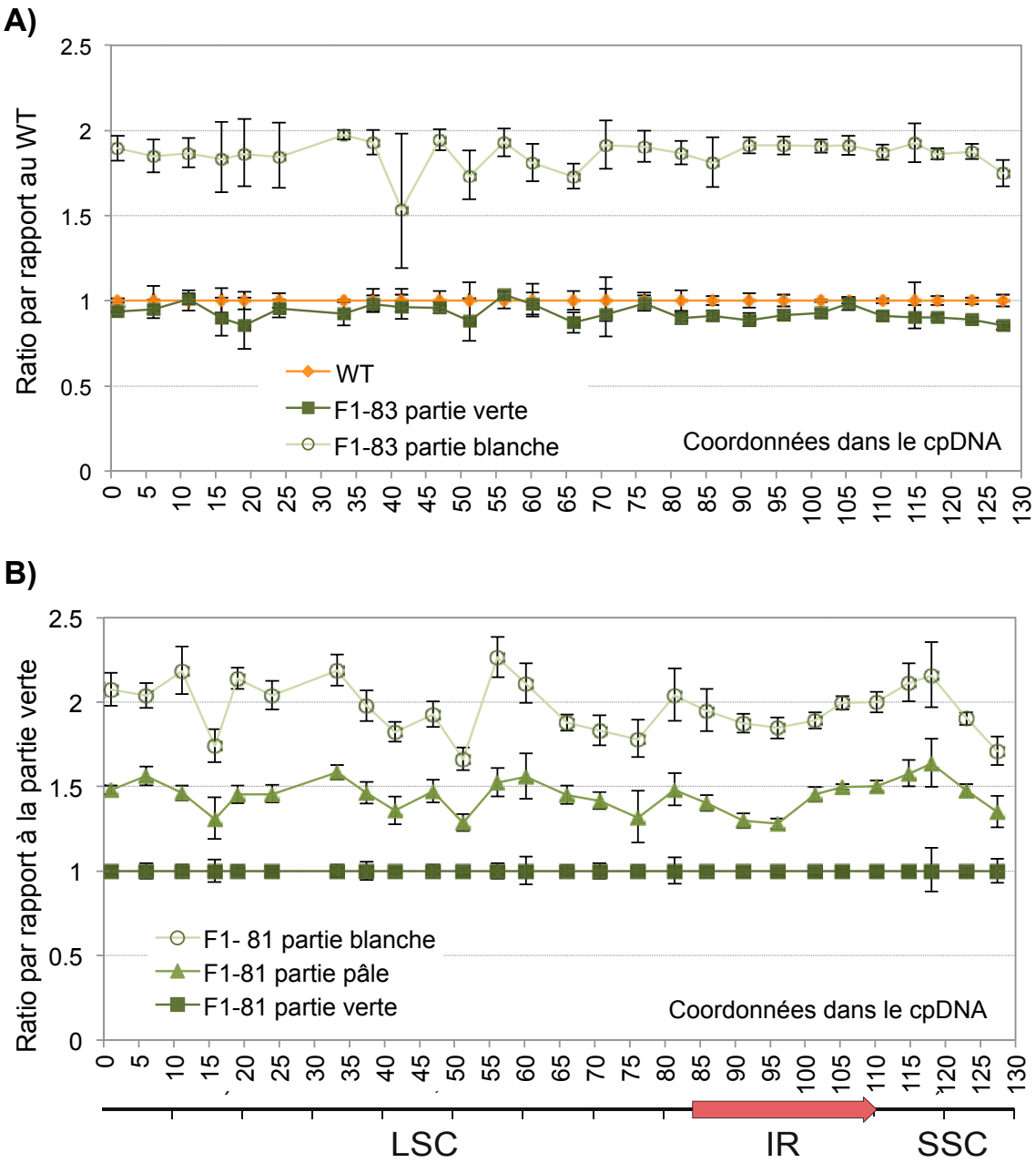


Figure 32. Analyse du cpDNA des plantes bigarrée F1-83 et F1-81.
A. Nombre de copies des différentes régions du cpDNA dans les parties verte et blanche d'une même feuille de la plante F1-83. **B.** Dans les régions vert pâle et blanches d'une feuille de la plante F1-81. Les localisations relatives des différentes régions du cpDNA sont représentées en dessous (LSC: large single copy region, SSC: small single copy region, IR: inverted repeat).

A)

B)

Figure 33. Plante bigarrée F1-81 et F1-83 et leur progéniture.
A. Feuilles, siliques et tiges de la plante bigarrée F1-81. **C.** Plantes provenant des graines de silique albinos des plantes F1-81 et F1-83 à différents stades de développement (cotylédons, deux feuilles, rosette). Les plantes peuvent être maintenues dans milieu MS contenant du sucrose.

10 kb environ. Dans les zones blanches des feuilles bicolores on observe que le cpDNA est deux fois plus abondant que dans les zones vertes ou que dans le WT (figure 32). Cette augmentation du nombre de copies du cpDNA pourrait être un effet compensatoire pour contrebalancer une déficience chloroplastique. Une mutation d'une séquence ou d'un gène impliqué dans la réplication pourrait aussi expliquer cette augmentation. Ces deux plantes ont développé au cours de leur croissance des hampes semi-albinos ou totalement albinos (figure 33A). Nous avons collecté les siliques bicolores qui semblaient présenter quelques graines viables et nous les avons mises à germer sur boite. Elles ont donné systématiquement des plantes albinos (Figure 33B). Si certaines des plantules ont vu leur croissance stoppée au stade cotylédons, certaines se sont développées à un stade avancé *in vitro* sur milieu contenant du sucre. Chez certaines plantes, les feuilles sont vert pâle et l'on observe quelques cellules vertes sur les cotylédons. Les plantes issues des hampes vertes des mêmes plantes F1-81 et F1-83 sont normales, et parmi 96 plantes criblées nous n'avons pas détecté d'autres plantes contenant des secteurs blancs. Cela suggère qu'il s'agit d'un effet clonal, surement dû au chloroplaste, et comme décrit pour d'autres mutants chloroplastiques, le tri ("sorting") des cellules contenants soit des chloroplastes mutants soit normaux se fait très rapidement (selon une discussion avec Stefan Greiner). En tout cas, il est intéressant d'avoir pu sélectionner des plantes totalement albinos et cela nous permettra peut-être de maintenir ce matériel et de l'étudier. Le séquençage par NGS ("new generation sequencing") du cpDNA et du mtDNA est prévu très prochainement, et devrait permettre de détecter la ou les mutations dans le cpDNA ou mtDNA responsables du phénotype. La nature des mutations devrait nous permettre de comprendre les mécanismes moléculaires responsables de leur origine, et des possibles fonctions de RECG1 dans ces processus.

2.5. Etude du double mutant *pol1B-2 recG1-1*

L'étude de doubles mutants permet d'apporter des informations supplémentaires sur les rôles de gènes impliqués dans des voies parallèles ou redondantes. Par exemple, comme décrit dans la Wallet *et al*., 2015, la double mutation r*ecG1 recA3* aggrave le phénotype de développement des plantes. La ségrégation de l'ADN est très affectée, suggérant que RECA3 et RECG1 ont des rôles dans des voies alternatives partiellement redondantes, par exemple dans la régression des fourches de réplication bloquées. Nous avons donc étudié la double mutation de *RECG1* avec un autre facteur de la recombinaison : POL1B.

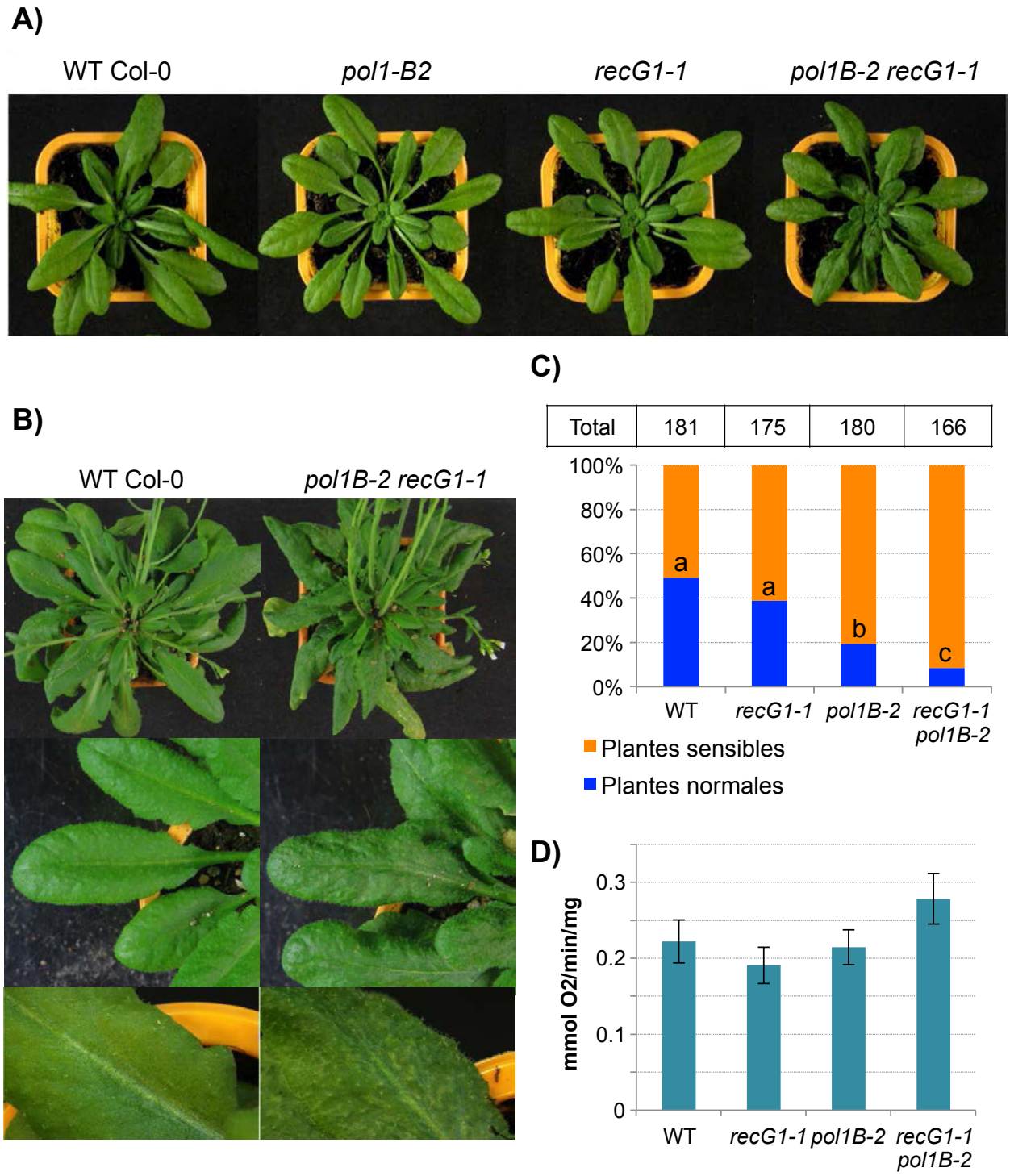


Figure 34. Comparaison des phénotypes des mutants *pol1B-2*, *recG1-1* et du double mutant. **A et B.** Photographies des mutants et zoom sur les feuilles gaufrées et tachetées du double mutant. **C.** Sensibilité des mutants à la ciprofloxacine (CIP): Pourcentage des plantes ayant développé leur premières feuilles à 14 jours de culture *in vitro* sur du milieu contenant 0,25 mM de CIP. Total: nombre total des plantes prises en compte. a, b et c représentent les groupes significativement différents entre eux, selon le test statistique du χ^2 **D.** Consommation d'oxygène par respiration de chaque mutant. Test réalisé sur des groupe de plantules de 10 jours cultivées *in vitro*.

Deux ADN polymérases codées par le noyau et adressées aux organelles ont été identifiées chez l'Arabidopsis (Elo *et al.*, 2003). Ces deux polymérases sont de type Pol I, et ont été appelées POL1A et POL1B (Parent *et al.*, 2011). Chez Arabidopsis, les deux polymérases sont redondantes pour la réplication des organelles, car des mutants individuels sont viables (Parent *et al.*, 2011; et résultats non-publiées du laboratoire). Mais le double mutant *pol1A pol1B* n'est pas viable. Au contraire, dans le maïs, les polymérases semblent être spécifiques à la réplication soit du cpDNA soit du mtDNA (Udy *et al.*, 2012). Cependant, chez Arabidopsis, POL1B semble être préférentiellement recruté pour la réparation par recombinaison (Parent *et al.*, 2011). Ainsi, au laboratoire il a été montré que les mutants *pol1B,* mais pas les mutants *pol1A*, sont déficients dans la réparation par recombinaison homologue de cassures induites dans le mtDNA par traitement à la ciprofloxacine (résultats non-publiées du laboratoire). Pour mieux comprendre le fonctionnement de RECG1 et ses possibles interactions avec POL1B, nous avons croisé les mutants de ces deux gènes. Bien qu'il ne présente pas le phénotype le plus marqué, le mutant *recG1-1* a été utilisé pour ce croisement, car il est dans le même fond génétique que le mutant *pol1B-2* (écotype Col-0).

2.5.1. Phénotype de développement

Les deux mutants individuels ne présentent pas de phénotype de développement. Mais le double mutant présente des feuilles déformées/gaufrées et tachetées (figures 34A et 34B). Ce phénotype est visible sur les 4 générations successives que nous avons mises en culture. Dans un premier temps, nous avons étudié l'effet de stress génotoxique par la ciprofloxacine sur ces plantes déficientes en RECG1 et en POL1B. Comme décrit dans d'autres travaux (Cappadocia *et al.*, 2010; Miller-Messmer *et al.*, 2012; Parent *et al.*, 2011) et dans ma publication, nous avons compté le nombre de plantules présentant des premières feuilles développées à 14 jours, pour le sauvage et chacun des mutant. Les résultats sont présentés pour la culture des plantules sur un milieu contenant 0,25 mM de ciprofloxacine (figure 34C). Le mutant *pol1B-2* est significativement plus affecté que le mutant *recG1-1,* qui est lui comparable au sauvage. Mais la double mutation provoque un effet synergique sur le développement des feuilles car la viabilité des plantes diminue de 40 % chez le double mutant en conditions de stress génotoxique. Nous avons fait des mesures de respiration de ces plantules à 10 jours de développement afin de détecter une éventuelle déficience mitochondriale. De façon surprenante, le double mutant présente un taux de respiration légèrement supérieur au sauvage et aux deux autres mutants individuels (figure 34D). Cette augmentation pourrait être due à un

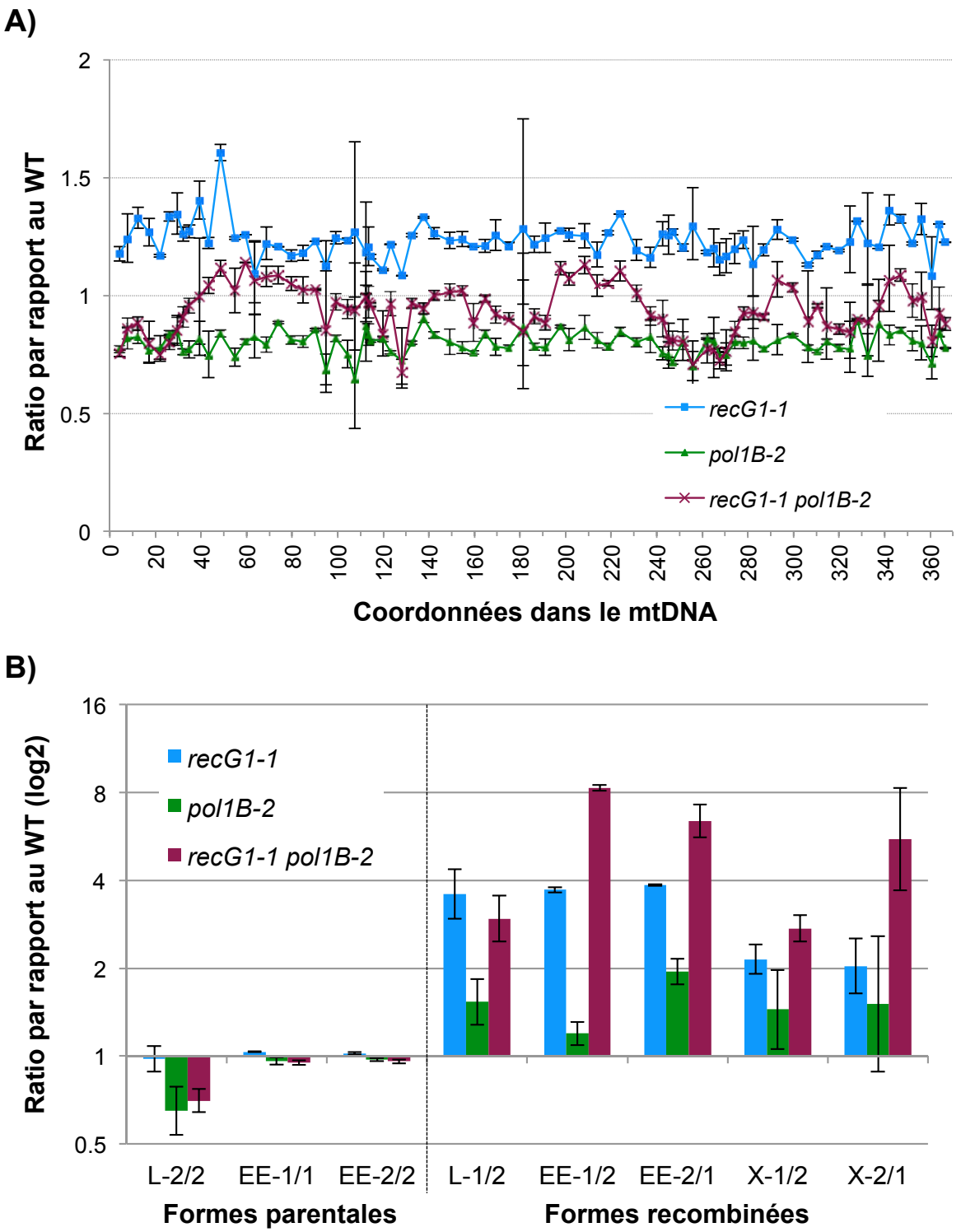


Figure 35. Analyse du mtDNA des doubles mutants *recG1 pol1B*.
A. Stœchiométrie relative des différentes régions du mtDNA. **B.** Abondance relative des forme parentales (1/1 et 2/2) et recombinées (1/2 et 2/1) de différentes séquences répétées. Les formes parentales n'ont pas été toutes testée. Résultats normalisés par rapport aux gènes mitochondriaux *COX2* et *ATP6*.

effet de compensation face à un dysfonctionnement mitochondrial ou chloroplastique, amenant les cellules à consommer plus d'oxygène.

2.5.2. Etude du mtDNA de *pol1B recG1*

Nous avons mesuré le nombre de copies des différentes régions du mtDNA dans les différents mutants. Globalement, l'effet n'est pas important sur la stœchiométrie des différentes régions du mtDNA (figure 35A). En effet, comparé au WT, le mutant *recG1-1* présente une très légère augmentation globale du nombre de copies du mtDNA, alors que chez *pol1B-2* c'est une légère diminution, comme il avait déjà été décrit (Parent *et al.*, 2011). La stœchiométrie du mtDNA du double mutant est un peu moins stable en fonction des régions, mais la différence maximale est d'un facteur de 0,3, encore dans la marge d'erreur technique attendue pour la qPCR (figure 35A). La stœchiométrie semble donc légèrement affectée mais n'expliquerait pas le phénotype de développement observé. Afin de tester si la recombinaison ectopique est augmentée dans le double mutant nous avons également analysé par qPCR la recombinaison impliquant quelques séquences répétées dont X, L et EE, pour lesquelles la recombinaison augmente dans des mutants comme RECA3 ou OSB1. Pour plusieurs séquences répétées, là encore, l'effet des deux mutations est légèrement synergique (figure 35C). Par exemple, les versions recombinées de la séquence EE sont 4 fois plus abondantes dans *recG1-1* que dans le WT. Chez le double mutant, elles sont 8 fois plus abondantes. La mutation de ces deux facteurs entraîne donc une augmentation modérée de la recombinaison ectopique. Cependant, ce phénomène est bien plus important dans des mutants comme *recA3* ou *recG1-2*. Là encore, ce taux de recombinaison n'explique pas le phénotype de développement des plantes.

2.5.3. Etude des fonctions chloroplastiques

RECG1 et POL1B sont toutes deux adressées aux mitochondries, mais aussi aux chloroplastes. Les zones décolorées pourraient donc refléter un dysfonctionnement des chloroplastes. Nous avons donc étudié la stœchiométrie du cpDNA. En effet RECG1 avec POL1B pourraient avoir un rôle important dans la réplication et la réparation du cpDNA. Les résultats ont montré que le nombre de copies du cpDNA est légèrement plus bas que dans le sauvage pour tous les mutants (figure 36A). Cependant ces variations ne sont pas assez importantes pour justifier le phénotype de développement, et sont encore dans la marge d'erreur technique attendue. Par la suite nous avons voulu tester le bon fonctionnement des chloroplastes. Dans une première approche j'ai effectué des immuno-détections avec

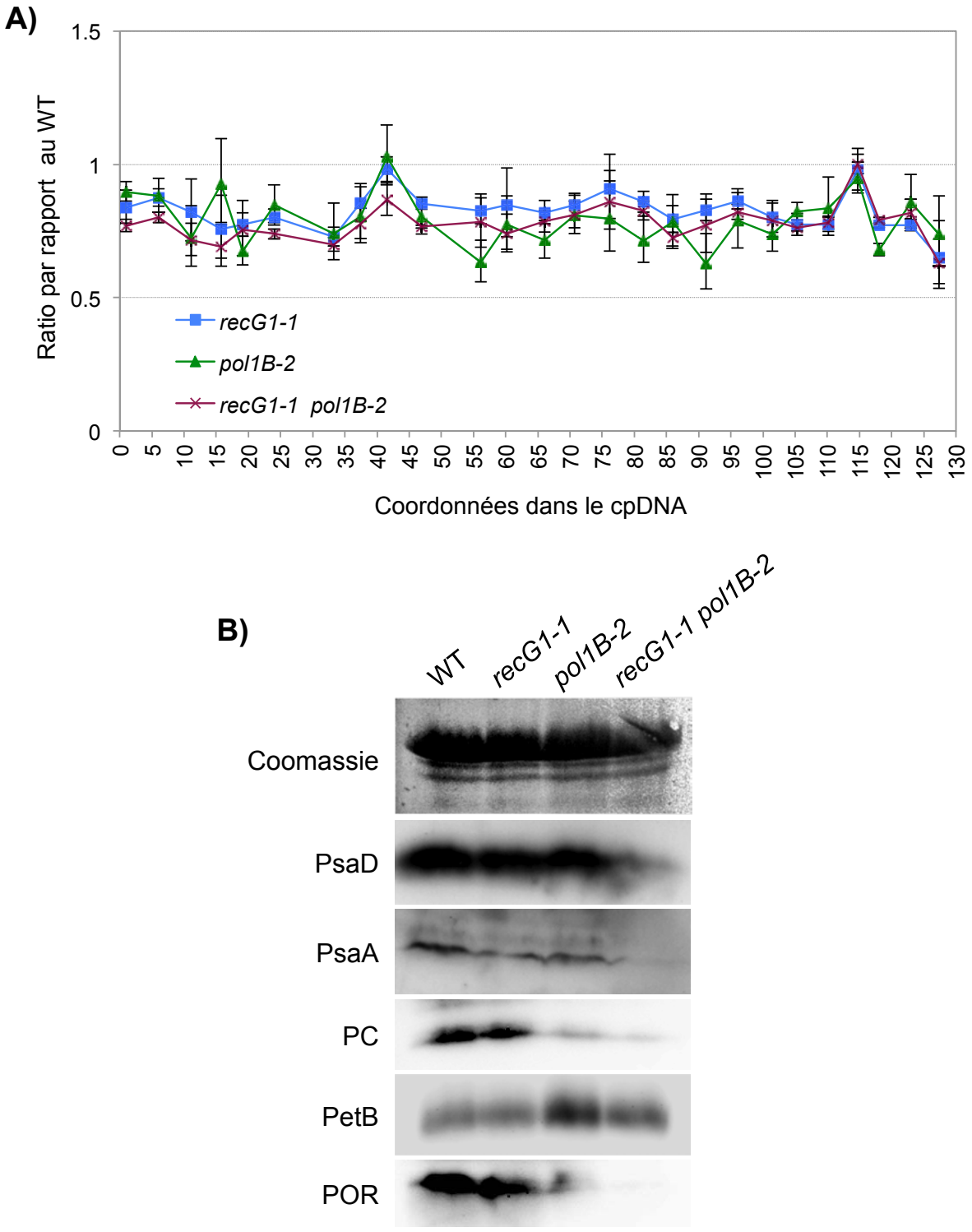


Figure 36. Analyse des effets de la double mutation *recG1 pol1B* sur le chloroplaste.
A. Nombre de copie des différentes région du cpDNA par qPCR. **B.** Immunodétection de différentes protéines chloroplastiques, encodées par le cpDNA (PsaA, PetB) et par le noyau (PsaD, PC, POR), dans des extraits totaux des mutants et du sauvage (WT). PsaD et PsaA: sous-unités du photosystème I. PC: plastocyanine. PetB: sous-unité du cytochrome *b6f*. POR: protochlorophyllide oxydoréductase.

plusieurs anticorps spécifiques de protéines essentielles des complexes photosynthétiques, sur des extraits totaux de mes mutants (figure 36B). PsaA et PsaD sont deux sous unités du photosystème I. PetB est une sous unité du complexe cytochrome *b6f*. La plastocyanine (PC) est responsable du transfert des électrons entre le cytochrome *b6f* et le photosystème I. La protochlorophyllide oxydoréductase (POR) catalyse la photo-réduction de protochlorophyllide en chlorophyllide. Toutes ces protéines sont codées par le noyau excepté PetB et PsaA qui sont codés par le cpDNA, et sont des éléments essentiels au fonctionnement de la chaine photosyntétique. Dans les extraits protéiques totaux, on peut voir que les protéines PsaD, PsaA, POR et PC sont moins abondantes dans le double mutant. L'abondance de POR semble aussi diminuée dans le mutant *pol1B*. Pour PetB en revanche, il n'y a pas de différences (figure 36B). Ces résultats restent à être répétés et confirmés par d'autres expériences, notamment l'analyse des complexes photosynthétiques par gel "Bleu-Native", mais il semble qu'il y ait en effet une diminution de protéines chloroplastiques dans le double mutant, par exemple à cause d'un mauvais assemblage des complexes photosynthétiques. Là encore, le séquençage NGS du cpDNA et mtDNA de ces mutants devrait nous éclaircir sur les éventuelles mutations affectant les génomes des organelles, et donner des indications sur les fonctions de POL1B et RECG1.

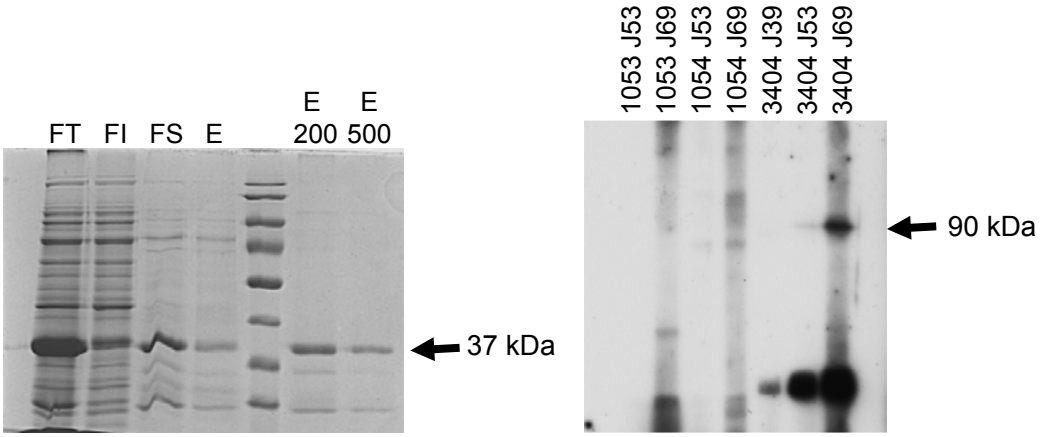


Figure 37. Expression et purification de la protéine recombinante TRCF.
A gauche, gel coloré au bleu de Coomassie avec les différentes fractions de la purification de TRCF recombinante. A droite, immunodétection de TRCF dans un extrait total de feuilles d'Arabidopsis à l'aide des sérums immuns des différents lapins récoltés aux jours J39, J53 et J69 après le début de l'immunisation. FT: fraction totale; FI: fraction insoluble; FS: fraction soluble; E: éluat; E200: élution à 200 mM imidazole; E500: élution à 500 mM imidazole; 1053, 1054, 3404: code numéro des lapins. Les flèches montrent les bandes correspondant aux tailles attendues pour la protéine TRCF recombinante (gel de gauche) et sauvage (gel de droite).

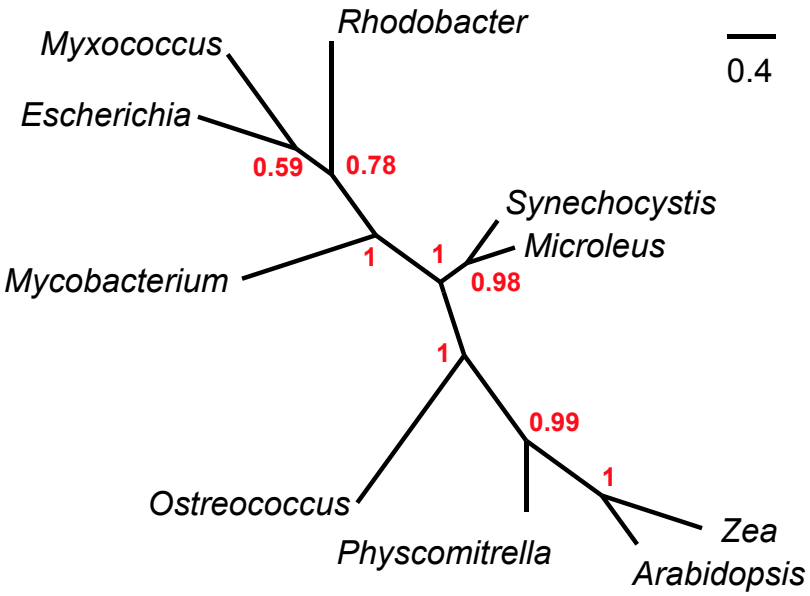


Figure 38. Arbre phylogénétique de TRCF.
Les séquences de plusieurs organismes représentant des grande lignées pertinentes de l'évolution des organelles des plantes ont été utilisées : protéobactéries (*E. coli)*, cyanobactéries (*Synechocystis*), algues (*Ostreococcus)*, mousses (*Physcomitrella*), angiospermes : monocotylédones (*Zea mays* ou maïs) et dicotylédones (*A. thaliana*). Les chiffres rouges sont les valeurs de bootstrap (entre 0 et 1), et montrent la crédibilité de la position de chaque branche.

3. Etude de TRCF

Chez les bactéries, TRCF (ou Mfd) est une DNA hélicase impliquée dans la réparation associée à la transcription, qui recrute les facteurs du NER. La protéine TRCF d'*E. coli* possède un domaine hélicase proche en séquence à celui de RecG. Chez Arabidopsis, nous avons identifié deux gènes codant pour des protéines contenant une séquence homologue à RecG. La première, RECG1, est bel et bien un homologue de RecG comme le traite le chapitre précédent. Mais l'analyse phylogénétique montre que la deuxième est une protéine homologue à TRCF. Les différents logiciels de prédiction la situent dans les plastes, mais c'est aussi le cas de nombreuses protéines dont l'adressage aux mitochondries a été démontré. Jusqu'à maintenant, aucun facteur homologue à une protéine impliquée dans le NER n'a été identifié dans les mitochondries d'animal ou de plantes, soulevant la question de l'existence de ce mécanisme de réparation dans les mitochondries. Si cette protéine s'avère avoir le même rôle dans les plantes que chez *E. coli*, cela pourrait être un premier indice de l'existence du NER dans les organelles des plantes. Des mutants d'insertion d'Arabidopsis étaient également disponibles, et nous les avons utilisés pour l'étude du rôle de TRCF dans la plante.

3.1. Fabrication d'anticorps

J'ai fait produire des anticorps contre TRCF d'Arabidopsis. Comme pour RECG1 j'ai cloné la séquence correspondant à la partie C-terminale comprenant les 300 derniers acides aminés de TRCF dans le vecteur p0GWA. La protéine exprimée s'est avérée être insoluble (figure 37). La solubilité n'était pas non plus optimale dans 8 M d'urée, mais s'est avérée acceptable en présence de 6 M de guanidium thiocyanate. Pendant la purification par affinité au nickel une grande partie de la protéine a été éluée avec peu d'imidazole (30 et 50 mM) dans les fractions de lavage, et le reste a été récupéré dans les fractions d'élution avec 200 mM et 500 mM d'imidazole. Les factions contenant la protéine ont été mélangées puis concentrées par ultrafiltration (Ultrafree, Millipore). Une concentration optimale en protéine, guanidium et imidazole a été atteinte pour permettre l'injection dans les lapins. Deux lapins (1053 et 1054) ont été immunisés par l'entreprise Covalab et un lapin (3404) par le service de production d'anticorps de l'Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC, Illkirch). Après avoir vérifié que le sérum pré-immun des lapins ne contenait pas d'anticorps pouvant interférer avec la détection de notre protéine, les animaux ont subi plusieurs injections réparties sur 2 mois. J'ai testé les différents sérums immuns sur un extrait

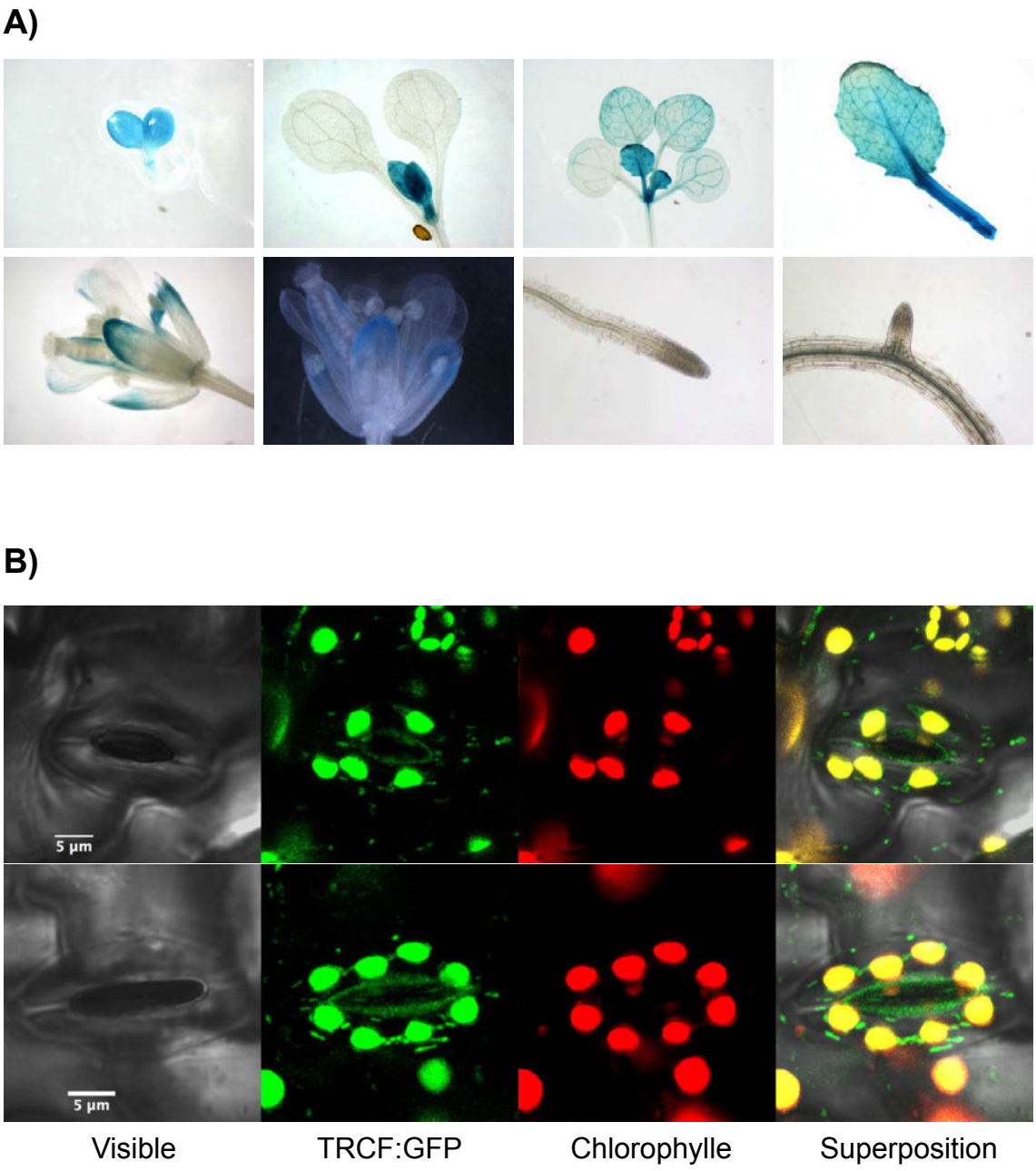


Figure 39. Localisation tissulaire et intra-cellulaire de la protéine TRCF.
A. Fusion du promoteur de *TRCF* au gène GUS. Coloration GUS de plantes pTRCF:GUS à différents stades de développement. Feuilles à 6 jours, 13 jours, 18 jours et 1 mois, fleurs et racines à 18 jours. **B.** Fluorescence de la protéine de fusion entre la partie N-terminale de TRCF et la eGFP. La fluorescence GFP co-localise avec la fluorescence naturelle de la chlorophylle et avec des structures pointillées caractéristiques des mitochondries. Les images correspondent à des cellules de garde de stomates des lignées transgéniques d'Arabidopsis exprimant la construction.

total de plantules d'Arabidopsis. L'anticorps du lapin 3404 était le seul utilisable, capable de reconnaitre TRCF, malgré le fait qu'il détecte aussi une protéine plus petite sur les extraits totaux d'Arabidopsis. La détection de la protéine dans des extraits totaux est difficile, car on observe beaucoup de bruit de fond.

3.2. Phylogénie de TRCF

Cette protéine est bien conservée dans les bactéries, chez les algues, les mousses et les angiospermes comme le montre l'arbre phylogénétique de la figure 38. Cet arbre suggère également que la protéine TRCF d'Arabidopsis descend de la protéine des cyanobactéries.

3.3. Localisation tissulaire et intracellulaire de TRCF

Pour visualiser les tissus dans lequel est exprimée TRCF, une construction promoteur:GUS a été réalisée en clonant une séquence de 381 nt en 5' du gène dans le vecteur pMDC162. La séquence couvre la région entre les 20 premiers nucléotides de la partie 5' UTR du gène jusqu'aux 136 derniers nucléotides du gène précédent. La construction a été transformée dans l'Arabidopsis par agro-transfection et des plantes résistantes à la kanamycine ont été sélectionnées. Plusieurs plantes de 5 lignées indépendantes ont été testées pour l'expression du gène GUS, par coloration au X-Gluc. Les résultats montrés sont représentatifs des observations faites sur plusieurs lignées. La coloration révèle une forte expression du gène GUS au niveau des méristèmes apicaux caulinaires, des jeunes feuilles ainsi que des sépales. Les feuilles plus âgées sont également colorées, mais de façon moins intense. Aucune coloration n'a été observée dans les cotylédons ainsi que dans les fleurs et dans tous les tissus de l'appareil reproducteur. Il en est de même pour les méristèmes racinaires et les racines, dans lesquelles le promoteur de *TRCF* ne semble pas être actif (figure 39A). Cela nous indique que le promoteur de *TRCF* est actif principalement dans les jeunes tissus chlorophylliens en formation. La localisation au niveau des méristèmes est aussi caractéristique des tissus riches en mitochondries.

Afin d'étudier la localisation intracellulaire de TRCF une construction codant pour les 93 premiers acides aminés de la partie N-terminale de TRCF fusionnée à la GFP a été réalisée. Des expériences d'expression transitoire par bombardement n'ont pas donné des résultats concluants, et des plantes transgéniques exprimant ces constructions de façon stable ont été préparées. Cependant, sur près d'une centaine de plantes issues de plusieurs lignées transformées indépendantes, seules quelques-unes exprimaient la GFP. Cela pourrait être la

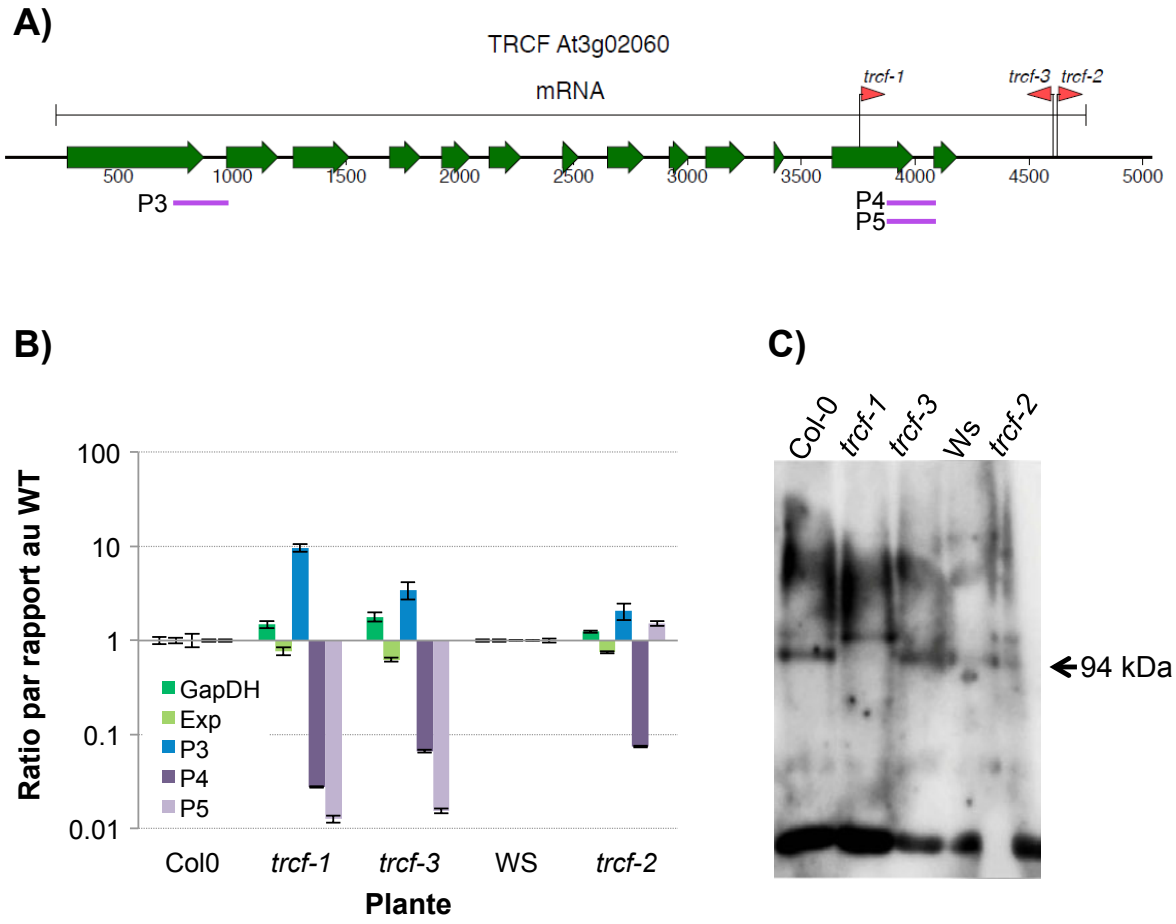


Figure 40. Mutants d'insertion *trcf*.
A. Carte du gène *TRCF* d'Arabidopsis et position des insertions T-DNA dans les différentes lignées. P3, P4 et P5 sont les séquence utilisées pour quantifier le transcrit par qRT-PCR. **B**. Quantification relative du transcrit de *TRCF* dans les différents mutants. Les résultats ont été normalisés par rapport aux gènes nucléaires de références *GAPDH* et *EXP*. L'expression du transcrit dans les mutants est comparée au sauvage de l'écotype correspondant, Col-0 pour *trcf-1* et *trcf-3*, WS pour *trcf-2*. **C.** Immunodétection de la protéine TRCF dans les différents mutants. Le test a été fait sur des extraits protéiques totaux des mutants *trcf* et des sauvages correspondants.

conséquence de phénomènes d'exctinction du gène ("silencing"). Les cellules observées montrent que la localisation de TRCF est chloroplastique, comme le proposent les sites de prédiction d'adressage, car on remarque que la fluorescence verte est superposée à la florescence naturelle rouge de la chlorophylle. De plus, des spots fluorescents nombreux et de petite taille ont été observés, représentatifs des mitochondries. Une acquisition dans le temps nous a permis de voir le mouvement de ces spots, caractéristique des mitochondries. Cependant la fluorescence était très faible, et nous n'avons pas pu faire d'expérience de co-localisation avec le MitoTracker ou autre marqueur des mitochondries. TRCF semble donc bien être adressée à la fois aux mitochondries et aux chloroplastes (figure 39B). Cependant, la localisation mitochondriale doit être encore confirmée.

3.4. Caractérisation des mutants

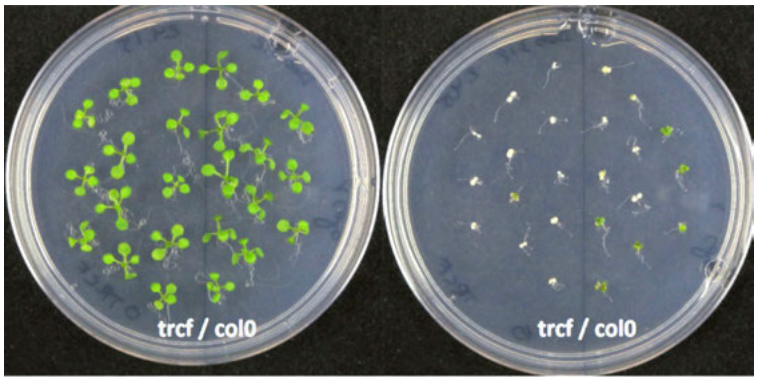
3.4.1. Localisation de l'insertion T-DNA

Pour mieux comprendre le rôle de TRCF dans Arabidopsis nous avons acquis trois mutants d'insertion de T-DNA. Aucun d'entre eux ne montre de phénotype de développement en conditions normales de culture. La carte du gène *TRCF* montrant la localisation des insertions est représentée dans la figure 40A. Le génotypage des lignées *trcf* mutantes a été fait par PCR utilisant des amorces qui encadrent la région d'insertion du T-DNA, et une amorce spécifique de la région "left border" du T-DNA. Le séquençage des produits d'amplification nous a permis de déterminer la position exacte de l'insertion T-DNA dans chacun des 3 mutants. Dans *trcf-1* l'insertion se situe dans le douzième et avant-dernier exon. Dans les mutants *trcf-2* et *trcf-3* les insertions sont localisées dans la région 3'-UTR du gène (figure 40A).

3.4.2. Quantification du transcrit de *TRCF* dans les mutants

La quantification du messager de *TRCF* a été effectuée par qRT-PCR. Les valeurs de quantification ont été normalisées par comparaison à des gènes dits "de ménage" ayant des niveaux d'expression constants dans les différents tissus de la plante. Les gènes de références utilisés étaient ceux codant l'actine (*ACT2*, At3g18780), la protéine à fonction inconnue EXP (At4g26410) et la glycéraldéhyde-3-phosphate déshydrogénase (GAPDH, At1g13440). D'après les expériences de qRT-PCR il y a une augmentation de 3 à 4 fois du transcrit de *TRCF* (figure 40B) dans le mutant *trcf-1*. Ce résultat a été vérifié dans trois expériences indépendantes. Ceci peut être expliqué par la présence de promoteurs forts dans la séquence

A)

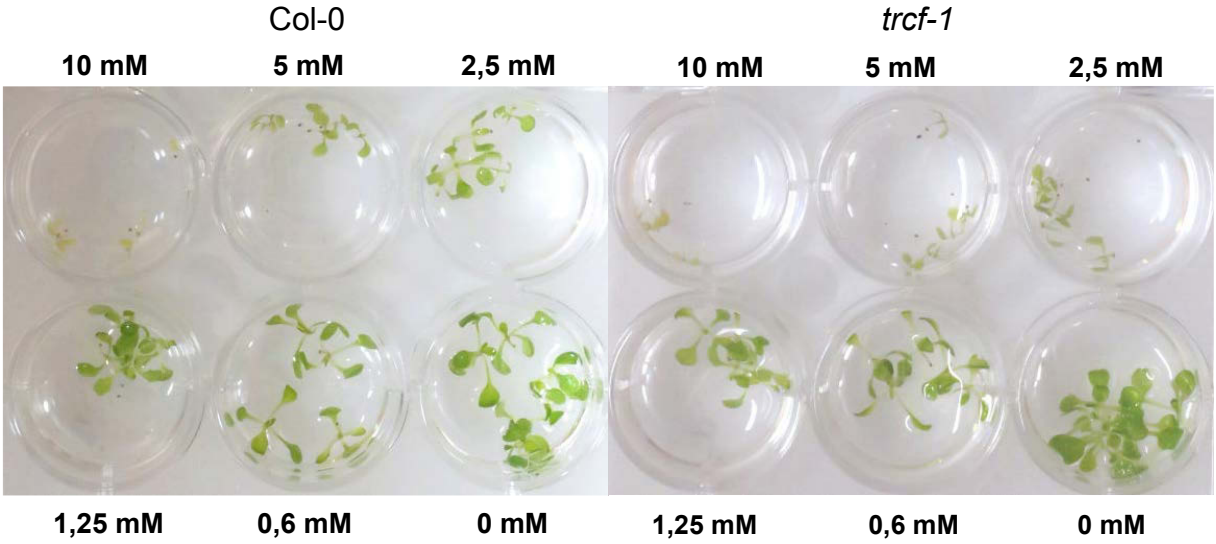


0 mM 10 mM

B)

	Col-0	*trcf-1*
0 mM	0%	0%
5 mM	5%	21%
10 mM	47%	64%
n=	81	78

C)



D)

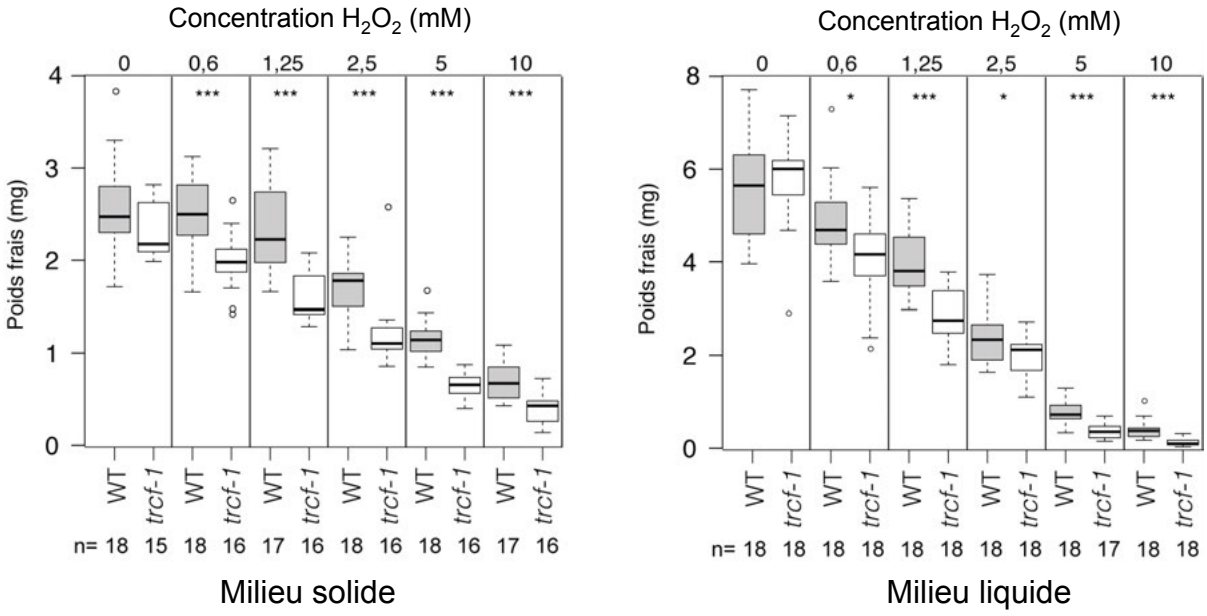


Figure 41: Tests de sensibilité au peroxyde d'hydrogène du mutant *trcf-1* par rapport au sauvage.

A et C. Plantules de 10 jours, 5 jours après exposition à l'H_2O_2 en milieu solide (A) ou en milieu liquide (C). **B.** Pourcentage de plantules albinos pour le témoin et le mutant à 0,5 et 10 mM. **D.** Box-plot des taux de croissance en milieu solide ou liquide. Chaque valeur correspond au poids moyen déterminée pour des pools de 5 plantules. Les astérisques indiquent la confiance statistique (T test, * pour P<0,01; *** pour P<0,001).

du T-DNA, qui peuvent avoir une influence positive sur le promoteur de TRCF. Par contre, en utilisant des amorces qui amplifient la région en 3' de l'insertion T-DNA, une baisse significative du niveau du transcrit a été observée (figure 40B). L'étude par Western blot et immunodétection avec l'anticorps dirigée contre TRCF a montré une protéine de la taille attendue (environ 90 kDa) détectée dans la plante sauvage, mais pas dans *trcf-1* (figure 40C). La mutant *trcf-1* serait bien un mutant knock-out. Les mutants *trcf-2* et *trcf-3* sont des knock-down, dont le transcrit reste tout de même exprimé à des niveaux non-négligeables. En effet on observe une diminution d'environ la moitié du taux de transcrit (75% pour *trcf-2* et 50% pour *trcf-3*). Par contre on n'observe pas dans ces mutants de diminution détectable de la protéine qui correspondrait à TRCF. Le mutant *trcf-1* sera le seul testé dans la plupart des études suivantes.

3.5. Etude des rôles de TRCF

3.5.1. Sensibilité au stress oxydative: peroxyde d'hydrogène

Pour tester le rôle de TRCF dans la réparation, nous avons testé la résistance des plantes au stress oxydatif. Pour cela nous avons cultivé des plantes en présence d'H_2O_2, soit en milieu gélosé soit en milieu liquide, puis comparé leur poids frais à celui du sauvage. L'expérience a été reproduite trois fois de façon indépendante. et dans les trois réplicas nous avons observé les mêmes résultats. L'H_2O_2 provoque une inhibition de la croissance observable dès les plus faibles concentrations (0,6 et 1,25 mM), et qui se poursuit jusqu'à des concentrations provoquant l'arrêt de la croissance et la mort de la plante (10 mM). Le mutant *trcf-1* est plus sensible que le sauvage au stress oxydatif provoqué par l'eau oxygénée, les différences pouvant aller jusqu'à ¼ du poids par rapport au sauvage. Cette sensibilité augmente avec la concentration en agent génotoxique. Ces résultats suggèrent que TRCF est impliquée dans la réparation de l'ADN ayant subi des dommages dû aux stress oxydatifs (figure 41).

3.5.2. Sensibilité aux UV-C

Les lésions induites par l'irradiation UV sont celles les plus ciblées par le NER. Nous avons donc testé la sensibilité du mutant *trcf-1* aux UV. Des plantules ont été traitées aux UV comme décrit dans "Méthodes". Les résultats obtenus montrent clairement une sensibilité accrue de *trcf-1* par rapport aux plantes sauvages contrôles. En effet les plantes montrent un retard de développement dès la dose minimale testée (750 Joules/m_2) (plus de 50% de poids frais en moins par rapport au non-traité), alors que le WT semble peu affecté (figures 42A et

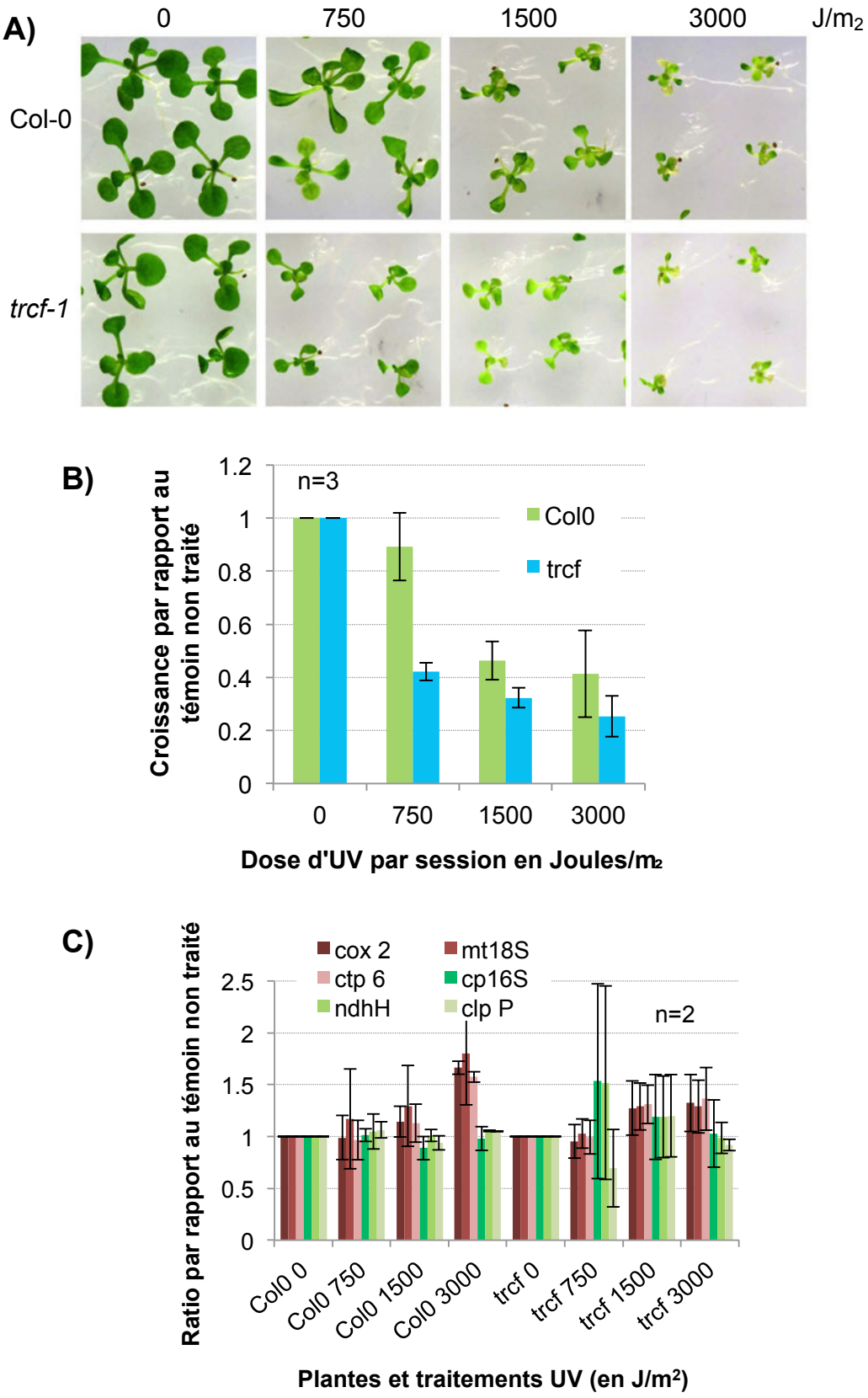


Figure 42. Tests de sensibilité aux UV-C du mutant *trcf-1* par rapport au sauvage. **A.** Plantes exposées à différentes doses d'UV-C. **B.** Taux de croissance des plantules à différents niveaux d'irradiation. Les plantes sont pesées à l'âge de 12 jours. **C.** Quantification relative du nombre de copies de séquences mitochondriales et chloroplastiques par qPCR. Résultats normalisés par rapport aux gènes nucléaires. En rouge les gènes mitochondriaux et en vert les gènes chloroplastiques.

42B). Les UV-C sont connus pour causer des dommages à l'ADN et plus particulièrement des dommages de type inter-pontage, comme les dimères de pyrimidine. Cette expérience nous conforte quant au possible rôle de TRCF dans la réparation de l'ADN des organelles.

3.5.3. Nombre de copies du mtDNA et du cpDNA dans les plantes stressées

Afin de déceler une éventuelle action de TRCF sur la régulation du nombre de copies des génomes des mitochondries et des chloroplastes, nous avons quantifié le nombre de copies de plusieurs gènes représentatifs par qPCR, en conditions normales et en conditions de stress. Le nombre de copies du mtDNA augmente (150%) lorsque les plantes témoins sont soumises aux UV, mais il en est de même pour *trcf-1*, et on n'observe aucune différence significative entre mutant et sauvage. Il en est de même du génome du chloroplaste pour lequel le nombre de copies ne varie pas significativement même en état de stress (figure 42C). Les valeurs ont été normalisées par rapport aux trois gènes nucléaires de référence actine (*ACT1*, At2g37620), polyubiquitine (*UBI10*, At4g05320) et ARN 18S (*n18S*, 18S RRNA). Les résultats suivant la même tendance lors du stress au peroxyde d'hydrogène, ceux-ci ne sont pas représentés.

3.5.4. Mesure de la recombinaison ectopique

La recombinaison homologue ectopique impliquant des courtes séquences répétées a aussi été testée par qPCR (figure 43). L'augmentation de ce type de recombinaison ectopique a été extensivement décrite comme un phénotype moléculaire caractéristique de mutants impliqués dans la régulation de la réparation par recombinaison dans la mitochondrie (Miller-Messmer *et al.*, 2012). Comme l'indique les résultats présentés dans la figure 43, il n'y a pas de différences significatives entre WT et *trcf-1*, en conditions normales de culture. En conditions de stress par les UV on remarque chez le WT une augmentation de 4 à 5 fois des formes recombinées mitochondriales X-2/1 et EE-2/1, résultant de la recombinaison impliquant les séquences répétées X (196 bp) et EE (127 bp) respectivement. Cependant, cette augmentation est aussi observée dans *trcf-1*. On observe aussi dans le WT et mutant une augmentation d'environ 10 fois de la forme R2-2/1, qui résulte de la recombinaison impliquant la séquence répétée R2 (126 bp) du chloroplaste. Cependant, les différences entre WT et *trcf-1* ne sont pas significatives. Les résultats obtenus avec les stress au peroxyde d'hydrogène montrent quant à eux une légère diminution des formes recombinées de R2 en condition de stress, mais aucune différence entre le témoin et le mutant. Ces résultats ne sont pas présentés. Ces résultats nous indiquent que la protéine TRCF d'Arabidopsis, bien que

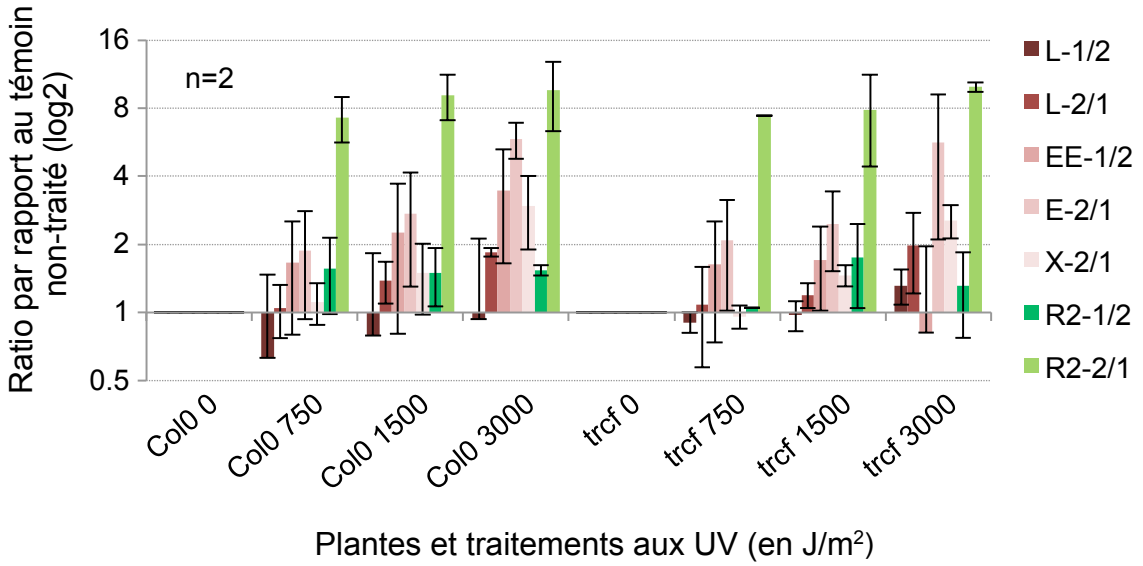


Figure 43. Effet de la mutation *trcf-1* sur la recombinaison ectopique du mtDNA. Quantification des formes recombinées des courtes séquences répétées mitochondriales (L, EE et X) et chloroplastique (R2), pour les plates traitées à 0, 750, 1500 et 3000 J/m^2.

contribuant à la résistance de la plante au stress oxydatif, ne semble pas jouer de rôle important dans la réparation par recombinaison homologue des génomes des organelles. Cela pourrait vouloir dire que ce facteur est bien un membre du NER, et n'est donc pas impliqué dans la réparation par recombinaison. De plus, la sensibilité accrue aux UV du mutant peut être un indice quant à l'existence du NER, qui répare préférentiellement les lésions UV.

Les possibilités d'étude sur TRCF ne sont pas terminées et c'est un doctorant arrivé en début d'année qui prendra le relais. Les perspectives à ce sujet sont discutées dans le chapitre suivant.

Discussion et Perspectives

1. Recombinaison *in organello*

La transformation de mitochondries de plantes permettrait d'obtenir des mutants de gènes mitochondriaux, d'étudier les séquences impliquées dans leur régulation, ou bien d'exprimer dans les mitochondries des séquences exogènes, comme par exemple des séquences responsables de CMS. Pour le moment aucun système ne permet d'acheminer de l'ADN exogène jusqu'aux mitochondries des plantes *in vivo* et de l'intégrer dans le génome. Mon travail portait sur l'intégration d'ADN exogène dans le mtDNA de mitochondries isolées. Malheureusement, le système de recombinaison *in organello* dans Arabidopsis demande une optimisation plus poussée. La qualité des mitochondries extraites de tissus verts d'Arabidopsis n'est pas optimale comparé aux mitochondries de tubercule de pomme de terre, ce qui inclut la variable de la qualité des mitochondries dans les expériences. Les étapes d'extraction sont longues et les mitochondries d'Arabidopsis sont peut-être plus fragiles. De plus, l'approche d'analyse par PCR crée trop d'artefacts, malgré l'essai de diverses polymérases provenant de différents fournisseurs. Et comme l'objectif initial était de pouvoir tester des mitochondries extraites de différents mutants de recombinaison, cela n'est pas possible sans une approche plus robuste et quantitative.

2. RECG1

L'étude de la protéine RECG1 d'Arabidopsis, homologue à RecG bactérienne montre qu'elle est doublement adressée aux organelles et qu'elle est exprimée de façon ubiquitaire dans la plante. Cette protéine est conservée dans les plantes à fleurs (Angiospermes), mousses et algues, et est phylogénétiquement proche de celle des cyanobactéries, ce qui fait penser qu'elle a été héritée de l'ancêtre symbiotique du chloroplaste. La conservation des cette protéine dans les plantes suggère qu'elle a conservé des fonctions importantes au cours de l'évolution. La structure de RecG est relativement conservée des bactéries à Arabidopsis comme le montre la modélisation structurale de RECG1. Nous avons montré que RECG1 a en effet conservé plusieurs de ses fonctions dans les plantes. Le rôle de RECG1 dans les mitochondries et les chloroplastes est discuté dans notre publication (Wallet *et al.*, 2015). J'en reprends ici les grandes lignes au regard des données supplémentaires qui n'ont pas été publiées.

2.1. RECG1 et son rôle dans les mitochondries

Je me suis intéressée dans un premier temps aux évènements qui affectent le mtDNA, et ces données constituent la plus grosse partie de mon travail qui a été publié. Nous avons montré que RECG1 a de multiples rôles dans les mitochondries d'Arabidopsis.

2.1.1. RECG1 a un rôle dans la réparation par HR, mais des pièces sont manquantes

Dans *E. coli*, RecG a un rôle important dans la réparation des DSB par HR. Nous avons montré que RECG1 complémente partiellement RecG pour la réparation après exposition aux UV-C. Ce premier indice suggère que RECG1 peut, comme RecG, intervenir dans la réparation.

Chez les plantes sauvages, le traitement à la CIP qui produit des DSB dans le mtDNA entraîne une augmentation de la recombinaison entre IRs (recombinaison ectopique). Cela suggère que la réparation de DSB mobilise des mécanismes de recombinaison homologue, et entraîne par conséquence une augmentation de la recombinaison ectopique. A l'inverse, le traitement du mutant d'Arabidopsis *recG1* à la CIP n'entraîne pas d'augmentation de la recombinaison ectopique. En d'autres termes, le mécanisme de réparation de DSB par recombinaison est moins actif dans le mutant. Ce phénotype moléculaire est également observé chez les mutants des gènes *RECA3*, *ODB1* (Janicka *et al.*, 2012; Miller-Messmer *et al.*, 2012) et *POL1B* (résultats non-publiés du laboratoire), impliqués dans la réparation par recombinaison dans les mitochondries de plantes. Cela suggère que RECG1 agit également dans ce mécanisme.

Les mutants *recG1* montrent une augmentation de la recombinaison ectopique entre les IR en conditions normales de développement. Ce phénotype moléculaire est aussi observé chez les mutants des gènes *RECA2*, *RECA3*, *OSB1* et *MSH1*, qui ont un effet de suppresseurs de la recombinaison ectopique dans le WT (Miller-Messmer *et al.*, 2012; Shedge *et al.*, 2010; Zaegel *et al.*, 2006). Cela suggère qu'en absence de ces protéines, une voie de réparation par recombinaison générant des erreurs est stimulée. La voie du SSA ou plus vraisemblablement celle du BIR pourrait donc être employée. RECG1 a donc un effet de suppresseur de la recombinaison ectopique et cela confirme son implication dans des mécanismes de réparation par recombinaison.

La mutation de *RECG1* a un léger effet sur le phénotype de développement, comparé aux phénotypes d'autres mutants comme *osb1*, *msh1* ou *recA2* (Miller-Messmer *et al.*, 2012; Shedge *et al.*, 2007; Zaegel *et al.*, 2006). Cela signifie que la présence RECG1 n'est pas essentielle au développement de la plante. Or la réparation de DSB par recombinaison RecA-dépendante fait intervenir des HJ et une protéine catalysant leur migration est indispensable pour la synthèse d'ADN. Soit RECG1 a une fonction non-essentielle dans la plante, et n'est utile qu'en conditions particulières (stress, stade de développement, etc.), soit un mécanisme alternatif peut prendre en charge la fonction normalement assumée par RECG1.

Un mécanisme alternatif à la réparation RecA-dépendante est la réparation par SSA, qui ne fait pas intervenir de HJ. Mais cette voie est connue pour être non-conservative et produire des délétions de larges portions du génome. Il est donc possible qu'une autre protéine intervienne dans la migration des HJ. Chez *E. coli*, une autre hélicase, RuvAB, prend en charge les migrations de HJ. Or on ne trouve pas d'homologue de cette protéine chez Arabidopsis. RadA est une protéine bactérienne dont la mutation est synergique avec celle de RecG pour la sensibilité aux stress génotoxiques. Son rôle dans la migration d'intermédiaires de recombinaison et l'élargissement des zones de recombinaison a été récemment montré dans *E. coli* (Cooper et Lovett, 2016). Un homologue de RadA adressé aux mitochondries d'Arabidopsis a été identifié au laboratoire (données non-publiées). Cette protéine pourrait permettre la migration des HJ à la place de RECG1. L'étude des doubles mutants *radA recG1* permettrait de tester cette hypothèse et de savoir si ces protéines ont un rôle redondant dans la migration des HJ. Un homologue de FANCM ainsi que de HTLF (RAD5A) et plusieurs RecQ ont été identifié dans Arabidospsis (Knoll *et al.*, 2012; Mannuss *et al.*, 2010). Ces hélicases prennent également en charge les HJ, mais aucune n'est prédite comme adressée aux organelles dans Arabidopsis.

RuvAB et RecG ont des fonctions différentes dans *E. coli* (Zhang *et al.*, 2010). En effet RuvAB recrute RuvC, une nucléase qui résout les HJ en catalysant leur clivage. Le fait que RECG1 ne complémente pas le mutant bactérien *ruvA* pour la réparation des lésions UV montre que RECG1 a bien le même rôle que RecG et n'assure pas les fonctions de RuvC. Cela semble normal, car RECG1 n'a pas de domaine d'interaction avec RuvC. RecG n'a pas de nucléase associée, et la résolvase RuvC, la seule connue dans *E. coli*, n'a pas d'homologue chez Arabidopsis. Pour le moment, aucune protéine avec cette fonction n'a été identifiée dans les mitochondries de plantes. Les deux résolvases d'Arabidopsis déjà identifiées, AtGEN1 et AtSEND1, sont de type eucaryotique et sont apparemment adressées au noyau (Bauknecht et

Kobbe, 2014). Cette fonction de résolvase est importante pour la voie du DSBR, mais des mécanismes alternatifs peuvent exister. En effet RECG1 pourrait simplement catalyser la migration de HJ jusqu'à a la fin de la molécule d'ADN et se détacher. Ou bien la résolution/dissolution des HJ pourrait se faire par l'action de topoisomérases, qui existent dans la mitochondrie. Les voies du SDSA et/ou du BIR sont également probablement favorisées pour la réparation RecA-dépendante des DSB dans les mitochondries.

2.1.2. RECG1 et son rôle dans la réplication

Comme RecG, RECG1 a probablement un rôle important dans la réparation associée à la réplication, mais aussi dans l'inhibition de la formation de certains intermédiaires de recombinaison qui peuvent ré-initier la réplication.

2.1.2.1. RECG1 et la régression de fourches de réplication

Chez *recG1*, la recombinaison ectopique augmente en conditions normales de croissance. Ce phénotype moléculaire est similaire à celui observé chez les plantes sauvages traitées à la CIP. Cela signifierait donc que l'absence de RECG1 provoque la création de DSB. Chez *E. coli* RecG catalyse la régression de fourches de réplication et permet le contournement de lésions qui bloquent la progression de la fourche sur le brin retard. Si RECG1 a conservé ce rôle dans Arabidopsis, sa mutation supprime cette possibilité. La ré-initiation de la réplication par invasion de brin, RecA-dépendante, serait la voie possible pour le sauvetage des fourches de réplication chez les mutants *recG1* (Wallet et *al.*, 2015, figure S6 de l'article). Cette voie de sauvetage des fourches provoque le détachement des brins néoformés inachevés, ce qui revient à provoquer des DSB dans le brin néoformé et à la ré-initiation de la réplication par invasion de brin (D-loop). Cela expliquerait donc le phénotype moléculaire de RECG1 en conditions normales de développement.

RecA permet également la régression des fourches de réplication en particulier lorsqu'il y a un SSG dans le brin retard (Cox, 2007). Les homologues de RecA dans Arabidopsis pourraient avoir conservé ce rôle. En effet, la mutation *recG1* a un faible impact sur le phénotype de développement des plantes, alors que la double mutation *recG1 recA3* donne des plantes très affectées (voir figure 7 de Wallet et *al.*, 2015). Cela suggère que RECA3 et RECG1 agissent dans des voies alternatives et redondantes de la réparation associée à la réplication. Les plantes arrivent malgré tout à survivre car la deuxième RecA-like mitochondriale, RECA2, peut probablement participer aussi au sauvetage de fourches de réplication (Miller-Messmer *et al.*, 2012).

2.1.2.2. RECG1 et l'inhibition de la réplication

RECG1 complémente RecG pour l'inhibition de la réplication de plasmides contenant l'origine de réplication ColE1, qui est initiée par un R-loop. Au même titre que RecG, RECG1 semble donc déstabiliser les intermédiaires de recombinaison en R-loop (Hong *et al.*, 1995). RECG1 complémente aussi la protéine bactérienne dans l'inhibition de la réplication pathologique qui a été observé quand les fourches de réplication qui progressent en sens opposées se rencontrent (Rudolph *et al.*, 2009; Rudolph *et al.*, 2013). Ce résultat suggère que RECG1 est capable de dissoudre des structures en 3'-flap qui peuvent être prises par PriA pour la ré-initiation de la réplication. Ces deux expériences montrent que RECG1 a probablement un rôle dans la déstabilisation d'intermédiaires de recombinaison qui peuvent initier la réplication.

Dans les mutants *recG1-2*, on observe l'apparition d'une molécule circulaire de 8 kb contenant le gène *ATP1* et se répliquant apparemment de façon autonome. Cette molécule est présente en nombre de copies plus grand que le reste du génome. C'est probablement l'augmentation de la recombinaison ectopique dans *recG1-2* qui favorise la recombinaison entre les séquences répétées. La relative proximité des deux mêmes séquences répétées EE (8 kb), permet cette circularisation dans le mutant (figures 19 et 20). Cette forme existe dans les plantes sauvages mais en des quantités infimes. Sa présence est donc inhibée par RECG1. Le fort taux de recombinaison dans le mutant permettrait de générer cette molécule de façon continue et donc d'augmenter sa concentration. Mais il est alors difficile d'expliquer comment cette molécule existe à un nombre de copies plus élevé que les molécules qui lui aurait donné naissance par recombinaison. Sa quantité élevée pourrait s'expliquer par un phénomène de réplication autonome qui n'est pas synchronisée avec le reste du génome. En effet, nous avons montré que RECG1 déstabilise certains intermédiaires de recombinaison dans les bactéries. Si cette fonction est conservée dans les plantes, il se pourrait que RECG1 déstabilise les intermédiaires de recombinaison pouvant initier la réplication d'une telle molécule. De plus, la présence du gène *ATP1* qui est hautement transcrit pourrait favoriser la réplication par R-loop ou par recombinaison. En absence de RECG1, la réplication de cet épisome ne serait donc plus inhibée. La ségrégation préférentielle de cette molécule pourrait également expliquer sa présence en quantités élevées.

Dans les doubles mutants *recA3 recG1*, on observe une différence de stœchiométrie entre les différentes régions du mtDNA. Les régions dont le nombre de copies est augmenté correspondent à des séquences situées la plupart du temps entre une paire d'IR en sens direct.

Dans ces plantes, la recombinaison entre IR augmente et donne lieu à des molécules pouvant se répliquer à haute fréquence et de façon autonome. Tous les doubles mutants ne présentent pas la même réorganisation moléculaire, ce qui montre que dans ces plantes le contrôle de la stœchiométrie est perdu. RECA3 et RECG1 agissent donc probablement dans des voies alternatives redondantes permettant de contrôler la stœchiométrie des différentes régions du mtDNA. Il est à noter que seulement certains des produits issus de la recombinaison entre IRs ont la capacité à se répliquer et ségréger de façon indépendante.

2.1.3. RECG1 et son rôle dans la ségrégation

La réintroduction du gène *RECG1* par rétrocroisement avec du pollen sauvage nous a permis d'observer l'effet de RECG1 sur les mitochondries maternellement héritées du mutant *recG1-2*. Nous étions, entre autres, intéressés par le devenir de la molécule circulaire répliquée de façon autonome. Dans le mutant *recG1-2* cette molécule était stable, et nous n'avons pas observé de ségrégation même dans les générations suivantes. Mais dans les plantes issues de rétrocroisement les différentes versions du mtDNA ont ségrégée. Cette ségrégation ne semble pas possible sans le gène RECG1. Il s'est avéré que dans toutes les catégories de plantes F1 la molécule circulaire avait été perdue. Les plantes ne présentant pas de phénotype de développement peuvent être divisées en deux catégories, selon leur version de cette région du mtDNA. Dans les plantes appelées "revertant", seule la version sauvage est présente. Chez les plantes appelées "shifted", le gène *ATP1* a été réintégré dans le génome par un second événement de recombinaison faisant intervenir une autre paire d'IRs. Nous avons alors obtenu des plantes contenant une version réarrangée mais stable du mtDNA. Il s'est avéré que ces différentes versions du mtDNA étaient déjà présentes dans le mutant *recG1-2*. Il semble donc que RECG1 est nécessaire à la ségrégation des différentes versions de cette région du mtDNA. RECG1 pourrait avoir un effet sur la déstabilisation de D-loop et R-loop, comme il a été montré par complémentation sur les bactéries, et par conséquence inhiber la réplication de cet épisome si ces structures sont intermédiaires de sa réplication. RECG1 pourrait également avoir un effet sur la structure du mtDNA. En effet, une grande partie de la séquence que contient l'épisome EE-2/1 migre avec les molécules de haut poids moléculaire. Il se pourrait donc que l'épisome soit présent sous forme de concatémères ou de formes circulaire caténaires associés avec le mtDNA principal. RECG1 permettrait la résolution de ces structures complexes, et leur dissociation, ce qui favoriserait la ségrégation de ces molécules. Cependant cette décaténation pourrait être assurée par les topoisomérase présentes dans les mitochondries. RECG1 pourrait également permettre une ségrégation sélective des

différentes formes et empêcher la transmission de cette petite molécule d'ADN fragmentant le génome.

Parmi les plantes F1 obtenues certaines, appelées "variegated" présentaient des phénotypes typiques de feuilles déformées et tachetées. Un phénotype similaire a été observé chez d'autres mutants comme *msh1* ou *osb1* (Arrieta-Montiel *et al.*, 2009; Zaegel *et al.*, 2006). Les doubles mutants *recG1 recA3* étaient aussi déformés à différents degrés, mais la réorganisation de l'ADN dans ces plantes doubles mutantes était tellement importante que nous n'avons pas pu faire de lien entre un phénotype de développement et un défaut particulier de la chaine respiratoire. Chez les *recG1-2* F1 déformées, les versions du mtDNA contenant le gène *ATP1* n'ont pas été retenues. Nous avons donc pu expliquer le phénotype de développement par la forte diminution du nombre de copies du gène *ATP1*, et par conséquence de la protéine. Nous avons donc pu lier directement ce phénotype avec la déficience en ATP1. Ces plantes présentaient parfois deux types de tissus, l'un affecté et l'autre normal. Cela montre que l'effet est clonal et que RECG1 agit sur la ségrégation à un stade précoce. Il est cependant difficile de savoir à quel moment du développement cette ségrégation est réalisée.

La complémentation par transformation du gène RECG1 dans le mutant *recG1-2* est une façon alternative de réintroduire RECG1 dans un cytoplasme de mutant. Cette expérience a confirmé que RECG1 permettait bien la ségrégation. En effet les plantes ne présentant pas de phénotype de développement sont des "revertants" qui ont perdu l'épisome et principalement retenu la version sauvage du mtDNA. Cependant, le pourcentage de plantes présentant le phénotype déformées et tachetées a été moins élevé que dans les expériences de rétrocroisement. De plus nous n'avons pas observé de plantes "shifted". Cela pourrait être la conséquence de la sélection de plantes transformantes sur antibiotique, qui pourrait défavoriser la croissance de plantes déficientes en ATP1 et celle des plantes "shifted" par un désavantage non-visible en conditions normales de cultures. Plus vraisemblablement, la ségrégation des versions "shifted" pourrait dépendre du nombre de copies de ces versions dans la plante maternelle. En effet nous avons montré que différentes versions du mtDNA étaient déjà présente dans le mutant, cependant nous ne savons pas précisément en quelles proportions elles sont présentes et si ces proportions sont constantes d'une plante à l'autre. D'autres facteurs, internes ou environnementaux pourraient également avoir un impact sur la ségrégation d'une ou l'autre version. L'étude d'un plus grand nombre de plante permettrait de confirmer encore ces résultats.

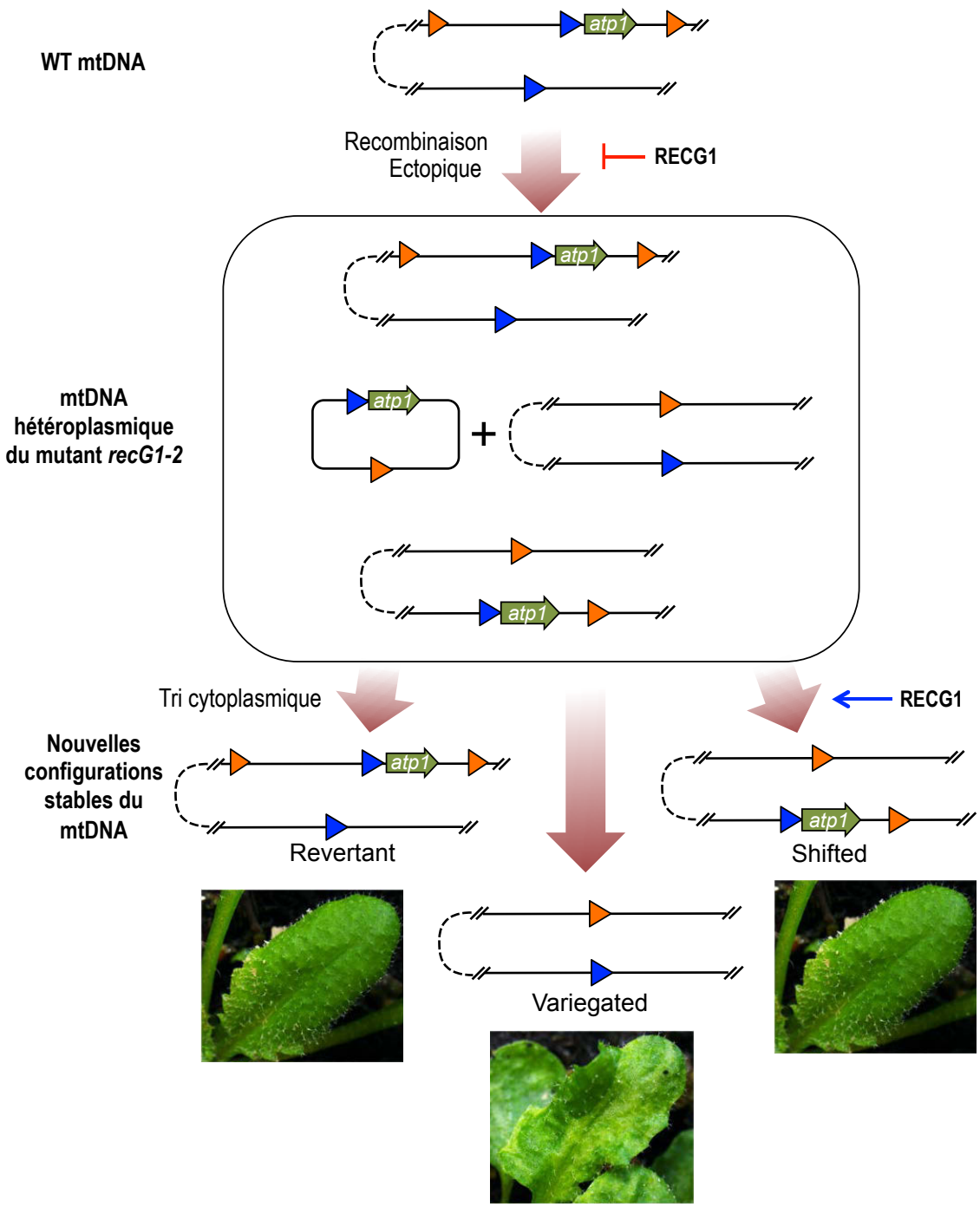


Figure 44. Schéma récapitulatif des évènements de recombinaison et de ségrégation liés à RECG1. La mutation de RECG1 entraine une augmentation de la recombinaison ectopique, ce qui provoque l'augmentation de l'hétéroplasmie du mtDNA , et provoque l'apparition d'une molécule autonome contenant *ATP1*. La réintroduction de RECG1 provoque la disparition de la molécule circulaire, et un "tri" des différentes configurations du mtDNA. Les plantes ayant héritées des formes contenant ATP1 sont normales. Les plantes ayant perdu *ATP1* sont affectées. Le gène *ATP1* est représenté en vert. Les flèches oranges représentent les séquences répétées EE. Les flèches bleues, les séquences répétées T entre lesquelles un événement de recombinaison secondaire permet l'insertion d'ATP1 à un autre endroit dans le génome.

Pour résumer, l'absence de RECG1 entraîne une augmentation de la recombinaison ectopique et donc une augmentation de l'hétéroplasmie dans le mutant *recG1-2*. La réintroduction de RECG1 permet un tri, et une distribution des différentes versions du mtDNA, tout en inhibant la réplication autonome de l'épisome EE-2/1. (figure 44).

2.2. RECG1 et son rôle dans les chloroplastes

RECG1 est doublement adressée aux mitochondries et aux chloroplastes mais jusqu'à présent nous n'avions pas observé d'effet significatif de la mutation de RECG1 sur la maintenance du cpDNA. Cependant, il a été montré dans la mousse *Physcomitrella patens* que l'homologue de RECG1 est impliqué dans le contrôle des évènements de recombinaisons entre séquences répétées dans la mitochondrie, mais aussi dans le chloroplaste (Odahara *et al.*, 2015). Chez Arabidopsis, le cpDNA est présent sous la forme d'un unique chromosome circulaire de 154,5 kb (Sato *et al.*, 1999), et ne possède pas de multiples IRs de plus de 100 bp, comme le mtDNA. Les séquences répétées appelées R2 de 76 pb que nous avons utilisées comme marqueur de la recombinaison ectopique dans le chloroplaste (voir figures 4 et 5 de Wallet et *al*., 2015) ne sont pas complétement identiques et ne sont donc pas un substrat optimal pour la recombinaison. Nos résultats sous-estiment probablement l'effet de la mutation *recG1* sur la recombinaison du cpDNA.

Il semblerait pourtant que RECG1 ait effectivement un effet sur le chloroplaste. En effet, le criblage d'un nombre plus important de plantes *recG1-2* F1 a permis de sélectionner des plantes d'un autre type présentant des bigarrures caractéristiques d'un disfonctionnement du chloroplaste. De plus, le phénotype observé sur les doubles mutants *pol1B recG1* n'a pas pu être relié à un effet sur le mtDNA. Nous n'avons pas non plus mesuré de changement de nombres de copies et de stœchiométrie du cpDNA pouvant expliquer de tels phénotypes. Chez les plantes F1 bigarrés on observe une délimitation franche entre les zones blanches et les zones vertes, montrant que l'effet est clonal. Cela signifie qu'un évènement se produit à un stade précoce dans une cellule, puis elle transmet son génome à l'ensemble des cellules provenant de sa réplication. Cet effet clonal semble indiquer que RECG1 pourrait aussi agir dans la ségrégation du cpDNA. Un séquençage par NGS du cpDNA des tissus albinos pourrait expliquer quel évènement se produit en absence de RECG1.

3. Le rôle de TRCF dans les organelles

3.1. TRCF et le NER dans les organelles

La présence d'une protéine homologue à l'une des protéines du TCR-NER d'*E. coli* n'a jamais encore été montré dans les mitochondries humaines ou végétales. L'homologue de TRCF d'Arabidopsis est donc un premier indice de la présence du NER. Cependant, il reste à montrer qu'elle a vraiment un rôle dans cette voie. Le NER prend préférentiellement en charge les lésions UV, mais il peut également exciser les bases oxydées. La sensibilité des mutants aux UV et aux stress oxydatifs pourrait être un premier indice montrant que les plantes *trcf* ne réparent pas ou peu les lésions induites par les UV. Les mutants *trcf* ne montrent pas de phénotype de développement en conditions normales de croissance, et sur plusieurs générations, ce qui peut paraître surprenant quand on connait l'importance du NER dans les organismes. Ce mécanisme, s'il existe, a peut-être une importance moindre dans les mitochondries, où d'autres mécanismes de réparation sont déjà présents. Pourtant, dans un génome où peut de séquences sont codantes, il serait pertinent qu'un mécanisme de réparation préférentielle des séquences transcrites existe. De plus, il a été montré que dans le mtDNA les séquences codantes évoluent très lentement, alors que les séquences non codantes évoluent rapidement (Allen *et al.*, 2007). Cette constatation suggère qu'il existe un mécanisme de surveillance préférentielle des séquences codantes. Cette protéine ne semble pas avoir de rôle dans le contrôle de la recombinaison ectopique ni dans la réparation par HR, car nous n'avons pas décelé de changement significatif dans le taux de recombinaison des IRs chez le mutant.

L'introduction de la protéine TRCF d'Arabidopsis dans le mutant bactérien *trcf* et sa complémentation dans la résistance aux UV serait un premier indice de la conservation de son activité NER dans les plantes. Cependant, il n'existe pour l'instant pas d'homologue connu aux protéines partenaire de la protéine TRCF bactérienne dans les plantes. Il est donc probable que la protéine TRCF de plante ait évoluée en perdant certains domaines d'interaction.

Dans les plantes, l'étude du nombre de mutations des séquences codantes, par rapport à celui des séquences non-codantes dans les mutants *trcf* d'Arbidopsis permettrait de savoir si un mécanisme de réparation préférentielle des séquences transcrites existe. Pour cela, une approche NGS adaptée permettrait d'avoir des informations précises.

3.2. Protéines partenaires

Dans les bactéries, TRCF détecte la RNAP bloquée et recrute UvrA via ses domaines D1a, D1b et D2 contenant une similarité à UvrB. Ces domaines d'interaction ne semblent pas être conservé dans la TRCF des plantes, et de plus il n'existe pas d'homologue à UvrAB chez Arabidopsis. Cependant, le domaine D4 d'interaction avec la RNAP est apparemment conservé dans la TRCF des plantes. TRCF pourrait donc interagir avec une RNAP et recruter d'autres facteurs. Les dicotylédones possèdent trois RNAP d'origine phagique, codées par le noyau, adressées aux organelles (Hedtke *et al.*, 1997). Deux d'entre elles, RPOTm et RPOTmp, sont adressées aux mitochondries. RPOTmp est également adressée aux chlorpolastes (Hedtke *et al.*, 1997, 2000). RPOTm est essentielle, car le mutant n'est pas viable, alors que RPOTmp semble impliquée dans la transcription d'un sous-groupe des gènes mitochondriaux (Kuhn *et al.*, 2009). Ces RNAP ne sont donc pas l'homologue de la RNAP bactérienne, et il est possible qu'il n'y ait pas d'interactions possibles avec TRCF. Des expériences d'immuno-précipitaiton et de double hybride permettraient de détecter les protéines partenaires et de savoir si elles peuvent interagir entre elles.

Cependant on peut se questionner sur l'importance de TRCF dans les mitochondries. En effet, les expériences de localisations intra-cellulaires ne sont pas claires. La fluorescence est surtout visible dans le chloroplaste, mais elle est moins forte dans les mitochondries, et parfois même difficile à visualiser. De plus, certaines protéines sont doublement adressées, mais ont une activité uniquement mitochondriale ou chloroplastique. C'est le cas de RECA2 qui est doublement adressé mais qui semble affecter uniquement la maintenance du mtDNA (Miller-Messmer *et al.*, 2012). Dès lors on peut se questionner sur le rôle que pourrait avoir cette protéine dans les chloroplastes.

3.3. Rôle de TRCF dans les chloroplastes

Deux RNAP d'origine phagique, RPOTp et RPOTmp, sont codées par le noyau et sont adressées au chloroplaste (Hedtke *et al.*, 1997, 2000). En addition, une polymérase d'origine bactérienne, codée par le cpDNA (PEP pour "plastid encoded polymerase"), transcrit les gènes dont les promoteurs sont de type bactérien (Hess et Borner, 1999). La probabilité est donc plus grande pour que TRCF puisse avoir des interactions avec cette polymérase et que son activité soit liée à cette protéine. De plus, il a été montré que dans les bactéries TRCF a aussi un rôle dans la résolution des conflits transcription/réplication (Trautinger *et al.*, 2005).

Dans le contexte du génome chloroplastique circulaire dont la transcription est très active, TRCF pourrait permettre la synchronisation entre réplication et transcription.

Il reste encore beaucoup à découvrir sur cette protéine et ses fonctions dans les organelles de plantes. Ce projet a été repris par un doctorant de l'équipe.

Conclusion

L'étude de RECG1 nous a apporté des informations importantes sur la ségrégation de sous-génomes qui peuvent coexister dans la mitochondrie, et nous a permis de formuler un modèle des mécanismes de ségrégation qui pourraient être en œuvre. De plus, le mutant *recG1* est un bon outil pour des futurs travaux. L'étude de l'effet d'autres protéines mitochondriales sur la transmission de l'épisome de *recG1-2* pourrait apporter des informations complémentaires sur leurs rôles dans la réplication et la ségrégation du génome. RECG1 est une pièce supplémentaire de la machinerie permettant le maintien et l'évolution du génome, et elle nous permettra de découvrir le rôle d'autres facteurs dans ces mécanismes.

L'identification de TRCF est aussi une première dans l'étude des organelles. Il reste cependant beaucoup de travail à effectuer pour découvrir si le NER est vraiment présent dans les mitochondries et les chloroplastes. L'étude de cette protéine apportera dans tous les cas de précieuses informations sur l'évolution de TRCF, et de son adaptation à de nouveaux partenaires protéiques.

Matériel et Méthodes

1. Matériel

1.1. Matériel bactérien

La souche DH5α, dérivant *d'Escherichia coli* K12 (Yanisch-Perron *et al.*, 1985) a été utilisée pour tous les clonages. Le taux de recombinaison réduit de cette souche permet un bon maintien de l'intégrité des plasmides transfectés dans les bactéries. De plus, le gène codant pour la β-galactosidase est désactivé par troncature de la partie N-terminale. La transfection par un plasmide portant la partie manquante rétablit l'activité de la β-galactosidase et permet une sélection blanc-bleu des bactéries transfectées.

DH5α : F^- *end*A1 *gln*V44 *thi*-1 *rec*A1 *rel*A1 *gyr*A96 *deo*R *nup*G Φ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169, *hsd*R17(r_K^- m_K^+), λ^-

La souche Rosetta(DE3)pLysS (Novagen), dérivant de la souche BL21 d'*E. coli*, a été utilisée pour la surexpression de protéines recombinantes. Cette souche permet d'induire l'expression de la protéine recombinante tout en évitant les fuites d'expression. La transcription est effectuée par l'ARN polymérase du phage T7 dont le gène, porté par l'ADN génomique de la bactérie, est sous le contrôle d'un promoteur inductible à l'IPTG. Le plasmide pLysS présent dans la souche porte le gène de résistance au chloramphénicol et code pour le lysozyme du phage T7 qui supprime l'expression basale de la T7 ARN polymérase en l'absence d'induction. Ce plasmide porte également des gènes de tRNA (ARN de transfert), correspondant à des codons rarement utilisés chez *E. coli*, ce qui permet d'optimiser l'expression des protéines eucaryotiques.

Rosetta(DE3)pLysS : F^- *omp*T *hsd*S_B(R_B^- m_B^-) *gal dcm* λ (DE3 [*lac*I *lac*UV5-T7 gene 1 ind1 sam7 nin5]) pLysSRARE (Cam^R)

La souche JM103, dérivant d'*E. coli,* et ses mutants a été utilisée lors des tests de complémentation. Les souches mutantes pour RecG ou RuvA, respectivement JM103 recG::kan (**Δ**recG) et JM103 ruvA::Tc (**Δ**ruvA) provenaient de la collection de Marc Bichara et ont été obtenues par insertion du gène de résistance à la kanamycine et à la tétracycline, respectivement.

JM103 : *end*A1 *gln*V44 *sbc*BC *rps*L *thi*-1 Δ(*lac-pro*AB) F'[*tra*D36 *pro*AB^+ *lac*I^q *lac*ZΔM15]

La souche bactérienne GV3101 d'*Agrobacterium tumefaciens* a été utilisée pour transformer les plantes. Cette souche est résistante à la rifampicine et possède le plasmide Ti pMP90RK qui porte un gène de résistance à la gentamycine et les gènes *vir* permettant le passage de l'ADN transféré (T-DNA pour "Transferred DNA") dans la cellule végétale.

1.2. Matériel végétal

Les mutants d'insertion T-DNA d'Arabidopsis (*Arabidopsis thaliana*) ont été cultivés à 21°C la journée et 18°C la nuit, avec un cycle diurne de 16 h. Le mutant *recG1-1* (WiscDsLox485488E5), dérivé de l'écotype Columbia 0 (Col-0), provenait de la collection de l'Université du Wisconsin-Madison, distribuée par le Nottingham Arabidopsis Stock Centre (NASC). Le mutant *recG1-2* (FLAG_630A09), dérivé de l'écotype Wassilewskija (Ws), provenait de la collection de l'INRA de Versailles (*Arabidopsis thaliana* Stock Centre, Institut Jean-Pierre Bourgin).

Le mutant de l'ADN polymérase POL1B (At3g20540) utilisé était le mutant *pol1B-2* (GK-291F05) de la collection GABI-Kat, distribuée par le NASC.

Les mutants TRCF dérivaient également de deux écotypes différents, (Ws) pour le mutant *trcf-2* (FLAG_411E05) et Col-0 pour les mutants *trcf-1* (SAIL_682_C02) et *trcf-3* (WiscDsLox485-488J1). Il s'agissait de mutants d'insertion T-DNA issues respectivement des collections de la "Syngenta Arabidopsis Insertion Library" et de l'Université du Wisconsin-Madison, et distribuées par le NASC.

Pour les tests *in vitro*, les graines ont été stérilisées (II.6.7.) puis déposées sur du milieu de culture solide. Après vernalisation pendant 48 h à 4°C, les cultures ont été développées dans une chambre à 21°C avec un cycle diurne de 16 h. Les plantes destinées à l'extraction de mitochondries ont été récoltées à 15 jours de culture.

Une culture de cellules d'Arabidopsis en suspension a été utilisée pour en extraire des mitochondries dans le cadre des tests de recombinaison *in organello*. Les suspensions ont été cultivées à 23°C, en lumière continue, sous une agitation de 130 rpm, dans le milieu de culture liquide. Chaque semaine, 10 mL de cellules ont été repiqués dans 100 mL de milieu frais. Pour extraire des mitochondries, des cultures de 5 jours ont été utilisées.

Milieu de culture solide : Micro- et macro-éléments M0255 Duchefa, saccharose 1% (p/v), pH 5,7 (ajusté avec KOH), agar 8 g/L.

Milieu de culture liquide : Murashige and Skoog Basal medium (Duchefa M0256 including Nitch vitamins) 4,41 g/L, saccharose 3% (p/v), acide α-Naphtylèneacétique (NAA) 5,4 µM, kinétine 0,23 µM, pH 5,6 (ajusté avec KOH).

Des feuilles de *Nicotiana benthamiana* ont été utilisées pour les tests d'expression transitoire des protéines de fusion GFP.

1.3. Vecteurs plasmidiques

pCK-GFP3

Le vecteur pCK-GFP3 (Menand *et al.*, 1998) possède l'origine de réplication colE1 et le gène de la ß-lactamase qui confère la résistance à l'ampicilline. Il est utilisé pour l'expression transitoire de séquences d'adressage fusionnées à la eGFP ("enhanced Green Fluorescent Protein") dans les expériences de localisation subcellulaire. La séquence de l'eGFP porte des mutations (L64F et S65T) permettant une meilleure fluorescence de la protéine (Youvan et Michel-Beyerle, 1996). Le vecteur pCK-GFP3 possède le promoteur 35S, deux séquences "enhancer" identiques, et le terminateur 35S du CaMV (*Cauliflower mosaic virus*), encadrant la séquence de la eGFP et permettant sa surexpression dans les plantes. Le site de clonage pour l'introduction d'une séquence d'intérêt se situe en phase, en 5' du gène de la eGFP.

pCKmRFP

Le vecteur pCKmRFP (Vermel *et al.*, 2002) est utilisé pour identifier les mitochondries dans les expériences de localisation subcellulaire. Il contient le gène DsRed1 qui code pour la RFP ("Red Fluorescent Protein"), fusionné à la séquence d'adressage aux mitochondries de la protéine COXIV de la levure. Ce plasmide est co-transfecté avec le pCK-GFP3 afin de localiser les mitochondries grâce à la fluorescence de la RFP.

pUCAP35S

Le vecteur pUCAP35S (van Engelen *et al.*, 1995) est un vecteur dérivé d'un pUC dans lequel les sites de restriction rares AscI et PacI ont été rajoutés aux extrémités de la cassette de clonage. Il est utilisé pour le clonage de séquences codant pour des protéines de fusion GFP qui seront ensuite transférées dans le vecteur binaire pBinPlus contenant ces deux mêmes sites de restriction. La cassette de clonage du vecteur pUCAP35S contient, comme pour le pCK-GFP3, le promoteur 35S, une séquence "enhancer" en deux exemplaires, le gène de la eGFP et le terminateur 35S du CaMV, ce qui permet une surexpression de la protéine de

fusion dans la plante, et une meilleure visualisation de la protéine de fusion GFP. Contrairement aux vecteurs binaires, pUCAP35S est de petite taille et se réplique en grand nombre de copies dans les bactéries grâce à son origine de réplication similaire à celle d'un pUC. L'on peut donc obtenir une grande quantité de plasmide pour le séquençage de la construction et l'excision de l'ensemble de la cassette d'expression en vue du transfert dans pBinPlus. Il porte le gène de résistance à l'ampicilline.

pBinPlus

Le vecteur pBinPlus (van Engelen *et al.*, 1995), dérivé du pBin19 (Bevan, 1984) est utilisé pour l'agrotransformation d'Arabidopsis. Un gène de résistance à la kanamycine sous le contrôle d'un promoteur bactérien permet la sélection de colonies transformées. Ce plasmide contient les séquences frontières du T-DNA, LB ("Left Border") et RB ("Right Border") qui permettent le transfert de séquences dans le génome nucléaire de la plante. Entre les séquences LB et RB, l'on trouve également les deux sites de restriction AscI et PacI qui sont eux-mêmes localisés au milieu du gène LacZ codant pour la ß-galactosidase. L'insertion d'une séquence d'intérêt entre ces deux sites de restriction coupe donc le gène LacZ et permet une sélection supplémentaire blanc/bleu des colonies positives lors du clonage. Un deuxième gène de résistance à la kanamycine sous contrôle du promoteur NOS (nopaline synthase) permet de sélectionner les plantes transformées.

pACYC184*Lac*Z

Le vecteur pACYC184*Lac*Z (Miller-Messmer *et al.*, 2012) est utilisé pour les expériences de complémentation fonctionnelle dans les bactéries mutantes. Il a été construit au laboratoire à partir de pACYC184 (Chang et Cohen, 1978) auquel a été rajouté le gène *lac*Z codant pour la β-galactosidase, à l'intérieur duquel se trouve une cassette de clonage multi-sites, permettant ainsi une sélection blanc/bleu des colonies transformées. Ce plasmide contient l'origine de réplication p15A, il est donc répliqué en faible nombre de copies et permet une expression de protéine plus proche des conditions physiologiques des bactéries. Il possède un gène de résistance au chloramphénicol, et sa cassette de clonage est sous le contrôle d'un promoteur bactérien inductible à l'IPTG, permettant ainsi de déclencher au moment voulu l'expression de la construction insérée.

pUC18

Le plasmide pUC18 (NEB) (Bevan, 1984) est habituellement utilisé pour les clonages de routine. Il contient une cassette de clonage multi-sites à l'intérieur du gène *lac*Z pour une sélection blanc/bleu, ainsi qu'un gène de résistance à l'ampicilline. Ce plasmide a été utilisé tel quel dans les expériences de complémentation bactérienne pour établir le rôle de RECG1 dans le contrôle de la réplication de molécules circulaires.

pBAD/Thio

Le vecteur pBAD/Thio (Invitrogen) contient le promoteur pBAD de l'operon bactérien *ara*BAD qui regroupe des gènes servant au transport et au métabolisme de l'arabinose (Guzman *et al.*, 1995). Il permet une expression de protéines hautement maitrisée dans *E. coli* et est habituellement utilisé pour l'expression de protéines toxiques ou pour en améliorer la solubilité. Ici, il a été utilisé comme le pUC18 comme témoin de la réplication dans les expériences de complémentation, et comme celui-ci il contient un gène de résistance à l'ampicilline.

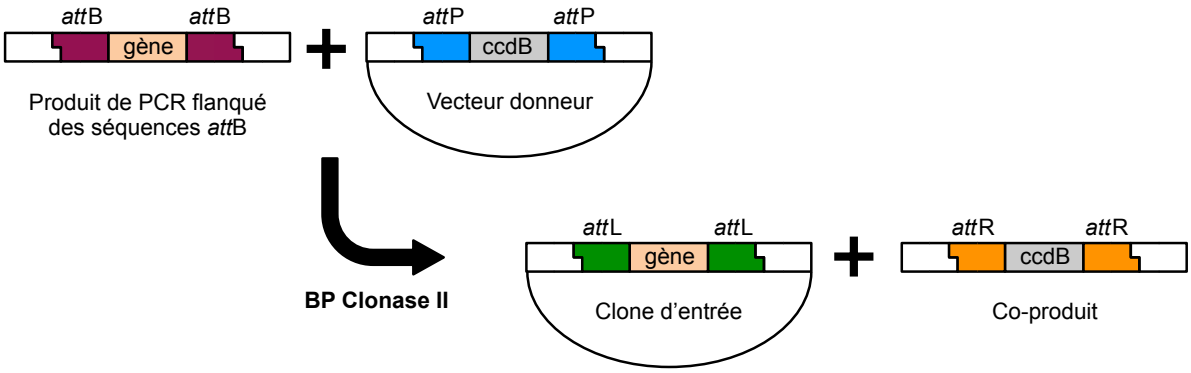
pGEM-T easy

Pour le clonage direct de produits PCR, en particulier en vue du séquençage, le vecteur pGEM-T easy (Promega) a été utilisé. C'est un vecteur commercial qui est vendu linéarisé avec une thymidine sortante à chaque extrémité 3'. Cela permet l'insertion de produits de PCR que les Taq polymérases synthétisent avec des adénosines sortantes à chaque extrémité 5'. Après amplification avec d'autres polymérases thermostables, qui ont généralement une activité de vérification ("proofreading") efficace, il est possible d'ajouter un A en 5' de chaque brin du fragment à insérer grâce à une post-incubation avec une Taq polymérase classique (Matériel et Méthodes 2.4.2.2.). Le vecteur pGEM-T easy est à haut nombre de copies. Il contient les promoteurs de la T7 et de la SP6 ARN polymérase flanquant la cassette de clonage multi-sites au milieu du gène de la β-galactosidase, permettant une sélection blanc/bleu des colonies transformées. Ce plasmide porte un gène de résistance à l'ampicilline.

pBluescript II SK+

Comme le pGEM-T, le pBluescript II SK+ permet une sélection blanc/bleu des colonies et possède le gène de résistance à l'ampicilline. C'est dans ce vecteur que nous avons reçu les constructions RECG1 et TRCF1 adaptées au code génétique d'*E. coli*.

A



B

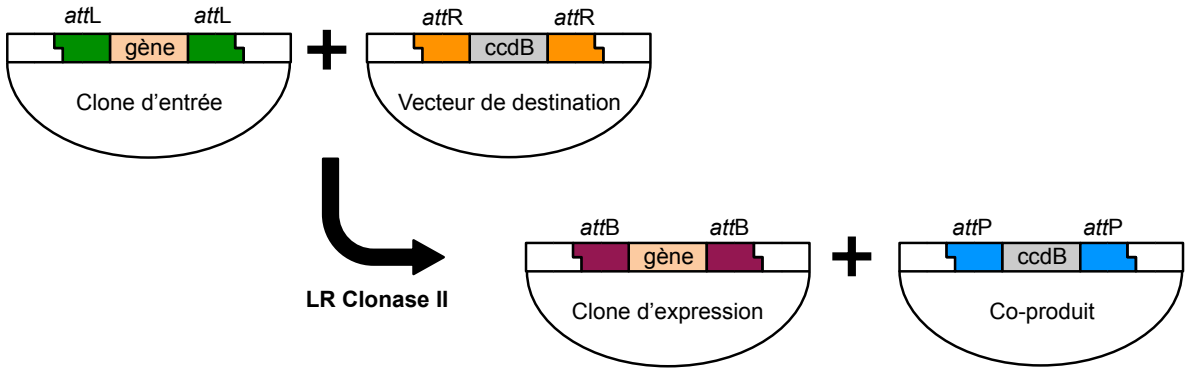


Figure 45. Clonage par le système Gateway.
A. Réaction de recombinaison BP permettant l'entrée de l'insert dans un vecteur donneur
B. Réaction de recombinaison LR permettant l'entrée de l'insert dans un vecteur de destination.

pDONRZeo

Pour la plupart des clonages, les vecteurs de la technologie Gateway (Invitrogen) (Hartley *et al.*, 2000) ont été utilisés. Cette méthode de clonage est basée sur le système de recombinaison du bactériophage lambda qui facilite son intégration à *E. coli*. Le système requiert d'une part, des séquences de recombinaison appelées "*att*" et d'autre part, un mélange de facteurs de recombinaison nommé "clonase". Cette technologie permet de générer un clone d'entrée possédant la séquence d'intérêt dans un vecteur donneur, et de transférer cette séquence dans n'importe quel vecteur de destination à l'aide d'une simple réaction de recombinaison (figure 45). Cela est utile par exemple pour exprimer une protéine dans différents systèmes et avec différentes étiquettes. D'une manière générale, les plasmides Gateway sont composés d'un gène de résistance à un antibiotique pour la sélection dans les bactéries, et d'une cassette de clonage composée des deux séquences *att*P (vecteur donneur) ou *att*R (vecteur de destination) flanquant un gène de résistance au chloramphénicol et le gène ccdB. La protéine associée au gène ccdB interfère avec l'ADN gyrase bactérienne (Bernard et Couturier, 1992), et permet ainsi d'éliminer toute bactérie transformée avec un plasmide n'ayant pas recombiné. Le pDONRZeo est un vecteur donneur qui possède un gène de résistance à la zéocine. Il possède les séquences *att*P et la recombinaison a lieu entre le fragment d'intérêt, flanqué des deux séquences *att*B, en présence du mélange d'enzymes BP clonase II (Invitrogen).

pMDC162

Le vecteur pMDC162 (Curtis et Grossniklaus, 2003) est un vecteur binaire de destination Gateway que nous avons utilisé pour les constructions associant un promoteur à étudier au gène de la ß-glucuronidase (GUS). Ici les séquences *att*R du vecteur se recombinent avec les séquences *att*L du plasmide donneur pDONRZeo recombinant en présence du mélange d'enzymes LR clonase, pour intégrer notre séquence d'intérêt, ici le promoteur putatif. Cette cassette de clonage par recombinaison est en 5' et en phase avec le gène GUS. Le plasmide confère une résistance à la kanamycine aux bactéries transformées. Les séquences frontières RB et LB du T-DNA permettant le transfert de l'information dans la plante via *Agrobacterium tumefaciens,* encadrent la cassette de clonage et le gène de résistance à l'hygromycine qui permettra la sélection des plantes transformées.

pGWB40

Le vecteur pGWB40 (Nakagawa *et al.*, 2007) a été utilisé pour la complémentation dans les plantes. Ce vecteur binaire de destination Gateway ne possède pas de promoteur en amont de la cassette de clonage, ce qui permet d'insérer le promoteur endogène correspondant au gène étudié. Il confère une résistance à la kanamycine aux bactéries transformées, et une résistance à l'hygromycine et à la kanamycine aux plantes transformées.

p0GWA

Le vecteur p0GWA (Busso *et al.*, 2005) est un vecteur de destination Gateway qui a été utilisé pour la production de protéines recombinantes dans *E. coli* afin de générer des anticorps. Ce vecteur permet d'ajouter une étiquette composée de 6 histidines (6His-tag) à l'extrémité C-terminale de la protéine recombinante. Ici la cassette de clonage contenant les séquences *att*R est flanquée des séquences promotrices du phage T7, permettant la transcription de l'insert par la T7 ARN polymérase. Ce promoteur est sous le contrôle de l'opéron lactose, inductible à l'IPTG. Le vecteur contient également un gène de résistance à l'ampicilline.

1.4. Gènes de synthèse

Pour plusieurs expériences, notamment la complémentation bactérienne, des gènes de synthèse fabriqués par Genecust ont été utilisés. Les séquences de l'ADN complémentaire (cDNA) des gènes RECG1 et TRCF, sans les séquences d'adressage, ont été adaptées à la fréquence d'utilisation des codons chez *E. coli* et clonées dans pBluescript II SK+.

1.5. Oligonucléotides

La liste des amorces utilisées pour les diverses PCR et qPCR est en annexe et en données supplémentaires dans Wallet *et al.* 2015.

1.6. Outils informatiques

Les comparaisons et alignements de séquences, ainsi que les assemblages informatiques des constructions et les analyses de restriction, ont été réalisés à l'aide du logiciel MacVector (MacVector, Inc.). Pour la phylogénie, les alignements de séquences ont été réalisés avec

ClustalW. Pour modeler la structure de RECG1, un alignement a d'abord été réalisé avec la protéine RecG cristallisée de *Thermotoga maritima* grâce à ClustalX.

Les principaux sites utilisés pour les recherches dans les bases de données et alignements de séquences sont :

NCBI : http://www.ncbi.nlm.nih.gov/

Base de données Arabidopsis : http://www.arabidopsis.org/

T-DNA express : http://signal.salk.edu/cgi-bin/tdnaexpress

Les logiciels de prédiction d'adressage aux organelles :

Predotar : https://urgi.versailles.inra.fr/predotar/predotar.html (Small *et al.*, 2004)

TargetP : http://www.cbs.dtu.dk/services/TargetP/ (Emanuelsson *et al.*, 2000)

SUBA3 : http://suba.plantenergy.uwa.edu.au/ (Tanz *et al.*, 2013)

Le logiciel pour créer les arbres phylogénétiques :

PhyMLv3 : http://www.phylogeny.fr/ (Dereeper *et al.*, 2008)

Le logiciel de modélisation des protéines (manipulé par Marc Bergdoll) :

Modeler : https://salilab.org/modeller/ (Sali et Blundell, 1993)

2. Méthodes

2.1. Bactéries

2.1.1. Préparation de bactéries compétentes

2.1.1.1. Electro-compétentes

Deux mL d'une préculture de la souche adéquate d'*E. coli* (DH5α ou TOP10) ayant poussé toute la nuit à 37°C sont utilisés pour inoculer 750 mL de milieu LB additionné de l'antibiotique correspondant au gène de résistance de la souche (acide nalidixique pour DH5α, streptomycine pour TOP10). La culture est incubée à 37°C sous agitation jusqu'à une A_{600}=0,4, puis placée dans la glace pendant 20 min, et centrifugée à 2700 *g* pendant 10 minutes à 4°C. Après deux lavages à l'eau froide stérile, le culot est lavé dans 50 mL de glycérol 10% (v/v) froid. Après une centrifugation de 12 min à 3000 *g*, le culot est repris dans du glycérol 10% de façon à ce que la suspension ait une A_{600} de 200. Cette préparation est

répartie en aliquotes de 50 µL qui sont congelées dans de l'azote liquide et conservées à -80°C.

Pour la préparation de cellules d'*Agrobacterium* électro-compétentes, la préculture pousse pendant 24 h à 28°C avec de la gentamycine à 15 µg/mL et de la rifampicine à 150 µg/mL. Soixante-dix µL de cette préculture sont utilisés pour inoculer 1 L de milieu LB avec antibiotiques, et la culture est incubée à 28°C sous agitation jusqu'à atteindre une A_{600}=0,4. Le protocole est ensuite le même que pour la préparation d'*E. coli* compétentes.

Milieu LB : bacto-tryptone 1% (p/v) ; extrait de levure 0,5% (p/v) ; NaCl 0,5% (p/v), pH ajusté à 7,0 (ajouter de l'agar 1% (p/v) pour obtenir du milieu LB solide).

2.1.1.2. Chimio-compétentes

Un mL d'une préculture de la souche adéquate d'*E. coli* est utilisé pour inoculer 500 mL de milieu LB additionné de l'antibiotique correspondant et de $MgCl_2$ 0,01 mM. La culture est incubée à 18°C sous agitation pendant plus de 24 h, jusqu'à atteindre une $A_{600}\leq$0,4. La culture est stoppée par incubation 10 min dans de la glace puis les bactéries sont centrifugées à 2700 *g* pendant 10 minutes. Le culot est repris dans 80 mL de tampon TB puis incubé 10 min sur glace. Les bactéries sont centrifugées une seconde fois 10 min à 2700 *g*, puis reprises dans du tampon TB. 1,4 mL de DMSO est ajouté avant d'ajuster le volume à 20 mL avec le tampon TB. Les bactéries sont incubées 10 min sur glace puis elles sont réparties en aliquotes et conservées à -80°C.

Tampon TB : PIPES 10 mM ; $CaCl_2$ 15 mM ; KCl 250 mM ; ajuster le pH à 6,7 puis ajouter $MnCl_2$ 55 mM et filtrer la solution.

2.1.2. Tests de complémentation de RECG1 dans *E. coli*

Pour les tests de complémentation, les souches JM103 sauvage et mutantes (Δ*rec*G et Δ*ruv*A) ont été utilisées. Le gène RECG1 synthétisé par Genecust a été transféré dans le plasmide pACYC184LacZ, entre les sites de restriction AscI et PacI. Cette construction a été transfectée dans les mutants bactériens Δ*rec*G et Δ*ruv*A, tandis que le vecteur vide a été transfecté dans les deux mutants et dans la souche sauvage, en tant que contrôle.

2.1.2.1. Tests de sensibilité aux UV

Pour les tests de sensibilité aux UV, 25 µL de préculture de bactéries ont été inoculés dans 5 mL de milieu LB supplémenté avec du glucose 1% (p/v) et du chloramphénicol 10 µg/mL. L'induction de l'expression de RECG1 a été déclenchée à A_{650}=0,2 par l'ajout de 0,25 mM

	Souches bactériennes				
Dose d'UV (Joules/m²)	**WT**	**Δ*recG***	**Δ*recG* *RECG1***	**Δ*ruvA***	**Δ*ruvA* *RECG1***
0	1E-04	1E-04	1E-04	1E-04	1E-04
5	1E-04	1E-04	1E-04	1E-03	1E-03
10	1E-03	1E-03	1E-03	1E-02	1E-02
20	1E-02	1E-02	1E-02	1E+00	1E+00

Figure 46. Dilutions des souches bactériennes utilisées lors de la complémentation, en fonction des doses d'UV appliquées.

d'IPTG et la culture a été incubée jusqu'à une A_{650}=0,4. Différentes dilutions des suspensions bactériennes ont été préparées, selon la dose d'UV appliquée, afin de faciliter le comptage des bactéries sur boite (figure 46). Cinquante µL de ces suspensions ont été étalés grâce à des billes de verre, sur boites de LB en triplicat, puis les boites ont été exposées à différentes doses d'UV-C (254 nm) sous le contrôle d'un radiomètre. Les bactéries ont ensuite été incubées 20 h à 37°C au noir, pour éviter que les mécanismes de réparation UV-dépendants ne se mettent en place. Le nombre de colonies sur chaque boite pour chaque dose d'UV a ensuite été compté et comparé au contrôle sans UV.

2.1.2.2. Tests de réplication

Pour les tests de réplication pathologique de l'ADN génomique, les souches JM103 sauvage et *Δrec*G ont été utilisées telles quelles. Pour les tests de réplication de plasmides exogènes, les souches ont été transfectées avec le plasmide pBad/thio (Invitrogen) ou pUC18 (NEB). Les cultures ont été incubées jusqu'à une A_{650}=1,2, puis elles ont été diluées 400 fois dans du milieu contenant du chloramphénicol à 10 µg/mL ainsi que 0,25 mM d'IPTG pour induire l'expression de la protéine RECG1. L'incubation sous agitation a été maintenue à 37°C, et des échantillons ont été prélevés dans la phase exponentielle à A_{650}=0,4 et à A_{650}=1,0. Un échantillon de la phase stationnaire a été prélevé plusieurs heures après que la A_{650} soit stable. L'ADN de ces différents échantillons a été extrait grâce au kit Gene Elute Bacteria (Sigma-Aldrich). Pour les tests sur la réplication du génome bactérien, une quantification relative des séquences *Ori*C, *Ter*A, *Ter*B, *Ydc*R et *Ydc*S a été effectuée par qPCR. Pour les tests sur la réplication de plasmides exogènes, la quantification par qPCR a été effectuée sur le gène de résistance à l'ampicilline présent dans pBAD/Thio et pUC18.

2.2. Acides nucléiques

2.2.1. Quantification

Les acides nucléiques ont été quantifiés par absorbance à 260 nm grâce à un spectrophotomètre NanoDrop (ThermoScientific). Pour une préparation d'ADN de bonne qualité, les rapports A_{260}/A_{280} et A_{260}/A_{230} représentatifs de la contamination en protéines et en impuretés (EDTA, carbohydrates, composées phénoliques…), sont respectivement supérieurs à 1,7 et à 2. Pour l'ARN pur le rapport A_{260}/A_{280} est proche de 2.

2.2.2. Analyse par électrophorèse

L'électrophorèse consiste en la migration de molécules (ADN, ARN, protéines) sous la contrainte d'un courant électrique, au sein d'une matrice (agarose, polyacrylamide) qui par la taille de sa maille permet de séparer les molécules selon leur taille. Les analyses d'ADN sont effectuées dans des gels d'agarose (0,7% à 1,5% (p/v) dans du tampon TAE) contenant du Bromure d'Ethidium (BET) à 0,5 µg/mL. Avant d'être chargés sur le gel, les échantillons sont préparés en ajoutant 1/5ème de volume de Tampon de charge 6x. La migration est réalisée dans du tampon TAE à 130 V. Le BET s'intercale entre les bases et permet la visualisation de l'ADN sur gel sous lumière UV. La taille des fragments d'ADN est estimée par comparaison avec une échelle d'ADN marqueur.

Gel d'agarose à x% : agarose x% (p/v) ; TAE 0,5x.

Tampon TAE : Tris 40 mM ; acide acétique 20 mM ; EDTA 1 mM.

Tampon de charge 6x : Tris-HCl 100 mM pH 7,6 ; Bleu de bromophénol 0,3% (p/v) ; glycérol 60% (v/v) ; EDTA 60 mM.

2.3. Méthodes relatives à l'ARN

2.3.1. Extraction d'ARN total

Lors de manipulations d'ARN, un soin particulier est apporté dans le choix d'un matériel propre et stérile, de solutions filtrées et autoclavées, dans le but d'éviter les contaminations par des RNases. Les ARN sont extraits de fleurs, de feuilles ou de plantules entières. Le matériel est tout d'abord congelé dans l'azote liquide et broyé à froid avec un broyeur TissueLyserII (Quiagen). Un mL de Trizol (Life Technologies) est ensuite ajouté et le mélange est vortexé jusqu'à décongélation complète. Puis 200 µL de chloroforme sont ajoutés avant 15 sec de votex et 15 min de centrifugation à 16000 *g*. A ce stade, le mélange est séparé en trois phases : la phase inférieure phénolique, l'interphase blanche et compacte contenant les protéines, et la phase supérieure aqueuse. La phase aqueuse contenant les ARN est prélevée en prenant grand soin de ne pas toucher l'interphase avec le cône. Afin de précipiter les ARN, 0,8% (v/v) d'isopropanol est ajouté, le tout est mélangé doucement et incubé 10 min à température ambiante. Après une centrifugation de 5 min à 16000 *g* et un lavage du culot à l'éthanol 70% (v/v), les ARN sont séchés et repris dans 60 µL d'eau Milli-Q stérile.

2.3.2. Traitement DNase

Afin d'éliminer les contaminations d'ADN, une phase de digestion à la DNase I est réalisée avant la réaction de transcription inverse. Cinq µg d'ARN total sont incubés dans 100 µL de milieu réactionnel contenant 10 µL de tampon DNase 10X (Thermo Scientific) et 5 µL de DNase sans RNase (Thermo Scientific). Le mélange est incubé pendant 45 min à 37°C. A la fin de la réaction, les ARN sont purifiés par une extraction au phénol/chloroforme, afin de se débarrasser de la DNase. Un volume de phénol/chloroforme (1:1) est ajouté à la réaction puis le tout est vortexé 15 sec et centrifugé 10 min à 16000 *g*. La phase aqueuse est récoltée et les ARN sont précipités avec 10% (v/v) d'acétate de sodium 3 M pH 5 et 2,5 volumes d'éthanol absolu. Après 1 h d'incubation à -20°C, les ARN sont sédimentés et lavés dans de l'éthanol 70% (v/v), puis séchés et repris dans de l'eau.

Tampon DNase 10X (Thermo Scientific) : Tris-HCl 100 mM pH 7,5 ; $MgCl_2$ 25 mM ; $CaCl_2$ 1 mM.

2.3.3. Réaction de transcription inverse (RT)

Pour quantifier les ARN, des brins d'ADN complémentaire (cDNA) sont d'abord synthétisés grâce à une ADN polymérase ARN-dépendante. Un premier mélange est réalisé avec :

- 1 µL d'oligo-dT 50 µM
- 1 µL d'hexamères aléatoires 200 ng/µL
- 1 µL de dNTP 10 mM
- entre 2 et 5 µg d'ARN traité à la DNase

Ce mélange est incubé 5 min à 65°C puis 5 min dans la glace. Les éléments suivants sont ensuite ajoutés pour atteindre un volume final de 20 µL :

- 4 µL de tampon First-strand 5x (Invitrogen)
- 1 µL de DTT 0,1 M
- 1 µL de RNase OUT (Invitrogen), un inhibiteur de RNase
- 1 µL de transcriptase inverse Superscript III (Invitrogen)

Le mélange est incubé 5 min à 25°C, puis 1 h à 50°C. Enfin, l'enzyme est inactivée 15 min à 70°C. Un µL de cDNA obtenu est utilisé pour l'amplification par PCR ou qPCR.

Tampon First-strand 5x (Invitrogen) : Tris-HCl 250 mM pH 8,3 ; KCl 375 mM ; $MgCl_2$ 15 mM.

Type de cycle	ADN polymérase				Cycles
	Phusion et Phire		Taq		
	Température	Temps	Température	Temps	
Dénaturation initiale	98 °C	30 s	94 °C	5 min	1
Dénaturation	98 °C	5 s	94 °C	30 s	25-35
Appareillement	X °C	10 s	X °C	30 s	
Elongation	72 °C	15 s/kb	72 °C	1 min/kb	
Elongation finale	72 °C 4°C	5 min pause	72 °C 4°C	5 min pause	1

Figure 47. Protocole de PCR selon les polymérases utilisées.
X dépend du Tm des amorces et il est compris entre 58 °C et 65 °C.

2.4. Méthodes relatives à l'ADN

2.4.1. Amplification par PCR

La réaction de polymérisation en chaîne (PCR) consiste en l'amplification exponentielle d'une séquence d'ADN double brin située entre deux amorces spécifiques. La PCR se divise en plusieurs étapes répétées qui ont lieu à différentes températures et pendant différents temps. La première incubation de 1 min à 98°C permet de dénaturer la matrice et casser la paroi bactérienne dans le cas de PCR sur colonies. S'ensuit une série de cycles (entre 25 et 35) composés de trois étapes : la dénaturation de l'ADN (5 sec à 98°C), l'hybridation des amorces (5 sec entre 58 et 70°C selon le Tm des amorces) et enfin l'élongation par une polymérase thermorésistante (72°C, temps variable selon le type de polymérase et la taille du fragment). Une phase d'élongation finale (5 min à 72°C) permet de finir l'élongation de tous les brins.
Plusieurs types d'ADN polymérases ADN-dépendantes ont été utilisés. Pour la synthèse de produits à cloner, l'enzyme Phusion (Thermo Scientific) à haut degré de fidélité a été utilisée. Elle possède une activité correctrice exonucléase 3'-5' similaire à celle de l'enzyme *Pfu* de *Pyrococcus furiosus*. Pour les génotypages, malgré son prix l'enzyme Phire Hot Start II (Thermo Scientific) a souvent été utilisée, pour sa robustesse vis-à-vis d'inhibiteurs, son haut rendement et sa rapidité, ce qui permet de diminuer les temps d'incubation pour chaque cycle de PCR (figure 47).

2.4.2. Techniques de clonage

2.4.2.1. Digestion

Les digestions des fragments d'ADN avec différentes enzymes de restrictions ont été réalisées selon les instructions des fabriquants.

2.4.2.2. Ligation

2.4.2.2.1. ***Ligation d'un fragment dans un vecteur linéarisé***

La ligation d'un fragment dans un plasmide linéarisé est réalisée dans un volume de 20 µL pendant 16 h à 4°C, ou 2 h à température ambiante. Le milieu réactionnel comprend :
- 20 ng de vecteur linéarisé
- l'ADN à insérer dans un rapport molaire de 1/1 à 1/5 avec le vecteur
- 2 µL de tampon de ligation 10x (Thermo Fischer Scientific)
- 1 U de T4 ADN ligase (Thermo Fischer Scientific)

Tampon de ligation 10x (Thermo Fischer Scientific) : Tris-HCl 400 mM ; $MgCl_2$ 100 mM ; DTT 100 mM ; ATP 5 mM ; pH 7,8.

2.4.2.2.2. A-tailing et ligation avec pGEM-T easy

L'enzyme Phusion, qui génère des extrémités franches a généralement été utilisée pour l'amplification de fragments à cloner. Pour insérer un tel fragment de PCR dans le vecteur pGEM-T easy, du "A-tailing" est nécessaire. Cela consiste en l'ajout de nucléotides adénosine sortants à l'extrémité 5' des brins d'ADN, par une activité *Taq* ADN polymérase classique. Le milieu réactionnel est composé de :

- 500 ng de produit de PCR purifié
- 1/10 (v/v) tampon PCR 10x
- 1 U de *Taq* ADN polymérase
- 0,2 mM ATP

Le mélange est incubé à 72°C pendant 10 min. Le produit de cette réaction est ensuite utilisé dans la réaction de ligation avec le vecteur pGEM-T easy linéarisé. Le milieu est composé de :

- 50 ng de vecteur pGEM-T easy (Promega)
- 2 µL de la réaction de A-tailing contenant le fragment de PCR (environ 100 ng d'ADN)
- 1/2 (v/v) tampon de ligation rapide 2x (Promega)
- 3 U de T4 ADN ligase (Promega)

La réaction a lieu toute la nuit à 4°C.

Tampon de PCR 10x : Tris-HCl 20 mM pH 8,3 ; KCl 50 mM ; $MgCl_2$ 2 mM.

Tampon de ligation rapide 2x (Promega) : Tris-HCl 20 mM pH 7,8 ; $MgCl_2$ 20 mM ; DTT 20 mM ; ATP 2 mM ; PEG 8000 10% (p/v).

2.4.2.3. Purification de l'ADN

2.4.2.3.1. Sur colonne de silice

Pour les clonages (vecteurs pGEM-T ou pDONRZeo) il est nécessaire d'avoir un produit de PCR unique et purifié. Dans le cas où l'on n'observe qu'une seule bande sur gel, le produit est purifié avec le kit PCR Cleanup (Macherey-Nagel). L'ADN est fixé sur une colonne de silice et il est élué avec de l'eau ou du tampon d'élution après plusieurs lavages avec une solution éthanolique.

Tampon d'élution: Tris-HCl 5 mM pH 8 ; EDTA 0,1 mM.

2.4.2.3.2. ***Sur gel d'agarose***

Pour les fragments digérés, ou pour les réactions de PCR présentant des produits multiples, l'ADN est purifié sur gel d'agarose (0,6% à 1,2% (p/v) dans du tampon TAE). Le fragment d'intérêt est localisé sous lumière UV puis excisé du gel et purifié grâce au kit PCR and Gel Cleanup (Macherey et Nagel). L'agarose contenant l'ADN est fondu dans un tampon contenant du perchlorate de sodium. L'ADN est ensuite fixé sur colonne de silice et il est élué avec du tampon TE après plusieurs lavages avec une solution éthanolique.

Tampon TAE : Tris 40 mM ; acide acétique 20 mM ; EDTA 1 mM.

Tampon TE : Tris-HCl 10 mM pH 8 ; EDTA 1 mM.

2.4.2.4. Réaction BP et LR

Pour les clonages effectués avec la technologie Gateway, l'insertion des séquences d'intérêt dans les vecteurs est effectuée par recombinaison, comme expliqué en I.3. Pour la réaction "BP" permettant de créer des clones d'entrée, le milieu réactionnel de 5 µL est composé de :

- la séquence d'intérêt linéaire, flanquée des séquences *att*B (environ 50 fmol)
- le vecteur donneur, contenant les séquences *att*P, en quantité équimolaire avec le produit à cloner (50 fmol, soit environ 150 ng pour pDONRZeo)
- 1µL de mélange enzymatique BP clonase II (Invitrogen)

Le milieu réactionnel est incubé entre 1 h et 16 h à 25°C. Pour stopper la réaction, 2 µg de protéinase K sont ajoutés, et le tout est incubé 10 min à 37°C. Un µL de cette réaction est ensuite utilisé pour transformer une souche d'*E. coli* sensible au produit du gène ccdB, comme par exemple les bactéries DH5□.

Pour la réaction LR permettant de créer des clones d'expression, le milieu réactionnel de 10 µL est composé de :

- 50 à 150 ng d'un vecteur d'entrée pENTR ou d'un vecteur donneur pDONR contenant la séquence d'intérêt flanquée des séquences *att*L
- 150 ng d'un vecteur de destination contenant les séquences *att*R
- 2 µL de mélange enzymatique LR clonase II (Invitrogen)

Le milieu réactionnel est incubé entre 1 h et 18 h à 25°C, puis 2 µg de protéinase K sont ajoutés et le tout est incubé 10 min à 37°C. Un à 2 µL de cette réaction sont utilisés pour transformer *E. coli*.

2.4.3. Transformation de bactéries

2.4.3.1. Transformation par électroporation

Les bactéries électrocompétentes, *E. coli* et *A. tumefaciens*, sont transformées selon des protocoles similaires. Cinquante µL de bactéries compétentes sont mises en contact avec 50 à 200 ng d'ADN plasmidique (jamais plus d'1 µL de réaction de ligation) dans une cuve d'électroporation. Ce mélange subit une décharge électrique de 25 µF, sous 2,5 kV et 200 Ω. Le choc électrique déstabilise la paroi bactérienne et permet à l'ADN d'être intégré à la cellule. Juste après le choc, 1 mL de milieu LB est ajouté puis le tout incubé à 37°C (*E. coli*) ou 28°C (*A. tumefaciens*) pendant 1 h. Enfin, les bactéries sont étalées sur des boites de Pétri contenant du milieu LB additionné de l'antibiotique approprié puis incubées soit une nuit à 37°C soit deux jours à 28°C.

Milieu LB : bacto-tryptone 1% (p/v) ; extrait de levure 0,5% (p/v) ; NaCl 0,5% (p/v) ; pH ajusté à 7,0 (ajouter 1% (p/v) d'agar pour obtenir du milieu LB solide).

2.4.3.2. Transformation par choc thermique

Cinquante µL de bactéries *E. coli* thermo-compétentes sont décongelés sur glace, puis incubés 30 min sur glace avec 1 à 50 ng de plasmide (2 µL de la réaction de ligation). Ce mélange subit un choc thermique de 1 min à 42°C, ce qui permet l'entrée d'ADN dans les cellules, puis il est incubé dans la glace pendant 2 min. 250 µL de milieu LB sont ensuite ajoutés et le tout est incubé à 37°C pendant 1 h, le temps que les membranes se régénèrent et que la bactérie aie le temps d'exprimer le gène de résistance à l'antibiotique conféré par le plasmide. Le mélange est ensuite étalé sur boite de Pétri contenant du milieu LB additionné de l'antibiotique approprié. Les boîtes sont incubées pour la nuit à 37°C.

Pour la transformation d'*A. tumefaciens* thermo-compétentes, le choc thermique est effectué à 37°C. L'incubation permettant d'exprimer le gène de résistance, puis la croissance sur boîte se font à 28°C.

Milieu LB : bacto-tryptone 1% (p/v) ; extrait de levure 0,5% (p/v) ; NaCl 0,5% (p/v) ; pH ajusté à 7,0 (ajouter 1% (p/v) d'agar pour obtenir du milieu LB solide).

2.4.4. Préparation d'ADN plasmidique

Pour la préparation d'ADN plasmidique, le kit NucleoSpin Plasmid QuickPure (Macherey-Nagel) a été utilisé. Les bactéries sont lysées dans un tampon alcalin contenant du SDS, puis l'ADN génomique est co-précipité avec les protéines par l'addition d'acétate de

sodium qui forme un précipité avec le SDS. L'ADN plasmidique est ensuite fixé sur colonne de silice, lavé dans une solution éthanolique, puis élué dans du tampon TE.
Pour la purification de grandes quantités d'ADN plasmidique, le kit NucleoSpin Xtra Midi (Macherey-Nagel) a été utilisé. Il fonctionne sur un principe similaire.

Tampon TE : Tris-HCl 10 mM pH 8 ; EDTA 1 mM

2.4.5. Séquençage

Les réactions de séquençage ont été réalisées par M. Alioua de la plateforme de séquençage de l'IBMP. La technique est basée sur la méthode de Sanger (Sanger *et al.*, 1977) et utilise la PCR asymétrique. Les acides nucléiques sont ensuite séparés par électrophorèse capillaire sur un appareil "ABI PRISM® 3100 Genetic Analyzer" de Applied Biosystems (Perkin Elmer).

2.4.6. Extraction d'ADN total

Pour extraire l'ADN total de plantes, un protocole au cetyl-trimethylammonium bromide (CTAB) a été utilisé. Le matériel végétal est broyé congelé, dans le broyeur TissueLyserII (Quiagen) à l'aide de billes métalliques (diamètre 3 mm, acier inox). Puis 400 µL de tampon d'extraction CTAB préchauffé à 65°C sont ajoutés dans les tubes et le tout est incubé à 65°C pendant 1 h. L'ajout de 400 µL de chloroforme, une agitation sur vortex, et une centrifugation de 10 min à 16000 *g* permettent une séparation du mélange en deux phases, avec une interphase solide et blanchâtre contenant les protéines. La phase supérieure est transférée dans un nouveau tube et l'ADN est précipité par l'ajout de 10% (v/v) d'acétate de sodium 3 M pH 5 et de 0,8% (v/v) d'isopropanol. Après centrifugation et lavage à l'éthanol 70% (v/v), le culot est repris dans du tampon TE 0,5x. La qualité de l'ADN est contrôlée sur gel d'agarose et la quantification se fait au spectrophotomètre NanoDrop.

Tampon d'extraction CTAB : CTAB 2% (p/v) ; NaCl 1,4 M ; EDTA 20 mM ; Tris-HCl 100 mM pH 8. A ajouter juste avant utilisation: ß-mercaptoéthanol 0,2% (v/v) ; RNase A 0,1 mg/mL.

Tampon TE : Tris-HCl 10 mM pH 8 ; EDTA 1 mM.

Un protocole permet également d'extraire l'ADN total de la phase inférieure phénolique restante après extraction d'ARN au Trizol (II.3.1). Cela permet d'obtenir ARN et ADN d'une même préparation du même échantillon. Après avoir suivi le protocole d'extraction d'ARN, la phase supérieure aqueuse est totalement retirée, et l'interphase ainsi que la phase inférieure

sont conservées. Puis, 300 µL d'éthanol absolu sont ajoutés à ces deux phases pour précipiter l'ADN. Après une incubation de 30 min à température ambiante et une centrifugation de 15 min à 16000 *g*, le culot d'ADN est lavé 2 fois avec de l'éthanol à 70% (v/v) puis séché et repris dans 200 µL de SDS 1% (p/v). Une extraction est réalisée en ajoutant 200 µL d'un mélange phénol/chloroforme (1:1). Après avoir vortexé et centrifugé le mélange, la phase supérieure est transférée dans un nouveau tube, puis l'ADN est précipité avec 10% (v/v) d'acétate de sodium 3 M pH 5 et 2,5 volumes d'éthanol absolu.

2.4.7. Réaction de PCR quantitative

La PCR quantitative (qPCR) repose sur la mesure d'une libération de fluorescence à chacun des cycles de PCR. L'aspect quantitatif réside dans la détermination du Ct (« Cycle Threshold ») ou Cp ("crossing point") qui correspond au nombre de cycles minimum nécessaires pour que la fluorescence dépasse le seuil du bruit de fond. Les qPCR sont réalisées dans des microplaques optiques de 384 puits et sur un appareil LightCycler 480 II (Roche). Le volume réactionnel de 6 µL est composé de 1x 480 SYBR Green I Master (Roche), 1,5 µM d'amorce directe et d'amorce réverse et environ 2 ng d'ADN total par réaction. Chaque PCR sur un échantillon d'ADN ou de cDNA est effectuée en triplicata. Les réactions de qPCR sont réalisées selon les conditions suivantes : une phase de préchauffage de 10 min à 95°C, puis 40 cycles de 15 sec à 95°C, 15 sec à 58°C et 15 sec à 72°C. A la fin de ces cycles d'amplification, une courbe de fusion est réalisée en chauffant les produits de 55°C à 95°C, pour ainsi vérifier la nature unique des amplicons et la spécificité des amorces. Les résultats sont normalisés par rapport à des gènes de contrôle endogènes. L'efficacité de PCR pour chaque couple d'amorces est calculée à l'aide d'une courbe de dilution d'ADN, grâce au programme de l'appareil Roche. Cette efficacité est ensuite utilisée pour établir la quantité d'ADN relative à la référence.

2.4.8. Détection d'ADN par hybridation avec une sonde radioactive (Southern blot)

2.4.8.1. Synthèse de la sonde radioactive

Le principe du Southern blot repose sur le transfert sur membrane des molécules d'ADN séparées sur gel d'agarose, puis sur la détection d'une séquence d'ADN spécifique par hybridation avec une sonde marquée radioactivement.

Le marquage radioactif d'une sonde ADN consiste en l'incorporation enzymatique d'un nucléotide radioactif. Pour cela, l'enzyme Exo-Klenow du kit DECAprime II (Ambion inc),

est utilisée. Elle est composée du fragment Klenow de l'ADN polymérase I d'*E. coli* qui a perdu son activité exonucléase 5'-3', et dont une mutation inactive également l'activité exonucléase 3'-5'. L'ADN servant de sonde est synthétisé par PCR et purifié sur colonne de silice. Puis, sont mélangés :

- 25 ng de cet ADN dans 12,5 µL
- 2,5 µL de solution de décamères aléatoires 10x.

L'ensemble est incubé 5 min à 99°C afin de dénaturer l'ADN. Les oligonucléotides aléatoires vont servir d'amorces pour l'Exo-Klenow. Le milieu est ensuite plongé dans de l'azote liquide afin d'empêcher la renaturation de l'ADN avec lui-même, puis transféré sur de la glace. Les éléments suivants sont ensuite ajoutés :

- 5 µL de tampon de réaction 1x sans dCTP
- 25 µCi de [□-^{32}P]dCTP (3000 Ci/mmol)
- 1,5 µL d'eau
- 1 µL d'enzyme Exo-Klenow

La réaction de polymérisation est conduite à 37°C pendant 10 min et arrêtée par l'ajout de 20 mM d'EDTA. L'ADN marqué est séparé des nucléotides non-incorporés par filtration sur "spin-column" de Sephadex G-50 (GE Healthcare). Cette sonde est dénaturée 5 min à 99°C avant utilisation.

2.4.8.2. Transfert d'ADN sur membrane et hybridation

Des échantillons d'ADN, digéré ou pas avec une enzyme de restriction, sont fractionnés par électrophorèse sur gel d'agarose à 0,7%, (p/v) pendant 2 h à 100V. Le gel subit un traitement afin de permettre le transfert de l'ADN sur membrane. Il est incubé dans une solution de dépurination, pendant 10 min, ce qui permet une fragmentation et un meilleur transfert des molécules de grande taille. Après un rinçage à l'eau distillée, le gel est incubé 30 min dans une solution de dénaturation afin de séparer les brins d'ADN. Enfin, le gel est incubé 30 min dans un tampon de neutralisation avant un dernier rinçage, puis il est équilibré 10 min dans du SSC 10x. Le transfert sur membrane s'effectue par capillarité. Deux bandes de papier Whatman® 3MM sont disposées de sorte que leurs extrémités trempent dans le tampon de transfert. Trois feuilles de la taille du gel sont posées en leur centre. Le gel est déposé par dessus, et recouvert d'une membrane Hybond N+ (GE Healthcare) en évitant les bulles d'air. On dépose ensuite une pile de papier absorbant pour établir un flux de liquide par capillarité. Ce flux permet le transfert de l'ADN sur la membrane. La membrane est ensuite

imbibée de SSC 2x avant pontage de l'ADN par la lumière UV (254 nm, 1200 joules/m, Stratalinker 2400™, Stratagene).

La membrane est ensuite pré-hybridée plus de 30 min à 65°C dans du tampon de Church. La sonde d'ADN radioactif dénaturé est ensuite ajoutée à raison de $2x10^5$ cpm/mL. L'hybridation se déroule sur une nuit à 65°C. Le lendemain, le tampon radioactif est enlevé et la membrane lavée au moins 3 fois 10 min à 65°C dans un excès de tampon SSC 2x, 0,1% (p/v) SDS. La radioactivité est détectée au Phosphorimager (Fujifilm FLA7000) ou par autoradiographie.

Pour pouvoir réutiliser la membrane, la sonde est déshybridée dans du SDS 0,1% (p/v) bouillant. Cette étape est répétée une fois et l'élimination de la radioactivité de la membrane est vérifiée au Phosphorimager .

Solution de dépurination : HCl 0,25 M.

Solution de dénaturation : NaOH 0,5 M ; NaCl 1,5 M.

Tampon de neutralisation : Tris-HCl 0,5 M pH 7,5 ; NaCl 1,5 M.

SSC 10x : NaCl 1,5 M ; citrate de sodium ($C_6H_5Na_3O_7$) 150 mM ; pH 7.

Tampon de Church : BSA 1% (p/v) ; EDTA 1 mM pH 8 ; Tampon Na-phosphate 0,5 M pH 7,2 ; SDS 7% (p/v).

2.5. Méthodes relatives aux protéines

2.5.1. Extraction de protéines à partir de matériel végétal

Les plantules ou feuilles sont broyées congelées, manuellement, avec un piston en plastique adapté aux microtubes de 1,5 mL Puis, selon la quantité de matériel, 250 à 500 µL de tampon dénaturant sont ajoutés. Une centrifugation de quelques minutes permet d'enlever les débris de broyage.

Tampon dénaturant : Tris-HCl 80 mM pH 6,8 ; urée 4 M ; SDS 2% (p/v) ; glycérol 10% (v/v) ; DTT 0,1 M à ajouter juste avant utilisation.

2.5.2. Précipitation (purification et concentration)

Pour purification ou concentration, les protéines sont précipitées à partir des extraits réalisés précédemment. Pour un volume de solution de protéines, 4 volumes de méthanol froid, et 1 volume de chloroforme froid sont ajoutés. Le mélange est vortexé et incubé 15 min sur glace, puis 3 volumes d'eau sont ajoutés. Après 30 sec d'agitation sur vortex et une incubation de 5 min sur glace, le mélange est centrifugé 5 min à 18000 *g*. Trois phases se

forment alors, la phase supérieure est enlevée en laissant l'interphase qui contient les protéines. Quatre volumes de méthanol sont enfin ajoutés et l'ensemble est centrifugé pendant 5 min. Le culot est séché, et resuspendu dans du tampon dénaturant ou directement dans du tampon de Laemmli.

Tampon dénaturant : Tris-HCl 80 mM pH 6,8 ; urée 4 M ; SDS 2% (p/v) ; glycérol 10% (v/v) ; DTT 0,1 M à ajouter juste avant utilisation.

Tampon de Laemmli 2x: Tris-HCl 125 mM pH 6,8 ; glycérol 20% (v/v) ; SDS 4% (p/v) ; ß-mercaptoethanol 2% (v/v) ; bleu de bromophénol 0,02% (p/v).

2.5.3. Séparation des protéines en conditions dénaturantes : SDS-PAGE

Les protéines sont fractionnées en fonction de leur taille par électrophorèse sur gel de polyacrylamide en conditions dénaturantes en présence de SDS (SDS-PAGE). Les mailles de la matrice sont définies par le rapport polyacrylamide/N,N'méthylène bisacrylamide (37,5/1 ou 29/1 selon la résolution voulue et la taille des protéines d'intérêt). Le gel de concentration formant la partie supérieure est à une concentration en acrylamide de 5% (p/v) dans un tampon Tris-HCl 125 mM pH 6,8, SDS 0,1% (p/v). Ce gel permet de concentrer les protéines afin qu'elles entrent de façon uniforme dans le gel de séparation. Celui-ci est composé d'acrylamide d'une concentration allant entre 8 et 15% (p/v) dans un tampon Tris-HCl 375 mM pH 8,8, SDS 0,1% (p/v). Les échantillons de protéines sont complétés avec 1 volume de tampon Laemmli 2x et incubés 5 min à 96°C avant d'être déposés sur gel. La migration s'effectue dans du tampon Tris-Glycine-SDS à ampérage constant de 30 mA (pour des gels de 12x10x0,1 cm). Après migration, les protéines peuvent être visualisées sur le gel par coloration pendant 30 min au bleu de Coomassie. Le gel est ensuite lavé avec la solution de décoloration jusqu'à apparition des bandes protéiques. La taille des protéines est estimée par comparaison avec une échelle de protéines marqueurs.

Tampon Laemmli 2x : Tris-HCl 125 mM pH 6,8 ; glycérol 20% (p/v) ; SDS 4% (p/v) ; ß-mercaptoethanol 2% (v/v); bleu de bromophénol 0,02% (p/v).

Tampon Tris-Glycine-SDS : Tris 25 mM ; Glycine 192 mM, SDS 0,1% (p/v), pH 8,3.

Solution de coloration : Acide acétique 10% (v/v) ; méthanol 40% (v/v) ; Coomassie Brilliant Blue R250 0,25% (p/v).

Solution de décoloration : Acide acétique 10% (v/v), éthanol 20% (v/v).

2.5.4. Quantification

La quantité de protéines est mesurée par test de Bradford (Bradford, 1976). Cette méthode colorimétrique repose sur le changement d'absorbance du réactif de Bradford de 450 nm à 595 nm quand celui-ci se fixe à des protéines. Un réactif commercial (protein assay dye reagent concentrate de Bio-Rad) a été utilisé. L'essai est composé de 200 µL de la solution commerciale, d'un volume compris entre 1 et 100 µL de solution de protéines et du complément en eau pour atteindre 1 mL. Le blanc est réalisé avec le tampon dans lequel les protéines sont en solution. La concentration en protéines est estimée en mesurant l'absorbance à 595 nm en présence du réactif et en comparant la valeur obtenue à une courbe standard établie avec des concentrations connues de BSA.

2.5.5. Surproduction dans *E. coli*

Les protéines possédant une étiquette de 6 histidines peuvent être purifiées par chromatographie d'affinité sur des billes d'agarose recouvertes d'acide nitrilotriacétique chélaté à un ion Ni^{2+} par quatre liaisons (Ni-NTA, Quiagen). L'ion Ni^{2+} peut établir 2 liaisons supplémentaires avec le groupement imidazole de deux histidines voisines. Les billes d'agarose Ni-NTA placées dans une colonne retiennent donc fortement les protéines ayant l'étiquette 6xhistidine. La protéine est ensuite éluée par compétition avec une solution d'imidazole.

Pour la surproduction, une préculture d'une souche de bactéries Rosetta(DE3)pLysS transformée avec un plasmide p0GWBA contenant une séquence codant pour une protéine à exprimer est inoculée à 1% (v/v) dans 250 mL de milieu LB préchauffé à 37°C comprenant les antibiotiques adéquats. Quand la culture atteint une A_{600}=0,4, l'expression est induite par ajout de 0,3 mM d'IPTG. Après 2 h d'induction à 37°C, les bactéries sont centrifugées pendant 20 min à 6000 *g* puis lavées dans de l'eau avant d'être resuspendues dans 30 mL de tampon de lyse additionné d'une pointe de spatule de lysozyme, afin de fragiliser la paroi bactérienne. Les bactéries sont ensuite lysées par dépression grâce à une presse de French (100 bars). La fraction insoluble est sédimentée par centrifugation à 20000 *g* pendant 10 min. Les deux protéines recombinantes que nous avons exprimées étaient toutes deux insolubles et nous les avons donc purifiées en conditions dénaturantes. Pour cela, le culot contenant la protéine d'intérêt est resuspendu dans 15 mL de tampon de lyse additionné d'urée 7 M pour RECG1 ou de guanidium 6 M pour TRCF. Sept-cents µL de billes d'agarose Ni-NTA (Quiagen) sont alors ajoutés et le tout est incubé sous rotation pendant 2 à 16 h à 4°C. La suspension est ensuite coulée dans une colonne et la résine Ni-NTA est rincée avec 10 mL de

tampon de resuspension. La colonne est ensuite lavée avec 10 mL de tampon de lavage, puis l'élution s'effectue par le passage consécutif de 1 mL de tampon d'élution 1 et 1 mL de tampon d'élution 2. Les fractions contenant la protéine d'intérêt sont identifiées par analyse d'une aliquote sur gel de polyacrylamide. Elles sont ensuite réunies et les protéines sont concentrées par ultrafiltration (Amicon Ultra 4 mL centrifugal filters, Millipore), diluées et re-concentrées afin de réduire la concentration en urée et en imidazole. La concentration de protéine finale est dosée par la méthode de Bradford.

Tampon de lyse : KH_2PO_4 50 mM ; NaCl 400 mM ; KCl 100 mM ; glycérol 10% (v/v) ; Tween20 0,1% (p/v) ; imidazole 10 mM ; pH 8.

Tampon de resuspension : tampon de lyse ; urée 7 M.

Tampon de lavage : tampon de lyse ; urée 7 M ; imidazole 20 mM.

Tampon d'élution 1 : tampon de lyse ; urée 7 M ; imidazole 190 mM ; pH 5.

Tampon d'élution 2 : tampon de lyse ; urée 7 M ; imidazole 490 mM ; pH 5.

2.5.6. Transfert semi-sec sur membrane de PVDF

Afin de réaliser des immunodétections, les protéines fractionnées sur gel de polyacrylamide sont transférées sur membrane de PVDF microporeuse de 0,45 µm (Immobilon P, Millipore). La membrane qui est très hydrophobe est d'abord activée avec de l'éthanol 100%, puis rincée à l'eau et équilibrée dans du tampon de transfert semi-sec. Un "sandwich" est formé avec, de l'anode vers la cathode, deux feuilles de papier filtre Whatman® 3 MM, le gel, la membrane et deux autres feuilles de papier Whatman® 3 MM, le tout imbibé au préalable de tampon de transfert semi-sec. L'électrotransfert est réalisé dans une cuve de transfert semi-sec (VWR) sous un ampérage de 0,8 mA/cm^2 de gel pendant 45 min. Après transfert, la membrane peut être colorée avec une solution de bleu de Coomassie, puis décolorée pour visualiser les protéines. Pour une immunodétection avec des anticorps la membrane est totalement décolorée avec du méthanol 100%.

Tampon de transfert semi-sec : Tris 48 mM ; glycine 39 mM ; pH 8.

Solution de coloration : acide acétique 7% (v/v) ; méthanol 50% (v/v) ; Coomassie brillant blue R-250 0,1% (p/v).

Solution de décoloration : acide acétique 7% (v/v) ; méthanol 50% (v/v).

2.5.7. Immunodétection

Après transfert, la membrane de PVDF est incubée avec un anticorps primaire spécifique à la protéine d'intérêt. Une seconde hybridation avec un anticorps secondaire, couplé à la

peroxydase du raifort, permet la détection par chimioluminescence de la protéine spécifique à l'anticorps primaire. Afin de saturer les sites de fixation aspécifiques, la membrane est incubée dans du tampon bloquant 1 h à température ambiante. Elle est ensuite incubée toute la nuit à 4 °C sous agitation avec une dilution d'anticorps primaire (de lapin) allant de 1/5000ème à 1/25000ème dans du tampon anticorps 1. Après 3 lavages avec du tampon TBS-T, pour éliminer les anticorps non-attachés, la membrane est incubée avec un anticorps secondaire de chèvre anti-IgG de lapin, 1 h à température ambiante sous agitation dans du tampon anticorps 2. Après 3 autres lavages avec du tampon TBS-T, la révélation est effectuée grâce aux kits ECL ("Enhanced Chemical Luminescence") Western Blotting detection system (Amersham), ou Lumi-Light Western blotting system (Roche) selon le bruit de fond et le seuil de détection désiré.

Pour effectuer une seconde hybridation avec un autre anticorps primaire, la membrane doit être nettoyée de toute trace d'anticorps primaire ayant déjà servi. Pour cela, elle est incubée 2 fois 10 min dans de la solution strip à température ambiante, puis 2 fois 10 min dans du PBS et enfin 2 fois 5 min dans du TBS-T. Le protocole d'immunodétection peut alors être ré-appliqué à partir de la phase d'incubation dans le tampon bloquant.

TBS-T : NaCl 137 mM ; KCl 2,7 mM ; Tris-HCl 19 mM ; pH 7,4 ; Tween20 0,2% (p/v)

Tampon bloquant : lait écrémé en poudre entre 2% et 5% (p/v) dans du TBS-T

Tampon anticorps 1 : lait écrémé en poudre 2% (p/v) dans du TBS-T

Tampon anticorps 2 : lait écrémé en poudre 1% (p/v) dans du TBS-T

Solution strip : 1,5% (p/v) glycine ; SDS 0,1% (p/v) ; Tween20 1% (v/v)

2.6. Méthodes relatives aux Plantes

2.6.1. Transformation d'Arabidopsis par *Agrobacterium tumefaciens*

2.6.1.1. Transformation par « floral-dip »

Les plantes d'Arabidopsis sont prêtes à être transformées quand la première silique apparaît. Pour stimuler la formation de nombreuses hampes florales, les premières à apparaître sont coupées au dessus du premier nœud. Lorsque les plantes atteignent le bon stade de développement pour la transformation, une préculture de 3 mL d'agrobactéries contenant un plasmide binaire avec la construction d'intérêt est cultivée sur la nuit à 28 °C sous agitation dans du milieu LB plus antibiotiques adéquats. Le lendemain, 500 mL de milieu sont ensemencés avec 0,7 mL de cette préculture, puis incubés à 28 °C sous agitation pour atteindre une A_{600}=0,8. Les bactéries sont sédimentées 20 min à 5000 *g*, resuspendues

dans 10 mL de $MgCl_2$ 10 mM, puis diluées à nouveau à A_{600}=0,8 dans du saccharose 5% (p/v), Silwet 0,02% (p/v). Les parties aériennes des plantes sont alors trempées 30 sec dans la suspension bactérienne. Les plantes sont ensuite placées sous cloche pour maintenir des conditions humides pendant 24 h. Le taux d'humidité est progressivement réduit jusqu'au retour des plantes en serre. Les Arabidopsis sont alors cultivées dans des conditions normales jusqu'à l'obtention des graines.

Milieu LB antibiotiques : milieu LB ; rifampicine 50 µg/mL ; kanamycine 50 µg/mL.

2.6.1.2. Sélection des transformants sur boîte

Les graines sont ensuite sélectionnées sur un milieu contenant l'antibiotique correspondant à la résistance codée par le plasmide. Pour une boîte de culture de 14,5 cm de diamètre, une quantité de graines correspondant à un volume de 150 µL est stérilisée en milieu liquide. Les graines sont d'abord trempées 2 min dans de l'éthanol à 70% (v/v), puis incubées 15 min dans de l'eau de javel (chlore actif 4% (v/v)) avec du Tween-20 4% (v/v) sous agitation lente. Elles sont ensuite rincées 4 fois dans 10 mL d'eau stérile, reprises dans 12 mL d'agarose 0,1% (p/v) fondu et refroidi à température ambiante, et étalées dans la boîte contenant le milieu de sélection. La densité de cette solution d'agarose permet une répartition homogène des graines sur le milieu de sélection. Pour lever la dormance (stratification) et synchroniser la germination, les boîtes sont incubées à 4°C pendant deux jours. Un protocole de sélection rapide a été utilisé (Harrison *et al.*, 2006): après stratification, les plantes sont exposées à la lumière continue pendant 8 h pour activer la germination. Elles sont ensuite placées à l'obscurité pendant 2 jours, puis remises en lumière continue pendant 24 à 48 h. Les plantes résistantes à la kanamycine se reconnaissent à leur couleur verte prononcée, tandis que les plantes résistantes à l'hygromycine ont un hypocotyle plus allongé que les autres. Les plantes résistantes sont transférées vers une nouvelle boîte contenant du milieu de sélection Arabidopsis additionné de sucrose 1% (p/v), jusqu'à ce qu'elles soient assez grandes pour être transférées en terre.

Milieu de sélection Arabidopsis : MS0255 (Duchefa) pH 5,6 ; agarose 0,8% (p/v) ; kanamycine 50 µg/mL ou hygromycine B 15 µg/mL ; carbénicilline 200 µg/mL ; céfotaxime 250 µg/mL. Ces deux derniers antibiotiques servent à éliminer les agrobactéries et champignons divers présents sur les graines.

2.6.2. Croisements entre deux plantes

Arabidopsis est une plante autogame. Pour obtenir des doubles mutants homozygotes, un croisement entre les deux lignées de mutants est nécessaire. L'on peut également réintégrer un allèle dans une plante homozygote par rétrocroisement entre un mutant et une plante sauvage. Comme les organelles sont maternellement transmises, cela permet d'étudier l'effet de la réintégration d'une protéine sur des mitochondries issues de plantes mutantes. Pour les croisements, seulement 3-4 fleurs sur le point de s'ouvrir sont conservées sur la hampe florale. Ces fleurs sont émasculées, en prenant garde de ne pas abimer le pistil, et les plantes sont gardées en serre la nuit afin que le pistil finisse sa maturation. Le matin, le pistil est mis en contact avec une fleur mature de la plante paternelle, ce qui permet un dépôt de pollen de la plante paternelle sur le pistil de la plante maternelle. Les siliques sont ramassées dès leur maturation, juste avant leur ouverture, et finissent de sécher en tube ouvert.

2.6.3. Transfection par méthode biolistique

La transfection biolistique utilise un canon à particules pour projeter sur le matériel végétal des particules de tungstène ou d'or sur lesquelles l'ADN est précipité (Sanford *et al.*, 1993). Les transfections sont effectuées sur des jeunes feuilles de tabac coupées au niveau du pétiole, et déposées face ventrale dans une boîte de Pétri contenant du milieu de germination solide. Vingt µL d'une suspension de billes de tungstène (diamètre moyen de 1 µm) à 30 mg/mL dans du glycérol à 50 %, sont additionnés de 3 µg d'ADN plasmidique. Dans le cas d'une cotransfection, 3 µg du plasmide pCK-mRFP sont également ajoutés. L'ensemble est vortexé énergiquement et additionné de 20 µL de $CaCl_2$, 2,5 M et 8 µL de spermidine 0,1 M. Puis le tout est vortexé pendant 30 min à 4 °C. Les particules sont finalement lavées 3 fois avec 200 µL d'éthanol absolu froid, et reprises dans 30 µL d'éthanol.

La transfection biolistique se fait à l'aide d'un appareil PDS-1000/He™ (Bio-Rad). La feuille de tabac disposée sur le milieu de germination est placée dans la chambre de l'appareil, sous vide (0,1 bar). Six µL de particules portant l'ADN sont déposés sur un disque "macrocarrier" et projetés sur la feuille à 1100 psi (disques de rupture calibrés). Les boîtes sont ensuite incubées la nuit au noir à 20°C. Pour observer les cellules au microscope, les zones de la feuille intactes mais proches des impacts du bombardement sont découpées et posées face supérieure contre la lamelle. Une goutte d'eau est déposée sur la face inférieure et la feuille est grattée délicatement avec la lamelle pour percer la cuticule et faciliter l'observation. Le microscope confocal utilisé est un microscope inversé à balayage laser Zeiss LSM 700.

Milieu de germination solide: MS0255 (Duchefa) pH 5,6 ; agarose 0,8% (p/v) ; sucrose 1% (p/v).

2.6.4. Coloration GUS

Pour déterminer les organes où le promoteur d'un gène est actif, des plants d'Arabidopsis sont transformés avec le gène de la ß-glucuronidase (GUS) d'*E. coli* sous le contrôle du promoteur du gène d'intérêt. La ß-glucuronidase agit sur le cyclohexylammonium (X-Gluc) qu'elle hydrolyse pour former un produit bleu (Jefferson *et al.*, 1987). Les tissus où le promoteur est actif sont donc colorés en bleu. Cette coloration est effectuée sur différents organes à différents stades de développement. Les tissus végétaux sont prélevés et fixés dans l'acétone glacé à 90% (v/v) pendant 15 min sur glace avant d'être rincés dans du tampon phosphate. Les tissus sont ensuite immergés dans le tampon de coloration et infiltrés sous vide pendant 30 min pour faire pénétrer la solution à l'intérieur du matériel végétal. Le tout est ensuite incubé à l'obscurité toute la nuit à 37°C pour permettre la réaction d'hydrolyse. Pour mieux visualiser la coloration, la chlorophylle est lavée de l'échantillon par incubation dans des solutions de concentration croissante en éthanol (15%, 30%, 50%, 70%, 85%, 95% (v/v)). Une fois décolorés, les tissus peuvent être conservés dans une solution de fixation.

Tampon phosphate 50 mM : $NaPO_4$ 50 mM pH 7,2 (68,4 mL de Na_2HPO_4 50 mM et 31,6 mL de NaH_2PO_4 50 mM pour 100 mL de tampon)

Tampon de coloration : tampon phosphate 50 mM pH 7,2 ; EDTA 10 mM ; Triton X-100 0,1% (v/v) ; $K_3Fe(CN)_6$ 1 mM ; $K_4Fe(CN)_6$ 1 mM ; X-Gluc 0,4 mg/mL (conditionné à 20 mg/mL dans du N-N dimethylformamide (DMF) et conservé à -80°C).

Solution de fixation : éthanol 50% (v/v) ; acide acétique 5% (v/v) ; formaldéhyde 3,7% (v/v).

2.6.5. Test de respiration

Pour estimer le taux de fonctionnalité des mitochondries dans les plantes, des mesures de respiration sont effectuées sur des plantules *in vitro* de 3 semaines, en utilisant une électrode à oxygène (Eurosep instrument, Hansatech). Le matériel végétal est pesé avant chaque mesure puis plongé dans du tampon de respiration. L'ajout d'un substrat de la chaîne respiratoire, le succinate (10 mM), choisi pour sa capacité à être directement pris en charge par le complexe II, suivi par l'ajout d'ADP (200 µM), substrat de la phosphorylation, stimulent la respiration. L'ajout de KCN permet de bloquer la chaîne respiratoire. La consommation d'O_2 observée est

alors uniquement due à l'oxydase alternative et/ou à d'autres phénomènes d'oxydation. Le taux de respiration est obtenu en retirant la consommation d'O_2 non-inhibée par le KCN.

Tampon de respiration : KH_2PO_4 10 mM ; saccharose 0,3 M ; KCl 10 mM ; $MgCl_2$ 5 mM ; BSA 0,1% (p/v) ; pH 7,2

2.6.6. Culture de plantules *in vitro* en milieu solide

Les graines mises à germer *in vitro* sont stérilisées au préalable. Pour avoir des graines stériles sèches et faciles à déposer une à une, le protocole de stérilisation suivant est utilisé. Les graines sont incubées 10 min sous agitation dans de l'éthanol 70% (v/v) additionné de Triton X-100 à 0,05% (v/v) avant d'être lavées à l'éthanol 70% (v/v). Elles sont ensuite rincées rapidement à l'éthanol absolu avant d'être séchées à l'air libre sous une hotte stérile. Les graines sont réparties sur du milieu de culture solide à l'aide d'un cure-dent. Dans le cas de cultures sur milieu génotoxique, les substances testées à différentes concentrations sont ajoutées dans le milieu fondu et tiédi avant qu'il ne soit coulé dans les boîtes.

Milieu de culture solide : MS0255 (Duchefa) pH 5,6 ; agarose 0,8% (p/v) ; sucrose 1% (p/v).

2.6.7. Tests de sensibilité aux stress génotoxiques

2.6.7.1. Exposition au peroxyde d'hydrogène

Les stress oxydatifs génotoxiques au peroxyde d'hydrogène (Sigma-Aldrich) sont effectués selon le protocole établi par Molinier *et al.* (Molinier *et al.*, 2004b) en adaptant le matériel et les concentrations. Les stress ont été testés à la fois sur milieu solide et liquide, puis comparés. Les concentrations utilisées sont les suivantes : 0; 0,3; 0,6; 1,2; 2,5; 5 mM. Des plantes de 5 jours ayant poussé sur du milieu gélosé sont ensuite transférées stérilement dans du milieu solide ou liquide contenant les différentes doses d'agent génotoxique. La croissance des plantes est observée 10 jours à 3 semaines plus tard.

2.6.7.2. Exposition aux UV-C

Pour les tests de résistance aux UV-C (280-100nm), des plantes de 5 jours sont irradiées à l'aide de l'UV Stratalinker 2400™ (Stratagene). Trois irradiations sont effectuées, en espaçant chacune d'elles de 48 h. Une fois les plantes irradiées à 0, 750, 1500 ou 3000 J/m², elles sont maintenues 6 h à l'obscurité afin d'éviter l'action des photolyases.

2.7. Méthodes relatives aux organelles

2.7.1. Extraction de mitochondries

2.7.1.1. A partir de fleurs, de rosettes, ou de plantules *in vitro* d'Arabidopsis

Le but est d'obtenir des mitochondries isolées intactes et purifiées, exemptes de matériel nucléaire ou cytosolique. Afin d'éviter une dégradation des mitochondries, toutes les étapes en partant du broyage sont réalisées en chambre froide, en minimisant les temps d'attente. Le matériel végétal est récolté et pesé puis broyé dans du tampon de broyage à l'aide d'un mortier en utilisant 20 mL de tampon pour 10 g de matériel végétal. Le broyat est filtré grâce à deux couches de gaze et une toile de nylon (96 µm de vide de maille). L'opération de broyage est répétée avec le matériel récupéré dans le filtre. Le filtrat subit ensuite plusieurs centrifugations différentielles à 4°C. Une première de 5 min à 2000 *g* permet d'éliminer les plus gros débris cellulaires et les noyaux. La deuxième de 10 min à 20000 *g* est réalisée sur le surnageant et permet d'obtenir un culot formé entre autres de mitochondries. Le culot est repris dans du tampon de lavage et resuspendu dans un homogénéisateur de Dounce. Après une deuxième série de centrifugations différentielles, le culot est encore une fois resuspendu dans du tampon de lavage. Les mitochondries sont purifiées sur gradient de Percoll discontinu par centrifugation de 45 min à 40000 *g*, sans utilisation de frein en fin de centrifugation. Le gradient est réalisé en superposant 3 solutions de percoll, à 50%, 25% et 18% (v/v), respectivement. Dans le gradient de densité qui se forme durant la centrifugation, les mitochondries intactes sont récupérées à l'interface 50-25%, dans le bas du tube, puis reprises dans du tampon de lavage et sédimentées à 20000 *g* pendant 10 min. Après un lavage final elles sont reprises dans un volume minimal de tampon de lavage.

Tampon de broyage : sucrose 0,3 M ; tetrasodiumpyrophosphate ($10H_2O$) 5 mM ; EDTA 2 mM ; KH_2PO_4 10 mM ; polyvinylpyrrolidone (PVP-40) 1% (p/v) ; pH 7,5 ajusté à l'acide phosphorique. A ajouter juste avant utilisation : BSA 1% (p/v) ; acide ascorbique 20 mM ; cystéine 5 mM.

Tampon de lavage : sucrose 0,3 M ; MOPS 10 mM ; EGTA 1 mM ; pH 7,4 ajusté avec du KOH

Solution de gradient x% de Percoll: sucrose 0,3 M ; MOPS 10 mM ; pH 7,2 ajusté avec du KOH ; Percoll x% (v/v).

2.7.1.2. A partir de protoplastes de cultures cellulaires d'Arabidopsis

Huit erlenmeyers contenant 200 mL de milieu de culture liquide d'Arabidopsis sont ensemencés avec 10 % (v/v) d'une préculture de cellules maintenues en permanence à l'institut. Après 5 jours de culture sous agitation et lumière continue, les cellules sont récupérées par filtration du milieu sur toile de nylon (100 µm de vide de maille). La paroi végétale est ensuite digérée par incubation dans du tampon de digestion (5 mL/g de cellules), sous agitation à 30°C dans l'obscurité. Une observation à la loupe binoculaire permet de vérifier la disparition de la paroi végétale. Les protoplastes obtenus sont sédimentés à 800 *g* pendant 10 min, puis délicatement lavés 2 fois avec le tampon I. Ils sont ensuite repris dans 40 mL de tampon de broyage et cassés mécaniquement par filtration à travers une toile de nylon de 30 µm de vide de maille fixée sur une seringue. A partir de ce stade, le protocole précédent (II.7.1) est appliqué en commençant à la première étape de centrifugation.

Tampon I : mannitol 0,45 M ; MES 3,5 mM ; pH 5,5 ajusté avec du KOH.

Tampon de digestion : tampon I avec pectolyase Y23 0,10% (p/v) (Onozuka) ; cellulase RS 1% (p/v) (Onozuka).

Tampon de broyage : sucrose 0,3 M ; Pyrophosphate 30 mM ; EDTA 2 mM ; BSA 0,3% (p/v) ; PVP 0,8% (p/v) ; pH 7,5 ajusté avec du HCl. A ajouter avant utilisation : Glycine 5 mM ; Cystéine 0,05% (p/v) ; ß-mercaptoéthanol 2 mM.

2.7.2. Tests de contrôle respiratoire des mitochondries

Comme pour les plantes entières (II.6.5.), le test se déroule avec une électrode à oxygène (Eurosep instrument, Hansatech). Les mitochondries sont reprises dans du tampon de respiration (30 à 200 µg de protéines/mL). Le succinate (10 mM), substrat de respiration, est d'abord ajouté au milieu, suivi de l'ADP (200 µM), substrat de la phosphorylation. La mesure de la consommation d'O_2 lors de ces différentes phases permet de vérifier que les mitochondries sont fonctionnelles.

Le contrôle respiratoire des mitochondries est un bon indicateur de leur qualité fonctionnelle. Il reflète le couplage entre la respiration, mesurée par la consommation d'O_2 dans l'électrode à oxygène, et la phosphorylation de l'ADP en ATP, grâce au gradient de protons généré au long de la chaîne respiratoire si les mitochondries sont intactes. Le contrôle respiratoire est estimé par mesure du rapport de la consommation d'O_2 entre la phase III, qui suit l'addition d'ADP, et la phase IV où tout l'ADP ajouté a été transformé en ATP. Un rapport de 3 à 5 reflète un bon état fonctionnel des mitochondries.

Tampon de respiration : KH_2PO_4 10 mM ; saccharose 0,3 M ; KCl 10 mM ; $MgCl_2$ 5 mM ; BSA 0,1% (p/v) ; pH 7,2

2.7.3. Tests d'import d'ADN dans les mitochondries isolées

2.7.3.1. Marquage de l'ADN à importer

Dans le cas des expériences *in organello*, le fragment d'ADN destiné à être importé est synthétisé par PCR avec des amorces spécifiques, puis une partie est marquée radioactivement, afin de suivre l'import dans les mitochondries. Pour cela, un cycle supplémentaire de PCR de 20 min est effectué en présence de 100 µCi de [α-^{32}P]dCTP (3000 Ci/mmole), des mêmes amorces, d'une ADN polymérase et d'un mélange de dATP, dGTP et dTTP (0,4 mM chacun). L'élongation est complétée par addition de 0,4 mM de dCTP non radioactif et incubation 5 min à 72°C. Le produit est purifié sur "spin-column" de Sephadex G-50 et l'on effectue ensuite un comptage pour évaluer l'efficacité du marquage.

2.7.3.2. Import et recombinaison *in organello*

L'objectif est d'observer s'il y a import d'ADN dans les mitochondries isolées et recombinaison avec le mtDNA. Pour cela, les mitochondries sont simplement mises en contact avec l'ADN dans du tampon d'import (environ 12 ng d'ADN pour 1 mg de protéines mitochondriales). La suspension est incubée pendant 40 min à 25°C.

Après la phase d'import, l'ADN non-importé à l'intérieur des mitochondries doit être éliminé. Pour cela, les mitochondries sont sédimentées à travers 1 mL d'un coussin de saccharose à 27% par centrifugation à 10000 *g* pendant 10 min. Une digestion à la DNase est ensuite effectuée en reprenant les mitochondries dans du tampon d'import additionné de 75 µg/mL de DNase I et 10 mM de $MgCl_2$, puis en incubant 20 min à 25°C. Après 2 lavages au tampon de lavage, l'on incube 90 min à 25°C dans du tampon de réplication, contenant les éléments nécessaires à la synthèse d'ADN. Le but cette étape est de promouvoir la recombinaison *in organello* entre l'ADN importé et le mtDNA. Les acides nucléiques sont finalement extraits en ajoutant 1 volume de phénol/chloroforme (1 :1), puis en vortexant 2 min. Après une centrifugation de 15 min à 13000 *g* la phase aqueuse est prélevée et les acides nucléiques sont précipités à l'éthanol. Pour permettre la quantification exacte de l'ADN et pour éviter les interférences dans les diverses réactions amenant à la PCR inverse, les ARN sont hydrolysés en utilisant un mélange des RNases T1 et A, pendant 30 min à 25°C. Cette étape est suivie d'une nouvelle extraction au phénol/chloroforme et d'une précipitation à l'éthanol.

Un essai de contrôle est réalisé en parallèle, en mettant en contact les mitochondries avec l'ADN mais en enchaînant immédiatement les étapes de purification sur coussin de saccharose 27%, de digestion à la DNase et d'extraction des acides nucléiques comme décrit précédemment. Ce contrôle qui ne laisse le temps ni à l'import ni à une éventuelle recombinaison *in organello* de se dérouler sert à mettre en évidence les artéfacts possibles et le bruit de fond lors des analyses PCR ultérieures.

Finalement, pour vérifier l'étape d'import, un contrôle est réalisé en parallèle en incubant des mitochondries avec un fragment d'ADN marqué radioactivement. La digestion à la DNase est ensuite effectuée. Après 2 lavages au tampon de lavage, l'ADN est extrait comme décrit précédemment, analysé sur gel d'agarose et transféré sur membrane de nylon (Hybond N+, GE Healthcare) pour autoradiographie.

Tampon d'import : saccharose 0,4 M ; phosphate de potassium 40 mM, pH 7,2.

Tampon de lavage: saccharose 0,4 M ; phosphate de potassium 40 mM ; EDTA 10 mM ; EGTA 10 mM, pH 7,2.

Coussin de saccharose 27% : saccharose 27% (p/v) ; phosphate de potassium 10 mM ; EDTA 1 mM ; BSA 0,1% (p/v), pH 7,2.

Tampon de réplication : saccharose 0,33 M ; KCl 90 mM ; Tricine 12 mM ; KH_2PO_4 5 mM ; EDTA 1,2 mM ; $MgCl_2$ 5 mM ; DTT 2 mM ; ADP 1 mM ; dNTP mix 50 µM ; succinate 10 mM, pH 7,2.

REFERENCES

Abdelnoor, R.V., Yule, R., Elo, A., Christensen, A.C., Meyer-Gauen, G., et Mackenzie, S.A. (2003). Substoichiometric shifting in the plant mitochondrial genome is influenced by a gene homologous to MutS. *Proc Natl Acad Sci U S A* **100,** 5968-5973.

Abdelnoor, R.V., Christensen, A.C., Mohammed, S., Munoz-Castillo, B., Moriyama, H., et Mackenzie, S.A. (2006). Mitochondrial genome dynamics in plants and animals: convergent gene fusions of a MutS homologue. *J Mol Evol* **63,** 165-173.

Adams, K.L., et Palmer, J.D. (2003). Evolution of mitochondrial gene content: gene loss and transfer to the nucleus. *Mol Phylogenet Evol* **29,** 380-395.

Ahmad, M., Jarillo, J.A., Klimczak, L.J., Landry, L.G., Peng, T., Last, R.L., et Cashmore, A.R. (1997). An enzyme similar to animal type II photolyases mediates photoreactivation in Arabidopsis. *Plant Cell* **9,** 199-207.

Allen, J.O., Fauron, C.M., Minx, P., Roark, L., Oddiraju, S., Lin, G.N., Meyer, L., Sun, H., Kim, K., Wang, C., Du, F., Xu, D., Gibson, M., Cifrese, J., Clifton, S.W., et Newton, K.J. (2007). Comparisons among two fertile and three male-sterile mitochondrial genomes of maize. *Genetics* **177,** 1173-1192.

Allinson, S.L., Dianova, II, et Dianov, G.L. (2001). DNA polymerase beta is the major dRP lyase involved in repair of oxidative base lesions in DNA by mammalian cell extracts. *EMBO J* **20,** 6919-6926.

Alverson, A.J., Rice, D.W., Dickinson, S., Barry, K., et Palmer, J.D. (2011). Origins and recombination of the bacterial-sized multichromosomal mitochondrial genome of cucumber. *Plant Cell* **23,** 2499-2513.

Andre, C., Levy, A., et Walbot, V. (1992). Small repeated sequences and the structure of plant mitochondrial genomes. *Trends Genet* **8,** 128-132.

Apel, K., et Hirt, H. (2004). Reactive oxygen species: metabolism, oxidative stress, and signal transduction. *Annu Rev Plant Biol* **55,** 373-399.

Araki, H. (2001). [DNA polymerase epsilon and relative factors in the S phase checkpoint]. *Tanpakushitsu Kakusan Koso* **46,** 1201-1207.

Arrieta-Montiel, M.P., Shedge, V., Davila, J., Christensen, A.C., et Mackenzie, S.A. (2009). Diversity of the Arabidopsis mitochondrial genome occurs via nuclear-controlled recombination activity. *Genetics* **183,** 1261-1268.

Au, K.G., Welsh, K., et Modrich, P. (1992). Initiation of methyl-directed mismatch repair. *J Biol Chem* **267,** 12142-12148.

Backert, S., Dorfel, P., Lurz, R., et Borner, T. (1996). Rolling-circle replication of mitochondrial DNA in the higher plant Chenopodium album (L.). *Mol Cell Biol* **16,** 6285-6294.

Backert, S., Nielsen, B.L., et Borner, T. (1997). The mystery of the rings: structure and replication of mitochondrial genomes from higher plants. *Trends in Plant Science* **2,** 477-483.

Backert, S., et Borner, T. (2000). Phage T4-like intermediates of DNA replication and recombination in the mitochondria of the higher plant Chenopodium album (L.). *Curr Genet* **37,** 304-314.

Bauknecht, M., et Kobbe, D. (2014). AtGEN1 and AtSEND1, two paralogs in Arabidopsis, possess holliday junction resolvase activity. *Plant Physiol* **166,** 202-216.

Beam, C.E., Saveson, C.J., et Lovett, S.T. (2002). Role for radA/sms in recombination intermediate processing in Escherichia coli. *J Bacteriol* **184,** 6836-6844.

Beerens, N., Hoeijmakers, J.H., Kanaar, R., Vermeulen, W., et Wyman, C. (2005). The CSB protein actively wraps DNA. *J Biol Chem* **280,** 4722-4729.
Bendich, A.J. (1993). Reaching for the ring: the study of mitochondrial genome structure. *Curr Genet* **24,** 279-290.
Bernard, P., et Couturier, M. (1992). Cell killing by the F plasmid CcdB protein involves poisoning of DNA-topoisomerase II complexes. *J Mol Biol* **226,** 735-745.
Betous, R., Glick, G.G., Zhao, R., et Cortez, D. (2013). Identification and characterization of SMARCAL1 protein complexes. *PLoS One* **8,** e63149.
Bevan, M. (1984). Binary Agrobacterium vectors for plant transformation. *Nucleic Acids Res* **12,** 8711-8721.
Blanck, S., Kobbe, D., Hartung, F., Fengler, K., Focke, M., et Puchta, H. (2009). A SRS2 homolog from Arabidopsis thaliana disrupts recombinogenic DNA intermediates and facilitates single strand annealing. *Nucleic Acids Res* **37,** 7163-7176.
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.
Bonnefoy, N., Remacle, C., et Fox, T.D. (2007). Genetic transformation of Saccharomyces cerevisiae and Chlamydomonas reinhardtii mitochondria. *Methods Cell Biol* **80,** 525-548.
Bradford, M.M. (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.
Brettel, K., et Byrdin, M. (2010). Reaction mechanisms of DNA photolyase. *Curr Opin Struct Biol* **20,** 693-701.
Britt, A. (2002). Repair of damaged bases. *Arabidopsis Book* **1,** e0005.
Britt, A.B. (1996). DNA Damage and Repair in Plants. *Annu Rev Plant Physiol Plant Mol Biol* **47,** 75-100.
Britt, A.B. (1999). Molecular genetics of DNA repair in higher plants. *Trends in Plant Science* **4,** 20-25.
Budar, F., et Pelletier, G. (2001). Male sterility in plants: occurrence, determinism, significance and use. *C R Acad Sci III* **324,** 543-550.
Burdett, V., Baitinger, C., Viswanathan, M., Lovett, S.T., et Modrich, P. (2001). In vivo requirement for RecJ, ExoVII, ExoI, and ExoX in methyl-directed mismatch repair. *Proc Natl Acad Sci U S A* **98,** 6765-6770.
Busso, D., Delagoutte-Busso, B., et Moras, D. (2005). Construction of a set Gateway-based destination vectors for high-throughput cloning and expression screening in Escherichia coli. *Anal Biochem* **343,** 313-321.
Cappadocia, L., Marechal, A., Parent, J.S., 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.
Cappelli, E., Taylor, R., Cevasco, M., Abbondandolo, A., Caldecott, K., et Frosina, G. (1997). Involvement of XRCC1 and DNA ligase III gene products in DNA base excision repair. *J Biol Chem* **272,** 23970-23975.
Caron, P., Choudjaye, J., Clouaire, T., Bugler, B., Daburon, V., Aguirrebengoa, M., Mangeat, T., Iacovoni, J.S., Alvarez-Quilon, A., Cortes-Ledesma, F., et Legube, G. (2015). Non-redundant Functions of ATM and DNA-PKcs in Response to DNA Double-Strand Breaks. *Cell Rep* **13,** 1598-1609.
Carrie, C., Kuhn, K., Murcha, M.W., Duncan, O., Small, I.D., O'Toole, N., et Whelan, J. (2009). Approaches to defining dual-targeted proteins in Arabidopsis. *Plant J* **57,** 1128-1139.

Champoux, J.J. (2001). DNA topoisomerases: structure, function, and mechanism. *Annu Rev Biochem* **70,** 369-413.
Chan, S.N., Vincent, S.D., et Lloyd, R.G. (1998). Recognition and manipulation of branched DNA by the RusA Holliday junction resolvase of Escherichia coli. *Nucleic Acids Res* **26,** 1560-1566.
Chang, A.C., et Cohen, S.N. (1978). Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J Bacteriol* **134,** 1141-1156.
Chaves, I., Nijman, R.M., Biernat, M.A., Bajek, M.I., Brand, K., da Silva, A.C., Saito, S., Yagita, K., Eker, A.P., and van der Horst, G.T. (2011). The Potorous CPD photolyase rescues a cryptochrome-deficient mammalian circadian clock. PloS one **6,** e23447.
Chen, H., Chu, P., Zhou, Y., Li, Y., Liu, J., Ding, Y., Tsang, E.W., Jiang, L., Wu, K., et Huang, S. (2012). Overexpression of AtOGG1, a DNA glycosylase/AP lyase, enhances seed longevity and abiotic stress tolerance in Arabidopsis. *J Exp Bot* **63,** 4107-4121.
Cho, Y., Mower, J.P., Qiu, Y.L., et Palmer, J.D. (2004). Mitochondrial substitution rates are extraordinarily elevated and variable in a genus of flowering plants. *Proc Natl Acad Sci U S A* **101,** 17741-17746.
Christensen, A.C. (2013). Plant mitochondrial genome evolution can be explained by DNA repair mechanisms. *Genome Biol Evol* **5,** 1079-1086.
Clark, M.A., et Shay, J.W. (1982). Mitochondrial transformation of mammalian cells. *Nature* **295,** 605-607.
Cohen, S.E., Godoy, V.G., et Walker, G.C. (2009). Transcriptional modulator NusA interacts with translesion DNA polymerases in Escherichia coli. *J Bacteriol* **191,** 665-672.
Cohen, S.E., et Walker, G.C. (2010). The transcription elongation factor NusA is required for stress-induced mutagenesis in Escherichia coli. *Curr Biol* **20,** 80-85.
Constantin, N., Dzantiev, L., Kadyrov, F.A., et Modrich, P. (2005). Human mismatch repair: reconstitution of a nick-directed bidirectional reaction. *J Biol Chem* **280,** 39752-39761.
Cooper, D.L., Lahue, R.S., et Modrich, P. (1993). Methyl-directed mismatch repair is bidirectional. *J Biol Chem* **268,** 11823-11829.
Cooper, D.L., Boyle, D.C., et Lovett, S.T. (2015). Genetic analysis of Escherichia coli RadA: functional motifs and genetic interactions. *Mol Microbiol* **95,** 769-779.
Cooper, D.L., et Lovett, S.T. (2016). Recombinational branch migration by the RadA/Sms paralog of RecA in Escherichia coli. *Elife* **5**.
Cordoba-Canero, D., Dubois, E., Ariza, R.R., Doutriaux, M.P., et Roldan-Arjona, T. (2010). Arabidopsis uracil DNA glycosylase (UNG) is required for base excision repair of uracil and increases plant sensitivity to 5-fluorouracil. *J Biol Chem* **285,** 7475-7483.
Cordoba-Canero, D., Roldan-Arjona, T., et Ariza, R.R. (2014). Arabidopsis ZDP DNA 3'-phosphatase and ARP endonuclease function in 8-oxoG repair initiated by FPG and OGG1 DNA glycosylases. *Plant J* **79,** 824-834.
Costa, R.M., Morgante, P.G., Berra, C.M., Nakabashi, M., Bruneau, D., Bouchez, D., Sweder, K.S., Van Sluys, M.A., et Menck, C.F. (2001). The participation of AtXPB1, the XPB/RAD25 homologue gene from Arabidopsis thaliana, in DNA repair and plant development. *Plant J* **28,** 385-395.
Costanzo, V. (2011). Brca2, Rad51 and Mre11: performing balancing acts on replication forks. *DNA Repair (Amst)* **10,** 1060-1065.
Cox, M.M. (2007). Motoring along with the bacterial RecA protein. *Nat Rev Mol Cell Biol* **8,** 127-138.

Cuneo, M.J., et London, R.E. (2010). Oxidation state of the XRCC1 N-terminal domain regulates DNA polymerase beta binding affinity. *Proc Natl Acad Sci U S A* **107,** 6805-6810.

Cupp, J.D., et Nielsen, B.L. (2013). Arabidopsis thaliana organellar DNA polymerase IB mutants exhibit reduced mtDNA levels with a decrease in mitochondrial area density. *Physiol Plant* **149,** 91-103.

Curtis, M.D., et Grossniklaus, U. (2003). A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol* **133,** 462-469.

Cvetkovska, M., et Vanlerberghe, G.C. (2012). Alternative oxidase modulates leaf mitochondrial concentrations of superoxide and nitric oxide. *New Phytol* **195,** 32-39.

Dany, A.L., et Tissier, A. (2001). A functional OGG1 homologue from Arabidopsis thaliana. *Mol Genet Genomics* **265,** 293-301.

Darracq, A., Varre, J.S., et Touzet, P. (2010). A scenario of mitochondrial genome evolution in maize based on rearrangement events. *BMC Genomics* **11,** 233.

Davila, J.I., Arrieta-Montiel, M.P., Wamboldt, Y., Cao, J., Hagmann, J., Shedge, V., Xu, Y.Z., Weigel, D., et Mackenzie, S.A. (2011). Double-strand break repair processes drive evolution of the mitochondrial genome in Arabidopsis. *BMC Biol* **9,** 64.

de Haas, J.M., Hille, J., Kors, F., van der Meer, B., Kool, A.J., Folkerts, O., et Nijkamp, H.J. (1991). Two potential Petunia hybrida mitochondrial DNA replication origins show structural and in vitro functional homology with the animal mitochondrial DNA heavy and light strand replication origins. *Curr Genet* **20,** 503-513.

Deaconescu, A.M., Chambers, A.L., Smith, A.J., Nickels, B.E., Hochschild, A., Savery, N.J., et Darst, S.A. (2006). Structural basis for bacterial transcription-coupled DNA repair. *Cell* **124,** 507-520.

Deaconescu, A.M., Savery, N., et Darst, S.A. (2007). The bacterial transcription repair coupling factor. *Curr Opin Struct Biol* **17,** 96-102.

Deaconescu, A.M. (2013). RNA polymerase between lesion bypass and DNA repair. *Cell Mol Life Sci* **70,** 4495-4509.

del Rio, L.A., Sandalio, L.M., Corpas, F.J., Palma, J.M., et Barroso, J.B. (2006). Reactive oxygen species and reactive nitrogen species in peroxisomes. Production, scavenging, and role in cell signaling. *Plant Physiol* **141,** 330-335.

Demple, B., et Linn, S. (1980). DNA N-glycosylases and UV repair. *Nature* **287,** 203-208.

Demple, B., et Harrison, L. (1994). Repair of oxidative damage to DNA: enzymology and biology. *Annu Rev Biochem* **63,** 915-948.

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J.F., Guindon, S., Lefort, V., Lescot, M., Claverie, J.M., et Gascuel, O. (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* **36,** W465-469.

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

Diray-Arce, J., Liu, B., Cupp, J.D., Hunt, T., et Nielsen, B.L. (2013). The Arabidopsis At1g30680 gene encodes a homologue to the phage T7 gp4 protein that has both DNA primase and DNA helicase activities. *BMC Plant Biol* **13,** 36.

Dodson, M.L., et Lloyd, R.S. (2002). Mechanistic comparisons among base excision repair glycosylases. *Free Radic Biol Med* **32,** 678-682.

Dzantiev, L., Constantin, N., Genschel, J., Iyer, R.R., Burgers, P.M., et Modrich, P. (2004). A defined human system that supports bidirectional mismatch-provoked excision. *Mol Cell* **15,** 31-41.

Edmondson, A.C., Song, D., Alvarez, L.A., Wall, M.K., Almond, D., McClellan, D.A., Maxwell, A., et Nielsen, B.L. (2005). Characterization of a mitochondrially targeted single-stranded DNA-binding protein in Arabidopsis thaliana. *Mol Genet Genomics* **273,** 115-122.

Elo, A., Lyznik, A., Gonzalez, D.O., Kachman, S.D., et Mackenzie, S.A. (2003). Nuclear genes that encode mitochondrial proteins for DNA and RNA metabolism are clustered in the Arabidopsis genome. *Plant Cell* **15,** 1619-1631.

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

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.

Epshtein, V. (2015). UvrD helicase: an old dog with a new trick: how one step backward leads to many steps forward. *Bioessays* **37,** 12-19.

Falck, J., Coates, J., et Jackson, S.P. (2005). Conserved modes of recruitment of ATM, ATR and DNA-PKcs to sites of DNA damage. *Nature* **434,** 605-611.

Fauron, C.M., Havlik, M., et Brettell, R.I. (1990). The mitochondrial genome organization of a maize fertile cmsT revertant line is generated through recombination between two sets of repeats. *Genetics* **124,** 423-428.

Fitch, M.E., Nakajima, S., Yasui, A., et Ford, J.M. (2003). In vivo recruitment of XPC to UV-induced cyclobutane pyrimidine dimers by the DDB2 gene product. *J Biol Chem* **278,** 46906-46910.

Foyer, C.H., et Shigeoka, S. (2011). Understanding oxidative stress and antioxidant functions to enhance photosynthesis. *Plant Physiol* **155,** 93-100.

Friedberg, E.C. (2005). Suffering in silence: the tolerance of DNA damage. *Nat Rev Mol Cell Biol* **6,** 943-953.

Fritsch, O., Benvenuto, G., Bowler, C., Molinier, J., et Hohn, B. (2004). The INO80 protein controls homologous recombination in Arabidopsis thaliana. *Mol Cell* **16,** 479-485.

Fromme, J.C., et Verdine, G.L. (2004). Base excision repair. *Adv Protein Chem* **69,** 1-41.

Frosina, G., Fortini, P., Rossi, O., Carrozzino, F., Raspaglio, G., Cox, L.S., Lane, D.P., Abbondandolo, A., et Dogliotti, E. (1996). Two pathways for base excision repair in mammalian cells. *J Biol Chem* **271,** 9573-9578.

Gallego, F., Fleck, O., Li, A., Wyrzykowska, J., et Tinland, B. (2000). AtRAD1, a plant homologue of human and yeast nucleotide excision repair endonucleases, is involved in dark repair of UV damages and recombination. *Plant J* **21,** 507-518.

Gao, M.J., et Murphy, T.M. (2001). Alternative forms of formamidopyrimidine-DNA glycosylase from Arabidopsis thaliana. *Photochem Photobiol* **73,** 128-134.

Garcia-Ortiz, M.V., Ariza, R.R., et Roldan-Arjona, T. (2001). An OGG1 orthologue encoding a functional 8-oxoguanine DNA glycosylase/lyase in Arabidopsis thaliana. *Plant Mol Biol* **47,** 795-804.

Gerhold, J.M., Aun, A., Sedman, T., Joers, P., et Sedman, J. (2010). Strand invasion structures in the inverted repeat of Candida albicans mitochondrial DNA reveal a role for homologous recombination in replication. *Mol Cell* **39,** 851-861.

Gill, S.S., et Tuteja, N. (2010). Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. *Plant Physiol Biochem* **48,** 909-930.

Gorbunova, V.V., et Levy, A.A. (1999). How plants make ends meet: DNA double-strand break repair. *Trends in Plant Science* **4,** 263-269.

Griffiths, L.M., Swartzlander, D., Meadows, K.L., Wilkinson, K.D., Corbett, A.H., et Doetsch, P.W. (2009). Dynamic compartmentalization of base excision repair proteins in response to nuclear and mitochondrial oxidative stress. *Mol Cell Biol* **29,** 794-807.

Grove, J.I., Harris, L., Buckman, C., et Lloyd, R.G. (2008). DNA double strand break repair and crossing over mediated by RuvABC resolvase and RecG translocase. *DNA Repair (Amst)* **7,** 1517-1530.

Gualberto, J.M., Mileshina, D., Wallet, C., Niazi, A.K., Weber-Lotfi, F., et Dietrich, A. (2014). The plant mitochondrial genome: dynamics and maintenance. *Biochimie* **100,** 107-120.

Gualberto, J.M., Le Ret, M., Beator, B., et Kuhn, K. (2015). The RAD52-like protein ODB1 is required for the efficient excision of two mitochondrial introns spliced via first-step hydrolysis. *Nucleic Acids Res* **43,** 6500-6510.

Gupta, S., Yeeles, J.T., et Marians, K.J. (2014). Regression of Replication Forks Stalled by Leading-strand Template Damage: I. BOTH RecG AND RuvAB CATALYZE REGRESSION, BUT RuvC CLEAVES THE HOLLIDAY JUNCTIONS FORMED BY RecG PREFERENTIALLY. *J Biol Chem* **289,** 28376-28387.

Gutman, B.L., et Niyogi, K.K. (2009). Evidence for base excision repair of oxidative DNA damage in chloroplasts of Arabidopsis thaliana. *J Biol Chem* **284,** 17006-17012.

Guzman, L.M., Belin, D., Carson, M.J., et Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* **177,** 4121-4130.

Hanson, M.R., et Bentolila, S. (2004). Interactions of mitochondrial and nuclear genes that affect male gametophyte development. *Plant Cell* **16 Suppl,** S154-169.

Hara, R., Selby, C.P., Liu, M., Price, D.H., et Sancar, A. (1999). Human transcription release factor 2 dissociates RNA polymerases I and II stalled at a cyclobutane thymine dimer. *J Biol Chem* **274,** 24779-24786.

Hargreaves, D., Rice, D.W., Sedelnikova, S.E., Artymiuk, P.J., Lloyd, R.G., et Rafferty, J.B. (1998). Crystal structure of E.coli RuvA with bound DNA Holliday junction at 6 A resolution. *Nat Struct Biol* **5,** 441-446.

Harrison, S.J., Mott, E.K., Parsley, K., Aspinall, S., Gray, J.C., et Cottage, A. (2006). A rapid and robust method of identifying transformed Arabidopsis thaliana seedlings following floral dip transformation. *Plant Methods* **2,** 19.

Hartley, J.L., Temple, G.F., et Brasch, M.A. (2000). DNA cloning using in vitro site-specific recombination. *Genome Res* **10,** 1788-1795.

Hays, J.B. (2002). Arabidopsis thaliana, a versatile model system for study of eukaryotic genome-maintenance functions. *DNA Repair (Amst)* **1,** 579-600.

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.

Hefner, E., Preuss, S.B., et Britt, A.B. (2003). Arabidopsis mutants sensitive to gamma radiation include the homologue of the human repair gene ERCC1. *J Exp Bot* **54,** 669-680.

Heller, R.C., et Marians, K.J. (2006). Replisome assembly and the direct restart of stalled replication forks. *Nat Rev Mol Cell Biol* **7,** 932-943.

Hess, W.R., et Borner, T. (1999). Organellar RNA polymerases of higher plants. *Int Rev Cytol* **190,** 1-59.

Heyer, W.D., Ehmsen, K.T., et Liu, J. (2010). Regulation of homologous recombination in eukaryotes. *Annu Rev Genet* **44,** 113-139.

Holliday, R. (2007). A mechanism for gene conversion in fungi (Reprinted). *Genetics Research* **89,** 285-307.

Hong, X., Cadwell, G.W., et Kogoma, T. (1995). Escherichia coli RecG and RecA proteins in R-loop formation. *EMBO J* **14,** 2385-2392.

Huertas, P. (2010). DNA resection in eukaryotes: deciding how to fix the break. *Nat Struct Mol Biol* **17,** 11-16.

Ibrahim, N., Handa, H., Cosset, A., Koulintchenko, M., Konstantinov, Y., Lightowlers, R.N., Dietrich, A., et Weber-Lotfi, F. (2011). DNA delivery to mitochondria: sequence specificity and energy enhancement. *Pharm Res* **28,** 2871-2882.

Iyer, R.R., Pluciennik, A., Burdett, V., et Modrich, P.L. (2006). DNA mismatch repair: functions and mechanisms. *Chem Rev* **106,** 302-323.

Janicka, S., Kuhn, K., Le Ret, M., Bonnard, G., Imbault, P., Augustyniak, H., et Gualberto, J.M. (2012). A RAD52-like single-stranded DNA binding protein affects mitochondrial DNA repair by recombination. *Plant J* **72,** 423-435.

Jansen, M.A.K., Gaba, V., et Greenberg, B.M. (1998). Higher plants and UV-B radiation: balancing damage, repair and acclimation (vol 3, pg 131, 1998). *Trends in Plant Science* **3,** 243-243.

Jefferson, R.A., Kavanagh, T.A., et Bevan, M.W. (1987). GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J* **6,** 3901-3907.

Jiang, C.Z., Yee, J., Mitchell, D.L., et Britt, A.B. (1997). Photorepair mutants of Arabidopsis. *Proc Natl Acad Sci U S A* **94,** 7441-7445.

Johansson, E., et Dixon, N. (2013). Replicative DNA polymerases. *Cold Spring Harb Perspect Biol* **5**.

Johne, R., Muller, H., Rector, A., van Ranst, M., et Stevens, H. (2009). Rolling-circle amplification of viral DNA genomes using phi29 polymerase. *Trends Microbiol* **17,** 205-211.

Kadyrov, F.A., Dzantiev, L., Constantin, N., et Modrich, P. (2006). Endonucleolytic function of MutLalpha in human mismatch repair. *Cell* **126,** 297-308.

Kelner, A. (1949). Photoreactivation of Ultraviolet-Irradiated Escherichia Coli, with Special Reference to the Dose-Reduction Principle and to Ultraviolet-Induced Mutation. *J Bacteriol* **58,** 511-522.

Killen, M.W., Stults, D.M., Wilson, W.A., et Pierce, A.J. (2012). Escherichia coli RecG functionally suppresses human Bloom syndrome phenotypes. *BMC Mol Biol* **13,** 33.

Kimura, S., et Sakaguchi, K. (2006). DNA repair in plants. *Chem Rev* **106,** 753-766.

Kleffmann, T., Russenberger, D., von Zychlinski, A., Christopher, W., Sjolander, K., Gruissem, W., et Baginsky, S. (2004). The Arabidopsis thaliana chloroplast proteome reveals pathway abundance and novel protein functions. *Curr Biol* **14,** 354-362.

Klein, M., Eckert-Ossenkopp, U., Schmiedeberg, I., Brandt, P., Unseld, M., Brennicke, A., et Schuster, W. (1994). Physical mapping of the mitochondrial genome of Arabidopsis thaliana by cosmid and YAC clones. *Plant J* **6,** 447-455.

Klungland, A., et Lindahl, T. (1997). Second pathway for completion of human DNA base excision-repair: reconstitution with purified proteins and requirement for DNase IV (FEN1). *EMBO J* **16,** 3341-3348.

Knoll, A., Higgins, J.D., Seeliger, K., Reha, S.J., Dangel, N.J., Bauknecht, M., Schropfer, S., Franklin, F.C., et Puchta, H. (2012). The Fanconi anemia ortholog FANCM ensures

ordered homologous recombination in both somatic and meiotic cells in Arabidopsis. *Plant Cell* **24,** 1448-1464.
Kogoma, T. (1997). Stable DNA replication: interplay between DNA replication, homologous recombination, and transcription. *Microbiol Mol Biol Rev* **61,** 212-238.
Koulintchenko, M., Konstantinov, Y., et Dietrich, A. (2003). Plant mitochondria actively import DNA via the permeability transition pore complex. *EMBO J* **22,** 1245-1254.
Koulintchenko, M., Temperley, R.J., Mason, P.A., Dietrich, A., et Lightowlers, R.N. (2006). Natural competence of mammalian mitochondria allows the molecular investigation of mitochondrial gene expression. *Hum Mol Genet* **15,** 143-154.
Kowalczykowski, S.C. (2000). Initiation of genetic recombination and recombination-dependent replication. *Trends Biochem Sci* **25,** 156-165.
Kowalczykowski, S.C. (2015). An Overview of the Molecular Mechanisms of Recombinational DNA Repair. *Cold Spring Harb Perspect Biol* **7**.
Krejci, L., Van Komen, S., Li, Y., Villemain, J., Reddy, M.S., Klein, H., Ellenberger, T., et Sung, P. (2003). DNA helicase Srs2 disrupts the Rad51 presynaptic filament. *Nature* **423,** 305-309.
Kreuzer, K.N. (2000). Recombination-dependent DNA replication in phage T4. *Trends Biochem Sci* **25,** 165-173.
Krokan, H.E., Standal, R., et Slupphaug, G. (1997). DNA glycosylases in the base excision repair of DNA. *Biochem J* **325 (Pt 1),** 1-16.
Kubo, T., et Newton, K.J. (2008). Angiosperm mitochondrial genomes and mutations. *Mitochondrion* **8,** 5-14.
Kuhn, K., Richter, U., Meyer, E.H., Delannoy, E., de Longevialle, A.F., O'Toole, N., Borner, T., Millar, A.H., Small, I.D., et Whelan, J. (2009). Phage-type RNA polymerase RPOTmp performs gene-specific transcription in mitochondria of Arabidopsis thaliana. *Plant Cell* **21,** 2762-2779.
Kühn, K., et Gualberto, J.M. (2012). Recombination in the Stability, Repair and Evolution of the Mitochondrial Genome **63,** 215-252.
Kunkel, T.A. (2004). DNA replication fidelity. *J Biol Chem* **279,** 16895-16898.
Kunkel, T.A., et Erie, D.A. (2015). Eukaryotic Mismatch Repair in Relation to DNA Replication. *Annu Rev Genet* **49,** 291-313.
Kuzmin, E.V., Duvick, D.N., et Newton, K.J. (2005). A mitochondrial mutator system in maize. *Plant Physiol* **137,** 779-789.

Landry, L.G., Stapleton, A.E., Lim, J., Hoffman, P., Hays, J.B., Walbot, V., et Last, R.L. (1997). An Arabidopsis photolyase mutant is hypersensitive to ultraviolet-B radiation. *Proc Natl Acad Sci U S A* **94,** 328-332.
Lee, Y.C., Cai, Y., Mu, H., Broyde, S., Amin, S., Chen, X., Min, J.H., et Geacintov, N.E. (2014). The relationships between XPC binding to conformationally diverse DNA adducts and their excision by the human NER system: is there a correlation? *DNA Repair (Amst)* **19,** 55-63.
Lees-Miller, S.P., et Meek, K. (2003). Repair of DNA double strand breaks by non-homologous end joining. *Biochimie* **85,** 1161-1173.
Lin, J.J., et Sancar, A. (1992). (A)BC excinuclease: the Escherichia coli nucleotide excision repair enzyme. *Mol Microbiol* **6,** 2219-2224.
Lindahl, T. (1993). Instability and decay of the primary structure of DNA. *Nature* **362,** 709-715.
Liu, Z., Hossain, G.S., Islas-Osuna, M.A., Mitchell, D.L., et Mount, D.W. (2000). Repair of UV damage in plants by nucleotide excision repair: Arabidopsis UVH1 DNA repair gene is a homolog of Saccharomyces cerevisiae Rad1. *Plant J* **21,** 519-528.

Liu, Z., Hall, J.D., et Mount, D.W. (2001). Arabidopsis UVH3 gene is a homolog of the Saccharomyces cerevisiae RAD2 and human XPG DNA repair genes. *Plant J* **26,** 329-338.

Liu, Z., Hong, S.W., Escobar, M., Vierling, E., Mitchell, D.L., Mount, D.W., et Hall, J.D. (2003). Arabidopsis UVH6, a homolog of human XPD and yeast RAD3 DNA repair genes, functions in DNA repair and is essential for plant growth. *Plant Physiol* **132,** 1405-1414.

Lloyd, R.G., et Buckman, C. (1991). Genetic analysis of the recG locus of Escherichia coli K-12 and of its role in recombination and DNA repair. *J Bacteriol* **173,** 1004-1011.

Lloyd, R.G., et Sharples, G.J. (1993). Dissociation of synthetic Holliday junctions by E. coli RecG protein. *EMBO J* **12,** 17-22.

Lohman, G.J., Tabor, S., et Nichols, N.M. (2011). DNA ligases. *Curr Protoc Mol Biol* **Chapter 3,** Unit3 14.

Lohman, T.M., Tomko, E.J., et Wu, C.G. (2008). Non-hexameric DNA helicases and translocases: mechanisms and regulation. *Nat Rev Mol Cell Biol* **9,** 391-401.

Lonsdale, D.M., Hodge, T.P., et Fauron, C.M. (1984). The physical map and organisation of the mitochondrial genome from the fertile cytoplasm of maize. *Nucleic Acids Res* **12,** 9249-9261.

Lovett, S.T., Hurley, R.L., Sutera, V.A., Jr., Aubuchon, R.H., et Lebedeva, M.A. (2002). Crossing over between regions of limited homology in Escherichia coli. RecA-dependent and RecA-independent pathways. *Genetics* **160,** 851-859.

Lyrawati, D., Trounson, A., et Cram, D. (2011). Expression of GFP in the mitochondrial compartment using DQAsome-mediated delivery of an artificial mini-mitochondrial genome. *Pharm Res* **28,** 2848-2862.

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.

Mahdi, A.A., Sharples, G.J., Mandal, T.N., et Lloyd, R.G. (1996). Holliday junction resolvases encoded by homologous rusA genes in Escherichia coli K-12 and phage 82. *J Mol Biol* **257,** 561-573.

Majeran, W., Friso, G., Asakura, Y., Qu, X., Huang, M., Ponnala, L., Watkins, K.P., Barkan, A., et van Wijk, K.J. (2012). Nucleoid-enriched proteomes in developing plastids and chloroplasts from maize leaves: a new conceptual framework for nucleoid functions. *Plant Physiol* **158,** 156-189.

Maki, H., et Sekiguchi, M. (1992). MutT protein specifically hydrolyses a potent mutagenic substrate for DNA synthesis. *Nature* **355,** 273-275.

Maluf, N.K., Fischer, C.J., et Lohman, T.M. (2003). A Dimer of Escherichia coli UvrD is the active form of the helicase in vitro. *J Mol Biol* **325,** 913-935.

Mannuss, A., Dukowic-Schulze, S., Suer, S., Hartung, F., Pacher, M., et Puchta, H. (2010). RAD5A, RECQ4A, and MUS81 have specific functions in homologous recombination and define different pathways of DNA repair in Arabidopsis thaliana. *Plant Cell* **22,** 3318-3330.

Marechal, A., Parent, J.S., Sabar, M., Veronneau-Lafortune, F., Abou-Rached, C., et Brisson, N. (2008). Overexpression of mtDNA-associated AtWhy2 compromises mitochondrial function. *BMC Plant Biol* **8,** 42.

Marechal, A., Parent, J.S., Veronneau-Lafortune, F., Joyeux, A., Lang, B.F., et Brisson, N. (2009). Whirly proteins maintain plastid genome stability in Arabidopsis. *Proc Natl Acad Sci U S A* **106,** 14693-14698.

Marechal, A., et Brisson, N. (2010). Recombination and the maintenance of plant organelle genome stability. *New Phytol* **186,** 299-317.
Mbantenkhu, M., Wang, X., Nardozzi, J.D., Wilkens, S., Hoffman, E., Patel, A., Cosgrove, M.S., et Chen, X.J. (2011). Mgm101 is a Rad52-related protein required for mitochondrial DNA recombination. *J Biol Chem* **286,** 42360-42370.
McGlynn, P., Al-Deib, A.A., Liu, J., Marians, K.J., et Lloyd, R.G. (1997). The DNA replication protein PriA and the recombination protein RecG bind D-loops. *J Mol Biol* **270,** 212-221.
McGlynn, P., et Lloyd, R.G. (2001). Rescue of stalled replication forks by RecG: simultaneous translocation on the leading and lagging strand templates supports an active DNA unwinding model of fork reversal and Holliday junction formation. *Proc Natl Acad Sci U S A* **98,** 8227-8234.
McGlynn, P., et Lloyd, R.G. (2002a). Genome stability and the processing of damaged replication forks by RecG. *Trends Genet* **18,** 413-419.
McGlynn, P., et Lloyd, R.G. (2002b). Recombinational repair and restart of damaged replication forks. *Nat Rev Mol Cell Biol* **3,** 859-870.
McVey, M., et Lee, S.E. (2008). MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends Genet* **24,** 529-538.
Mellon, I., Spivak, G., et Hanawalt, P.C. (1987). Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene. *Cell* **51,** 241-249.
Mellon, I., et Hanawalt, P.C. (1989). Induction of the Escherichia coli lactose operon selectively increases repair of its transcribed DNA strand. *Nature* **342,** 95-98.
Menand, B., Marechal-Drouard, L., Sakamoto, W., Dietrich, A., et Wintz, H. (1998). A single gene of chloroplast origin codes for mitochondrial and chloroplastic methionyl-tRNA synthetase in Arabidopsis thaliana. *Proc Natl Acad Sci U S A* **95,** 11014-11019.
Michel, B., et Leach, D. (2012). Homologous Recombination-Enzymes and Pathways. *EcoSal Plus* **5**.
Mileshina, D., Koulintchenko, M., Konstantinov, Y., et Dietrich, A. (2011). Transfection of plant mitochondria and in organello gene integration. *Nucleic Acids Res* **39,** e115.
Miller-Messmer, M., Kuhn, K., Bichara, M., Le Ret, M., Imbault, P., et Gualberto, J.M. (2012). RecA-dependent DNA repair results in increased heteroplasmy of the Arabidopsis mitochondrial genome. *Plant Physiol* **159,** 211-226.
Mitchell, D.L., et Nairn, R.S. (1989). The biology of the (6-4) photoproduct. *Photochem Photobiol* **49,** 805-819.
Modrich, P. (2006). Mechanisms in eukaryotic mismatch repair. *J Biol Chem* **281,** 30305-30309.
Molinier, J., Ramos, C., Fritsch, O., et Hohn, B. (2004a). CENTRIN2 modulates homologous recombination and nucleotide excision repair in Arabidopsis. *Plant Cell* **16,** 1633-1643.
Molinier, J., Stamm, M.E., et Hohn, B. (2004b). SNM-dependent recombinational repair of oxidatively induced DNA damage in Arabidopsis thaliana. *EMBO Rep* **5,** 994-999.
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.
Morgante, P.G., Berra, C.M., Nakabashi, M., Costa, R.M., Menck, C.F., et Van Sluys, M.A. (2005). Functional XPB/RAD25 redundancy in Arabidopsis genome: characterization of AtXPB2 and expression analysis. *Gene* **344,** 93-103.

Morimatsu, K., et Kowalczykowski, S.C. (2003). RecFOR proteins load RecA protein onto gapped DNA to accelerate DNA strand exchange: a universal step of recombinational repair. *Mol Cell* **11,** 1337-1347.

Morimatsu, K., et Kowalczykowski, S.C. (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, J.P., Case, A.L., Floro, E.R., et Willis, J.H. (2012). Evidence against equimolarity of large repeat arrangements and a predominant master circle structure of the mitochondrial genome from a monkeyflower (Mimulus guttatus) lineage with cryptic CMS. *Genome Biol Evol* **4,** 670-686.

Nakabeppu, Y., Miyata, T., Kondo, H., Iwanaga, S., et Sekiguchi, M. (1984). Structure and expression of the alkA gene of Escherichia coli involved in adaptive response to alkylating agents. *J Biol Chem* **259,** 13730-13736.

Nakagawa, T., Kurose, T., Hino, T., Tanaka, K., Kawamukai, M., Niwa, Y., Toyooka, K., Matsuoka, K., Jinbo, T., et Kimura, T. (2007). Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation. *J Biosci Bioeng* **104,** 34-41.

Nakajima, S., Sugiyama, M., Iwai, S., Hitomi, K., Otoshi, E., Kim, S.T., Jiang, C.Z., Todo, T., Britt, A.B., et Yamamoto, K. (1998). Cloning and characterization of a gene (UVR3) required for photorepair of 6-4 photoproducts in Arabidopsis thaliana. *Nucleic Acids Res* **26,** 638-644.

Nghiem, Y., Cabrera, M., Cupples, C.G., et Miller, J.H. (1988). The mutY gene: a mutator locus in Escherichia coli that generates G.C----T.A transversions. *Proc Natl Acad Sci U S A* **85,** 2709-2713.

Odahara, M., Masuda, Y., Sato, M., Wakazaki, M., Harada, C., Toyooka, K., et Sekine, Y. (2015). RECG maintains plastid and mitochondrial genome stability by suppressing extensive recombination between short dispersed repeats. *PLoS Genet* **11,** e1005080.

Oldenburg, D.J., et Bendich, A.J. (1996). Size and Structure of Replicating Mitochondrial DNA in Cultured Tobacco Cells. *Plant Cell* **8,** 447-461.

Oldenburg, D.J., et Bendich, A.J. (1998). The structure of mitochondrial DNA from the liverwort, Marchantia polymorpha. *J Mol Biol* **276,** 745-758.

Oldenburg, D.J., et Bendich, A.J. (2001). Mitochondrial DNA from the liverwort Marchantia polymorpha: circularly permuted linear molecules, head-to-tail concatemers, and a 5' protein. *J Mol Biol* **310,** 549-562.

Olsson, M., et Lindahl, T. (1980). Repair of alkylated DNA in Escherichia coli. Methyl group transfer from O6-methylguanine to a protein cysteine residue. *J Biol Chem* **255,** 10569-10571.

Orren, D.K., et Sancar, A. (1989). The (A)BC excinuclease of Escherichia coli has only the UvrB and UvrC subunits in the incision complex. *Proc Natl Acad Sci U S A* **86,** 5237-5241.

Orren, D.K., Selby, C.P., Hearst, J.E., et Sancar, A. (1992). Post-incision steps of nucleotide excision repair in Escherichia coli. Disassembly of the UvrBC-DNA complex by helicase II and DNA polymerase I. *J Biol Chem* **267,** 780-788.

Pages, V., et Fuchs, R.P. (2003). Uncoupling of leading- and lagging-strand DNA replication during lesion bypass *in vivo*. *Science* **300,** 1300-1303.

Palmer, J.D., et Shields, C.R. (1984). Tripartite Structure of the Brassica-Campestris Mitochondrial Genome. *Nature* **307,** 437-440.

Palmer, J.D., et Herbon, L.A. (1988). Plant mitochondrial DNA evolves rapidly in structure, but slowly in sequence. *J Mol Evol* **28,** 87-97.

Parent, J.S., Lepage, E., et Brisson, N. (2011). Divergent Roles for the Two PolI-Like Organelle DNA Polymerases of Arabidopsis. *Plant Physiol* **156,** 254-262.

Park, J.S., Marr, M.T., et Roberts, J.W. (2002). E. coli Transcription repair coupling factor (Mfd protein) rescues arrested complexes by promoting forward translocation. *Cell* **109,** 757-767.

Pavlov, Y.I., Mian, I.M., et Kunkel, T.A. (2003). Evidence for preferential mismatch repair of lagging strand DNA replication errors in yeast. *Curr Biol* **13,** 744-748.

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.

Putnam, C.D., Clancy, S.B., Tsuruta, H., Gonzalez, S., Wetmur, J.G., et Tainer, J.A. (2001). Structure and mechanism of the RuvB Holliday junction branch migration motor. *J Mol Biol* **311,** 297-310.

Rafferty, J.B., Sedelnikova, S.E., Hargreaves, D., Artymiuk, P.J., Baker, P.J., Sharples, G.J., Mahdi, A.A., Lloyd, R.G., et Rice, D.W. (1996). Crystal structure of DNA recombination protein RuvA and a model for its binding to the Holliday junction. *Science* **274,** 415-421.

Ray Chaudhuri, A., Hashimoto, Y., Herrador, R., Neelsen, K.J., Fachinetti, D., Bermejo, R., Cocito, A., Costanzo, V., et Lopes, M. (2012). Topoisomerase I poisoning results in PARP-mediated replication fork reversal. *Nat Struct Mol Biol* **19,** 417-423.

Rissel, D., Losch, J., et Peiter, E. (2014). The nuclear protein Poly(ADP-ribose) polymerase 3 (AtPARP3) is required for seed storability in Arabidopsis thaliana. *Plant Biol (Stuttg)* **16,** 1058-1064.

Roe, S.M., Barlow, T., Brown, T., Oram, M., Keeley, A., Tsaneva, I.R., et Pearl, L.H. (1998). Crystal structure of an octameric RuvA-Holliday junction complex. *Mol Cell* **2,** 361-372.

Roldan-Arjona, T., Garcia-Ortiz, M.V., Ruiz-Rubio, M., et Ariza, R.R. (2000). cDNA cloning, expression and functional characterization of an Arabidopsis thaliana homologue of the Escherichia coli DNA repair enzyme endonuclease III. *Plant Mol Biol* **44,** 43-52.

Rudolph, C.J., Upton, A.L., et Lloyd, R.G. (2009). Replication fork collisions cause pathological chromosomal amplification in cells lacking RecG DNA translocase. *Mol Microbiol* **74,** 940-955.

Rudolph, C.J., Upton, A.L., Stockum, A., Nieduszynski, C.A., et Lloyd, R.G. (2013). Avoiding chromosome pathology when replication forks collide. *Nature* **500,** 608-611.

Sakamoto, W., Kondo, H., Murata, M., et Motoyoshi, F. (1996). Altered mitochondrial gene expression in a maternal distorted leaf mutant of Arabidopsis induced by chloroplast mutator. *Plant Cell* **8,** 1377-1390.

Sali, A., et Blundell, T.L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* **234,** 779-815.

Samach, A., Melamed-Bessudo, C., Avivi-Ragolski, N., Pietrokovski, S., et Levy, A.A. (2011). Identification of plant RAD52 homologs and characterization of the Arabidopsis thaliana RAD52-like genes. *Plant Cell* **23,** 4266-4279.

Sancar, A. (1996). No "End of History" for photolyases. *Science* **272,** 48-49.
Sancar, A. (2003). Structure and function of DNA photolyase and cryptochrome blue-light photoreceptors. *Chem Rev* **103,** 2203-2237.
Sancar, A., et Reardon, J.T. (2004). Nucleotide excision repair in E. coli and man. *Adv Protein Chem* **69,** 43-71.
Sanford, J.C., Smith, F.D., et Russell, J.A. (1993). Optimizing the biolistic process for different biological applications. *Methods Enzymol* **217,** 483-509.
Sanger, F., Nicklen, S., et Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74,** 5463-5467.
Santerre, A., et Britt, A.B. (1994). Cloning of a 3-methyladenine-DNA glycosylase from Arabidopsis thaliana. *Proc Natl Acad Sci U S A* **91,** 2240-2244.
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.
Schofield, M.J., et Hsieh, P. (2003). DNA mismatch repair: molecular mechanisms and biological function. *Annu Rev Microbiol* **57,** 579-608.
Schormann, N., Ricciardi, R., et Chattopadhyay, D. (2014). Uracil-DNA glycosylases-structural and functional perspectives on an essential family of DNA repair enzymes. *Protein Sci* **23,** 1667-1685.
Schropfer, S., Kobbe, D., Hartung, F., Knoll, A., et Puchta, H. (2014). Defining the roles of the N-terminal region and the helicase activity of RECQ4A in DNA repair and homologous recombination in Arabidopsis. *Nucleic Acids Res* **42,** 1684-1697.
Sedgwick, B., et Vaughan, P. (1991). Widespread adaptive response against environmental methylating agents in microorganisms. *Mutat Res* **250,** 211-221.
Sedgwick, B., Robins, P., et Lindahl, T. (2006). Direct removal of alkylation damage from DNA by AlkB and related DNA dioxygenases. *Methods Enzymol* **408,** 108-120.
Seeliger, K., Dukowic-Schulze, S., Wurz-Wildersinn, R., Pacher, M., et Puchta, H. (2012). BRCA2 is a mediator of RAD51- and DMC1-facilitated homologous recombination in Arabidopsis thaliana. *New Phytol* **193,** 364-375.
Selby, C.P., Witkin, E.M., 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.
Selby, C.P., et Sancar, A. (1995). Structure and function of transcription-repair coupling factor. II. Catalytic properties. *J Biol Chem* **270,** 4890-4895.
Sen, D., Patel, G., et Patel, S.S. (2016). Homologous DNA strand exchange activity of the human mitochondrial DNA helicase TWINKLE. *Nucleic Acids Res* 10.1093/nar/gkw098.
Shaked, H., Avivi-Ragolsky, N., et Levy, A.A. (2006). Involvement of the Arabidopsis SWI2/SNF2 chromatin remodeling gene family in DNA damage response and recombination. *Genetics* **173,** 985-994.
Shedge, V., Arrieta-Montiel, M., Christensen, A.C., et Mackenzie, S.A. (2007). Plant mitochondrial recombination surveillance requires unusual RecA and MutS homologs. *Plant Cell* **19,** 1251-1264.
Shedge, V., Davila, J., Arrieta-Montiel, M.P., Mohammed, S., et Mackenzie, S.A. (2010). Extensive rearrangement of the Arabidopsis mitochondrial genome elicits cellular conditions for thermotolerance. *Plant Physiol* **152,** 1960-1970.
Shereda, R.D., Kozlov, A.G., Lohman, T.M., Cox, M.M., et Keck, J.L. (2008). SSB as an organizer/mobilizer of genome maintenance complexes. *Crit Rev Biochem Mol Biol* **43,** 289-318.
Shutt, T.E., et Gray, M.W. (2006). Twinkle, the mitochondrial replicative DNA helicase, is widespread in the eukaryotic radiation and may also be the mitochondrial DNA primase in most eukaryotes. *J Mol Evol* **62,** 588-599.

Singleton, M.R., Scaife, S., et Wigley, D.B. (2001). Structural analysis of DNA replication fork reversal by RecG. *Cell* **107,** 79-89.

Singleton, M.R., Dillingham, M.S., et Wigley, D.B. (2007). Structure and mechanism of helicases and nucleic acid translocases. *Annu Rev Biochem* **76,** 23-50.

Sloan, D.B., Alverson, A.J., Wu, M., Palmer, J.D., et Taylor, D.R. (2012). Recent acceleration of plastid sequence and structural evolution coincides with extreme mitochondrial divergence in the angiosperm genus Silene. *Genome Biol Evol* **4,** 294-306.

Sloan, D.B. (2013). One ring to rule them all? Genome sequencing provides new insights into the 'master circle' model of plant mitochondrial DNA structure. *New Phytol* **200,** 978-985.

Slupphaug, G., Eftedal, I., Kavli, B., Bharati, S., Helle, N.M., Haug, T., Levine, D.W., et Krokan, H.E. (1995). Properties of a recombinant human uracil-DNA glycosylase from the UNG gene and evidence that UNG encodes the major uracil-DNA glycosylase. *Biochemistry* **34,** 128-138.

Small, I., Suffolk, R., et Leaver, C.J. (1989). Evolution of plant mitochondrial genomes via substoichiometric intermediates. *Cell* **58,** 69-76.

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

Small, I.D., Isaac, P.G., et Leaver, C.J. (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.

Song, J., Keppler, B.D., Wise, R.R., et Bent, A.F. (2015). PARP2 Is the Predominant Poly(ADP-Ribose) Polymerase in Arabidopsis DNA Damage and Immune Responses. *PLoS Genet* **11,** e1005200.

Steinum, A.L., et Seeberg, E. (1986). Nucleotide sequence of the tag gene from Escherichia coli. *Nucleic Acids Res* **14,** 3763-3772.

Sukhanova, M., Khodyreva, S., et Lavrik, O. (2010). Poly(ADP-ribose) polymerase 1 regulates activity of DNA polymerase beta in long patch base excision repair. *Mutat Res* **685,** 80-89.

Sunderland, P.A., West, C.E., Waterworth, W.M., et Bray, C.M. (2006). An evolutionarily conserved translation initiation mechanism regulates nuclear or mitochondrial targeting of DNA ligase 1 in Arabidopsis thaliana. *Plant J* **47,** 356-367.

Sutherland, J.C., et Sutherland, B.M. (1975). Human photoreactivating enzyme action spectrum and safelight conditions. *Biophys J* **15,** 435-440.

Swartzlander, D.B., Griffiths, L.M., Lee, J., Degtyareva, N.P., Doetsch, P.W., et Corbett, A.H. (2010). Regulation of base excision repair: Ntg1 nuclear and mitochondrial dynamic localization in response to genotoxic stress. *Nucleic Acids Res* **38,** 3963-3974.

Symington, L.S. (2014). End resection at double-strand breaks: mechanism and regulation. *Cold Spring Harb Perspect Biol* **6**.

Talpaert-Borle, M., et Liuzzi, M. (1982). Base-excision repair in carrot cells. Partial purification and characterization of uracil-DNA glycosylase and apurinic/apyrimidinic endodeoxyribonuclease. *Eur J Biochem* **124,** 435-440.

Tanz, S.K., Castleden, I., Hooper, C.M., Vacher, M., Small, I., et Millar, H.A. (2013). SUBA3: a database for integrating experimentation and prediction to define the SUBcellular location of proteins in Arabidopsis. *Nucleic Acids Res* **41,** D1185-1191.

Tchou, J., Kasai, H., Shibutani, S., Chung, M.H., Laval, J., Grollman, A.P., et Nishimura, S. (1991). 8-oxoguanine (8-hydroxyguanine) DNA glycosylase and its substrate specificity. *Proc Natl Acad Sci U S A* **88,** 4690-4694.

Temple, M.D., Perrone, G.G., et Dawes, I.W. (2005). Complex cellular responses to reactive oxygen species. *Trends Cell Biol* **15,** 319-326.

Touzet, P., et Meyer, E.H. (2014). Cytoplasmic male sterility and mitochondrial metabolism in plants. *Mitochondrion* **19 Pt B,** 166-171.

Trautinger, B.W., Jaktaji, R.P., Rusakova, E., et Lloyd, R.G. (2005). RNA polymerase modulators and DNA repair activities resolve conflicts between DNA replication and transcription. *Mol Cell* **19,** 247-258.

Udy, D.B., Belcher, S., Williams-Carrier, R., Gualberto, J.M., et Barkan, A. (2012). Effects of reduced chloroplast gene copy number on chloroplast gene expression in maize. *Plant Physiol* **160,** 1420-1431.

Unseld, M., Marienfeld, J.R., 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, F.A., Molthoff, J.W., Conner, A.J., Nap, J.P., Pereira, A., et Stiekema, W.J. (1995). pBINPLUS: an improved plant transformation vector based on pBIN19. *Transgenic Res* **4,** 288-290.

Varshney, U., Hutcheon, T., et van de Sande, J.H. (1988). Sequence analysis, expression, and conservation of Escherichia coli uracil DNA glycosylase and its gene (ung). *J Biol Chem* **263,** 7776-7784.

Vaughan, P., Sedgwick, B., Hall, J., Gannon, J., et Lindahl, T. (1991). Environmental mutagens that induce the adaptive response to alkylating agents in Escherichia coli. *Carcinogenesis* **12,** 263-268.

Verhoeven, E.E., van Kesteren, M., Moolenaar, G.F., Visse, R., et Goosen, N. (2000). Catalytic sites for 3' and 5' incision of Escherichia coli nucleotide excision repair are both located in UvrC. *J Biol Chem* **275,** 5120-5123.

Vermel, M., Guermann, B., Delage, L., Grienenberger, J.M., Marechal-Drouard, L., et Gualberto, J.M. (2002). A family of RRM-type RNA-binding proteins specific to plant mitochondria. *Proc Natl Acad Sci U S A* **99,** 5866-5871.

Vincent, S.D., Mahdi, A.A., et Lloyd, R.G. (1996). The RecG branch migration protein of Escherichia coli dissociates R-loops. *J Mol Biol* **264,** 713-721.

Vonarx, E.J., Mitchell, H.L., Karthikeyan, R., Chatterjee, I., et Kunz, B.A. (1998). DNA repair in higher plants. *Mutat Res* **400,** 187-200.

Wagle, M.A., Martinville, L.E., et D'Souza, G.G. (2011). The utility of an isolated mitochondrial fraction in the preparation of liposomes for the specific delivery of bioactives to mitochondria in live mammalian cells. *Pharm Res* **28,** 2790-2796.

Wall, M.K., Mitchenall, L.A., et Maxwell, A. (2004). Arabidopsis thaliana DNA gyrase is targeted to chloroplasts and mitochondria. *Proc Natl Acad Sci U S A*.

Wallace, D.C. (1992). Mitochondrial genetics: a paradigm for aging and degenerative diseases? *Science* **256,** 628-632.

Wallet, C., Le Ret, M., Bergdoll, M., Bichara, M., Dietrich, A., et Gualberto, J.M. (2015). The RECG1 DNA Translocase Is a Key Factor in Recombination Surveillance, Repair, and Segregation of the Mitochondrial DNA in Arabidopsis. *Plant Cell* 10.1105/tpc.15.00680.

Wang, D.Y., Zhang, Q., Liu, Y., Lin, Z.F., Zhang, S.X., Sun, M.X., et Sodmergen. (2010). The levels of male gametic mitochondrial DNA are highly regulated in angiosperms with regard to mitochondrial inheritance. *Plant Cell* **22,** 2402-2416.

Whitby, M.C., Ryder, L., et Lloyd, R.G. (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.

Wigley, D.B. (2013). Bacterial DNA repair: recent insights into the mechanism of RecBCD, AddAB and AdnAB. *Nat Rev Microbiol* **11,** 9-13.

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.

Wyatt, M.D., et Pittman, D.L. (2006). Methylating agents and DNA repair responses: Methylated bases and sources of strand breaks. *Chem Res Toxicol* **19,** 1580-1594.

Xu, Y.Z., Arrieta-Montiel, M.P., Virdi, K.S., de Paula, W.B., Widhalm, J.R., Basset, G.J., Davila, J.I., Elthon, T.E., Elowsky, C.G., Sato, S.J., Clemente, T.E., et Mackenzie, S.A. (2011). MutS HOMOLOG1 is a nucleoid protein that alters mitochondrial and plastid properties and plant response to high light. *Plant Cell* **23,** 3428-3441.

Yamada, K., Kunishima, N., Mayanagi, K., Ohnishi, T., Nishino, T., Iwasaki, H., Shinagawa, H., et Morikawa, K. (2001). Crystal structure of the Holliday junction migration motor protein RuvB from Thermus thermophilus HB8. *Proc Natl Acad Sci U S A* **98,** 1442-1447.

Yamada, Y., et Harashima, H. (2013). Enhancement in selective mitochondrial association by direct modification of a mitochondrial targeting signal peptide on a liposomal based nanocarrier. *Mitochondrion* **13,** 526-532.

Yanisch-Perron, C., Vieira, J., et Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33,** 103-119.

Ye, J., Osborne, A.R., Groll, M., et Rapoport, T.A. (2004). RecA-like motor ATPases--lessons from structures. *Biochim Biophys Acta* **1659,** 1-18.

Yeeles, J.T., et Dillingham, M.S. (2010). The processing of double-stranded DNA breaks for recombinational repair by helicase-nuclease complexes. *DNA Repair (Amst)* **9,** 276-285.

Yeeles, J.T., et Marians, K.J. (2011). The Escherichia coli replisome is inherently DNA damage tolerant. *Science* **334,** 235-238.

Yeeles, J.T., Poli, J., Marians, K.J., et Pasero, P. (2013). Rescuing stalled or damaged replication forks. *Cold Spring Harb Perspect Biol* **5,** a012815.

Yoon, J.H., Park, J., Conde, J., Wakamiya, M., Prakash, L., et Prakash, S. (2015). Rev1 promotes replication through UV lesions in conjunction with DNA polymerases eta, iota, and kappa but not DNA polymerase zeta. *Genes Dev* **29,** 2588-2602.

Youvan, D.C., et Michel-Beyerle, M.E. (1996). Structure and fluorescence mechanism of GFP. *Nat Biotechnol* **14,** 1219-1220.

Zaegel, V., Guermann, B., Le Ret, M., Andres, C., Meyer, D., Erhardt, M., Canaday, J., Gualberto, J.M., 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.

Zhang, J., Mahdi, A.A., Briggs, G.S., et Lloyd, R.G. (2010). Promoting and avoiding recombination: contrasting activities of the Escherichia coli RuvABC Holliday junction resolvase and RecG DNA translocase. *Genetics* **185,** 23-37.

Zhou, J., Liu, L., et Chen, J. (2010). Mitochondrial DNA heteroplasmy in Candida glabrata after mitochondrial transformation. *Eukaryot Cell* **9,** 806-814.

Zhu, A., Guo, W., Jain, K., et Mower, J.P. (2014). Unprecedented heterogeneity in the synonymous substitution rate within a plant genome. *Mol Biol Evol* **31,** 1228-1236.

Zou, Y., et Van Houten, B. (1999). Strand opening by the UvrA(2)B complex allows dynamic recognition of DNA damage. *EMBO J* **18,** 4889-4901.

Amorces pour l'import *in organello*		
Amorce	Séquence (5' vers 3')	Fonction
ATP9-5 dir	AGAGAAGGGCAGCATTTATGAG	Construction import 2,1 kb
ATP9-3 rev	CATTCCTCAGATTACCCAGGAG	Construction import 2,1 kb
P794	TGCAAAATCAATAGGTGCCGGA	Construction import 1,3 kb
P1393	ATTCAATCCAGCCGCAGGTTCC	Construction import 1,3 kb
P442	CAAAATCTTCCTTGTTAGTTACTC	Construction import 0,7 kb
P396	CGAGAGTTACTCCACGTTGAT	Construction import 0,7 kb
Poly dir	AACGACGGCCAGTGAATTGTA	Construction MCS et PCR directe 3' (d2)
Poly rev	AAGCTCGAAATTAACCCTCAC	Construction MCS et PCR directe 5' (r1) et inverse (d3)
P395	TAGAGGTTGTCATTGAGTTGA	PCR directe 5' (d1)
ATP9 D3	AATGAAGAGGTGGGCTATGAAG	PCR directe 3' (r2)
ATP9 inv4	ATACACCATGGATGGATACATACGAT	PCR inverse contrôle
ATP9 inv6	AGATCAGGTCAGTTTGGAGGAG	PCR inverse contrôle
ATP9 inv2	TCACTTTCTCATTCCAGCACAC	PCR inverse (r3)
Amorces pour l'étude de *pol1B recG1*		
Amorce	Séquence (5' vers 3')	Fonction
RecG1-1F	GGAATCTCGTGTGTAATGGTTTC	vérification insertion T-DNA dans-RecG1
RecG1-1R	TCTACAGACACGTCAATCGAGC	vérification insertion T-DNA dans-RecG1
A127	AACGTCCGCAATGTGTTATTAAGTTGTC	vérification insertion T-DNA dans-RecG1
PolG1-R2	AGCAAGATGTGCCAGAATCC	vérification insertion T-DNA dans Pol1B
PolG1-F	GTTAGCTTGAATGGAGATGCAATC	vérification insertion T-DNA dans Pol1B
LB-1	CGCCTATAAATACGACGGATCGTA	vérification insertion T-DNA dans Pol1B
Amorce pour l'étude de *trcf*		
Amorce	Séquence (5' vers 3')	Fonction
RecG2-F2	AGTAGTAGGTCTGTTGCTTGTG	Séquençage de l'insertion tDNA sur le mutant trcf-1 + vérification insertion tDNA
RecG2-R2	AGCTTGAACACCTTCTTACTC	Séquençage de l'insertion tDNA sur le mutant trcf-1 + vérification insertion tDNA
RecG2-F	CCCACACTTGTCTACATGCTTC	Séquençage de l'insertion tDNA sur les mutants trcf-2 et trcf-3
RecG2-R	TACCCAATAGATTCGCAAACAG	Séquençage de l'insertion tDNA sur les mutants trcf-2 et trcf-3
SAIL-LB3	TAGCATCTGAATTTCATAACCAATCTCGATACAC	vérification insertion T-DNA dans TRCF
Couple P1-F	GAAATGTCGGTATCGATCTCTTCT	qRT-PCR pour quantification du transcrit de TRCF
Couple P1-R	GAGGTCATATGGAACCGAAAAA	qRT-PCR pour quantification du transcrit de TRCF
Couple P2-F	TGGATTCCGGGATGCTAGT	qRT-PCR pour quantification du transcrit de TRCF
Couple P2-R	GGAATGACGAAAGATGGGTTT	qRT-PCR pour quantification du transcrit de TRCF

Couple P3-F	TGAGCCCCTCGAATATGTCT	qRT-PCR pour quantification du transcrit de TRCF
Couple P3-R	GCCGTTTAGTCTCATTTGGAAG	qRT-PCR pour quantification du transcrit de TRCF
Couple P4-R	AGCAGGAAGTGATGCATGAAGTTC	qRT-PCR pour quantification du transcrit de TRCF
Couple P4-F	TTCCATGACTTGTGACGTTTACCG	qRT-PCR pour quantification du transcrit de TRCF
Couple P5-R	GCATGAAGTTCTGACAAGCACTGG	qRT-PCR pour quantification du transcrit de TRCF
Couple P5-F	GTGACGTTTACCGAAGCTCCTTG	qRT-PCR pour quantification du transcrit de TRCF
Actine-F	CTTGCACCAAGCAGCATGAA	gènes de références pour qRT-PCR
Actine-R	CCGATCCAGACACTGTACTTCCTT	gènes de références pour qRT-PCR
EXP-F	GAGCTGAAGTGGCTTCAATGAC	gènes de références pour qRT-PCR
EXP-R	GGTCCGACATACCCATGATCC	gènes de références pour qRT-PCR
GapDH-F	TTGGTGACAACAGGTCAAGCA	gènes de références pour qRT-PCR
GapDH-R	AAACTTGTCGCTCAATGCAATC	gènes de références pour qRT-PCR
RecG2:GFP-F	TAGGTACCTCTGGGGTTCTTGTTAAAG	Construction GFP TRCF
RecG2:GFP-R	ACGGATCCGGTCTAGCAGTCTCTCTCTTG	Construction GFP TRCF
RecG2 d1	GAAATCCTCGTCTTGGTTGAAG	Construction GUS TRCF
RecG2 r3	AGAAGAGATGATGACTTTCGGG	Construction GUS TRCF
RecG2 Cter F	ATGGAGAGATTTGCGCAAGGAA	Construction anticorps TRCF
RecG2 Cter R	ATATTTGATAAGAGCAGGAAGT	Construction anticorps TRCF
Amorces supplémentaires de PCR et qPCR: voir la publication*		

UNIVERSITÉ DE STRASBOURG

Clémentine WALLET

L'hélicase RECG1, un facteur-clé dans le maintien et la ségrégation de l'ADN mitochondrial d'*Arabidopsis thaliana*

Résumé

L'ADN mitochondrial (mtDNA) des plantes est caractérisé par les activités de recombinaison qui modulent sa structure. Ces activités sont nécessaires à son maintien et contribuent à son évolution rapide. Des facteurs contrôlant la recombinaison sont donc indispensables à la stabilité du mtDNA des plantes. Au cours de ma thèse j'ai identifié et caractérisé deux ADN hélicases présentes dans les organelles d'Arabidopsis thaliana. L'une d'entre elles est homologue à une hélicase bactérienne impliquée dans la réparation couplée à la transcription. Son rôle précis dans les organelles des plantes reste à déterminer. La deuxième hélicase, l'hélicase RECG1, a des rôles dans la réparation par recombinaison, la surveillance de la recombinaison ectopique impliquant des courtes séquences répétées, mais aussi dans la ségrégation du mtDNA. En effet, nous avons observé qu'en absence de RECG1 il y a une perte du contrôle de la recombinaison qui a pour conséquence la création de versions alternatives du mtDNA par recombinaison. L'analyse de leur ségrégation, induite par RECG1, nous a permis de modéliser comment de nouvelles configurations stables du mtDNA sont générées par le changement de stœchiométrie entre sous-génomes. Ce travail a permis de mieux comprendre les mécanismes de recombinaison et de ségrégation du mtDNA d'Arabidopsis.

Mots clés: ADN mitochondrial, Arabidopsis, recombinaison, réparation, ségrégation, hélicase RECG1.

Résumé en anglais

The mitochondrial DNA (mtDNA) of flowering plants is characterized by the recombination activities that modulate its structure. These activities are required for the mtDNA maintenance, and drive its rapid structural evolution. The factors that control recombination are therefore essential for plant mtDNA stability. During my PhD, I identified and characterized two DNA helicases that are present in the organelles of Arabidopsis thaliana. One is the homologue of a bacterial helicase involved in transcription-coupled repair. Its role in the plant organelles is still not determined. The other one, the RECG1 helicase, has roles in recombination-dependent repair, the surveillance of ectopic recombination involving short repeated sequences, and also the segregation of the mtDNA. We have found that in the absence of RECG1 there is loss of recombination control resulting in the occurrence of alternative versions of the mtDNA generated by recombination. The analysis of their segregation, induced by RECG1, allowed us to build a model to how new stable mtDNA configurations are generated by the stoichiometric shift of mtDNA sub-genomes. This work allowed us to better understand the recombination and segregation mechanisms that modulate the Arabidopsis mtDNA.

Mots clés: mitochondrial DNA, Arabidopsis, recombination, repair, segregation, helicase, RECG1.