

Université de Montréal

**Identification et caractérisation de facteurs impliqués dans la
réplication et la stabilité des génomes des organelles de plantes**

par

Jean-Sébastien Parent

Département de Biochimie

Faculté de Médecine

Thèse présentée à la Faculté des Études Supérieures en vue de l'obtention
du grade de Docteur en biochimie

Novembre 2010

© Jean-Sébastien Parent, 2010

Université de Montréal
Faculté des études supérieures

Cette thèse intitulée :

**Identification et caractérisation de facteurs impliqués dans la
réplication et la stabilité des génomes des organelles de plantes**

présentée par :

Jean-Sébastien Parent

a été évaluée par un jury composé des personnes suivantes :

Dr. Guy Boileau, président-rapporteur

Dr. Normand Brisson, directeur de recherche

Dr. Elliot Drobetsky, membre du jury

Dr. William Zerges, examinateur externe

Dr. Marc St-Arnaud, représentant du doyen de la FES

Résumé

Comparativement au génome contenu dans le noyau de la cellule de plante, nos connaissances des génomes des deux organelles de cette cellule, soit le plastide et la mitochondrie, sont encore très limitées. En effet, un nombre très restreint de facteurs impliqués dans la réplication et la réparation de l'ADN de ces compartiments ont été identifiés à ce jour. Au cours de notre étude, nous avons démontré l'implication de la famille de protéines Whirly dans le maintien de la stabilité des génomes des organelles. Des plantes mutantes pour des gènes *Whirly* chez *Arabidopsis thaliana* et *Zea mays* montrent en effet une augmentation du nombre de molécules d'ADN réarrangées dans les plastides. Ces nouvelles molécules sont le résultat d'une forme de recombinaison illégitime nommée *microhomology-mediated break-induced replication* qui, en temps normal, se produit rarement dans le plastide. Chez un mutant d'*Arabidopsis* ne possédant plus de protéines Whirly dans les plastides, ces molécules d'ADN peuvent même être amplifiées jusqu'à cinquante fois par rapport au niveau de l'ADN sauvage et causer un phénotype de variéation.

L'étude des mutants des gènes *Whirly* a mené à la mise au point d'un test de sensibilité à un antibiotique, la ciprofloxacine, qui cause des bris double brin spécifiquement au niveau de l'ADN des organelles. Le mutant d'*Arabidopsis* ne contenant plus de protéines Whirly dans les plastides est plus sensible à ce stress que la plante sauvage. L'agent chimique induit en effet une augmentation du nombre de réarrangements dans le génome du plastide. Bien qu'un autre mutant ne possédant plus de protéines Whirly dans les mitochondries ne soit pas plus sensible à la ciprofloxacine, on retrouve néanmoins plus de réarrangements dans son ADN mitochondrial que dans celui de la plante sauvage. Ces résultats suggèrent donc une implication pour les protéines Whirly dans la réparation des bris double brin de l'ADN des organelles de plantes.

Notre étude de la stabilité des génomes des organelles a ensuite conduit à la famille des protéines homologues des polymérases de l'ADN de type I bactérienne. Plusieurs groupes ont en effet suggéré que ces enzymes étaient responsables de la synthèse de l'ADN dans les plastides et les mitochondries. Nous avons apporté la preuve génétique de ce lien

grâce à des mutants des deux gènes *Poll* d'*Arabidopsis*, qui encodent des protéines hautement similaires. La mutation simultanée des deux gènes est létale et les simples mutants possèdent moins d'ADN dans les organelles des plantes en bas âge, confirmant leur implication dans la réplication de l'ADN. De plus, les mutants du gène *PollB*, mais non ceux de *PollA*, sont hypersensibles à la ciprofloxacine, suggérant une fonction dans la réparation des bris de l'ADN. En accord avec ce résultat, la mutation combinée du gène *PollB* et des gènes des protéines Whirly du plastide produit des plantes avec un phénotype très sévère. En définitive, l'identification de deux nouveaux facteurs impliqués dans le métabolisme de l'ADN des organelles nous permet de proposer un modèle simple pour le maintien de ces deux génomes.

Mots clés : Biologie végétale, maintien du génome, réplication de l'ADN, réparation de l'ADN, plastide, mitochondrie, protéines Whirly, polymérase de l'ADN.

Abstract

Compared to the nuclear genome, very little is known about the genomes of the two plant cytoplasmic organelles, the plastid and the mitochondria. Indeed, very few factors involved in either the replication or the repair of these genomes have been identified. Here we show the implication of the Whirly protein family in the maintenance of organellar DNA. Indeed, mutations in *Whirly* genes lead to DNA rearrangements in both *Arabidopsis thaliana* and *Zea mays* plastids. These rearrangements are the product of microhomology-mediated break-induced replication that rarely occurs in wild-type plants but increases in absence of Whirly proteins. In a mutant plant devoid of plastidial Whirly proteins, these new DNA molecules can be amplified up to fifty times the normal DNA level and cause a variegated phenotype.

In the course of the study of the Whirly mutant plants, we developed a strategy, based on the use of the antibiotic ciprofloxacin, to induce DNA double-strand breaks specifically in plant organelles. The *Arabidopsis* mutant plants without Whirly proteins in the plastids are more sensitive to the antibiotic ciprofloxacin than wild-type plants. Accordingly, there is a much larger increase in the number of rearranged DNA molecules in the plastids of the mutant plants than in the control plants. Surprisingly, while the mutant plants devoid of Whirly proteins in the mitochondria do not show increased sensitivity to the drug, they do accumulate more rearrangements in their mitochondrial DNA compared to wild-type plants. These results suggest that the Whirly proteins are involved in the repair of DNA double-strand breaks in the plant organelle genomes.

Our study of the plant organelle genome stability has lead us to a family of proteins homologous to the DNA polymerase I in bacteria. This family has been proposed to be responsible for most of the DNA-synthesis activity in the plant organelles. We bring genetic proof to support this hypothesis using mutants of the two *Poll* genes of *Arabidopsis*. The combined mutation of both genes is lethal and the single mutations cause a decrease in the relative DNA levels in the organelles, thus confirming the involvement of both genes in DNA replication. Interestingly, mutants of the *PollB* but not *PollA* gene shows increase sensitivity to ciprofloxacin suggesting a function in DNA repair.

In line with these results, a cross between a *PolIB* mutant and the mutant of plastid *Whirly* genes resulted in plants with severe growth defects and numerous rearrangements in the plastid DNA. In conclusion, we have identified two factors involved in the metabolism of organelle DNA and proposed a simple model of how these genomes are maintained in the plant cell.

Keywords: Plant biology, maintenance of the genome, DNA replication, DNA repair, plastid, mitochondria, Whirly proteins, DNA polymerase.

Table des matières

Résumé	iii
Abstract	v
Table des matières	vii
Liste des tableaux	ix
Liste des figures	x
Remerciements	xvi
CHAPITRE 1. Revue de la littérature	1
1.1 Origine et fonctionnement des organelles chez les plantes	2
1.1.1 Origine endosymbiotique des organelles.....	3
1.1.2 Organisation des génomes des organelles.....	5
1.1.3 Importation des protéines nécessaires aux organelles.....	11
1.2 Métabolisme de l'ADN chez les organelles	12
1.2.1 Réplication de l'ADN	12
1.2.2 Réparation de l'ADN.....	16
1.2.2.1 Réparation de mésappariements.....	17
1.2.2.2 Photoréactivation	18
1.2.2.3 Excision de nucléotides	19
1.2.2.4 Excision de bases.....	20
1.2.2.5 Recombinaison.....	21
1.2.2.5.1 Recombinaison non-homologue.....	21
1.2.2.5.2 Recombinaison homologue	24
1.2.2.5.3 Protéines impliquées dans la recombinaison de l'ADN des organelles.....	27
1.3 Les protéines Whirly	32
1.3.1 Découverte de la famille de protéines Whirly.....	32
1.3.2 Structure et conservation des protéines Whirly.....	33
1.3.3 Localisation intracellulaire.....	34
1.3.4 Rôles nucléaires	35
1.3.5 Rôles dans les organelles	36
1.4 Les polymérase de l'ADN des organelles	38
1.4.1 Isolation de l'activité de synthèse de l'ADN	38
1.4.2 Clonage des gènes des polymérase de l'ADN	39
1.4.3 Localisation intracellulaire.....	42
1.5 Survol du projet de recherche	42
1.5.1 Objectifs	42
1.5.2 Approche expérimentale	43
CHAPITRE 2. Premier article	46
2.1 Mise en contexte	47
2.2 Abstract	49
2.3 Introduction	49

2.4 Results.....	51
2.5 Discussion	55
2.6 Materials and Methods	59
2.7 Acknowledgments.....	60
2.8 References	61
2.9 Figure Legends	64
2.10 Supporting Information.....	76
2.11 Supporting Information Materials and Methods	77
2.12 References.....	83
2.13 Supporting Information Figure Legends.....	84
CHAPITRE 3. Deuxième article	106
3.1 Mise en contexte	107
3.2 Abstract.....	109
3.3 Introduction	109
3.4 Results.....	112
3.5 Discussion	120
3.6 Methods	126
3.7 Acknowledgments.....	131
3.8 References	131
3.9 Figure Legends	140
3.10 Supplemental Data	157
CHAPITRE 4. Troisième article	195
4.1 Mise en contexte	196
4.2 Summary	198
4.3 Introduction	198
4.4 Results.....	200
4.5 Discussion	205
4.6 Experimental procedures.....	209
4.7 Acknowledgements	211
4.8 References	211
4.9 Figures legends	215
4.10 Supporting Information.....	222
Chaptitre 5. Discussion.....	226
5.1 Survol.....	226
5.2 La réplication et le maintien des génomes des organelles de plante	227
5.3 Le rôle des protéines Whirly dans la stabilité des génomes.....	231
5.4 Le rôle des polymérase de l'ADN de type I dans la réplication et la réparation de l'ADN des organelles.....	234
5.5 La stabilité des génomes et l'évolution	237
Chaptitre 6. Conclusion et perspectives	239
Références	241

Liste des tableaux

Tableau I. Percentage of plants showing a variegated phenotype

Tableau II. Characteristics of amplified regions in variegated plants

Tableau III. Representative recombination events between short direct repeats in *Arabidopsis* plants lacking *ptWhirlies*

Tableau IV. Short direct repeat-mediated recombination frequency in *Arabidopsis*

Tableau V. Recombination events mediated by short direct repeats in *Arabidopsis*

Tableau VI. Recombination events mediated by short direct repeats in the maize plastid genome

Tableau VII. Representative DNA Rearrangements Observed in *Arabidopsis* Organellar Genomes

Tableau VIII. Data Collection and Refinement Statistics

Tableau IX. Analysis of the Binding of St-WHY2 to Four Different DNA Sequences by EMSA

Tableau X. DNA Rearrangement Products Identified in the Mitochondrial and Plastid Genomes

Tableau XI. Related DNA Rearrangements Detected in the Mitochondrial and Plastid Genomes

Tableau XIII. DNA Rearrangement Products Detected in the Mitochondrial and Plastid Genomes and Used as Probes in DNA Gel Blots

Tableau XIV. Oligonucleotides Used in this Study

Liste des figures

Figure 1. Origines des endosymbiotes sur une échelle de temps.

Figure 2. Transfert des gènes de l'endosymbiote vers le noyau de l'hôte.

Figure 3. Conformations du génome du chloroplaste créées par la réplication dépendante de la recombinaison.

Figure 4. Modèles de la réplication de l'ADN des organelles.

Figure 5. Modèles des mécanismes de réparation non conservateurs des bris double brin de l'ADN.

Figure 6. Modèles des mécanismes de réparation conservateurs des bris double brin de l'ADN.

Figure 7. Arbre phylogénétique des polymérases de plantes de type I.

Figure 8. AtWhy1 and AtWhy3 are involved in the biogenesis of chloroplasts.

Figure 9. Variegated plants contain rearranged amplified plastid DNA regions.

Figure 10. Illegitimate recombination is increased in the absence of *Arabidopsis* ptWhirlies.

Figure 11. Model for the appearance of variegation in plants lacking Whirlies.

Figure 12. ssDNA-binding activity of AtWhy1/3 in *Arabidopsis* ptWhirlies mutants.

Figure 13. Co-localization of StWhy1 with chloroplast DNA.

Figure 14. AtWhy1 and AtWhy3 interact with chloroplast DNA.

Figure 15. Long (16 hrs) and short (4 hrs) exposure of panel 2 from Figure 9A.

Figure 16. A gene map of the ptDNA region containing the amplified rearranged regions found in Var A and Var B plants was drawn using the web-based tool OGDRAW.

Figure 17. Absence of subgenomic amplicons in *ZmWhy1* maize mutants.

Figure 18. Levels of plastid 23S and 16S rRNA in *Arabidopsis* ptWhirlies mutants.

Figure 19. PCR Strategy to Detect DNA Rearrangements.

Figure 20. DNA Rearrangements Accumulate in Plastids Following Treatment with the Gyrase Inhibitor Ciprofloxacin.

Figure 21. DNA Rearrangements Accumulate in Mitochondria Following Treatment with the Gyrase Inhibitor Ciprofloxacin.

Figure 22. Crystal Structures of St-WHY2 in the Free Form and Bound to ERE32 at 2.2-Å Resolution.

Figure 23. Mechanism of ssDNA Binding of Whirly Proteins.

Figure 24. WHY2 Binds ssDNA with Limited Sequence Specificity.

Figure 25. WHY2 Binds Single-Stranded Overhangs, Destabilizes dsDNA, and Protects ssDNA against Nuclease Degradation.

Figure 26. Model for the Repair of Organellar Double-Strand Breaks in the Absence or Presence of Whirly Proteins.

Figure 27. Effect of Ciprofloxacin Treatment on Arabidopsis Plant Survival.

Figure 28. DNA Rearrangements Accumulate Primarily in Plastids of Plants Lacking Plastid-Targeted Whirlies and in the Mitochondria of Plants Lacking Mitochondria-Targeted Whirlies.

Figure 29. DNA Gel Blot Detection of DNA Rearrangements in Plastids Following Plant Treatment with Ciprofloxacin.

Figure 30. DNA Gel Blot Detection of DNA Rearrangements in Mitochondria Following Plant Treatment with Ciprofloxacin.

Figure 31. Analysis of Mitochondrial Recombination in Plants Treated with Ciprofloxacin.

Figure 32. Analysis of Mitochondrial Recombination in Plants Treated with Novobiocin.

Figure 33. Effects of Bleomycin on Plastid and Mitochondria DNA Rearrangements of Plants lacking Plastid- or Mitochondria-Targeted Whirlies.

Figure 34. Effects of Hydroxyurea on Plastid and Mitochondria DNA Rearrangements of Plants lacking Plastid- or Mitochondria-Targeted Whirlies.

Figure 35. Analysis of Mitochondrial Recombination in WT Plants Treated with Bleomycin or Hydroxyurea.

Figure 36. Single-Stranded DNA is Stabilized on the Edges of the β - Sheets of Whirly Proteins.

Figure 37. Modeling of a St-WHY2-RNA Complex.

Figure 38. Molecular Characterization of why2 Lines.

Figure 39. Isolation of *PollA* and *PollB* mutant lines.

Figure 40. Organelle DNA is less abundant in seedlings of *polla-1* and *pollb-1* mutants.

Figure 41. *PollB* mutants exhibit an increased sensitivity to ciprofloxacin (CIP) and are impaired in their ability to process DNA DSBs.

Figure 42. A *pollb-1/why1why3* triple mutant yields a pale-green dwarf phenotype and gives rise to abundant DNA rearrangements in plastids.

Figure 43. Organelle DNA is less abundant in seedlings of *polla-2*.

Figure 44. Ciprofloxacin does not generate DSB in the nuclear genome.

Figure 45. Crosses between *Poll* and *Whirly* mutants.

Liste des abréviations

ADN : Acide désoxyribonucléique

kb : Kilobase

ADNmt : ADN mitochondrial

RDR : Réplication dépendante de la recombinaison

ADNpt : ADN plastidique

Ori : Origine de réplication

D-loop : Boucle de déplacement

Pol : Polymérase de l'ADN

VIGS : *Virus-induced gene silencing*

MMR : *Mismatch repair*

UV : Rayons ultra-violets

Plastome : Génome plastidique

NER : Réparation par excision de nucléotides

RPA : *Replication Protein A*

BER : Réparation par excision de bases

ADNsb : ADN simple brin

NHEJ : Jonction d'extrémités non-homologues

MMEJ : Microhomology-mediated end joining

HR : Recombinaison homologue

DSBR : Réparation des bris double brin

SDSA : *Synthesis-dependent strand annealing*

BIR : Réplication induite par les bris

MMBIR : *Microhomology-mediated break-induced replication*

ATP : Adénosine triphosphate

WT : Type sauvage

OSB : Organellar Single-stranded DNA-Binding proteins

SSB : *Single-strand DNA binding*

RPA : *Replication protein A*

kDa : Kilodalton

ptWhirly : Whirly ciblant le plastide

mtWhirly : Whirly ciblant la mitochondrie

ChIP : Immunoprécipitation de la chromatine

ADNc : ADN complémentaire

ARNr : ARN ribosomal

ddNTP : didéoxyribonucléotides triphosphates

chondriome : génome mitochondrial

It is in the admission of ignorance and the admission of uncertainty that there is a hope for the continuous motion of human beings in some direction that doesn't get confined, permanently blocked, as it has so many times before in various periods in the history of man. (Richard P. Feynman)

Remerciements

Mes premiers remerciements vont au Dr. Normand Brisson que j'ai eu la chance de côtoyer ces sept dernières années. Normand, merci pour m'avoir accueilli dans ton groupe et d'avoir mis toutes les ressources nécessaires à ma disposition pour mener à bien (ou dans le mur!) mes nombreux projets. Ton calme désarment m'a appris une autre signification de la patience et de la persévérance. J'espère faire une longue carrière en science simplement pour un jour comprendre comment ton laboratoire a pu me transformer du tout au tout à travers les années. Par-dessus tout, merci pour m'avoir communiqué cette passion tranquille et néanmoins brûlante qu'est la curiosité. Je quitte ton laboratoire avec suffisamment de bons souvenirs pour être nostalgique le restant de mes jours.

Merci à mes deux vieux compagnons de route, Alexandre Maréchal et Laurent Cappadocia. Alexandre d'abord, pour m'avoir pris sous ton aile et avoir partagé généreusement ton expérience. Ensemble, nous avons vécu les succès mais aussi les frustrations. Nous avons bâti quelque chose qui me procure une fierté qui dépasse largement les épreuves que nous avons traversées. Avec toi, travailler le jour (comme la nuit) était un plaisir. Merci Laurent pour ta persévérance exemplaire mais aussi ta patience infini. Combien de fois est-ce que j'ai pu te poser la même question sur une ligne de commande enfantine et combien de fois m'as-tu donné la même réponse avec ton calme et tes mots posés? Tu es un personnage encore plus incroyable que tout ce que tu n'as jamais pu interpréter sur scène. J'espère qu'on se retrouvera dans l'avenir car je ne retrouverai jamais une personne aussi zen et généreuse que tu peux l'être.

Un merci très spécial à Étienne Lepage qui, plus que personne d'autre, m'a amené à me dépasser à tous les niveaux, que ce soit personnel ou professionnel. Même si nous n'avons travaillé ensemble que peu de temps (relativement à la longueur de mon doctorat), il me semble qu'on se connaît depuis toujours. J'aurais aimé qu'on poursuive notre travail encore des années pour ainsi profiter plus longtemps de notre amitié et de l'énergie inépuisable avec laquelle tu «infectes» tous les gens qui t'entourent. Qui sait? Peut-être

serons-nous collaborateurs dans le futur. Je ne peux pas penser à personne de mieux placé pour reprendre nos projets et les amener là où personne n'a jamais osé rêver aller.

Merci aux personnes d'expérience dont Rocio Gonzalez-Lamothe, Mathieu Coinçon et Mohammed Sabar d'avoir partager leur sagesse avec moi. Merci aux Drs. Franz Lang et Kamal Bouarab pour leurs conseils précieux. Merci à mon partenaire de hockey, Thierry Vincent, pour le plaisir qu'on a eu sur la glace et à l'extérieur. Merci également à Éric Zampini d'avoir «souffert» avec moi le temps d'un stage estival et de s'être réengagé au labo malgré tout. Thierry et Éric, nos bras meurtris vous tendent le flambeau...

Je remercie tous mes amis du département qui m'ont apporté la motivation pour travailler tous les matins pendant toutes ces années. Merci également aux employés du département de simplifier la vie des étudiants qui ont mille choses à faire en plus de leurs études. Une pensée spéciale va aux membres du CA de l'AECSBUM 2009-2010 avec qui j'ai passé une année d'enfer.

Je tiens à remercier mes amis que se sont toujours intéressés à mes recherches et plus important encore ont réussi à me les faire oubliées dans les moments de fête, de détente et les soirées de hockey. Tous ces moments m'ont permis de rester sain d'esprit. Finalement, cette thèse n'aurait pu voir le jour sans les encouragements et le support indéfectible de ma famille. Je vous aime fort, merci d'être là pour moi. Le plus gros merci de tous va à Rachel, ma douce moitié, pour m'avoir supporté à chaque heure de chaque jour pendant toutes ces années.

Merci à tous!!!

CHAPITRE 1. Revue de la littérature

Le présent travail porte sur le maintien des génomes contenus dans les organelles présentes dans les cellules de plantes, soit ceux de la mitochondrie et du plastide (ou plaste). Plus spécifiquement, ce travail concerne deux familles de gènes bien conservées chez les végétaux, la famille nommée «Whirly» et la famille de polymérase de l'ADN de type I. Les protéines encodées par ces gènes se retrouvent dans les mitochondries et les plastides en plus d'avoir un rôle dans le maintien de l'acide désoxyribonucléique (ADN) propre à ces organelles. Une introduction générale sur les mitochondries et les plastides présente l'origine des organelles et leur fonctionnement à la section 1.1 alors que la section 1.2 fait un résumé des connaissances de leur métabolisme de l'ADN. La description des deux familles de protéines faisant l'objet de la présente étude ainsi que la revue de la littérature pertinente se retrouvent à la section 1.3 dans le cas des Whirly et à la section 1.4 dans le cas des polymérase de type I. Finalement, un survol du projet de recherche est présenté à la section 1.5.

1.1 Origine et fonctionnement des organelles chez les plantes

À la différence des autres organismes vivants, les végétaux ont la particularité de devoir assurer le maintien de trois génomes différents présents dans chaque cellule afin d'assurer la réplication et la propagation de leur matériel génétique. En effet, en plus de leur génome nucléaire, les cellules de plantes possèdent deux autres génomes qui sont contenus dans les organelles que sont les mitochondries et les plastides. Comme les cellules animales, les cellules végétales contiennent des mitochondries qui sont le centre de la phosphorylation oxydative qui produit de l'énergie grâce à la combustion des composés organiques. En plus de cette fonction vitale, les mitochondries des plantes jouent un rôle essentiel dans la synthèse de certains acides aminés ainsi que la production de tetrapyrroles et de co-facteurs de vitamines (revue dans (1)).

De son côté, le plastide, et plus spécifiquement le chloroplaste, est l'organelle qui caractérise le règne végétal et qu'on ne retrouve pas dans le règne animal. Cette organelle peut jouer différents rôles dans une cellule, ce qui requiert une certaine différenciation de plusieurs types de plastes spécialisés à partir d'un type pluripotent (revue dans (2)). Les plastides pluripotents, ou proplastides, se retrouvent en effet dans les tissus réplicatifs (méristématiques) de la plante et sont à l'origine de toutes les catégories de plastes. On distingue ces différentes catégories par leur pigment (chlorophylle, carotène, etc.) ou alors par les molécules qu'elles entreposent (amidon, lipides, etc.). Le chloroplaste est sans aucun doute le plus connu des plastides, car il est le siège de la photosynthèse qui convertit l'énergie lumineuse en énergie chimique, essentielle à la vie sur terre. La particularité des mitochondries et des plastides de posséder un génome qui leur est propre est expliquée par le fait qu'ils ont tout deux une origine bactérienne. C'est à dire que les deux étaient autrefois des organismes vivants à part entière, qu'ils ont été absorbés par un autre organisme qui les a soumis à son avantage (voir section 1.1.1). Dans les sections suivantes, il sera également discuté de la nouvelle organisation des génomes de ces anciens organismes (1.1.2), de leur nouvelle machinerie d'importation de protéines (1.1.3) ainsi que du maintien de leur génome (1.1.4).

1.1.1 Origine endosymbiotique des organelles

La première formulation de l'hypothèse selon laquelle la mitochondrie était à l'origine une bactérie date de plus d'un siècle (3). Aujourd'hui, il est généralement accepté que la mitochondrie moderne est le résultat de l'invasion d'un hôte de type archea par une α -protéobactérie il y a plus de 1,5 milliard d'années (Figure 1) (revue dans (4) et (5)). Il n'est cependant pas encore établi si cette invasion s'est produite avant ou après l'apparition du noyau et de la branche des eucaryotes. Les deux organismes auraient évolué en symbiose et modifié considérablement leur génome respectif. Étant donné les grandes similarités de séquences dans l'ADN des mitochondries d'organismes évolutivement éloignés, l'hypothèse actuellement retenue stipule que tous les eucaryotes résultent de cette première endosymbiose. Bien que cette symbiose ait évolué différemment à travers les branches des eucaryotes, la tendance générale a été la perte du matériel génétique appartenant à l'ancienne α -protéobactérie au profit du génome nucléaire. Le génome des mitochondries modernes n'encode plus qu'un faible pourcentage des protéines qui leur sont nécessaires. Les mitochondries végétales ont néanmoins conservé une plus large partie du génome ancestral (plus de 200 kilobases (kb)) que les mitochondries animales (environ 16 kb chez l'humain) (revue dans (6)).

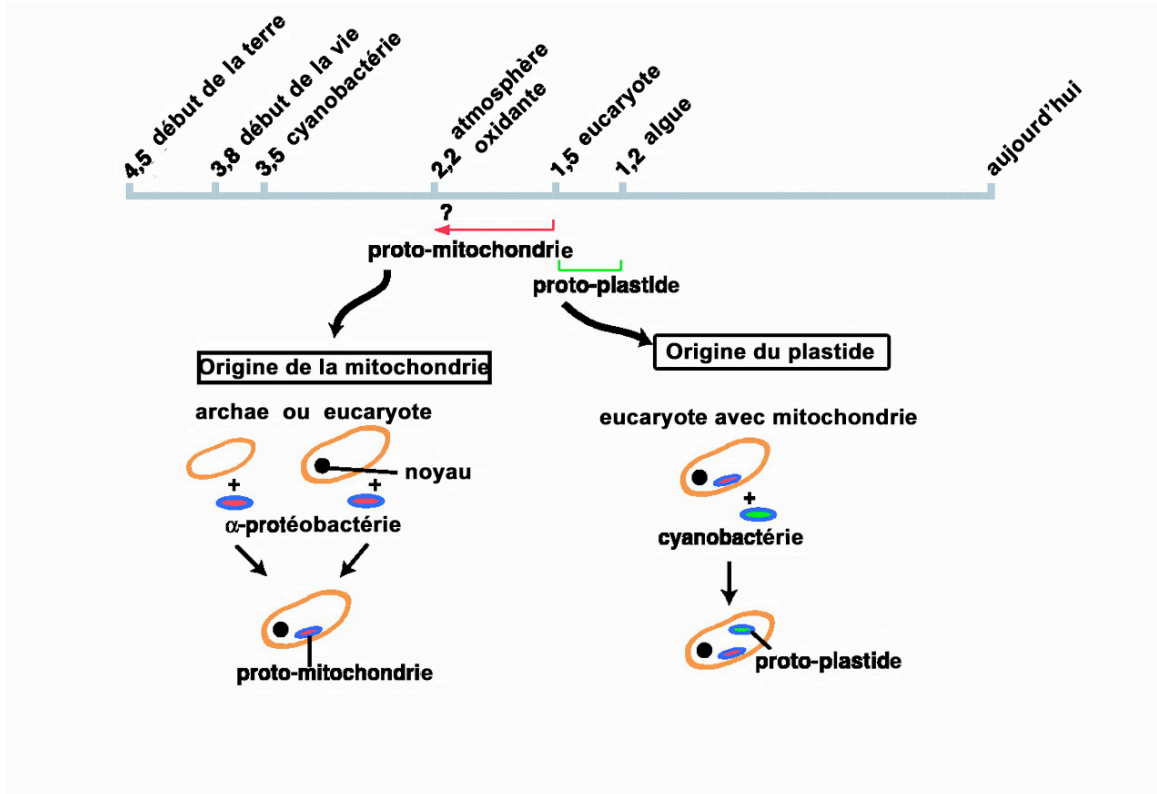


Figure 1. Origines des endosymbiotes sur une échelle de temps. Les α -protéobactéries et les mitochondries sont représentées en rouge. Les cyanobactéries et les plastides sont représentés en vert. L'échelle de temps est graduée en milliard d'années. (Figure adaptée de (5))

De leur côté, les plastes seraient apparus après la mitochondrie, entre 1,5 et 1,2 milliard d'années avant notre ère (Figure 1) (revue dans (5)). C'est à ce moment qu'un eucaryote aurait tenté d'ingérer une cyanobactérie phototrophe, mais cette dernière aurait survécu à l'intérieur du cytoplasme. Cet événement représente une endosymbiose primaire par opposition à endosymbiose secondaire qui consiste à l'assimilation par un eucaryote d'un endosymbiote primaire (7). Des produits d'endosymbiose tertiaire ont même été rapportés, mais, en ce qui concerne le présent travail, toutes les algues vertes et les plantes terrestres proviennent d'une endosymbiose primaire. Ces différents événements d'endosymbiose contrastent néanmoins avec l'invasion unique qui serait à l'origine de toutes les mitochondries modernes. Cependant, comme dans le cas de la mitochondrie, l'ancêtre des plastides a transféré une grande partie de son génome dans le noyau de son hôte et ces changements majeurs ont amené le besoin d'une nouvelle organisation de ces génomes.

1.1.2 Organisation des génomes des organelles

Bien que l'origine bactérienne des organelles soit aujourd'hui admise, les mécanismes par lesquels les procaryotes d'origine ont été convertis en organelles restent incertains. Il est néanmoins évident que les deux organismes bactériens ont transféré la majeure partie de leur génome d'origine vers le noyau de leur hôte (revue dans (8)). Ces transferts se sont produits à différents moments au cours de l'évolution et, bien que les génomes des organelles modernes soient très réduits, il semble que ce genre de transfert ait encore lieu aujourd'hui. Une hypothèse populaire stipule que la migration du matériel génétique vers le noyau se produise lorsque l'intégrité membranaire des organelles est compromise, par exemple lors d'un stress, lors de la division ou bien lors de la dégradation de l'organelle durant la gamétogenèse. L'hypothèse de la rupture des organelles est très probable puisqu'elle expliquerait par le fait même pourquoi le transfert de gènes du plastide vers le noyau ne peut être observé en laboratoire chez l'algue unicellulaire *Chlamydomonas reinhardtii* contrairement à la plante de tabac (*Nicotiana tabacum*) (Figure 2) (9). En effet, la rupture du chloroplaste unique de l'algue unicellulaire signifierait la mort de l'organisme et l'ADN doit donc être transféré dans un autre contexte dans ces cas. Lors de

la rupture de la membrane des organelles, les acides nucléiques sont libérés dans le cytosol pourraient alors être importés et intégrés dans le noyau par divers mécanismes. Une fois que les gènes ont été transférés au noyau et qu'ils mutent suffisamment pour être exprimés, traduits et ciblés vers les organelles, la copie originale se retrouve sans contrainte sélective et pourra éventuellement disparaître (revue dans (8)). Cette perte de contrainte sélective explique pourquoi les deux organelles ont perdu la grande majorité de leurs gènes alors que ceux-ci se retrouvent maintenant dans le noyau.

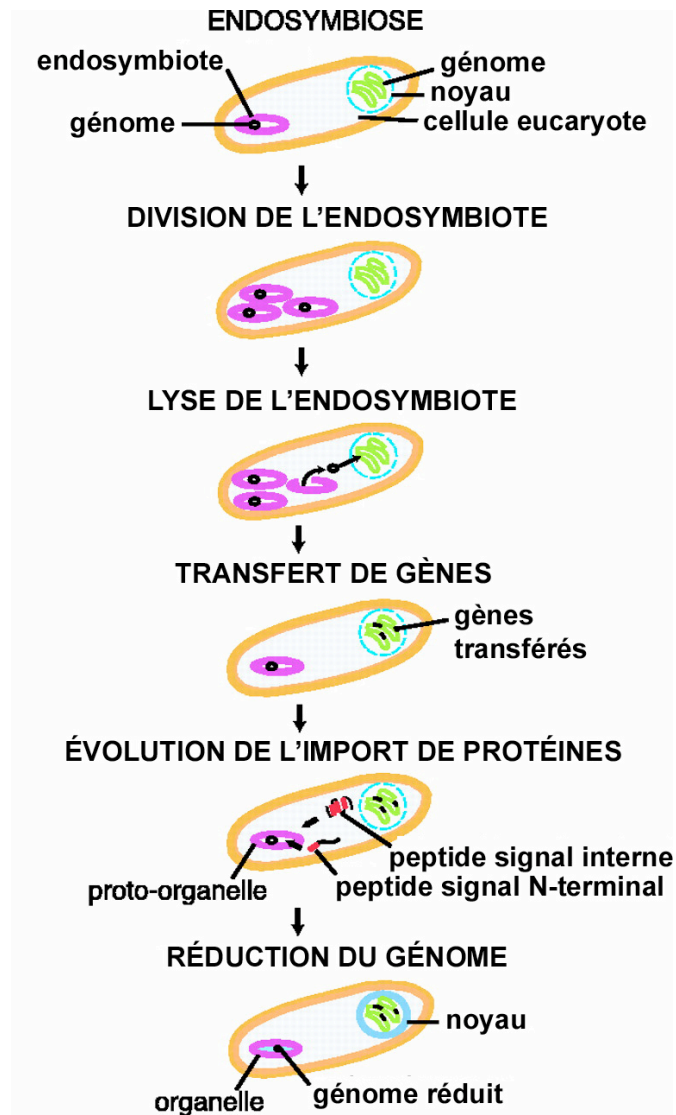


Figure 2. Transfert des gènes de l'endosymbiote vers le noyau de l'hôte. Les endosymbiotes sont représentés en rose et l'hôte en orange. Le génome d'un endosymbiote se retrouve dans le cytosol suite à la lyse de la membrane externe et est ensuite importé dans le noyau. Les nouveaux gènes nucléaires évoluent et produisent des protéines capables d'être importées dans les organelles. La redondance des gènes se trouvant dans les deux compartiments mène alors à l'élimination de la copie de l'endosymbiote. (Figure adaptée de (5))

Deux tendances ont contribué à façonner le génome des mitochondries des plantes. Comme il a été mentionné précédemment, la grande majorité des gènes d'origine ont été perdus au profit du noyau causant ainsi la réduction du génome. Des différences considérables existent entre différentes plantes à fleurs, ou angiospermes, quant à la présence des gènes des protéines du ribosome, ce qui suggère que beaucoup de transferts de gènes ont eu lieu récemment (revue dans (6)). La deuxième tendance a été l'expansion des régions intergéniques qui confère aux angiospermes les plus grands génomes mitochondriaux connus (environ de 200 à 2 500 kb (revue dans (10)). L'ADN mitochondrial (ADNmt) des angiospermes est également caractérisé par la présence de régions répétées soit directes ou inversées. Ces différentes répétitions seraient à l'origine de la grande variété de formes des molécules d'ADN qu'on observe chez la mitochondrie (revue dans (11) et (12)). En effet, la forme et la taille de ces différentes molécules sont grandement variables. On retrouve la forme circulaire classique, mais également des formes linéaires branchées de taille égale, plus grande et même plus petite qu'un génome complet. La combinaison d'une réplication dépendante de la recombinaison (RDR) avec les différentes répétitions dispersées dans l'ADNmt serait à l'origine de ces différentes formes (13).

Les différentes copies du génome sont réparties en petits groupes qui, liés à certaines protéines spécifiques, forment les nucléoïdes (14). Ceux-ci varient en nombre et en taille et leur composition exacte en protéines est encore inconnue. Ces différentes protéines semblent impliquées dans le métabolisme de l'ADN ainsi que dans d'autres voies métaboliques de la mitochondrie. Elles sont également responsables de l'ancrage des nucléoïdes à la membrane interne.

Comme la mitochondrie, le plastide ne possède plus qu'un nombre très limité de gènes (environ 80 chez les plantes à fleurs). À quelques exceptions près, il ne reste plus que des gènes codant pour des protéines impliquées dans les voies de la traduction et de la photorespiration (revue dans (15)). Comparé à la mitochondrie, le génome du plastide est plus petit en termes absolus pour un nombre semblable de gènes. La différence réside dans les régions intergéniques qui sont plus modestes en taille dans le cas de l'ADN plastidique (ADNpt). Alors que le génome mitochondrial contient de multiples répétitions,

le génome du plaste ne contient généralement qu'une seule paire de répétitions inversées. Ces deux segments d'ADN en séparent deux autres, uniques et nommés la grande région codante et la petite région codante. Les répétitions inversées sont à l'origine des deux isomères du génome qui sont observés et qui diffèrent seulement par rapport à l'orientation de la petite région codante qui se trouve entre les deux séquences répétées. Il a été proposé dans les années 1980 qu'une réaction baptisée «flip-flop» était à l'origine de ces deux isomères (16). Cette hypothèse était compatible avec les données expérimentales de l'époque ainsi que l'idée largement répandue que le génome du plaste devait être sous la forme d'un cercle fermé.

Des données plus récentes et l'observation de formes complexes linéaires et branchées ont mené à une autre hypothèse selon laquelle la RDR serait à la fois à l'origine des formes complexes et des deux isomères (Figure 3) (17). Comme chez la mitochondrie, l'ADN des plastides est regroupé en énormes complexes d'acides nucléiques et de protéines nommés nucléoïdes (revue dans (15)). Le nombre et la taille de ces complexes restent hautement variables et semblent dépendre du type de cellules et de l'état général du métabolisme. Il semble que dans le cas des plastides, les nucléoïdes soient ancrés à la membrane interne, mais également aux thylakoïdes ce qui pourrait avoir pour but de rapprocher les gènes associés à la photosynthèse de l'endroit où ils doivent être transcrits et traduits (revue dans (15)). Les deux organelles de plante possèdent donc plusieurs copies d'un génome réduit qui n'encode que quelques gènes et qui est organisé en énormes complexes de protéines et d'acides nucléiques. Le reste des protéines requises au bon fonctionnement de ces compartiments étant encodées par le noyau, les organelles nécessitent alors un système d'importation efficace des protéines.

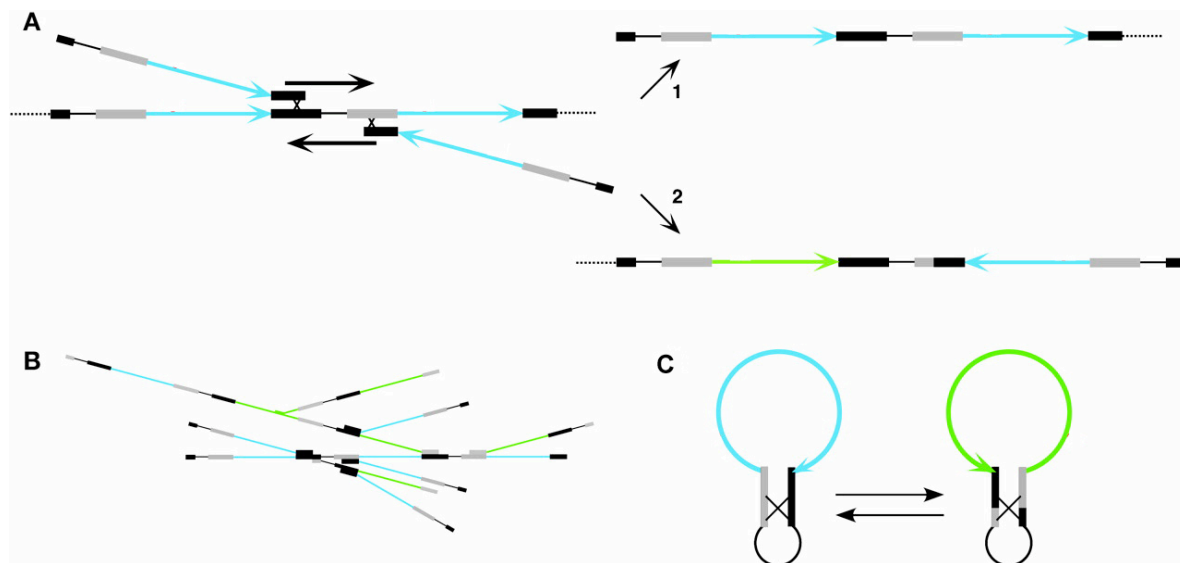


Figure 3. Conformations du génome du plastide créées par la réplication dépendante de la recombinaison. Les lignes noires et grises représentent les deux séquences répétées et inversées du plastome. (A) La fin d'une molécule d'ADN représentant l'équivalent d'un génome initie une réaction de réplication en utilisant une autre molécule comme matrice. La réaction allant vers la droite entraîne le produit de la voie 1, un concatémère tête-à-queue. La réaction allant vers la gauche se fait grâce à l'appariement d'une séquence répétée d'une molécule avec la séquence inversée d'une autre molécule. Le produit 2 possède donc une nouvelle copie du génome (en vert), mais dans une orientation inversée par rapport à la première (en bleu). (B) Structure branchée et complexe qui résulte de multiples réactions de réplication dépendante de la recombinaison. (C) Explication de la réaction de «flip-flop» du génome du chloroplaste par recombinaison intramoléculaire. (Figure adaptée de (17))

1.1.3 Importation des protéines nécessaires aux organelles

La grande majorité des protéines nécessaires au bon fonctionnement des organelles sont encodées par l'ADN du noyau et synthétisées dans le cytoplasme. Lorsqu'un gène appartenant à une organelle est transféré au noyau, il se retrouve d'abord sous forme d'ADN inerte. Il lui faut premièrement acquérir des éléments favorisant sa transcription dans le contexte nucléaire (18). La protéine peut ensuite être produite et se retrouvera fort probablement dans le cytosol. Pour que cette protéine retourne vers l'organelle dont elle origine, il lui faut donc acquérir une séquence de ciblage (ou peptide signal). Dans le cas de la mitochondrie comme dans celui du plastide, la séquence de ciblage peut varier, mais elle sera classiquement composée d'un mélange d'acides aminés chargés positivement, hydrophobes et hydroxylés (sérine et thréonine) (revue dans (19) et (20)). Cette séquence est souvent retrouvée à l'extrémité N-terminale de la protéine, mais il peut s'agir d'une séquence interne dans certains cas.

Une fois qu'elle possède toutes les caractéristiques nécessaires, la protéine sera synthétisée dans le cytosol sous forme de pré-protéine non-repliée qui sera prise en charge par des chaperones (revue dans (19) et (20)). La séquence de ciblage lui permettra d'être reconnue par les complexes de translocation. Dans le cas de la mitochondrie, la pré-protéine est d'abord reconnue par le complexe de translocase de la membrane externe de la mitochondrie (TOM) qui lui permet de pénétrer la première membrane. La pré-protéine peut alors être insérée dans la membrane externe ou bien être relâchée dans l'espace intermembranaire. Alternativement, le complexe de translocase de la membrane interne de la mitochondrie (TIM) peut permettre la pénétration de la deuxième membrane. La pré-protéine se retrouvera donc dans la membrane interne ou bien dans la matrice mitochondriale. Dans tous les cas, un clivage de la séquence de ciblage par une peptidase donne à la pré-protéine sa forme mature.

Le cas de l'importation des protéines dans les plastides comporte beaucoup de ressemblances à celui de la mitochondrie. La pré-protéine doit également passer par les translocases de la membrane externe (TOC) et de la membrane interne (TIC). Différents signaux compris dans le peptide transite dictent la localisation précise de la protéine dans

l'organelle ainsi que le(s) site(s) de coupure pour les peptidases. La grande ressemblance entre les deux systèmes d'importation a permis à l'organisme de cibler une même protéine à la fois vers les mitochondries et les plastides (revue dans (21)). Il est d'ailleurs très intéressant de constater qu'une majeure partie de ces protéines ciblées vers les deux organelles sont impliquées dans le métabolisme de l'ADN.

1.2 Métabolisme de l'ADN chez les organelles

Les organelles des plantes possèdent leur propre génome encodant pour moins d'une centaine de protéines qui sont néanmoins essentielles. Vu l'existence de ces génomes, on retrouvera donc des activités de réplication, de recombinaison et de réparation de l'ADN à l'intérieur des mitochondries et des plastides. Ici s'arrête le consensus à ce sujet, car les multiples conformations de l'ADN des organelles qui ont été observées ont donné lieu à autant d'hypothèses quant à leur mode de réplication et de leur maintien (revue dans (13, 17, 22)). Pour compliquer les choses, très peu de facteurs impliqués dans ces divers processus ont été identifiés à ce jour. L'identification et l'étude des protéines impliquées dans les voies du métabolisme de l'ADN pourraient jeter un peu de lumière sur la provenance et la fonction des différentes molécules d'ADN que l'on retrouve dans ces organelles. Étant donné l'importance de la réplication et de la réparation de l'ADN pour n'importe quel organisme, il est surprenant que si peu de facteurs impliqués dans ces processus aient été identifiés chez les organelles de plante.

1.2.1 Réplication de l'ADN

Puisque les organelles des plantes ont une origine bactérienne, il a longtemps été assumé qu'elles possédaient un génome parfaitement circulaire comme leurs ancêtres prokaryotes (revue dans (22)). Une telle organisation impliquerait un modèle de réplication bactérienne nécessitant deux origines de réplication (Ori). Des facteurs de reconnaissance de ces origines provoqueraient alors l'ouverture simultanée de l'ADN double brin pour former deux boucles de déplacement (D-loop) et donner naissance à deux fourches de réplication unidirectionnelles (Figure 4A). La rencontre des deux

fourches crée alors une bulle de réplication bidirectionnelle qui complète la duplication du génome. En accord avec ce modèle, les premières études de visualisation de l'ADN des chloroplastes en microscopie électronique ont mis en évidence des structures associées aux boucles de déplacement (23).

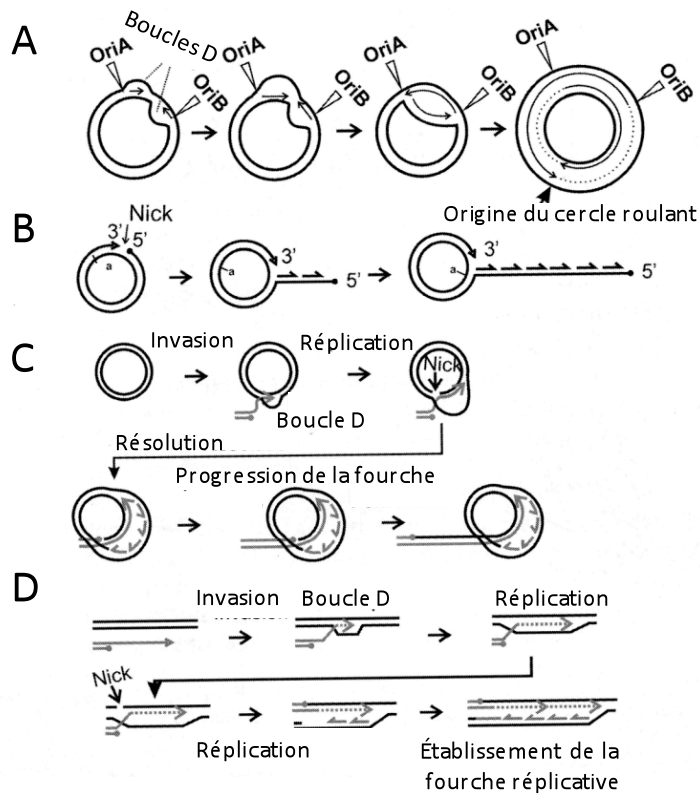


Figure 4. Modèles de la réplication de l'ADN des organelles. Représentations schématiques des différents modes de réplication qui permettent d'expliquer les différentes formes des molécules d'ADN contenues dans les organelles. (A) Modèle de réplication utilisant les origines de rélications conservées (OriA et OriB). (B) Modèle du cercle roulant initié par une brisure simple brin (nick). (C) Modèle de réplication dépendante de la recombinaison utilisant un cercle fermé comme matrice. (D) Modèle de réplication dépendante de la recombinaison utilisant une molécule linéaire comme matrice et résultant en une nouvelle molécule d'ADN branchée. (Figure adaptée de (22))

Par contre, ces dernières études ont également rapporté l'existence d'autres conformations de l'ADN comme des cercles roulants (Figure 4B), des dimères circulaires ainsi que des molécules linéaires dont la taille est inférieure à celle d'un génome complet. Par la suite, d'autres méthodes de visualisation de l'ADN ont également révélé la grande variété de conformations (circulaire, linéaire et branché) que le génome du plastide peut adopter (24, 25). De même, chez la mitochondrie de la plante, plusieurs formes de molécules d'ADN ont été observées indiquant un mécanisme de réplication certainement plus complexe que ce qui avait initialement été prédit (revue dans (11, 26)). Le modèle de réplication de l'ADN des organelles prévalant aujourd'hui propose l'existence de plusieurs modes réplicatifs dont le mode de réplication bactérienne (modèle de Cairns), le cercle roulant et un mode dépendant de la recombinaison (Figure 4C et D) comme pour la réplication de l'ADN des phages (revue dans (11) et (22)).

Les quelques études qui ont apporté des précisions sur les mécanismes de réplication de l'ADN des organelles n'ont pas permis de trancher la question du mode réplicatif prévalant. Certains groupes se sont penchés sur les Ori du génome du plastide qui semblent être relativement conservées à travers les espèces (revue dans (22)). Pourtant, la mutation des Ori du plastide de la plante de tabac semble n'avoir qu'une influence mineure sur la réplication (27, 28) suggérant ainsi l'existence de mécanismes alternatifs de démarrage de la réplication. Bien que l'identité des polymérases de l'ADN (Pol) reste à découvrir, plusieurs prédictions ont été faites. Une famille de protéines partageant une homologie forte avec les Pol bactériennes de type I est retrouvée chez les plantes (voir section 1.4) (29-31). Chaque angiosperme semble posséder deux isoformes de cette polymérase bactérienne et il semble que les deux protéines se retrouvent dans les deux organelles (31-33). Différents tests *in vitro* effectués avec des protéines recombinantes montrent que ces Pol seraient suffisamment processives pour répliquer efficacement le génome des organelles (29-31, 34). L'existence de Pol bactériennes supporterait donc le modèle de la réplication de l'ADN utilisant des origines de réplication, mais des études plus poussées seront requises avant de pouvoir tirer une conclusion définitive.

Contrairement à la mitochondrie animale, les organelles de plantes contiennent des gyrases qui sont des topoisomérases de type II d'origine bactérienne. Chez les bactéries,

ces enzymes contrôlent le surenroulement de l'ADN et aident donc à la réplication. La présence de ces enzymes chez les plantes rend ces dernières sensible aux antibiotiques ciblant les gyrases (35). Certains de ces antibiotiques, nommés quinolones, bloquent les gyrases au milieu de la catalyse du surenroulement, empêchant la religation, ce qui a comme conséquence d'introduire des bris double brin dans l'ADN (36). Une gyrase fonctionnelle est généralement retrouvée sous forme d'homotétramère composé de deux sous-unités A et deux sous-unités B. Chez *Arabidopsis thaliana*, on retrouve un gène codant pour la sous-unité A et cette protéine est ciblée à la fois vers la mitochondrie et le plastide (35). Il existe deux gènes encodant pour des sous-unités B, l'un d'eux produisant une protéine ciblée vers la mitochondrie et l'autre vers le plastide. La mutation d'un seul gène codant pour ces protéines conduit à une mortalité embryonnaire dans le cas de la gyrase A et à une mortalité chez les jeunes pousses ou des plantes infertiles dans le cas des gyrases B, montrant ainsi l'importance de ces protéines.

Pour contourner le problème de la mortalité de la mutation des gènes des gyrases, un groupe de chercheurs a utilisé la technique du *virus-induced gene silencing* (VIGS) dans les feuilles de la plante *Nicotiana benthamiana* pour neutraliser les messagers de ces gènes (37). La diminution de la quantité de gyrases présente dans les organelles a eu pour effet de perturber la séparation et la bonne ségrégation des nucléoïdes aussi bien dans les plastides que dans les mitochondries. Ces problèmes entraînent alors des irrégularités au niveau de la division des organelles qui se retrouvent en plus petit nombre dans les régions affectées. Des effets similaires sont observés chez des mutants des gènes des gyrases de la bactérie *Escherichia coli* (38). Ces résultats montrent d'une manière plus précise l'importance des gyrases pour la bonne réplication de l'ADN dans les organelles de plantes. Il est également logique de penser que les gyrases, comme les polymérase de l'ADN, occuperont une place centrale dans les voies de la réparation de l'ADN.

1.2.2 Réparation de l'ADN

Le fait que l'ADN de la mitochondrie et du chloroplaste soit à proximité des chaînes de transport d'électrons caractéristique de ces deux compartiments soumet les deux génomes à la présence constante de molécules oxydatives potentiellement mutagènes De

plus, puisque les organelles se transmettent de façon asexuée dans la plupart des plantes à fleurs, la correction des mutations par la recombinaison entre les allèles parentales est impossible. Le postulat connu sous le nom du «Rochet de Muller» stipule en effet que les organismes asexués devraient accumuler des mutations de façon irréversible (39). Il est donc surprenant de constater que la mitochondrie et le plastide accumulent en fait moins de mutations silencieuses que le génome nucléaire (40). Ce paradoxe suggère l'existence de mécanismes de corrections et de réparations très efficaces de l'ADN (41, 42). Cependant, très peu d'études ont porté sur le maintien des génomes des organelles de plantes et conséquemment, très peu de facteurs ont été identifiés. Plusieurs indices portent néanmoins à croire que les organelles contiendraient les éléments nécessaires pour pouvoir effectuer la réparation de mésappariements, la photoréactivation, l'excision de nucléotides et de bases ainsi que la recombinaison.

1.2.2.1 Réparation de mésappariements

La réparation de mésappariements, aussi appelée *mismatch repair* (MMR), permet d'identifier les mauvais appariements Watson-Crick des bases azotées dans les duplex d'ADN ou bien des boucles protubérantes dues à des insertions/délétions sur un brin et de les corriger. Ce mécanisme a été très bien caractérisé chez la bactérie *E. coli* où une multitude de facteurs nommés Mut ont été identifiés. Alors que certains facteurs Mut se spécialisent dans la reconnaissance des mésappariements, d'autres coupent la double hélice d'ADN pour permettre l'extraction des bases fautives et leur remplacement (revue dans (43)).

Chez *E. coli*, la reconnaissance de l'ADN endommagé se fait grâce au facteur MutS qui recrutera lui-même les autres facteurs de réparation, dont MutL et MutH qui couperont la molécule d'ADN au site du dommage. L'excision des bases, la synthèse de nouvel ADN et la ligation dépendent alors d'une multitude d'exonucléases en plus des Pol et ligases classiques. Du côté de la mitochondrie animale, la présence de mécanismes de réparation de mésappariements a longtemps été débattue (revue dans (44)). Ceci était dû au fait que les protéines homologues des facteurs Mut bactérien se retrouvent dans le noyau des cellules et non dans la mitochondrie. Néanmoins, une activité de réparation de

mésappariements a été détectée dans des extraits de mitochondrie de foie de rat (45). Récemment, il a même été démontré qu'une nouvelle protéine nommée YB-1 est associée à une telle activité de réparation dans la mitochondrie humaine (46). Il semble donc qu'un tel mécanisme existe mais utilise des facteurs différents. Il est intéressant de constater que chez la levure, on retrouve la protéine MSH1 (MutS Homologue 1) dans la mitochondrie (47). Cette protéine joue un rôle vital dans le maintien du génome de la mitochondrie de *Saccharomyces cerevisiae* et l'étude de mutations du domaine de MMR de la protéine suggère qu'elle est également impliquée dans d'autres voies de réparation (48). Une étude récente montre en effet que la protéine MSH1 de levure est impliquée dans le mécanisme d'excision de bases (49).

Chez les plantes, on retrouve également des homologues des protéines MutS bactérienne qui sont impliquées dans le MMR mais aussi dans l'inhibition de la recombinaison non-homologue ou homéologue (dans le cas des homologies imparfaites) dans le noyau (50, 51). En ce qui a trait aux organelles, il a été démontré que, comme c'est le cas pour la levure, la protéine MSH1 est présente au niveau des organelles (52). La mutation du gène encodant cette protéine chez *Arabidopsis* entraîne des réarrangements importants au niveau de l'ADNmt qui ont des effets impressionnants sur le développement de la plante (voir la section 1.1.4.2.5). Cependant, aucune fonction classique de réparation de mésappariements n'a été trouvée pour MSH1, introduisant un doute sur l'existence d'un tel mécanisme dans les organelles de plantes. Il serait même envisageable que les organelles de plantes ne possèdent plus de voie de réparation de mésappariements comme c'est le cas de leurs ancêtres bactériens.

1.2.2.2 Photoréactivation

De par leur nature, les plantes sont constamment soumises aux radiations du soleil, notamment aux rayons ultra-violet (UV) de type B (290 à 320 nanomètres) qui peuvent causer des dommages à l'ADN (53). Le résultat de l'absorption de ces rayonnements UV par la molécule d'ADN est l'apparition de dimères de pyrimidines et de pyrimidine-(6'-4')-pyrimidinones (revue dans (54)). Ces deux types de dommage sont très nocifs puisqu'ils peuvent interrompre à la fois la réplication et la transcription de l'ADN. Le génome

nucléaire des plantes est protégé de ce type de dommage par des enzymes nommés photo-lyases qui peuvent couper le lien chimique entre les dimères de pyrimidines en utilisant l'énergie lumineuse (revue dans (54)). Bien qu'aucune photo-lyase dirigée vers les organelles n'ai été trouvée dans le génome d'*Arabidopsis*, une étude a mis en évidence la capacité du génome plastidique (plastome) à réparer les dommages causés par les rayons UV-B (55). De plus, chez l'algue verte *Chlamydomonas reinhardtii*, un gène encodant une photo-lyase (*PHR1*) retrouvée dans le chloroplaste a été découvert (56). La mutation de ce gène entraîne des défauts de réparation de l'ADNpt. Une fois de plus, ces données suggèrent l'existence de mécanismes de réparation complexes dont les acteurs restent majoritairement inconnus à ce jour.

1.2.2.3 Excision de nucléotides

Une autre voie, différente de la photoréactivation, est disponible pour réparer un dimère de pyrimidine. En effet, la réparation par excision de nucléotides (NER) permet de réparer une telle lésion mais aussi une variété d'autres dommages qui causent des changements de structure dans l'ADN (revue dans (57, 58)). Ce mécanisme peut être divisé en deux voies, l'une réparant les dommages dans le génome en entier et l'autre associé spécifiquement à la transcription (54, 58). Les deux voies diffèrent dans la reconnaissance des lésions; dans la réparation globale chez l'humain, celle-ci est faite par le complexe formé des protéines XPC et d'un homologue de Rad23 alors que c'est la polymérase de l'ARN II qui fait la reconnaissance dans la voie associée à la transcription. Il y aura alors ouverture de l'ADN au site du dommage grâce à des protéines comme XPA ainsi que la *Replication Protein A* (RPA). Ensuite, un oligonucléotide de 20 à 30 nucléotides contenant la région endommagée sera retiré par un complexe d'endonucléases comme XPF et ERCC1 (revue dans (54)). Il y aura alors synthèse d'un nouveau segment d'ADN en utilisant le brin d'ADN intact comme matrice et une étape de ligation vient ensuite clore le processus (revue dans (54, 57, 58)). Jusqu'à ce jour, aucune preuve ne vient appuyer l'existence d'une telle voie de réparation dans les organelles de plantes (54) ou même dans la mitochondrie animale (59).

1.2.2.4 Excision de bases

La réparation par excision de bases (BER) est un mécanisme qui permet d'extraire des bases modifiées par oxydation ou radiation et de corriger la séquence de l'ADN. Ce mécanisme implique d'abord une glycosylase d'ADN qui coupe la base azotée problématique (revue dans (44, 54)). Par la suite, une endonucléase AP vient faire une coupure simple brin au site où la base a été enlevée. Deux voies sont alors disponibles pour procéder selon qu'on synthétise un court fragment d'ADN ou un long fragment. Dans la mitochondrie animale, où le BER est un mécanisme de réparation très important, la synthèse d'ADN à partir de l'extrémité 3' est réalisée par la polymérase de l'ADN γ dans les deux voies (revue dans (59)). L'ADNsb non-apparié en 5' est alors enlevé par des enzymes qui sont elles spécifiques à chacune des deux voies et la ligation est effectuée par la Ligase III.

Du côté des végétaux, certains facteurs impliqués dans ces voies chez les mitochondries et les plastides ont récemment été identifiés, suggérant du même coup la présence de ces mécanismes dans les organelles de plantes. Dans la mitochondrie d'*Arabidopsis*, une activité uracil-ADN glycosylase a été isolée et a été associée au gène At3g18630 (60). Une activité uracil-ADN glycosylase a également été isolée à partir des chloroplastes de la plante *Zea mays* (maïs) (61). Une étude plus récente a confirmé la présence de mécanismes BER dans les plastides d'*Arabidopsis* (62). Cette étude identifie trois protéines co-localisant avec les nucléoïdes du chloroplaste comme facteurs potentiels des voies de la BER. Cependant la mutation, même simultanée, des trois gènes encodant deux endonucléases glycosylases (*AtNTH1* et *AtNTH2*) et une endonucléase AP (*ARP*) ne provoque aucune sensibilité vis-à-vis des stress oxydatifs. Cette absence de phénotype suggère la présence de plusieurs mécanismes de réparation des bases modifiées et/ou l'implication d'une multitude de facteurs aux fonctions similaires capables de compenser pour la perte de certains gènes.

1.2.2.5 Recombinaison

Les bris double brin sont parmi les dommages les plus sévères que peut subir l'ADN. Lorsque la double hélice se brise, non seulement la réplication et la transcription sont interrompues, mais la survie même de la cellule est compromise. La réparation de ces dommages passe par un mécanisme de recombinaison qui peut être homologue ou non-homologue (revue dans (54, 63)). Dans le cas des organelles de plantes, il semble certain que la recombinaison homologue est présente chez la mitochondrie comme chez le plastide. L'existence de réparation non-homologue, comme la jonction d'extrémités non-homologues (NHEJ) dans les organelles a récemment été démontrée grâce à l'induction d'une endonucléase dans le plastide d'*Arabidopsis* (64). La recombinaison a une importance particulière pour le génome de la mitochondrie de plante qui est très recombinogénique (65). Le génome de la mitochondrie contient en effet un grand nombre de séquences répétées qui peuvent produire différentes molécules réarrangées (revue dans (65) et (63)). Cette flexibilité au niveau de l'ADNmt conférerait un avantage pour les plantes qui auraient la capacité de s'adapter plus facilement à des variations brusques de leur environnement (66).

1.2.2.5.1 Recombinaison non-homologue

La NHEJ est un phénomène universel de réparation des bris double brin de l'ADN. Dans sa forme la plus simple, la NHEJ consiste à joindre deux extrémités d'ADN laissées libres à la suite d'un bris de la double hélice (Figure 5) (revue dans (67)). Les deux extrémités jointes peuvent être franches ou bien utiliser des microhomologies d'un à quatre nucléotides. La jonction se fait donc de façon quasi aléatoire et est donc susceptible aux erreurs qui peuvent introduire des réarrangements dans le génome. Les étapes les mieux caractérisées de ce phénomène chez les eucaryotes sont celles de la reconnaissance des extrémités par le complexe Ku, composé des protéines Ku70 et Ku80 ainsi que le recrutement d'une protéine kinase dépendante de l'ADN (DNA-PK) (revue dans (68)). Chez *Arabidopsis*, les homologues AtKu70 et AtKu80 ont été identifiés en plus d'une ligase de type IV et des protéines AtXRCC4, AtMRE11 et AtRAD50 qui sont toutes impliquées dans la NHEJ (69-71). Toutes ces protéines sont cependant localisées au noyau et aucun facteur

impliqué dans la NHEJ au niveau de organelles n'a été identifié à ce jour. Une étude de la réparation des bris double brin dans l'ADNpt a montré l'absence de réparation de type NHEJ chez le tabac, mettant en doute l'existence d'un tel mécanisme dans les organelles (72). Cependant, l'induction d'une endonucléase ciblée vers le plastide d'*Arabidopsis* a permis d'examiner comment cet organelle répare les coupure cohésive (64). Il semble que les voies du NHEJ et du *microhomology-mediated end joining* (MMEJ) (revue dans (68)) soient privilégiées.

Le MMEJ est une autre forme de jonction possible mais qui implique une séquence de microhomologie, d'où son nom. Ce mécanisme est en effet caractérisé par une courte séquence répétée de 5 à 25 nucléotides à chacune des deux extrémités (Figure 5). Ce processus est très similaire à celui qu'on nomme hybridation de simples brins (*single-strand annealing*, SSA) et il pourrait en fait s'agir du même mécanisme (68). Malgré le fait qu'il soit guidé par une certaine homologie, ce processus reste hautement susceptible aux erreurs, c'est à dire qu'il crée souvent des réarrangements dans la séquence de l'ADN. Dans l'article de Kwon *et al.* les mécanismes de NHEJ et du MMEJ provoquent des délétions de larges fragments d'ADNpt de part et d'autre du site de coupure de l'endonucléase (64).

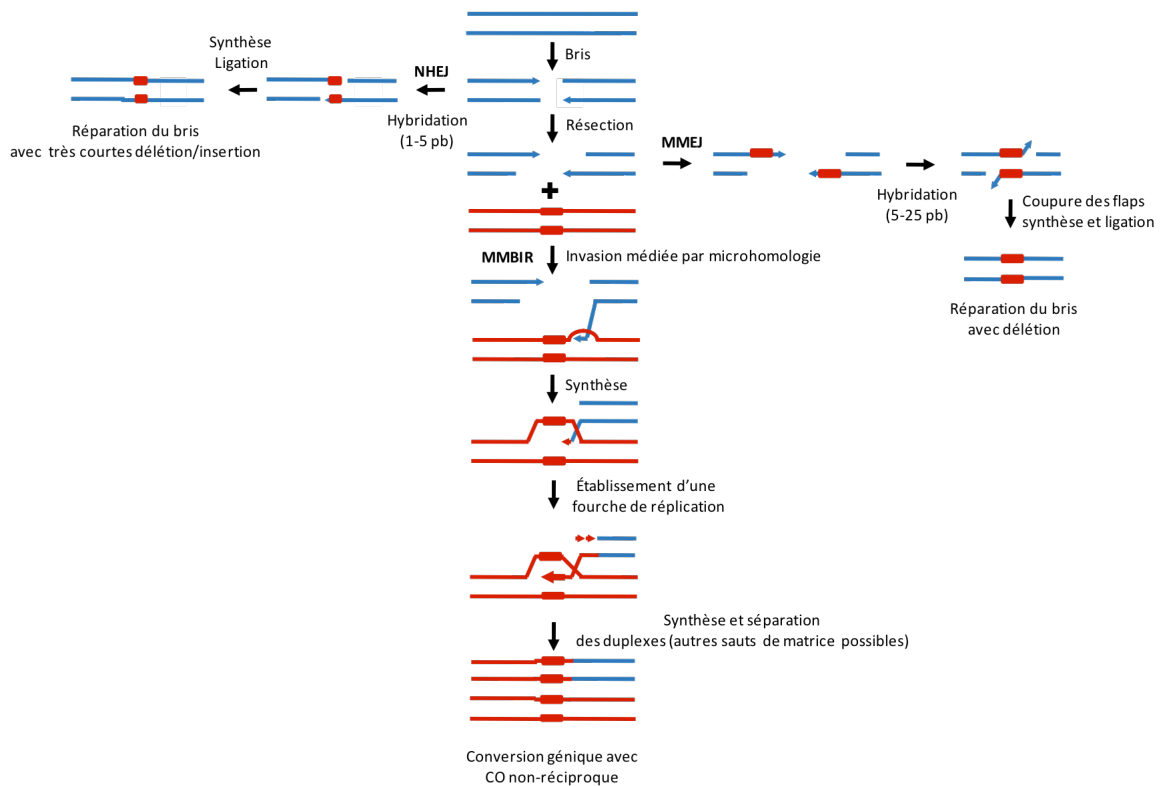


Figure 5. Modèles des mécanismes de réparation non-conservateurs des bris double brin de l'ADN. Les deux extrémités d'un bris double brin peuvent être immédiatement recollées ensemble ce qui donne lieu à la jonction d'extrémités non-homologues (NHEJ). Alternativement, si il y a résection des brins en 5' à chacune des extrémités, une séquence de microhomologie peut être trouvée et guide alors la jonction des deux extrémités (MMEJ). Quand l'un des deux brins 3' envahit une autre molécule d'ADN sur une séquence de microhomologie, il peut y avoir établissement d'une fourche de réplication qui entrainera une recombinaison non-homologue (MMBIR). CO; *Crossing-over* (Adaptée de (73)).

1.2.2.5.2 Recombinaison homologue

La méthode la plus fiable de réparation des bris double brin de l'ADN reste la recombinaison homologue (HR). Le modèle proposé par Szostak *et al.* illustre comment un bris double brin peut être réparé à l'aide d'un phénomène nommé «réparation de bris double brin par recombinaison homologue» (DSBR) (Figure 6) (74). Ce mécanisme implique le transfert d'informations d'une molécule d'ADN à une autre et est souvent associée au transfert réciproque (*crossing-over*). Lorsque deux extrémités d'ADN ne sont pas directement jointes par NHEJ, le brin 5' de chaque extrémité va subir une résection qui laissera une extrémité 3'OH libre. Ces extrémités simple brin sont très réactives puisqu'elles peuvent envahir des duplex d'ADN et amorcer la synthèse de nouvel ADN (revue dans (63)). Lorsque l'une des deux extrémités s'apparie avec une autre molécule d'ADN à une région d'homologie, une réaction de synthèse de l'ADN s'engage en utilisant une autre molécule comme matrice. Par exemple, dans le cas de la recombinaison méiotique, c'est l'ADN du chromosome sœur qui sert de matrice. La synthèse peut alors se poursuivre sur un brin seulement jusqu'à ce que la réplication atteigne la région homologue à la deuxième extrémité du bris initial. Une nouvelle fourche de réplication peut alors démarrer avec la seconde extrémité 3'OH et ainsi former la double jonction de Holliday. Une fois la synthèse terminée, il y a résolution du complexe avec deux résultats possibles selon qu'il y a *crossing-over* ou non. Ce type de recombinaison classique a été abondamment étudié chez les végétaux autant pour l'ADN du noyau que celui de la mitochondrie et du plastide (revue dans (54), (12) et (22) respectivement).

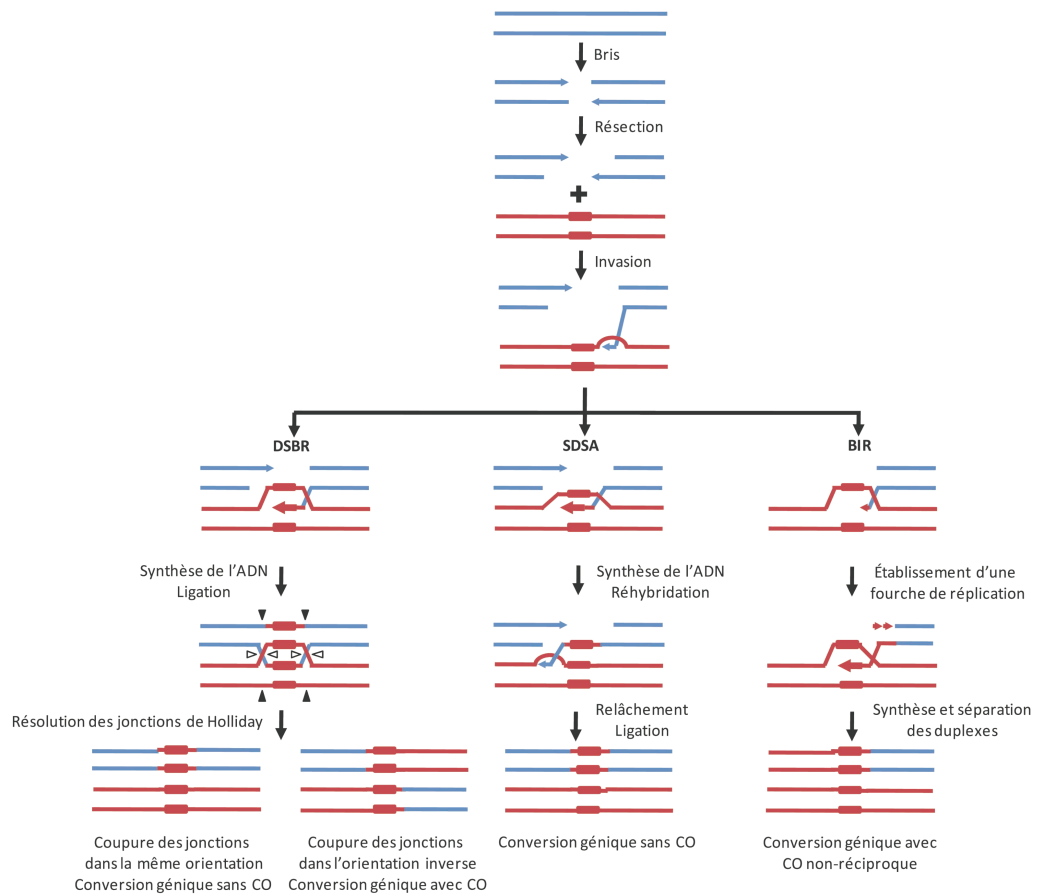


Figure 6. Modèles des mécanismes de réparation conservateurs des bris double brin de l'ADN. Plusieurs types de réactions de recombinaison peuvent être utilisés pour réparer une cassure double brin dans l'ADN de façon conservatrice. La réparation de bris double brin (DSBR) par recombinaison homologue peut générer deux produits différents dépendant de la résolution de la double jonction de Holliday. Seul le produit sans *crossing-over* (CO) est possible lors de la réaction de *synthesis-dependent strand annealing* (SDSA). La réparation induite par une brisure (BIR) entraîne obligatoirement un *crossing-over* non-réciproque (Adaptée de (73)).

Un autre modèle proposé par Szostak *et al.*, le *synthesis-dependent strand annealing* (SDSA) implique une jonction de Holliday simple (Figure 6) (74). La différence entre ce mécanisme et la DSBR réside dans le fait que la deuxième fourche de réplication ne se forme jamais ce qui anéantit la possibilité de *crossing-over* (revue dans (63)). Ceci pourrait être dû à l'action de protéines topoisomérase ou hélicase qui fermentaient l'unique boucle de réplication au fur et à mesure que celle-ci progresse ou bien à une bulle de réplication très petite. Le deuxième brin d'ADN sera donc complété en utilisant le premier brin comme matrice et non grâce à l'invasion d'une autre molécule d'ADN. Ce type de recombinaison a déjà été rapporté pour le chloroplaste de *C. reinhardtii* (75), ce qui suggère que ce mécanisme pourrait bien exister dans les organelles de plantes.

Un dernier mécanisme, indépendant de la machinerie de recombinaison homologue classique, qui permet de réparer les bris double brin d'ADN de façon conservatrice se nomme la réplication induite par les bris (BIR) (Figure 6) (revue dans (76)). Contrairement aux mécanismes DSBR et SDSA, la BIR implique une fourche de réplication complète avec non seulement la synthèse du brin précoce (*leading*) mais aussi celle du brin retardé (*lagging*). Encore une fois, le brin d'ADN 3'OH d'une extrémité du bris envahit une molécule d'ADN double brin ou bien une région d'ADN devenue simple brin à une région d'homologie et la boucle D devient une fourche de réplication (revue dans (77)). Ce mécanisme ne fait donc pas intervenir la seconde extrémité du bris double brin qui pourra donc être perdue si elle n'est pas elle-même réparée. Cette indépendance envers une seconde extrémité en fait un mécanisme de prédilection pour redémarrer une fourche de réplication arrêtée chez les eucaryotes (revue dans (78)). En effet, chez la levure *Saccharomyces cerevisiae*, la BIR est utilisée pour la réparation des dommages à l'ADN, mais aussi des fourches de réplifications arrêtées et même l'élongation des télomères en absence de télomérase (79, 80). Chez l'humain, le processus de BIR peut passer par plusieurs cycles de séparation et ré-invasion des deux duplex d'ADN (77). Ces cycles successifs donnent cependant plusieurs opportunités à l'ADN envahissant de s'aligner avec une molécule d'ADN ne partageant qu'une homologie partielle et donc donner lieu à une recombinaison non-homologue. Lorsque la BIR est guidée par une séquence de microhomologie, on parle alors de «*microhomology-mediated break-induced replication*»

(MMBIR), un processus responsable d'une grande quantité de réarrangements du génome chez l'être humain (Figure 5) (revue dans (81, 82)). L'existence de tels mécanismes chez la plante reste encore à prouver et les protéines potentiellement impliquées dans de telles voies restent à identifier.

1.2.2.5.3 Protéines impliquées dans la recombinaison de l'ADN des organelles

A) MSH1

Une des premières mutations identifiées pour causer des réarrangements au niveau des organelles d'*Arabidopsis* a été isolée il y a plus de 30 ans (83). Cette allèle mutante a d'abord été nommée «*chloroplast mutator*» ou *CHM* en raison du phénotype de la plante mutante qui développe des régions de feuilles sans chlorophylle. Il a cependant été révélé que la mutation causait en fait des réarrangements dans l'ADNmt qui seraient donc indirectement responsables du phénotype observé dans le développement des chloroplastes (84). Plusieurs études subséquentes ont répertorié les différents types de réarrangements détectables dans le génome de la mitochondrie du mutant (revue dans (63)). Tous ces produits sont le résultat de recombinaisons dites «homologues» du fait qu'elles utilisent de grandes régions d'homologies (supérieures à 200 nucléotides) malgré le fait qu'elles produisent des réarrangements. Le gène mutant a finalement été cloné et identifié comme un homologue du gène *MutS* de *E. coli* dont la protéine est impliquée dans la réparation de mésappariements (voir section 1.1.4.2.1) qu'on retrouve également dans la mitochondrie de la levure *S. cerevisiae* {Chi, 1994 #164}. La protéine, nommée MSH1, est prédite pour être localisée à la fois dans les mitochondries et les chloroplastes (85) bien qu'à ce jour, on ne lui ait trouvé aucun rôle dans le plastide (86).

Chez la bactérie, la protéine MutS reconnaît les sites de mésappariements dans le duplex d'ADN (revue dans (87) et (43)). Pour compléter la réparation des mésappariements dans l'ADN, MutS doit être secondé dans ses fonctions par les protéines MutH et MutL qui n'ont pas d'homologue qui pourrait se retrouver dans les organelles de plantes. La protéine MSH1 d'*Arabidopsis* contient six domaines dont trois sans motif identifiable et donc sans aucune fonction prédite (86). Il serait donc possible que cette protéine de plante intègre les fonctions de plusieurs protéines bactériennes impliquées dans la

réparation des mésappariements. Le rôle précis de cette protéine dans la recombinaison de l'ADN reste obscur, elle pourrait cependant reconnaître les mésappariements à l'intérieur d'un hétéroduplexe lors de l'échange de brins et faire avorter le processus (63).

B) Les homologues de RecA

Les recombinases bactériennes de la famille RecA sont essentielles pour la recombinaison homologue chez les procaryotes (revue dans (88)). Ces protéines catalysent l'échange de brins d'ADN entre deux duplex en utilisant l'énergie de l'adénosine triphosphate (ATP) (89). Cette étape est critique pour le bon alignement des brins d'ADN au tout début des processus de recombinaison DSBR et SDSA. Comme les protéines RecA sont retrouvées chez les bactéries, il est surprenant de constater que trois homologues (RecA1, RecA2 et RecA3) sont encodés dans le génome nucléaire d'*Arabidopsis* (90). Toutefois, ces trois protéines possèdent un peptide signal permettant de cibler les organelles.

La mutation du gène *RecA3* ne semble pas causer de phénotype visible, mais mène à des réarrangements de l'ADNmt (91). Il est intéressant de constater que ces molécules réarrangées sont identiques à certaines des molécules observées chez le mutant *msh1*, mais d'autres semblent spécifiques à l'un ou l'autre des mutants (65). Des résultats similaires ont été obtenus avec un mutant du gène *RecA1* de la mousse photosynthétique *Physcomitrella patens* (92). L'organisme mutant est affecté dans son développement et dans la morphologie de ses mitochondries. La plus grande sévérité de cette mutation par rapport à celle du gène *RecA3* chez *Arabidopsis* pourrait être expliquée par la présence de gènes homologues chez la plante alors que RecA1 est unique chez *P. patens*. Au niveau de l'ADNmt, des molécules réarrangées font leur apparition à la suite d'évènements de recombinaison entre de courtes séquences (<100 nucléotides) répétées directes ou indirectes. Encore une fois, la protéine RecA montre un rôle anti-réarrangement et semble donc promouvoir la stabilité du génome. Aucun phénotype n'a cependant été observé concernant le chloroplaste.

Une étude récente portant sur le plastide d'*Arabidopsis* a montré qu'une mutation réduisant l'expression du gène *RecA1* ne mène à aucun phénotype visible (93). La protéine RecA1 étant prédite pour être localisée dans les chloroplastes, la conformation physique

adoptée par l'ADNpt a été étudiée chez la plante mutante. L'ADN des plantes mutantes montrent moins de structures branchées complexes au niveau des nucléoïdes comparativement à l'ADN des plantes de type sauvage (WT). Ces superstructures représenteraient des molécules d'ADN en réplication/transcription actives présentent surtout dans les jeunes tissus de la plante (94). De plus, les plantes *reca1* montrent une sensibilité accrue à la ciprofloxacine, un antibiotique de la famille des quinolones qui induit des bris double brin dans l'ADN par l'intermédiaire des gyrases (95). Bien que cette étude ne donne aucun indice quant à l'impact direct de la protéine RecA1 sur l'ADNpt, elle souligne une fois de plus l'importance de ces protéines pour le maintien de la stabilité des génomes des organelles de plantes (revue dans (63)). Il semble donc que la famille des protéines RecA soit intimement impliquée dans le maintien de la stabilité des génomes des organelles de plante comme c'est le cas chez les procaryotes.

C) La famille des OSB

Les protéines OSB (*Organellar Single-stranded DNA-Binding protein*) de la pomme de terre (*Solanum tuberosum*) ont été isolées grâce à leur grande affinité pour l'ADNsb (96). Il existe quatre protéines de cette famille chez *Arabidopsis* : OSB1 qui se retrouve dans la mitochondrie, OSB2 dans le chloroplaste, OSB3 prédites pour être dans les deux organelles et finalement OSB4 dont l'existence n'a pas été confirmée (97). Alors qu'une mutation des gènes *OSB2* ou *OSB3* n'a aucun effet sur la plante, la mutation du gène *OSB1* entraîne un phénotype visible. En effet, à partir de la quatrième génération, un phénotype de variégation, caractérisé par l'apparition de taches jaunes et/ou blanches sur les feuilles, apparaît chez certains individus et gagne en sévérité avec les générations subséquentes (97). Ce trait a pour cause des réarrangements au niveau de l'ADNmt produits par des événements de recombinaison entre de courtes séquences (>200 nucléotides) homologues. Bien qu'il s'agisse du même type de réarrangements que dans le cas des mutants *msh1* et *reca3*, la sévérité du phénotype est intermédiaire, indiquant fort probablement un nouveau mode d'action pour les OSBs dans la protection des génomes des organelles.

D) Les protéines liant l'ADN simple brin

La double hélice est la conformation privilégiée de l'ADN car elle est très stable et limite ainsi la dégradation et la modification des bases. Cependant, toutes les transactions cellulaires impliquant l'ADN le laisseront temporairement sous forme simple brin et donc susceptible aux attaques par des agents chimiques ainsi qu'aux mauvais appariements. C'est pourquoi toutes les cellules vivantes sont dotées de protéines liant l'ADNsb qui stabilisent l'état transitoire de l'ADN et protègent contre les mauvais appariements et les dommages (revue dans (59)). La protéine liant l'ADNsb la plus étudiée est certainement la protéine *single-strand DNA binding* (SSB) de la bactérie *E. coli*. Cette protéine possède un domaine OB (*oligonucleotide binding fold*) qui permet la liaison de l'ADNsb. Elle est considérée comme le prototype des protéines SSB procaryotes.

Les protéines SSB sont impliquées dans pratiquement toutes les réactions du métabolisme de l'ADN des bactéries, incluant la réplication et la réparation (revue dans (98)). Ces protéines lient l'ADNsb de façon coopérative et permettent, dans un premier temps, sa protection et ensuite, recrutent les différents facteurs nécessaires à son maintien. En effet, lors de la réplication, la protéine SSB interagit avec certaines sous-unités de la DNA polymérase III et cette interaction est vitale pour l'organisme (99). La protéines SSB interagit également avec la primase dont l'activité de synthèse des amorces d'ARN est essentielle à la réplication (100). Dans le contexte de la réparation de l'ADN, les protéines SSB bactériennes interagissent directement avec plusieurs facteurs importants. Parmi ces protéines, on retrouve les hélicases RecQ (101) et RecG (102) qui sont essentielles à la recombinaison homologue ainsi que PriA impliquée dans le redémarrage de la fourche de réplication (103). La protéine SSB peut également recruter à l'ADN l'exonucléase ExoI qui est importante pour la réparation (104), tout comme la protéine RecO qui elle recrutera la protéine RecA, centrale au phénomène de la recombinaison homologue chez les procaryotes (104, 105). Finalement, la protéine SSB est aussi impliquée dans le BER à travers son interaction avec l'uracil glycosylase de l'ADN (UDG) (106).

Chez les eucaryotes, la principale protéine liant l'ADNsb se nomme *Replication protein A* (RPA). Contrairement aux SSB procaryotes très homogènes et qui ne contiennent qu'un

seul domaine de liaison à l'ADN, les protéines RPA sont variées, peuvent contenir plusieurs domaines de liaison et forment des hétéro-complexes (revue dans (59)). Le complexe des protéines RPA humaines a été isolé avec le système de réplication du virus simien 40 car elles sont recrutées *in vitro* par l'antigène T du virus (107). Dans la cellule, les protéines RPA sont observées sous forme de foci au moment du démarrage de la réplication, suggérant un rôle dans la formation du complexe de pré-initiation (revue dans (59)). Similairement à ce qui est observé chez les bactérie, les protéines RPA interagissent directement avec la primase et la polymérase de l'ADN α et provoquent ainsi le démarrage de la synthèse de l'ADN (108). Ces protéines sont également importantes pour la dégradation des amorces d'ARN produites lors de la réplication en recrutant spécifiquement certaines nucléases (109).

Comme pour les SSB bactériennes, la fonction des protéines RPA dans la réparation de l'ADN est d'abord la protection physique de l'ADNsb contre les dommages et les mauvais appariements (revue dans (59)). En effet, lors d'un bris double brin de l'ADN, les extrémités 5' libres subissent une résection et le brin 3' ainsi découvert est lié et protégé par le complexe RPA. Le complexe peut ensuite recruter les protéines Rad51 (110) et Rad52 (111), essentielles à la recombinaison homologue, grâce à des interactions directes. Il est également connu que les protéines RPA agissent sur la voie du NER d'abord en stabilisant l'ADNsb produit dans les premières étapes du mécanisme (112) et ensuite en interagissant avec le facteur XPA qui reconnaît la lésion et contribue à ouvrir l'ADN tout autour (113).

Alors que l'on croyait que l'ADNsb du noyau des cellules eucaryotes était gardé uniquement par les protéines RPA, il a récemment été découvert qu'un homologue de la protéine SSB bactérienne protégeait le génome de la cellule humaine (114). Cette protéine est en effet phosphorylée par la kinase ATM en cas de dommage à l'ADN ce qui provoque l'apparition de foci de hSSB1 dans le noyau de la cellule. De plus, des cellules déficientes en protéine hSSB1 sont plus sensibles aux radiations ionisantes et montrent une plus grande instabilité génomique comparée aux cellules normales. Toutes les cellules eucaryotes possèdent également un homologue des protéines SSB bactériennes qu'on retrouve au niveau de la mitochondrie et que l'on nomme mtSSB. Ces protéines sont

importantes pour la réplication de l'ADNmt chez la levure (115), chez l'insecte modèle *Drosophila melanogaster* (116) ainsi que chez l'humain (117). On retrouve deux homologues de SSB chez la plante *A. thaliana*. L'une de ces protéines a été testée *in vitro* et stimule l'activité d'échange de brins de l'enzyme RecA de *E. coli*, ce qui suggère une implication dans le phénomène de recombinaison homologue (118). Il n'est cependant pas certain qu'un homologue des protéines SSB se retrouve au niveau des plastides et donc les multiples fonctions de protection de l'ADNsb seraient exécutées par une autre famille de protéines.

1.3 Les protéines Whirly

Le principal sujet de la présente étude est la famille de protéines nommées Whirly. Ces protéines ont été découvertes dans le laboratoire du Dr. Normand Brisson à l'Université de Montréal et ce groupe poursuit la recherche sur ces facteurs encore à ce jour (section 1.3.1). L'intérêt pour ces protéines vient de leur conservation à travers le règne végétal ainsi que de leur structure tridimensionnelle particulière (section 1.3.2). De plus, de nombreuses études ont prouvé la présence des protéines Whirly dans plusieurs compartiments subcellulaires (section 1.3.3) et on leur a attribué une multitude de fonctions dans ces différentes localisations (sections 1.3.4 et 1.3.5).

1.3.1 Découverte de la famille de protéines Whirly

Le premier membre de la famille de protéines Whirly a été isolé à partir d'un extrait de pomme de terre grâce à une colonne d'affinité pour l'ADNsb (119). L'oligonucléotide utilisé lors de cette purification correspondait à l'élément activateur, nommé ERE, du promoteur de gène de défense *PR-10a* qui est activé dans les tubercules par une blessure, un éliciteur ou une infection par un pathogène (120). Un signal de retardement sur gel d'une sonde d'ADNsb portant la séquence ERE et qui corrèle avec l'expression de *PR-10a* a été attribué à un facteur de transcription nommé PBF-2 (121). Une colonne d'ADNsb a alors été conçue pour isoler ce facteur et une élution utilisant une forte concentration de

chlorure de sodium a permis de détacher une protéine de 24 kilodalton (kDa), qui constitue le facteur PBF-2 et qui a été nommée StWhy1 (*Solanum tuberosum* Whirly1) (119). Cette protéine est le prototype de la nouvelle famille de protéines spécifiques aux plantes et liant l'ADN (revue dans (122)). Comme il est relativement rare pour un facteur de transcription d'avoir une affinité beaucoup plus grande pour l'ADN sous forme de simple brin par rapport au double brin, l'hypothèse proposée est que StWhy1 stabilise la conformation ouverte de l'ADN et favorise ainsi la transcription. Cette affinité particulière pour l'ADNsb a amené un intérêt quant aux déterminants structuraux de cette protéine qui semblent être partagés par tous les membres de la famille.

1.3.2 Structure et conservation des protéines Whirly

La séquence primaire de la protéine StWhy1 a permis de retrouver les homologues Whirly à travers le règne végétal (123). En effet, on a retrouvé chez tous les angiospermes avec suffisamment d'informations disponibles au moins deux membres de la famille Whirly (revue dans (122)). Ces deux membres se distinguent principalement l'un de l'autre par leur région N-terminale, qui contient une séquence de ciblage aux organelles. Une des deux protéines est prédite pour se retrouver dans la mitochondrie et l'autre dans le chloroplaste. Il est intéressant de noter que deux algues vertes unicellulaires, *C. reinhardtii* et *Ostreococcus tauri* ne possèdent qu'une seule protéine Whirly et que celle-ci serait pour être localisée dans les deux organelles, ce qui suggère que les protéines Whirly auraient acquis une spécialisation pour leur compartiment plus tard dans l'évolution.

Suivant le peptide signal, on retrouve le domaine d'activation potentiel des protéines Whirly. En effet, les protéines Whirly retrouvées dans le plastide (ptWhirly), possèdent une courte région de poly-glutamine, poly-sérine ou poly-proline qui dans le cas de StWhy1 a été montré pour jouer un rôle dans l'activation de la transcription (123). Il est intéressant de constater que les protéines Whirly retrouvées dans la mitochondrie (mtWhirly) ne contiennent pas de telles régions, indiquant une fois encore la spécialisation des deux types de protéines ciblées vers des organelles différentes. Le domaine de transactivation est suivi par le domaine de liaison à l'ADN Whirly (revue dans (122)). Cette partie est la mieux conservée à travers toutes les protéines de la famille. La

structure tridimensionnelle de ce domaine unique de liaison à l'ADN a même été élucidée par cristallographie aux rayons X, ce qui a révélé une structure où se superposent deux planchers β antiparallèles à 90 degrés l'un par rapport à l'autre (123, 124). Cette technique a également révélé l'organisation quaternaire de la protéine qui forme des tétramères alors que quatre sous-unités identiques se lient les unes aux autres tels les pales d'un moulin-à-vent grâce à trois hélices α . C'est cette forme qui a d'ailleurs inspiré le nom «Whirly», contraction du mot anglais *whirlygig*, qui signifie moulin à vent. Plusieurs résidus très conservés sont essentiels au domaine de liaison à l'ADN, particulièrement le motif KGKAAL situé dans le premier plancher β (123).

Le domaine Whirly est suivi d'une région C-terminale peu conservée chez cette famille de protéines. Aucun motif n'est identifiable dans ce domaine et aucune structure particulière ne lui est prédite. Il est cependant intéressant de noter que les derniers six acides aminés de la protéine sont presque parfaitement conservés, indiquant un rôle potentiellement important pour la protéine. Ceci rappelle le cas des protéines SSB (*Single Strand Binding proteins*) des prokaryotes chez qui les derniers résidus servent au recrutement des différents interacteurs (revue dans (98)). La queue C-terminale des Whirly pourrait donc jouer le même rôle que celui des protéines SSB et recruter les bons interacteurs aux bons endroits sur l'ADN. L'identité de ces interacteurs potentiels sera cependant très différente selon que l'interaction ait lieu dans le noyau ou dans les organelles. En effet, l'existence d'une séquence de ciblage aux organelles à l'extrémité N-terminale pose une question importante sur la localisation intracellulaire des protéines Whirly.

1.3.3 Localisation intracellulaire

Au début de l'étude des protéines Whirly, plusieurs indices portaient à croire à une localisation nucléaire de ces protéines. En effet, la protéine prototype de la famille, StWhy1, a été purifiée à partir d'un extrait de noyaux de tubercules grâce à la séquence activatrice du promoteur du gène nucléaire *PR-10a* (119). De plus, la production de cette protéine dans les protoplastes de pomme de terre active l'expression d'un gène de luciférase placé à la suite du promoteur de *PR-10a*. *In vivo*, une expérience d'immunoprécipitation de la chromatine (ChIp) a démontré l'association de StWhy1 avec

le promoteur de *PR-10a* dans les tubercules de pomme de terre (125). Tous ces résultats suggéraient un rôle de facteur de transcription nucléaire pour la protéine. Il est donc surprenant, suite au clonage de l'ADN complémentaire (ADNc) du gène *StWhy1*, de réaliser que la protéine est précédée d'un peptide signal ciblant le plastide (revue dans (122)). Puisque la protéine *StWhy1* naissante est prédite pour faire environ 30 kDa, mais qu'elle apparaît toujours sur gel dénaturant à une hauteur de d'environ 24 kDa, il semble donc que le peptide signal soit coupé très rapidement et que la protéine se retrouve dans le plastide. De plus, une étude sur la localisation intracellulaire des protéines Whirly d'*Arabidopsis* a montré que les constructions Whirly-GFP (Green Fluorescent Protein) se retrouvaient aux plastides dans le cas de *AtWhy1* et *AtWhy3* et dans les mitochondries pour *AtWhy2* (126).

Ces résultats sont donc en contradiction avec l'hypothèse du rôle de facteur de transcription nucléaire. Cependant, dans une étude subséquente, le même groupe a réussi à prouver la double localisation d'une protéine Whirly de l'orge, *HvWhy1* (*Hordeum vulgare* Whirly1) (127). Un autre résultat intéressant rapporté par la même étude montre que la production dans les cellules de l'épiderme de l'oignon (*Allium cepa*) de constructions de *AtWhy1* ou *HvWhy1* fusionnés à la GFP, mais sans peptide signal entraîne une localisation nucléaire. Il semble donc que les protéines Whirly soient d'abord dirigées vers les organelles où elles se font enlever leur partie N-terminale et ensuite, certaines d'entre elles peuvent être relocalisées au noyau par un mécanisme encore inconnu (128). Il devient donc envisageable que ces protéines possèdent différents rôles qui soient dépendants de leur localisation intracellulaire.

1.3.4 Rôles nucléaires

La relation entre la liaison du promoteur du gène *PR-10a* de pomme de terre par *StWhy1* et le niveau d'expression du gène a été abondamment étudiée (119, 123, 125). Étant donné que la protéine *PR-10a* a été impliquée dans la réponse de défense contre les pathogènes, il semble donc y avoir un rôle pour les protéines Whirly dans ce processus. L'implication de ces protéines dans la réponse de défense a été confirmée chez *Arabidopsis* où une mutation du gène *AtWhy1*, provoquant une mutation ponctuelle dans

la protéine, rend la plante plus susceptible à une infection par le pathogène *Hyaloperonospora arabidopsidis* (125). De plus, la protéine AtWhy1 WT semble requise pour qu'il y ait induction complète de la réponse systémique acquise, un système permettant aux plantes de s'immuniser contre des infections subséquentes.

Une étude récente montre que les protéines AtWhy1 et AtWhy3 seraient des répresseurs transcriptionnels du gène *AtKP1* qui encodent une kinésine (129). La surexpression de deux transgènes comprenant les deux protéines sans leur peptide signal entraîne une diminution de l'expression du gène *AtKP1*. Cette étude constitue une nouvelle preuve supportant un rôle de facteur de transcription nucléaire pour certaines protéines de la famille Whirly. De plus, la nécessité de retirer le peptide de ciblage aux plastides pour voir un effet nucléaire vient confirmer les résultats de localisation des Whirly obtenus par Grabowski *et al.* (127).

En plus d'un rôle de répresseur de la transcription, la protéine Why1 d'*Arabidopsis* possède un rôle de maintien de l'homéostasie des télomères (130). Les télomères sont des répétitions d'une courte séquence d'ADNs qui protègent les extrémités des chromosomes des eucaryotes contre l'érosion et la protéine AtWhy1 a été démontré comme ayant une grande affinité pour cette séquence (131). De plus, lors d'un test *in vitro* d'élongation de télomères, l'ajout de protéines AtWhy1 recombinantes a pour effet d'augmenter la longueur des molécules d'ADN (130). L'étude démontre également que l'activité télomérase à l'intérieur de deux lignées mutantes pour le gène *AtWhy1* est modifiée. Il semble donc que la protéine AtWhy1 aie un deuxième rôle à l'intérieur du noyau de la cellule de plante. En conclusion, plusieurs groupes de chercheurs ont trouvé une grande variété de rôles pour les protéines Whirly pour le seul compartiment du noyau.

1.3.5 Rôles dans les organelles

Le fait même que toutes les plantes à fleurs possèdent des protéines Whirly ciblant spécifiquement soit la mitochondrie ou le plastide suggère un rôle important pour ces protéines dans les organelles. Malheureusement, peu d'études se sont intéressées à ces

rôles potentiels. Une première étude a été réalisée par notre groupe avec pour sujet la protéine *Why2* d'*Arabidopsis*, une protéine de la classe des mtWhirly. Pour cette étude, des plantes surexprimant la protéine ont été isolées (132). Ces plantes montraient un grave phénotype de retard de croissance et de feuillage jaune qui corrélait avec une diminution importante de la quantité d'ADNmt. Nous avons également démontré par ChIp que la protéine *AtWhy2* interagissait avec l'ADNmt peu importe qu'il s'agisse d'une séquence codant pour un gène ou une séquence intergénique. Ce résultat suggère une absence de spécificité pour une séquence d'ADN particulière et suggère l'existence d'un autre rôle dans la mitochondrie que celui de facteur de transcription.

Une autre étude a tenté d'élucider le rôle de la protéine de maïs *ZmWhy1* (*Zea mays* Whirly1) un membre de la classe des ptWhirly (133). La protéine a été isolée lors d'une expérience de co-immunoprécipitation utilisant un anticorps contre le facteur d'épissage CRS1 (134). Les deux lignées mutantes pour le gène *ZmWhy1* obtenues montraient de graves phénotypes au niveau de la coloration des feuilles et au niveau du développement de la plante qui meurt avant la deuxième semaine (133). Les deux mutations en question sont des insertions de transposons à l'intérieur du gène *ZmWhy1*. La coloration blanche et jaune des feuilles des mutants *ZmWhy1-1* et *ZmWhy1-2* suggère un dérèglement au niveau de la biogénèse des chloroplastes. Ce phénotype a été attribué à une forte diminution de la quantité d'ARN ribosomal (ARNr) essentiel au fonctionnement des cellules. Puisque *ZmWhy1* a été isolé avec un complexe d'épissage de l'ARN, l'identification des ARN immunoprécipités avec la protéine a été entreprise. Certains ARN identifiés étaient enrichis dans le précipité, mais ne correspondaient cependant pas aux ARNr. Le mécanisme par lequel la présence de *ZmWhy1* affecte les ARNr reste conséquemment à découvrir. Comme pour l'étude de *AtWhy2* (132), une expérience ChIp a été réalisée sur les plants de maïs WT et a démontrée que la protéine *ZmWhy1* interagissait avec toutes les régions de l'ADNpt (133). L'étude de Prikryl *et al.* a donc démontré qu'une protéine Whirly interagissait de façon spécifique avec certains ARN et d'une façon non spécifique avec l'ADN. Il semble donc que les protéines Whirly peuvent avoir différents rôles à l'intérieur des organelles comme dans le noyau.

1.4 Les polymérase de l'ADN des organelles

Il est surprenant de constater que les enzymes responsables de la réplication de l'ADN dans les organelles de plantes n'ont jamais été formellement identifiées. Plusieurs études ont été faites pour isoler l'activité de synthèse de l'ADN des deux organelles et elles ont établi les différentes caractéristiques propres aux polymérases de l'ADN contenues dans différents extraits de plantes (section 1.4.1). Suite au séquençage des premiers génomes de plantes, une famille de polymérases bactériennes de type I a été identifiée comme le candidat le plus probable pour être responsable de la réplication des génomes des organelles. Certains des gènes encodant ces polymérases ont été clonés et l'activité des protéines recombinantes a été mesurée pour plusieurs d'entre elles confirmant la capacité de ces protéines à pouvoir répliquer de l'ADN (section 1.4.2). Il semble également que ces protéines localisent simultanément dans les deux organelles, ce qui en fait des candidats encore plus intéressants (section 1.4.3). La preuve génétique du lien entre les polymérases de type I et l'activité de synthèse de l'ADN à l'intérieur des organelles n'a cependant toujours pas été produite.

1.4.1 Isolation de l'activité de synthèse de l'ADN

Depuis longtemps, plusieurs groupes de recherche se sont intéressés à la synthèse de l'ADN dans les organelles de plantes, spécialement dans le cas du chloroplaste, unique aux plantes. Une activité de synthèse d'ADN a donc été isolée à partir de chloroplaste de *Euglena gracilis* (135), *C. reinhartii* (136), *Spinacia oleracea* (épinard) (137, 138) et *Pisum sativum* (pois) (139). Les polymérases isolées au cours de ces études partageaient plusieurs caractéristiques comme une haute masse moléculaire (87 à 110 kDa) et une insensibilité à un inhibiteur puissant des polymérases α (caractéristiques du noyau des eucaryotes), l'aphidicoline (140). Cette résistance à un inhibiteur classique des Pol α a poussé plusieurs groupes à faire le rapprochement entre la Pol du chloroplaste et la polymérase γ (Pol γ) de la mitochondrie animale, également insensible à l'aphidicoline. Chez l'épinard, le polypeptide associé à la synthèse de l'ADN se verra attribué une activité

3'-5' exonucléase caractéristique des Pol capables de corriger des erreurs lors de la polymérisation de l'ADN (140).

La purification de la Pol de la mitochondrie de plantes a été faite à partir d'extraits de germes de blé (*Triticum aestivum*) (141). L'activité de synthèse identifiée dans cette étude sera plus tard associée à la fraction membranaire des mitochondries par l'entremise de l'ADNmt, confirmant qu'il s'agit bien de l'activité de réplication de l'ADN des mitochondries (142). La Pol de la mitochondrie a également été purifiée de la plante à fleurs *Chenopodium album* et on lui a associé une activité 3'-5' exonucléase comme dans le cas de la Pol du chloroplaste de l'épinard (143). Les Pol des deux organelles ont également été purifiées parallèlement à partir de cultures cellulaires de soya (*Glycine max*) (144). Le résultat de la comparaison des deux enzymes est frappant puisqu'elles possèdent les mêmes caractéristiques biochimiques. Les deux Pol atteignent leur performance maximale dans les mêmes conditions, en plus d'être inhibées par les mêmes molécules. Cette dernière étude trace encore les parallèles entre les Pol des organelles de plantes et les Pol des mitochondries animales. Cependant, l'étude révèle une différence majeure entre les deux types de Pol alors que celles de plantes sont quasi insensibles aux didéoxyribonucléotides (ddNTP) auxquels les Poly sont particulièrement sensibles (144). Cette particularité des Pol des organelles a donc amené un grand intérêt pour l'identification précise et le clonage des gènes correspondants.

1.4.2 Clonage des gènes des polymérases de l'ADN

La première Pol des organelles clonée a été celle du riz (*Oryza sativa*). Une banque de données contenant une collection de séquences exprimées chez la plante a été criblée pour trouver un homologue de la polymérase de type I (PolI) de la bactérie *E. coli* (29). Un gène a été identifié et nommé *OsPolI-like* (ou *OsPolIA*). L'analyse de la séquence de la protéine encodée montre une similarité plus prononcée pour les Pol bactériennes de type I que pour les Poly des animaux. L'étude de l'expression de ce gène montre qu'il est sollicité dans les tissus en division active, suggérant un rôle important dans la réplication de l'ADN. La protéine *OsPolIA* est localisée dans le plastide et comme les Pol isolées biochimiquement lors d'études antérieures, la protéine recombinante *OsPolIA* est

insensible à l'aphidicoline ainsi qu'aux ddNTP, indiquant qu'il pourrait bien s'agir d'une seule et même famille de protéines. Une étude subséquente a montré la présence d'une activité 3'-5' exonucléase (34), semblable à celle des Pol isolées du chloroplaste de l'épinard (140) et de la mitochondrie de *C. album* (143).

Une recherche sur les banques de données d'*Arabidopsis* montre que cette plante possède également deux homologues de Pol bactériennes (AtPolIA et AtPolIB). Comme l'homologue du riz *OsPolIA*, ces deux gènes sont exprimés dans les tissus en réplication active (30). Un gène homologue à *OsPolIA* a également été identifié chez *N. tabacum* grâce à un criblage d'une banque de séquences exprimées (31). Une expérience de buvardage de type Southern a ensuite révélé la présence d'un deuxième gène homologue dans l'ADN génomique. Dans un arbre phylogénétique, les deux protéines NtPolIA et B, comme les autres Pol de plantes, se retrouvent dans une branche spécialisée entre les Pol animales et les Pol bactériennes mais plus proches de ces dernières (Figure 7). Dans la même étude, l'activité de la protéine recombinante NtPolIB a été comparée à l'activité de synthèse de l'ADN des extraits de plastides et de mitochondries de *N. tabacum* (31). Il est encore une fois très intéressant de constater que la protéine recombinante ainsi que les extraits de plastides et de mitochondries ont les mêmes préférences au niveau des substrats, sont inhibés par des agents intercalants (bromure d'éthidium et actinomycine D) de façon similaire et se montrent tous résistants à l'aphidicoline et aux ddNTP. Tous ces résultats semblent indiquer que les Pol responsables de la réplication de l'ADN des organelles sont belles et bien les homologues des Pol. Comme chaque plante semble posséder deux Pol de ce type, il est possible que chacune d'elle soit spécialisée pour un des deux organelles. L'étude de la localisation subcellulaire de ces protéines prend donc toute son importance.

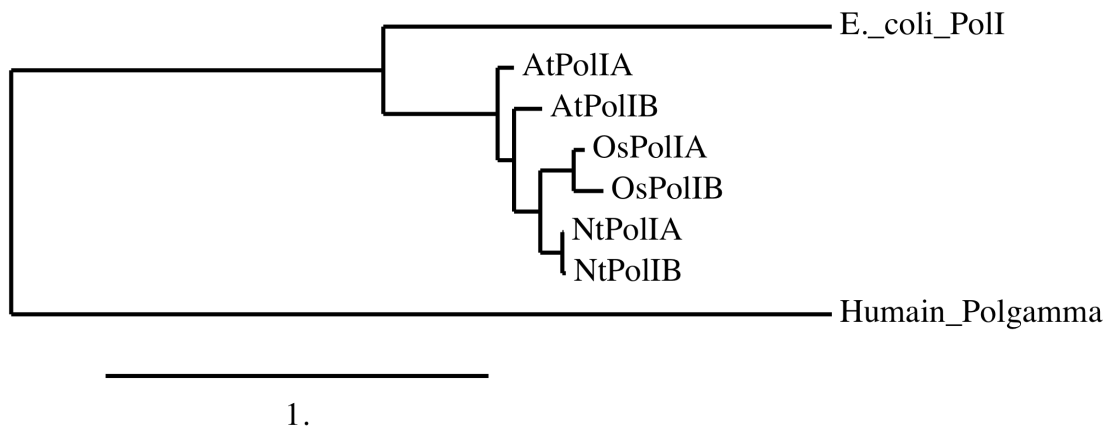


Figure 7. Arbre phylogénétique des polymérase de plantes de type I. Alignement des séquences en acides aminés de six différentes PolI de plantes (At; *Arabidopsis thaliana*, Os; *Oryza sativa*, Nt; *Nicotiana tabacum*) avec les séquences de la PolI de *E. coli* et la polymérase γ de l'humain. L'arbre a été construit avec le programme phylogeny.fr (145).

1.4.3 Localisation intracellulaire

Puisque les plantes à fleurs semblent toutes posséder deux gènes codant pour des Poll, il se peut que chacun soit spécifique pour un seul compartiment. Le premier membre de la famille a effectivement été rapporté comme étant localisé spécifiquement au niveau des plastides (29). Par contre, une expérience de localisation subcellulaire des deux Poll d'*Arabidopsis* fusionnées à la protéine GFP a révélée une localisation dans les chloroplastes pour AtPollA et une double localisation pour AtPollB (32). Cependant, le même groupe a démontré par la suite qu'en utilisant différents sites d'initiation de la traduction le messager du gène AtPollA pouvait produire une protéine qui cible soit les mitochondries ou les plastides (33, 85). Dans le cas des Pol de *N. tabacum*, les expériences de localisation subcellulaire utilisant des fusions GFP ainsi que les buvardages de type Western à la suite d'un fractionnement des cellules ont montré la double localisation des protéines (31). Il semble donc que les deux protéines qu'on retrouve dans les plantes à fleurs se retrouvent simultanément dans les deux organelles.

1.5 Survol du projet de recherche

1.5.1 Objectifs

Le présent travail vise d'abord à contribuer à une meilleure connaissance des mécanismes de maintien de l'ADN dans les organelles de plantes. Ce but très général a été poursuivie à travers deux objectifs plus précis : L'étude des protéines de la famille Whirly et l'étude des polymérase de l'ADN bactériennes à l'intérieur des organelles. Les deux sections précédentes (1.3 et 1.4) illustrent l'état des connaissances vis-à-vis ces deux familles au moment de démarrer cette étude. Bien que les deux volets de cette thèse portent sur des protéines très différentes, l'étude de ces deux familles jette un peu de lumière sur le métabolisme de l'ADN des organelles.

Comme mentionné à la section 1.3, les protéines Whirly ont d'abord été identifiées comme des facteurs de transcription du noyau de la cellule. Cependant, ces protéines sont

invariablement synthétisées avec un peptide signal à leur extrémité N-terminale. Elles sont donc toutes prédites pour se retrouver soit dans la mitochondrie, soit dans le plastide. Le mécanisme grâce auquel les protéines Whirly pourraient être transportées au noyau suite à leur entrée dans les organelles est encore inconnu à ce jour. Ce manque de donné rend impossible la prédiction des membres de la famille qui posséderont un rôle nucléaire. Par contre, il est clair que chaque plante possède au moins un Whirly dans la mitochondrie et un autre dans le plastide. Cependant, très peu d'études se sont penchées sur le rôle potentiel de ces protéines dans les organelles. Notre étude de la famille Whirly ne repose sur aucune hypothèse préalable quant au rôle de ces protéines si ce n'est leur présence constitutive dans les organelles et leur grande affinité pour les acides nucléiques démontrée à maintes reprises. Nous proposons ainsi de découvrir le ou les rôle(s) des protéines Whirly dans les organelles de plantes en espérant découvrir une nouvelle facette du métabolisme de l'ADN dans ces compartiments.

Les polymérases de l'ADN sont, par définition, des enzymes essentiels pour le maintien de l'ADN pour n'importe quel organisme. Beaucoup d'hypothèses ont été émises par rapport à l'identité des polymérases responsables de la réplication de l'ADN dans les organelles de plantes (voir section 1.4). Cependant, aucune preuve définitive n'a permis de révéler les protéines responsables. Les différents indices accumulés à ce jour suggèrent fortement que les enzymes en question font partie d'une famille de protéines ressemblant à la polymérase de l'ADN de type I chez *E. coli*. Notre second objectif est donc de confirmer le rôle des polymérases de type bactériennes dans le métabolisme de l'ADN des organelles de plantes.

1.5.2 Approche expérimentale

Pour atteindre notre premier objectif de découvrir le rôle des Whirly dans les organelles, nous avons profité de la plante modèle *Arabidopsis thaliana* pour utiliser une approche de génétique inverse. Le Chapitre 2 présente les résultats d'une telle approche pour l'étude des deux gènes encodant des protéines qui se retrouvent au plastide, *AtWhy1* et *AtWhy3*. La mutation combinée de ces deux gènes entraîne un dérèglement de la recombinaison au niveau de l'ADN du plastide qui peut mener à un phénotype de variégation. Le type de

recombinaison en question est la réparation induite par bris et guidée par microhomologies. Le même type de dérèglement a également été détecté dans des plants de maïs mutés pour le gène *ZmWhy1* qui encode une protéine Whirly qui se retrouve également dans le plastide. Cette confirmation dans un deuxième organisme a permis d'étendre nos conclusions. Étant donné le lien entre le type de réarrangements que nous avons mesuré et les bris dans l'ADN chez d'autres organismes, nous proposons un rôle pour ces protéines dans la réparation des bris double brin.

Dans le but de tester notre hypothèse, il a été nécessaire de trouver un moyen d'introduire des bris dans l'ADN des organelles de plantes. Le Chapitre 3 introduit l'antibiotique nommé ciprofloxacine comme un agent pouvant introduire des bris double brin dans les mitochondries et les plastides de plantes. En accord avec notre hypothèse, le double mutant des gènes *AtWhy1* et *AtWhy3* est plus sensible à cet agent qu'une plante de type sauvage. L'antibiotique induit une augmentation du nombre de réarrangements dans le génome des organelles de plantes. L'apparition de réarrangements est plus marquée chez le double mutant mais également chez les plantes mutantes pour le gène *AtWhy2*, un Whirly mitochondriale. Ce résultat prouve donc la conservation du rôle des protéines Whirly dans les deux organelles. Puisque nous proposons un rôle dans la réparation de l'ADN pour les protéines Whirly, il devient alors nécessaire de prouver que ces protéines peuvent interagir avec n'importe quelle séquence d'ADN. La structure tridimensionnelle d'une protéine Whirly en complexe avec un brin d'ADN a donc été résolue par cristallographie et diffraction de rayons X. Les résultats sont venus confirmer le manque de spécificité de la protéine pour la séquence d'ADN, ce qui vient appuyer notre modèle. Les protéines Whirly deviennent donc les premiers facteurs prouvés pour être impliqués dans la réparation des bris double brin de l'ADN des organelles.

Notre second objectif visait à confirmer la participation d'une famille de polymérase de l'ADN bactériennes dans le métabolisme des organelles. Notre approche a été, une fois de plus, la génétique inverse et les résultats sont rapportés au Chapitre 4. Nous avons obtenus deux lignées mutantes pour deux gènes d'*Arabidopsis thaliana*, *PollA* et *PollB*. Aucun phénotype n'est observable chez ces mutants sauf pour un allèle de *PollB* qui induit une croissance ralentie par rapport à la plante sauvage. La combinaison d'une mutation de

PollA avec une mutation de *PollB* est cependant létale, appuyant un rôle essentiel pour cette famille de gènes. Pour lier ces mutations avec un défaut dans la réplication de l'ADN nous avons mesuré le nombre de copies relatives des génomes des organelles par rapport au génome nucléaire. Tous les mutants possèdent moins de copies du génome des organelles lorsque mesuré à trois et quatre jours après la germination. Nous avons également soumis les mutants à la ciprofloxacine et les mutants de *PollB*, et non ceux de *PollA*, se sont révélés plus sensibles. Il semble donc que les deux polymérases se partagent la tâche de la réplication de l'ADN mais que l'enzyme *PollB* se soit spécialisée dans la réparation de l'ADN. Cette conclusion est appuyée par le croisement des mutants *PollB* avec le double mutant *AtWhy1/AtWhy3* qui résulte en une plante naine, pâle et extrêmement sensible aux bris double brin.

CHAPITRE 2. Premier article

Whirly Proteins Maintain Plastid Genome Stability in *Arabidopsis*

Article publié dans la revue Proceedings of the National Academy of Sciences of the United States of America, volume 106, pages 14693-14698, le 25 août 2009.

2.1 Mise en contexte

Au moment des travaux rapportés dans cet article, très peu de choses étaient connues sur le rôle des protéines Whirly à l'intérieur des organelles de plantes. Une étude antérieure de notre groupe avait démontré que la surexpression de la protéine mitochondriale AtWhy2 entraînait une diminution de l'ADNmt (132). Cet effet a été associé à la capacité de cette protéine à se lier à l'ADN de façon non spécifique et donc d'entraver le métabolisme normal de l'ADN. Une autre étude portant sur la protéine plastidique ZmWhy1 du maïs (133) a utilisé des mutants du gène correspondant qui sont également sujet de l'article présenté ici. Le groupe du Dr. Alice Barkan a d'abord isolé ces plantes mutantes qui montraient de sévères problèmes développementaux. Ces problèmes ont été associés à une forte diminution de la quantité d'ARNr au niveau des plastides. Cette étude n'a cependant pas pu démontrer une implication directe de la protéine ZmWhy1 dans le métabolisme de l'ARNr. Les auteurs ont donc conclu que les effets observés chez les mutants devaient être causés par un rôle plus général des protéines Whirly dans le métabolisme de l'ARN ou de l'ADN (réplication, transcription, etc.) (133).

Cette étude a débuté avec l'isolation de mutants pour les gènes *AtWhy1* et *AtWhy3*, encodant des ptWhirly chez *Arabidopsis*. Une fois les deux mutants isolés et croisés, j'ai collaboré avec mon collègue, Alexandre Maréchal, pour le restant du projet. Ma contribution a été l'isolation, la caractérisation et les croisements des différents mutants d'*Arabidopsis* et de maïs. J'ai également réalisé les expériences de PCR pour détecter les réarrangements de l'ADN. AM a isolé les lignées variégues, réalisé les buvardages de types Southern, l'étude statistique des plantes variégues, le clonage et l'analyse des produits de PCR, le retardement sur gel et il a supervisé Félix Véronneau-Lafortune qui effectuait un stage au laboratoire. FVL a réalisé l'expérience d'immunoprécipitation de la chromatine. Alexandre Joyeux a réalisé les images de microscopie confocale. AM, B. Franz Lang, Normand Brisson et moi-même avons écrit et révisé cet article.

Whirly Proteins Maintain Plastid Genome Stability in *Arabidopsis*

Alexandre Maréchal^{*}, Jean-Sébastien Parent^{*}, Félix Véronneau-Lafortune,
Alexandre Joyeux, B. Franz Lang and Normand Brisson

***These authors contributed equally to this work.**

Department of Biochemistry, Université de Montréal, PO Box 6128, Station
Centre-ville, Montréal, Québec, Canada, H3C 3J7.

Corresponding author: Normand Brisson

Department of Biochemistry, Université de Montréal, PO Box 6128, Station Centre-ville,
Montréal, Québec, Canada, H3C 3J7 Phone : 514-343-6984 Fax : 514-343-2210

BIOLOGICAL SCIENCES: Plant Biology

20 text pages, 4 figures, and 4 tables.

Author contributions: A.M., J-S. P., A.J., B.F.L. and N.B. designed research; A.M., J-S. P., F.
V.-L. and A.J. performed research; A.M., J.-S. P., A.J., B.F.L. and N.B. analyzed data; and
A.M., J-S. P., A.J., B.F.L. and N.B. wrote the paper.

The authors declare no conflict of interest.

2.2 Abstract

Maintenance of genome stability is essential for the accurate propagation of genetic information and for cell growth and survival. Organisms have therefore developed efficient strategies to prevent DNA lesions and rearrangements. Much of the information concerning these strategies has been obtained through the study of bacterial and nuclear genomes. Comparatively, little is known about how organelle genomes maintain a stable structure. Here we report that the plastid-localized Whirly single-stranded DNA binding proteins are required for plastid genome stability in *Arabidopsis*. We show that a double knockout of the genes *AtWhy1* and *AtWhy3* leads to the appearance of plants with variegated green/white/yellow leaves, symptomatic of non-functional chloroplasts. This variegation is maternally-inherited, indicating defects in the plastid genome. Indeed, in all variegated lines examined, reorganized regions of plastid DNA are amplified as circular and/or head-tail concatemers. All amplified regions are delimited by short direct repeats of 10 to 18 bp, strongly suggesting that these regions result from illegitimate recombination between repeated sequences. This type of recombination occurs frequently in plants lacking both Whirlies, to a lesser extent in single knockout plants and rarely in wild-type individuals. Maize mutants for the *ZmWhy1* Whirly protein also show an increase in the frequency of illegitimate recombination. We propose a model where Whirlies contribute to plastid genome stability by protecting against illegitimate repeat-mediated recombination.

2.3 Introduction

Plastids play diverse and essential roles in plants. Despite this central importance, surprisingly little is known about even the most basic aspects of the plastid genome structure, maintenance and propagation. For example, while the textbook depiction of plastid DNA (ptDNA) is that of a genome-sized circular DNA molecule, recent evidence suggests instead that most of the ptDNA is organized in concatenated, branched linear forms, with T4 phage-like features (1). This change of perception of plastid genome

architecture requires a re-evaluation of the current rolling-circle model for plastid genome replication. It is now considered that a recombination-dependent replication process is responsible for the branched, multi-genomic structures present in plastids (1) due to its exposure to radiation and reactive oxygen species, the plastid genome is expected to accumulate mutations at a high rate. This stresses the importance of efficient DNA-replication, recombination and repair (DNA-RRR) pathways in these organelles (2). However, to date the mechanisms and enzymes involved in these pathways remain poorly characterized.

Evidence for recombination in plastid genomes abound in the literature (2). An example is the recombination between the large inverted repeat sequences present in many plastid genomes (3). This flip-flop recombination is responsible for the two isomers of ptDNA, which differ only with respect to the orientation of the single-copy regions. More direct evidence of recombination comes from plastid transformation experiments, which demonstrate that foreign DNA is integrated into ptDNA by homologous recombination (4).

Homologues of bacterial genes involved in DNA-RRR are present in the nuclear genome of plants, and some of their encoded proteins are targeted to plastids. These include the *recA* homologs *RECA1* (5) and *RECA2*, whose disruption is lethal in *Arabidopsis* (6), a Rec Q-like DNA helicase from rice (7), and genes for a gyrase A-like and two gyrase B-like subunits in *Arabidopsis* (8). Recently, two homologs of OSB1, a single-stranded DNA (ssDNA) binding protein that regulates recombination in mitochondria (9), were shown to localize to plastids. However, no role has yet been ascribed to these proteins. Finally, homologs of RPA, another ssDNA-binding protein family which plays an essential role in mammalian DNA-RRR, have recently been identified in plants. One member of this family is targeted to plastids (10).

Similar to many DNA-RRR proteins, Whirlies form a small family of ssDNA-binding proteins (11). They are involved in a variety of phenomena, ranging from pathogen defence (12) to telomeric homeostasis (13). In *Arabidopsis*, three Whirly genes are present and their proteins localize to organelles; AtWhy1 and AtWhy3 are targeted to plastids and AtWhy2 to mitochondria (14, 15).

Recent evidence indicates that Whirlies bind organelle DNA without apparent sequence-specificity *in vivo*. In *Arabidopsis*, AtWhy2 binds to many regions of the mitochondrial genome with no obvious sequence consensus (15). Similarly, in maize, the plastid-localized ZmWhy1 interacts with DNA from throughout the plastid genome (16). Knockdown mutations of *ZmWhy1* lead to ivory or pale-green plants, indicating that this Whirly is involved in chloroplast biogenesis. This phenotype was attributed to a defect in the maturation of the *atpF* and *23S* ribosomal RNAs, but the authors did not rule out participation of ZmWhy1 in DNA recombination or repair.

To better understand the role of plastid-targeted Whirlies (ptWhirlies), we characterized an *Arabidopsis* double knockout (KO) line of the *AtWhy1* and *AtWhy3* genes (KO1/3). Variegation patterns, which appear on leaves in about 4.6% of the progeny, correlated with the selective rearrangement and amplification of large regions of the plastid genome. We show that the rearrangements are produced by illegitimate recombination at short direct repeats that border the amplified regions in intact ptDNA. We suggest that AtWhy1 and AtWhy3 function as anti-recombination proteins, contributing to safeguard plastid genome integrity.

2.4 Results

Arabidopsis ptWhirlies are Involved in Chloroplast Biogenesis

To investigate the role of Whirlies in plastids, we obtained mutant plants that no longer produce the AtWhy1 (KO1) and/or AtWhy3 (KO3) proteins (Figure 8A and B). Only one Whirly is detected in the single KOs, and no Whirlies are detected in KO1/3 extracts. We then monitored the ssDNA-binding activity of Whirlies by electrophoretic mobility shift assay. As shown in Figure 12, a strong DNA-binding signal is observed with Col-0 (wild-type) proteins, whereas less intense signals are detected with KO1 and KO3, respectively. No signal was observed in KO1/3, confirming the absence of ptWhirlies in these plants.

The KO1 and KO3 plants have no apparent phenotype. Interestingly, while most KO1/3 plants also have a wild-type (WT) appearance, some individuals have a smaller size and a variegated phenotype with white/yellow sectors on some leaves (Figure 8C, upper panel).

In these sectors, a strong diminution in chlorophyll autofluorescence is observed (Figure 8C, lower panel). Examination of plastids by electron microscopy reveals that in the white sectors, thylakoid stacks are replaced by large round vesicles, indicating that plastid development is compromised (Figure 8D, right panels). By contrast, plastids from green sectors of the same leaf appear normal (Figure 8D, left panels). We then evaluated the frequency of sectoried individuals in large populations of Col-0 and Whirly mutant plants (Tableau I). While no variegated individuals were recovered from Col-0, KO1 and KO3 populations, 4.6 % of KO1/3 plants had at least one variegated leaf sector (Tableau I). These data indicate that elimination of both ptWhirlies triggers changes that ultimately lead to strong interference with chloroplast development and function.

The Variegated Phenotype of KO1/3 Plants is Maternally-Inherited

The severity of the variegation phenotype in the KO1/3 population is variable. While some plants exhibit chlorosis on a single leaf, others have most of their leaves covered by yellow/white sectors (data not shown and Figure 8E). Phenotypic differences were also observed between different variegated individuals, suggesting that the defects responsible for variegation differ from one plant to another (Figure 8E). From this first generation of variegated KO1/3 plants, we selected two lines with a strong variegation phenotype (Var A and Var B) and set out to define the molecular basis of impaired plastid development.

Maternal inheritance of variegation is often linked to modifications of organelle genomes (9, 17, 18). Crosses were performed between variegated line Var B and Col-0 plants to determine the inheritance mode of the variegation. When variegated plants were used as male parents, no variegation was observed in the heterozygous F1 progeny. However, when Var B plants were used as female in the same cross, variegation was found in 46% of the progeny (47 out of 103 F1 plants). This indicates that variegation is maternally inherited and that reintroduction of Whirlies is unable to completely rescue the plastid defect, suggesting that the plastids in variegated sectors are irreversibly damaged, most likely at the genetic level.

Amplification of Reorganized ptDNA Regions in Variegated Plants

Since Whirlies bind DNA in organelles ((15, 16); Figure 13 and 14), it is likely that variegation in KO plants results from mutations in their ptDNA. We thus searched for ptDNA rearrangements by DNA hybridization. Although no change was detected when comparing HindIII-digested Col-0 to non-variegated (green) KO1/3 DNA, unique amplified regions were identified in digested ptDNA of the Var A and Var B lines (Figure 9A and B). Amplified DNA was estimated to be 10 to 25 times more abundant than WT ptDNA. In addition, for both variegated lines, new HindIII fragments (~5.8 kb in Var A and ~6.3 kb in Var B) were detected using probes located at the extremities of the amplified regions. These new bands indicate that the amplified regions are reorganized either as circular molecules and/or head-tail concatemers (Figure 9C).

Reorganization of ptDNA is Caused by Recombination Between Short Direct Repeats

To map the extremities of the amplified DNA, we designed PCR oligonucleotides close to the ends of the amplified regions of Var A and Var B, facing opposite directions (outward-facing PCR; Figure 9C). In a WT plastid genome, these primers would yield no product. However in a rearranged genome where circular or head-tail concatemers are present, we predict a new fragment containing both extremities of the amplified regions. As expected, when PCR is performed with DNA from Col-0 and green KO1/3 plants, no product is observed with both Var A and Var B primers (Figure 9D). However, when the same primers are used on Var A and Var B DNA, amplification products appear in both cases. Interestingly, more amplified products are detected in variegated *versus* non-variegated leaves of the same plants, suggesting that the appearance of defective chloroplasts is related to the abundance of reorganized ptDNA molecules (Figure 9D; compare leaves 4 and 8).

The position of the ends of amplified regions was determined by cloning and sequencing the PCR products. These positions matched perfectly with the HindIII digestion patterns of the reorganized regions (Figure 9A, B, C and Tableau II). Remarkably, the regions of WT plastid genome that correspond to the amplified regions in Var A and Var B are bordered by short direct repeats of 10 and 14 bp, respectively (Tableau II). In the PCR fragments, both ends of the amplified regions are joined by a single repeat, indicating that

recombination occurred between these short sequences. Analysis of three additional variegated lines also revealed unique amplified regions with ends joined by recombination at short repeats (Tableau II). Interestingly, in the Var C and Var E lines, the bordering repeats are not identical and carry a few mismatches. Altogether, our data strongly suggest that variegation is induced by an overabundance of amplified recombined ptDNA regions.

Additional DNA hybridization experiments were performed to determine the arrangement of the amplified DNA. DNA from Col-0, green KO1/3 and Var A plants was digested either with XhoI and PstI restriction enzymes that cut on both sides of the Var A region in WT ptDNA, or with KpnI which cleaves once in this region (Figure 9E and F). A probe specific for the Var A amplified region was used for detection. Intact Col-0 and green KO1/3 DNA migrated within the unresolved compression zone (above ~ 10 kb). Six bands were detected in undigested Var A DNA, including a fast-migrating band of ~ 4.5 kb. When DNA was digested with XhoI and PstI, the Var A specific bands were unaffected (Figure 9F), indicating that these DNA molecules are extrachromosomal. Finally, after KpnI digestion, all Var A specific bands collapsed into a single 6.25 kb band, showing that these forms are concatemers of the same repeat unit. This also suggests that the fast migrating band in undigested Var A DNA corresponds to a circular molecule, possibly a monomeric one. Similar results were obtained with the Var B line (data not shown). We thus conclude that amplified ptDNA is present as both circular and/or head-tail concatemers in variegated plants.

Illegitimate Recombination is Increased in the Absence of *Arabidopsis* ptWhirlies

Our results suggest that the absence of ptWhirlies induces plastid genome instability through an increase in recombination between short direct repeats. One could therefore expect to detect low-level recombination events in non-variegated Whirly mutant plants. We tested this hypothesis using the outward-facing PCR approach described above, on each of four independent DNA pools from Col-0, KO1, KO3 and green KO1/3 plants. Representative results of the PCR amplification are shown in Figure 10 (panels A, B and C). Rearranged products were observed in all genotypes.

Cloning and sequencing of rearranged DNA confirmed that illegitimate recombination is strongly increased in green KO1/3 plants, where 40 different recombination products are identified out of 30 PCR reactions (Tableaus III, IV and V). Recombination products were also detected in single KOs, although less frequently, indicating that the depletion of *AtWhy1* or *AtWhy3* is sufficient to increase spurious recombination. Surprisingly, two recombination events were detected in Col-0 samples. These were present in all genotypes tested (Tableau III and V), suggesting the presence of a small sub-population of rearranged ptDNA even in WT plants.

Illegitimate Recombination is Increased in Maize Whirly Mutants

We verified in three maize *ZmWhy1* mutant lines whether mutation of monocot ptWhirlies also affects plastid genome stability. The lines *ZmWhy1-1*, *ZmWhy1-2* and the heteroallelic progeny of complementation crosses *ZmWhy1-2/-1* all have a reduced level of the plastid-localized protein ZmWhy1. These lines exhibit ivory (*ZmWhy1-1*), pale green (*ZmWhy1-2*) and intermediate (*ZmWhy1-2/-1*) leaf phenotypes (16). Our Southern hybridizations did not reveal amplified plastid genome regions in any of the three lines (Figure 17), indicating that the defect in chloroplast biogenesis in *ZmWhy1* mutants is not linked to the presence of amplified ptDNA regions. Outward-facing PCR analysis of the B73 control maize line revealed a single short-repeat mediated illegitimate recombination event for the 19 primer pairs tested (Tableau VI). The same primers revealed 3 events in lines *ZmWhy1-2* and *ZmWhy1-2/-1*, respectively. Illegitimate recombination was highest in the most severe mutant line (*ZmWhy1-1*), with a total of 14 different events (Tableau VI). These results suggest that Whirlies are also involved in stabilizing the plastid genome in maize.

2.5 Discussion

Whirlies Prevent Short Homology-Dependent Illegitimate Recombination

Accumulation of reorganized DNA through illegitimate recombination has also been observed in bacteria. In *E. coli*, stress can induce selective amplification of genome regions conferring tolerance to the applied stress, a phenomenon called adaptive amplification.

These regions (~10 to 30 kb) are first duplicated through illegitimate recombination between microhomologous repeats (5-15 bp), and are then further replicated through a mechanism requiring the homologous recombination machinery (reviewed in (19)). Another recombination process, called 'short homology-dependent illegitimate recombination' exists in bacteria (reviewed in (20)). It also requires microhomology (3-20 bp) at recombination sites, and its occurrence is increased by various DNA-damaging agents. Similarly, microhomology-mediated recombination has recently been proposed to trigger segmental DNA duplications that are linked to some genetic diseases in humans and to copy number variations (CNV) among individual genomes (21-23). It was also observed in yeast cells undergoing segmental duplications (24). Interestingly, in bacteria, DNA-RRR proteins such as the subunits of the UvrAB complex, the RecQ helicase, the histone-like H-NS and the single-stranded DNA binding protein SSB can suppress microhomology-mediated recombination (25-29). However, the genetics requirements for this suppression may vary with the system used to monitor recombination. We propose a similar role for Whirlies in plastids, where they contribute to maintain genome stability by preventing accumulation of illegitimate recombination products.

A Model for the Variegation in Arabidopsis Plants lacking ptWhirlies

The direct repeat sequences at Arabidopsis ptDNA junctions are too short to serve as substrates for homologous recombination ((30, 31)). Recently, an alternative model was proposed to explain how microhomologies may generate large DNA rearrangements in *E. coli*, yeast and human (21). In this model, the collapse of a replication fork leads to the generation of a 3' protruding ssDNA end. This 3' tail may anneal to a microhomologous ssDNA sequence at a different location, allowing re-initiation of replication at this new site. When such template switching occurs between two different DNA molecules, either duplications (switch to a position behind the fork) or deletions (switch to a position ahead of the fork) will occur. Alternatively, switching on the same DNA molecule behind a stalled replication fork leads to rolling-circle replication which produces concatemers of subgenomic regions, and eventually circular DNA products (Figure 9E). In addition to duplication/circularization events that were detected by DNA hybridization and outward-facing PCR, deletions were also identified using inward-facing PCR in Arabidopsis Whirly

mutants (Tableau V). In summary, the above replication-based mechanism might explain the variety of microhomology-mediated rearrangements observed in the plastid genome of Whirly mutants.

Based on our results, we propose the following model for the emergence of variegation in KO1/3 plants (Figure 11). In Col-0, the genome surveillance machinery repairs DNA lesions and, with the help of Whirlies, prevents accumulation of illegitimate DNA recombination products, which are rarely detected in these plants. In the single KO plants, the absence of one ptWhirly is enough to induce a low-level accumulation of rearranged molecules. However, when both Whirlies are eliminated, aberrant molecules accumulate to high levels. Some of these molecules, most likely circularized products, could replicate independently of the main plastid genome and eventually become overabundant. In each of the five variegated lines analysed in detail (Tableau II) only one rearranged molecule is amplified. Although a certain level of these molecules can be tolerated by the plastid, as demonstrated by their presence in green tissue of variegated plants (Figure 9D), their accumulation eventually gives rise to plastids that are non-functional, thus leading to the appearance of variegated leaf sectors. No deletions resulting in variegation were detected, suggesting that their accumulation is either lethal and/or that they are more easily eliminated by gene conversion (32).

Conservation of the Role of ptWhirlies in the Maintenance of Plastid Genome Stability

The finding that illegitimate recombination mediated by short repeats is also enhanced in a severely affected maize Whirly mutant line suggests a conserved function for Whirlies in maintaining plastid genome stability. *ZmWhy1* hypomorphic mutants show a less clear trend than the *ZmWhy1-1* mutant, which might be due to the presence of a higher residual level of ZmWhy1 protein in these lines (16). This situation resembles that of the single KO1 and KO3 *Arabidopsis* Whirly mutants, which show less illegitimate recombination than the double KO. The low level of ZmWhy1 protein present in all *ZmWhy1* mutants may also explain why no amplified ptDNA regions are found in these plants. In fact, the level of residual ptWhirly protein in the maize mutants appears similar to that present in the *Arabidopsis* KO1 mutant, which also does not contain amplicons. Amplified regions were only observed in KO1/3, the only line analysed which is completely

devoid of ptWhirlies, suggesting that a low level of Whirlies is sufficient to prevent amplification of recombined ptDNA molecules.

The absence of detectable amplicons and the severe first generation phenotypes in the maize Whirly mutants indicate that a different mechanism contributes to the formation of defective chloroplasts in this species. Actually, the *ZmWhy1* mutants exhibit a reduced content of plastid ribosomes, with the albino *ZmWhy1-1* seedlings almost completely lacking plastid rRNA (16). This ribosome deficiency was proposed to result from a defect in the biogenesis of the large ribosomal subunit due to an aberrant 23S rRNA metabolism, suggesting a role for Whirlies in plastid RNA metabolism (16). Although we cannot rule out such a role in *Arabidopsis*, it is unlikely that the more subtle variegation phenotype observed in this species is due to aberrant rRNA metabolism as no major changes are detected in 23S and 16S rRNA levels in the *Arabidopsis* ptWhirly mutants (Figure 18). Thus, Whirlies appear to form a flexible family of single-stranded nucleic acid binding proteins that can fulfill a variety of roles, depending on the cellular context and/or plant species.

Characteristics of the Short Direct Repeats

Some of the direct repeats reported here appear to be particularly prone to illegitimate recombination as they are used repeatedly and at multiple different positions of the plastid genome (see related events in Tableaux V and VI). This suggests that a specific sequence motif or structure promotes their use as substrates. Homopolymeric A/T stretches appear to be over-represented among the direct repeats (Tableaux II, III, V and VI). This type of sequence produces unusual structures called slipped-strand DNA, which can cause replication stalling (reviewed in (33)). Stalled replication forks can be reinitiated by template-switching, and can also generate double-strand ends (reviewed in (34)), potentially lead to an increase in illegitimate recombination. Alternatively, these poly A/T sequences could have a greater tendency to melt, providing easier access to invading ssDNA 3' tails and thus favouring ptDNA rearrangements.

Whirlies Could Prevent Illegitimate Recombination by Protecting ssDNA

Many ssDNA binding proteins bind non-specifically to DNA and play important roles in recombination processes. Bacterial SSBs and their eukaryotic nuclear counterparts such as

RPA play essential protective roles in genome biology by protecting ssDNA from damage and preventing spurious DNA annealing. They also serve to recruit a large number of genome maintenance proteins (reviewed in (35, 36)). In *E. coli*, SSB also suppresses short homology-dependent illegitimate recombination (27, 28). It is thus tempting to speculate that in plastids, Whirlies play a role similar to that of SSB in *E. coli*. This is supported by our data (Figure 14 and (15)) and those of Prikryl *et al.* (16), indicating that Whirlies bind DNA in a non-sequence specific manner. In addition, similar to bacterial SSBs, AtWhy1 protects ssDNA against degradation by nucleases *in vitro* (Cappadocia, L., Maréchal, A., Lepage, E., Sygush, J. and Brisson, N., unpublished results). We therefore propose that in plastids, the absence of Whirlies enhances the availability of free ssDNA. This might lead to increased DNA damage due to the vulnerability of the exposed DNA, with a subsequent increase in recombination-repair mechanisms. More importantly, the damaged ssDNA might increase the rate of illegitimate recombination by increasing the frequency of collapsed replication forks (reviewed in (37)). In any case, stresses able to induce DNA damage would also increase ssDNA availability, as happens after processing of double-strand breaks, leading to increased recombination (28).

Our experiments define Whirlies as new and important components of the plastid genome maintenance machinery. As a follow-up, it will be interesting to determine if stresses that alter the integrity of ptDNA also lead to increased illegitimate recombination both in Whirly mutants and in plants deficient for putative plastid DNA-RRR proteins. This will help precise the function of Whirlies in DNA protection/repair processes and will permit identification of new regulators of plastid genome stability.

2.6 Materials and Methods

Mutant characterization

The SALK Institute Genomic Analysis Laboratory provided the sequence-indexed T-DNA insertion line (38). The Seattle TILLING Project (39) provided plants with mutations in the *AtWhy3* gene. The maize mutant lines *ZmWhy1-1*, *ZmWhy1-2* and *ZmWhy1-2/-1* were obtained from Dr. Alice Barkan (16).

Antibody production

Recombinant AtWhy1 was purified as described previously (40). Rabbits were immunized and antiserum was collected. For protein gel blot analysis, the antiserum was used at a concentration of 1:4000.

DNA gel blots

DNA was isolated from plants using a Cetyl trimethylammonium bromide DNA extraction protocol (41). Running of the samples and blotting of the gels was performed as described (15).

Detection of DNA rearrangements by PCR in *Arabidopsis*

PCR reactions were conducted using the Taq polymerase from Genscript and a series of outward-facing oligonucleotides spaced by approximately 10-20 kb. Inward-facing primers spaced by 10-20 kb were used to detect deletion events. Thirty reactions were performed on each of four independent DNA samples from Col-0, KO1, KO3 and green KO1/3 plants and analysed by gel electrophoresis. All visible DNA bands were isolated, cloned and sequenced. See Supplementary Information.

2.7 Acknowledgments

We thank Dr. C.S. Hardtke for his help with the *Arabidopsis* crosses, R.M. Gaafar for help with genotyping, S. Grondin, P. Tan and A. Moreau for technical assistance and L. Forget for the sequencing of ptDNA. We thank Dr. A. Barkan for the generous gift of the *ZmWhy1* mutant lines. A.M. was supported by scholarships from the Natural Science and Engineering Research Council of Canada (NSERC) and from the Fonds Québécois de la Recherche sur la Nature et les Technologies. This research was supported by a grant from NSERC.

2.8 References

1. Bendich AJ (2004) Circular chloroplast chromosomes: the grand illusion. *Plant Cell* 16(7):1661-1666 .
2. Day A, Madesis, P. (2007) DNA replication, recombination, and repair in plastids. *Cell and Molecular Biology of Plastids, Topics in Current Genetics*, ed Bock R (Springer-Verlag Berlin Heidelberg, Leipzig), Vol 19, pp 65-119.
3. Palmer JD (1983) Chloroplast DNA exists in 2 orientations. *Nature* 301:92-93.
4. Staub JM & Maliga P (1992) Long regions of homologous DNA are incorporated into the tobacco plastid genome by transformation. *Plant Cell* 4(1):39-45 .
5. Cao J, Combs C, & Jagendorf AT (1997) The chloroplast-located homolog of bacterial DNA recombinase. *Plant Cell Physiol* 38(12):1319-1325 .
6. Shedge V, Arrieta-Montiel M, Christensen AC, & Mackenzie SA (2007) Plant mitochondrial recombination surveillance requires unusual RecA and MutS homologs. *Plant Cell* 19(4):1251-1264 .
7. Saotome A, *et al.* (2006) Characterization of four RecQ homologues from rice (*Oryza sativa* L. cv. Nipponbare). *Biochem Biophys Res Commun* 345(4):1283-1291 .
8. Wall MK, Mitchenall LA, & Maxwell A (2004) *Arabidopsis thaliana* DNA gyrase is targeted to chloroplasts and mitochondria. *Proc Natl Acad Sci U S A* 101(20):7821-7826 .
9. Zaegel V, *et al.* (2006) The plant-specific ssDNA binding protein OSB1 is involved in the stoichiometric transmission of mitochondrial DNA in *Arabidopsis*. *Plant Cell* 18(12):3548-3563 .
10. Ishibashi T, Kimura S, & Sakaguchi K (2006) A higher plant has three different types of RPA heterotrimeric complex. *J Biochem* 139(1):99-104 .
11. Desveaux D, Maréchal A, & Brisson N (2005) Whirly transcription factors: defense gene regulation and beyond. *Trends Plant Sci* 10(2):95-102 .
12. Desveaux D, Després C, Joyeux A, Subramaniam R, & Brisson N (2000) PBF-2 is a novel single-stranded DNA binding factor implicated in PR-10a gene activation in potato. *Plant Cell* 12(8):1477-1489 .

13. Yoo HH, Kwon C, Lee MM, & Chung IK (2007) Single-stranded DNA binding factor AtWHY1 modulates telomere length homeostasis in Arabidopsis. *Plant J* 49(3):442-451 .
14. Krause K, *et al.* (2005) DNA-binding proteins of the Whirly family in Arabidopsis thaliana are targeted to the organelles. *FEBS Lett* 579(17):3707-3712 .
15. Maréchal A, *et al.* (2008) Overexpression of mtDNA-associated AtWhy2 compromises mitochondrial function. *BMC Plant Biol* 8:42 .
16. Prikryl J, Watkins KP, Friso G, van Wijk KJ, & Barkan A (2008) A member of the Whirly family is a multifunctional RNA- and DNA-binding protein that is essential for chloroplast biogenesis. *Nucleic Acids Res* 36(16):5152-5165 .
17. Chang TL, *et al.* (1996) Characterization of primary lesions caused by the plastome mutator of *Oenothera*. *Curr Genet* 30(6):522-530 .
18. Martinez-Zapater JM, Gil P, Capel J, & Somerville CR (1992) Mutations at the Arabidopsis CHM locus promote rearrangements of the mitochondrial genome. *Plant Cell* 4(8):889-899 .
19. Hastings PJ (2007) Adaptive amplification. *Crit Rev Biochem Mol Biol* 42(4):271-283 .
20. Ehrlich SD (1989) Illegitimate recombination in bacteria. *Mobile DNA*, ed Berg DE, M-Howe, S.M.), pp 799-831.
21. Hastings PJ, Ira G, & Lupski JR (2009) A microhomology-mediated break-induced replication model for the origin of human copy number variation. *PLoS Genet* 5(1):e1000327 .
22. Redon R, *et al.* (2006) Global variation in copy number in the human genome. *Nature* 444(7118):444-454 .
23. Lee JA, Carvalho CM, & Lupski JR (2007) A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. *Cell* 131(7):1235-1247 .
24. Payen C, Koszul R, Dujon B, & Fischer G (2008) Segmental duplications arise from Pol32-dependent repair of broken forks through two alternative replication-based mechanisms. *PLoS Genet* 4(9):e1000175 .

25. Hanada K, Iwasaki M, Ihashi S, & Ikeda H (2000) UvrA and UvrB suppress illegitimate recombination: synergistic action with RecQ helicase. *Proc Natl Acad Sci U S A* 97(11):5989-5994 .
26. Hanada K, *et al.* (1997) RecQ DNA helicase is a suppressor of illegitimate recombination in *Escherichia coli*. *Proc Natl Acad Sci U S A* 94(8):3860-3865 .
27. Mukaihara T & Enomoto M (1997) Deletion formation between the two *Salmonella typhimurium* flagellin genes encoded on the mini F plasmid: *Escherichia coli* *ssb* alleles enhance deletion rates and change hot-spot preference for deletion endpoints. *Genetics* 145(3):563-572 .
28. Reddy M & Gowrishankar J (1997) Identification and characterization of *ssb* and *uup* mutants with increased frequency of precise excision of transposon Tn10 derivatives: nucleotide sequence of *uup* in *Escherichia coli*. *J Bacteriol* 179(9):2892-2899 .
29. Shanado Y, Hanada K, & Ikeda H (2001) Suppression of gamma ray-induced illegitimate recombination in *Escherichia coli* by the DNA-binding protein H-NS. *Mol Genet Genomics* 265(2):242-248 .
30. Lovett ST, Hurley RL, Sutura VA, Jr., Aubuchon RH, & Lebedeva MA (2002) Crossing over between regions of limited homology in *Escherichia coli*. RecA-dependent and RecA-independent pathways. *Genetics* 160(3):851-859 .
31. Liskay RM, Letsou A, & Stachelek JL (1987) Homology requirement for efficient gene conversion between duplicated chromosomal sequences in mammalian cells. *Genetics* 115(1):161-167 .
32. Khakhlova O & Bock R (2006) Elimination of deleterious mutations in plastid genomes by gene conversion. *Plant J* 46(1):85-94 .
33. Mirkin EV & Mirkin SM (2007) Replication fork stalling at natural impediments. *Microbiol Mol Biol Rev* 71(1):13-35 .
34. Michel B, Grompone G, Flores MJ, & Bidnenko V (2004) Multiple pathways process stalled replication forks. *Proc Natl Acad Sci U S A* 101(35):12783-12788 .
35. Zou Y, Liu Y, Wu X, & Shell SM (2006) Functions of human replication protein A (RPA): from DNA replication to DNA damage and stress responses. *J Cell Physiol* 208(2):267-273 .

36. Shereda RD, Kozlov AG, Lohman TM, Cox MM, & Keck JL (2008) SSB as an organizer/mobilizer of genome maintenance complexes. *Crit Rev Biochem Mol Biol* 43(5):289-318 .
37. Persky NS & Lovett ST (2008) Mechanisms of recombination: lessons from E. coli. *Crit Rev Biochem Mol Biol* 43(6):347-370 .
38. Alonso JM, *et al.* (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* 301(5633):653-657 .
39. Till BJ, *et al.* (2003) Large-scale discovery of induced point mutations with high-throughput TILLING. *Genome Res* 13(3):524-530 .
40. Desveaux D, *et al.* (2004) A "Whirly" transcription factor is required for salicylic acid-dependent disease resistance in *Arabidopsis*. *Dev Cell* 6(2):229-240 .
41. Weigel D & Glazebrook J (2002) *Arabidopsis : a laboratory manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.) pp xii, 354 p.
42. Sato S, Nakamura Y, Kaneko T, Asamizu E, & Tabata S (1999) Complete structure of the chloroplast genome of *Arabidopsis thaliana*. *DNA Res* 6(5):283-290 .
43. Kunnimalaiyaan M & Nielsen BL (1997) Fine mapping of replication origins (ori A and ori B) in *Nicotiana tabacum* chloroplast DNA. *Nucleic Acids Res* 25(18):3681-3686.

2.9 Figure Legends

Figure 8. AtWhy1 and AtWhy3 are involved in the biogenesis of chloroplasts. (A) Physical maps of the *AtWhy1* (AT1G14410) and *AtWhy3* (AT2G02740) genes. The position of the T-DNA insertion in the KO1 line is indicated. KO3 is a TILLING line with a mutation that changes a TGG codon to a TGA stop codon in the *AtWhy3* gene. An asterisk indicates the position of the mutation. (B) Protein gel blot analysis of simple and double ptWhirlies

knockout plants. Crude plastid proteins were separated by SDS-PAGE on a 15 % polyacrylamide gel. Whirlies were detected using an anti-AtWhy1/3 antibody. A section of the blot stained with Ponceau red is presented in the lower panel as a loading control. (C) In the upper panel, 4 week-old individuals of the indicated genotypes are shown. In the lower panel, fluorescence of chlorophyll was visualized by confocal microscopy. (D) Transmission electron microscopy of sections from green (left panels) and white (right panels) sectors of variegated leaves of KO1/3 plants. The bar represents 10 μm in the upper panels and 2 μm in the lower panels. (E) Variegation phenotype varies between independent lines. Four week-old individuals from the two variegated lines Var A and Var B are shown.

Figure 9. Variegated plants contain rearranged amplified plastid DNA regions. (A) DNA gel blot (10 μg /lane) of total leaf DNA digested with HindIII and hybridized with the probes indicated below the gels. The probe numbers refer to the nucleotides of the published *Arabidopsis* chloroplast genome (146). Expected fragments from restriction analysis of Col-0 ptDNA and the size of new fragments observed in variegated lines are presented below the probes. A lower exposition of the second panel allowing better visualization of the amplified bands in VarA is presented in Figure 15. (B) Restriction map of the reorganized regions in variegated lines. The red and green arrows represent the amplified regions in Var A and Var B respectively. Probes used in DNA gel blot experiments are represented as blue lines. The HindIII restriction sites are represented as vertical black lines and the horizontal black line corresponds to ptDNA. The expected restriction fragments in base pairs (bp) are shown underneath the ptDNA. A gene map of this region of *Arabidopsis* ptDNA is presented in Figure 16. (C) Schematic of the possible arrangements of the reorganized ptDNA in variegated lines. A head-tail dimer and a monomeric circular molecule are represented for Var A (red) and Var B (green). Oligonucleotides used for the PCR-amplification of the junctions of reorganized ptDNA are represented by small black arrows. (D) PCR amplification of fragments containing the junctions of reorganized ptDNA in Var A and B plants (bottom panel). DNA from leaves 4 (variegated) and 8 (non-variegated) was isolated (upper panel). The plastidial *ycf2* gene was used as a loading control. (E) DNA gel blot analysis showing the arrangement of

amplified ptDNA in the Var A line. DNA from the indicated genotypes was digested with the indicated restriction enzymes and separated on an agarose gel. The DNA was hybridized with the probe depicted in F. A 9.2 kb band corresponding to the WT DNA fragment appears in all samples digested with XhoI and PstI restriction enzymes. A band of 9.4 kb expected from digestion of the WT plastid genome with KpnI was found in all genotypes. The asterisks indicate putative circular molecules. (F) Restriction map of the reorganized regions in the Var A line. The red arrow represents the amplified region in Var A. The probe used is represented as a blue line. A portion of ptDNA is represented as a black horizontal line. The restriction sites are indicated by vertical black lines. K = KpnI, P = PstI, X = XhoI. A circular monomer is represented on the right with the expected linear digestion product of this molecule.

Figure 10. Illegitimate recombination is increased in the absence of *Arabidopsis* ptWhirlies. For each genotype, PCR reactions using outward or inward-facing PCR primers were performed on 4 pools of DNA from 4 different plants. Reactions were run on agarose gels containing ethidium bromide. (A-C) Representative PCR reactions are shown. The oligonucleotides used are indicated above each panel. Individual bands (white numbers) were cut, cloned and sequenced. Each band represents a unique recombination product (Tableau III). The asterisks indicate non-recombinant products arising from non-specific hybridization of the 69633F primer at positions 58720-58733 of the plastid genome. (D) The plastidial *ycf2* gene was used as a loading control.

Figure 11. Model for the appearance of variegation in plants lacking Whirlies. Very low levels of rearranged products are found in Col-0 plants. These products are more abundant in single Whirly KO plants. Rearranged product levels are highest in KO1/3 plants. In these plants, some rearranged molecules can reach 10-25 times the levels of the normal plastome and thus compromise the development and function of chloroplasts.

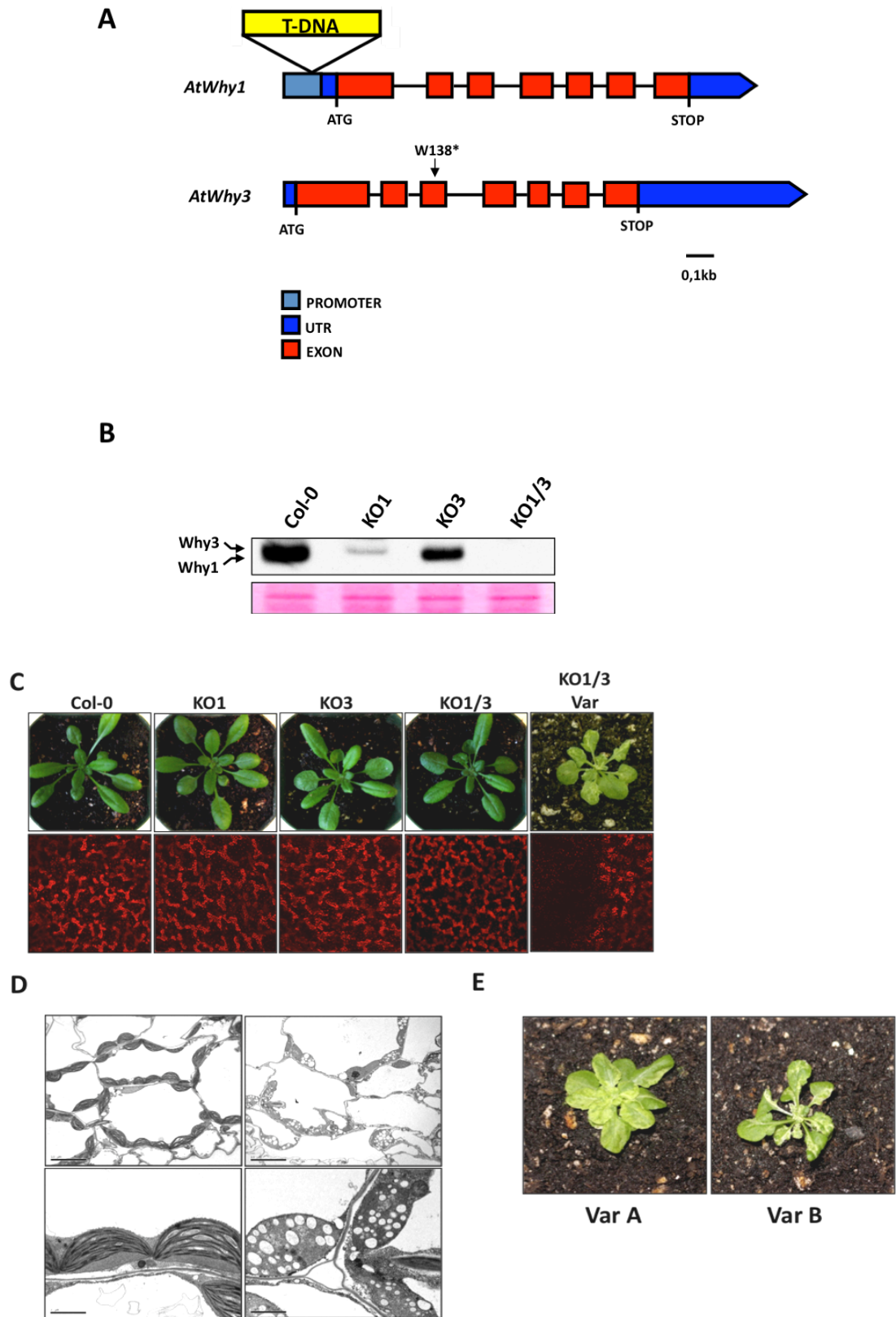
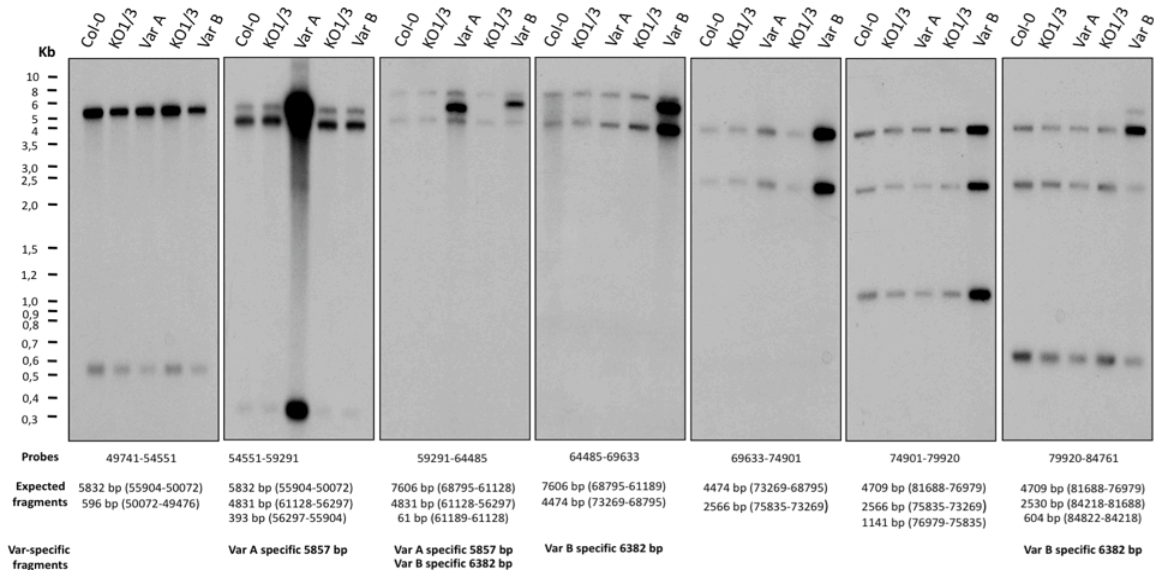


Figure 8. AtWhy1 and AtWhy3 are involved in the biogenesis of chloroplasts.

A



B

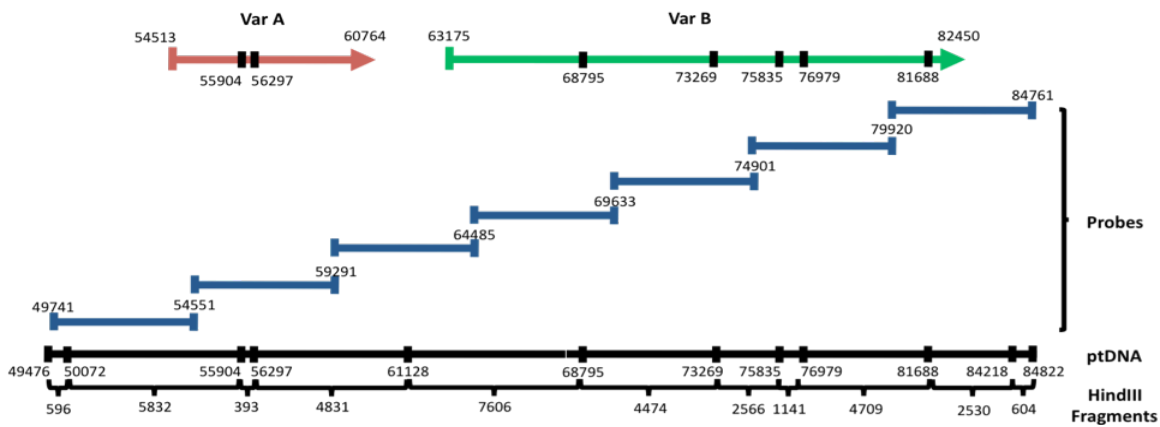


Figure 9. Variegated plants contain rearranged amplified plastid DNA regions.

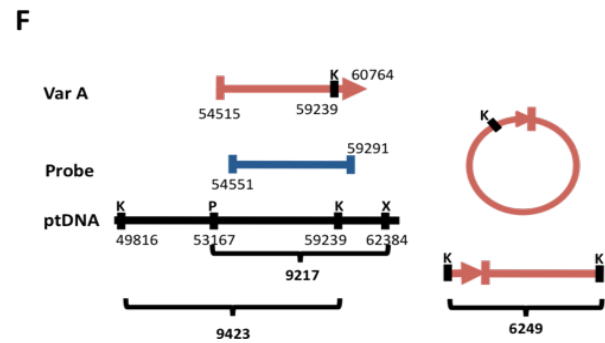
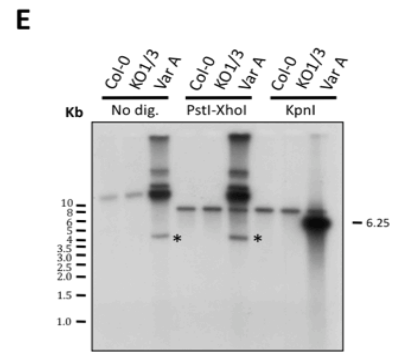
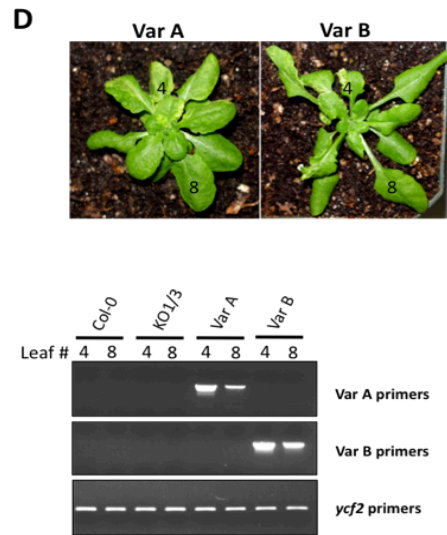
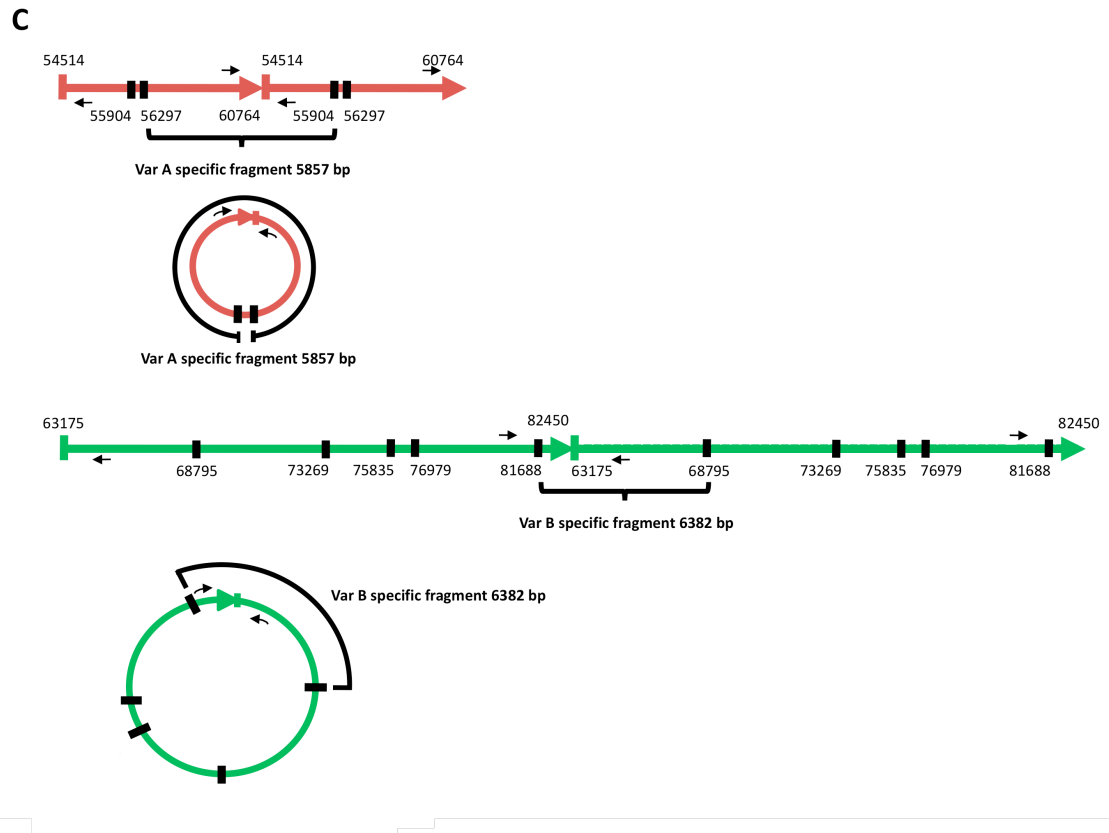


Figure 9 (suite). Variegated plants contain rearranged amplified plastid DNA regions.

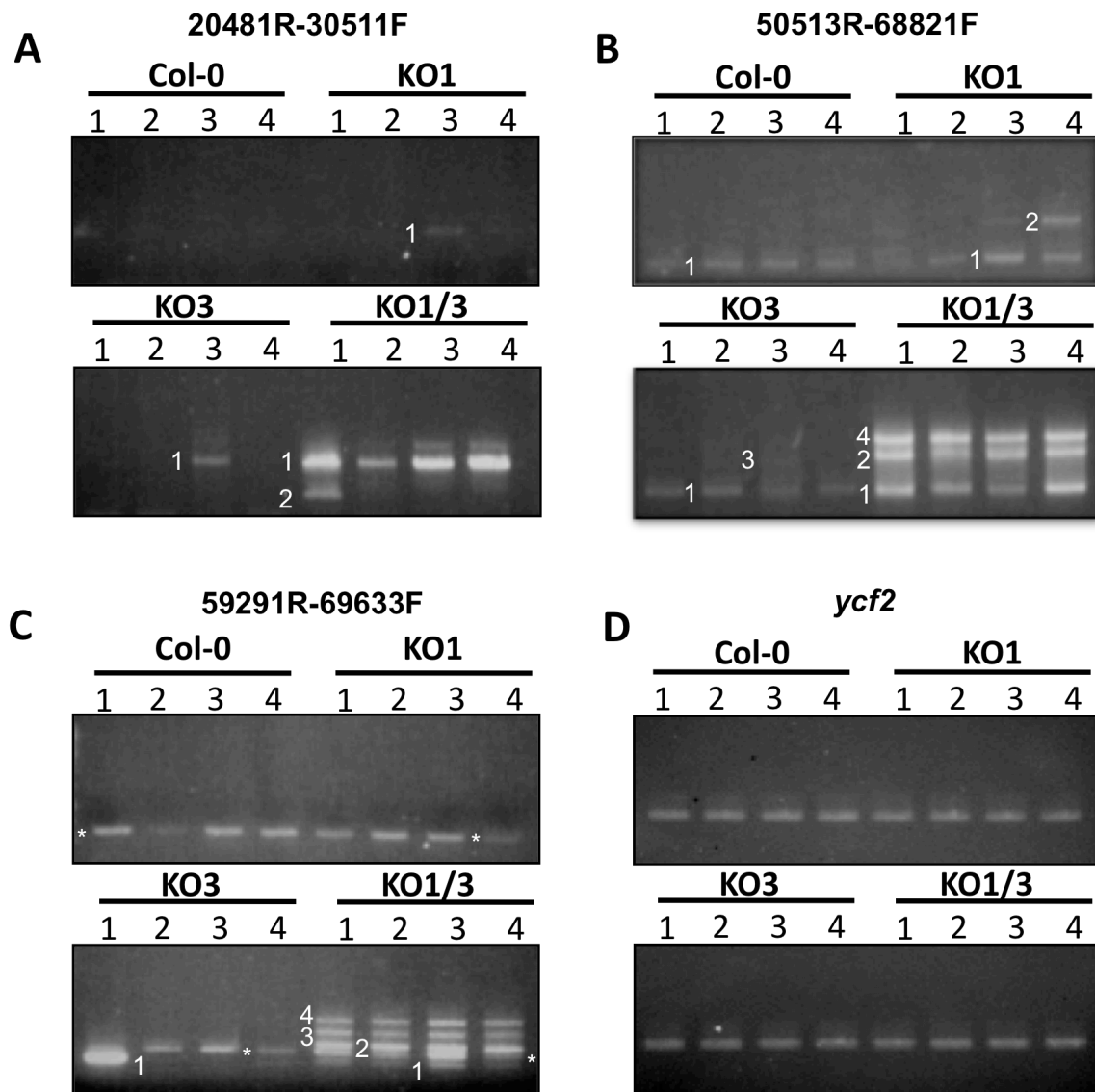


Figure 10. Illegitimate recombination is increased in the absence of *Arabidopsis* ptWhirlies.

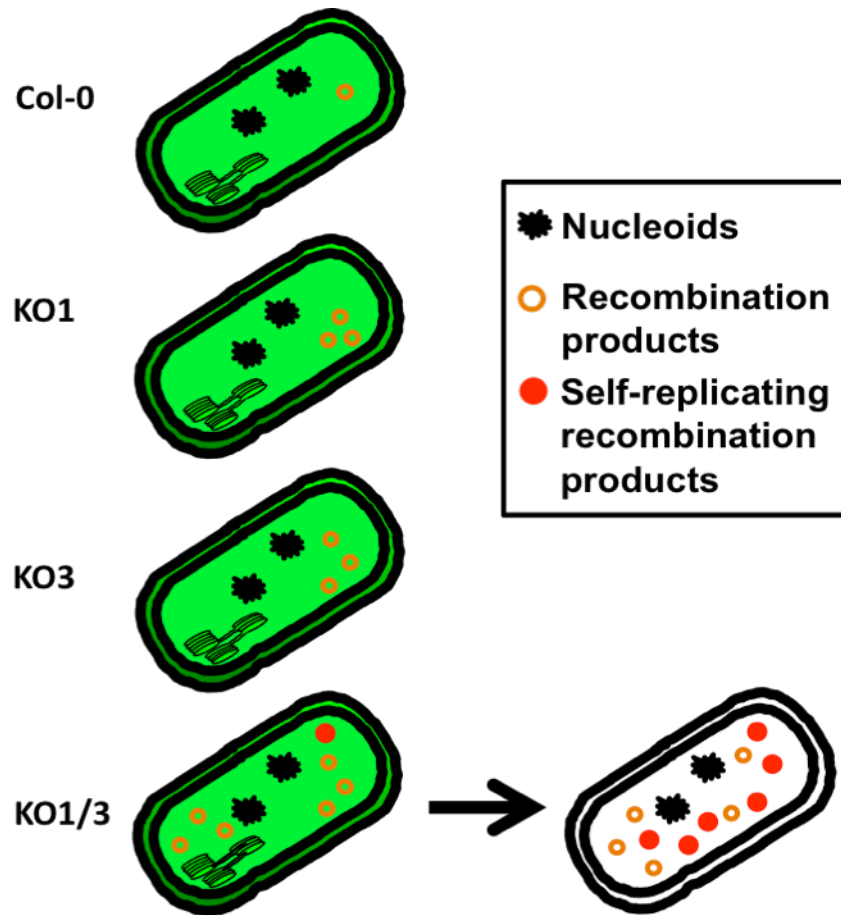


Figure 11. Model for the appearance of variegation in plants lacking Whirlies.

Tableau I. Percentage of plants showing a variegated phenotype

Genotype	Plants Observed	Variegated Plants	% Variegation
Col-0	2287	0	0
KO1	1367	0	0
KO3	2736	0	0
KO1/3	2879	133	4.6

Tableau II. Characteristics of amplified regions in variegated plants

Plant line	Length of amplified region	Direct repeats (DR) positions ¹	Short direct repeat sequences and junction	DR length and (mismatches)
Var A	6250 bp	54514 recombinant 60764	atctcattagCCTTTTTTTTTTcgatatttca agcattcattCCTTTTTTTTTTcgatatttca agcattcattCCTTTTTTTTTTatcttctatc	(0) 10 bp (0)
Var B	19275 bp	63165 recombinant 82450	ttttgtagttTTCCTTTTTTTTTTTTTtcaattttg tagttctttaTTCCTTTTTTTTTTTTTtcaattttg tagttctttaTTCCTTTTTTTTTTTTTcagttgctac	(0) 14 bp (0)
Var C	25940 bp	9992 recombinant 35932	gttctagttgCTCTTGAATACCTTCTTTcaaaaagctt attctcttatCTCTTGAATACCTTCTTTcaaaaagctt attctcttatCTCTTGAATTCATTCGTTgcagatccaa	(0) 18 bp (3)
Var D	19216 bp	61921 recombinant 80507	tgaagcagttGTTAAAATTCCTTATgatatgcaac cccatgataaGTTAAAATTCCTTATgatatgcaac cccatgataaGTTAAAATTCCTTATtgttctataa	(0) 15 bp (0)
Var E	46310 bp	124537 recombinant 6019	tcaataatgtAAAAGTATACTGTCTCTTGgtttagactaa tatattacaaAAAAGTGTAG-GACTCTTGgtttagactaa tatattacaaAAAAGTGTAG-GACTCTTGtaaaatagaa	(3) 18 bp (0)

¹ The positions of the direct repeats refer to the numbers of the nucleotides in the published Arabidopsis Chloroplast Genome Sequence (38)

Tableau III. Representative recombination events between short direct repeats in *Arabidopsis* plants lacking ptWhirlies.

Reaction	Band #	Plant lines	DR positions	Short direct repeat sequences	DR length and (mismatches)
A	1	KO1, KO3 KO1/3	20366 recombinant 31005	aatatatataCAAAAATTTTCTTTTTTTacacttctta gttaagttatCAAAAATTTGCTTTTTTTacacttctta gttaagttatCAAAAATTTGCTTTTTTTtttttagaa	(1) 18 bp (0)
A	2*	KO1/3	30300 recombinant 30827	gtcatttatgCAAAAAAAAAAAtggttatgta attggtctaaCATAAAAAAAAAAtggttatgta attggtctaaCATAAAAAAAAAagaaaaaag	(1) 14 bp (0)
B	1	Col-0 KO1, KO3 KO1/3	50427 recombinant 69424	aattccaatcGGAGTGGATTGGgcaagggata tctaaaaaatGGAGTTGATTGGgcaagggata tctaaaaaatGGAGTTGATTGGattgcacca	(1) 12 bp (0)
B	2	KO1 KO1/3	49181 recombinant 68999	tttttatttcCCCCACACCTTTTTTatataaaatt ctctaaccttCCCCACACCTTTTTTatataaaatt ctctaaccttCCCCACACTATTTTTtgctaggta	(0) 15 bp (3)
B	3	KO3	49972 recombinant 70402	gttttttttaCTTTTTTTTTTTTTAttattgtatc cccaaagtgTCTTTTTTTTTTTTTAttattgtatc cccaaagtgTCTTTTTTTTTTTTTAcggtgtgaaa	(0) 15 bp (0)
B	4	KO1/3	48553 recombinant 68998	tagaattgtaTCCCCCCTTCATTTATTGCTttccgatctt actctaacctTCCCCCCTTCATTTATTGCTttccgatctt actctaacctTCCCCACCCTATTTTTTGCTaggatatttc	(0) 21 bp (5)
C	1	KO3 KO1/3	59043 recombinant 69807	tttttgtaccTATTTTTTTATTctatttctat ttctcttcaaTATTTTTTTATTctatttctat ttctcttcaaTATTTTTTTATTtttatattga	(0) 12 bp (0)
C	2	KO1/3	59287 Recombinant 70228	acttaccctcTATTTTTTGTGCCTTtagtaggct ccgtacaggcTTTTTTTTGTGCCTTtagtaggct ccgtacaggcTTTTTTTTGTGCATTgcatacggct	(1) 14 bp (1)
C	3	KO1/3	59028 recombinant 70225	tatttagtttGGCTTTTTTTTGTacctattttt taaccgtacaGGCTTTTTTTTGTacctattttt taaccgtacaGGCTTTTTTTTGTgcattgcata	(0) 12 bp (0)

2*: Non-specific annealing of primer 20481REV at positions 30367-30346 yielded this product.

Tableau IV. Short direct repeat-mediated recombination frequency in Arabidopsis

Plant line	PCR reactions	Recombination events
Col-0	30	2
KO1	30	7
KO3	30	6
KO1/3	30	40

2.10 Supporting Information

Plastid Whirlies are Associated with Nucleoids

It has been previously shown that Whirlies of *Arabidopsis* localize to either plastids or mitochondria (1). An AtWhy1-green fluorescent protein (GFP) fusion produces punctuate foci of fluorescence in the plastids of transformed potato protoplasts, a pattern that is reminiscent of nucleoids (condensed DNA-protein complexes containing the complete plastid genome). Yet, co-localization of AtWhy1-GFP with plastid DNA has so far not been demonstrated. Therefore, we produced transgenic tobacco plants expressing a GFP fusion protein of StWhy1 (*Solanum tuberosum* Whirly 1), a close homolog of AtWhy1 (68% amino acid identity). Localization of StWhy1-GFP was monitored by laser-scanning confocal microscopy (LSCM). Examination of StWhy1-GFP fluorescence in transgenic leaf guard cells indicated that StWhy1 was also present in speckles within chloroplasts (Figure 13A). To determine if ptWhirlies are associated with nucleoids, we transiently transfected protoplasts produced from tobacco mesophyll leaf cells with StWhy1-GFP. As shown in the upper right panel of Figure 13B, StWhy1-GFP appeared as punctuate foci of fluorescence distributed in the chloroplasts. The transformed protoplasts were subsequently stained with the DNA-dye Syto85 which allowed the visualisation of nucleoids (lower left panel of Figure 13B). Co-localization of Syto85 fluorescence with that of StWhy1-GFP in chloroplasts of tobacco mesophyll protoplasts (lower right panel of Figure 13B) confirmed that StWhy1 co-localizes with DNA in plastids.

Plastidial Whirlies are Associated with ptDNA in a Non-Sequence Specific Manner

To characterize the DNA regions that are bound by AtWhy1 and AtWhy3, we performed DNA immunoprecipitation experiments followed by PCR detection of precise regions distributed all over the plastid genome : inside genes, around the determined transcription start sites of genes, in isolated regions that are devoid of any coding sequence and inside putative replication origins (Figure 14A). No differences could be observed for the immunoprecipitation of all types of regions between Col-0, KO1 and KO3 extracts, whereas no ptDNA could be immunoprecipitated using the KO1/3 extracts

(Figure 14B). To control for the specificity of interaction, we tested that the sonication regimen to which ptDNA was subjected yielded fragments of an average length of 500-1000 base pairs (Figure 14C). This indicated that for any given region immunoprecipitated specifically, it can be assumed that there is at least one Whirly binding-site within a maximum of 1200 bp relative to the center of the amplified region. Additionally, we used primers to amplify a mitochondrial DNA region in the *atp9* gene as a negative control since *AtWhy1* and *AtWhy3* exclusively localize to chloroplasts and thus should not interact with the mitochondrial genome. We found that we could only amplify similar background levels of mitochondrial DNA for all genotypes, confirming the lack of specific interaction between *AtWhy1/3* and mitochondrial DNA *in vivo* (Figure 14D). These results indicate that ptWhirlies are able to bind many if not all ptDNA regions *in vivo*.

2.11 Supporting Information Materials and Methods

Arabidopsis mutant characterization

The SALK Institute Genomic Analysis Laboratory provided the sequence-indexed T-DNA insertion line SALK_099937(2). The insertion inside the *AtWhy1* gene was mapped at nucleotide -102 relative to the initial ATG. The Seattle TILLING Project (3) provided plants with mutations in the *AtWhy3* gene. One line isolated and identified as il68D8 changed the TGG codon 138 to a TGA stop codon. The heterozygous M2 plants were backcrossed with Col-0 eight consecutive times to eliminate background mutations. The homozygous progeny obtained was used for experiments. These *AtWhy3* mutant plants (KO3) were crossed with *AtWhy1* mutant plants (KO1) and the progeny was genotyped to isolate double mutants. The insertion mutant was genotyped using PCR primers specific for the wild-type or mutant allele. The KO3 point mutation was detected using PCR amplification with dCAPS primer W138for (CAGCTGGTGTTCGTC AATATGCCTG) coupled with primer W138rev (GCCCAAGGCTAACTAGATTACCGAT) and digestion of the product with the MvaI restriction enzyme (Fermentas).

Chlorophyll autofluorescence in *Arabidopsis*

Mature leaf disks were taken from 4 week old plants and mounted on glass slides. A confocal microscope OLYMPUS IX71 was used with a 488 nm laser. Chlorophyll autofluorescence was visualised through a 660 nm long pass filter.

Transmission electron microscopy

Leaf mesophyll tissue from 4 week old plants was cut into 1-2 mm² pieces which were washed twice with 0.2 M sodium cacodylate buffer pH 7.4 for 10 minutes. Tissue was fixed by incubation for 6 hours in cacodylate buffer containing 1 % EM-grade glutaraldehyde. Samples were coated with epoxy resin, cut into 80 nm sections and mounted onto nickel/formvar grids. Staining of the samples was done using uranyl acetate and lead citrate. Observations were done using a JEOL (JEM1230) transmission electron microscope at 80 kV. Photographs were taken using a Gatan DualVision camera.

Intracellular localization of StWhy1

The StWhy1-GFP construct was produced by inserting the Emerald GFP (Clontech) sequence downstream of *StWhy1* into pBluescript (Stratagene). The StWhy1-GFP fusion comprising all 274 amino acids of StWhy1 was then inserted into the pBin19 vector which contains 2 copies of the CaMV35S promoter. All constructs were electroporated into *Agrobacterium tumefaciens* strain LBA4404 and transgenic plants were obtained as described (4). For transient expression, the StWhy1-GFP fusion in pBin19 was transferred to the pBI223 vector which also contains the CaMV35S promoter. Leaf mesophyll protoplasts were isolated and transformed as described. GFP fluorescence (500-530 nm) was visualised with a Leica DM IRB/E laser-scanning confocal microscope using a 488 nm laser excitation source. Chlorophyll was excited using a 543 nm laser and its autofluorescence was visualised at 575-630 nm. For DNA staining, protoplasts were incubated with 5 µM Syto85 (Molecular Probes) for 10-30 min and then washed with fresh culture medium. Syto85 fluorescence (570-600 nm) was visualised using a 568 nm laser excitation source. GFP and Syto85 images were collected sequentially, and no fluorescence cross-talk was observed under our conditions (data not shown). Pseudocoloring of the images, maximal projections, and image overlays were done using

the Leica confocal software. Hand-made thin sections of fresh leaves from 4 weeks-old plants were examined in sterile water. Confocal microscopy was carried out using an Olympus FV300 microscope.

Plastid DNA immunoprecipitation

Leaf tissue was fixed in 1 % formaldehyde for 15 minutes under vacuum. Glycine was added to 0.125 M to titrate the remaining formaldehyde and fixed tissue was washed 3 times with distilled water. Crude plastids were prepared by grinding in chloroplast extraction buffer (0.33M sucrose, 10 mM MES-NaOH pH 4.6, 10 mM NaCl, 1 mM EDTA, 10 mM DTT) using a mortar and pestle. The ground tissue was filtered through 2 layers of Miracloth (Calbiochem) and spun at 1000 g for 5 minutes in a microcentrifuge to pellet most plastids. The pellet was resuspended in IPP buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 % NP-40, 0.5 % sodium deoxycholate and 1 tablet of Complete Protease Inhibitor (Roche)), lysis was carried out on ice for 2 minutes and the lysate was spun at 11000 g for 15 minutes. Supernatant was sonicated using a Branson microtip sonicator. Sonicated material was precleared by incubating with protein-A-agarose beads (Roche) for 1 hour at 4°C on a rotating platform. The precleared material was spun at 11000 g for 15 minutes and the supernatant was transferred to a fresh tube. Proteins were quantified and an 100 µg aliquot was kept as input material while 500 µg was used for immunoprecipitation using 1/200 rabbit polyclonal anti-AtWhy1 sera in 1 mL final volume. Immunoprecipitations were carried out for 1 hour at 4°C before adding 50 µl of protein-A-agarose beads. Complexes were allowed to form overnight. Beads were pelleted at 100 g for 1 minute and were washed 2 times 5 minutes with IPP buffer, 2 times 5 minutes with low salt buffer (50 mM Tris-HCl pH 7.5, 0.1 % NP-40, 0.05% sodium deoxycholate), 2 times 5 minutes with high salt buffer (50 mM Tris-HCl pH 7.5, 500 mM NaCl, 0.1 % NP-40, 0.05% sodium deoxycholate), 2 times 5 minutes with LiCl buffer (50 mM Tris-HCl pH 7.5, 0.1 % NP-40, 0.05 % sodium deoxycholate, 250 mM LiCl) and 2 times 5 minutes with TE pH 7.5 (10 mM Tris-HCl pH 7.5, 1 mM EDTA). Beads were pelleted at 11000 g for 20 seconds and resuspended in TE buffer containing 1 % SDS. Complexes were eluted from the beads by heating at 65°C for 15 minutes. Beads were spun down at 11000 g for 5 minutes and the supernatant transferred to a new tube. One tenth volume of 2 M NaCl was added and the

eluted complexes were incubated at 65°C overnight for de-crosslinking. Glycogen was added to a final concentration of 0.5 µg/µl and two volumes of 100 % ethanol was added. DNA was precipitated at –80°C for 1 hour. Tubes were spun at 11000 g for 20 minutes and DNA pellets were resuspended in TE buffer. DNA was subjected to a phenol/chloroform extraction and to a chloroform extraction before being reprecipitated as above. After a final wash with 70% ethanol, DNA was resuspended in water. This DNA was used for PCR reactions. To further control the specificity of the experiment, we determined the size of DNA fragments following sonication and found that they were between 500 and 1200 bp in length. For the determination of the specificity of the interaction between AtWhy1 and AtWhy3 and plastid DNA, mitochondrial DNA immunoprecipitation was performed as described previously (5).

Electrophoretic mobility shift assays

Five hundred mg of leaves taken from 4 week old plants were ground in 5 mL of ice-cold chloroplast isolation buffer (0.33 M sucrose, 10 mM MES-NaOH pH 4.6, 10 mM NaCl, 1 mM EDTA, 10 mM DTT). The resulting slurry was filtered through 2 layers of Miracloth (Calbiochem). The suspension was then centrifuged at 3000 rpm in a microcentrifuge for 5 minutes. The crude plastids were resuspended in IPP buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 % NP-40, 0.5 % sodium deoxycholate, 1X Complete protease inhibitor cocktail (Roche)) and incubated 2 minutes on ice. The lysed plastids were centrifuged at maximal speed in a microcentrifuge for 15 min and the supernatant was taken. Twenty µg of the crude plastid extracts were used for ssDNA binding assays. The same extracts were used in the protein gel blot presented Figure 1B thereby providing a loading control. For electrophoretic mobility shift assays, crude plastid proteins were incubated in IPP buffer with 30000 counts per minute of end-labelled modified 4XPB (6) oligonucleotide (5'-TGTCATTTTGTGCATTTTGTGCATTTTGTCA-3') for 30 minutes at room temperature. The anti-AtWhy1/3 antibody or preimmune serum were added to the Col-0 extracts before adding the oligonucleotide. The reactions were then separated on a 5.4% polyacrylamide Tris-Borate-EDTA gel as described (6).

Isolation of RNA

RNA was isolated from 4 week old plants using TRIZOL[®] reagent (Invitrogen) according to manufacturer's instructions. Running of the samples was performed as described (7). For the Var A and Var B samples, only highly variegated tissue was used.

Detection of rearranged DNA by PCR in maize

PCR reactions were conducted using the Ex Taq polymerase from Takara Bio Inc. according to the manufacturer's instructions. The search for rearranged products was performed using a series of outward-facing oligonucleotides spaced by approximately 5-50 kb. A total of 19 PCR reactions were performed on each of two independent DNA samples from B73 inbred line (WT) and each of the *ZmWhy1* mutant lines and analysed by gel electrophoresis. The DNA samples for all the plants were pre-adjusted to the same amount of ptDNA using semi-quantitative amplification of the maize *rpoB* gene sequence before performing the outward-facing PCR. All visible DNA bands were isolated, cloned and sequenced.

Oligonucleotides used in this study

The following oligonucleotides were used to identify rearranged products in *Arabidopsis*:

5291REV, TCC CAA CGA GCC GTT TAT CGA ; 9121REV, AAG ATA GAT CGG GCC AAC TCG ; 9640FOR, GGT TCA AAT CCT ATT GGA CGC; 19181FOR, CAG CAT CTG ATG TGC GTA CAG ; 19309FOR, TCG TCT GCT AAG ACA CGA CCA ; 20481, FOR GCG TTG ATC TAA TTG CCA CCG ; 20481REV, CGG TGG CAA TTA GAT CAA CGC ; 30511FOR, GCT GCC TCC TTG AAA GAG AGA ; 30511REV, TCT CTC TTT CAA GGA GGC AGC ; 31138REV, AAT AGC CTA GCG CAT CGA ACG ; 40130FOR, TGA CAT GTT CCT CCT CTT CCC ; 40130REV, GGG AAG AGG AGG AAC ATG TCA ; 48481FOR, AGG GCA AAA ATA CCC CAA GGT ; 50251FOR, GGC CCC TTC CTA ATT GGA GAT ; 50513FOR, GGT TGT TGT GGT AGA CGT CTC ; 50513REV, GAG ACG TCT ACC ACA ACA ACC ; 54551REV, TGG GAC GCA TAA CCG GAT ATG ; 58730FOR, CCT CCC TTG CTT GCA TAT GTA; 59291FOR, AGG CCT AGT CTT TCC GGC AAT ; 59291REV, ATT GCC GGA AAG ACT AGG CCT ; 60182REV, ATA TAA AGG ACA CGC CGG GCA ; 65482REV, GCA TCT GCA CTA TCT CAG GAG ; 68821FOR, CTC CGA GCT ATA TAT CCG CGT; 68821REV, ACG CGG ATA TAT AGC TCG GAG; 69633FOR, ACA CCC GAG TAC ATG TTC CTC; 69633REV, GAG GAA CAT GTA CTC GGG TGT; 69941REV, AGC AAC AGA AGC CCA AGC TCA; 71221REV, TTC TTG TTC CTG AAC GGG CTT ; 79920FOR, CCA TGC CGT TCG GAA GTG ATT ; 79920REV, AAT CAC TTC CGA ACG GCA TGG ; 84761FOR, TAT GTA TGG CCG TGC CTA AGG ; 84761REV, CCT TAG GCA CGG CCA TAC ATA ; 95260FOR, AAC CTG CTA GTG GAG GAA GAC; 95260REV, GTC TTC CTC CAC TAG CAG GTT ; 104831FOR, CTG AAT CCA TGG GCA GGC AAG ; 105983REV, GCG TTA CTC AAG CCG ACA TTC ; 111330FOR, TGC GGT TAA TCC CGC TGT TGA ;

111330REV, TCA ACA GCG GGA TTA ACC GCA ; 115987FOR, AAA CCA CTC ATG CCC GGC AAT ;
 116007FOR, AAA CCA CTC ATG CCC GGC AAT ; 116007REV, ATT GCC GGG CAT GAG TGG TTT ;
 121141FOR, AAC GTC ATC CGG TTA CCG GTT.

The following oligonucleotides were used to identify rearranged products in maize:

Zm33FOR, GGG TAT TTC CGG CTT TCC TTC ; Zm33REV, GAA GGA AAG CCG GAA ATA CCC ;
 Zm4761FOR, TCT CCT CAT ACG GCT CGA GAA ; Zm4761REV, TTC TCG AGC CGT ATG AGG AGA ;
 Zm10113REV, TTC CAC GTG GTA GAA CCT CCT ; Zm15001REV, TAC CCC TTG GAA ATG GGG TAC ;
 Zm20058FOR, CCA ACT TAG GGG GGT ACG AAT ; Zm20058REV, ATT CGT ACC CCC CTA AGT TGG ;
 Zm25060FOR, GCA TGT CCA GTG ACT CAT GTG ; Zm25060REV, CAC ATG AGT CAC TGG ACA TGC ;
 Zm29971REV, AAC AAG AGA AGC ACG AGC TCC, Zm35230FOR, GAA AGG GAG TGT GTG CGA GTT ;
 Zm40020FOR, CCC AAG GTA TGG AAT CCT AGG ; Zm40020REV, CCT AGG ATT CCA TAC CTT GGG ;
 Zm45152FOR, TTT CGT TCT AGC GCC CGG AAA; Zm45152REV, TTT CCG GGC GCT AGA ACG AAA ;
 Zm50141FOR, AGA GGT GAG ATT AGG TGC CCT ; Zm50141REV, AGG GCA CCT AAT CTC ACC TCT ;
 Zm54971FOR, TGC CAA CAA TCC GAG GCT GTA ; Zm54971REV, TAC AGC CTC GGA TTG TTG GCA;
 Zm59880FOR, CGG GCC TAT TCA TTA GCT CCT ; Zm59880REV, AGG AGC TAA TGA ATA GGC CCG ;
 Zm64883FOR, GGG GTC GTG GAA TGC TTT TCT ; Zm64883REV, AGA AAA GCA TTC CAC GAC CCC ;
 Zm69981FOR, ATT AGC CAA CCG CCA AGG GAA ; Zm69981REV, TTC CCT TGG CGG TTG GCT AAT ;
 Zm74851FOR, GCG TAA GTC CCC TAT CTA GGA ; Zm85072FOR, ACC TCT CCG GAT CCT CGA ATT ;
 Zm89967FOR, GAA AAG CAA CGA CTG GAG TGG ; Zm89967REV, CCA CTC CAG TCG TTG CTT TTC ;
 Zm100051REV, TCG ACG ATG AAG CTT ATC CCC ; Zm105133FOR, ATT CGC CCT CCC CCT ACA TAT ;
 Zm105133REV, ATA TGT AGG GGG AGG GCG AAT ; Zm110101FOR, CTG GAA TAT AGA GCG GAC
 TGG ; Zm129981FOR, GAG CAG GCT ACC ATG AGA CAA ; Zm129981REV, TTG TCT CAT GGT AGC
 CTG CTC ; Zm135011FOR, TCA GTG ACT TTG GCA CTG GAC.

The following oligonucleotides were used to generate probes for DNA gel blots in
Arabidopsis:

421FOR, AGC AGC TAG GTC TAG AGG GAA ; 5291REV, TCC CAA CGA GCC GTT TAT CGA ; 10230FOR,
 GCA ATC GTT GAC CTC TTG CCA ; 15461REV, CAA TAC GGG CTC GGT GTC ATT ; 30511FOR, GCT GCC
 TCC TTG AAA GAG AGA ; 35850REV, ACC AAC CAT CAG GAG ACG CAA ; 40130FOR, TGA CAT GTT
 CCT CCT CTT CCC ; 44916REV, GTA CGG AAA GAG AGG GAT TCG ; 49741FOR, CCT TAC GTA AAG
 GCC ACC CTA ; 54551REV, TGG GAC GCA TAA CCG GAT ATG ; 54551FOR, CAT ATC CGG TTA TGC
 GTC CCA ; 59291REV, ATT GCC GGA AAG ACT AGG CCT ; 59291FOR, AGG CCT AGT CTT TCC GGC
 AAT ; 64485REV, CGT TGC TGT GTC AGA AGA AGG ; 69633FOR, ACA CCC GAG TAC ATG TTC CTC ;
 74901FOR, CAG AGA ATG GGG TCT GTC ATC ; 79920REV, AAT CAC TTC CGA ACG GCA TGG ;
 79920FOR, CCA TGC CGT TCG GAA GTG ATT ; 84761REV, CCT TAG GCA CGG CCA TAC ATA ;
 84761FOR, TAT GTA TGG CCG TGC CTA AG ; 90150REV, TCC ATC TGC CCT TTG TCA ACG ;
 100770FOR, GAA GTC ATC AGT TCG AGC CTG ; 105983REV, GCG TTA CTC AAG CCG ACA TTC ;
 105983FOR, GAA TGT CGG CTT GAG TAA CGC ; 111330REV, TCA ACA GCG GGA TTA ACC GCA ;
 111330FOR, TGC GGT TAA TCC CGC TGT TGA ; 116007REV, ATT GCC GGG CAT GAG TGG TTT.

The following oligonucleotides were used to determine the ends of amplified regions in
Arabidopsis variegated lines:

56061REV, AGA ACA CCA GGT AGT GAG ACC ; 60211FOR, GAG GTC AAG GGG CTA TTC CTT ;
 63990REV, CAT GGA CTA GCT GTA CCT ACC ; 81360FOR, CGC AGG TCT TAC TGT AAC TGG ;
 10230REV, TGG CAA GAG GTC AAC GAT TGC ; 30511FOR, GCT GCC TCC TTG AAA GAG AGA ;
 62161REV, TAG CAG GGT CTG GAG CAA GAA ; 75961FOR, TTA CTG GTG TGG TTC TGG GTG ;
 120601FOR, GCG CTT CGG AAT TCA TCT CAC ; 4341FOR, CCG AGT ACT CTA CCG TTG AGT.

The following oligonucleotides were used for adjusting DNA levels in the PCR experiments:

YCF2FOR, GAT CTC TGA GAG CTG TTT CCG ; YCF2REV, TGT TTC GCC TCT TAC TCG GAG ;
 ZmRpoBFOR, GCT TGG CGG AAG AAC TTG AGA ; ZmRpoBREV, TGT TTC GCC TCT TAC TCG GAG.

The following oligonucleotides were used to generate probes for DNA gel blots in maize:

Zm4761FOR, TCT CCT CAT ACG GCT CGA GAA ; Zm10113REV, TTC CAC GTG GTA GAA CCT CCT ;
 Zm10113FOR, AGG AGG TTC TAC CAC GTG GAA ; Zm15001REV, TAC CCC TTG GAA ATG GGG TAC ;
 Zm15001FOR, GTA CCC CAT TTC CAA GGG GTA ; Zm20058REV, ATT CGT ACC CCC CTA AGT TGG ;
 Zm25060FOR, GCA TGT CCA GTG ACT CAT GTG ; Zm29971REV, AAC AAG AGA AGC ACG AGC TCC ;
 Zm29971FOR, GGA GCT CGT GCT TCT CTT GTT ; Zm35230REV, AAC TCG CAC ACA CTC CCT TTC ;
 Zm35230FOR, GAA AGG GAG TGT GTG CGA GTT ; Zm40020REV, CCT AGG ATT CCA TAC CTT GGG ;
 Zm40020FOR, CCC AAG GTA TGG AAT CCT AGG ; Zm45152REV, TTT CCG GGC GCT AGA ACG AAA ;
 Zm45152FOR, TTT CGT TCT AGC GCC CGG AAA ; Zm50141REV, AGG GCA CCT AAT CTC ACC TCT ;
 Zm50141FOR, AGA GGT GAG ATT AGG TGC CCT ; Zm54971REV, TAC AGC CTC GGA TTG TTG GCA ;
 Zm59880FOR, CGG GCC TAT TCA TTA GCT CCT ; Zm64883REV, AGA AAA GCA TTC CAC GAC CCC ;
 Zm69981FOR, ATT AGC CAA CCG CCA AGG GAA ; Zm74851REV, TCC TAG ATA GGG GAC TTA CGC ;
 Zm74851FOR, GCG TAA GTC CCC TAT CTA GGA ; Zm79860REV, TGA AGG GAA AAT CCT GCC GAG ;
 Zm85072FOR, ACC TCT CCG GAT CCT CGA ATT ; Zm89967REV, CCA CTC CAG TCG TTG CTT TTC ;
 Zm95042FOR, GTG GCG AAC TCC AGG CTA ATA ; Zm100051REV, TCG ACG ATG AAG CTT ATC CCC ;
 Zm100051FOR, GGG GAT AAG CTT CAT CGT CGA ; Zm105133REV, ATA TGT AGG GGG AGG GCG
 AAT ; Zm105133FOR, ATT CGC CCT CCC CCT ACA TAT ; Zm110101REV, CCA GTC CGC TCT ATA TTC
 CAG ; ZmMt255542FOR, GCA GTA TTG GAA CCA TCC TCG ; ZmMt259720REV, CCC ATG ATG TGG
 TAA AGG GAG.

2.12 References

1. Krause K, *et al.* (2005) DNA-binding proteins of the Whirly family in *Arabidopsis thaliana* are targeted to the organelles. *FEBS Lett* 579(17):3707-3712 .
2. Alonso JM, *et al.* (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* 301(5633):653-657 .
3. Till BJ, *et al.* (2003) Large-scale discovery of induced point mutations with high-throughput TILLING. *Genome Res* 13(3):524-530 .

4. Matton DP, Prescott G, Bertrand C, Camirand A, & Brisson N (1993) Identification of cis-acting elements involved in the regulation of the pathogenesis-related gene STH-2 in potato. *Plant Mol Biol* 22(2):279-291 .
5. Maréchal A, *et al.* (2008) Overexpression of mtDNA-associated AtWhy2 compromises mitochondrial function. *BMC Plant Biol* 8:42 .
6. Desveaux D, *et al.* (2004) A "Whirly" transcription factor is required for salicylic acid-dependent disease resistance in Arabidopsis. *Dev Cell* 6(2):229-240 .
7. Sambrook J, Fritsch EF, & Maniatis T (1989) *Molecular cloning : a laboratory manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.) 2nd Ed.
8. Sato S, Nakamura Y, Kaneko T, Asamizu E, & Tabata S (1999) Complete structure of the chloroplast genome of Arabidopsis thaliana. *DNA Res* 6(5):283-290 .
9. Lohse M, Drechsel O, & Bock R (2007) OrganellarGenomeDRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. *Curr Genet* 52(5-6):267-274 .
10. Maier RM, Neckermann K, Igloi GL, & Kossel H (1995) Complete sequence of the maize chloroplast genome: gene content, hotspots of divergence and fine tuning of genetic information by transcript editing. *J Mol Biol* 251(5):614-628 .
11. Prikryl J, Watkins KP, Friso G, van Wijk KJ, & Barkan A (2008) A member of the Whirly family is a multifunctional RNA- and DNA-binding protein that is essential for chloroplast biogenesis. *Nucleic Acids Res* 36(16):5152-5165 .
12. Clifton SW, *et al.* (2004) Sequence and comparative analysis of the maize NB mitochondrial genome. *Plant Physiol* 136(3):3486-3503 .

2.13 Supporting Information Figure Legends

Figure 12. ssDNA-binding activity of AtWhy1/3 in Arabidopsis ptWhirlies mutants. Electrophoretic mobility shift assay was performed using 20 µg of crude plastid proteins isolated from plants of the indicated genotypes and a radiolabeled probe of 32

nucleotides (5'-TGTCATTTTTGTCATTTTTGTCATTTTTGTCA-3'). As a control, an anti-AtWhy1/3 antibody (α -Why1/3) or preimmune serum (PI) was preincubated with Col-0 extracts before adding the probe. Addition of the antibody eliminated the signal in the Col-0 extract, confirming that this signal corresponds to a complex between Whirlies and ssDNA.

Figure 13. Co-localization of StWhy1 with chloroplast DNA. (A) LSCM of StWhy1-GFP tobacco leaf guard cells. Upper left panel is chlorophyll autofluorescence pseudo-coloured in red. Upper right panel is GFP fluorescence pseudo-coloured in green. Bottom left panel is overlay of chlorophyll and GFP fluorescence. Bottom right panel is corresponding phase-contrast image. (B) LSCM of a tobacco leaf mesophyll protoplast transiently expressing StWhy1-GFP and stained with the DNA dye Syto85. Upper left panel is chlorophyll autofluorescence pseudo-coloured in red. Upper right panel is GFP fluorescence pseudo-coloured in green. Bottom left panel is Syto85 fluorescence pseudo-coloured in blue. Bottom right panel is overlay of all three images. Maximum projections are shown. Scale bars = 8 μ m.

Figure 14. AtWhy1 and AtWhy3 interact with chloroplast DNA. (A) Position of the amplified regions on the plastid genome. (B) PCR amplification of ptDNA regions following immunoprecipitation on crude plastid extracts of the indicated genotypes. Inputs represent 20 % of the total DNA used in the immunoprecipitation. Representative regions are shown here indicating specific interaction with AtWhy1/3. (C) Total plastid DNA following sonication was purified and used as a template for PCR. Oligonucleotides designed to amplify regions of increasing length were used to verify the efficiency of the sonication regimen. (D) Immunoprecipitation on total crude organelles was performed and plastid or mitochondrial DNA was assessed by PCR using oligonucleotides designed to amplify part of the plastid *psbA* or mitochondrial *atp9* genes.

Figure 15. Long (16 hrs) and short (4 hrs) exposure of panel 2 from Figure 9A. DNA gel blot (10 μ g/lane) of total leaf DNA digested with HindIII and hybridized with the probes indicated below the gel. The probe numbers refer to the nucleotides of the published *Arabidopsis* chloroplast genome (8). Expected fragments from restriction analysis of Col-0

ptDNA and the size of new fragments observed in variegated lines are presented below the probes.

Figure 16. A gene map of the ptDNA region containing the amplified rearranged regions found in Var A and Var B plants was drawn using the web-based tool OGDRAW (9).

Figure 17. Absence of subgenomic amplicons in *ZmWhy1* maize mutants. (A-E) DNA gel blots (10 µg/lane) of total leaf DNA digested with HindIII and hybridized with the probes indicated below the gel. The probe numbers refer to the nucleotides of the published maize chloroplast genome (10). Expected fragments from restriction analysis of WT ptDNA are presented below the probes. B73 inbred line was used as the source for WT DNA. For each of the three *ZmWhy1* mutants, a segregating green plant issued from heterozygous seeds was used as a control (g). The pale green *ZmWhy1-2* (pg), the yellow (intermediate) *ZmWhy1-2/-1* (y) and the albino *ZmWhy1-1* (al) lines were previously described (11). The corresponding ethidium bromide staining of total digested DNA is presented below each panel. (E) The ptDNA level in the albino mutant was slightly lower than in the other lines where no significant difference was observed compared to the WT. To verify this, a membrane was hybridized with plastid probe 95042-100051, partially stripped and re-blotted with mitochondrial probe 255542-259720 (12). The residual signal from the plastid genome is marked by an asterisk. This confirmed that the level of ptDNA was specifically diminished in the albino *ZmWhy1-1* mutant.

Figure 18. Levels of plastid 23S and 16S rRNA in *Arabidopsis* ptWhirlies mutants. Total RNA was extracted from 100 mg of mature leaf tissue from the indicated plants. Ten µg of purified RNA was migrated on a 1.2 % agarose gel and stained with ethidium bromide.

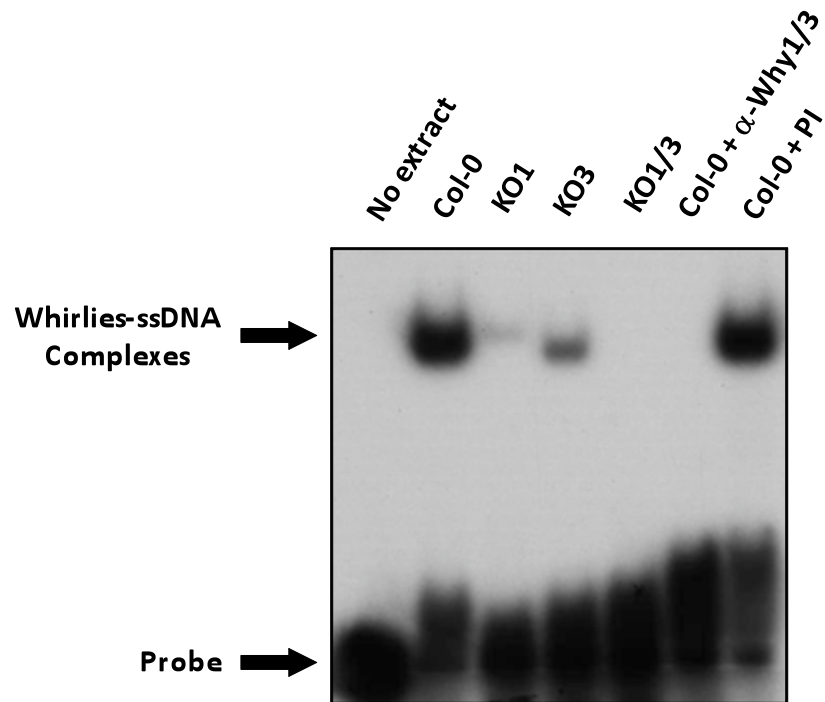


Figure 12. ssDNA-binding activity of AtWhy1/3 in *Arabidopsis* ptWhirlies mutants.

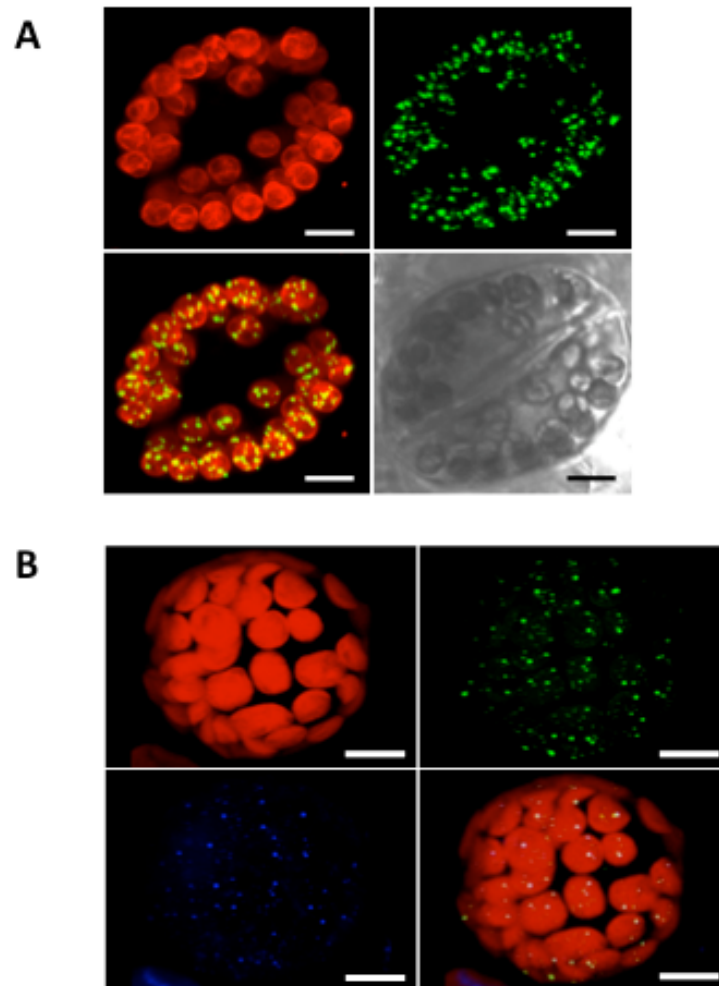


Figure 13. Co-localization of StWhy1 with chloroplast DNA.

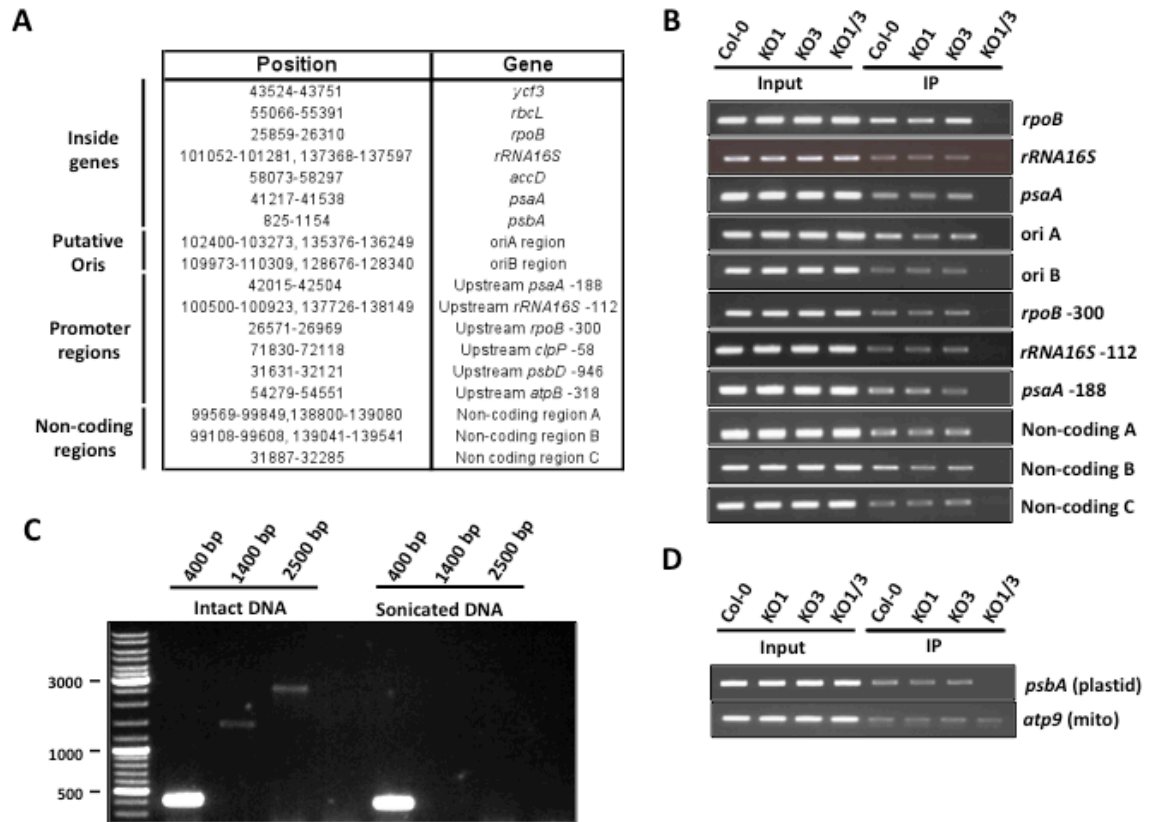


Figure 14. AtWhy1 and AtWhy3 interact with chloroplast DNA.

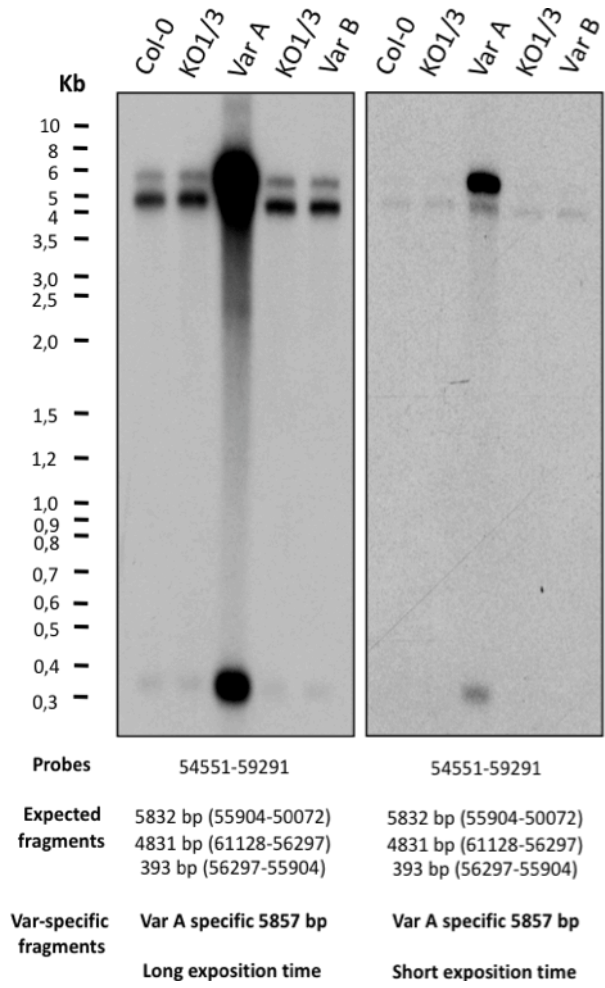


Figure 15. Long (16 hrs) and short (4 hrs) exposure of panel 2 from Figure 9A.

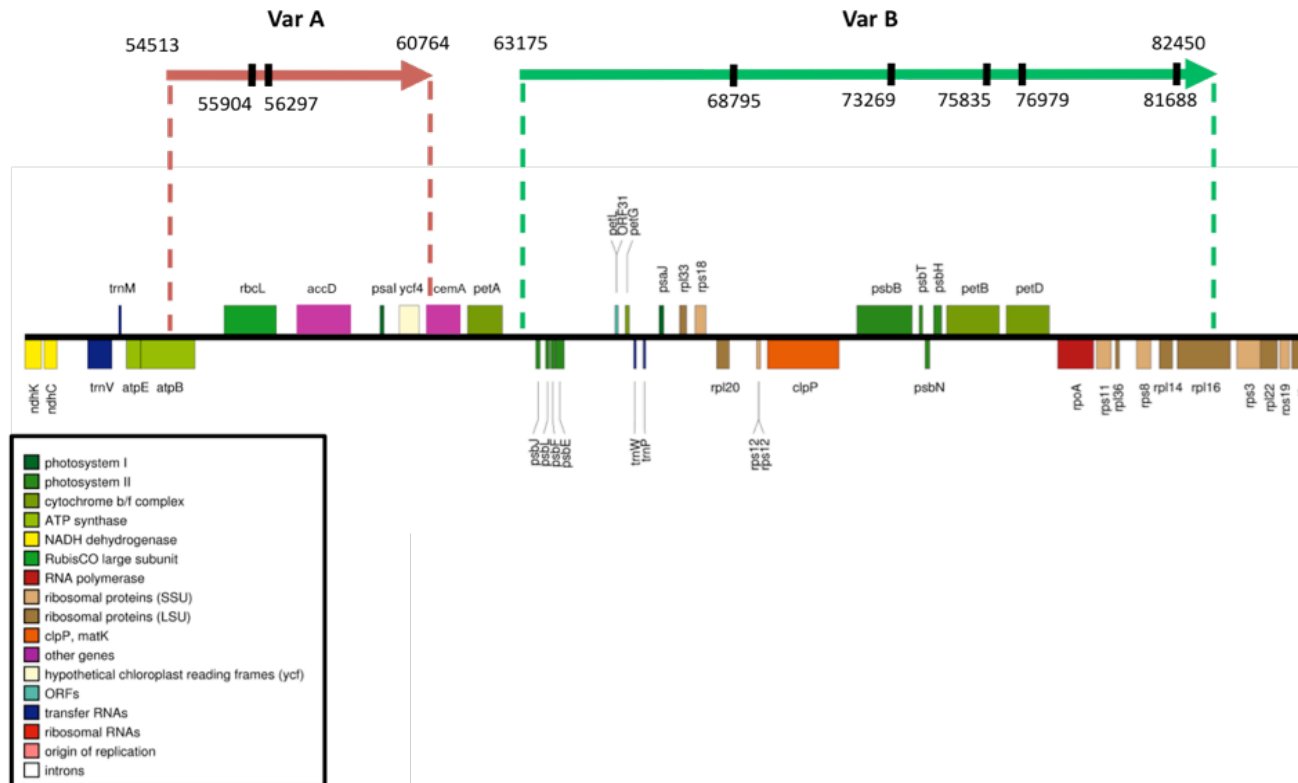


Figure 16. A gene map of the ptDNA region containing the amplified rearranged regions found in Var A and Var B plants was drawn using the web-based tool OGDRAW (9).

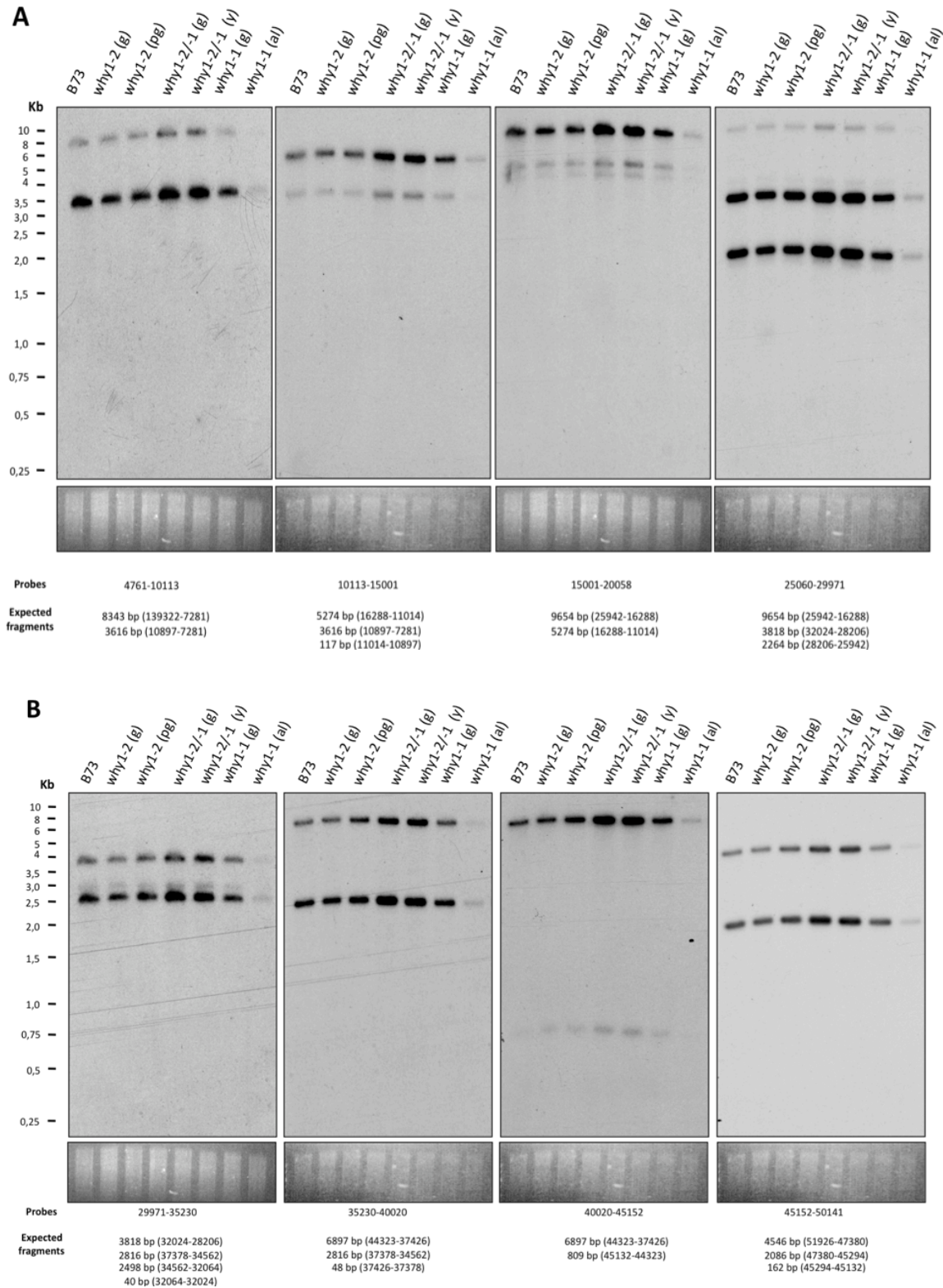


Figure 17. Absence of subgenomic amplicons in *ZmWhy1* maize mutants.

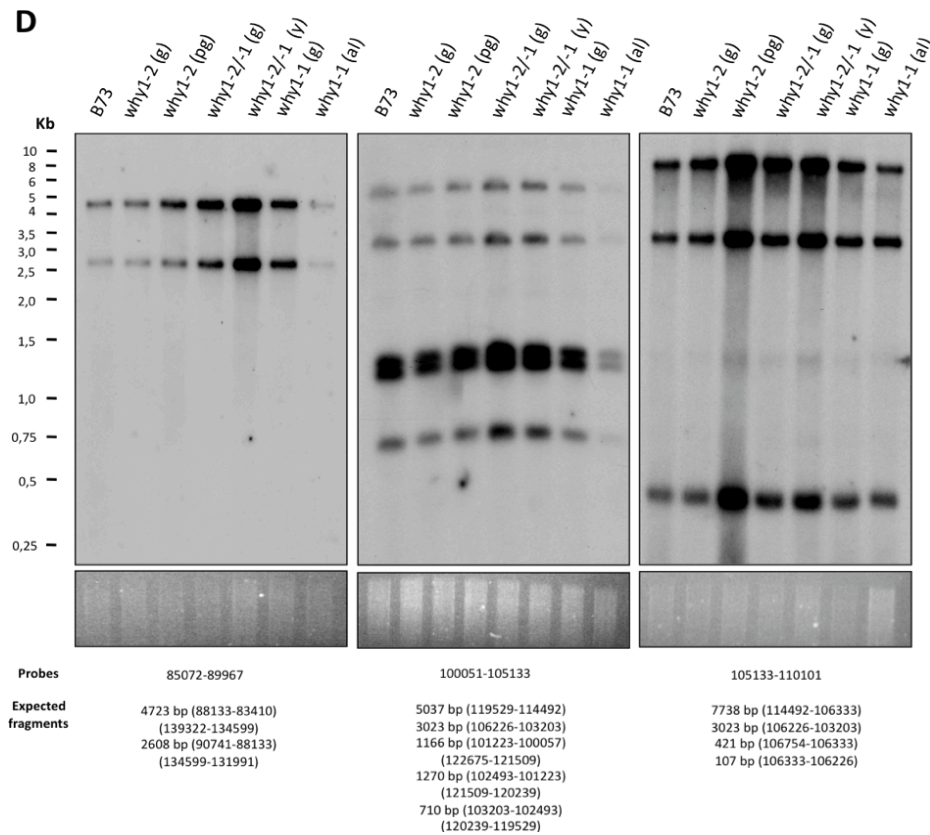
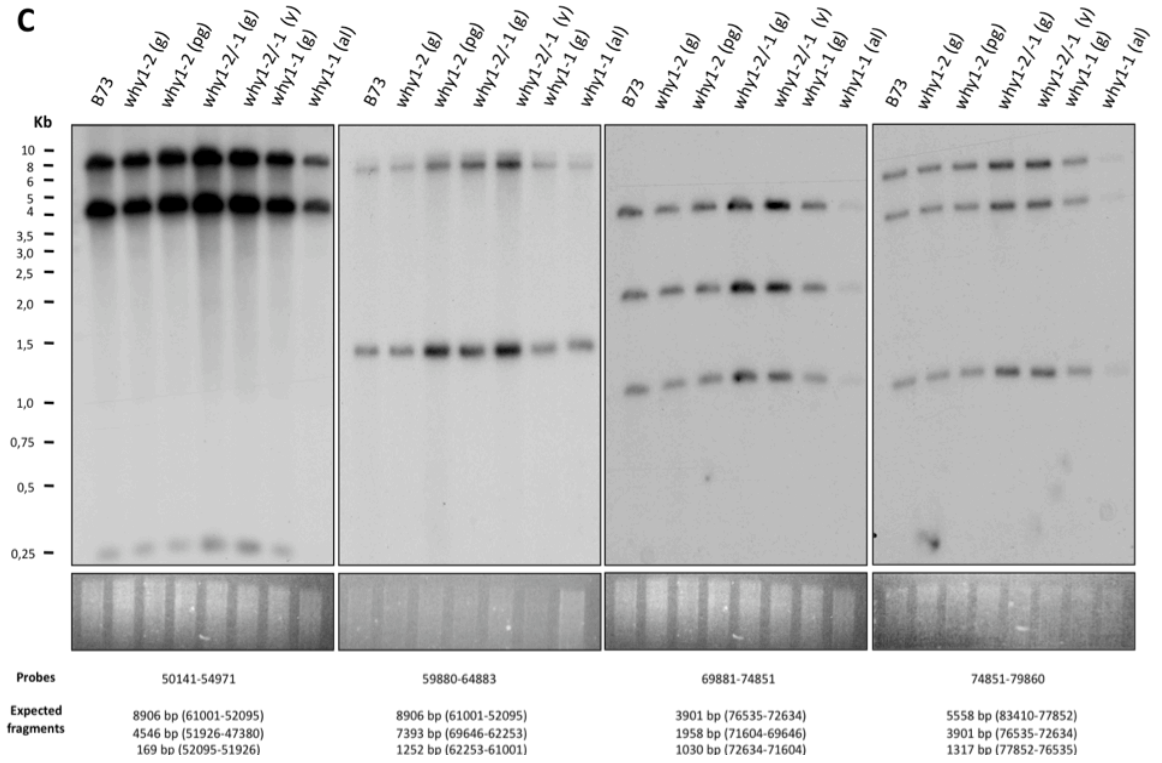


Figure 17 (suite). Absence of subgenomic amplicons in *ZmWhy1* maize mutants.

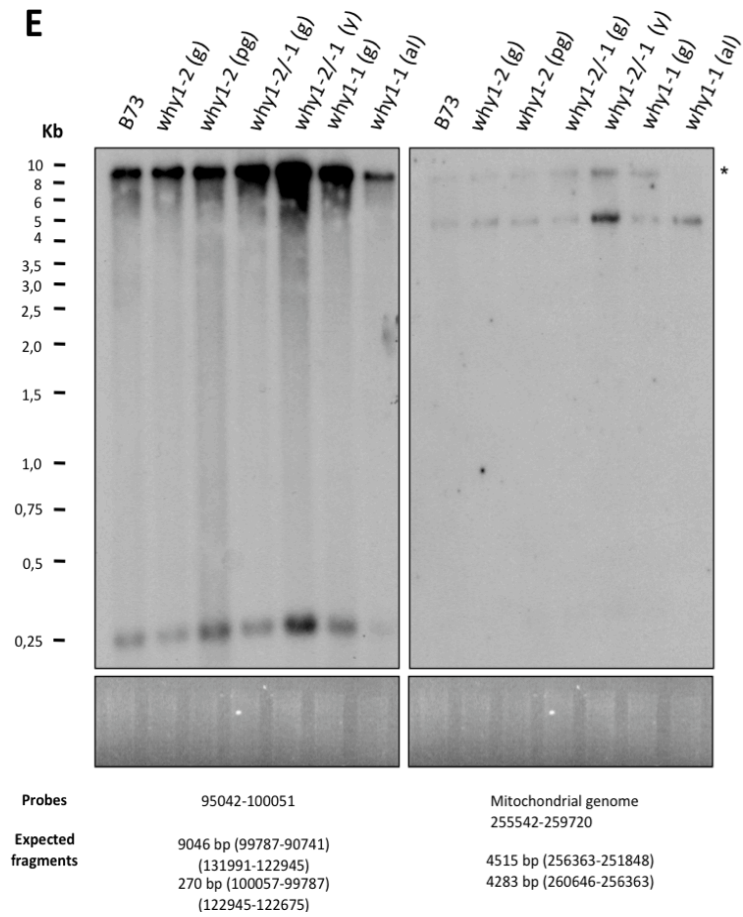


Figure 17 (suite). Absence of subgenomic amplicons in *ZmWhy1* maize mutants.

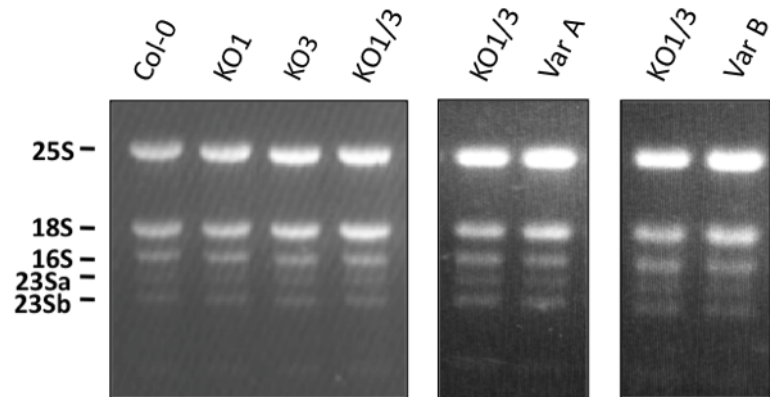


Figure 18. Levels of plastid 23S and 16S rRNA in *Arabidopsis* ptWhirlies mutants.

Tableau V. Recombination events mediated by short direct repeats in Arabidopsis

Event	Related ¹ Events	Reaction	Plant lines	DR positions	Short direct repeat sequences and junction	DR length and (mismatches)
1		5291REV DUP/CIR ² 20481FOR	KO1/3	5204 recombinant 21835	aaaagtttgaTTAGTTATGGGatcataaaaa agtaggttttTTAGTTATGGGatcataaaaa agtaggttttTTAGTTATGGGcctagcaaaaa	(0) 11 bp (0)
2		19138FOR DEL ³ 31172REV	KO1/3	19766 recombinant 30894	ccatatctcaATAGATTGgcgtaatttt ccatatctcaATAGATTGcctcatctca actttccttgATAGATTGcctcatctca	(0) 8 bp (0)
3		20481REV DUP/CIR 30511FOR	KO1,KO 3 KO1/3	20366 recombinant 31005	aatatatataCAAAAATTTTCTTTTTTTacacttctta gttaagttatCAAAAATTTGCTTTTTTTacacttctta gttaagttatCAAAAATTTGCTTTTTTTtttttagaa	(1) 18 bp (0)
4 ⁴		20481REV DUP/CIR 30511FOR	KO1/3	30300 recombinant 30827	gtcatttatgCAAAAAAAAAAAAAAtggttatgta attggtctaaCATAAAAAAAAAAAAAAtggttatgta attggtctaaCATAAAAAAAAAAAAAAgaaaaaaag	(1) 14 bp (0)
5		30511REV DUP/CIR 40130FOR	KO1/3	29711 recombinant 40160	gaccctttctTTTGCTTTATgaaaaaagaa aaaaccaagaTTTGCTTTATgaaaaaagaa aaaaccaagaTTTGCTTTATctggatttaa	(0) 10 bp (0)
6	23	40130REV DUP/CIR 50251FOR	KO1/3	38989 recombinant 50299	cccctggatgCAGGAATAGCgacatggact aatgccaaaaCAGGAATAGCgacatggact aatgccaaaaCAGGAATAGCacttgatatt	(0) 10 bp (0)
7		40130REV DUP/CIR 50251FOR	KO1/3	39109 recombinant 50758	cttggtttccATTTTGTTGtaggtgtaac gatagattgaATTTTGTTGtaggtgtaac gatagattgaATTTTGTTGaatttaatta	(0) 10 bp (0)

8		40130REV DUP/CIR 50251FOR	KO1/3	39266 recombinant 51611	cacctcctcgAGTAaacgcttcca cattctacctAGCAaacgcttcca cattctacctAGCAttgagtagac	(1) 4 bp (0)
9		40130REV DUP/CIR 50251FOR	KO1	39411 recombinant 50322	gtccacagaaAAAAttattgctaa tgatattagtAAAAttattgctaa tgatattagtAAAaatgccaga	(0) 3bp (0)
10		40130REV DUP/CIR 50251FOR	KO3	39586 recombinant 50325	ttaatgccataAATGCCtttcaaatcc tattagtaaaAATGCCtttcaaatcc tattagtaaaAATGCCcagaaaatat	(0) 6 bp (0)
11	25,35	50251FOR DEL 60182REV	KO1/3	50780 recombinant 59406	tttaattagaTTTCTTTTTTTTATTTTTTAAAcaggccgtg tttaattagaTTTCTTTTTTTTATTTTTTAAAaacttcgatt tcatatcaccTTTTTTTTTTTATTTTTTAAAAaacttcgatt	(0) 20 bp (2)
12		50251FOR DEL 60182REV	KO1/3	50847 recombinant 59726	gggatacaaaaAAAAAAAAAAGAAAGTattatgaact gggatacaaaaAAAAAAAAAGGAAAGTcatattttct ctttgaaaacAAAAAAAAAGGAAAGTcatattttct	(1) 15 bp (0)
13	14,18,21,22,24	50251FOR DEL 60182REV	KO1/3	50474 recombinant 59953	aggttaatcgAATCATTTTTTTTTCgtttggttg aggttaatcgAATCATTTTTTTTTCcacaagggat cgtcccagcaAATCATTTTTTTTTCcacaagggat	(0) 14 bp (0)
14	13,18,21,22,24	50513REV DUP/CIR 68421FOR	KO1/3	50488 recombinant 70515	aaattcagggtAATCGAATCATTTTTTTTTCgtttggttg tttcaatataTAATCGAATCATTTTTTTTTCgtttggttg tttcaatataTAATCGAA---TTTTTTTTTaatctaaaa	(0) 19 bp (3)
15		50461REV DUP/CIR 68821FOR	Col-0 KO1,KO3 KO1/3	50427 recombinant 69424	aattccaatcGGAGTGGATTGGgcaagggata tctaaaaaatGGAGTTGATTGGgcaagggata tctaaaaaatGGAGTTGATTGGattgcacca	(1) 12 bp (0)
16	17	50461REV DUP/CIR 68821FOR	KO1/3	48553 recombinant 68998	tagaattgtaTCCCCCCTTCATTTATTGCTttccgatctt actctaacctTCCCCCCTTCATTTATTGCTttccgatctt actctaacctTCCCC <u>CCACT</u> ATTTTTTTGCTaggtattttc	(0) 21 bp (5)

17	16	50461REV DUP/CIR 68821FOR	KO1 KO1/3	49181 recombinant 68999	tttttatttcCCCCACACCTTTTTTatataaaatt ctctaacccttCCCCACACCTTTTTTatataaaatt ctctaacccttCCCCAC <u>CACT</u> ATTTTTttgctaggta	(0) 15 bp (3)
18	13,14,21,22	50461REV DUP/CIR 68821FOR	KO1	50482 recombinant 70117	caaaaattcaGGTTAATCGAATCATTtttttttcggtt cgaactagcggGTTGAT-GGATCATTtttttttcggtt cgaactagcggGTTGAT-GGATCATTaccctgatga	(3) 15 bp (0)
19		50461REV DUP/CIR 68821FOR	KO1/3	50243 recombinant 70143	tacctgattcATAACTAGAAAgtttctccgg ccctgatgatATAA-TAGAAAgtttctccg ccctgatgatATAA-TAGAAAagggttttct	(1) 10 bp (0)
20		50461REV DUP/CIR 68821FOR	KO3	49972 recombinant 70402	gttttttttaCTTTTTTTTTTTTTTAttattgtatc cccaaagtgtCTTTTTTTTTTTTTTAttattgtatc cccaaagtgtCTTTTTTTTTTTTTTAcggtgtgaaa	(0) 15 bp (0)
21	13,14,18,22,24	50461REV DUP/CIR 59633FOR	Col-0 KO1,KO3 KO1/3	50489 recombinant 59938	aggttaatcgaATCATTTTTTTTTCgtttggtgt cgcccagcaAATCATTTTTTTTTCgtttggtgt cgcccagcaAATCATTTTTTTTTCcacaagggat	(0) 14 bp (0)
22	13,14,18,21,24	50461REV DUP/CIR 59633FOR	KO1/3	50489 recombinant 59936	tcaggttaatCGAATCATTTTTTTTTCgtttggttg tccgtcccagCGAATCATTTTTTTTTCgtttggttg tccgtcccagCAATCATTTTTTTTTCcacaagggat	(0) 16 bp (1)
23	6	50461REV DUP/CIR 59633FOR	KO1/3	50309 recombinant 60067	atgccaaaacAGGAATAGcacttgatat accgaaaagaAAGAATAGcacttgatat accgaaaagaAGGAATAGtgcgtatttt	(0) 8 bp (0)
24	13,14,21,22	50461REV DUP/CIR 59633FOR	KO1/3	50495 recombinant 60081	gaatcattttTTTTCGTTGGTtggttggt tagtgcgtatTTTTCGTTGGGttggttggt tagtgcgtatTTTTCGTTGGGgatttcctgg	(1) 11 bp (0)
25	11,32,35	54551REV DUP/CIR 58730FOR	KO1/3	54523 recombinant 59387	ctcattagccTTTTTTTTCGTATTTTcatttttagca catatcacctTTTTTTTTCGTATTTTcatttttagca catatcacctTTTTTTTTT-TTATTTTaaaaacttcg	(0) 16 bp (2)

26		59311REV DUP/CIR 69633FOR	KO1/3	58721 recombinant 70186	attagactagACAAACAAAAAAAAgttcattttc agacgggataACAAACAAAAAAAAgttcattttc agacgggataACAA <u>AAAA</u> AAAAAAAAatagataaat	(0) 14 bp (1)
27	28	59311REV DUP/CIR 69633FOR	KO1/3	59287 recombinant 70228	acttaccctcT <u>ATTTTT</u> TGTGCCTTtagtaggcct ccgtacaggcTTTTTTTTGTGCCTTtagtaggcct ccgtacaggcTTTTTTTTGTGC <u>ATT</u> gcatacggct	(1) 14 bp (1)
28	27	59311REV DUP/CIR 69633FOR	KO1/3	59028 recombinant 70225	tatttagtttGGCTTTTTTTGTacctatTTTT taaccgtacaGGCTTTTTTTGTacctatTTTT taaccgtacaGGCTTTTTTTGTgcattgcata	(0) 12 bp (0)
29	30, 36	59311REV DUP/CIR 69633FOR	KO1/3	59039 recombinant 69901	tttttttgtaCCTATTTTTTtattctattt tgaaatcgctCCTATTTTTTtattctattt tgaaatcgctCCTATTTTTTattgaaccgc	(0) 10 bp (0)
30	29,35,36	59311REV DUP/CIR 69633FOR	KO3 KO1/3	59043 recombinant 69807	ttttgtaccTATTTTTTTATTctatttctat ttctcttcaaTATTTTTTTATTctatttctat ttctcttcaaTATTTTTTTATTtttatattga	(0) 12 bp (0)
31		59311FOR DEL 69633REV	KO1/3	59307 recombinant 68984	tctttccggcAATTGCAATGGCTTctttatttct tctttccggcAATTGCAATGGCTTttgctactct gtataaaaaAATT <u>CCA</u> ATGGCTTttgctactct	(0) 14 bp (1)
32	25,36	59311FOR DEL 69633REV	KO1/3	59380 recombinant 68522	accgaatcatATC <u>AC</u> CTT <u>TTTTTTTT</u> tattttaa accgaatcatATCCCTTATTTTTTTatgtcattcg atttctccaATCCCTTATTTTTTTatgtcattcg	(2) 16 bp (0)
33		59739FOR DEL 69946REV	KO1/3	58973 recombinant 69303	acatctagttGTAGTTGATTTTGAAAGATGAataagtcatt acatctagttGTAGTTGATTTTGAAAGATGattaggtcatt tacagtaacgGTAT <u>TT</u> CATTTTGAA <u>AT</u> TGattaggtcatt	(0) 21 bp (4)
34		59739FOR DEL 69946REV	KO1/3	59590 recombinant 69740	gtatataactATAGGGGATCATATAAAggaaggagat gtatataactATAGGGGATCATATAAA <u>aat</u> gggctg aatggttggcATAC <u>GGA</u> ATCATATAAA <u>aat</u> gggctg	(0) 17 bp (2)

35	11,25,30,32	59739FOR DEL 69946REV	KO1/3	59390 recombinant 69822	atcacctttttTTTTTTTATTTTaaaaacttcg atcacctttttTTTTTTTATTTTtatattgaag ctcttcaataTTTTTTTATTTTtatattgaag	(0) 12 bp (0)
36	30,32	59739FOR DEL 69946REV	KO1/3	59030 recombinant 68524	tttttgtaccTATTTTTTTATtctatttcta tttttgtaccTATTTTTTTATgtcattcgaa ccaatcccctTATTTTTTTATgtcattcgaa	(0) 11 bp (0)
37		59739FOR DEL 69946REV	KO1/3	59074 recombinant 69714	ctattctataTATTTCTATTAGGTTGTATAtaatattag ctattctataTATTTCTATTAGGTTGTTAatggttggca tggctgtctgTATTTCTAAATAAGTTGTTAatggttggca	(1) 20 bp (2)
38		69633REV DUP/CIR 79920FOR	KO1/3	69468 recombinant 80032	atgtaaaccaTAAGTTTCAGAcacaatagag tcacaaattgTAAGTTACAGAcacaatagag tcacaaattgTAAGTTACAGAtatctaattt	(1) 11 bp (0)
39		79920REV DUP/CIR 111330FOR	KO1/3	78747 recombinant 111896	aagctcttcgCATcgcaatgcct ggtaaccataCATcgcaatgcct ggtaaccataCATgaagaggaaa	(0) 3 bp (0)
40		79920REV DUP/CIR 111330FOR	KO1/3	76183 recombinant 111428	tgggcagcaaAGGGTGTTCCTCTTcttgtacccc ccacaaagagAGAGTGTTCCTCTTcttgtacccc ccacaaagagAGAGTGTTCCTATTaaaaaggccg	(1) 14 bp (1)
41		79920REV DUP/CIR 111330FOR	KO1/3	78830 recombinant 111688	ataataaagaCGCTTACTGTCTCTTTttgattcaac ggctaaacccCGCTTACTGTCTCTTTttgattcaac ggctaaacccCGCTTAAATGTCTTTTgagcaagagc	(0) 16 bp (2)
42		84761REV DUP/CIR 95260FOR	KO1/3	84370 recombinant 95699	acggcgacgaAGAATCAAAGTCtactatatt ctggagtgggAGAATCAAAGTCtactatatt ctggagtgggAGAGTCAGAGTCgaaaagagga	(0) 13 bp (2)
43		105983REV DUP/CIR 116007FOR	KO1/3	104982 recombinant 116127	gtgggagagcAAAAAAGcgtcgtgctg tcctgccaaGAAAAAAGcgtcgtgctc tcctgccaaGAAAAAAGtcgagcgcca	(0) 8 bp (0)

44		105983REV DUP/CIR 116007FOR	KO1 KO1/3	105322 recombinant 116009	<p>tggcttggttAAGGGAACCCACCGgagccgtagc</p> <p>ccggcaatgcAAGGGAACCCACCGgagccgtagc</p> <p>ccggcaatgcAAGGGA<u>CCAT</u>Cgataaaatagt</p>	(0) 14 bp (2)
----	--	-----------------------------------	--------------	---------------------------------	--	---------------------

¹ Related events implicate at least part of a direct repeat that is found independently in another recombination event.

² DUP/CIR = Illegitimate recombination event yielding a duplicated/circularized product and detected using outward-facing PCR.

³ DEL = Illegitimate recombination event yielding a deletion-containing product and detected using inward-facing PCR.

⁴ Non-specific annealing of primer 20481REV at positions 30367-30346 yielded this product.

Tableau VI. Recombination events mediated by short direct repeats in the maize plastid genome

Events	Related Events	Reaction	Plant lines	DR positions ²	Short direct repeat sequences and junctions	DR length and (mismatches)
1		15001REV DUP/CIR ¹ 50141FOR	Why1-1	14923 recombinant 50999	aattcgactcCCAACCCTtttttttaga aattagataaCCAACCCTtttttttaga aattagataaCCAACCCTgctgcattgt	(0) 8 bp (0)
2		15001REV DUP/CIR 50141FOR	Why1-1	14492 recombinant 50341	aggagtgagaTGACTTTC ¹ CCAAAatttttcata aacaaatatgTGATTTTCCAAAatttttcata aacaaatatgTGATTTTCCAAAatggaaaata	(1) 12 bp (0)
3		25060REV DUP/CIR 35230FOR	Why1-1/Why1-2	24850 recombinant 36373	tggcgaagtcACAAGGccctccactt aacaagctgtACAAGGccctccactt aacaagctgtACAAGGagctctagga	(0) 6 bp (0)
4		25060REV DUP/CIR 35230FOR	Why1-2	25015 recombinant 35708	ggggtagaatTTg ¹ ttgattctc ataaaaggaaTTg ¹ ttgattctc ataaaaggaaTTctatcaattt	(0) 2 bp (0)
5		25060REV DUP/CIR 35230FOR	Why1-1	24796 recombinant 35785	tttcccctcaCAAATAAAGgcttgggcta cagagcaaacCAAATAAAGgcttgggcta cagagcaaacCAAATAAAGaaacaacttt	(0) 10 bp (0)
6		25060REV DUP/CIR 35230FOR	Why1-1	24258 recombinant 35542	cccgagtaccCGGGAAAAAGCagaatatttg atgctttaaaCGGGAAATAGCagaatatttg atgctttaaaCGGGAAATAGCtagttagaa	(1) 11 bp (0)
7		29971REV DUP/CIR 54971FOR	Why1-1	29627 recombinant 55252	tttctgaagAGGCGGATAagatattagg ctgcctaataAGGCGGATAagatattagg ctgcctaataAGGCGGATActtccgatcc	(0) 9 bp (0)

8		29971REV DUP/CIR 54971FOR	Why1-1	29507 recombinant 55692	attattggagTCGATAcacaactcac gtatctaacTCGATAcacaactcac gtatctaacTCGATAaaggcaggcg	(0) 6 bp (0)
9		29971REV DUP/CIR 54971FOR	Why1-1	29434 recombinant 55437	gtagatcggTTTTTTTCATTcttcaagaac ttcctcaataTTTTTTTCATTcttcaagaac ttcctcaataTTTTTTTCATTaattactccg	(0) 11 bp (0)
10	11	29971REV DUP/CIR 54971FOR	Why1-1 Why1-2	26297 recombinant 56132	ctaaggaatcAAAAAAAAAGGAAaaattggggtt ataat ttt caAAAAAAAAAGGAAaaattggggtt ataat ttt caAAAAAAAAAGGAAtttgtcgaaa	(0) 12 bp (0)
11	10	29971REV DUP/CIR 54971FOR	WT	26691 recombinant 56131	tctaaggaatCAAAAAAAAAaggaaaaattg acataat ttt CAAAAAAAAAaggaaaaattg acataat ttt CAAAAAAAAAaggaatttgt	(0) 9 bp (0)
12		40020REV DUP/CIR 45152FOR	Why1-1/Why1-2	39662 recombinant 45567	ggcgcgcaAGGAAAAC TATAcccgaaatcc ttaatgctagAGGAAAAC TATAcccgaaatcc ttaatgctagAGGAAAACGATAgctccaagaa	(0) 12 bp (1)
13		45152REV DUP/CIR 59880FOR	Why1-2	44350 recombinant 60179	tccgcaattCCGAATCGCCCTgtagaatggc caaaaagctgCCGAATCGCCCTgtagaatggc caaaaagctgCCGAATGGCC Tatttcttgcg	(0) 12 bp (2)
14		45152REV DUP/CIR 59880FOR	Why1-1	45065 recombinant 60562	agtattcactAGAAA--ATTTCCTCCTTTCTTttttttcttc atgaaaaaaaaAGAAAGCATTGCCCTCCTTTCTTttttttcttc atgaaaaaaaaAGAAAGCATTGCC <u>T</u> CTTTCTTatatactt g ta	(3) 22 bp (1)
15		45152REV DUP/CIR 59880FOR	Why1-1	44815 recombinant 60324	gattcctaaaTTTTGTCCcatatcgtgg ttogaatttgTTTTGTCCcatatcgtgg ttogaatttgTTTTGTCCaagt g agata	(0) 8 bp (0)

16		45152REV DUP/CIR 59880FOR	Why1-1	45075 recombinant 61177	tttctccttTCTTTTTTtcttcttct aaat tttggg TCTTTTTTtcttcttct aaat tttggg TCTTTTTTgtttaaccg
17		54971REV DUP/CIR 64883FOR	Why1-1	54631 recombinant 65071	ataagatttAATTCATTTCTTCAATtgcctcct ccgaaatccAATTCATTTCTTCAATtgcctcct ccgaaatccAATTCATTT-TTCAATggggtagat
18		59880REV DUP/CIR 74851FOR	Why1-1	59266 recombinant 76157	tatttcggcAATTGCAATGACTTccttatttct aagaggaaaaAATTGCAATGACTTccttatttct aagaggaaaaAATTGCAATGATTTaagcgagaa
19		59880REV DUP/CIR 74851FOR	Why1-1	58828 recombinant 74980	caactgggtaTTCTATTCCACTTCTACTTTTTAaattaagg ccaattcaatTTTTATTCCACTTCTACTTTTTAaattaagg ccaattcaatTTTTATT <u>CACTT</u> --ACTTTTTAttatagtatg

¹ DUP/CIR = Illegitimate recombination event yielding a duplicated/circularized product and detected using outward-facing PCR.

² The positions of the direct repeats refer to the numbers of the nucleotides in the published Maize Chloroplast Genome Sequence (10).

Tableau VII. Frequency of recombination events mediated by short direct repeats in the maize plastid genome

Plant line	PCR reactions	Recombination events
WT	19	1
Why1-2	19	3
Why1-2/-1	19	3
Why1-1	19	14

CHAPITRE 3. Deuxième article

Crystal Structures of DNA-Whirly Complexes and Their Role in *Arabidopsis* Organelle Genome Repair

Article publié dans la revue Plant Cell, volume 22, pages 1849-1867, le 15 juin 2010.

3.1 Mise en contexte

Suite aux résultats rapportés au Chapitre 2, plusieurs hypothèses ont été formulées quant aux rôles des protéines Whirly dans les organelles de plantes. Étant donné l'apparition de produits de recombinaison illégitime en absence de protéine Whirly dans les plastides, il est naturel de penser que ces protéines sont impliquées dans la réparation des dommages à l'ADN. L'étude présentée dans ce chapitre décrit l'utilisation d'un stress chimique pour provoquer des bris double brin spécifiquement dans l'ADN des mitochondries et des plastides. Nous avons donc pu soumettre nos différentes plantes mutantes à ce stress et étudier leurs réactions. L'article rapporte également la résolution de la structure tridimensionnelle d'une protéine Whirly avec un ADNsb par cristallographie aux rayons X. Tous ces résultats ont apporté une compréhension plus approfondie du rôle des protéines Whirly dans le maintien de l'ADN des organelles de plante.

Dans cet article, la partie concernant la biologie structurale ainsi que plusieurs buvardage de type Southern et clonage de produits de PCR ont été réalisés par Laurent Cappadocia. Alexandre Maréchal a participé au clonage de produits de PCR et a supervisé le travail d'Étienne Lepage dans la mise au point de stress chimiques. Mon travail a été l'isolation et la préparation des plantes mutantes ainsi que la planification et la réalisation des réactions de PCR. J'ai également participé au processus de rédaction de l'article avec tous les auteurs mais dans une proportion moindre que LC et AM.

Crystal Structures of DNA-Whirly Complexes and their Role in Arabidopsis Organelle Genome Repair

Laurent Cappadocia^a, Alexandre Maréchal^a, Jean-Sébastien Parent^a, Étienne Lepage^a, Jurgén Sygusch^{a,1} and Normand Brisson^{a,1}

^aDepartment of Biochemistry, Université de Montréal, C.P. 6128, Succ. Centre-Ville,

Montréal, Qc, Canada H3C 3J7

Running Title: **Whirlies in Organelle Genome Repair**

¹To whom correspondence should be addressed.

Estimate of article length: 14.3 pages

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: Normand Brisson.

Journal URL : www.plantcell.org

Copyright American Society of Plant Biologists

3.2 Abstract

DNA double-strand breaks are highly detrimental to all organisms and need to be quickly and accurately repaired. Although several proteins are known to maintain plastid and mitochondrial genome stability in plants, little is known about the mechanisms of DNA repair in these organelles and the roles of specific proteins. Here, using ciprofloxacin as a DNA damaging agent specific to the organelles, we show that plastids and mitochondria can repair DNA double-strand breaks through an error-prone pathway similar to the microhomology-mediated break-induced replication observed in humans, yeast, and bacteria. This pathway is negatively regulated by the single-stranded DNA (ssDNA) binding proteins from the Whirly family, thus indicating that these proteins could contribute to the accurate repair of plant organelle genomes. To understand the role of Whirly proteins in this process, we solved the crystal structures of several Whirly-DNA complexes. These reveal a nonsequence-specific ssDNA binding mechanism in which DNA is stabilized between domains of adjacent subunits and rendered unavailable for duplex formation and/or protein interactions. Our results suggest a model in which the binding of Whirly proteins to ssDNA would favor accurate repair of DNA double-strand breaks over an error-prone microhomology-mediated break-induced replication repair pathway.

3.3 Introduction

Plant cells must ensure the accurate replication and faithful transmission of three different genomes located in the nucleus, mitochondria, and plastids. Although the maintenance of nuclear genome stability in eukaryotes is the subject of an intense research effort, the systems that prevent DNA lesions and rearrangements in organelle genomes remain poorly understood. This is unfortunate as the small genomes of plant organelles encode a variety of proteins with critical functions in oxidative phosphorylation or photosynthesis (Knoop, 2004; Raubeson and Jansen, 2005) and any substantial alteration in their DNA could have devastating consequences for plant growth and development (Newton et al., 2004).

Despite the fact that plant organelle genomes are constantly exposed to reactive by-products of the electron transport machineries, they are surprisingly more stable than the nuclear genome at the nucleotide level (Wolfe et al., 1987; Drouin et al., 2008). Contrasting with their slow rate of nucleotide substitution, plant organelle genome organization is complex and highly dynamic. Indeed, DNA rearrangements resulting from spurious recombination between repeated sequences occur frequently in mitochondria and plastids (André et al., 1992; Maul et al., 2002). The prevention and repair of mutations as well as the regulation of recombination events require extensive genome maintenance systems. Plant organelle genomes encode very few, if any, of the proteins involved in these systems (Mackenzie, 2005; Day and Madesis, 2007). Their maintenance is thus mainly ensured by nuclear-encoded DNA replication, recombination, and repair proteins that are targeted to plastids and/or mitochondria. In recent years, some of these proteins have been identified, including the *Arabidopsis thaliana* and *Physcomitrella patens* RecA homologs (Shedge et al., 2007; Odahara et al., 2009), the *Arabidopsis* MutS-like MSH1 (Abdelnoor et al., 2003), the *Arabidopsis* organelle single-stranded DNA binding protein OSB1 (Zaegel et al., 2006), and the plastid-targeted Whirlies of *Arabidopsis* and maize (*Zea mays*) (Maréchal et al., 2009).

Whirlies form a small family of proteins present mainly in the plant kingdom (Desveaux et al., 2005). Different members of this family can localize to the nucleus, plastids, or mitochondria (Desveaux et al., 2005; Krause et al., 2005; Grabowski et al., 2008; Maréchal et al., 2008, 2009; Prikryl et al., 2008). Whirly proteins preferentially bind single-stranded DNA (ssDNA) (Desveaux et al., 2000, 2002, 2005) and perform numerous activities related to DNA metabolism, including the regulation of transcription (Desveaux et al., 2000, 2004; Xiong et al., 2009) and modulation of telomere length (Yoo et al., 2007). Interestingly, the maize Whirly protein WHY1 interacts with DNA from throughout the plastid genome but also binds to a subset of plastid RNAs and participates in chloroplast RNA metabolism (Prikryl et al., 2008). Resolution of the crystallographic structure of the *Solanum tuberosum* WHY1 protein revealed that Whirly proteins form tetramers resembling whirligigs. However, the precise mode of single-stranded nucleic acid binding, as well as the sequence specificity of Whirlies, has so far remained elusive (Desveaux et al., 2002).

The nuclear genome of *Arabidopsis* encodes three Whirly proteins: WHY1 and WHY3, which are targeted to plastids, and WHY2, which localizes to mitochondria (Krause et al., 2005; Maréchal et al., 2008, 2009). Recently, we demonstrated that in the absence of WHY1 and WHY3, the plastid genome of *Arabidopsis* becomes unstable and eventually accumulates rearranged DNA molecules causing chloroplast defects (Maréchal et al., 2009). This phenomenon is not unique to *Arabidopsis*, as rearranged DNA molecules also accumulate in the plastids of maize lines with reduced levels of Zm-WHY1 (Maréchal et al., 2009). The rearranged molecules correspond to deletion, duplication, and/or circularization events that are mediated by DNA microhomologies (Maréchal et al., 2009). Interestingly, microhomology-mediated DNA rearrangements (MHMRs) are also present in the nuclear genomes of humans, yeast, and bacteria (Hastings et al., 2009a). In humans, these MHMRs are associated with numerous copy number variants, as well as with several diseases (Hastings et al., 2009a, 2009b; Lupski, 2009; Conrad et al., 2010), and appear following DNA-related stresses, notably in response to double-strand breaks (DSBs) (Hastings et al., 2009a).

The similarity of the DNA rearrangements observed in plastids of Whirly mutants and in other species led us to suspect that the plastid genome rearrangements could be due to the error-prone repair of DNA following DNA damage and that the role of Whirly proteins in organelles could be to promote the accurate repair of DNA. We tested this hypothesis by treating wild-type and Whirly mutant plants with ciprofloxacin, a DSB-inducing genotoxic agent that targets the enzymes gyrases, which are localized in plant organelles (Wall et al., 2004). Our results indicate that DSBs lead to an accumulation of DNA rearrangements, including MHMRs, in both plastid and mitochondrial genomes. Importantly, plants lacking plastid or mitochondrial Whirlies accumulate higher levels of MHMRs than do the wild-type controls, suggesting that Whirlies might be components of the organelle repair machinery. To gain insights into the mechanisms by which Whirlies prevent DSB-induced rearrangements, we obtained several crystal structures of a mitochondrial Whirly protein in complex with ssDNA oligonucleotides. These structures reveal how Whirlies bind and protect the single-stranded form of DNA in a nonsequence-specific fashion. We propose a model in which Whirlies help maintain genome integrity

under DNA stresses by binding to ssDNA regions of the genome and favoring conservative over error-prone repair pathways.

3.4 Results

DSB Induces MHMR Accumulation in the Plastids of *Arabidopsis*

To test the hypothesis that Whirlies are important for the accurate repair of DNA lesions, Columbia-0 (Col-0; wild-type) and a *why1why3* double knockout (formerly called KO1/3; Maréchal et al., 2009) line of *Arabidopsis* were treated with the quinolone ciprofloxacin. Quinolones are inhibitors of organelle-localized DNA gyrases (Wall et al., 2004) and induce DSBs in DNA (Gellert et al., 1977; Sugino et al., 1977; Snyder and Drlica, 1979; Chen et al., 1996; Rowan et al., 2010). The effect of ciprofloxacin on *Arabidopsis* wild-type plant survival was monitored and a LD50 (dose required to kill half the members of the tested population) of 0.5 μ M ciprofloxacin was calculated (see Figure 27). The error-prone repair of plastid DNA was assayed by monitoring the formation of DNA rearrangements using the PCR approach illustrated in Figure 19 and primer pairs scattered throughout the plastid genome (Maréchal et al., 2009).

As expected, treatment of *Arabidopsis* with ciprofloxacin induces leaf etiolation/variegation and inhibits seedling growth (Figure 20A) (Wall et al., 2004). At 0.25 μ M ciprofloxacin, *Arabidopsis* wild-type plants displayed only slight growth delay and only 9% of the plants displayed etiolation/variegation (Figure 20B). At this same concentration, *why1why3* plants displayed greater sensitivity to ciprofloxacin, with 88% of *why1why3* plants displaying etiolation/variegation (Figure 20B). At 0.75 μ M ciprofloxacin, both genotypes were seriously affected and true leaves were unable to develop. We then monitored the DNA rearrangements in the plastid genomes of these plants using the PCR approach described in Figure 19. Figure 20C shows that whereas DNA rearrangements were almost undetectable in untreated wild-type plants, they were abundant and diverse in untreated *why1why3* plants, in agreement with our previous findings (Maréchal et al., 2009). Treatment of plants with 0.25 μ M ciprofloxacin led to a stronger increase in the number of DNA rearrangements in *why1why3* (13 additional PCR products) compared

with wild-type plants (four additional PCR products) (Figure 20D). At 0.75 μM ciprofloxacin, both genotypes accumulated high levels of rearranged DNA molecules (Figure 20D).

To ensure that the PCR product accumulation was due to DSB induction rather than DNA gyrase inhibition, wild-type and *why1why3* plants were treated with novobiocin, a DNA gyrase inhibitor that also targets organelle-localized DNA gyrases (Wall et al., 2004) but does not produce DSBs (Gellert et al., 1976, 1977; Sugino et al., 1977). Figures 20E and 20F show that novobiocin inhibited seedling growth to the same extent in both wild-type and *why1why3* lines, with very little effect on leaf etiolation/variegation. However, no increase in the diversity or abundance of DNA rearranged molecules could be observed upon treatment with 100 μM novobiocin, the highest concentration tested, in either wild-type or *why1why3* plants (Figures 20G and 20H). This confirmed that DSB induction, rather than DNA gyrase inhibition, leads to DNA rearrangements in ciprofloxacin-treated plants.

DSBs Induce MHMR Accumulation in the Mitochondria of *Arabidopsis*

In contrast with *Arabidopsis* WHY1 and WHY3, no function in DNA metabolism has yet been ascribed to the mitochondria-localized protein WHY2. To verify if this protein has a role in maintaining mitochondrial genome stability under DNA stress conditions, we repeated the experiment with gyrase inhibitors on *why2* knockout plants. No phenotypic differences could be observed between *why2-1* (formerly called KO2; Maréchal et al., 2008) and wild-type plants in the presence or absence of ciprofloxacin (Figures 21A and 21B). We then monitored rearrangements in mitochondrial DNA following ciprofloxacin treatment using the PCR approach described above. Figures 3C and 3D show that rearranged DNA molecules are not abundant in wild-type or *why2-1* plants at 0 μM ciprofloxacin concentration. However, treatment with 0.25 μM ciprofloxacin led to a large increase in rearranged products for *why2-1* plants (18 additional PCR products) compared with wild-type plants (one additional PCR product) (Figure 20D). At 0.75 μM ciprofloxacin, rearranged DNA molecules accumulate to high levels in both wild-type and *why2-1* plants.

Figures 21E and 21F show that *why2-1* and wild-type plants were similarly affected by novobiocin. No rearranged DNA molecules could be detected in plants treated with either

0 or 100 μ M novobiocin (Figures 21G and 21H), thus confirming that mitochondrial DNA rearrangements in ciprofloxacin-treated plants depend on DSB induction. These results indicate that, upon induction of DSBs, *why2-1* plants accumulate DNA rearrangements in the mitochondrial genome more readily than do wild-type plants.

DNA Rearrangements Are Mediated by Microhomologies

Cloning and sequencing of PCR products such as those shown in Figures 20C, 20G, 21C, and 21G led to the identification of 191 DNA rearrangement events (see Tableau X), some of which are presented in Tableau VII. The DNA rearrangements were classified in two groups: MHMR products that had at least 5 bp of microhomology at the junction and nonhomologous end joining (NHEJ)-like products that had <5 bp of microhomology at the junction (McVey and Lee, 2008). MHMR and NHEJ-like products constituted 83 and 17%, respectively, of all DNA rearrangements and were present in both deletion and duplication/circularization events (see Tableau X). Although MHMR junctions varied in sequence, they were frequently A/T rich and often contained homopolymeric tracts (Tableau VII). Some MHMR junctions appeared recurrently in different genotypes and/or in different plant treatments (see Tableau X) and were counted as independent events. This result suggests that DNA rearrangement hot spots exist. However, these may not all represent biologically relevant hot spots as DSBs could have a nonrandom distribution due to the binding of DNA gyrases at preferred locations (Fisher et al., 1981). The 49 mitochondrial MHMR junctions had a mean size of 12.6 ± 5.2 bp, whereas the 109 plastid MHMR junctions had a mean size of 14.5 ± 5.5 bp. Globally, the size of the MHMR junctions varied from 5 to 37 bp. No major difference was observed between the mean size of plastid junctions from ciprofloxacin-treated and untreated plants and that of mitochondrial junctions from ciprofloxacin-treated plants. This, plus the fact that MHMRs were more abundant in Whirly-deficient than in wild-type organelles (see Tableau VIII), indicates that a microhomology-dependent repair pathway is common to both organelles and is more active in Whirly-deficient plants. Importantly, rearrangements in the plastid genome of *why2-1* mutant plants and in the mitochondria of *why1why3* mutant plants were at wild-type levels (see Figure 28), thereby demonstrating the specificity of Whirly

proteins for their respective organelle. Altogether, our data suggest that Whirlies could prevent break-induced DNA rearrangements of *Arabidopsis* organelle genomes.

DNA Rearrangements Mediated by Microhomologies Are Present as Substoichiometric Molecules in the Mitochondrial and Plastid Genomes

DNA gel blot analysis was performed on HindIII-digested DNA from wild-type and *why1why3* plants using isolated PCR products as probes (see Figure 29 and Tableau IX). In untreated plants, only the bands corresponding to the wild-type genome were visible in wild-type or *why1why3* plants. Additional bands were detected in wild-type plants treated with 0.75 μ M ciprofloxacin and in *why1why3* plants treated with 0.25 or 0.75 μ M ciprofloxacin. The fact that the additional bands always have a lower intensity than the band corresponding to the main genome suggests that they are present in substoichiometric amounts compared with the main genome. The DNA gel blot analysis of undigested DNA (see Figure 29 and Tableau IX) revealed the existence of subgenomic-length molecules in wild-type and *why1why3* plants treated with 0.75 μ M ciprofloxacin.

A similar analysis performed on *why2-1* and wild-type plants (see Figure 30 and Tableau IX) indicated that the mitochondria of ciprofloxacin-treated plants also contain additional DNA molecules that are present at substoichiometric levels. Analysis of undigested DNA (see Figure 29 and Tableau IX) further revealed the existence of subgenomic-length molecules in mitochondria. Together with our previous observations, these results suggest that substoichiometric DNA rearrangements accumulate in wild-type, *why2-1*, and *why1why3* genotypes following ciprofloxacin treatment and that some of these rearrangements exist as subgenomic-length molecules.

Mitochondrial Recombination Mediated by Short Repeats (50 to 1000 Nucleotides) Is Not Affected in Whirly Mutants

We next investigated the effect of gyrase inhibitors and Whirly mutants on recombination mediated by short repeats (50 to 1000 nucleotides). This type of recombination is often detected in plant mitochondria (Mackenzie, 2005; Maréchal and Brisson, 2010; Woloszynska, 2010) and is modulated by a variety of proteins involved in maintaining mitochondrial genome stability, such as OSB1 (Zaegel et al., 2006), RecA3 (Shedge et al.,

2007), RecA1 (Odahara et al., 2009), and Msh1 (Abdelnoor et al., 2003). DNA from wild-type, *why2-1*, and *why1why3* plants were analyzed by DNA gel blot and probed with mitochondrial repeats as described by Arrieta-Montiel et al. (2009). DNA gel blot analyses revealed that both ciprofloxacin (see Figure 31) and novobiocin (see Figure 32) trigger mitochondrial recombination mediated by short repeated sequences. However, different recombination patterns were obtained for ciprofloxacin- and novobiocin-treated plants. This was expected since inhibition of DNA gyrase by these compounds leads to different outcomes. Indeed, ciprofloxacin triggers recombination-dependent repair of DSB (López and Blázquez, 2009), whereas novobiocin triggers recombination-dependent restart of stalled replication forks (Woelfle et al., 1993). Interestingly, the recombination patterns were similar for all plant genotypes. This suggests that Whirlies do not modulate the mitochondrial recombination mediated by short repeated sequences (50 to 1000 nucleotides).

The DNA Rearrangements Are Induced Mainly by Stresses Targeting the Organelles

We next treated wild-type, *why2-1*, and *why1why3* plants with bleomycin and hydroxyurea. In *Arabidopsis*, bleomycin triggers a DSB stress response (Menke et al., 2001; Tamura et al., 2002; Chen et al., 2008), whereas hydroxyurea induces a replication stress response (Ferreira et al., 1994; Roa et al., 2009). Treating wild-type and Whirly mutant plants with bleomycin (see Figure 33) or hydroxyurea (see Figure 34) did not result in an increase in plastid or mitochondrial MHMR. Furthermore, no increase in mitochondrial recombination mediated by short repeated sequences was observed upon treatment of *Arabidopsis* plants with these compounds (see Figure 35). These results suggest that bleomycin and hydroxyurea have a weaker effect than do ciprofloxacin and novobiocin on the formation of DNA rearrangements in *Arabidopsis* organelle genomes. As shown in other species (Baugnet-Mahieu et al., 1971; Shen et al., 1995; Morel et al., 2008), this could be due to the fact that bleomycin and hydroxyurea target both the nucleus and the organelles, whereas compounds such as ciprofloxacin and novobiocin target mainly the organelles (Ye and Sayre, 1990). However, we cannot exclude the possibility that DNA damage induced by bleomycin and hydroxyurea in the organelles is repaired by alternative mechanisms.

The Crystal Structure of St-Why2 Reveals How ssDNA Is Bound

To understand how Whirly proteins prevent MHMR accumulation in the organelle genomes of *Arabidopsis*, we sought to determine how these proteins bind ssDNA. We thus solved the crystal structure of St-WHY248-216, a close homolog of At-WHY2 that also localizes to mitochondria (Vermel et al., 2002), in the free form and bound to ssDNA (Cappadocia et al., 2008). The structures were obtained at 2.2- to 2.7-Å resolution (Tableau VIII). The St-WHY248-216 construct, hereafter called WHY2, encompasses the entire ssDNA binding domain but lacks the mitochondria targeting sequence and the acidic/aromatic C terminus (Figure 22A). These regions, which are not required for high affinity binding to the ssDNA (Cappadocia et al., 2008), are predicted to be disordered in solution. For the ssDNA sequence, we chose ERE32, a 32-nucleotide ssDNA derived from the elicitor response element (ERE) because the interaction between this sequence and plant Whirlies is well documented (Desveaux et al., 2000, 2004). WHY2 and WHY2-ERE32 crystallized in different chemical environments, yet they both gave rise to isomorphous crystals of the F432 space group. In the two crystals, WHY2 tetramers are disposed around fourfold crystallographic axes (Figures 22B and 22C). In the WHY2-ERE32 structure, electron density corresponding to ssDNA was clearly visible (Figure 22C), enabling us to build a nine-nucleotide model. Although we could observe electron density for most nucleobases, we were unable to assign unambiguously the ERE32 sequence into the electron density. We thus modeled the thymine-rich ERE32 oligonucleotide as an all-thymine oligonucleotide.

Shown in Figure 22C is the electron density corresponding to ssDNA that was observed primarily on the edges and in between the β -sheets of adjacent protomers. Such a positioning of the DNA was unexpected. Indeed, while most single-stranded nucleic acid proteins use the core of their β -sheets as a primary binding platform (Horvath, 2008), WHY2-ssDNA interaction relies mainly on the binding of the DNA between properly positioned domains. As a consequence, binding to ssDNA exploits, but also depends on, the fourfold symmetry of the Whirly protein. The root mean square deviation (RMSD) of merely 0.5 Å for all C atoms between the free and the DNA-bound forms of WHY2

indicates that the protein does not undergo major conformational changes upon DNA binding. This is consistent with a need for properly positioned ssDNA binding residues.

Figure 36 shows that whereas good electron density corresponding to ssDNA is observed near the edges of the β -sheets, this density fades out on the top of the β -sheets before reappearing near the other edge of the β -sheets. This suggests that the ssDNA is disordered and/or adopts multiple conformations on the top of the β -sheets. We propose that, in this region, long DNA molecules, such as the ERE32, either thread from one ssDNA binding site to the other or enter/leave the complex.

WHY2 Binds ssDNA through a Conserved Mechanism

Figure 23A indicates that ssDNA is maintained in an extended conformation in each ssDNA binding site. The single-strand preference of WHY2 arises from steric impediment for the binding of a second strand at several positions on the binding surface. The presence of abrupt twists in the DNA backbone, such as between nucleotides 2 and 3, also argues against binding of a double-stranded DNA (dsDNA) helix. The mode of ssDNA binding is dominated by stacking and hydrophobic interactions between adjacent nucleobases and between nucleobases and aromatic/hydrophobic protein residues, consistent with the salt-resistant binding of St-WHY2 to ssDNA (Vermel et al., 2002). The polar contribution to DNA binding involves seven hydrogen bonds plus one water-mediated interaction (Figure 23A). Few sequence-specific interactions between St-WHY2 and ERE32 were observed (see below). Indeed, most of the nucleobases have their edges, containing the sequence-specific binding moieties, exposed to the solvent, whereas the faces of the nucleobases make intimate contact with residues of the protein surface.

A structure-based sequence alignment reveals that most of the ssDNA binding residues are conserved among different plastid- and mitochondria-targeted Whirlies (Figure 23B). Structural alignment of St-WHY1 (PDB 1L3A) with St-WHY2 also indicates a strong structure conservation both at the monomeric level with an RMSD of 1.4 Å for 159 matched C atoms and at the tetrameric level with an RMSD of 1.9 Å for 642 matched C atoms. Furthermore, structural superposition of St-WHY1 with St-WHY2 reveals that the key ssDNA binding residues are properly positioned in St-WHY1 to contact ssDNA. This

supports the hypothesis that plastid- and mitochondria-targeted Whirlies have closely related ssDNA binding interfaces and is consistent with a common role of Whirly proteins in both organelles. Importantly, ssDNA binding residues are also conserved in *Arabidopsis* proteins, suggesting that the mechanism of ssDNA binding of St-WHY2 is similar to the ones of *Arabidopsis* WHY1, WHY2, and WHY3 (Figure 23B).

WHY2 Binds Nucleobases That Differ in Size and in Functional Groups

The local disorder of the electron density around the nucleobases at their junction with the β -sheets, as well as the scarcity of sequence-specific contacts between St-WHY2 and ERE32, suggests that WHY2 could bind other ssDNA sequences with high affinity. We thus assayed the binding of St-WHY2 to different A/T-rich ssDNA, as plant organelle genomes are generally A/T rich (Fauron et al., 2004; Ravi et al., 2008). Quantitative electrophoretic mobility shift assays (EMSAs) revealed that, despite their different nucleotide sequences, the oligonucleotides ERE32, cERE32, rcERE32, and dT32 all interacted with WHY2 with similar nanomolar affinities (Figure 24A, Tableau IX). To gain structural insights on how WHY2 could accommodate these DNA sequences, we obtained the crystal structure of these oligonucleotides bound to St-WHY2 (Tableau VIII) and compared them to that of WHY2-ERE32. Comparison of the unbiased difference electron densities of the four oligonucleotides reveals that they adopt similar conformations in the ssDNA binding site (Figure 24B), suggesting that they interact with WHY2 through a common mechanism. The structures also reveal how WHY2 can accommodate bases that differ in size and in functional groups. Indeed, Ser-62 and Asp-142, which make sequence-specific interactions with T7 and T3 in the WHY2-dT32 and WHY2-ERE32 structures, respectively, make equivalent interactions with A7 and A3 in the structures of WHY2-cERE32 and WHY2-rcERE32 (Figures 24C and 24D). These compensating interactions demonstrate the capacity of WHY2 to bind different DNA sequences with high affinity.

WHY2 Binds to Single-Stranded Overhangs, Melts dsDNA, and Protects ssDNA

An important step in the repair of DSBs is the rapid resection of DNA by exonucleases to generate DNA with single-stranded overhangs (Persky and Lovett, 2008). Because the choice of a repair pathway depends on the identity of the repair proteins that bind resected DNA ends, we verified that Whirly proteins could bind ssDNA overhangs. Figure

25A indicates that St-WHY2 can bind 16, but not 8, nucleotide-long 5' or 3' overhangs. Intriguingly, ssDNA molecules of the same size as these overhangs are not bound with high affinity by Whirlies (Yoo et al., 2007). This suggests either that the dsDNA moiety contributes to the high-affinity binding or that high-affinity binding is achieved through partially melting of the dsDNA moiety. As Whirlies were shown to have much less affinity for dsDNA than ssDNA (Desveaux et al., 2000; Prikryl et al., 2008), and since our structural data suggest that dsDNA is a poor substrate for Whirlies, we tested whether Whirlies could melt dsDNA to bind additional nucleotides. We incubated a radiolabeled DNA duplex with WHY2 and monitored the WHY2-induced denaturation of dsDNA by electrophoresis. Figure 25B shows that more ssDNA is produced as the concentration of WHY2 is increased. This indicates that WHY2 can destabilize DNA duplexes, probably by binding with a higher affinity to the single-stranded form of DNA and driving the equilibrium toward the formation of ssDNA. This could explain the capacity of WHY2 to bind short ssDNA overhangs with high affinity.

In addition to its ssDNA binding capacity, WHY2 effectively protects DNA against nuclease degradation. As shown in Figure 25C, phage M13mp18 ssDNA preincubated with a saturating amount of WHY2 is protected against digestion by mung bean (*Vigna radiata*) nuclease, whereas free M13mp18 ssDNA is readily digested. Also, Figure 25D shows that WHY2 can protect the single-stranded overhang of a duplex against degradation by the exonuclease activity of T4 DNA polymerase. The fact that WHY2 protects ssDNA against both mung bean nuclease and T4 DNA polymerase likely suggests that WHY2 can prevent these nucleases from accessing the ssDNA.

3.5 Discussion

Binding of Whirly Proteins to Nucleic Acids

There have been conflicting reports concerning the sequence specificity of Whirly proteins. Indeed, Whirlies were reported to bind three nuclear sequences that share little sequence similarity with each other: an ERE element (Desveaux et al., 2000, 2004), an *Arabidopsis* telomeric repeat (Yoo et al., 2007), and a distal element upstream of the

Arabidopsis kinesin gene KP1 (Xiong et al., 2009). In plant organelles, Whirly proteins were shown to bind many regions of the plastid or the mitochondrial genome without obvious sequence consensus in vivo (Maréchal et al., 2008, 2009; Prikryl et al., 2008). This work demonstrates that Whirlies can bind ssDNA molecules that differ in nucleotide sequence using a binding mechanism involving mainly contacts with nonsequence-specific moieties of the ssDNA and allowing Whirlies to establish equivalent interactions with different nucleobases. Because Whirlies fused to green fluorescent protein accumulate abundantly in the organelles, whereas they are difficult to detect in the nucleus (Krause et al., 2005), we propose that, in the nucleus, the low concentration of Whirly proteins would enable them to bind only a limited number of sites that have a pronounced single-stranded character (e.g., A/T-rich sequences of transcribed regions or telomeric repeats). In the organelles, the abundance of Whirly proteins and their limited sequence specificity would enable them to bind all available ssDNA.

In addition to ssDNA, Whirlies were shown to bind RNA in maize (Prikryl et al., 2008). Furthermore, a close structural neighbor of plant Whirlies also binds RNA (Schumacher et al., 2006). Our crystallographic models are consistent with a role of Whirly proteins in RNA metabolism. Indeed, modeling of ssRNA in place of ssDNA leads only to small clashes, thus suggesting that RNA could be accommodated in an ssDNA-like conformation (see Figure 37). However, RNA appears to be bound less strongly than DNA because Whirlies were found to be associated with DNA from throughout the genome but with only a subset of RNA molecules in vivo (Prikryl et al., 2008).

DSBs Can Be Repaired through a Microhomology-Mediated Break-Induced Replication Mechanism in Plant Organelles

Plants treated with ciprofloxacin accumulate rearranged DNA molecules in both plastids and mitochondria. Most of the DNA rearrangements contain microhomologies at their junction, suggesting that at least part of the ciprofloxacin-induced DSBs are repaired through an error-prone microhomology-dependent pathway. We previously proposed that a microhomology-mediated break-induced replication (MMBIR) pathway could explain the accumulation of MHMR in both untreated *why1why3 Arabidopsis* and *Why1* knockdown maize plants (Maréchal et al., 2009). The MMBIR pathway relies on the

microhomology-dependent restart of DNA synthesis on a different template following collapse of the replication fork (Hastings et al., 2009a). Microhomologies are sufficient to initiate DNA synthesis as certain DNA polymerases can efficiently use mismatched primers or primers as short as 2 to 3 bp as substrates (Cannistraro and Taylor, 2007; Hastings et al., 2009a). Our present finding that MHMR events that appear in untreated plants are increased upon ciprofloxacin treatment suggests that a MMBIR pathway is involved in the error-prone repair of DSBs in organelles.

The use of a MMBIR pathway is consistent with our observation of both duplications/circularization and deletion events, whereas pathways such as microhomology-mediated end joining or single-strand annealing can account only for deletion events (McVey and Lee, 2008). A NHEJ repair pathway could be involved in the formation of NHEJ-like events. However, there is little support for the existence of such a repair pathway in plant organelles, particularly in plastids (Odom et al., 2008; Kohl and Bock, 2009). Furthermore, whereas MHMRs are detected in the absence or presence of ciprofloxacin, all but one NHEJ-like event are detected in ciprofloxacin-treated plants, thus raising the possibility that these events could be by-products of the ciprofloxacin-mediated DNA gyrase inhibition (Marvo et al., 1983). The fact that mismatches are detected in 91 out of 158 MHMR junctions relative to one or both parental strands (see Tableau X) is consistent with the low-specificity requirements of the MMBIR pairing process. Finally, MMBIR can account for complex cases of DNA rearrangements (Hastings et al., 2009a), including the formation of molecules containing multiple MHMRs, which have also been observed in our study (Tableau VII). Our results also suggest that MHMR products exist at substoichiometric levels. These DNA variants could accumulate to higher levels through a process known as substoichiometric shifting, which is well documented for mitochondrial DNA rearrangements (Small et al., 1987; Janska et al., 1998; Maréchal and Brisson, 2010; Woloszynska, 2010). The fact that, in *why1why3* plants, MHMR products present at low abundance in the chloroplasts of green sectors accumulate to high levels in the variegated sectors (Maréchal et al., 2009) suggests that this substoichiometric shifting could also occur in chloroplasts. A substoichiometric shift may also facilitate the transmission of rearranged molecules. Indeed, in the *why1why3* plants,

variegation could readily be inherited provided that the female parent displayed a strong variegation phenotype (Maréchal et al., 2009). In this study, however, no increase in DNA rearrangements was observed in the progeny of plants treated with 0.25 μ M ciprofloxacin. It is thus possible that DNA rearrangements, when present at substoichiometric levels, have poor transmission efficiency. Alternatively, germ cells may be less sensitive to ciprofloxacin treatment than are somatic cells. Globally, the characteristics of the MHMR events obtained in this study support the idea that MMBIR acts as a DNA repair pathway in plant organelles (Figure 26).

Effects of DSBs and Whirly Mutants on Mitochondrial Recombination Mediated by Short Repeated Sequences (50 to 1000 Nucleotides)

Plants treated with ciprofloxacin or novobiocin displayed an increase in mitochondrial recombination mediated by short repeated sequences. As the plastid genome of *Arabidopsis* does not contain short repeated sequences capable of supporting aberrant homologous recombination (Maréchal and Brisson, 2010), we were unable to determine if ciprofloxacin or novobiocin affected homologous recombination in plastids. Whirlies do not appear to modulate mitochondrial DNA recombination mediated by short repeated sequences. This was expected for the plastid-localized proteins WHY1 and WHY3 and is consistent with our previous observations that DNA recombination products mediated by short repeats do not accumulate in untreated *Arabidopsis* plants lacking the mitochondria-localized WHY2 (Maréchal et al., 2008). As At-WHY2 appears to modulate only the DSB-induced mitochondrial MMBIR pathway, it is possible that these repair responses are mechanistically distinct.

However, we cannot rule out a role for WHY2 in homologous recombination because other proteins may function redundantly with the Whirlies in this process. Indeed, many different ssDNA binding proteins are located in mitochondria, including the OSB proteins (Zaegel et al., 2006) and the mitochondrial ssDNA binding proteins (Edmondson et al., 2005).

Whirly Proteins Share Common Characteristics with Eubacterial ssDNA Binding Proteins

The eubacterial ssDNA binding proteins SSBs are critical players during DNA replication, recombination, and repair processes. Often envisioned as simple ssDNA-coating proteins, they actually work as organizational scaffolds that recruit genome maintenance complexes when and where they are most required (Cox, 2007; Shereda et al., 2008). These proteins share a number of characteristics with the Whirly proteins even though they are not structurally related: (1) they are both tetrameric proteins that bind ssDNA with high affinity and modest sequence preference, (2) they use mainly hydrophobic/aromatic residues to contact ssDNA and, as a result, their binding to ssDNA is salt resistant, (3) they can destabilize dsDNA, (4) they can prevent degradation of ssDNA by nucleases, and (5) they both possess a conserved C-terminal acidic/aromatic tail. Usually tethered to the protein surface, the acidic/aromatic C terminus of the eubacterial ssDNA binding proteins SSBs is believed to be freed and rendered available for interacting with target proteins upon ssDNA binding (Lu and Keck, 2008; Shereda et al., 2008, 2009). Interestingly, structural comparison between St-WHY1 and St-WHY2 revealed that the C terminus in WHY1 and the ssDNA in WHY2 both interact with the same residues in the β -sheets. This demonstrates that the C terminus and the ssDNA can compete for the same binding surface on the Whirly proteins and that their interaction with the Whirly domain is mutually exclusive. The competition between ssDNA and an acidic/aromatic C-terminal tail is consistent with our previous results showing that a St-WHY1 protein mutated at its C terminus has an increased affinity for ssDNA (Desveaux et al., 2005). Our results are thus consistent with a model in which, upon binding of Whirly proteins to ssDNA, the acidic-aromatic C-terminal tail would become available for interacting with target proteins in an SSB-like mode.

Possible Roles for Whirly Proteins in the Repair of DSBs

Whirlies could maintain genome stability by favoring the repair of DSBs by error-free homologous recombination and/or by restricting the DSB repair by error-prone MMBIR. Based on our results, we propose that Whirlies could bind and protect resected DNA ends in a sequence-independent manner at break sites (Figure 26). Upon binding to resected DNA ends, Whirlies would prevent annealing of ssDNA overhangs to microhomologous

sequences present in stretches of ssDNA elsewhere in the genome. Alternatively, or concomitantly, Whirlies could prevent MMBIR by binding to these ssDNA stretches, which may occur in many places in the genome, including in DNA secondary structures, in replication forks, in stalled transcription complexes, and in promoter regions (Hastings et al., 2009a). In addition, Whirlies might regulate the access of repair proteins to ssDNA by competing with them for binding to ssDNA, thereby influencing the choice of a repair pathway. Whirlies could even promote the accurate repair of DSBs by recruiting proteins involved in DNA repair through their acidic/aromatic C terminus in an SSB-like fashion.

No accumulation of rearranged DNA molecules could be detected in the mitochondria of *why2-1* plants in absence of ciprofloxacin treatment. This could suggest that mitochondrial DNA is less subject to genotoxic stresses under normal growth conditions or that WHY2 fulfils a different function than do WHY1 and WHY3 in *Arabidopsis*. A partial functional redundancy by the OSBs and mitochondrial SSBs could also explain why, in the absence of DNA stresses, *why2-1* plants do not show an increase in genome instability, contrary to what is observed in plastids of *Arabidopsis* and maize plants lacking plastid-localized Whirlies (Maréchal et al., 2009). This idea is consistent with the apparent lack of SSB homologs targeted to plastids. A functional redundancy could also explain our observation that, upon DSB induction, DNA rearrangements accumulate more readily in plastids of *why1why3* plants than in mitochondria of *why2-1* plants. Further work is needed to determine if WHY2 works redundantly with other mitochondria-localized ssDNA binding proteins or if the protection against MHMR observed in *why2-1* mutants could be due to a more general role of WHY2 in DNA metabolism.

Interestingly, TIF1, a distant homolog of the plant Whirlies (Saha et al., 2001) safeguards chromosomes from DNA damage in both nuclei of the ciliate *Tetrahymena thermophyla*. TIF1-deficient cells are hypersensitive to genotoxic stress and are defective in the activation of the intra-S phase checkpoint mediated by the sensor/transducer kinase ATR (Yakisich et al., 2006). These studies suggest that a role of Whirlies in the maintenance of genome stability may be conserved between plants and ciliates.

Our finding that MHMR products accumulate at high concentrations of ciprofloxacin even in the presence of Whirly proteins suggests that accurate repair of DNA may be bypassed

if severe DNA damage occurs. Interestingly, in other organisms, MHMR was detected mainly when canonical repair pathways such as NHEJ or homologous recombination were overloaded or challenged, thus leading to the idea that microhomology-mediated DNA repair acts as a backup DNA repair pathway (Wang et al., 2003). In *Arabidopsis* organelles, multiple DSBs may cause depletion of repair proteins due to their recruitment to multiple damage sites, allowing the accumulation of MHMRs. Our observation that MHMR products accumulate to high levels at 0.75 μ M ciprofloxacin both in the absence and the presence of Whirly proteins is consistent with this idea. However, the use of an error-prone repair pathway under DNA stress may be more than a backup pathway. Indeed, as MHMR enables DNA rearrangements including copy number variation (Hastings et al., 2009a), it promotes rapid changes in the genome. Some of these changes may confer selective advantages to plants under stress. A similar mechanism, termed stress-induced mutation, exists in bacteria and is also dependent on error-prone repair of DSBs (Ponder et al., 2005). The fact that certain plants maintain DNA rearrangements mediated by microhomologies in their organellar genome provides indirect support for their selective advantage (Ogihara et al., 1988; Kanno et al., 1993; Moeykens et al., 1995).

3.6 Methods

Plant Material and Growth Conditions

Sterilized seeds of *Arabidopsis thaliana* ecotype Col-0 (wild type), the *why1why3* double mutant, and the *why2* single mutants were germinated on Murashige and Skoog basal media (Sigma-Aldrich) containing ciprofloxacin or novobiocin at the indicated concentrations and collected after 3 weeks of growth under long-day conditions (16 h light/8 h dark). Bleomycin and hydroxyurea treatments were performed as described (Chen et al., 2008; Roa et al., 2009). The *why2-1* and *why1why3* mutant lines (formerly called KO2 and KO1/3, respectively) were described previously (Maréchal et al., 2008, 2009). An additional *Arabidopsis* line lacking WHY2 (*why2-2*; SALK_016156) was tested for its sensitivity to the DNA damaging agents and the generation of MHMR products during ciprofloxacin treatment (see Figure 38C) and similar results were obtained. A protein gel

blot was performed using α -AtWhy2 polyclonal antibodies (Maréchal et al., 2008) to verify that both *why2* lines are completely devoid of WHY2 protein (see Figure 38).

Detection of DNA Rearrangements

DNA was isolated from plants using a cetyl trimethylammonium bromide DNA extraction protocol (Weigel and Glazebrook, 2002). PCR was conducted using the Taq polymerase (Genscript) according to the manufacturer's instructions. DNA rearrangement events were detected using both outward- and inward-facing PCR primers spaced by 5 to 30 kb. Thirteen and 10 PCR primer pairs scattered in the plastid genome were used for ciprofloxacin- and novobiocin-treated plants, respectively. Twenty and 10 PCR primer pairs scattered in the mitochondrial genome were used for ciprofloxacin- and novobiocin-treated plants, respectively. The sequences of the PCR primers are listed in Tableau XIV. PCR was performed on DNA samples from wild-type, *why2-1*, and *why1why3* plants and analyzed by gel electrophoresis. All visible DNA bands were isolated and cloned into the pDrive vector (Qiagen). The DH5 strain was used for transformation, and the plasmids from two randomly picked colonies were sequenced. When both sequencing products were identical, only one product was considered, whereas when sequencing products differed, both products were considered. Sequencing of visible bands confirmed that >96% of all cloned PCR products were due to DNA rearrangements, with the remaining events being caused by nonspecific annealing of the oligonucleotides at undesired locations. DNA rearrangements are listed in Tableau X.

Analysis of DNA Rearrangements by DNA Gel Blots

To evaluate the relative abundance of MHMRs compared with normal organelle genomes, rearranged DNA molecules were amplified by PCR using the oligonucleotides listed in Tableau XIV. Cloning and sequencing of PCR products confirmed the occurrence of DNA rearrangements (see Tableau X). PCR products were internally radiolabeled using [γ -³²P]dATP (6000 Ci/mmol; Perkin-Elmer Life Science) and Klenow polymerase (Fermentas) and were used as probes. DNA gel blots were performed as described previously (Maréchal et al., 2009). Detection of recombination mediated by short repeats (50 to 1000 nucleotides) was performed as described (Arrieta-Montiel et al., 2009). Briefly, PCR primers listed in Tableau XIV were used to amplify three mitochondrial repeats. The

repeats were internally radiolabeled using [-32P]dCTP (3000 Ci/mmol; Perkin-Elmer Life Science) and Klenow polymerase (Fermentas) and were used as probes.

St-WHY2 Cloning, Expression, and Purification

Details concerning cloning, expression, and purification of St-WHY2 are reported elsewhere (Cappadocia et al., 2008). Briefly, WHY248-216 was amplified from *Solanum tuberosum* total RNA and cloned into the plasmid vector pET-21a (Novagen), which encodes a C-terminal hexahistidine tag. St-WHY2 was expressed in *Escherichia coli* BL21 (DE3) strain with 1 mM isopropyl β -D-thiogalactopyranoside for 3 h at 37°C. The cells were lysed by alumina grinding and the lysate resuspended in 20 mM sodium phosphate buffer, pH 7.5, containing 500 mM NaCl and 25 mM imidazole. The recombinant protein was purified by applying the supernatant to a HiTrap Chelating column (GE Healthcare) followed by size exclusion chromatography on a Superdex 200 16/60 (GE Healthcare) in 20 mM Tris-HCl, pH 8.0, 100 mM NaCl.

Data Collection and Structure Determination

St-WHY2 in the free form and St-WHY2-ssDNA complexes were crystallized as previously described (Cappadocia et al., 2008). Briefly, protein-DNA complexes were crystallized by the hanging drop vapor diffusion method using a precipitant solution containing 100 mM Tris/HCl, pH 8.0, 13.5 to 25% (w/v) PEG6000, and 1 to 2 M LiCl. The free protein was crystallized by the hanging drop vapor diffusion method using a precipitant solution containing 100 mM MOPS, pH 7.0, 21 to 27% (w/v) PEG1000, and 100 to 400 mM $\text{NH}_4\text{H}_2\text{PO}_4$. Diffraction data were collected using ADSC Quantum 315 CCD detectors at beamlines X25 and X29 at the National Synchrotron Light Source (NSLS) at the Brookhaven National Laboratory or using a MAR-225 CCD detector at beamline 22-ID at the Advanced Photon Source at the Argonne National Laboratory. Initial phases for the St-WHY2-ERE32 complex were obtained by molecular replacement using St-WHY1 (PDB 1L3A) as a search template. Phases were improved by iterative cycles of model building with Coot (Emsley and Cowtan, 2004) and refinement with CNS (Brünger et al., 1998) and Phenix (Adams et al., 2002). DNA nucleotides were modeled one nucleobase at a time from difference electron density maps. The refined structure of St-WHY2-ERE32 was used as a molecular replacement search template for structural determination of data sets collected from

crystal of other St-WHY2-ssDNA complexes. Clear density was visible for protein residues 55 to 215 in St-WHY2. Since we could not attribute unambiguously the specific sequences of the ERE32, rcERE32, and cERE32 oligonucleotides to the electron density, we assigned these oligonucleotides as all-thymine, all-adenine, and all-adenine oligonucleotides, respectively. The DNA occupancy in each structure was calculated from the temperature factors and it ranges from 0.7 to 0.8. Test data sets were randomly selected from the observed reflections prior to refinement. The figures were prepared using PyMOL (DeLano, 2002).

EMSAs

The apparent dissociation constants for St-WHY2 bound to DNA were determined by EMSA. The ssDNA oligonucleotides were obtained from Integrated DNA Technologies and radiolabeled at the 5' end using [³²P]ATP (6000 Ci/mmol; Perkin-Elmer Life Science) and T4 polynucleotide kinase (Fermentas) following the manufacturer's instructions. Binding reactions were done for 30 min at 20°C by incubating 15,000 cpm radiolabeled ssDNA (<100 pM) with the serially diluted St-WHY2 protein in a buffer containing 10 mM Tris-HCl, pH 8.0, 150 mM KCl, 0.5 mM EDTA, pH 8.0, and 15% (v/v) glycerol. Reactions were loaded on 10% (w/w) 29:1 acrylamide:bisacrylamide gels with 4.5 mM Tris, 4.5 mM boric acid, and 1 mM EDTA, pH 8.0, at 20°C under 150 V constant voltage for 20 min and then subjected to autoradiography. The autoradiograms were scanned and the intensity of the bands corresponding to bound and free radiolabeled ssDNA were quantified using ImageJ (NIH; <http://rsb.info.nih.gov/ij/>). The data was plotted and analyzed using Prism 5 Demo (GraphPad Software). The apparent dissociation constants were calculated by fitting the plot of the fraction of ssDNA bound versus protein concentration to the Hill equation.

The binding of St-WHY2 to ssDNA overhangs was also assessed by EMSA using DNA duplex formed by annealing two of the following oligonucleotides: YG, YC, YG8_3, YC8_5, YG16_3, and YC16_5. The sequences of these oligonucleotides are shown in Supplemental Tableau 4 online.

DNA Destabilization Assay

Oligonucleotide R2 was radiolabeled at the 5' end using [³²P]ATP (6000 Ci/mmol; Perkin-Elmer Life Science) and T4 polynucleotide kinase (Fermentas) following the manufacturer's instructions and annealed with oligonucleotide R1. The annealed duplex was incubated with St-WHY2 or BSA for 30 min at 20°C. The reactions were stopped by the addition of a 40-fold excess of unlabeled R2 and SDS to a final concentration of 0.2%. Reactions were separated on 7.5% (w/w) 29:1 acrylamide:bisacrylamide gels with 4.5 mM Tris, 4.5 mM boric acid, and 1 mM EDTA, pH 8.0, at 20°C under 130 V constant voltage for 30 min. The sequences of oligonucleotides R1 and R2 are shown in Tableau XIV.

Mung Bean Nuclease Protection Assay

Two hundred nanograms of M13mp18 ssDNA (USB) either alone or prebound for 30 min at 20°C with St-WHY2 at a 1:10 protein/nucleotide ratio were incubated with 1U of mung bean (*Vigna radiata*) nuclease (New England Biolabs) for 0, 15, 30, or 60 min at 30°C. Reactions were stopped by incubation on ice and addition of SDS and EDTA to final concentrations of 0.1% (w/v) and 1.5 mM, respectively. Reactions were loaded on 0.7% (w/w) agarose gels and migrated under 100 V of constant voltage.

T4 DNA Polymerase Protection Assay

ssDNA or a dsDNA with a 3' overhang of 16 nucleotides (3'-dsDNA) either alone or prebound for 30 min at 20°C with 1 μM St-WHY2 was incubated with 0.02 units of T4 DNA Polymerase (T4; Fermentas) for 0, 15, 30, or 60 min at 20°C. Reactions were stopped by incubation on ice and addition of SDS to a final concentration of 0.1% (w/v). Reactions were separated on 10% (w/w) 29:1 acrylamide:bisacrylamide gels with 4.5 mM Tris, 4.5 mM boric acid, and 1 mM EDTA, pH 8.0, at 20°C under 150 V constant voltage for 20 min and then subjected to autoradiography.

Accession Numbers

Sequence data from this article can be found in the *Arabidopsis* Genome Initiative, GenBank/EMBL, or The Institute for Genomic Research database under the following accession numbers: St-WHY2 (HM234504), At-WHY1 (AT1G14410), At-WHY2 (AT1G71260), and At-WHY3 (AT2G02740). The accession numbers for the mitochondrial

and plastid genomes are NC_001284 (Unsel'd et al., 1997) and NC_000932 (Sato et al., 1999), respectively. The atomic coordinates of St-WHY2 in the free form, St-WHY2-ERE32, St-WHY2-dT32, St-WHY2-cERE32, and St-WHY2-rcERE32, have been deposited in the Protein Data Bank (www.rcsb.org) under ID codes 3N1H, 3N1I, 3N1J, 3N1K, and 3N1L, respectively. The ID code for the crystal structure of St-WHY1 in the free form is 1L3A.

3.7 Acknowledgments

The assistance of Mathieu Coinçon, Sandra Grondin, and Bruno Piché is gratefully acknowledged. We thank James Omichinski for valuable discussions. We also thank the High Throughput Crystallization Service at the Hauptman-Woodward Institute, Buffalo, NY, for determining the initial crystallization conditions. Research carried out at the NSLS and at the APS was supported by the U.S. Department of Energy, Division of Materials Sciences and Division of Chemical Sciences, and by the U.S. Department of Energy, Office of Science, respectively. Assistance by X25, X29, and 22ID beamline personnel is gratefully appreciated. L.C. and A.M. were supported by fellowships from the Natural Sciences and Engineering Research Council of Canada (NSERC) and from the Fonds québécois de la recherche sur la nature et les technologies (FQRNT). This research was supported by grants from NSERC and FQRNT to both N.B. and J.S. and from the Canadian Institutes of Health Research to J.S.

3.8 References

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

Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchettini, J.C., Sauter, N.K., and Terwilliger, T.C. (2002). PHENIX: Building new software for automated crystallographic structure determination. *Acta Crystallogr. D Biol. Crystallogr.* 58: 1948–1954.

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

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.

Baugnet-Mahieu, L., Goutier, R., and Baes, C. (1971). Differential response of mitochondrial and nuclear DNA syntheses to hydroxyurea in normal and regenerating rat liver. *Biochem. Pharmacol.* 20: 141–149.

Brünger, A.T., et al. (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* 54: 905–921.

Cannistraro, V.J. and Taylor, J.S. (2007). Ability of polymerase eta and T7 DNA polymerase to bypass bulge structures. *J. Biol. Chem.* 282: 11188–11196.

Cappadocia, L., Sygusch, J., and Brisson, N. (2008). Purification, crystallization and preliminary X-ray diffraction analysis of the Whirly domain of StWhy2 in complex with single-stranded DNA. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 64: 1056–1059.

Chen, C.R., Malik, M., Snyder, M., and Drlica, K. (1996). DNA gyrase and topoisomerase IV on the bacterial chromosome: Quinolone-induced DNA cleavage. *J. Mol. Biol.* 258: 627–637.

Chen, I.P., Mannuss, A., Orel, N., Heitzeberg, F., and Puchta, H. (2008). A homolog of ScRAD5 is involved in DNA repair and homologous recombination in *Arabidopsis*. *Plant Physiol.* 146: 1786–1796.

Conrad, D.F., Bird, C., Blackburne, B., Lindsay, S., Mamanova, L., Lee, C., Turner, D.J., and Hurles, M.E. (2010). Mutation spectrum revealed by breakpoint sequencing of human germline CNVs. *Nat. Genet.* 42: 385–391.

Cox, M.M. (2007). Regulation of bacterial RecA protein function. *Crit. Rev. Biochem. Mol. Biol.* 42: 41–63.

Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall, W.B., III, Snoeyink, J., Richardson, J.S., and Richardson, D.C. (2007). MolProbity: All-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* 35: W375–383.

Day, A. and Madesis, P. (2007). DNA replication, recombination and repair in plastids. In *Cell and Molecular Biology of Plastids. Topics in Current Genetics, Vol. 19*, R. Bock, ed (Berlin, Heidelberg, Germany: Springer), pp. 65–119.

DeLano, W.L. (2002). *The Pymol Molecular Graphics System*. (San Carlos, CA: DeLano Scientific).

Desveaux, D., Allard, J., Brisson, N., and Sygusch, J. (2002). A new family of plant transcription factors displays a novel ssDNA-binding surface. *Nat. Struct. Biol.* 9: 512–517.

Desveaux, D., Després, C., Joyeux, A., Subramaniam, R., and Brisson, N. (2000). PBF-2 is a novel single-stranded DNA binding factor implicated in PR-10a gene activation in potato. *Plant Cell* 12: 1477–1489.

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

Desveaux, D., Subramaniam, R., Després, C., Mess, J.N., Lévesque, C., Fobert, P.R., Dangl, J.L., and Brisson, N. (2004). A "Whirly" transcription factor is required for salicylic acid-dependent disease resistance in *Arabidopsis*. *Dev. Cell* 6: 229–240.

Drouin, G., Daoud, H., and Xia, J. (2008). Relative rates of synonymous substitutions in the mitochondrial, chloroplast and nuclear genomes of seed plants. *Mol. Phylogenet. Evol.* 49: 827–831.

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

Emsley, P. and Cowtan, K. (2004). Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* 60: 2126–2132.

Fauron, C., Allen, J., Clifton, S., and Newton, K. (2004). Plant mitochondrial genomes. In *Molecular Biology and Biotechnology of Plant Organelles*, H. Daniell and C. Chase, eds (Dordrecht, The Netherlands: Springer), pp. 151–177.

Ferreira, P., Hemerly, A., de Almeida Engler, J., Bergounioux, C., Burssens, S., Van Montagu, M., Engler, G., and Inzé, D. (1994). Three discrete classes of *Arabidopsis* cyclins are expressed during different intervals of the cell cycle. *Proc. Natl. Acad. Sci. USA* 91: 11313–11317.

Fisher, L.M., Mizuuchi, K., O'Dea, M.H., Ohmori, H., and Gellert, M. (1981). Site-specific interaction of DNA gyrase with DNA. *Proc. Natl. Acad. Sci. USA* 78: 4165–4169.

Gellert, M., Mizuuchi, K., O'Dea, M.H., Itoh, T., and Tomizawa, J.I. (1977). Nalidixic acid resistance: A second genetic character involved in DNA gyrase activity. *Proc. Natl. Acad. Sci. USA* 74: 4772–4776.

Gellert, M., O'Dea, M.H., Itoh, T., and Tomizawa, J. (1976). Novobiocin and coumermycin inhibit DNA supercoiling catalyzed by DNA gyrase. *Proc. Natl. Acad. Sci. USA* 73: 4474–4478.

Gouet, P., Robert, X., and Courcelle, E. (2003). ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res.* 31: 3320–3323.

Grabowski, E., Miao, Y., Mulisch, M., and Krupinska, K. (2008). Single-stranded DNA-binding protein Whirly1 in barley leaves is located in plastids and the nucleus of the same cell. *Plant Physiol.* 147: 1800–1804.

Hastings, P.J., Ira, G., and Lupski, J.R. (2009a). A microhomology-mediated break-induced replication model for the origin of human copy number variation. *PLoS Genet.* 5: e1000327.

Hastings, P.J., Lupski, J.R., Rosenberg, S.M., and Ira, G. (2009b). Mechanisms of change in gene copy number. *Nat. Rev. Genet.* 10: 551–564.

Horvath, M.P. (2008). Single-stranded nucleic acid-binding proteins. In *Protein-Nucleic Acid Interactions: Structural Biology*, P.A. Rice and C.C. Correll, eds (London: Royal Society of Chemistry), pp. 91–128.

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.[Abstract/Free Full Text]

Kanno, A., Watanabe, N., Nakamura, I., and Hirai, A. (1993). Variations in chloroplast DNA from rice (*Oryza sativa*): Differences between deletions mediated by short direct-repeat sequences within a single-species. *Theor. Appl. Genet.* 86: 579–584.

Knoop, V. (2004). The mitochondrial DNA of land plants: Peculiarities in phylogenetic perspective. *Curr. Genet.* 46: 123–139.

Kohl, S. and Bock, R. (2009). Transposition of a bacterial insertion sequence in chloroplasts. *Plant J.* 58: 423–436.

Krause, K., Kilbiński, I., Mulisch, M., Rödiger, A., Schäfer, A., and Krupinska, K. (2005). DNA-binding proteins of the Whirly family in *Arabidopsis thaliana* are targeted to the organelles. *FEBS Lett.* 579: 3707–3712.

López, E. and Blázquez, J. (2009). Effect of subinhibitory concentrations of antibiotics on intrachromosomal homologous recombination in *Escherichia coli*. *Antimicrob. Agents Chemother.* 53: 3411–3415.

Lu, D. and Keck, J.L. (2008). Structural basis of *Escherichia coli* single-stranded DNA-binding protein stimulation of exonuclease I. *Proc. Natl. Acad. Sci. USA* 105: 9169–9174.

Lupski, J.R. (2009). Genomic disorders ten years on. *Genome Med.* 1: 42.

Mackenzie, S.A. (2005). The mitochondrial genome of higher plants: A target for natural adaptation. In *Plant Diversity and Evolution: Genotypic and Phenotypic Variation in Higher Plants*, R. Henry, ed (London: CABI Publishers), pp. 69–80.

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

Maréchal, A., Parent, J.S., Sabar, M., Véronneau-Lafortune, F., Abou-Rached, C., and Brisson, N. (2008). Overexpression of mtDNA-associated AtWhy2 compromises mitochondrial function. *BMC Plant Biol.* 8: 42.

Maréchal, A., Parent, J.S., Véronneau-Lafortune, F., Joyeux, A., Lang, B.F., and Brisson, N. (2009). Whirly proteins maintain plastid genome stability in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* 106: 14693–14698.

Marvo, S.L., King, S.R., and Jaskunas, S.R. (1983). Role of short regions of homology in intermolecular illegitimate recombination events. *Proc. Natl. Acad. Sci. USA* 80: 2452–2456.

Maul, J.E., Lilly, J.W., Cui, L., dePamphilis, C.W., Miller, W., Harris, E.H., and Stern, D.B. (2002). The *Chlamydomonas reinhardtii* plastid chromosome: Islands of genes in a sea of repeats. *Plant Cell* 14 2659–2679.

McVey, M. and Lee, S.E. (2008). MMEJ repair of double-strand breaks (director's cut): Deleted sequences and alternative endings. *Trends Genet.* 24: 529–538.

Menke, M., Chen, I., Angelis, K.J., and Schubert, I. (2001). DNA damage and repair in *Arabidopsis thaliana* as measured by the comet assay after treatment with different classes of genotoxins. *Mutat. Res.* 493: 87–93.

Moeykens, C.A., Mackenzie, S.A., and Shoemaker, R.C. (1995). Mitochondrial genome diversity in soybean: Repeats and rearrangements. *Plant Mol. Biol.* 29: 245–254.

Morel, F., Renoux, M., Lachaume, P., and Alziari, S. (2008). Bleomycin-induced double-strand breaks in mitochondrial DNA of *Drosophila* cells are repaired. *Mutat. Res.* 637: 111–117.

Newton, K.J., Gabay-Laughnan, S., and Paepe, R.D. (2004). Mitochondrial mutation in plants. In *Plant Mitochondria*, D.A. Day, A.H. Millar, and J. Whelan, eds (Dordrecht, The Netherlands: Kluwer Academic Publishers), pp. 121–141.

Notredame, C., Higgins, D.G., and Heringa, J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* 302: 205–217.

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.

Odom, O.W., Baek, K.H., Dani, R.N., and Herrin, D.L. (2008). *Chlamydomonas* chloroplasts can use short dispersed repeats and multiple pathways to repair a double-strand break in the genome. *Plant J.* 53: 842–853.

Ogihara, Y., Terachi, T., and Sasakuma, T. (1988). Intramolecular recombination of chloroplast genome mediated by short direct-repeat sequences in wheat species. *Proc. Natl. Acad. Sci. USA* 85: 8573–8577.

Persky, N.S. and Lovett, S.T. (2008). Mechanisms of recombination: Lessons from *E. coli*. *Crit. Rev. Biochem. Mol. Biol.* 43: 347–370.

Ponder, R.G., Fonville, N.C., and Rosenberg, S.M. (2005). A switch from high-fidelity to error-prone DNA double-strand break repair underlies stress-induced mutation. *Mol. Cell* 19: 791–804.

Prikryl, J., Watkins, K.P., Friso, G., Wijk, K.J., and Barkan, A. (2008). A member of the Whirly family is a multifunctional RNA- and DNA-binding protein that is essential for chloroplast biogenesis. *Nucleic Acids Res.* 36: 5152–5165.

Raubeson, L.A. and Jansen, R.K. (2005). Chloroplast genomes of plants. In *Plant Diversity and Evolution: Genotypic and Phenotypic Variation in Higher Plants*, R. Henry, ed (London: CABI Publishers), pp. 45–68.

Ravi, V., Khurana, J.P., Takagi, A.K., and Khurana, P. (2008). An update on chloroplast genomes. *Plant Syst. Evol.* 271: 101–122.

Roa, H., Lang, J., Culligan, K.M., Keller, M., Holec, S., Cognat, V., Montane, M.H., Houlné, G., and Chabouté, M.E. (2009). Ribonucleotide reductase regulation in response to genotoxic stress in *Arabidopsis*. *Plant Physiol.* 151: 461–471.

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.

Saha, S., Nicholson, A., and Kapler, G.M. (2001). Cloning and biochemical analysis of the tetrahymena origin binding protein TIF1: Competitive DNA binding in vitro and in vivo to critical rDNA replication determinants. *J. Biol. Chem.* 276: 45417–45426.

Sato, S., Nakamura, Y., Kaneko, T., Asamizu, E., and Tabata, S. (1999). Complete structure of the chloroplast genome of *Arabidopsis thaliana*. *DNA Res.* 6: 283–290.

Schumacher, M.A., Karamooz, E., Zíková, A., Trantírek, L., and Lukes, J. (2006). Crystal structures of *T. brucei* MRP1/MRP2 guide-RNA binding complex reveal RNA matchmaking mechanism. *Cell* 126: 701–711.

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.

Shen, C.C., Wertelecki, W., Driggers, W.J., LeDoux, S.P., and Wilson, G.L. (1995). Repair of mitochondrial DNA damage induced by bleomycin in human cells. *Mutat. Res.* 337: 19–23.

Shereda, R.D., Kozlov, A.G., Lohman, T.M., Cox, M.M., and Keck, J.L. (2008). SSB as an organizer/mobilizer of genome maintenance complexes. *Crit. Rev. Biochem. Mol. Biol.* 43: 289–318.

Shereda, R.D., Reiter, N.J., Butcher, S.E., and Keck, J.L. (2009). Identification of the SSB binding site on *E. coli* RecQ reveals a conserved surface for binding SSB's C terminus. *J. Mol. Biol.* 386: 612–625.

Small, I.D., Isaac, P.G., and 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.

Snyder, M. and Drlica, K. (1979). DNA gyrase on the bacterial chromosome: DNA cleavage induced by oxolinic acid. *J. Mol. Biol.* 131: 287–302.

Sugino, A., Peebles, C.L., Kreuzer, K.N., and Cozzarelli, N.R. (1977). Mechanism of action of nalidixic acid: Purification of *Escherichia coli* nalA gene product and its

relationship to DNA gyrase and a novel nicking-closing enzyme. *Proc. Natl. Acad. Sci. USA* 74: 4767–4771.

Tamura, K., Adachi, Y., Chiba, K., Oguchi, K., and Takahashi, H. (2002). Identification of Ku70 and Ku80 homologues in *Arabidopsis thaliana*: Evidence for a role in the repair of DNA double-strand breaks. *Plant J.* 29: 771–781.

Unsel, M., Marienfeld, J.R., Brandt, P., and Brennicke, A. (1997). The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. *Nat. Genet.* 15: 57–61.

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.

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.

Wang, H., Perrault, A.R., Takeda, Y., Qin, W., and Iliakis, G. (2003). Biochemical evidence for Ku-independent backup pathways of NHEJ. *Nucleic Acids Res.* 31: 5377–5388.

Weigel, D. and Glazebrook, J. (2002). *Arabidopsis: A Laboratory Manual*. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).

Woelfle, M.A., Thompson, R.J., and Mosig, G. (1993). Roles of novobiocin-sensitive topoisomerases in chloroplast DNA replication in *Chlamydomonas reinhardtii*. *Nucleic Acids Res.* 21: 4231–4238.

Wolfe, K.H., Li, W.H., and Sharp, P.M. (1987). Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc. Natl. Acad. Sci. USA* 84: 9054–9058.

Woloszynska, M. (2010). Heteroplasmy and stoichiometric complexity of plant mitochondrial genomes—Though this be madness, yet there's method in't. *J. Exp. Bot.* 61: 657–671.

Xiong, J.Y., Lai, C.X., Qu, Z., Yang, X.Y., Qin, X.H., and Liu, G.Q. (2009). Recruitment of AtWHY1 and AtWHY3 by a distal element upstream of the kinesin gene AtKP1 to mediate transcriptional repression. *Plant Mol. Biol.* 71: 437–449.

Yakisich, J.S., Sandoval, P.Y., Morrison, T.L., and Kapler, G.M. (2006). TIF1 activates the intra-S-phase checkpoint response in the diploid micronucleus and amitotic polyploid macronucleus of *Tetrahymena*. *Mol. Biol. Cell* 17: 5185–5197.

Ye, J. and Sayre, R.T. (1990). Reduction of chloroplast DNA content in *Solanum nigrum* suspension cells by treatment with chloroplast DNA synthesis inhibitors. *Plant Physiol.* 94: 1477–1483.

Yoo, H.H., Kwon, C., Lee, M.M., and Chung, I.K. (2007). Single-stranded DNA binding factor AtWHY1 modulates telomere length homeostasis in *Arabidopsis*. *Plant J.* 49: 442–451.

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.

3.9 Figure Legends

Figure 19. PCR Strategy to Detect DNA Rearrangements.

Outward-facing PCR primers are used to monitor DNA duplication/circularization events, whereas inward-facing PCR primers detect deletions. In both cases, PCR amplification occurs only if a DNA rearrangement brings together the annealing sites of the primers. Gray areas represent repeated sequences.

Figure 20. DNA Rearrangements Accumulate in Plastids Following Treatment with the Gyrase Inhibitor Ciprofloxacin.

(A) and (E) Phenotypic effects of various concentrations of ciprofloxacin (A) or novobiocin (E) on wild-type (WT) and *why1why3 Arabidopsis* plants. Plants were grown for 3 weeks

on solid media containing the indicated concentrations of ciprofloxacin (CIP) or novobiocin (NOV).

(B) and (F) Histograms showing the average \pm SD of plants with etiolated/variegated leaves at each ciprofloxacin (B) or novobiocin (F) concentration for wild-type and *why1why3* genotypes. At least three independent experiments were done. Plants with partially or fully white first true leaves were scored as etiolated/variegated. ND, etiolated/variegated leaves could not be counted at 0.75 μ M ciprofloxacin and 100 μ M novobiocin because they were very small or absent.

(C) and (G) Electrophoretic analysis of representative PCR performed with 13 outward- or inward-facing plastid genome-directed PCR primers on total leaf DNA of wild-type and *why1why3* plants treated with ciprofloxacin (C) or with 10 outward- or inward-facing plastid genome-directed PCR primers on total leaf DNA of wild-type and *why1why3* plants treated with novobiocin (G). Low cycle amplification of the YCF2 plastid gene was used as a loading control. The oligonucleotides used for each PCR are indicated. Individual bands were cut from the gel, cloned, and sequenced. DNA rearrangements are listed in Tableau X.

(D) and (H) Histograms showing the number of PCR products in wild-type and *why1why3* plants as a function of ciprofloxacin (D) or novobiocin (H) concentration.

Figure 21. DNA Rearrangements Accumulate in Mitochondria Following Treatment with the Gyrase Inhibitor Ciprofloxacin.

(A) and (E) Phenotypic effects of various concentrations of ciprofloxacin (A) or novobiocin (E) on wild-type (WT) and *why2-1 Arabidopsis* plants. Plants were grown for 3 weeks on solid media containing the indicated concentrations of ciprofloxacin (CIP) or novobiocin (NOV).

(B) and (F) Histograms showing the average \pm SD of plants with etiolated/variegated leaves at each ciprofloxacin (B) or novobiocin (F) concentration for wild-type and *why2-1* genotypes. At least three independent experiments were done. Plants with partially or fully white first true leaves were scored as etiolated/variegated. ND, etiolated/variegated

leaves could not be counted at 0.75 μM ciprofloxacin and 100 μM novobiocin because they were very small or absent.

(C) and (G) Electrophoretic analysis of representative PCR performed with 20 outward- or inward-facing mitochondrial genome-directed PCR primers on total leaf DNA of wild-type and *why2-1* plants treated with ciprofloxacin (C) or with 10 outward- or inward-facing mitochondrial genome-directed PCR primers on total leaf DNA of wild-type and *why2-1* plants treated with novobiocin (G). Low cycle amplification of the COX1 mitochondrial gene was used as a loading control. The oligonucleotides used for each PCR are indicated. Individual bands were cut from the gel, cloned, and sequenced. DNA rearrangements are listed in Tableau X.

(D) and (H) Histograms showing the number of PCR products in wild-type and *why2-1* plants as a function of ciprofloxacin (D) or novobiocin (H) concentration.

Figure 22. Crystal Structures of St-WHY2 in the Free Form and Bound to ERE32 at 2.2-Å Resolution.

(A) Schematic representation of St-WHY2. Filled boxes indicate the position of the mitochondria transit peptide (mTP), the Whirly domain, and the acidic/aromatic C-terminal tail (CT). A green dotted line represents the construct used for structure determination.

(B) Overall view of St-WHY2 in the free form in cartoon representation. The tetramer was generated by applying crystallographic fourfold symmetry. Protomers are colored in yellow, orange, pink, and green.

(C) Surface representation of the protein moiety in the St-WHY2-ERE32 complex. The tetramer was generated by applying the crystallographic symmetry along the fourfold axis. Difference electron density was calculated from a $F_o - F_c$ simulated annealing omit map encompassing DNA, contoured at 3.0, and colored in green. The density was carved at 8 Å around the protein model.

Figure 23. Mechanism of ssDNA Binding of Whirly Proteins.

(A) Protein–DNA interactions in the St-WHY2-ERE32 complex. The DNA and the DNA-interacting residues are in stick representation with carbon atoms colored in yellow and in gray, respectively. Nucleotides and protein residues are labeled. Asterisks, residues that contact DNA through their main-chain. +, residues that contact the DNA through both its main chain and side chain. Hydrogen bonds are represented as yellow dashed lines. A water molecule is represented as a red sphere.

(B) Sequence alignment of the Whirly domain of Whirly proteins from *Arabidopsis* (At-WHY1, At-WHY2, and At-WHY3) and *S. tuberosum* (St-WHY1 and St-WHY2) with schematic secondary structural elements from St-WHY2. Secondary structure conformations are denoted at the top of the sequence alignment. α , α -helix; β , β -strand; α_310 , α_310 helix. Similarity above 70% is depicted in yellow, whereas perfect conservation is depicted in red. Blue stars underneath the sequence alignment indicate residues of St-WHY2 that interact with ssDNA. The alignments were made using T-coffee (Notredame et al., 2000) and the figure prepared using ESPript (Gouet et al., 2003).

Figure 24. WHY2 Binds ssDNA with Limited Sequence Specificity.

(A) Representative EMSA results showing the binding of St-WHY2 to four different ssDNA sequences. Increasing amounts of WHY2 were incubated with target oligonucleotides ERE32, dT32, rcERE32, or cERE32 and the complexes resolved on a 10% (w/v) polyacrylamide gel. The sequences of these oligonucleotides can be found in Tableau XIV.

(B) Crystal structures of four different ssDNA sequences bound by St-WHY2. DNA molecules are presented as stick models with carbon atoms colored in yellow. Fo-Fc simulated annealing omit maps encompassing the entire DNA are contoured at 2.5 (colored in gray) or at 5 (colored in green). DNA molecules are presented in the same order as in (A).

(C) and (D) Interactions between St-WHY2 and the edges of T3 (left panel) and A3 (right panel) (C) or the edges of T7 (left panel) and A7 (right panel) (D) in the WHY2-ERE32 and WHY2-rcERE32 structures, respectively. The representation and the orientation of the molecule are similar, thus revealing that compensating interactions enable WHY2 to bind

DNA nucleobases that differ in size and in functional groups at these positions. A red sphere corresponds to a water molecule.

Figure 25. WHY2 Binds Single-Stranded Overhangs, Destabilizes dsDNA, and Protects ssDNA against Nuclease Degradation.

(A) Representative EMSA results showing the binding of St-WHY2 to DNA duplexes with or without single-stranded dT8 or dT16 overhangs. Fifty nanomolar of WHY2 were incubated with target radiolabeled oligonucleotides and the complexes resolved on a 10% (w/v) polyacrylamide gel. Diagrams at the bottom of the gels illustrate the DNA used in the assay. An asterisk indicates the strand that is radiolabeled.

(B) WHY2 destabilizes a DNA duplex. A DNA duplex with one strand radiolabeled was incubated with increasing amounts of St-WHY2 or BSA. After protein denaturation, DNA was resolved on a 7.5% acrylamide gel.

(C) WHY2 protects DNA against mung bean nuclease degradation. Phage M13mp18 ssDNA either alone or prebound with St-WHY2 at a 1:10 protein/nucleotide ratio was incubated with mung bean nuclease for the indicated amount of time. After protein denaturation, DNA was resolved on an agarose gel. Black/white inverted images are shown. M represents the molecular weight markers.

(D) WHY2 protects DNA against the exonuclease activity of T4 DNA Polymerase. Radiolabeled ssDNA or dsDNA with a 16-nucleotide 3'-overhang was complexed with St-WHY2 and then incubated with T4 DNA polymerase for the indicated amount of time. After protein denaturation, DNA was resolved on a 10% acrylamide gel. ssDNA, YG16_3; 3'-dsDNA, YG16_3-YC duplex; dsDNA, YG-YC duplex.

Figure 26. Model for the Repair of Organellar Double-Strand Breaks in the Absence or Presence of Whirly Proteins.

Upon formation of a DSB (1), the 5' end of the broken DNA molecule is resected from the break, exposing a 3' tail (2). At this step, the break can be repaired through homologous recombination in a Whirly-independent manner. Alternatively, if the Whirlies are absent or the repair machinery is overloaded due to numerous DSBs, the 3' tail can anneal to any

exposed ssDNA through microhomologies (3a). A D-loop forms and DNA polymerization proceeds from the microhomology junction (4a). A replication fork is established and lagging strand synthesis initiates while leading strand synthesis continues (5a). DNA synthesis continues until the end of the chromosome is reached (6a). Alternatively, if the Whirlies are present and the DSB level is low (3b), Whirlies could bind and protect ssDNA (either the 3' tail and/or any exposed ssDNA), thereby promoting homologous recombination and accurate DNA repair. Arrowheads represent 3' ends; a box symbolizes the microhomology between broken and unbroken DNA molecules; dashed arrows in tandem represent lagging strand synthesis.

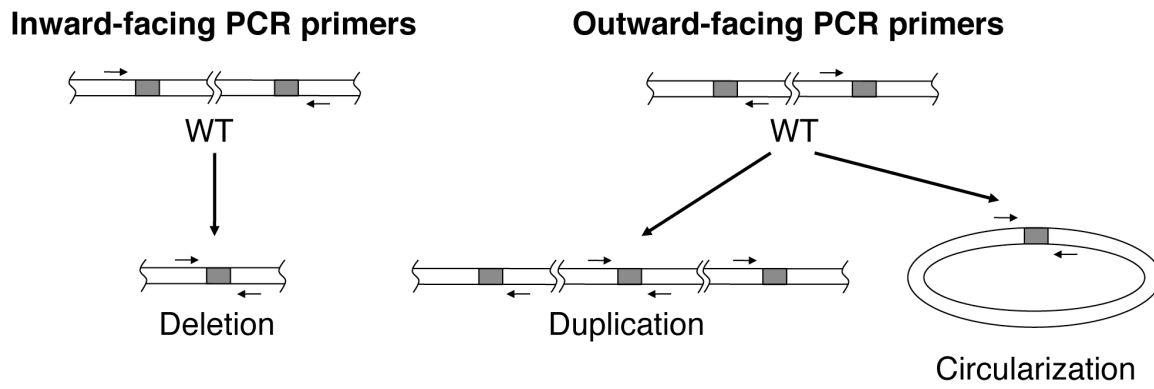


Figure 19. PCR Strategy to Detect DNA Rearrangements.

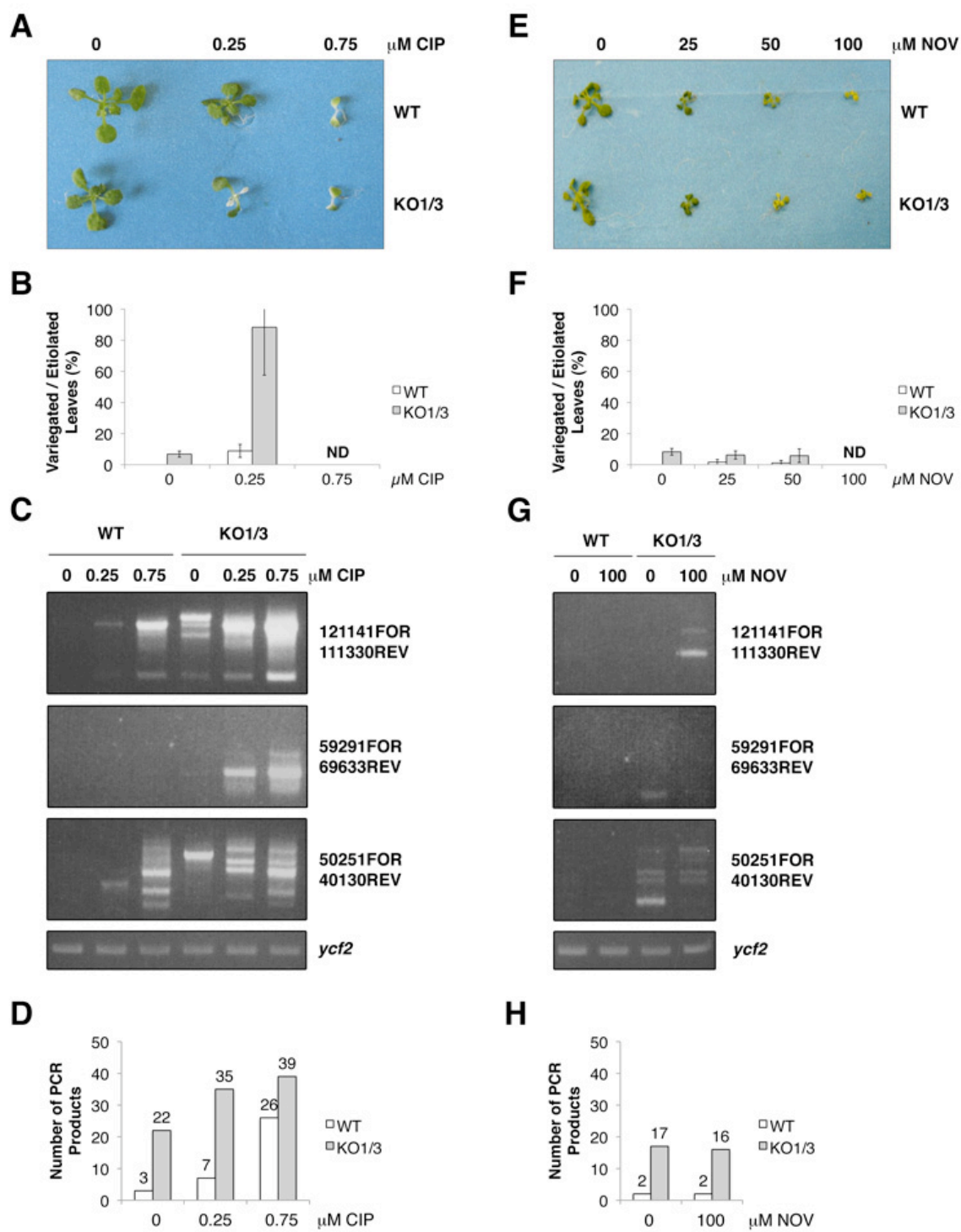


Figure 20. DNA Rearrangements Accumulate in Plastids Following Treatment with the Gyrase Inhibitor Ciprofloxacin.

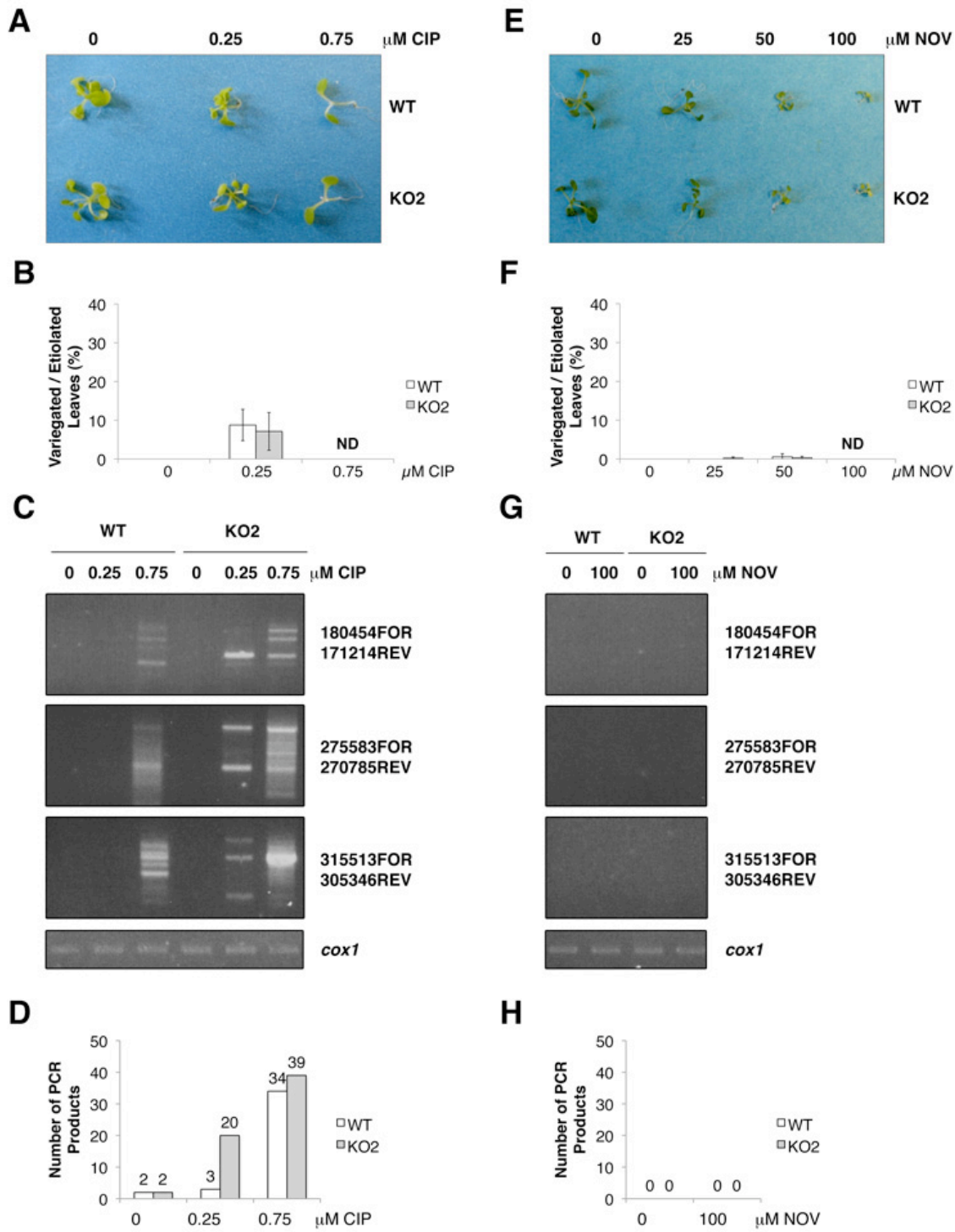


Figure 21. DNA Rearrangements Accumulate in Mitochondria Following Treatment with the Gyrase Inhibitor Ciprofloxacin.

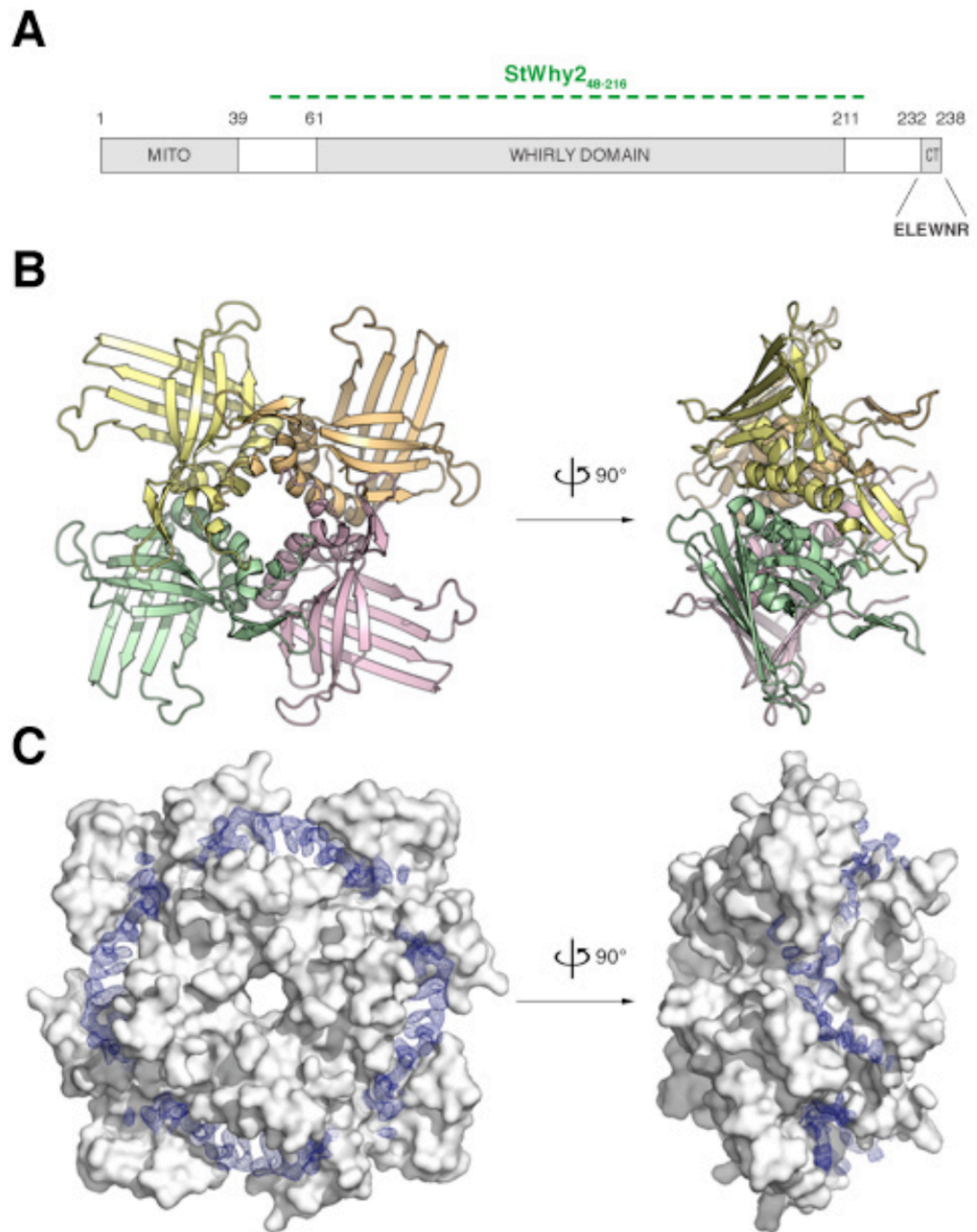
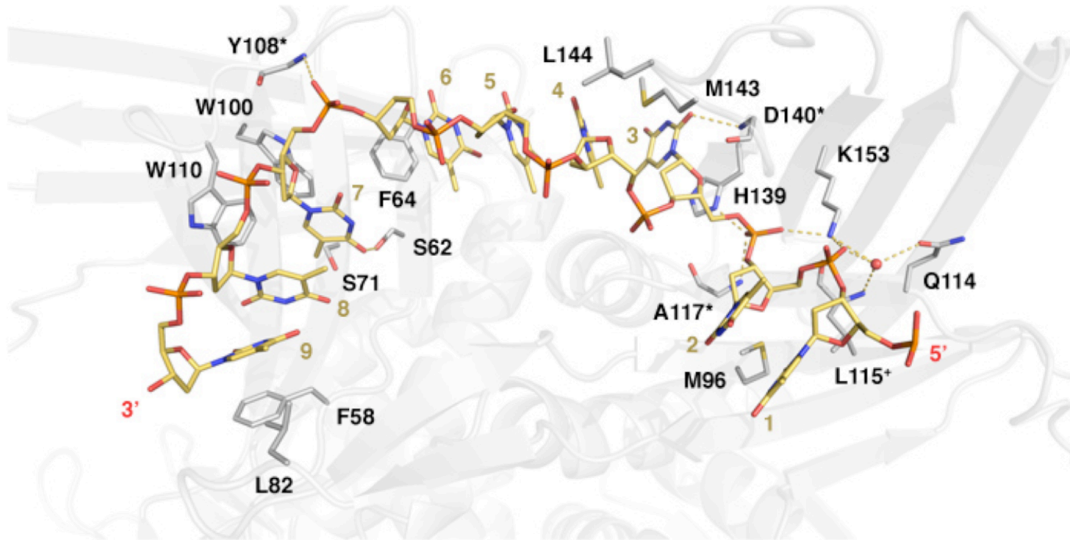


Figure 22. Crystal Structures of St-WHY2 in the Free Form and Bound to ERE32 at 2.2-Å Resolution.

A



B

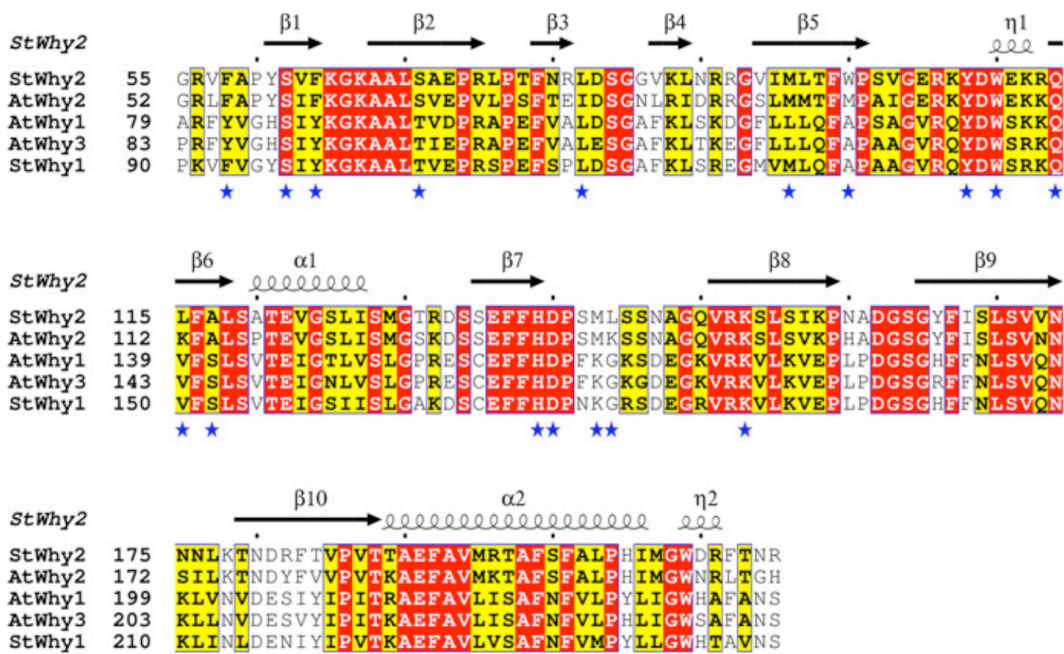


Figure 23. Mechanism of ssDNA Binding of Whirly Proteins.

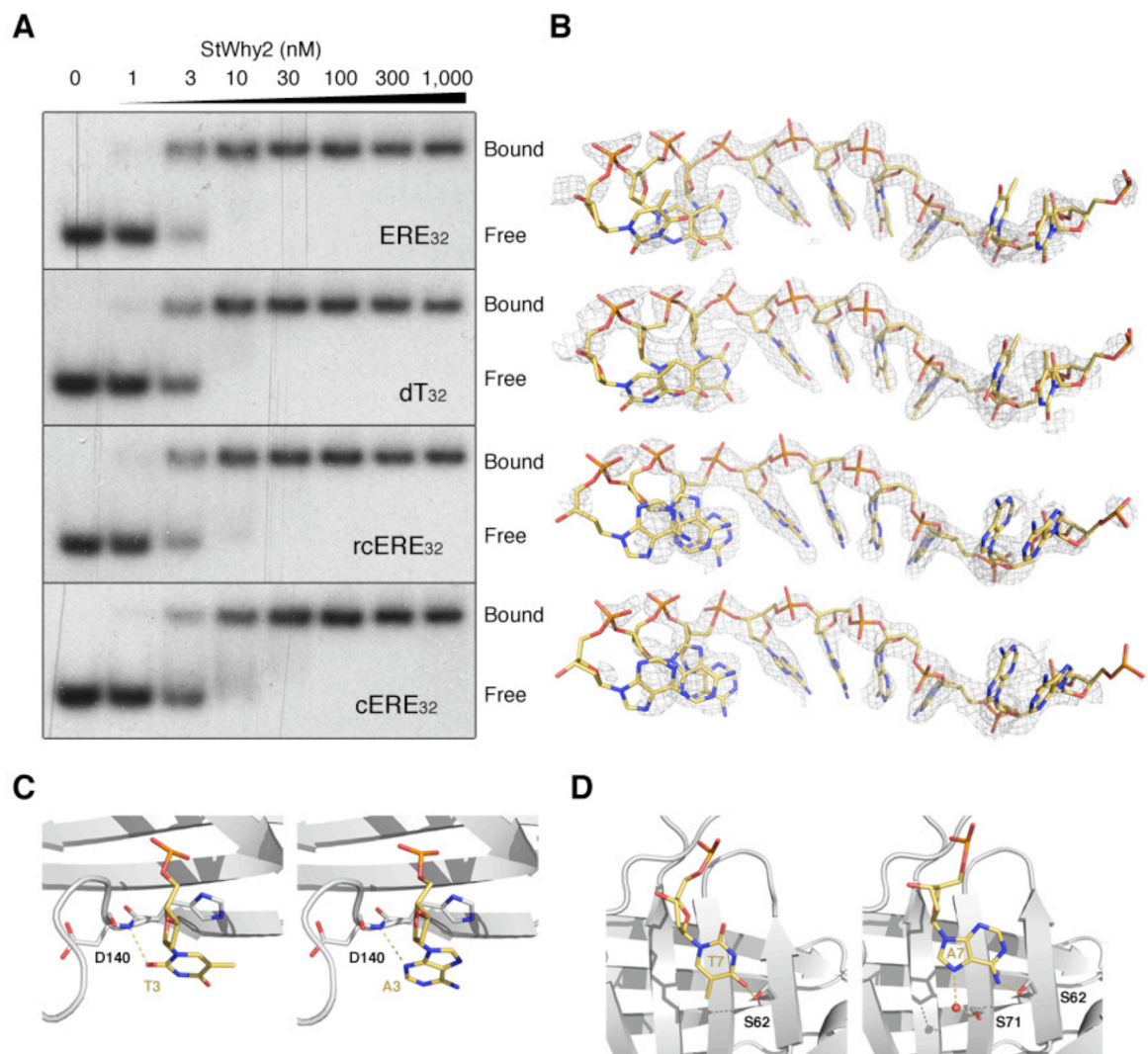


Figure 24. WHY2 Binds ssDNA with Limited Sequence Specificity.

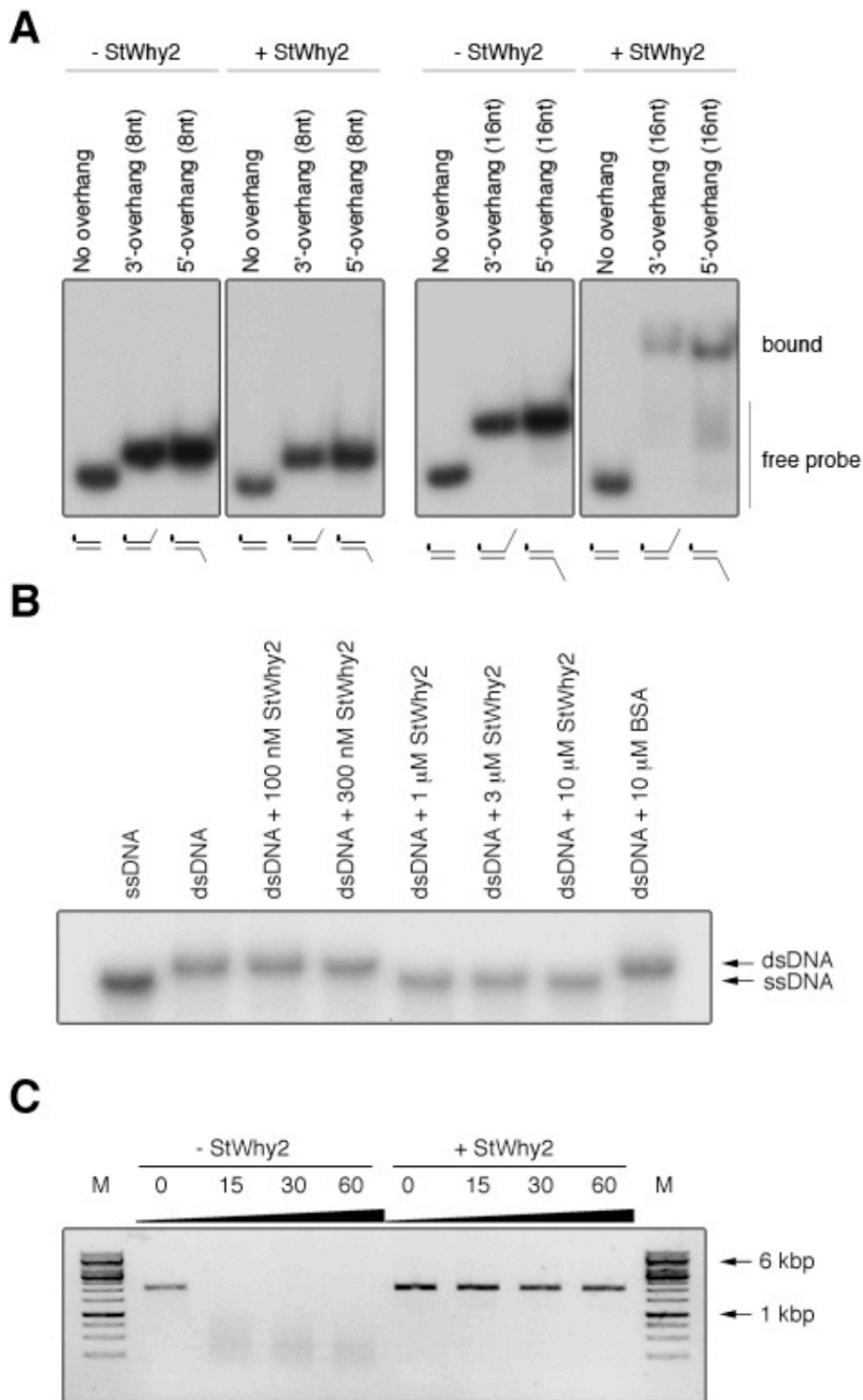


Figure 25. WHY2 Binds Single-Stranded Overhangs, Destabilizes dsDNA, and Protects ssDNA against Nuclease Degradation.



Figure 26. Model for the Repair of Organellar Double-Strand Breaks in the Absence or Presence of Whirly Proteins.

Tableau VII. Representative DNA Rearrangements Observed in *Arabidopsis* Organellar Genomes

Event	Plant Treatment ^a	Plant Genotype	Junction Endpoints ^b	Junction Sequences ^c	Length of Junctions (Mismatches)	Junction Types ^d
NHEJ-like DNA junction						
M3-2	CIP 0.75 μM	Col-0	55001 Recombinant 65518	cgaaggagc <u>cataccggaa</u> tgagccgctg <u>cataccggaa</u> tgagccgctg tactgtggca	0 bp	NHEJ-like
MHMR DNA junction						
M3-3	CIP 0.75 μM	<i>why2</i>	53421 Recombinant 65476	gctccgtcca <u>TGGAGCGTATG</u> <u>atcttccttg</u> <u>ctctaaagct</u> <u>TGGAGCGTATG</u> <u>atcttccttg</u> <u>ctctaaagct</u> <u>TGGAGCGTATG</u> gtaacggta	(0) 11 bp (0)	MHMR
MHMR DNA junction containing a mismatch in one parental strand						
M6-1	CIP 0.25 μM	<i>why2</i>	234779 Recombinant 246725	gatgaggatg <u>AAGTAAAGGAATT</u> <u>actctggttg</u> <u>acgaaagtga</u> <u>AAGTAAAGGAATT</u> <u>actctggttg</u> <u>acgaaagtga</u> <u>AAGTAAAGGGATT</u> gtctggcct	(0) 13 bp (1)	MHMR*
MHMR DNA junction containing a mismatch in both parental strands						
M7-7	CIP 0.75 μM	Col-0	304900 Recombinant 316104	tggaatgctc <u>TCGTTGTTAAAAAA</u> <u>ctacgagcgg</u> <u>caggggggat</u> <u>TCTTTGTCGATAAAA</u> <u>ctacgagcgg</u> <u>caggggggat</u> <u>TCTTTGTCGATAGAA</u> ggcatcgcg	(4) 15 bp (1)	MHMR*
Same MHMR DNA junction containing different mismatches						
M3-6	CIP 0.75 μM	<i>why2</i>	54234 Recombinant 65096	taggaatctt <u>AGGTTACCTCCAG</u> <u>ccttctgaag</u> <u>ccactaggtta</u> <u>AGGTTACCTCCAG</u> <u>ccttctgaag</u> <u>ccactaggtta</u> <u>AGGTTACTTCCAG</u> agcggagtat	(0) 13 bp (1)	MHMR*
M3-7	CIP 0.75 μM	<i>why2</i>	54234 Recombinant 65096	taggaatctt <u>AGGTTACCTCCAG</u> <u>ccttctgaag</u> <u>ccactaggtta</u> <u>AGGTTACTTCCAG</u> <u>ccttctgaag</u> <u>ccactaggtta</u> <u>AGGTTACTTCCAG</u> agcggagtat	(1) 13 bp (0)	MHMR*
MHMR DNA junctions with homopolymeric tracts						
M12-1	CIP 0.25 μM	Col-0	29609 Recombinant 80994	tttttcccct <u>CCCCCCCCCA</u> <u>aaaaaaaaaa</u> <u>attagacgtc</u> <u>CCCCCCCCCA</u> <u>aaaaaaaaaa</u> <u>attagacgtc</u> <u>CCCCCCCCCA</u> tgctacggtt	(0) 10 bp (0)	MHMR
P10-14	CIP 0.75 μM	<i>why1 why3</i>	59719 Recombinant 50843	tactttgaaa <u>ACAAAAAAAAA</u> <u>gaaagtcatt</u> <u>gagtggggat</u> <u>ACAAAAAAAAA</u> <u>gaaagtcatt</u> <u>gagtggggat</u> <u>ACAAAAAAAAA</u> aaagaaagta	(0) 10 bp (0)	MHMR
Molecules containing multiple MHMR						
P6-3a	CIP 0 μM	<i>why1 why3</i>	19786 Recombinant 110850	catatctcaa <u>TAGATTGGCGT-AATTTT</u> <u>ctaccgcatg</u> <u>caaaattttg</u> <u>TAGATTGGCGT-AATTTT</u> <u>ctaccgcatg</u> <u>caaaattttg</u> <u>TAGATTGGTATGAATTTT</u> gagaaatgca	(0) 18 bp (3)	MHMR*
P6-3b	CIP 0 μM	<i>why1 why3</i>	19867 Recombinant 122003	taattatggg <u>AGAAGTTGATCGTTAAAAATGAT</u> <u>aatttatgca</u> <u>gccagccctt</u> <u>AGAAGTTGATCGTTAAAAATGAT</u> <u>aatttatgca</u> <u>gccagccctt</u> <u>AGAAGGT-ATCGTTAAAA-GAT</u> catcaattcc	(0) 22 bp (3)	MHMR*

A complete list of DNA rearrangements can be found in Supplemental Data set 1 online.^a Treatment of plants with ciprofloxacin (CIP) or novobiocin (NOV) at the indicated concentrations.^b The junction endpoint is the position of the first nucleotide preceding the DNA junction. The accession numbers for the mitochondrial and plastid genomes are NC_001284 (147) and NC_000932 (146), respectively.^c Alignment of the sequenced DNA junction (center) with homologous sequences (top and bottom). Mismatches are underlined.^d Product types were classified according to junction length with NHEJ-like for junction of less than five nucleotides, MHMR for junctions of five or more nucleotides and MHMR* for junctions of five or more nucleotides containing at least one mismatch.

Tableau VIII. Data Collection and Refinement Statistics

Protein	WHY2 Free Form	WHY2 -ERE ₃₂	WHY2 -dT ₃₂	WHY2 -cERE ₃₂	WHY2 -rcERE ₃₂
Beamline/Wavelength	NSLS- X25/1.08 Å	NSLS- X29/1.08 Å	NSLS- X25/1.08 Å	APS-22ID/1.10 Å	APS-22ID/1.10 Å
Space group	F432	F432	F432	F432	F432
Cell dimensions :					
<i>a</i> (Å)	164.58	167.16	166.51	166.93	166.67
Resolution (Å)	50–2.20 (2.28–2.20)	50–2.20 (2.28–2.20)	50–2.65 (2.74–2.65)	50–2.70 (2.80–2.70)	50–2.35 (2.43–2.35)
Total Reflections	380467	382748	118110	178903	253462
Unique Reflections	10208	10665	6188	5865	8716
R _{sym} (%)	7.6 (92)	5.1 (56)	11.8 (87)	9.0 (68)	8.8 (75)
I / σI	16.9 (3.8)	23.1 (3.8)	9.1 (3.0)	12.7 (2.2)	16.5 (2.3)
Completeness (%)	100.0 (99.9)	99.8 (98.7)	100.0 (100.0)	99.6 (96.0)	99.5 (95.4)
Multiplicity	37.3 (19.5)	35.9 (9.9)	19.1 (12.7)	30.5 (5.9)	29.1 (6.5)
Refinement Statistics					
Resolution (Å)	50–2.20	50–2.20	50–2.65	50–2.70	50–2.35
Reflections (total/test) ^a	9873/532	9843/752	6012/603	5625/628	8437/900
R _{work} / R _{free} (%)	20.32/25.08	22.67/26.18	21.15/25.79	20.75/25.71	21.70/26.26
Number of atoms					
Protein	1267	1267	1267	1267	1267
DNA/Phosphate	10	180	180	189	189
Water	69	36	60	56	60
B-factors					
Protein	50.9	52.5	53.5	60.7	61.8
DNA/Phosphate	56.8	62.9	61.2	72.6	75.0
Water	53.5	46.1	52.8	58.4	62.0
RMSDs					
Bond lengths (Å)	0.013	0.008	0.006	0.006	0.010
Bond angles (°)	1.414	1.099	0.997	0.953	1.321
Ramachandran^b					
Favored (%)	96.8	97.5	95.0	97.5	95.6
Outliers (%)	0	0	0	0	0

Values in parentheses are for highest-resolution shell.

$R_{\text{sym}} = \sum_{hkl} \sum_i |I_{hkl,i} - \langle I_{hkl} \rangle| / \sum_{hkl,i} \langle I_{hkl} \rangle$, where $I_{hkl,i}$ is the intensity of an individual measurement of the reflection with Miller indices hkl and $\langle I_{hkl} \rangle$ is the mean intensity of that reflection.

$R_{\text{work}} = \sum_{hkl} | |F_o| - |F_c| | / \sum_{hkl} |F_o|$, where $|F_o|$ is the observed structure-factor amplitude and $|F_c|$ is the calculated structure-factor amplitude.

R_{free} is the R factor based on at least 500 test reflections that were excluded from the refinement.

^a Reflections with $F_o > 0$

^b MolProbity analysis (Davis et al., 2007)

Tableau IX. Analysis of the Binding of St-WHY2 to Four Different DNA Sequences by EMSA

DNA name	DNA sequence	$K_{d,app}$ (nM)	n^a
ERE ₃₂	TGTCATTTTGTTCATTTTGTTCATTTTGTCA	3.35 ± 0.95	5
dT ₃₂	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	4.45 ± 0.87	5
rcERE ₃₂	ACAGTAAAAACAGTAAAAACAGTAAAAACAGT	4.29 ± 0.68	3
cERE ₃₂	TGACAAAAATGACAAAAATGACAAAAATGACA	5.73 ± 1.33	5

^a: number of independent measurements (each independent measurement was done in duplicates)

3.10 Supplemental Data

Figure 27. Effect of Ciprofloxacin Treatment on *Arabidopsis* Plant Survival.

Graph showing the average percentage \pm SD of WT *Arabidopsis* plants that developed first true leaves as a function of ciprofloxacin concentration. A LD50 of 0.5 μ M ciprofloxacin was thus obtained for WT plants. More than 100 plants were examined for each ciprofloxacin concentration.

Figure 28. DNA Rearrangements Accumulate Primarily in Plastids of Plants Lacking Plastid-Targeted Whirlies and in the Mitochondria of Plants Lacking Mitochondria-Targeted Whirlies.

(A) and (C) Electrophoretic analysis of representative PCR performed with 8 outward- or inward-facing (A) mitochondria genome-directed and (C) plastid genome-directed PCR primers on total leaf DNA of WT, *why2-1* and *why1why3* plants treated with ciprofloxacin. Low cycle amplifications of the *cox1* mitochondrial gene or the *ycf2* plastid gene were used as loading controls. The oligonucleotides used for each PCR are indicated.

(B) and (D) Histograms showing the number of PCR products in the mitochondrial (B) and plastid (D) genome of WT, *why2-1* and *why1why3* plants as a function of ciprofloxacin concentration.

Figure 29. DNA Gel Blot Detection of DNA Rearrangements in Plastids Following Plant Treatment with Ciprofloxacin.

(A) Upper panel: Schematic representation of the microhomology-mediated DNA rearrangement occurring between two loci of the chloroplast genome of *Arabidopsis* (accession number NC_000932). The DNA rearrangement was isolated by PCR using the primers 4434REV and 15643FOR. The PCR product was used as a probe for the DNA gel blot presented in the middle and lower panels and is represented by a thick gray line. The HindIII fragments corresponding to the bands in the DNA gel blot in the middle panel are delimited by a "H". The black boxes designate the microhomologous DNA sequences that mediated the rearrangement.

Middle panel: DNA gel blots of total DNA from WT and *why1why3* plants treated with ciprofloxacin. The total DNA was digested with HindIII and hybridized with the probe described in the upper panel. The predicted sizes of the bands corresponding to the WT and rearranged DNA molecules are indicated. R designates the rearranged DNA molecules.

Lower panel: Detection of subgenome-length molecules in DNA gel blots of total undigested DNA from WT and *why1why3* plants treated with ciprofloxacin using the same probe described in the upper panel.

(B) Same as in (A) but using as a probe a PCR product amplified with primers 49386REV and 68817FOR.

Figure 30. DNA Gel Blot Detection of DNA Rearrangements in Mitochondria Following Plant Treatment with Ciprofloxacin.

(A) Upper panel: Schematic representation of the microhomology-mediated DNA rearrangement occurring between two loci of the mitochondria genome of *Arabidopsis* (accession number NC_001284). The DNA rearrangement was isolated by PCR using the primers 125781REV and 180454FOR. The PCR product was used as a probe for the DNA gel blot presented in the middle and lower panels and is represented by a thick gray line. The HindIII fragments corresponding to the bands in the DNA gel blot in the middle panel are delimited by a "H". The black boxes designate the microhomologous DNA sequences that mediated the rearrangement. The 2.3 and 2.6 kb bands both contain a 2 kb perfect repeat that could undergo DNA recombination. However, since a HindIII site interrupts the repeat, DNA recombination is not expected to generate new bands.

Middle panel: DNA gel blots of total DNA from WT and *why2-1* plants treated with ciprofloxacin. The total DNA was digested with HindIII and hybridized with the probe described in the upper panel. The predicted sizes of the bands corresponding to the WT and rearranged DNA molecules are indicated. R designates the rearranged DNA molecules.

Lower panel: Detection of subgenome-length molecules in DNA gel blots of total undigested DNA from WT and *why2-1* plants treated with ciprofloxacin using the same probe described in the upper panel.

(B) Same as in (A) but using as a probe a PCR product amplified with primers 187018REV and 349054FOR. The bands at 3.8 or 6.6 kb contain no repeats longer than 50 bp, suggesting that recombination cannot account for the apparition of new bands.

Figure 31. Analysis of Mitochondrial Recombination in Plants Treated with Ciprofloxacin.

DNA gel blots of total DNA from WT, *why2-1* and *why1why3* plants treated with ciprofloxacin. DNA gel blots were hybridized with probes encompassing mitochondrial repeats A, D and F. Total DNA was digested with HindIII or BamHI. Arrows indicate the expected parental forms (black arrow) and the recombination products (gray arrow). The predicted size of the expected bands are given. (R) designates a recombination product. In the case of the probe encompassing the repeat F, two recombination events can be detected upon BamHI digestion leading to the apparition of (R1) and (R2) recombination products. R1 recombination is mediated by the repeat F whereas R2 is mediated by a 206 bp repeat. The size of the repeats is given in parenthesis below each gel.

Figure 32. Analysis of Mitochondrial Recombination in Plants Treated with Novobiocin.

DNA gel blots of total DNA from WT, *why2-1* and *why1why3* plants treated with novobiocin. DNA gel blots were hybridized with probes encompassing mitochondrial repeats A, D and F. Total DNA was digested with HindIII. Arrows indicate the expected parental forms (black arrow) and the recombination products (gray arrow). The predicted sizes of the expected bands are given. (R) designates a recombination product. The size of the repeats is given in parenthesis below each gel.

Figure 33. Effects of Bleomycin on Plastid and Mitochondria DNA Rearrangements of Plants lacking Plastid- or Mitochondria-Targeted Whirlies.

(A) and (D) Phenotypic effects of various concentrations of bleomycin on (A) WT and *why1why3* or (D) WT and *why2-1 Arabidopsis* plants. Plants were grown for three weeks on solid media containing the indicated concentrations of bleomycin.

(B) and (E) Electrophoretic analysis of representative PCR performed with 8 outward- or inward-facing (B) plastid genome-directed or (E) mitochondria genome-directed PCR primers on total leaf DNA of WT, *why2-1* and *why1why3* plants treated with bleomycin.

Low cycle amplification of the *ycf2* plastid gene or the *cox1* mitochondrial gene was used as a loading control. The oligonucleotides used for each PCR are indicated.

(C) and (F) Histograms showing the number of PCR products in the (C) plastid or (F) mitochondrial genome of WT, *why2-1* and *why1why3* plants as a function of bleomycin concentration.

Figure 34. Effects of Hydroxyurea on Plastid and Mitochondria DNA Rearrangements of Plants lacking Plastid- or Mitochondria-Targeted Whirlies.

(A) and (D) Phenotypic effects of various concentrations of hydroxyurea (HU) on (A) WT and *why1why3* or (D) WT and *why2-1 Arabidopsis* plants. Plants were grown for three weeks on solid media containing the indicated concentrations of hydroxyurea.

(B) and (E) Electrophoretic analysis of representative PCR performed with 8 outward- or inward-facing (B) plastid genome-directed or (E) mitochondria genome-directed PCR primers on total leaf DNA of WT, *why2-1* and *why1why3* plants treated with hydroxyurea. Low cycle amplification of the *ycf2* plastid gene or the *cox1* mitochondrial gene was used as a loading control. The oligonucleotides used for each PCR are indicated.

(C) and (F) Histograms showing the number of PCR products in the (C) plastid or (F) mitochondrial genome of WT, *why2-1* and *why1why3* plants as a function of hydroxyurea concentration.

Figure 35. Analysis of Mitochondrial Recombination in WT Plants Treated with Bleomycin or Hydroxyurea.

DNA gel blots of total DNA from WT plants treated with ciprofloxacin, bleomycin or hydroxyurea. DNA gel blots were hybridized with probes encompassing mitochondrial repeats A, D and F. Total DNA was digested with HindIII. Arrows indicate the expected parental forms (black arrow) and the recombination products (gray arrow). The predicted sizes of the expected bands are given. (R) designates a recombination product. The size of the repeats is given in parenthesis bellow each gel.

Figure 36. Single-Stranded DNA is Stabilized on the Edges of the β - Sheets of Whirly Proteins.

Crystal structure of St-WHY2 bound to ERE32. St-WHY2 is in cartoon representation while DNA molecules are presented as stick models with carbon atoms colored in yellow. An Fo-Fc simulated annealing omit map encompassing the entire DNA is contoured at 2.5σ and colored in green. A yellow dashed line indicates the possible path of ssDNA on top of the β -sheets.

Figure 37. Modeling of a St-WHY2-RNA Complex.

The crystal structure of the St-WHY2-ERE32 complex was used as a template to model RNA. The modeling was performed using CNS (Brünger et al., 1998) and Pymol (DeLano, 2002). The RNA is in stick representation with carbon atoms colored in yellow. The protein is in cartoon representation. Nucleotides residues are labeled.

Figure 38. Molecular Characterization of *why2* Lines.

(A) Physical map of the At-*why2* (AT1G71260) gene. The positions of the T-DNA insertions in the *why2-1* and *why2-2* lines are indicated.

(B) WHY2 protein is undetectable in leaf protein extracts of *why2-1* and *why2-2* lines by protein gel blotting with an α -AtWhy2 antibody. A section of the blot stained with Ponceau red is presented below as a loading control.

(C) Electrophoretic analysis of representative PCR performed with 8 outward- or inward-facing mitochondria genome-directed PCR primers on total leaf DNA of WT, *why2-1* and *why2-2* plants treated with ciprofloxacin. Low cycle amplification of the *cox1* mitochondrial gene was used as a loading control. The oligonucleotides used for each PCR are indicated.

(D) Histogram showing the number of PCR products in WT, *why2-1* and *why2-2* plants as a function of ciprofloxacin concentration.

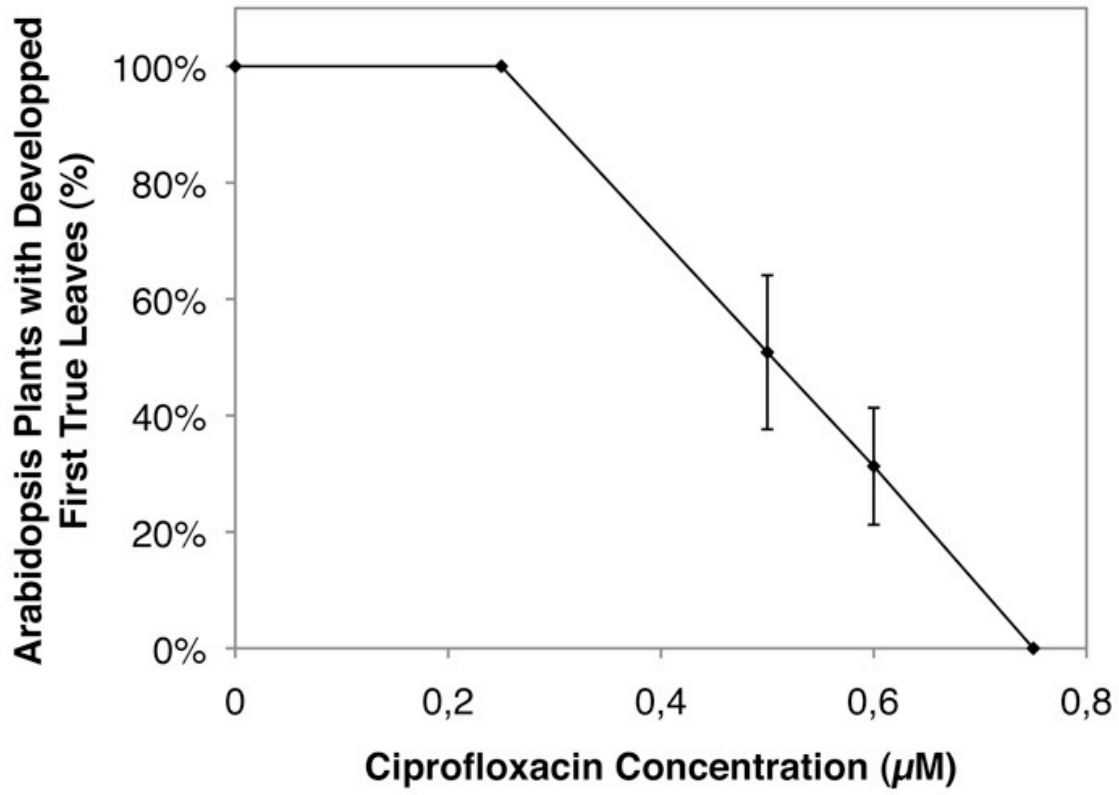


Figure 27. Effect of Ciprofloxacin Treatment on *Arabidopsis* Plant Survival.

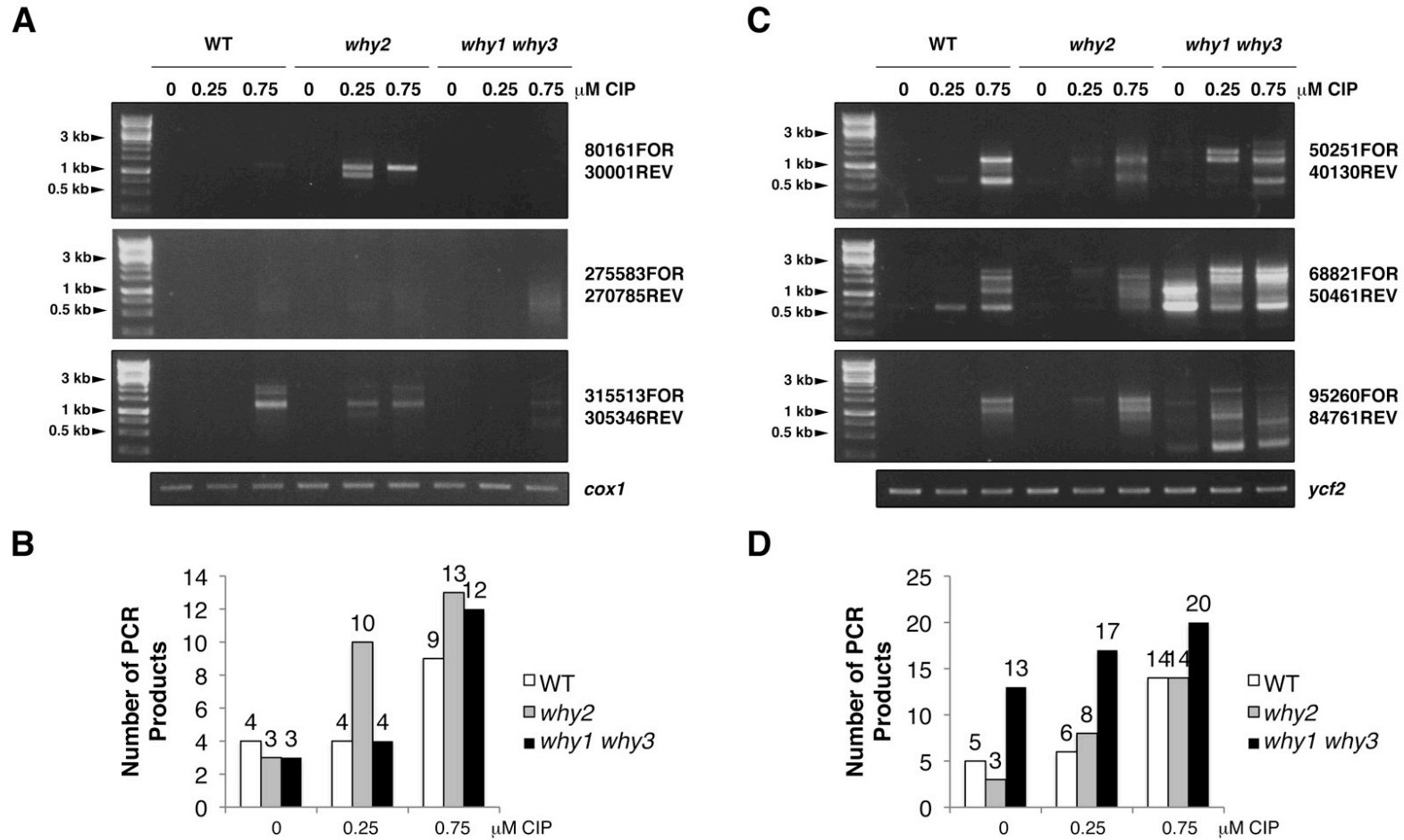


Figure 28. DNA Rearrangements Accumulate Primarily in Plastids of Plants Lacking Plastid-Targeted Whirlies and in the Mitochondria of Plants Lacking Mitochondria-Targeted Whirlies.

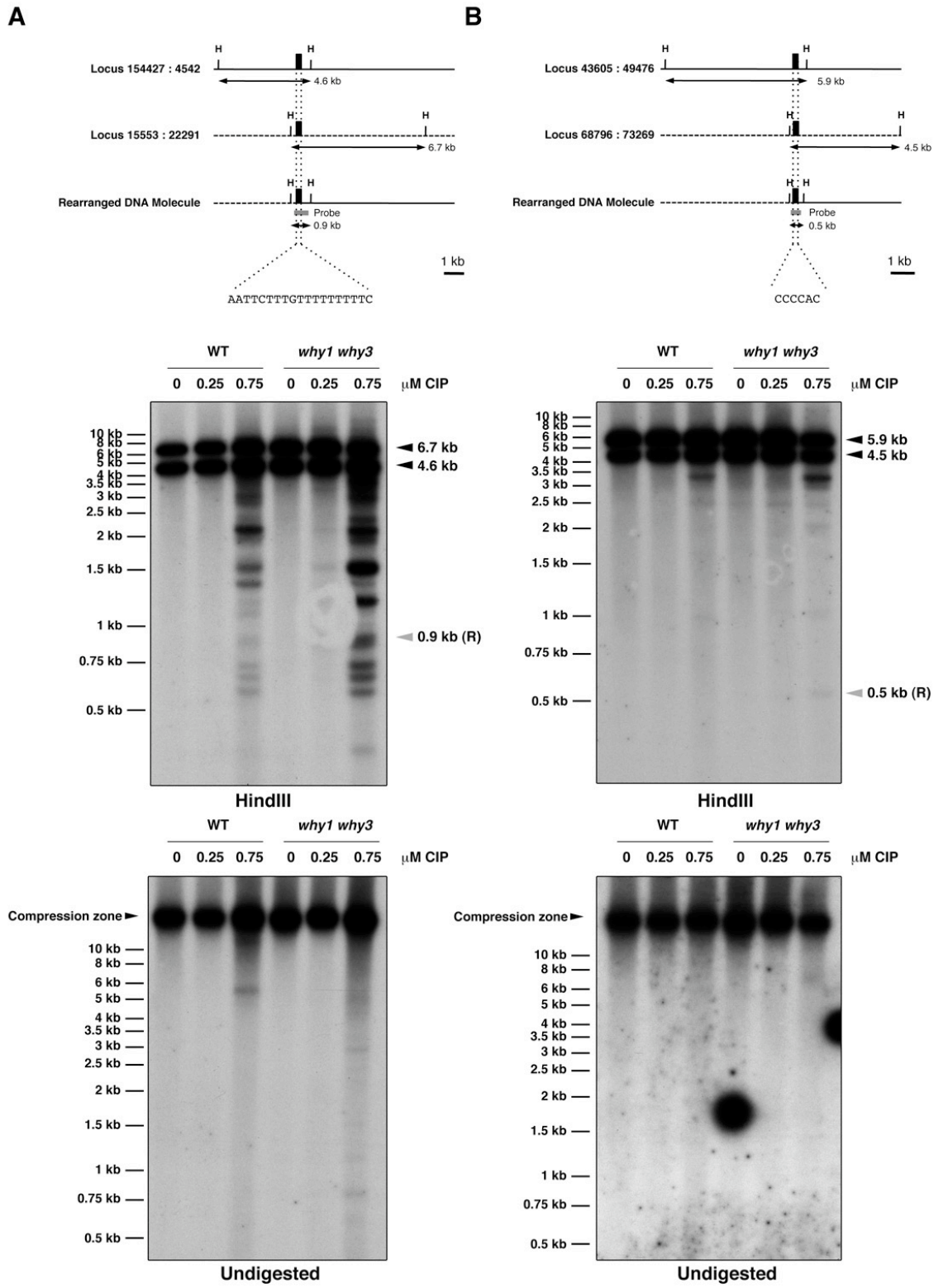


Figure 29. DNA Gel Blot Detection of DNA Rearrangements in Plastids Following Plant Treatment with Ciprofloxacin.

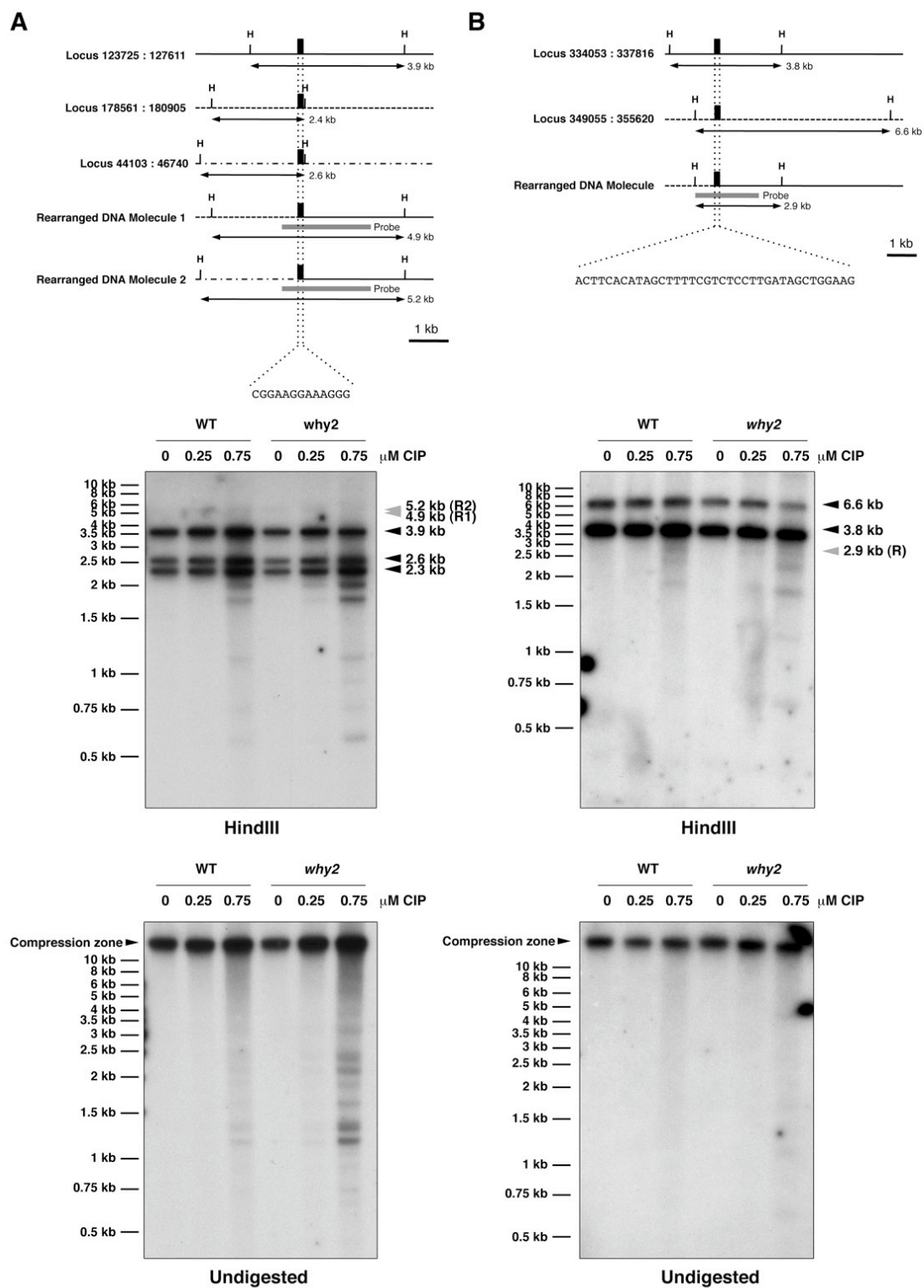


Figure 30. DNA Gel Blot Detection of DNA Rearrangements in Mitochondria Following Plant Treatment with Ciprofloxacin.

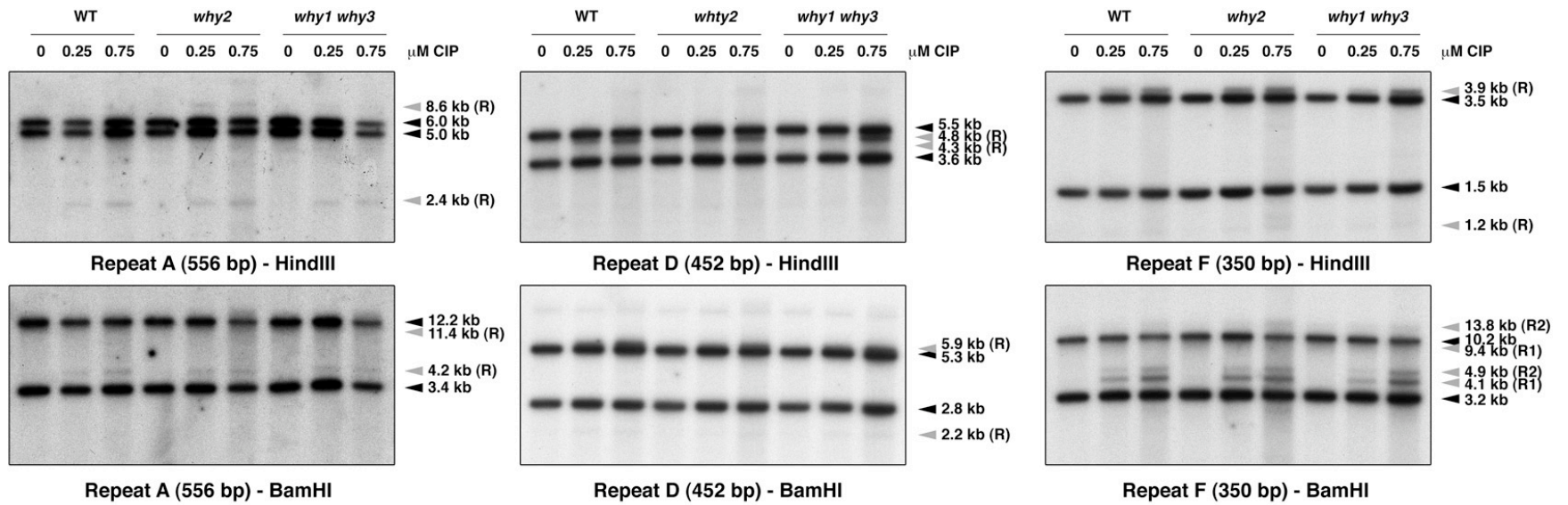


Figure 31. Analysis of Mitochondrial Recombination in Plants Treated with Ciprofloxacin.

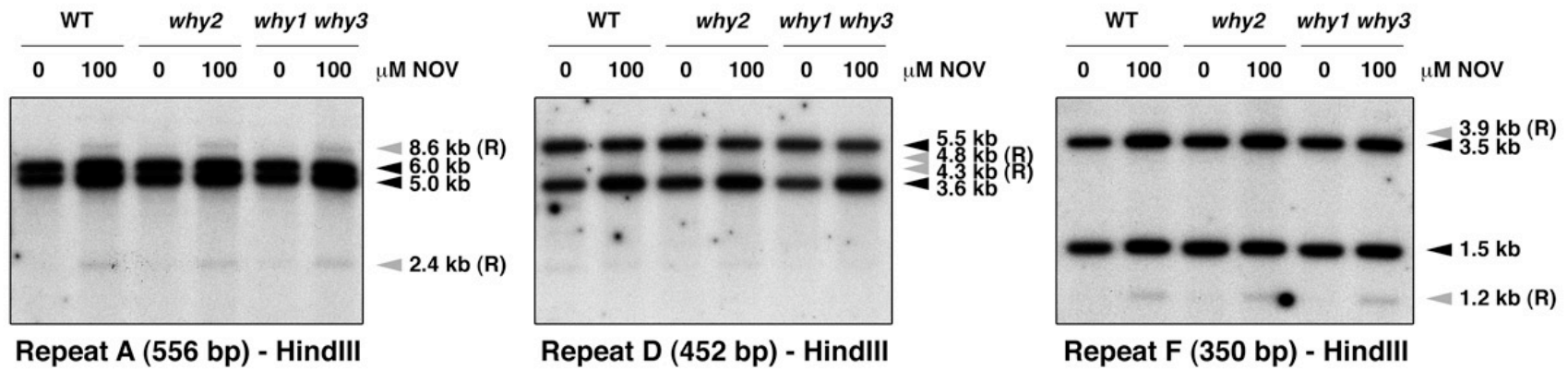


Figure 32. Analysis of Mitochondrial Recombination in Plants Treated with Novobiocin.

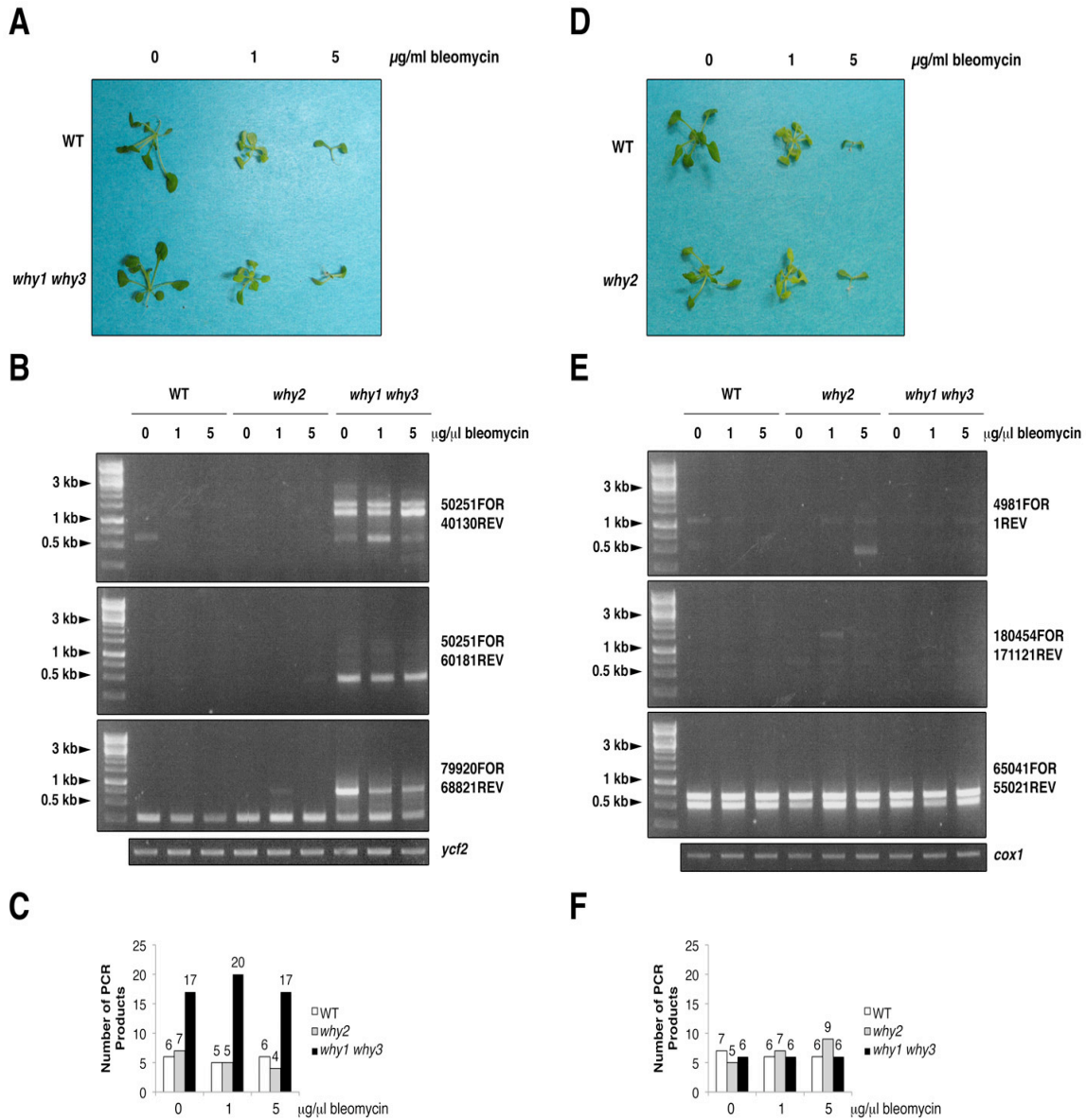


Figure 33. Effects of Bleomycin on Plastid and Mitochondria DNA Rearrangements of Plants lacking Plastid- or Mitochondria-Targeted Whirlies.

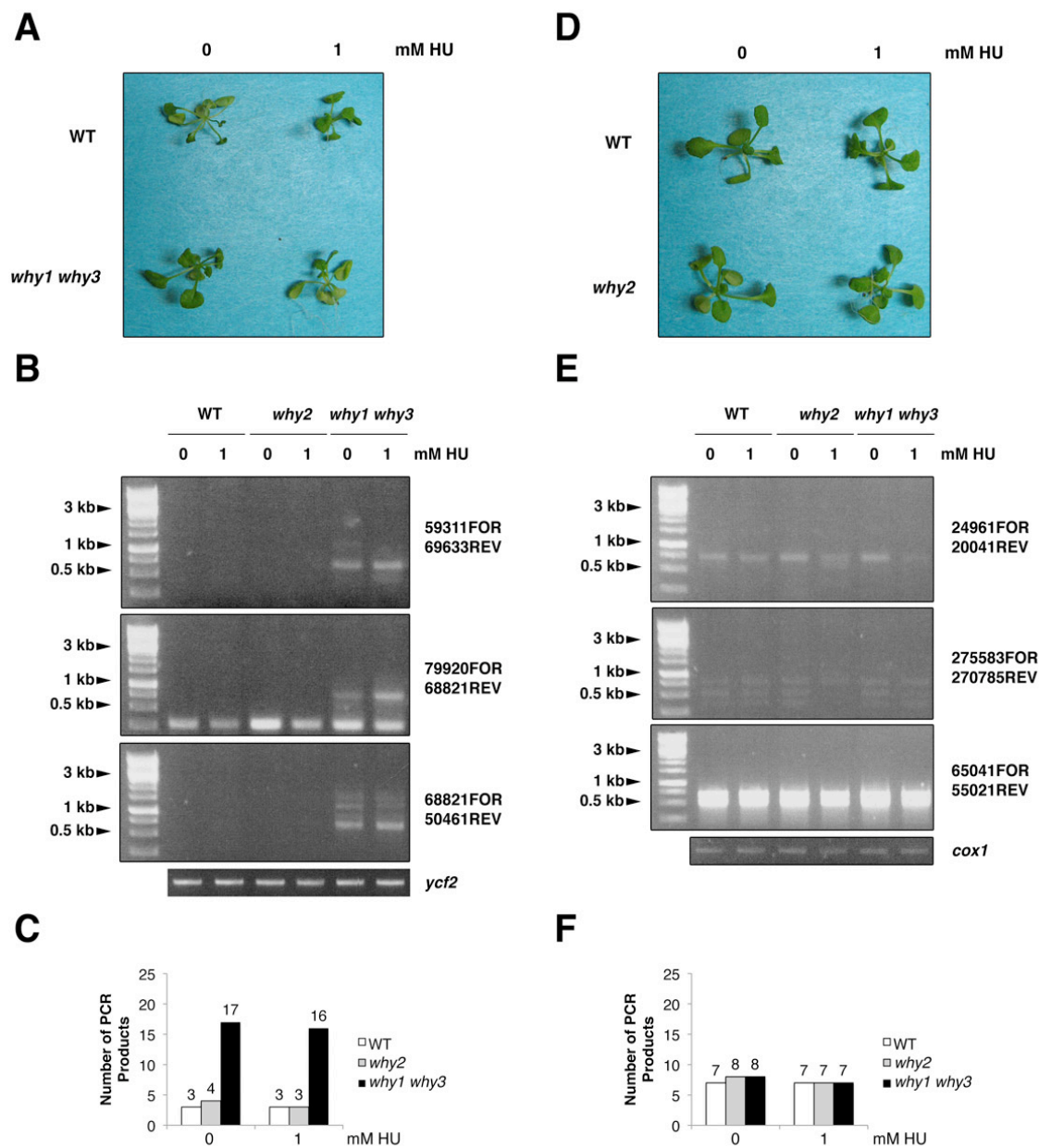


Figure 34. Effects of Hydroxyurea on Plastid and Mitochondria DNA Rearrangements of Plants lacking Plastid- or Mitochondria-Targeted Whirlies.

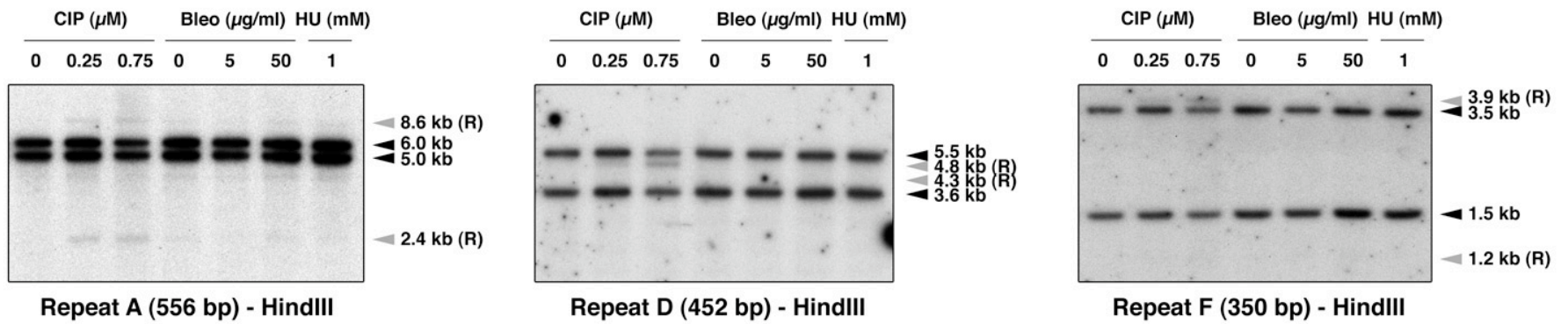


Figure 35. Analysis of Mitochondrial Recombination in WT Plants Treated with Bleomycin or Hydroxyurea.

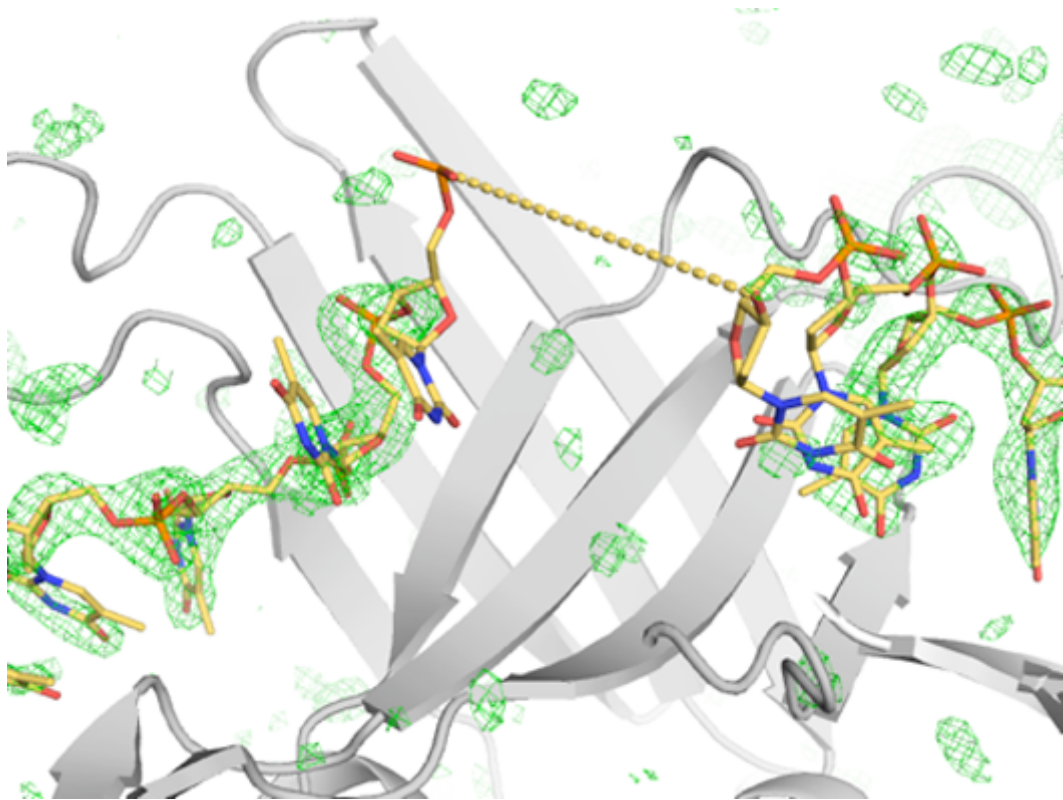


Figure 36. Single-Stranded DNA is Stabilized on the Edges of the β - Sheets of Whirly Proteins.

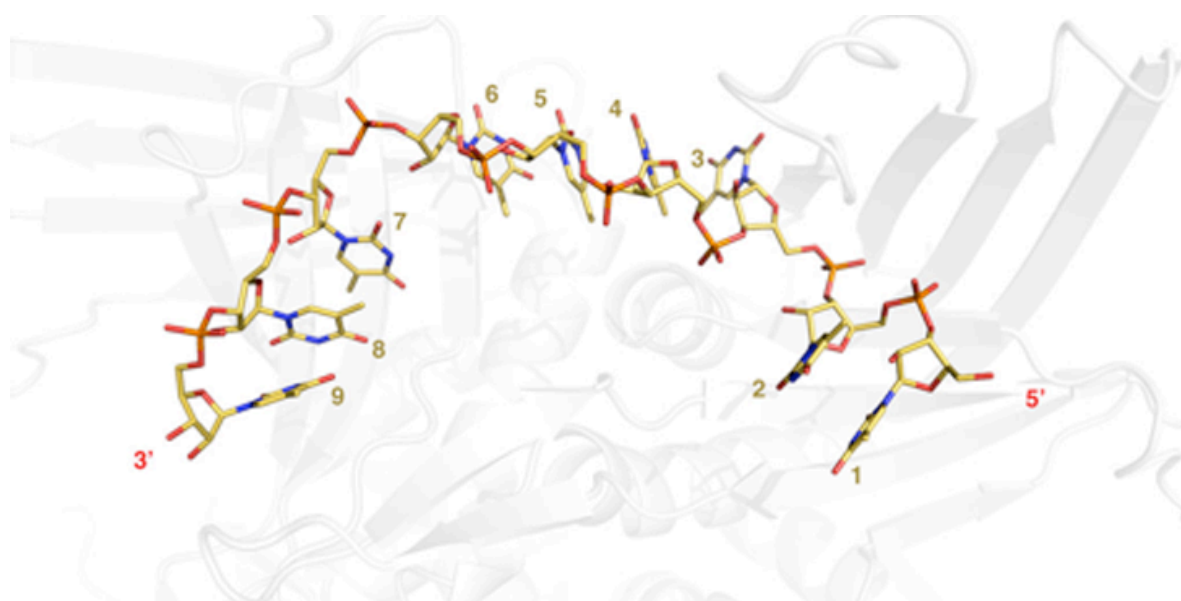


Figure 37. Modeling of a St-WHY2-RNA Complex.

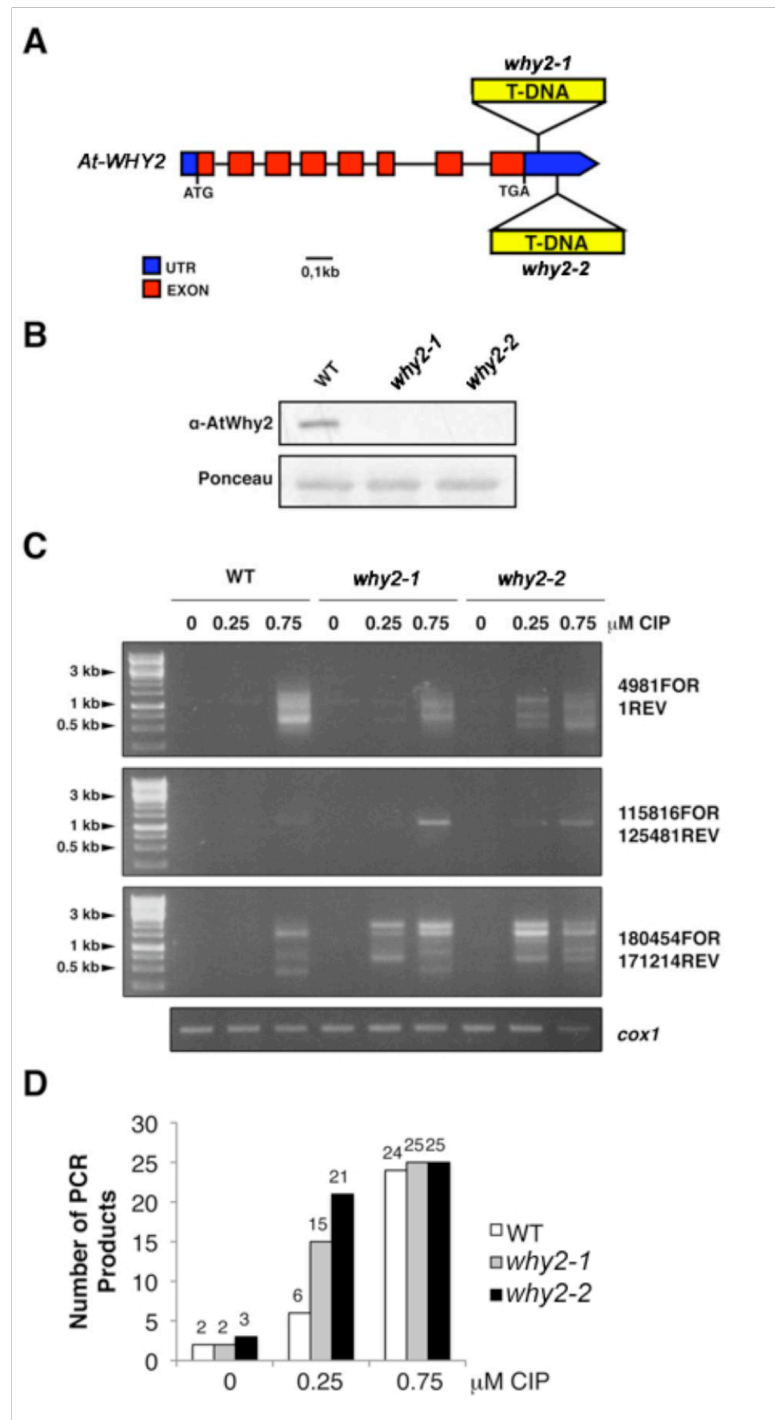


Figure 38. Molecular Characterization of why2 Lines.

Tableau X. DNA Rearrangement Products Identified in the Mitochondrial and Plastid Genomes

Event	Plant Treatment ^a	Plant Genotype	Junction Endpoints ^b	Junction Sequences ^c	Length of Junctions (Mismatches)	Junction Types ^d
Mitochondrial reaction 1 with primers 20041REV and 30001FOR^{e,f}						
Mitochondrial reaction 2 with primers 30001REV and 50341FOR						
Mitochondrial reaction 3 with primers 55021REV and 65041FOR						
M3-1	CIP 0.75 μM	Col-0	54739 Recombinant 65547	aaaaaaagg G aaagaactac gtcgcgtgtac G aaagaactac gtcgcgtgtac G ttcatggcct	1 bp	NHEJ-like
M3-2	CIP 0.75 μM	Col-0	55001 Recombinant 65518	cgaaggagc cataccggaa tgagccgctg cataccggaa tgagccgctg tactgtggca	0 bp	NHEJ-like
M3-3	CIP 0.75 μM	<i>why2-1</i>	53421 Recombinant 65476	gctcgcgtcca TGGAGCGTATG attttccttg ctctaaagct TGGAGCGTATG attttccttg ctctaaagct TGGAGCGTATG gtaacggtta	(0) 11 bp (0)	MHMR
M3-4	CIP 0.75 μM	<i>why2-1</i>	54638 Recombinant 65742	ccttatttagc ctaatatctt atctccttct ctaatatctt atctccttct tcctttgact	0 bp	NHEJ-like
M3-5	CIP 0.75 μM	<i>why2-1</i>	54463 Recombinant 65476	tcgtgccttc gtcctaggct ctctaaagct gtcctaggct ctctaaagct tggagcgtat	0 bp	NHEJ-like
M3-6	CIP 0.75 μM	<i>why2-1</i>	54234 Recombinant 65096	taggaatctt AGGTTACCTCCAG ccttctgaag ccactaggta AGGTTACCTCCAG ccttctgaag ccactaggta AGGTTACTTCCAG agcggagtat	(0) 13 bp (1)	MHMR*
M3-7	CIP 0.75 μM	<i>why2-1</i>	54234 Recombinant 65096	taggaatctt AGGTTACCTCCAG ccttctgaag ccactaggta AGGTTACTTCCAG ccttctgaag ccactaggta AGGTTACTTCCAG agcggagtat	(1) 13 bp (0)	MHMR*
M3-8	CIP 0.75 μM	<i>why2-1</i>	54393 Recombinant 65074	cgcgacttgg GAAACCCCT tcttcattcg ctgagctgta GAAACCCCT tcttcattcg ctgagctgta GAAACCCCT gatccactag	(0) 9 bp (0)	MHMR
M3-9	CIP 0.75 μM	<i>why2-1</i>	54483 Recombinant 65089	tgccttcgctc CTAGGCTCGGCTACTTC tttctcaaata ccttgatcca CTAGGCTCGGCTACTTC tttctcgaata ccttgatcca CTAGGTAAGGTACTTC cagagcggagt	(0) 17 bp (4)	MHMR*
Mitochondrial reaction 4 with primers 130413REV and 151514FOR						
Mitochondrial reaction 5 with primers 180454REV and 200500FOR						
M5-1	CIP 0.75 μM	<i>why2-1</i>	179517 Recombinant 200771	atgacgggtt ACCCGG ttcaaggcctt aaggtaagag ACCCGG ttcaaggcctt aaggtaagag ACCCGG gaaaggcggg	(0) 6 bp (0)	MHMR
Mitochondrial reaction 6 with primers 235252REV and 245966FOR						
M6-1	CIP 0.25 μM	<i>why2-1</i>	234779 Recombinant 246725	gatgaggatg AAGTAAAGGAATT actctggttg acgaaagtga AAGTAAAGGAATT actctggttg acgaaagtga AAGTAAAGGATT gtctggocct	(0) 13 bp (1)	MHMR*

Tableau X (suite). DNA Rearrangement Products Identified in the Mitochondrial and Plastid Genomes

Mitochondrial reaction 7 with primers 305346REV and 315513FOR ^{e f}							
M7-1	CIP 0.25 μM	<i>why2-1</i>	305245 Recombinant 316714	cgaagataca AGCGCCCG gtccttcttct ttaaagtccat AGCGCCCG gtccttcttct ttaaagtccat AGCGCCCG attcगतacac	(0) 8 bp (0)	MHMR	
M7-2	CIP 0.25 μM	<i>why2-1</i>	304953 Recombinant 316326	gataaaggta AGGTCACACT ctcaactagta taacaggcgt AGATCACACT ctcaactagta taacaggcgt AGATCACACT ggaactggag	(1) 10 bp (0)	MHMR*	
M7-3	CIP 0.25 μM	<i>why2-1</i>	305304 Recombinant 315835	aggttttgcc TCACAGAAGGAGTGG gaaattgggg aaagactacg TCACTGAAG-AGTGG gaaattgggg aaagactacg TCACTGAAG-AGTGG tttgaagaac	(2) 14 bp (0)	MHMR*	
M7-4	CIP 0.75 μM	Col-0	304556 Recombinant 316456	agaaggagaa AGGCTATTAACATAAAAAAAAAATATAA agtccaaatg atgccctctc AGGCTAATGAAAGCAAAAAAAAAATAGAA agtccaaatg atgccctctc AGGCTAATGAAAGCAAAAAAAAAAGCAGAA ggccttact	(6) 28 bp (2)	MHMR*	
M7-5	CIP 0.75 μM	Col-0	305222 Recombinant 317005	cgataactcc tatcccgaag agctaggcag tatcccgaag agctaggcag atagatagac	(0) 0 bp	NHEJ- like	
M7-6	CIP 0.75 μM	Col-0	305245 Recombinant 316714	cgaagataca AGCGCCCG gtccttcttct ttaaagtccat AGCGCCCG gtccttcttct ttaaagtccat AGCGCCCG attcगतacac	(0) 8 bp (0)	MHMR	
M7-7	CIP 0.75 μM	Col-0	304900 Recombinant 316104	tggaatgctc TCGTTGTTAAAAA ctacgagcgg caggggggat TCTTGTCGATAAAA ctacgagcgg caggggggat TCTTGTCGATAGAA ggcattcgcg	(4) 15 bp (1)	MHMR*	
M7-8	CIP 0.75 μM	Col-0	305105 Recombinant 316064	ccagacctaa CGGCTA caagaataaa tagcaaacac CGGCTA caagaataaa tagcaaacac CGGCTA gataagtggg	(0) 6 bp (0)	MHMR	
M7-9	CIP 0.75 μM	<i>why2-1</i>	305257 Recombinant 316731	cggtcctott CTCT tgtgtcgaag gattcgatca CTCT tgtgtcgaag gattcgatca CTCT atcgaataga	(0) 4 bp	NHEJ- like	
M7-10	CIP 0.75 μM	<i>why2-1</i>	305245 Recombinant 316714	cgaagataca AGCGCCCG gtccttcttct ttaaagtccat AGCGCCCG gtccttcttct ttaaagtccat AGCGCCCG attcगतacac	(0) 8 bp (0)	MHMR	
M7-11	CIP 0.75 μM	<i>why2-1</i>	305252 Recombinant 316371	agatacaagc GCCCGGTCCTCT tctcttgtgt gacgaacaat GTCGGCACTCT tctcttgtgt gacgaacaat GTCGGCACTCT ggtcaagtga	(3) 12 bp (0)	MHMR*	
M7-12	CIP 0.75 μM	<i>why2-1</i>	305132 Recombinant 315565	aaaagtacag GCAGAGCCAA acccatccta attcगतatct GCAGAGCCAA acccatccta attcगतatct GCAGAGACAA ctggaacaga	(0) 10 bp (1)	MHMR*	
Mitochondrial reaction 8 with primers 165904REV and 200500FOR							
Mitochondrial reaction 9 with primers 155034REV and 200500FOR							
Mitochondrial reaction 10 with primers 10081REV and 15001FOR							

Tableau X (suite). DNA Rearrangement Products Identified in the Mitochondrial and Plastid Genomes

M10-1	CIP 0.75 μ M	Col-0	9505 Recombinant 16023	atccttagctt G agcctacctt aagtgaaggg G agcctacctt aagtgaaggg G cgctatccgg	1 bp	NHEJ- like
M10-2	CIP 0.75 μ M	<i>why2-1</i>	9725 Recombinant 16003	ggttgctact GCGAAGCTA-TAGTGAAGGTGC acgggggtctt tcctagagta GCGAAGCTACAAGTGAAGGTGC acgggggtctt tcctagagta GCGAAACTACAAGTGAAGGGGC gctatccgga	(2) 22 bp (2)	MHMR*
M10-3	CIP 0.75 μ M	<i>why2-1</i>	9511 Recombinant 15480	gcttgagcct A ccttcgcaca acgaagggtc A ccttcgcaca acgaagggtc A tccagaagcg	1 bp	NHEJ- like
Mitochondrial reaction 11 with primers 20461REV and 24961FOR						
M11-1	CIP 0.75 μ M	Col-0	19795 Recombinant 25300	gaatccgagg ATTTGCTTGGATTGTCCGCT tcatctcccc atatcttacc ATTTCTTCGGAATTGTCCGCT tcatctcccc atatcttacc ATTTCTTCGGAATTGTCCGCT catagcacga	(4) 21 bp (0)	MHMR*
M11-2	CIP 0.75 μ M	<i>why2-1</i>	19317 Recombinant 25309	caagacgata GATAGGCCCG ttatgaaaaa catttcttcg GAATTGGCCG ttatgaaaaa catttcttcg GAATTGTCG ctcatagcac	(5) 10 bp (1)	MHMR*
M11-3	CIP 0.75 μ M	<i>why2-1</i>	19795 Recombinant 25300	gaatccgagg ATTTGCTTGGATTGTCCGCT tcatctcccc atatcttacc ATTTCTTCGGAATTGTCCGCT tcatctcccc atatcttacc ATTTCTTCGGAATTGTCCGCT catagcacga	(4) 21 bp (0)	MHMR*
Mitochondrial reaction 12 with primers 30001REV and 80161FOR						
M12-1	CIP 0.25 μ M	Col-0	29609 Recombinant 80994	tttttccct CCCCCCCCCA aaaaaaaaaa attagacgtc CCCCCCCCCA aaaaaaaaaag attagacgtc CCCCCCCCCA tgctacggtt	(0) 10 bp (0)	MHMR
M12-2	CIP 0.25 μ M	<i>why2-1</i>	29609 Recombinant 80994	tttttccct CCCCCCCCCA aaaaaaaaaa attagacgtc CCCCCCCCCA aaaaaaaaaag attagacgtc CCCCCCCCCA tgctacggtt	(0) 10 bp (0)	MHMR
M12-3	CIP 0.25 μ M	<i>why2-1</i>	29610 Recombinant 80713	acttttttcc CCT-CCCCCCCCCA aaaaaaaaaag ccggggcaaa CCT-CCCCCCCCCA aaaaaaaaaag ccggggcaaa CCA <u>CCCCCCT</u> CCAA gaggaattgt	(0) 14 bp (3)	MHMR*
M12-4	CIP 0.75 μ M	Col-0	29609 Recombinant 80994	tttttccct CCCCCCCCCA aaaaaaaaaa attagacgtc CCCCCCCCCA aaaaaaaaaag attagacgtc CCCCCCCCCA tgctacggtt	(0) 10 bp (0)	MHMR
M12-5	CIP 0.75 μ M	<i>why2-1</i>	29609 Recombinant 80994	tttttccct CCCCCCCCCA aaaaaaaaaa attagacgtc CCCCCCCCCA aaaaaaaaaag attagacgtc CCCCCCCCCA tgctacggtt	(0) 10 bp (0)	MHMR
M12-6	CIP 0.75 μ M	<i>why2-1</i>	29610 Recombinant 80713	acttttttcc CCT-CCCCCCCCCA aaaaaaaaaag ccggggcaaa CCT-CCCCCCCCCA aaaaaaaaaag ccggggcaaa CCA <u>CCCCCCT</u> CCAA gaggaattgt	(0) 14 bp (3)	MHMR*
Mitochondrial reaction 13 with primers 84901REV and 151514FOR^f						
Mitochondrial reaction 14 with primers 210239REV and 245966FOR^f						
Mitochondrial reaction 15 with primers 1REV and 4981FOR^f						

Tableau X (suite). DNA Rearrangement Products Identified in the Mitochondrial and Plastid Genome

M15-1	CIP 0.25 μ M	<i>why2-1</i>	366037 Recombinant 5462	ctaactgata ggccttggt ctttacctat ggccttggt ctttacctat cttgggtgag	0 bp	NHEJ- like
M15-2	CIP 0.25 μ M	<i>why2-1</i>	366764 Recombinant 5145	gcgaatagaa GGCTT-TTGCCAC ctcctcaac gattottctg GGCTTGTGCCAC ctcctcaac gattottctg GGCTTGTGCCAC taaagaaaga	(1) 13 bp (0)	MHMR*
M15-3	CIP 0.75 μ M	Col-0	366038 Recombinant 5171	taactgatag gccctggctg agaaagagtt gccctggctg agaaagagtt ctgccttcta	0 bp	NHEJ- like
M15-4	CIP 0.75 μ M	Col-0	366042 Recombinant 5080	tgataggccc tggctggacc ccgtaggaaa tggctggacc ccgtaggaaa agccgttgat	0 bp	NHEJ- like
M15-5	CIP 0.75 μ M	Col-0	366617 Recombinant 5215	gccgacttta AGTCAAAGGAGAAAGAGTG cctactgcca gcttcccagc AGTCAAAGGAGAAAGAGTG cctactgcca gcttcccagc AGTCAA <u>CA</u> AGAAAGAGTG tttctcccct	(0) 19 bp (3)	MHMR*
M15-6	CIP 0.75 μ M	<i>why2-1</i>	366041 Recombinant 5298	actgataggc C ctggctggac cacatctgct C ctggctggac cacatctgct C tttccaggac	1 bp	NHEJ- like
M15-7	CIP 0.75 μ M	<i>why2-1</i>	366856 Recombinant 5607	agaagaggtc TCAAGGAGCC tctgaattt aaactcccat TCAAGGAGCC tctgaattt aaactcccat TCAAGG <u>GC</u> CC aatcattcat	(1) 10 bp (1)	MHMR*
M15-8	CIP 0.75 μ M	<i>why2-1</i>	366856 Recombinant 5577	aggtctcaac GAGCC tctgaattt tttttataa GAGCC tctgaattt tttttataa GAGCC attggcagga	(0) 5 bp (0)	MHMR
Mitochondrial reaction 16 with primers 125781REV and 115816FOR^f – Deletion Events						
M16-1	CIP 0.25 μ M	<i>why2-1</i>	125708 Recombinant 116865	taatgctttc TCTGCTTCCCTCT tggtttcggt ggttgctgag TCTGCTTCCCTCT tggtttcggt ggttgctgag TCTGCTTCCCTCT ctatcaaagc	(0) 14 bp (0)	MHMR
M16-2	CIP 0.75 μ M	Col-0	125359 Recombinant 116280	gggtacgtcc CCTATC gcctttcatt gagcaggtaa CCTACC gcctttcatt gagcaggtaa CCTACC ttctaaacca	(1) 6 bp (0)	MHMR*
M16-3	CIP 0.75 μ M	<i>why2-1</i>	125529 Recombinant 116981	ttagctgagc ATCAATCCCAT tctttgcgga caggtatgaa ATCAATCCCAT tctttgcgga caggtatgaa ATCAATCCCAT gttggtcgta	(0) 11 bp (0)	MHMR
M16-4	CIP 0.75 μ M	<i>why2-1</i>	125708 Recombinant 116865	taatgctttc TCTGCTTCCCTCT tggtttcggt ggttgctgag TCTGCTTCCCTCT tggtttcggt ggttgctgag TCTGCTTCCCTCT ctatcaaagc	(0) 14 bp (0)	MHMR
Mitochondrial reaction 17 with primers 135180REV and 140053FOR^f						
Mitochondrial reaction 18 with primers 171214REV and 180454FOR^f						
M18-1	CIP 0.25 μ M	<i>why2-1</i>	170939 Recombinant 181121	attctataga TA <u>CT</u> CTGGTAG gtaaaagcact ggacaggtag TAGCTCTGGTAG gtaaaagcact ggacaggtag TAGCTCTGGTAG agcgaggagc	(1) 12 bp (0)	MHMR*

Tableau X (suite). DNA Rearrangement Products Identified in the Mitochondrial and Plastid Genomes

M18-2	CIP 0.25 μ M	<i>why2-1</i>	171149 Recombinant 181066	aaccttcgtc TCAATCTTCATAA tcaatztatg tcacaaacta TCAATCTTCATAA tcaatztatg tcacaaacta TCAAT_TTCATAA gagaagacga	(0) 13 bp (1)	MHMR*
M18-3	CIP 0.75 μ M	Col-0	170678 Recombinant 181451	gttttggatg gagctactgg attctccgtt gagctactgg attctccgtt tcgatattaa	0 bp	NHEJ- like
M18-4	CIP 0.75 μ M	Col-0	170609 Recombinant 180538	ttgtgatgac TGAACATGA-AACGATCGATATA gatcggcgta agtcaaaactg TGAACATGA-AACGATCGATATA gatcggcgta agtcaaaactg TGAACAT_CAGAAGGATCGATATA ttggtaatag	(0) 22 bp (3)	MHMR*
M18-5	CIP 0.75 μ M	<i>why2-1</i>	170939 Recombinant 181121	attctataga TA_CCTCTGGTAG gtaaagcact ggacaggtag TAGCTCTGGTAG gtaaagcact ggacaggtag TAGCTCTGGTAG agcgaggac	(1) 12 bp (0)	MHMR*
M18-6	CIP 0.75 μ M	<i>why2-1</i>	171149 Recombinant 181066	aaccttcgtc TCAATCTTCATAA tcaatztatg tcacaaacta TCAATCTTCATAA tcaatztatg tcacaaacta TCAAT_TTCATAA gagaagacga	(0) 13 bp (1)	MHMR*
Mitochondrial reaction 19 with primers 245966REV and 239584FOR^f – Deletion Event						
M19-1	CIP 0.75 μ M	<i>why2-1</i>	244727 Recombinant 240924	gtaaccccct AAAAGGAAACTCACCACCAAAGA_ aaaaggtagg aaaaaaagaa AAAAGGAAACTCACCACCAAAGAT aaaaggtagg aaaaaaagaa AAAAGGAAACTCACCACCAAAGAT cagtccaata	(1) 25 bp (0)	MHMR*
Mitochondrial reaction 20 with primers 270785REV and 275583FOR^f						
M20-1	CIP 0.25 μ M	<i>why2-1</i>	269741 Recombinant 276747	attgaagaaa GGAAGAAAGCCTGTAAA ctaccttact tccggtcatt GGTAGAAAGCCTGTAAA ctaccttact tccggtcatt GGTAGAAAGCCTGTAAA tggaaatttat	(1) 17 bp (0)	MHMR*
M20-2	CIP 0.25 μ M	<i>why2-1</i>	269069 Recombinant 275816	gtagtcccgct CCGGATC aatagattgg tctcccaaac CCGGATC aatagattgg tctcccaaac CCGGATC cgcaagtcgg	(0) 7 bp (0)	MHMR
M20-3	CIP 0.25 μ M	<i>why2-1</i>	270780 Recombinant 276407	tccaagccta ATAGGA acatcgctag tattccttcc ATAGGA acatcgctag tattccttcc ATAGGA ggacaccttg	(0) 6 bp (0)	MHMR
M20-4	CIP 0.75 μ M	Col-0	269800 Recombinant 276776	cagtaaaacta CCTCTCCTTCAGT gttttttcct gaatztatct CCTCTCCTTCAGT gttttttcct gaatztatct CCTCTCCTTCAGT cgagttccct	(1) 14 bp (0)	MHMR*
M20-5	CIP 0.75 μ M	Col-0	270679 Recombinant 276212	aggacggaac caccgaggaa tttggtggtt caccgaggaa tttggtggtt ttccacgacc	0 bp	NHEJ- like
M20-6	CIP 0.75 μ M	Col-0	270497 Recombinant 276011	cattcatggc C ttttttgttt gagatgcgtg C ttttttgttt gagatgcgtg C gcacgggtga	1 bp	NHEJ- like
M20-7	CIP 0.75 μ M	Col-0	270716 Recombinant 275872	agtaggatgt atactgtgct tgctacctcc atactgtgct tgctacctcc caacagacgt	0 bp	NHEJ- like

Tableau X (suite). DNA Rearrangement Products Identified in the Mitochondrial and Plastid Genomes

M20-8	CIP 0.75 μ M	Col-0	270715 Recombinant 275738	tagtaggatg tatactgtgc ggaatatggc tatactgtgc ggaatatggc caccgगतग	0 bp	NHEJ- like
M20-9	CIP 0.75 μ M	<i>why2-1</i>	269741 Recombinant 276747	attgaagaaa GGAAGAAAGCTGTAAA ctaccttact tccggtcatt GGTAGAAAGCTGTAAA ctaccttact tccggtcatt GGTAGAAAGCTGTAAA tggaaatttat	(1) 17 bp (0)	MHMR*
M20-10	CIP 0.75 μ M	<i>why2-1</i>	270561 Recombinant 276533	ctgaggcgtt CCTCTCAGTCA taca ^{aaaa} aag atctcctacg CCTCTCAGTCA taca ^{aaaa} aag atctcctacg CCTCTCAGTCA aaagagtgcg	(0) 11 bp (0)	MHMR
M20-11	CIP 0.75 μ M	<i>why2-1</i>	270682 Recombinant 276415	ggacggaacc AC cgaggaatac ccatagggagg AC cgaggaatac ccatagggagg AC accttggttc	2 bp	NHEJ- like
M20-12	CIP 0.75 μ M	<i>why2-1</i>	270521 Recombinant 276184	ttctttgatt GATCTC cctttcgat tcttccttaa GATCTC cctttcgat tcttccttaa GATCTC aaccaaccac	(0) 6 bp (0)	MHMR
M20-13	CIP 0.75 μ M	<i>why2-1</i>	270135 Recombinant 275738	gatgtggaag CAC atcttagcaa ggaatatggc CAC atcttagcaa ggaatatggc CAC cggatagagc	3 bp	NHEJ- like
M20-14	CIP 0.75 μ M	<i>why2-1</i>	270445 Recombinant 275819	gcatttatca GATCAGC tcccagagta cccaaaccg GATCAGC tcccagagta cccaaaccg GATCAGC aagtcgggtc	(0) 7 bp (1)	MHMR*
M20-15	CIP 0.75 μ M	<i>why2-1</i>	270679 Recombinant 275979	aggacggaac caccgaggaa agaaaggctg caccgaggaa agaaaggctg tatcggctag	0 bp	NHEJ- like
Plastid reaction 1 with primers 31138REV and 19181FOR¹ – Deletion Events						
P1-1	CIP 0.25 μ M	Col-0	30922 Recombinant 20410	aaatagtagt atccctgtct atgatagtta atccctgtct atgatagtta ccaaaagatt	0 bp	NHEJ- like
P1-2	CIP 0.25 μ M	Col-0	31024 Recombinant 20347	gttaagttat CAAAAATTTGCTTTTTT ttttttagaa aatatatata CAAAAATTTCTTTTTT ttttttagaa aatatatata CAAAAATTTCTTTTTT acacttctta	(1) 18 bp (0)	MHMR*
P1-3	CIP 0.25 μ M	<i>why1why3</i>	31024 Recombinant 20347	gttaagttat CAAAAATTTGCTTTTTT ttttttagaa aatatatata CAAAAATTTCTTTTTT ttttttagaa aatatatata CAAAAATTTCTTTTTT acacttctta	(1) 18 bp (0)	MHMR*
P1-4	CIP 0.75 μ M	<i>why1why3</i>	31024 Recombinant 20347	gttaagttat CAAAAATTTGCTTTTTT ttttttagaa aatatatata CAAAAATTTCTTTTTT ttttttagaa aatatatata CAAAAATTTCTTTTTT acacttctta	(1) 18 bp (0)	MHMR*
P1-5	NOV 0 μ M	<i>why1why3</i>	31024 Recombinant 20347	gttaagttat CAAAAATTTGCTTTTTT ttttttagaa aatatatata CAAAAATTTCTTTTTT ttttttagaa aatatatata CAAAAATTTCTTTTTT acacttctta	(1) 18 bp (0)	MHMR*

Tableau X (suite). DNA Rearrangement Products Identified in the Mitochondrial and Plastid Genomes

P1-6	NOV 100 μ M	<i>why1why3</i>	31024 Recombinant 20347	gttaagttat CAAAAATTTGCTTTTTTT ttttttagaa aatatatata CAAAAATTTCTTTTTTT ttttttagaa aatatatata CAAAAATTTCTTTTTTT acacttctta	(1) 18 bp (0)	MHMR*
P1-7	NOV 100 μ M	<i>why1why3</i>	30618 Recombinant 19900	tataatgaaa TGAAATGG tatgactagc caattcctaa TGAAATGG tatgactagc caattcctaa TGAAATGG atgtagcagt	(0) 8 bp (0)	MHMR
Plastid reaction 2 with primers 40130REV and 50251FOR^{e f}						
P2-1	CIP 0 μ M	Col-0	39586 Recombinant 50325	ttaatgccat AATGCC tttcaaatcc tattagtaaa AATGCC tttcaaatcc tattagtaaa AATGCC cagaaaatat	(0) 6 bp (0)	MHMR
P2-2	CIP 0 μ M	<i>why1why3</i>	38759 Recombinant 50687	ttgcgatagc TAAATGATGATGT gccataticg cttcgTTTTT TAAATGATGATGT gccataticg cttcgTTTTT TAAATGATGACGT atcaaaaaat	(0) 13 bp (1)	MHMR*
P2-3	CIP 0 μ M	<i>why1why3</i>	38989 Recombinant 50299	cccctggatg CAGGAATAGC gacatggact aatgccaaaa CAGGAATAGC gacatggact aatgccaaaa CAGGAATAGC acttgatatt	(0) 10 bp (0)	MHMR
P2-4	CIP 0.25 μ M	Col-0	39586 Recombinant 50325	ttaatgccat AATGCC tttcaaatcc tattagtaaa AATGCC tttcaaatcc tattagtaaa AATGCC cagaaaatat	(0) 6 bp (0)	MHMR
P2-5	CIP 0.25 μ M	<i>why1why3</i>	39109 Recombinant 50758	cttggtttcc ATTTTGGTTG taggtgtaac gatagattga ATTTTGGTTG taggtgtaac gatagattga ATTTTGGTTG aattaatta	(0) 10 bp (0)	MHMR
P2-6	CIP 0.25 μ M	<i>why1why3</i>	38989 Recombinant 50299	cccctggatg CAGGAATAGC gacatggact aatgccaaaa CAGGAATAGC gacatggact aatgccaaaa CAGGAATAGC acttgatatt	(0) 10 bp (0)	MHMR
P2-7	CIP 0.25 μ M	<i>why1why3</i>	39517 Recombinant 50285	tcatgtgcgg TAGCAATACCAAACCA aatacgacga ctccggaaat TAGCAATACCAAACCA aatacgacga ctccggaaat TAGAATGCCAAAACA ggaatagcac	(0) 16 bp (3)	MHMR*
P2-8	CIP 0.25 μ M	<i>why1why3</i>	39582 Recombinant 50289	tgaaatcct AATGCCATAA tgcctttcaa ggaaattaga AATGCCATAA tgcctttcaa ggaaattaga AATGCCAAA caggaatagc	(0) 10 bp (1)	MHMR*
P2-9	CIP 0.75 μ M	Col-0	37846 Recombinant 50396	ccagaaaact GC caataaaat aaaaatacc GC caataaaat aaaaatacc GC ttagtcaatt	2 bp	NHEJ- like
P2-10	CIP 0.75 μ M	Col-0	39655 Recombinant 50363	attcttgcta AGAAGAACGCCCATG-TTGTGGCAA ttcacccag gcagaaacat AGACGAACGCCCATG-TTGTGGCAA ttcacccag gcagaaacat AGACGAACCTTATGAATGTGGAAA aaataccgc	(1) 24 bp (5)	MHMR*
P2-11	CIP 0.75 μ M	Col-0	39192 Recombinant 50308	ttcattagta C gtaaaccgat acaggaatag C gtaaaccgat acaggaatag C acttgatatt	1 bp	NHEJ- like
P2-12	CIP 0.75 μ M	Col-0	39586 Recombinant 50325	ttaatgccat AATGCC tttcaaatcc tattagtaaa AATGCC tttcaaatcc tattagtaaa AATGCC cagaaaatat	(0) 6 bp (0)	MHMR

Tableau X (suite). DNA Rearrangement Products Identified in the Mitochondrial and Plastid Genomes

P2-13	CIP 0.75 μ M	Col-0	39845 Recombinant 50392	gagcacctag gaaaaaaga tggaaaaaat gaaaaaaga tggaaaaaat acccgcttag	0 bp	NHEJ- like
P2-14	CIP 0.75 μ M	<i>why1why3</i>	37647 Recombinant 50326	aagaacatcc ATGCCCAGA ctgataaact attagtaaaa ATGCCCAGA ctgataaact attagtaaaa ATGCCCAGA aaatatacata	(0) 9 bp (0)	MHMR
P2-15	CIP 0.75 μ M	<i>why1why3</i>	39109 Recombinant 50758	cttggtttcc ATTTTGGTTG taggtgtaac gatagattga ATTTTGGTTG taggtgtaac gatagattga ATTTTGGTTG aatttaatta	(0) 10 bp (0)	MHMR
P2-16	CIP 0.75 μ M	<i>why1why3</i>	38989 Recombinant 50299	cccctggatg CAGGAATAGC gacatggact aatgcaaaaa CAGGAATAGC gacatggact aatgcaaaaa CAGGAATAGC acttgatatt	(0) 10 bp (0)	MHMR
P2-17	CIP 0.75 μ M	<i>why1why3</i>	39536 Recombinant 50834	tacgacgagt AGTGGGG tcctgagcta accaagattg AGTGGGG tcctgagcta accaagattg AGTGGGG atacaaaaa	(0) 7 bp (0)	MHMR
P2-18	CIP 0.75 μ M	<i>why1why3</i>	39195 Recombinant 50397	ttcattagta CGTA aaccgattgt aaaatacccg CGTA aaccgattgt aaaatacccg CTTA gtcaattcca	4 bp	NHEJ- like
P2-19	CIP 0.75 μ M	<i>why1why3</i>	39613 Recombinant 50397	tagccattat CCTA ctgcaataat aaaatacccg CCTA ctgcaataat aaaatacccg CTTA gtcaattcca	4 bp	NHEJ- like
P2-20	CIP 0.75 μ M	<i>why1why3</i>	39586 Recombinant 50325	ttaatgccat AATGCC tttcaaatcc tattagtaaa AATGCC tttcaaatcc tattagtaaa AATGCC cagaaaaat	(0) 6 bp (0)	MHMR
P2-21	CIP 0.75 μ M	<i>why1why3</i>	39655 Recombinant 50363	attcttgcta AGAAGAAGCCCATG-TTGTGGCAA ttccaccag gcagaaacat AGACGAACGCCCATG-TTGTGGCAA ttccaccag gcagaaacat AGACGAACCTTATGAATGTGGA AA aaataccgc	(1) 24 bp (5)	MHMR*
P2-22	NOV 0 μ M	<i>why1why3</i>	38889 Recombinant 50299	cccctggatg CAGGAATAGC gacatggact aatgcaaaaa CAGGAATAGC gacatggact aatgcaaaaa CAGGAATAGC acttgatatt	(0) 10 bp (0)	MHMR
P2-23	NOV 0 μ M	<i>why1why3</i>	39109 Recombinant 50758	cttggtttcc ATTTTGGTTG taggtgtaac gatagattga ATTTTGGTTG taggtgtaac gatagattga ATTTTGGTTG aatttaatta	(0) 10 bp (0)	MHMR
P2-24	NOV 0 μ M	<i>why1why3</i>	39602 Recombinant 50362	caaatcctcc TAG ccattatcct agcagaaaca TAG ccattatcct agcagaaaca TAG acgaactcct	3 bp	NHEJ- like
P2-25	NOV 100 μ M	<i>why1why3</i>	38889 Recombinant 50299	cccctggatg CAGGAATAGC gacatggact aatgcaaaaa CAGGAATAGC gacatggact aatgcaaaaa CAGGAATAGC acttgatatt	(0) 10 bp (0)	MHMR
Plastid reaction 3 with primers 84761REV and 95260FOR^{g f}						
P3-1	CIP 0.25 μ M	<i>why1why3</i>	83151 Recombinant 95532	ggaaattctg GTAATAGCAATG ttgagttttc gcttgtttga GTAATAGCAATG ttgagttttc gcttgtttga GTAATAGCAATG agattcccga	(0) 12 bp (0)	MHMR

Tableau X (suite). DNA Rearrangement Products Identified in the Mitochondrial and Plastid Genomes

P3-2	CIP 0.25 μ M	<i>why1why3</i>	142740 Recombinant 36	aaaatcagtc <u>ATTCACAA</u> – <u>CCACTG</u> atgaaggatt gcgatggtga <u>ATTCACAATCCACTG</u> atgaaggatt gcgatggtga <u>ATTCACAATCCACTG</u> ccttaatcca	(1) 15 bp (0)	MHMR*
P3-3	CIP 0.25 μ M	<i>why1why3</i>	84368 Recombinant 95699	acggcgacga <u>AGAATCAAAGTC</u> tcaactatatt ctggagtggg <u>AGAATCAAAGTC</u> tcaactatatt ctggagtggg <u>AGAGTCAGAGTC</u> gaaaagagga	(0) 12 bp (2)	MHMR*
P3-4	CIP 0.25 μ M	<i>why1why3</i>	84661 Recombinant 95481	tttaatgtgg <u>CCGATTT</u> cccctctttt aattacatat <u>CCGATTT</u> cccctctttt aattacatat <u>CCGATTT</u> gacctatgga	(0) 7 bp (0)	MHMR
P3-5	CIP 0.75 μ M	Col-0	84464 Recombinant 96061	gtcgacggat gctcctatta gtcgacggat ccacctccat cctcccttca ccacctccat	0 bp	NHEJ- like
P3-6	CIP 0.75 μ M	<i>why1why3</i>	83151 Recombinant 95532	ggaaattctg <u>GTAATAGCAATG</u> ttgagttttc gcttgtttga <u>GTAATAGCAATG</u> ttgagttttc gcttgtttga <u>GTAATAGCAATG</u> agattcccga	(0) 12 bp (0)	MHMR
P3-7	CIP 0.75 μ M	<i>why1why3</i>	83532 Recombinant 95421	ttagatttct <u>GTTTTCAT</u> tgctgcatcc cagcatataa <u>GTTTTCAT</u> tgctgcatcc cagcatataa <u>GTTATCAT</u> gctcgcatat	(0) 8 bp (1)	MHMR*
P3-8	CIP 0.75 μ M	<i>why1why3</i>	142740 Recombinant 36	aaaatcagtc <u>ATTCACAA</u> – <u>CCACTG</u> atgaaggatt gcgatggtga <u>ATTCACAATCCACTG</u> atgaaggatt gcgatggtga <u>ATTCACAATCCACTG</u> ccttaatcca	(1) 15 bp (0)	MHMR*
P3-9	CIP 0.75 μ M	<i>why1why3</i>	143168 Recombinant 153988	tccataggtc <u>AAATCGG</u> atatgtaatt aaaagagggg <u>AAATCGG</u> atatgtaatt aaaagagggg <u>AAATCGG</u> ccacattaaa	(0) 7 bp (0)	MHMR
Plastid reaction 4 with primers 105961REV and 116007FOR^f						
P4-1	CIP 0.25 μ M	<i>why1why3</i>	105738 Recombinant 116787	agctaaggcc <u>CCTAAATGACCGCT</u> cagtgataaa agaaaatgat <u>CCTAAATGACCGCT</u> cagtgataaa agaaaatgat <u>CCTATTTGACCGCT</u> gtacattgct	(0) 14 bp (2)	MHMR*
P4-2	CIP 0.25 μ M	<i>why1why3</i>	105452 Recombinant 116276	ggaggtccga <u>ACCGACTGATGTTGAA</u> aaatcagcgg aattacgttg <u>ACCGGGAGATGTTGAA</u> aaatcagcgg aattacgttg <u>ACCGGGAGATGTTGAA</u> gctgcataaa	(3) 16 bp (0)	MHMR*
P4-3	CIP 0.25 μ M	<i>why1why3</i>	105322 Recombinant 116009	tggttggtt <u>AAGGGAACCCACG</u> gagccgtagc cgggcaatgc <u>AAGGGAAGCCATCG</u> gagccgtagc cgggcaatgc <u>AAGGGAAGCCATCG</u> ataaaatagt	(2) 14 bp (0)	MHMR*
P4-4	CIP 0.75 μ M	Col-0	105739 Recombinant 116394	ctaatgacc <u>GCTC</u> agtgataaag caaccatag <u>GCTC</u> agtgataaag caaccatag <u>GCTC</u> ccatttttaa	4 bp	NHEJ- like
P4-5	CIP 0.75 μ M	<i>why1why3</i>	105452 Recombinant 116276	ggaggtccga <u>ACCGACTGATGTTGAA</u> aaatcagcgg aattacgttg <u>ACCGGGAGATGTTGAA</u> aaatcagcgg aattacgttg <u>ACCGGGAGATGTTGAA</u> gctgcataaa	(3) 16 bp (0)	MHMR*
P4-6	CIP 0.75 μ M	<i>why1why3</i>	105821 Recombinant 116393	gtgcgtaata <u>GCTC</u> actgatcgag caaccatag <u>GCTC</u> actgatcgag caaccatag <u>GCTC</u> ccatttttaa	4 bp	NHEJ- like

Tableau X (suite). DNA Rearrangement Products Identified in the Mitochondrial and Plastid Genomes

Plastid reaction 5 with primers 69633REV and 59291FOR ^{e f} – Deletion Events								
P5-1	CIP 0.25 μM	<i>why1why3</i>	68984	tctttccggc	AATTGCAATGGCTT	ctttatttct	(1)	MHMR*
			Recombinant 59308	gtataaaaaa	AATTCCAATGGCTT	ctttatttct	14 bp (0)	
P5-2	CIP 0.75 μM	<i>why1why3</i>	68984	tctttccggc	AATTGCAATGGCTT	ctttatttct	(1)	MHMR*
			Recombinant 59308	gtataaaaaa	AATTCCAATGGCTT	ctttatttct	14 bp (0)	
P5-3	NOV 0 μM	<i>why1why3</i>	68522	atttcttcca	ATCCCTTATTTTTTTT	atgtcattcg	(2)	MHMR*
			Recombinant 59380	accgaatcat	ATCACCTTTTTTTTTT	atgtcattcg	16 bp (0)	
P5-4	NOV 100 μM	<i>why1why3</i>	68984	tctttccggc	AATTGCAATGGCTT	ctttatttct	(1)	MHMR*
			Recombinant 59308	gtataaaaaa	AATTCCAATGGCTT	ctttatttct	14 bp (0)	
Plastid reaction 6 with primers 111330REV and 121141FOR ^{e f}								
P6-1	CIP 0 μM	<i>why1why3</i>	109932	atgcaacaac	AAGATGTTATTTGTAAACAAGTAGTTTTGT	tggttggtta	(4)	MHMR*
			Recombinant 121745	tagatggacg	AAGATTTTCTTTG-AA-AAGTAGTTTTGT	tggttggtta	27 bp (0)	
P6-2	CIP 0 μM	<i>why1why3</i>	111043	cggataaatag	AAAGTCTTGTTATT	aagtccaaaa	(0)	MHMR*
			Recombinant 122213	actcttacat	AAAGTCTTGTTATT	aagtccaaaa	14 bp (1)	
P6-3a	CIP 0 μM	<i>why1why3</i>	19786	catatctcaa	TAGATTGGCGT-AATTTTT	ctaccgcatg	(0)	MHMR*
			Recombinant 110850	caaaattttg	TAGATTGGCGT-AATTTTT	ctaccgcatg	18 bp (3)	
P6-3b	CIP 0 μM	<i>why1why3</i>	19867	taattatggg	AGAAGTTGATCGTTAAAATGAT	aatttatgca	(0)	MHMR*
			Recombinant 122003	gccagccctt	AGAAGTTGATCGTTAAAATGAT	aatttatgca	22 bp (3)	
P6-4	CIP 0 μM	<i>why1why3</i>	111244	tagaaggagc	CACTTTTTTTTCCACT	gtaatttaga	(2)	MHMR*
			Recombinant 121304	aatttcagag	CAGTTTTTTTTTCACT	gtaatttaga	16 bp (0)	
P6-5	CIP 0.25 μM	Col-0	109932	atgcaacaac	AAGATGTTATTTGTAAACAAGTAGTTTTGT	tggttggtta	(4)	MHMR*
			Recombinant 121745	tagatggacg	AAGATTTTCTTTG-AA-AAGTAGTTTTGT	tggttggtta	27 bp (0)	
P6-6	CIP 0.25 μM	Col-0	109826	cctgaataat	CTCATTTTTTC	aattattcaa	(0)	MHMR
			Recombinant 121273	ggaatgaga	CTCATTTTTTC	aattattcaa	10 bp (0)	
P6-7	CIP 0.25 μM	Col-0	111043	cggataaatag	AAAGTCTTGTTATT	aagtccaaaa	(0)	MHMR*
			Recombinant 122213	actcttacat	AAAGTCTTGTTATT	aagtccaaaa	14 bp (1)	
P6-8	CIP 0.25 μM	Col-0	111244	tagaaggagc	CACTTTTTTTTCCACT	gtaatttaga	(2)	MHMR*
			Recombinant 121304	aatttcagag	CAGTTTTTTTTTCACT	gtaatttaga	16 bp (0)	

Tableau X (suite). DNA Rearrangement Products Identified in the Mitochondrial and Plastid Genomes

P6-9	CIP 0.25 μ M	<i>why1why3</i>	109932 Recombinant 121745	atgcaacaac AAGATGTTATTTGTAACAAGTAGTTTTGT tggttggtta tagatggacg AAGATTTTCCTTG-AA-AAGTAGTTTTGT tggttggtta tagatggacg AAGATTTTCCTTG-AA-AAGTAGTTTTGT cccatcggt	(4) 27 bp (0)	MHMR*
P6-10	CIP 0.25 μ M	<i>why1why3</i>	111244 Recombinant 121304	tagaaggagc CACTTTTTTTTCCACT gtaatttaga aatttcagag CAGTTTTTTTTTCACT gtaatttaga aatttcagag CAGTTTTTTTTTCACT catatataac	(2) 16 bp (0)	MHMR*
P6-11	CIP 0.75 μ M	Col-0	111043 Recombinant 122213	cggataatag AAAGTCTTGTATT aagtccaaaa actccttacat AAAGTCTTGTATT aagtccaaaa actccttacat AAAGTCTTGTATT gataattcaa	(0) 14 bp (1)	MHMR*
P6-12	CIP 0.75 μ M	Col-0	111244 Recombinant 121304	tagaaggagc CACTTTTTTTTCCACT gtaatttaga aatttcagag CAGTTTTTTTTTCACT gtaatttaga aatttcagag CAGTTTTTTTTTCACT catatataac	(2) 16 bp (0)	MHMR*
P6-13	CIP 0.75 μ M	<i>why1why3</i>	109932 Recombinant 121745	atgcaacaac AAGATGTTATTTGTAACAAGTAGTTTTGT tggttggtta tagatggacg AAGATTTTCCTTG-AA-AAGTAGTTTTGT tggttggtta tagatggacg AAGATTTTCCTTG-AA-AAGTAGTTTTGT cccatcggt	(4) 27 bp (0)	MHMR*
P6-14	CIP 0.75 μ M	<i>why1why3</i>	111244 Recombinant 121304	tagaaggagc CACTTTTTTTTCCACT gtaatttaga aatttcagag CAGTTTTTTTTTCACT gtaatttaga aatttcagag CAGTTTTTTTTTCACT catatataac	(2) 16 bp (0)	MHMR*
P6-15	NOV 0 μ M	<i>why1why3</i>	111189 Recombinant 121305	atTTTTtatta AGTTTTTTTTT cttcttcttt atttcagagc AGTTTTTTTTT cttcttcttt atttcagagc AGTTTTTTTTT tcactcatat	(0) 10 bp (0)	MHMR
Plastid reaction 7 with primers 50513REV and 59311FOR[†]						
P7-1	CIP 0.25 μ M	<i>why1why3</i>	50489 Recombinant 59938	aggttaatcg AATCATTTTTTTTC gtttggttgt cgtcccagca AATCATTTTTTTTC gtttggttgt cgtcccagca AATCATTTTTTTTC cacaagggat	(0) 14 bp (0)	MHMR
P7-2	CIP 0.75 μ M	Col-0	50489 Recombinant 59938	aggttaatcg AATCATTTTTTTTC gtttggttgt cgtcccagca AATCATTTTTTTTC gtttggttgt cgtcccagca AATCATTTTTTTTC cacaagggat	(0) 14 bp (0)	MHMR
P7-3	CIP 0.75 μ M	<i>why1why3</i>	50489 Recombinant 59938	aggttaatcg AATCATTTTTTTTC gtttggttgt cgtcccagca AATCATTTTTTTTC gtttggttgt cgtcccagca AATCATTTTTTTTC cacaagggat	(0) 14 bp (0)	MHMR
P7-4	CIP 0.75 μ M	<i>why1why3</i>	50462 Recombinant 59705	ctcttgagtc AAAACAAAAA ttcagggttaa agttactttg AAAACAAAAA ttcagggttaa agttactttg AAAACAAAAA aaaggaaagt	(0) 10 bp (0)	MHMR
P7-5	NOV 0 μ M	<i>why1why3</i>	50489 Recombinant 59938	aggttaatcg AATCATTTTTTTTC gtttggttgt cgtcccagca AATCATTTTTTTTC gtttggttgt cgtcccagca AATCATTTTTTTTC cacaagggat	(0) 14 bp (0)	MHMR
Plastid reaction 8 with primers 68821REV and 79920FOR[†]						
P8-1	CIP 0 μ M	<i>why1why3</i>	68809 Recombinant 81088	cctcgaaagc TTGATGCAAA taaacgaagt atctcttctt TTAATGCAAA taaacgaagt atctcttctt TTAATGCAAA ggacgaagta	(1) 10 bp (0)	MHMR*

Tableau X (suite). DNA Rearrangement Products Identified in the Mitochondrial and Plastid Genomes

P8-2	CIP 0 μ M	<i>why1why3</i>	68714 Recombinant 80522	ggcgaaaatc TCTTTTCTTTTA ccctatccc atcaacatgt TCTTTTCTTTTA ccctatccc atcaacatgt TCTTTTCTTTTA tttatatata	(0) 13 bp (0)	MHMR
P8-3	CIP 0.25 μ M	<i>why1why3</i>	68714 Recombinant 80522	ggcgaaaatc TCTTTTCTTTTA ccctatccc atcaacatgt TCTTTTCTTTTA ccctatccc atcaacatgt TCTTTTCTTTTA tttatatata	(0) 13 bp (0)	MHMR
P8-4	CIP 0.75 μ M	Col-0	68733 Recombinant 81329	cgatgagcc G aaactaaagc aatactcagg G aaactaaagc aatactcagg G gatcctttcc	1 bp	NHEJ- like
P8-5	CIP 0.75 μ M	Col-0	68714 Recombinant 80522	ggcgaaaatc TCTTTTCTTTTA ccctatccc atcaacatgt TCTTTTCTTTTA ccctatccc atcaacatgt TCTTTTCTTTTA tttatatata	(0) 13 bp (0)	MHMR
P8-6	CIP 0.75 μ M	<i>why1why3</i>	68733 Recombinant 82098	cgatgagcc G aaactaaagc tatattcatg G aaactaaagc tatattcatg G gttccgctcgt	1 bp	NHEJ- like
P8-7	CIP 0.75 μ M	<i>why1why3</i>	68714 Recombinant 80522	ggcgaaaatc TCTTTTCTTTTA ccctatccc atcaacatgt TCTTTTCTTTTA ccctatccc atcaacatgt TCTTTTCTTTTA tttatatata	(0) 13 bp (0)	MHMR
P8-8	NOV 0 μ M	<i>why1why3</i>	68522 Recombinant 80652	atccccttat TTTTT atgtcattcg tcaatacctg TTTTT atgtcattcg tcaatacctg TTTTT taatccgata	(0) 6 bp (0)	MHMR
P8-9	NOV 0 μ M	<i>why1why3</i>	68714 Recombinant 80522	ggcgaaaatc TCTTTTCTTTTA ccctatccc atcaacatgt TCTTTTCTTTTA ccctatccc atcaacatgt TCTTTTCTTTTA tttatatata	(0) 13 bp (0)	MHMR
P8-10	NOV 100 μ M	Col-0	68714 Recombinant 80522	ggcgaaaatc TCTTTTCTTTTA ccctatccc atcaacatgt TCTTTTCTTTTA ccctatccc atcaacatgt TCTTTTCTTTTA tttatatata	(0) 13 bp (0)	MHMR
P8-11	NOV 100 μ M	<i>why1why3</i>	68819 Recombinant 81100	atgcaaataa ACGAAGT tttgttctac aatgcaaagg ACGAAGT tttgttctac aatgcaaagg ACGAAGT aaaaaaataat	(0) 7 bp (0)	MHMR
P8-12	NOV 100 μ M	<i>why1why3</i>	68714 Recombinant 80522	ggcgaaaatc TCTTTTCTTTTA ccctatccc atcaacatgt TCTTTTCTTTTA ccctatccc atcaacatgt TCTTTTCTTTTA tttatatata	(0) 13 bp (0)	MHMR
Plastid reaction 9 with primers 10230REV and 30511FOR[†]						
P9-1	CIP 0 μ M	<i>why1why3</i>	9940 Recombinant 30859	agaataatcc GTTTCGTTTTT atactttctc agagggttaa GTTTCGTTTTT atactttctc agagggttaa GTTTCGTTTTT ttttactttc	(0) 12 bp (0)	MHMR
P9-2	CIP 0 μ M	<i>why1why3</i>	9948 Recombinant 30859	agaataatcc GTTTCGTTTTT-- <u>A</u> CTTTC tctgaagta agagggttaa GTTTCGTTTTTTTTACTTTC tctgaagta agagggttaa GTTTCGTTTTTTTTACTTTC cttgatagat	(3) 22 bp (0)	MHMR*
P9-3	CIP 0 μ M	<i>why1why3</i>	9895 Recombinant 30750	taaaaaatta ATTAAATTAAGAGTAAA attgactaaa aagtggaaat ATTAAATTAAGAGTAAA attgactaaa aagtggaaat ATTAAATTA <u>AAA</u> AATAA taaaaaaa	(0) 19 bp (2)	MHMR*

Tableau X (suite). DNA Rearrangement Products Identified in the Mitochondrial and Plastid Genomes

P9-4	CIP 0.25 μ M	Col-0	9948 Recombinant 30859	agaataatcc GTTTCGTTTTTT--ATACTTTC tcttgaagta agagggttaa GTTTCGTTTTTTTTTACTTTC tcttgaagta agagggttaa GTTTCGTTTTTTTTTACTTTC cttgatagat	(3) 22 bp (0)	MHMR*
P9-5	CIP 0.25 μ M	<i>why1why3</i>	9476 Recombinant 31012	aaattcacia TTGCCCTTTTTTTTTATTAGAGAAAATTTGTT atattagaat tatcaaaaat TTGCCCTTTTTTTTTATTAGAGAAAATTTGTT atattagaat tatcaaaaat TTG-CTTTTTTTTTTTTTAGAAAATATTGTT cagaggataa	(0) 32 bp (7)	MHMR*
P9-6	CIP 0.25 μ M	<i>why1why3</i>	9948 Recombinant 30864	agaataatcc GTTTCGTTTTTTATACTTTC tcttgaagta gttaagtttc GTTTCGTTTTTTATACTTTC tcttgaagta gttaagtttc GTTT--TTTTTT-TACTTTC cttgatagat	(0) 20 bp (3)	MHMR*
P9-7a	CIP 0.25 μ M	<i>why1why3</i>	9584 Recombinant 30885	ttttaacaat AGGAAAGTAAAAA gcgaaaagcg caatctatca AGGAAAGTAAAAA gcgaaaagcg caatctatca AGGAAAGTAAAAA aacgaaactt	(0) 16 bp (0)	MHMR
P9-7b	CIP 0.25 μ M	<i>why1why3</i>	31015 Recombinant 30829	ttttctaata AAAAAAAA gcaaattttt ggtctaacat AAAAAAAA gcaaattttt ggtctaacat AAAAAAAA aaagaaaaa	(0) 9 bp (0)	MHMR
P9-8	CIP 0.75 μ M	Col-0	9472 Recombinant 31015	cacaattgcc CTTTTTTTTTATTAGAGAAAATT tggtatatta caaaaatttg CTTTTTTTTTATTAGAGAAAATT tggtatatta caaaaatttg CTTTTTTTTT-TTTTAGAAAATT attggtcaga	(0) 23 bp (3)	MHMR*
P9-9	CIP 0.75 μ M	Col-0	9933 Recombinant 31046	tcaaaaatta AGAATAATCCGTTTC gttttttata tattggtcag AGGATAATCCGTTTC gttttttata tattggtcag AGGATAATCCGTA T C tcttcatcac	(1) 15 bp (1)	MHMR*
P9-10	CIP 0.75 μ M	<i>why1why3</i>	9933 Recombinant 31046	tcaaaaatta AGAATAATCCGTTTC gttttttata tattggtcag AGGATAATCCGTTTC gttttttata tattggtcag AGGATAATCCGTA T C tcttcatcac	(1) 15 bp (1)	MHMR*
P9-11	CIP 0.75 μ M	<i>why1why3</i>	9948 Recombinant 30859	agaataatcc GTTTCGTTTTTT--ATACTTTC tcttgaagta agagggttaa GTTTCGTTTTTTTTTACTTTC tcttgaagta agagggttaa GTTTCGTTTTTTTTTACTTTC cttgatagat	(3) 22 bp (0)	MHMR*
P9-12	CIP 0.75 μ M	<i>why1why3</i>	9883 Recombinant 30757	ctagaatttc TTAAAAAATAAATAAAA ttaaagagta aatattaaaa TTAAAAAATAAATAAAA ttaaagagta aatattaaaa TTAAAAAATAAATAAAA aaaatctaaa	(2) 18 bp (0)	MHMR*
Plastid reaction 10 with primers 60182REV and 50251FOR¹ – Deletion Events						
P10-1	CIP 0 μ M	<i>why1why3</i>	59060 Recombinant 50621	gctttttttg TACCTATTTTTTTATTCTATTTCTATTTCTATT cggttctatt atctgattaa TAACTAGTAATTTTTTTATTCTGTTTCTATT cggttctatt atctgattaa TAACTAGTAATTTTTTTATTCTGTTTATGTT taaaaaatt	(7) 33 bp (2)	MHMR*
P10-2	CIP 0 μ M	<i>why1why3</i>	59724 Recombinant 51123	gagttacttt GAAAACAAAAAAGGAAA gtcataatatt tagaaataaa GAATACAAAAAAGGAAA gtcataatatt tagaaataaa GAATACAAAAAAGTAAA aaaatttatc	(1) 19 bp (1)	MHMR*
P10-3	CIP 0 μ M	<i>why1why3</i>	59726 Recombinant 50847	ctttgaaaac AAAAAAAGGAAAGT catatattct gggatacaaa AAAAAAAGGAAAGT catatattct gggatacaaa AAAAAAAGGAAAGT attatgaact	(0) 15 bp (1)	MHMR*
P10-4	CIP 0 μ M	<i>why1why3</i>	59953 Recombinant 50474	cgtcccagca AATCATTTTTTTTC cacaagggat aggttaatcg AATCATTTTTTTTC cacaagggat aggttaatcg AATCATTTTTTTTC gtttggtgtg	(0) 14 bp (0)	MHMR

Tableau X (suite). DNA Rearrangement Products Identified in the Mitochondrial and Plastid Genomes

P10-5	CIP 0.25 μ M	<i>why1why3</i>	59724 Recombinant 51123	gagttacttt <u>GAAA</u> ACAAAAAAAAAGGAAA gtcatatattt tagaaataaa GAATACAAAAAAAAAGGAAA gtcatatattt tagaaataaa GAATACAAAAAAAAAGTAAA aaaatttatc	(2) 19 bp (2)	MHMR*
P10-6	CIP 0.25 μ M	<i>why1why3</i>	59726 Recombinant 50847	ctttgaaaac AAAAAAAAAAGGAAAGT catatatttct gggatacaaa AAAAAAAAAAGGAAAGT catatatttct gggatacaaa AAAAAAAAAAGAAAGT attatgaact	(0) 15 bp (1)	MHMR*
P10-7	CIP 0.25 μ M	<i>why1why3</i>	59726 Recombinant 50845	ctttgaaaac AAAAAAAA-GGAAAGT catatatttct tggggataca AAAAAAAA-GGAAAGT catatatttct tggggataca AAAAAAAAAGAAAGT attatgaact	(0) 16 bp (2)	MHMR*
P10-8	CIP 0.25 μ M	<i>why1why3</i>	60097 Recombinant 50836	tattttctgt TGGGGAT ttctctggaaa caagattgag TGGGGAT ttctctggaaa caagattgag TGGGGAT acaaaaaaaaa	(0) 7 bp (0)	MHMR
P10-9	CIP 0.25 μ M	<i>why1why3</i>	59953 Recombinant 50474	cgtcccagca AATCATTTTTTTTC cacaagggat aggttaatcg AATCATTTTTTTTC cacaagggat aggttaatcg AATCATTTTTTTTC gtttggttgt	(0) 14 bp (0)	MHMR
P10-10	CIP 0.75 μ M	Col-0	59846 Recombinant 51200	aattttctgct GGGC cttttatccta ttcttagtta GGGC cttttatccta ttcttagtta GGGC tatacggact	(0) 4 bp	NHEJ- like
P10-11	CIP 0.75 μ M	Col-0	59953 Recombinant 50474	cgtcccagca AATCATTTTTTTTC cacaagggat aggttaatcg AATCATTTTTTTTC cacaagggat aggttaatcg AATCATTTTTTTTC gtttggttgt	(0) 14 bp (0)	MHMR
P10-12	CIP 0.75 μ M	<i>why1why3</i>	58619 Recombinant 50471	catattttctt <u>TCGAATCATTTTTTTT</u> tatcgatatt ttcagggttaa <u>TCGAATCATTTTTTTT</u> tatcgatatt ttcagggttaa <u>TCGAATCATTTTTTTT</u> cgtttggttg	(1) 16 bp (0)	MHMR*
P10-13	CIP 0.75 μ M	<i>why1why3</i>	59398 Recombinant 50631	atcacctttt TTTTTT attttaaaa taactagtaa TTTTTT attttaaaaac taactagtaa TTTTTT tatttctggt	(0) 7 bp (0)	MHMR
P10-14	CIP 0.75 μ M	<i>why1why3</i>	59719 Recombinant 50843	tactttgaaa ACAAAAAAAA gaaagtcacat gagtggggat ACAAAAAAAA gaaagtcacat gagtggggat ACAAAAAAAA aaagaaagta	(0) 10 bp (0)	MHMR
P10-15	CIP 0.75 μ M	<i>why1why3</i>	59953 Recombinant 50474	cgtcccagca AATCATTTTTTTTC cacaagggat aggttaatcg AATCATTTTTTTTC cacaagggat aggttaatcg AATCATTTTTTTTC gtttggttgt	(0) 14 bp (0)	MHMR
P10-16	NOV 0 μ M	<i>why1why3</i>	59724 Recombinant 51123	gagttacttt <u>GAAA</u> ACAAAAAAAAAGGAAA gtcatatattt tagaaataaa GAATACAAAAAAAAAGGAAA gtcatatattt tagaaataaa GAATACAAAAAAAAAGTAAA aaaatttatc	(1) 19 bp (1)	MHMR*
P10-17	NOV 0 μ M	<i>why1why3</i>	59726 Recombinant 50847	ctttgaaaac AAAAAAAAAAGGAAAGT catatatttct ggggatacaaa AAAAAAAAAAGGAAAGT catatatttct ggggatacaaa AAAAAAAAAAGAAAGT attatgaact	(0) 15 bp (1)	MHMR*
P10-18	NOV 0 μ M	<i>why1why3</i>	59953 Recombinant 50474	cgtcccagca AATCATTTTTTTTC cacaagggat aggttaatcg AATCATTTTTTTTC cacaagggat aggttaatcg AATCATTTTTTTTC gtttggttgt	(0) 14 bp (0)	MHMR

Tableau X (suite). DNA Rearrangement Products Identified in the Mitochondrial and Plastid Genomes

P10-19	NOV 100 μ M	<i>why1why3</i>	59724 Recombinant 51123	gagttacttt <u>GAAA</u> ACAAAAAAAAAGGAAA gtcatatattt tagaaataaa <u>GAATAC</u> AAAAAAAAAGGAAA gtcatatattt tagaaataaa <u>GAATAC</u> AAAAAAAAAGTAAA aaaatttatc	(1) 19 bp (1)	MHMR*
P10-20	NOV 100 μ M	<i>why1why3</i>	59953 Recombinant 50474	cgtcccagca <u>AATCAT</u> TTTTTTC cacaagggat aggttaatcg <u>AATCAT</u> TTTTTTC cacaagggat aggttaatcg <u>AATCAT</u> TTTTTTC gtttggtgt	(0) 14 bp (0)	MHMR
Plastid reaction 11 with primers 31381REV and 19309FOR						
Plastid reaction 12 with primers 69633REV and 79920FOR						
Plastid reaction 13 with primers 50513REV and 68821FOR						
P13-1	CIP 0 μ M	Col-0	50427 Recombinant 69421	agtcaattcc <u>AATCGGAGTGG</u> ATTGG gcaagggata ttttctaaaa <u>AATCGGAGTGG</u> ATTGG gcaagggata ttttctaaaa <u>AAT</u> -GGAGTTGATTGG atttgcacca	(1) 16 bp (1)	MHMR*
P13-2	CIP 0 μ M	<i>why1why3</i>	50427 Recombinant 69421	agtcaattcc <u>AATCGGAGTGG</u> ATTGG gcaagggata ttttctaaaa <u>AATCGGAGTGG</u> ATTGG gcaagggata ttttctaaaa <u>AAT</u> -GGAGTTGATTGG atttgcacca	(1) 16 bp (1)	MHMR*
P13-3	CIP 0 μ M	<i>why1why3</i>	50394 Recombinant 69096	aatgtggaaa <u>AAATA</u> cccgcttagt agaattggat <u>AAATA</u> cccgcttagt agaattggat <u>AAATA</u> gtgggttcca	(0) 5 bp (0)	MHMR
P13-4	CIP 0.75 μ M	Col-0	50427 Recombinant 69421	agtcaattcc <u>AATCGGAGTGG</u> ATTGG gcaagggata ttttctaaaa <u>AATCGGAGTGG</u> ATTGG gcaagggata ttttctaaaa <u>AAT</u> -GGAGTTGATTGG atttgcacca	(1) 16 bp (1)	MHMR*
P13-5	CIP 0.75 μ M	<i>why1why3</i>	50427 Recombinant 69421	agtcaattcc <u>AATCGGAGTGG</u> ATTGG gcaagggata ttttctaaaa <u>AATCGGAGTGG</u> ATTGG gcaagggata ttttctaaaa <u>AAT</u> -GGAGTTGATTGG atttgcacca	(0) 16 bp (2)	MHMR*

^a Treatment of plants with ciprofloxacin (CIP) or novobiocin (NOV) at the indicated concentrations.

^b The junction endpoint is the position of the first nucleotide preceding the DNA junction. The accession numbers for the mitochondrial and plastid genomes are NC_001284 and NC_000932, respectively.

^c Alignment of the sequenced DNA junctions (center) with homologous sequences (top and bottom). Mismatches are underlined.

^d Product types were classified according to junction length with NHEJ-like for junction of less than five nucleotides, MHMR for junctions of five or more nucleotides and MHMR* for junctions of five or more nucleotides containing at least one mismatch.

^e These reactions were done twice with a ciprofloxacin treatment. Only nonredundant products were included in the table.

^f These reactions were done with plants treated with ciprofloxacin and with plants treated with novobiocin.

^g These primers anneal in the Large Inverted Repeats.

Tableau XI. Related DNA Rearrangements Detected in the Mitochondrial and Plastid Genomes

Same event in different genotypes or in different treatments
M7-1, M7-10
M11-1, M11-3
M12-1, M12-2, M12-4, M12-5
M12-3, M12-6
M16-1, M16-4
M18-2, M18-6
M18-1, M18-5
M20-1, M20-9
P1-2, P1-3, P1-4, P1-5, P1-6
P2-1, P2-4, P2-12, P2-20
P2-3, P2-6, P2-16, P2-22, P2-25
P2-5, P2-15, P2-23
P2-10, P2-21
P3-1, P3-6
P3-2, P3-8
P4-2, P4-5
P5-1, P5-2, P5-4
P6-1, P6-5, P6-9, P6-13
P6-2, P6-7, P6-11
P6-4, P6-8, P6-10, P6-12, P6-14
P7-1, P7-2, P7-3, P7-5
P8-2, P8-3, P8-5, P8-7, P8-9, P8-10, P8-12
P8-4, P8-6
P9-2, P9-4, P9-11
P9-9, P9-10
P10-2, P10-16, P10-19
P10-3, P10-6, P10-17
P10-4, P10-9, P10-11, P10-15, P10-18, P10-20
P13-1, P13-2, P13,4
Same event but with different mismatches
M3-6 vs M3-7
P9-6 vs (P9-2, P9-4 and P9-11)
(P10-2, P10-16 and P10-19) vs P10-5
P13-5 vs (P13-1, P13-2 and P13-4)
Events with at least one nearby breakpoint (less than 10 nts)
M15-1, M15-3, M15-4, M15-6
M20-7, M20-8
P9-1 vs (P9-2, P9-4 and P9-11)
P9-5 vs P9-8
(P10-4, P10-9, P10-11, P10-15, P10-18 and P10-20) vs P10-12
P10-8 vs P10-14

Tableau XII. DNA Rearrangement Statistics

Treatment	Type of DNA Rearrangement	Plastid Events ^a		Mitochondrial Events ^b	
		Col-0	KO1/3	Col-0	KO2
CIP 0 μ M	NHEJ-like	0	0	0	0
	MHMR	2	18	0	0
	Total Events	2	18	0	0
CIP 0.25 μ M	NHEJ-like	1	0	0	1
	MHMR	7	26	1	13
	Total Events	8	26	1	14
CIP 0.75 μ M	NHEJ-like	7	4	11	8
	MHMR	10	26	10	25
	Total Events	17	30	21	33
NOV 0 μ M	NHEJ-like	0	1	0	0
	MHMR	0	11	0	0
	Total Events	0	12	0	0
NOV 100 μ M	NHEJ-like	0	0	0	0
	MHMR	1	8	0	0
	Total Events	1	8	0	0

NHEJ-like are DNA junctions with four or less homologous base pairs. MHMR are DNA junctions with five or more homologous base pair. CIP relates to plant treated with ciprofloxacin whereas NOV relates to plants treated with novobiocin. DNA rearrangements containing multiple DNA junctions were counted as multiple events.

^a Plastid events are based on 13 primer pairs for ciprofloxacin treatment and 10 primer pairs for novobiocin treatment.

^b Mitochondrial events are based on 20 primer pairs for ciprofloxacin treatment and 10 primer pairs for novobiocin treatment.

Tableau XIII. DNA Rearrangement Products Detected in the Mitochondrial and Plastid Genomes and Used as Probes in DNA Gel Blots

Plant Treatment ^a	Plant Genotype	Junction Endpoints ^b	Junction Sequences ^c	Length of Junctions (Mismatches)	Junction Types ^d	Sizes of Hybridization Targets ^e
Plastid reaction A with primers 4434REV and 15643FOR						
CIP 0.75 μ M	KO1/3	4108 Recombinant 16052	cccaaattgg <u>AATTCTTT</u> -TTTTTTTTTC <u>gacatcaaaa</u> aataaagcta <u>AATTCTTT</u> TTTTTTTTTC <u>gacatcaaaa</u> aataaagcta <u>AATTCTTT</u> TTTTTTTTTC caaaataatg	(1) 19 bp (0)	MHMR*	4594 nt 913 nt (R) 6739 nt
Plastid reaction B with primers 49386REV and 68817FOR						
CIP 0.75 μ M	KO1/3	49165 Recombinant 69006	tttttatttc <u>CCCCAC</u> <u>acctttttta</u> ctctaacctt <u>CCCCAC</u> <u>acctttttta</u> ctctaacctt <u>CCCCAC</u> cactatTTTT	(0) 6 bp (0)	MHMR	5872 nt 514 nt (R) 4474 nt
Mitochondrial reaction A with primers 125781REV and 180454FOR						
CIP 0.75 μ M	KO2	125068 Recombinant 180866 46701	acttacctag <u>CGGAAGGAAAGGG</u> <u>ctaggaccta</u> tggagactga <u>CGGAAGGAAAGGG</u> <u>ctaggaccta</u> tggagactga <u>CGGAAGGAAAGGG</u> gttcctcct tggagactga <u>CGGAAGGAAAGGG</u> gttcctcct	(0) 13 bp (0) (0)	MHMR	3887 nt 4862 or 5155 nt (R) 2345 nt 2638 nt
Mitochondrial reaction B with primers 187018REV and 349054FOR						
CIP 0.75 μ M	KO2	335715 Recombinant 349869	cgttgatttg <u>ACTTCACATAGCTTTTCGTCCTTGATAGCTGGAAG</u> <u>agaacgttga</u> ttcttctttt <u>ACTTCACATAGCTTTTCGTCCTTGATAGCTGGAAG</u> <u>agaacgttga</u> ttcttctttt <u>ACTTCACATAGCTTTTCGTCCTTGATAGCTGGAAG</u> ttctcaaaa	(0) 37 bp (0)	MHMR	3764 nt 2914 nt (R) 6566 nt

^a Treatment of plants with ciprofloxacin (CIP) or novobiocin (NOV) at the indicated concentrations.

^b The junction endpoint is the position of the first nucleotide preceding the DNA junction. The accession numbers for the mitochondrial and plastid genomes are NC_001284 and NC_000932, respectively.

^c Alignment of the sequenced DNA junctions (center) with homologous sequences (top and bottom). Mismatches are underlined.

^d Product types were classified according to junction length with NHEJ-like for junction of less than five nucleotides, MHMR for junctions of five or more nucleotides and MHMR* for junctions of five or more nucleotides containing at least one mismatch.

^e Sizes of HindIII-digested DNA molecules that are expected to hybridize with the PCR product. R indicates the recombinant form.

Tableau XIV. Oligonucleotides Used in this Study

Reaction	Oligonucleotide Name	Oligonucleotide Sequence
Mitochondrial PCR primers		
1	20041REV	GCG CAT ACC TTA TCC GAT CGT
	30001FOR	CTT CCG CTA TTG GTG GAC TGT
2	30001REV	ACA GTC CAC CAA TAG CGG AAG
	50341FOR	GAG CTA GAT AGC ACG GCT TTG
3	55021REV	AAG GCT AGT CGT CAA CGA GCA
	65041FOR	GGA TCT GAG CTG TAG AAA CCC
4	130413REV	GCA CAG TCG AGT GCT TTA CCT
	151514FOR	CTG GAT CTC GCT CAT GTC TCT
5	180454REV	GTT TAC ATG GGG CTG GTA GGT
	200500FOR	CCC ACT GAG TTG GTA AGC AAG
6	235252REV	AAA CAG TGC GGT GAG CTC TAG
	245966FOR	GCA AGA TGA TCA GTC CGA GAG
7	305346REV	GTT TCG CAG CAC GGA GTA TAC
	315513FOR	CAT CTG AGG CAC CTC TAT TGG
8	165904REV	TGT ACT GGC TTC CGC ACT ACT
	200500FOR	CCC ACT GAG TTG GTA AGC AAG
9	155034REV	CTG ATG GTG TTC GTG GTA GGT
	200500FOR	CCC ACT GAG TTG GTA AGC AAG
10	10081REV	CAT GTC GAA GGA ACT CGG CAA
	15001FOR	CCG GAA GGT ATA GTG TGA AGG
11	20461REV	GCG CAT ACC TTA TCC GAT CGT
	24961FOR	TGA TTC CGT CTG ACT CCT CCT
12	30001REV	ACA GTC CAC CAA TAG CGG AAG
	80161FOR	ACG TGC AAG TTT CCC TGC ATG
13	84901REV	CAA AAC ACT CCA GAT CCG CAC
	151514FOR	CTG GAT CTC GCT CAT GTC TCT
14	210239REV	GGA AGG CTT TGT CTA CGG TCT
	245966FOR	GCA AGA TGA TCA GTC CGA GAG
15	1REV	GGC TAA CCT GTT TCG AAC GGA
	4981FOR	CCA GGA AAG CCT AGC TAC TTC
16	125781REV	CCA GAC TCG GAT GCT AGA GAA
	115816FOR	TTA CCT ACC CGG CTA CTG ACA
17	135180REV	ATC TCT AAG TGC TCG CAG CAC
	140053FOR	TTC CCT ATT CCT CGC CTA CTC
18	171214REV	CAT TCT AGC CCG AGA GGA ACT
	180454FOR	ACC TAC CAG CCC CAT GTA AAC
19	245966REV	CTC TCG GAC TGA TCA TCT TGC
	239584FOR	AGG GCT ACC TTT CTT CGA CTG
20	270785REV	ACA CAA ACT GAC CGA GCT AGC

	275583FOR	GCA CAG TGG ATT CAT CTA GGG
<i>cox1</i>	cox1REV	GTA ACG TCC GTT CCG TGA TCT
	cox1FOR	GCT AGC TCA TGG CAG GAA ATC
A	125781REV	CCA GAC TCG GAT GCT AGA GAA
	180454FOR	ACC TAC CAG CCC CAT GTA AAC
B	187018REV	AAG AAG CTA CCT CCT TCC GTG
	349054FOR	AAT CCG ATA GCT CTA CCA GCC
Repeat A	AFOR	CAC TGA ACC GAC TTG AAT CTG AAC
	AREV	GAA AGC CAT TTG CAC CAG GTT TAT
Repeat D	DFOR	AGT GAT CTG TTC ATC TAA CTC A
	DREV	TAC TAC TAC CTC GTC CAT TG
Repeat F	FFOR	CAC GAG GAA TGG AAA GAA ACA T
	FREV	GCG CAC AAA CCA CTC TAA AG

Plastid PCR primers

1	31138REV	AAT AGC CTA GCG CAT CGA ACG
	19181FOR	CAG CAT CTG ATG TGC GTA CAG
2	40130REV	GGG AAG AGG AGG AAC ATG TCA
	50251FOR	GGC CCC TTC CTA ATT GGA GAT
3	84761REV	CCT TAG GCA CGG CCA TAC ATA
	95260FOR	AAC CTG CTA GTG GAG GAA GAC
4	105961REV	GCG TTA CTC AAG CCG ACA TTC
	116007FOR	AAA CCA CTC ATG CCC GGC AAT
5	69633REV	GAG GAA CAT GTA CTC GGG TGT
	59291FOR	AGG CCT AGT CTT TCC GGG AAT
6	111330REV	TCA ACA GCG GGA TTA ACC GCA
	121141FOR	AAC GTC ATC CGG TTA CCG GTT
7	50513REV	GAG ACG TCT ACC ACA ACA ACC
	59311FOR	AGG CCT AGT CTT TCC GGC AAT
8	68821REV	ACG CGG ATA TAT AGC TCG GAG
	79920FOR	CCA TGC CGT TCG GAA GTG ATT
9	10230REV	TGG CAA GAG GTC AAC GAT TGC
	30511FOR	GCT GCC TCC TTG AAA GAG AGA
10	60182REV	ATA TAA AGG ACA CGC CGG GCA
	50251FOR	GGC CCC TTC CTA ATT GGA GAT
11	31381REV	CCG ATG ACT TAC GCC TTA CCA
	19309FOR	TCG TCT GCT AAG ACA CGA CCA
12	69633REV	GAG GAA CAT GTA CTC GGG TGT
	79920FOR	CCA TGC CGT TCG GAA GTG ATT
13	50513REV	GAG ACG TCT ACC ACA ACA ACC
	68821FOR	CTC CGA GCT ATA TAT CCG CGT
A	4434REV	TTT ATC CCT GAC GCA CAC ACG
	15643FOR	CCA TTT CCT CGT ACC ATG TCC
B	49386REV	TGT ACG CAG TCC GCA TAT TGG
	68817FOR	GTT TTG TTC TAC GCC TCC GAG

<i>ycf2</i>	ycf2REV	TGT	TTC	GCC	TCT	TAC	TCG	GAG
	ycf2FOR	GAT	CTC	TGA	GAG	CTG	TTT	CCG

Oligonucleotides Used for *in vitro* Assays.

YG	GGG	CGG	CGG	GGG	CGG	GGC	GG			
YC	CCG	CCC	CGC	CCC	CGC	CGC	CC			
YG8_3	GGG	CGG	CGG	GGG	CGG	GGC	GGT	TTT	TTT	T
YC8_5	TTT	TTT	TTC	CGC	CCC	GCC	CCC	GCC	GCC	C
YG16_3	GGG	CGG	CGG	GGG	CGG	GGC	GGT	TTT	TTT	
	TTT	TTT	TTT							
YC16_5	TTT	TTT	TTT	TTT	TTT	TCC	GCC	CCG	CCC	
	CCG	CCG	CCC							
R1	TAT	CGA	ATC	CGT	CTA	GTC	AAC	GCT	GCC	
	GAC	ATC	TAC	GAT	TGA	CGT	TTG	GGC	TCC	
	TCG	CAA	GTC	ATT	GG					
R2	CCA	ATG	ACT	TGC	GAG	GAG	CCC	AAA	CGT	
	CAA	TCG	TAG	ATG	TCG	GCA	GCG	TTG	ACT	
	AGA	CGG	ATT	CGA	TA					

CHAPITRE 4. Troisième article

Divergent roles for the two Poll-like organelle DNA polymerases of *Arabidopsis thaliana*

Article soumis à la revue *The Plant Journal*, le 20 décembre 2010.

4.1 Mise en contexte

Plusieurs modèles illustrant la réplication de l'ADN des organelles de plantes ont été proposés (revue dans (11, 22)). Il est cependant décevant de constater que ces modèles n'incluent aucun facteur protéique car même la polymérase responsable de la synthèse de l'ADN n'a pas encore été identifiée. Plusieurs études récentes ont cependant proposé que la famille de protéines homologues des polymérases de l'ADN bactérienne de type I (ou Poll) sont les enzymes responsables de cette activité de synthèse (29-31). Cette hypothèse n'est cependant basée que sur des mesures *in vitro* utilisant des protéines recombinantes et aucun argument *in vivo* n'a été apporté. Nous avons donc entrepris l'isolation et la caractérisation de mutants des deux gènes de cette famille chez *Arabidopsis thaliana* dans le but de tester cette hypothèse.

Cette étude a débuté avec l'isolation de mutants pour les gènes *PollA* et *PollB*, encodant les Poll chez *Arabidopsis*. Mon rôle spécifique a été l'isolation et la caractérisation des lignées mutantes ainsi que l'évaluation des niveaux d'ADN à l'intérieur des organelles. Étienne Lepage a participé à toutes les étapes de la réalisation de l'étude. Spécifiquement, É.L. s'est chargé de faire croître les plantes *in vitro* en présence ou absence de stress, il a également développé le protocole permettant d'évaluer la présence de bris double brin dans l'ADN des organelles. É.L. a réalisé les figures, puis écrit et révisé cet article avec Normand Brisson et moi-même.

Divergent roles for the two Poll-like organelle DNA polymerases of *Arabidopsis thaliana*

Jean-Sébastien Parent^{1,2}, Etienne Lepage¹ and Normand Brisson^{*}

Department of Biochemistry, Université de Montréal, C.P. 6128, PO Box Centre-Ville,
Montréal, Qc, Canada H3C 3J7

¹These authors contributed equally to this work.

²Current address : Institut Jean-Pierre Bourgin, Institut National de la Recherche
Agronomique, 78026 Versailles Cedex, France.

*For correspondence

Suggested running title: **Roles of Poll-like DNA polymerases in *Arabidopsis***

Key words: DNA polymerase, DNA replication, DNA repair, plastid, mitochondria,
Whirly, *Arabidopsis thaliana*

Accession number and polymorphism

PollA gene: At1g50840

PollB gene: At3g20540

polla-1 : SALK_022638

polla-2 : SALK_065221

pollb-1 : SALK_134274

pollb-2 : WiscDsLoxHs021_09D

Total word count: 5802

4.2 Summary

DNA polymerases play a central role in the process of DNA replication. Yet, the proteins in charge of the replication of plant organelle DNA have not been unambiguously identified. There are however many indications that a family of proteins homologous to bacterial DNA polymerase I is implicated in organelle DNA replication. Here, we have isolated mutant lines of the *PollA* and *PollB* genes of *Arabidopsis thaliana* to test this hypothesis. We find that mutation of both genes is lethal, thus confirming an essential and redundant role for these two proteins. However, the mutation of a single gene is sufficient to cause a reduction in the levels of DNA in both mitochondria and plastids. We also demonstrate that *pollb*, but not *polla* mutant lines, are hypersensitive to ciprofloxacin, an antibiotic that specifically induces DNA double-strand breaks in plant organelles, suggesting a function for *PollB* in DNA repair. In agreement with this result, a cross between *pollb* and a plastid Whirly mutant line yielded plants with high levels of DNA rearrangements and severe growth defects, indicating impairments in plastid DNA repair pathways. Taken together, this work provides further evidences for the involvement of the plant *Poll*-like genes in organelle DNA replication and suggests an additional role for *PollB* in DNA repair.

4.3 Introduction

Faithful replication of DNA is one of the most crucial biochemical steps in living organisms. Plant cells are particular in this regard because of their need to maintain genomes in three separate compartments: the nucleus, the plastid and the mitochondria. While nuclear DNA replication and repair have been extensively studied, little is known about these processes in plant organelles. For example, the DNA polymerases responsible for the replication of the genome of mitochondria and plastids have not yet been clearly identified.

In higher plants, DNA polymerase activity has been purified from chloroplast extracts of *Spinacia oleracea* (spinach) (Sala et al., 1980; Spencer and Whitfeld, 1969) and *Pisum sativum* (pea) (McKown and Tewari, 1984), while a mitochondrial DNA polymerase activity has been purified from *Triticum aestivum* (wheat) (Castroviejo et al., 1979) and the flowering plant *Chenopodium album* (Meissner et al., 1993). The DNA-polymerase activities from these various sources have similar characteristics, including insensitivity to aphidicolin, a known inhibitor of nuclear eukaryotic α DNA polymerases, a broad alkaline pH optimal activity and the requirement for a divalent metal cation (Mg^{2+} or Mn^{2+}). This suggests that plant organelles, like animal mitochondria, possess their own specialized enzymes, with distinct characteristics, for the replication of their genome (Sala et al., 1980).

Isolation of DNA polymerases from the chloroplast and mitochondria of *Glycine max* (soybean) and *Nicotiana tabacum* (tobacco) cell cultures allowed further comparison of the two enzymes contained in these compartments (Heinhorst et al., 1990; Sakai, 1999). Interestingly, the biochemical properties and molecular weights of these enzymes were shown to be very similar. They reach maximal activity under the same experimental conditions and are sensitive to the same chemical agents, suggesting that they are at least homologous proteins. These enzymes were also shown to share characteristics with the γ DNA polymerases (Poly) found in the mitochondria of animals (Heinhorst et al., 1990; Sakai, 1999), such as insensitivity to aphidicolin, requirement for KCl and template preference. Another important characteristic shared by these purified enzymes and Poly is that they possess a 3'-5' exonuclease activity, as demonstrated for the DNA polymerases from spinach chloroplasts (Keim and Mosbaugh, 1991) and *C. album* mitochondria (Meissner et al., 1993).

More recently, genes encoding proteins targeted to organelles and with similarity to the gene coding for *Escherichia coli* DNA polymerase I (Poll) were identified in *Oryza sativa* (rice) (Kimura et al., 2002), *Arabidopsis thaliana* (Mori et al., 2005) and tobacco (Ono et al., 2007). Full-length cDNAs were cloned and regions corresponding to the DNA polymerase domains of these enzymes were produced in bacteria and tested for template preference, biochemical properties and sensitivity to inhibitors. Like the enzymes purified from plants,

the recombinant proteins are insensitive to aphidicolin and dideoxynucleotides (ddNTP) (Kimura et al., 2002; Mori et al., 2005; Ono et al., 2007), suggesting that these DNA *Poll*-like enzymes are responsible for DNA replication in plants organelles.

It therefore appears that a conserved family of bacteria-like DNA polymerases is responsible for the replication of DNA inside plant organelles. Indeed, every plant genome that has been sequenced to date possesses two *Poll-like* genes which, in *Arabidopsis* and tobacco, encode proteins that localize to both organelles, thus suggesting an important conserved function (Christensen et al., 2005; Elo et al., 2003; Ono et al., 2007; Wamboldt et al., 2009). In addition, the recombinant proteins have a high processivity, which would allow them to replicate complete organelle genomes (Kimura et al., 2002; Mori et al., 2005; Ono et al., 2007). All these characteristics make the *Poll* proteins ideal candidates for a role in organelle DNA replication. However, no formal proof has yet been obtained that confirms this function *in vivo*. Here we report the isolation of mutant lines for the two *Arabidopsis* *Poll-like* genes *At1g50840* and *At3g20540* (named *PollIA* and *PollIB* henceforth). While no visible phenotype was observed with two mutant alleles of *PollIA*, a strong mutant allele of *PollIB* yielded plants that grew more slowly than wild-type (WT) plants. More important, we were unable to recover plants bearing homozygous mutations in both genes, indicating that they play an essential, at least partially redundant role in *Arabidopsis*. All mutants tested showed reduced levels of organelle DNA relative to nuclear DNA, suggesting that they are involved in genome replication. Interestingly, only the mutants for *PollIB* showed increase sensitivity to the organelle-specific DNA double-strand break-inducing agent ciprofloxacin, suggesting that one of the two *Poll* proteins has specialized in DNA repair.

4.4 Results

***PollIA* and *PollIB* share a common essential function in organelles**

To provide genetic evidence for the implication of the *Arabidopsis* *PollIA* and *PollIB* genes in organelle DNA replication, we have obtained two T-DNA insertion mutant lines for each gene. Figure 1a shows the position of the insertion in each line. Figure 1b shows *PollIA*

mRNA levels, measured by qRT-PCR, in the *polla* lines. The two mutations in *PollA* caused a marked reduction in the expression level of the gene. The *polla-1* line (SALK_022638) showed a 536 fold reduction and the *polla-2* line (SALK_065221) a 12 fold reduction in mRNA level as compared to WT. Mutations in the *PollB* gene had a different effect depending on the region of the gene that was amplified. Figure 1c shows the results of two qRT-PCR reactions using primers located either in the middle of the gene or at the 3' end (see Figure 1a). The *pollb-1* line (SALK_134274) showed a 33 fold reduction in expression in the mid-gene section but only a 3 fold reduction near the end of the gene. The *pollb-2* line (WiscDsLoxHs021_09D) resulted in a 1.5 fold reduction of the mid-gene section and a 365 fold reduction at the 3' end. This difference in signal level between two parts of the same messenger could be due to differential processing of the mRNA caused by the T-DNA insertion, or to leaky transcription from the 35S promoter inside the T-DNA insert. This could lead to different results when using probes placed on either side of the insertion site (as it is the case for *pollb-2*).

The phenotype of the two *polla* lines did not differ from WT (Figure 1d). Interestingly, while the *pollb-1* mutation also yielded no differential phenotype, the *pollb-2* line showed growth retardation (Figure 1d). This difference in phenotype compared to *pollb-1* might be attributed to the lower expression of *PollB* (maximal reduction of 365 fold vs 33 fold for the *pollb-1* line) or to the site of the insertion (5th exon vs 1st intron). We then attempted to cross *polla-1* with *pollb-1*, and *polla-2* with *pollb-2* to obtain mutant plants devoid of Poll proteins. We were unable to isolate double homozygous mutant plants in the progeny of double heterozygous plants. By testing the germination of seeds from the cross on MS medium, we did not observe any aborted seed and we concluded that the double homozygous individuals simply do not produce mature seeds. We went a step further for the *polla-1/pollb-1* cross and isolated one plant homozygous for *pollb-1*, but heterozygous for *polla-1*. We then tested 48 plants from the progeny of this plant and again could not recover any double homozygous plant. The significance of this result was confirmed by a χ^2 test ($p=0$). These results indicate that at least one intact *Poll* gene is essential for plant viability. Therefore, in the rest of our work we focused on the analysis of single mutant plants.

Mutants for the *Poll* genes have less organelle DNA

Mutants for DNA polymerase genes are expected to have a reduced capacity to synthesize and accumulate DNA. We therefore measured the relative level of organelle to nuclear DNA in the *Poll* mutants in order to evaluate plastid genome copy number in *Arabidopsis* (Zoschke et al., 2007; Rowan et al., 2009). WT, *polla-1* and *pollb-1* mutant lines were grown on MS medium, whole seedlings (including the roots) were harvested at 3, 4 and 5 days after germination and their total DNA isolated. The level of DNA was measured by qPCR using primers specific for a given DNA sequence inside the nucleus, the mitochondria or the plastid. The levels for the mitochondria and plastid DNA (mtDNA and ptDNA) are reported as a ratio to the nuclear DNA level. Figure 2a shows that, in 3 and 4 day old seedlings, the relative level of mtDNA is lower in both *polla-1* and *pollb-1* as compared to the WT plants, but that this difference is less pronounced by day 5. A similar trend is observed when testing ptDNA, although in this case the difference between WT and *pollb-1* is still pronounced at day 5 (Figure 2b). Similar results were obtained with the *polla-2* line (Figure S1 in Supporting Information). The ratios for the WT to mutant levels are presented in Supplementary Table 1. However, it was not possible to test the *pollb-2* line at these early time points due to its growth retardation defect. These results show that PollA and PollB are both required for maximal DNA accumulation in *Arabidopsis* plastids and mitochondria.

Mutants for the *PollB* but not *PollA* gene are impaired in their capacity to repair DNA double-strand breaks

DNA-synthesis activity is required for the repair of DNA damages by mechanisms such as homologous recombination (HR). To test for a role in DNA repair, the *Poll* mutants were treated with ciprofloxacin, a molecule that induces DNA double-strand breaks (DSBs) specifically in plant organelle DNA (Wall et al. 2004) and was recently shown to affect ptDNA conformation (Rowan et al. 2010). Interestingly, only the *pollb* mutants showed increased sensitivity to ciprofloxacin when compared to WT plants (Figure 3a). At 0.25 μ M of ciprofloxacin, *pollb-2* plants were more sensitive to the drug than *pollb-1* plants. In this condition, the *pollb-1* plants were able to produce small first true leaves with white regions while *pollb-2* plants produced abnormal first true leaves that were completely

white (Figure 3b). At 0.5 μ M ciprofloxacin very few true leaves were observed for either *pollb* alleles. Figure 3c and d present, for each line and condition tested, the percentage of plants showing first true leaves with white regions or no first true leaves, respectively.

These results suggest that in addition to a role in DNA replication, PolIB is also required for the repair of DNA. In that case one can predict that the DSBs induced by ciprofloxacin will accumulate to higher levels in *pollb-1* plants than in WT or *polla-1* plants. This was tested by measuring the level of DSBs in plastid and mitochondrial DNA using a PCR approach (Yakes and Van Houten, 1997). Figure 3e and f show a clear, dose dependent decrease in the relative amplification of a long DNA fragment when using *pollb-1* DNA isolated from plants treated with ciprofloxacin as template. This indicates the presence of a higher level of DSBs in the organelle genomes of these plants as compared to WT and *polla-1* plants. Although a similar trend was observed when comparing the results for ptDNA (Figure 3e) and mtDNA (Figure 3f), the difference with WT DNA was less marked when testing mtDNA. This difference was nevertheless significant ($p < 0.05$ in Student t-test). To confirm that the DSBs induced by ciprofloxacin are specific to organellar DNA, we designed a similar experiment with nuclear probes and found, as expected, no increase in DSBs in WT and *pollb-1* plants treated with 0.5 μ M ciprofloxacin (Figure S2 in Supporting Information). Taken together, these results indicate a role for PolIB in DNA repair in both the mitochondria and the plastid.

Genetic interaction between *PolIB* and *Whirly* genes

The above results suggest that PolIB is specifically recruited at DSB sites and contribute to their repair. In prokaryotes, the recruitment of repair proteins to the DNA is orchestrated by the single-stranded DNA-binding proteins (SSBs) (reviewed in (Shereda et al., 2008)). We have recently reported that the single-stranded DNA binding proteins Whirly are required for accurate repair of DSBs in plant organelles and proposed that their function in this process is comparable to that of SSBs in prokaryotes (Cappadocia et al., 2010; Maréchal et al., 2009). To test whether a genetic interaction exists between the Whirly and the Poll genes, we crossed the *polla-1* and the *pollb-1* mutants with the mitochondria (*why2-1*) or the plastid (*why1why3*) Whirly mutants reported previously (Cappadocia et al., 2010; Maréchal et al., 2009). All but one of the crosses yielded plants with no

difference in phenotype when compared to WT plants, except for the occasional appearance of variegation in lines bearing the *why1why3* mutations (Maréchal et al., 2009) (Figure 4a and S3 in Supplementary Information). Interestingly, the *pollb-1/why1why3* homozygous triple mutant exhibited a severe growth defect and pale leaves, indicating that a synergistic interaction exists between plastid Whirlies and PollB.

The *why1why3* mutant line of *Arabidopsis* has been shown to promote the repair of DSBs through an error-prone microhomology-mediated break-induced replication (MMBIR) pathway (Cappadocia et al., 2010) that leads to major genome rearrangements (see (Hastings et al., 2009; Maréchal and Brisson, 2010) for reviews). About 5% of the *why1why3* plants also exhibit a variegated leaf phenotype that correlates with selectively amplified genomic regions (Maréchal et al., 2009). As the results in Figure 3e show that the amount of DSBs is increased in ptDNA from ciprofloxacin-treated *pollb-1* plants, it suggests that the severe phenotype of the *pollb-1/why1why3* plants is also linked to an accumulation of rearranged DNA products in the plastid genome. This was tested by monitoring the extent of genome rearrangements in the mutant lines using eight primer pairs employed in our previous study to survey DNA rearrangements in Whirly mutant plants (Cappadocia et al., 2010). These primers are either in an opposite orientation or placed far apart (more than 5 kb) from each other (Figure 4c) and amplify respectively junctions of duplicated/circular or deleted DNA regions produced by the MMBIR pathway in plastids. Figure 4b shows the results of representative PCR reactions with three primer pairs. A graphic representation of the main rearranged molecule is illustrated for each reaction in Figure 4d. Untreated WT and *pollb-1* plants show no detectable rearrangements, while the same genotypes treated with 0.25 μ M ciprofloxacin show only few rearrangements (Fig. 4b). The *why1why3* plants contain a constitutive level of rearranged DNA molecules and, as reported before (Cappadocia et al., 2010), this level increases noticeably after treatment with 0.25 μ M ciprofloxacin. The level of DNA rearrangements is however markedly higher in the untreated triple mutant as compared to untreated *why1why3* plants. This level is actually similar to that of the treated *why1why3* plants and appears to further increase upon treatment with the drug.

Therefore, the severe phenotype of the *polb1-1/why1why3* plants correlates with increased rearrangements of the plastid genome.

4.5 Discussion

An essential role for the *Poll* genes in higher plants

Plant organelle DNA replication is still a highly misunderstood process (Bendich, 2007; Nielsen et al., 2010). This is largely due to our ignorance concerning some of the most basic factors, such as DNA polymerases, that contribute to this fundamental process in plants. Many studies have suggested that the Poll-like DNA polymerases found in higher plants are responsible for most if not all the DNA-synthesis activity found in mitochondria and plastids (Kimura et al., 2002; Mori et al., 2005; Ono et al., 2007). This is based on the biochemical properties of the recombinant proteins, which indicate that these enzymes have sufficient processivity to be able to replicate both genomes. In addition, the measured in vitro characteristics of these enzymes correspond almost perfectly to those measured ex vivo from the purified DNA-synthesis activity of mitochondria and plastids (Heinhorst et al., 1990; McKown and Tewari, 1984; Meissner et al., 1993; Sala et al., 1980; Spencer and Whitfeld, 1969). All these elements, plus the fact that they are present in both organelles (Christensen et al., 2005; Elo et al., 2003; Mori et al., 2005; Ono et al., 2007; Wamboldt et al., 2009), suggest that the Poll DNA polymerases are responsible for DNA replication in plant organelles.

In the present study, we bring further arguments to support a role for Poll proteins in the replication of plastid and mitochondria DNA. Using *Arabidopsis* mutant plants for the *PollA* and *PollB* genes, we show that concurrent mutation of the two genes is lethal, indicating that they share an essential function in plants. Furthermore, each gene is able to partially compensate for the function of the other gene as all single Poll mutants tested are viable. However, the fact that the *pollb-2* line displays a growth retardation phenotype suggests that *PollA* cannot entirely compensate for the function of *PollB*, or that *PollB* carries another function not shared by *PollA*.

A role for the *Poll* proteins in DNA replication is also supported by our results showing that, as compared to WT plants, the relative level of organelle DNA is reduced in plastids and mitochondria of both *polla-1* and *pollb-1* plants. The reduction is most pronounced at 3 days post-germination, which is consistent with expression data showing that the *Poll* genes are most highly expressed in young, meristematic tissues (Kimura et al., 2002; Mori et al., 2005; Ono et al., 2007). Also, in line with our observations, previous studies have shown that DNA replication in organelles is highly active in 3 day old Arabidopsis seedlings (Fujie et al., 1994; Fujie et al., 1993). Our results therefore provide a functional link between the *Poll* genes and DNA accumulation inside plant mitochondria and plastids. Although we cannot rule out another role for *Poll* proteins in DNA metabolism, such as in DNA degradation (Rowan et al., 2009), our results are more easily interpreted in view of the current literature indicating that these polymerases are actually involved in DNA synthesis.

Interestingly, the ratios of organelle to nuclear DNA appear to decrease in WT plants between day 3 and 5 (Figure 2 and S1). This could either mean that nuclear DNA ploidy is increasing or that organelle DNA is less abundant in cells of older seedlings. However, since WT and mutant plants of the same age are being compared, any variations in nuclear or organellar DNA levels should not affect our results. It is interesting that a previous study reported a similar trend for ptDNA ratios in aging cotyledons of Arabidopsis (Zoschke et al., 2007), although different samples were used in the two studies (whole seedling with roots vs. cotyledon alone).

A specific role for PolIB in DNA DSB repair

DNA repair activity is essential for correcting DNA damages caused by exogenous stress or that occur naturally during replication. Many mechanisms used to repair DNA, such as HR and MMBIR, involve a DNA-synthesis step. Since, as mentioned above, *PollA* does not fully compensate for the loss of *PolIB* in the *pollb-2* line, there is a possibility that *PolIB* could be involved in an additional process, such as synthesis of DNA during repair. This was tested by inducing DSBs in the organelle genome of all the *Poll* mutants and by measuring the effect of this stress on plant phenotype and on the level of DSBs. The fact that the

growth of the *pollb* lines, but not of the *polla* lines, was more affected by the genotoxic stress provides a first evidence for a role of PolIB in DNA repair. Indeed, increased sensitivity to ciprofloxacin has also been reported for mutants of Whirly proteins (Cappadocia et al. 2010) and cpRecA (Rowan et al. 2010), which are both implicated in the maintenance of plastid genome stability. A role in DNA repair for PolIB is further supported by the observation that the *pollb-1* line accumulates higher levels of DSBs than the *polla-1* line or WT plants.

The higher levels of DSBs in plastids suggest a specialized role for PolIB in this organelle. However, it could also simply reflect the presence of alternate repair mechanisms more active in the mitochondria. For example homologous recombination between short repeats (>100 bp and <1000 bp, reviewed in (Maréchal and Brisson, 2010)), a process very active in plant mitochondria, would provide an efficient mean for the repair of DSBs. Also, we cannot exclude the possibility that ciprofloxacin could be less active in mitochondria. The different response of the *polla* and *pollb* mutant lines to the genotoxic stress is surprising, considering that the two Poll proteins have over 72% identity at the amino acid sequence level (Mori et al. 2005) and that they are nearly identical regarding their 3'-5' exonuclease and DNA polymerase domains. Therefore our results indicate that PolIB has either evolved a specific repair activity or that it is specifically recruited to damage sites. This additional function in DNA repair might also explain why one of the *PolIB* mutants shows a growth retardation phenotype while both *PolIA* mutants are undistinguishable from WT plants. However, we cannot exclude that PolIA is involved in some other DNA repair mechanisms that do not involve DSBs.

A genetic interaction between *PolIB* and plastid *Whirly* genes

Previous studies have proposed that Whirly proteins play a role in organelle DNA replication (Maréchal et al., 2008; Prikryl et al., 2008). More recently, these proteins were shown to be involved in the maintenance of plastid genome stability (Maréchal et al. 2009) and the repair of DSBs in plastids and mitochondria (Cappadocia et al., 2010). It is therefore interesting that a strong synergistic interaction exists between *PolIB* and the plastid *Whirly* genes. This suggests that PolIB and the plastid Whirlies function in separate, but complementary pathways. The large increase of MMBIR rearrangements in ptDNA

observed in the triple mutant in the absence of genotoxic stress supports this conclusion and indicates that PolIB is not required or essential for DNA repair by MMBIR. This is reminiscent of the situation in yeast where break-induced replication (BIR) as well as MMBIR are dependent on a nonessential subunit of Pol δ called POL32 (Lydeard et al., 2007; Payen et al., 2008) and in *E. coli* where it is dependent on DinB (Galhardo et al., 2009). Our results suggest that, in the absence of PolIB, conservative repair of DSBs is reduced, leading to the accumulation of substrates (DSBs) that are repaired by the error prone MMBIR in absence of Whirlies. It also indicates that another DNA polymerase is responsible for the synthesis of rearranged DNA molecules, at least regarding the MMBIR process.

The fact that no phenotype is observed when *pollb-1* is crossed with the mitochondrial Whirly mutant allele *why2-1* supports once again the presence of an alternative DNA repair mechanism in this organelle. This is in agreement with our previous results where no visible or molecular phenotype were observed in the *why2-1* mutant unless an exogenous stress was applied (Cappadocia et al., 2010). This contrast with the situation with the *why1why3* plastid Whirly mutant where all the plants showed DNA rearrangements and 4.6% were variegated (Maréchal et al., 2009).

Taken together, our results indicate that the action of PolIB in DNA repair most likely involves a classical repair mechanism, such as HR. It is thus expected that reduced expression or inactivation of *PolIB* in *Arabidopsis* would lead to an increase in unattended DSBs and/or to slower repairs, even in the absence of exogenous stress. These breaks or arrested replication forks would be more likely to be corrected by other mechanisms, such as break-induced replication. In the absence of Whirly proteins, however, the breaks would also be repaired through error-prone mechanisms such as MMBIR, leading to the accumulation of rearranged DNA molecules such as those observed in *pollb-1/why1why3* and that could account for the severe phenotype of this line. It will be interesting to confirm this model in the future by further genetic crosses between the *pollb* alleles and other mutant alleles for genes involved in organelle DNA repair.

4.6 Experimental procedures

Plant material

Arabidopsis thaliana (ecotype Columbia-0) mutant lines *polla-1* (SALK_022638), *polla-2* (SALK_065221), *pollb-1* (SALK_134274) and *pollb-2* (WiscDsLoxHs021_09D) were obtained from the Arabidopsis Biological Resource Center (ABRC). Seeds were either sown on soil or sterilized and placed on petri dishes containing Murashige and Skoog (MS) basal media (Sigma-Aldrich, www.sigmaaldrich.com) supplemented with 1 % sucrose and 0.8 % agar with or without ciprofloxacin (Sigma-Aldrich). After three days of vernalization at 4°C, the seeds were placed under light (16 h day/8 h dark) at an average temperature of 22°C. Whole plants grown on MS medium were harvested at 3, 4, 5 (for seedlings), 14 (5 to 6 true leaves) or 21 (mature plants) days after germination by pulling the plants (including the roots) carefully from the agar and instantly frozen in liquid nitrogen. Plants were pooled to make 75 mg samples.

Nucleic acids isolation

DNA was isolated with a cetyl trimethylammonium bromide (CTAB) protocol (Weigel and Glazebrook, 2002). RNA was isolated using TRIzol® Reagent (Invitrogen, www.invitrogen.com) according to the manufacturer instructions. RNA pellets were resuspended in 30 µL of nuclease-free water and kept at -80°C. cDNA was synthesized using the First Strand cDNA Synthesis kit of Fermentas (www.fermentas.com) according to the manufacturer instructions. cDNA reactions were diluted 10 fold and kept at -20°C.

Quantitative PCR

Real-time qPCR reactions were done using the SYBR® Green Master Mix (SABiosciences, www.sabiosciences.com) according to the manufacturer instructions. Primers used for qRT-PCR experiments are as follow : 5'Tubu-qRT, CCT TCT CGG TGT TCC CTT CAC C, 3'Tubu-qRT, GGG AAC CTC AGA CAG CAA GTC AC, 5'PolIIA-qRTa, ACC GCG GTC GGG CTT TCT AGA, 3'PolIIA-qRTa, GCA CGT GAC TTT GAC GCC GGA, 5'PolIB-qRTb, GCG GAG ATC GCA AAG GAC ATA GTT GT, 3'PolIB-qRTb, TCC TTC AAC TAC GGC CGG TCT CT, 5'PolIB-qRTc, GAC GCT TAT CGG CTA GAA GGC CA and 3'PolIB-qRTc, CTT GAG TGG AAG TCT CCA CCA GCT. Primers used for qPCR experiments are as follow : 5'ncDNA-qPCR, GTT GAA GCC TCC GTT

CCC TGC TA, 3'ncDNA-qPCR, CTC TTC CAC CGT GCA TGG CTT GT, 5'mtDNA-qPCR, CCT GAT TCT GCG CGT AGA AGA CCT , 3'mtDNA-qPCR, AGG CGT AAG CGC AGC AGT TAG A, 5'ptDNA-qPCR, CCC TCT CTC TCG TAG TGT GGG GAA and 3'ptDNA-qPCR, TCG AAA GGG GTT ACC CCA TGA ATG G. The amplification efficiency of each reaction was determined using dilution of template cDNA or DNA. Calibration runs determined the use of 1/10 dilution for cDNA and 1/100 for DNA. Melting curves confirmed the presence of a unique PCR product in each case. Amplification was carried out by using the LightCycler®480 (Roche, www.roche-applied-science.com) and all the data were analyzed using the LightCycler®480 Software version 1.5.

Long-amplification PCR

Unprocessed DNA DSBs were quantified using the technique described in Yakes and Van Houten (1997). Briefly, DNA was amplified using GeneAmp® XL PCR kit (Applied Biosystems, www.appliedbiosystems.com) following the manufacturer recommendations. For mitochondria, plastids and the nucleus we amplified a long (about 20 kb) and a short (about 200 bp) PCR fragment. It is expected that the long amplification will be more susceptible to breaks than the short amplification. The primers used for ptDNA were as follow: 5'ptLongAmp, CCG GGA CTC GAA CCC GGA ACT A, 3'ptLongAmp, CTT CCT CCC GAG TTG AGA CCC A, 5'ptShortAmp, CTC CAG TGC ATT TCG CCC TCT GA and 3'ptShortAmp, ACG TAC ACG TCA CGG GCA TCC T. The primers used for mtDNA were as follow: 5'mtLongAmp, CCT ATC AGT AGC TAG GCC CTG GTC C, 3'mtLongAmp, CCT AAC TAC ACC GAC CTC CGA GTC C, 5'mtShortAmp, CCT GAT TCT GCG CGT AGA AGA CCT and 3'mtShortAmp, AGG CGT AAG CGC AGC AGT TAG A. The primers used for ncDNA were as follow: 5'ncLongAmp, TCG CGG AGA TCT TTC TCT CGC AC, 3'ncLongAmp, CTG CCT CCA CCA ACT CGT TCA C, 5'ncShortAmp, GTT GAA GCC TCC GTT CCC TGC TA and 3'ncShortAmp, CTC TTC CAC CGT GCA TGG CTT GT. DNA quantification of the reactions was done using PicoGreen® Reagent (Invitrogen) following the manufacturer instructions. Fluorescence was measured using a Perkin Elmer/Packard Fusion Alpha-FP Microplate Analyzer (PerkinElmer, www.perkinelmer.com). The level of PCR products for the long

amplification was reported on the level of the short product for each sample and normalized to 100 with WT untreated signal.

Detection of DNA rearrangements

Total DNA from MS medium-grown 21 day old plants was used to survey ptDNA rearrangements. DNA was equilibrated with a low-cycle PCR amplification of a Ycf2 DNA fragment. The PCR strategy and primers used were described in Cappadocia et al. (2010). PCR reactions products were migrated on ethidium bromide stained agarose gels.

4.7 Acknowledgements

The assistance of Alexandre Maréchal in isolating and crossing mutant lines is gratefully acknowledged. We thank B. Franz Lang and members of the Brisson lab for valuable discussions. E.L. was supported by a scholarship from the Natural Sciences and Engineering Research Council of Canada (NSERC). This research was supported by a grant from NSERC.

4.8 References

Bendich, A.J. (2007) The size and form of chromosomes are constant in the nucleus, but highly variable in bacteria, mitochondria and chloroplasts. *Bioessays*, 29, 474-483.

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.

Castroviejo, M., Tharaud, D., Tarrago-Litvak, L. and Litvak, S. (1979) Multiple deoxyribonucleic acid polymerases from quiescent wheat embryos. Purification and characterization of three enzymes from the soluble cytoplasm and one from purified mitochondria. *Biochem J*, 181, 183-191.

Christensen, A.C., Lyznik, A., Mohammed, S., Elowsky, C.G., Elo, A., Yule, R. and Mackenzie, S.A. (2005) Dual-domain, dual-targeting organellar protein presequences in *Arabidopsis* can use non-AUG start codons. *Plant Cell*, 17, 2805-2816.

Elo, A., Lyznik, A., Gonzalez, D.O., Kachman, S.D. and 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.

Fujie, M., Kuroiwa, H., Kawano, S., Mutoh, S. and Kuroiwa, T. (1994) Behavior of organelles and their nucleoids in the shoot apical meristem during leaf development in *Arabidopsis thaliana* L. *Planta*, 194, 395-405.

Fujie, M., Kuroiwa, H., Suzuki, T., Kawano, S. and Kuroiwa, T. (1993) Organelle DNA synthesis in the quiescent center of *Arabidopsis thaliana* (col.). *J Exp Bot*, 44, 689-693.

Galhardo, R.S., Do, R., Yamada, M., Friedberg, E.C., Hastings, P.J., Nohmi, T. and Rosenberg, S.M. (2009) DinB upregulation is the sole role of the SOS response in stress-induced mutagenesis in *Escherichia coli*. *Genetics*, 182, 55-68.

Hastings, P.J., Ira, G. and Lupski, J.R. (2009) A microhomology-mediated break-induced replication model for the origin of human copy number variation. *PLoS Genet*, 5, e1000327.

Heinhorst, S., Cannon, G.C. and Weissbach, A. (1990) Chloroplast and mitochondrial DNA polymerases from cultured soybean cells. *Plant Physiol*, 92, 939-945.

Keim, C.A. and Mosbaugh, D.W. (1991) Identification and characterization of a 3' to 5' exonuclease associated with spinach chloroplast DNA polymerase. *Biochemistry*, 30, 11109-11118.

Kimura, S., Uchiyama, Y., Kasai, N., Namekawa, S., Saotome, A., Ueda, T., Ando, T., Ishibashi, T., Oshige, M., Furukawa, T., Yamamoto, T., Hashimoto, J. and Sakaguchi, K. (2002) A novel DNA polymerase homologous to *Escherichia coli* DNA polymerase I from a higher plant, rice (*Oryza sativa* L.). *Nucleic Acids Res*, 30, 1585-1592.

Lydeard, J.R., Jain, S., Yamaguchi, M. and Haber, J.E. (2007) Break-induced replication and telomerase-independent telomere maintenance require Pol32. *Nature*, 448, 820-823.

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

Maréchal, A., Parent, J.S., Sabar, M., Véronneau-Lafortune, F., Abou-Rached, C. and Brisson, N. (2008) Overexpression of mtDNA-associated AtWhy2 compromises mitochondrial function. *BMC Plant Biol*, 8, 42.

Maréchal, A., Parent, J.S., Véronneau-Lafortune, F., Joyeux, A., Lang, B.F. and Brisson, N. (2009) Whirly proteins maintain plastid genome stability in Arabidopsis. *Proc Natl Acad Sci U S A*, 106, 14693-14698.

McKown, R.L. and Tewari, K.K. (1984) Purification and properties of a pea chloroplast DNA polymerase. *Proc Natl Acad Sci U S A*, 81, 2354-2358.

Meissner, K., Heinhorst, S., Cannon, G.C. and Börner, T. (1993) Purification and characterization of a gamma-like DNA polymerase from *Chenopodium album* L. *Nucleic Acids Res*, 21, 4893-4899.

Mori, Y., Kimura, S., Saotome, A., Kasai, N., Sakaguchi, N., Uchiyama, Y., Ishibashi, T., Yamamoto, T., Chiku, H. and Sakaguchi, K. (2005) Plastid DNA polymerases from higher plants, *Arabidopsis thaliana*. *Biochem Biophys Res Commun*, 334, 43-50.

Nielsen, B.L., Cupp, J.D. and Brammer, J. (2010) Mechanisms for maintenance, replication, and repair of the chloroplast genome in plants. *J Exp Bot*, 61, 2535-2537.

Ono, Y., Sakai, A., Takechi, K., Takio, S., Takusagawa, M. and Takano, H. (2007) NtPoll-like1 and NtPoll-like2, bacterial DNA polymerase I homologs isolated from BY-2 cultured tobacco cells, encode DNA polymerases engaged in DNA replication in both plastids and mitochondria. *Plant Cell Physiol*, 48, 1679-1692.

Payen, C., Koszul, R., Dujon, B. and Fischer, G. (2008) Segmental duplications arise from

Pol32-dependent repair of broken forks through two alternative replication-based mechanisms. *PLoS Genet*, 4, e1000175.

Prikryl, J., Watkins, K.P., Friso, G., van Wijk, K.J. and Barkan, A. (2008) A member of the Whirly family is a multifunctional RNA- and DNA-binding protein that is essential for chloroplast biogenesis. *Nucleic Acids Res*, 36, 5152-5165.

Rowan, B.A., Oldenburg, D.J. and Bendich, A.J. (2009) A multiple-method approach reveals a declining amount of chloroplast DNA during development in *Arabidopsis*. *BMC Plant Biol*, 9, 3.

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.

Sakai, A., Suzuki, T., Nagata, N., Sasaki, N., Miyazawa, Y., Saito, C., Inada, N., Nishimura, Y., T., Kuroiwa. (1999) Comparative analysis of DNA synthesis activity in plastid-nuclei and mitochondrial-nuclei simultaneously isolated from cultured tobacco cells. *Plant Science*, 140, 9-19.

Sala, F., Amileni, A.R., Parisi, B. and Spadari, S. (1980) A gamma-like DNA polymerase in spinach chloroplasts. *Eur J Biochem*, 112, 211-217.

Shereda, R.D., Kozlov, A.G., Lohman, T.M., Cox, M.M. and Keck, J.L. (2008) SSB as an organizer/mobilizer of genome maintenance complexes. *Crit Rev Biochem Mol Biol*, 43, 289-318.

Spencer, D. and Whitfeld, P.R. (1969) The characteristics of spinach chloroplast DNA polymerase. *Arch Biochem Biophys*, 132, 477-488.

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

Wamboldt, Y., Mohammed, S., Elowsky, C., Wittgren, C., de Paula, W.B. and Mackenzie, S.A. (2009) Participation of leaky ribosome scanning in protein dual targeting by

alternative translation initiation in higher plants. *Plant Cell*, 21, 157-167.

Weigel, D. and Glazebrook, J. (2002) *Arabidopsis: a laboratory manual*: Cold Spring Harbor Laboratory Press.

Yakes, F.M. and Van Houten, B. (1997) Mitochondrial DNA damage is more extensive and persists longer than nuclear DNA damage in human cells following oxidative stress. *Proc Natl Acad Sci U S A*, 94, 514-519.

Zoschke, R., Liere, K. and Borner, T. (2007) From seedling to mature plant: arabidopsis plastidial genome copy number, RNA accumulation and transcription are differentially regulated during leaf development. *Plant J*, 50, 710-722.

4.9 Figures legends

Figure 39. Isolation of *PollA* and *PollB* mutant lines.

(a) Schematic representation of *PollA* and *PollB* genes. Exons are denoted by blue boxes, the 3'UTR by a red box and the 5'UTR by a white box. The pale blue lines represent introns. The T-DNA insertions are numbered and indicated by black triangles. Half-arrows stand for the primers used for the qRT-PCR experiments.

(b) Histogram presenting the results of the qRT-PCR experiments measuring the level of expression of the *PollA* gene inside the *polla* mutant plants grown on MS medium for 14 days. Expression levels are represented as percentages of WT level. The experiments were done in technical and biological triplicates. The error bars represent the standard error of the mean on the biological triplicates.

(c) Histogram presenting the results of the qRT-PCR experiments measuring the level of expression of the *PollB* gene inside the *pollb* mutant plants grown on MS medium for 14 days. Expression levels are represented as percentages of WT level. The experiments were

done in technical and biological triplicates. The error bars represent the standard error of the mean on the biological triplicates.

(d) Representative photographs of four week old *Arabidopsis* plants for WT and *Poll* mutant lines.

Figure 40. Organelle DNA is less abundant in seedlings of *polla-1* and *pollb-1* mutants.

(a) Histogram presenting the results of the qPCR analysis of mitochondrial genome copy number relative to nuclear genome copy number for 3, 4 and 5 day old seedlings. The experiments were done in technical and biological triplicates. The error bars represent the standard error of the mean on the biological triplicate.

(b) Histogram presenting the results of the qPCR analysis of plastid genome copy number relative to nuclear genome copy number for 3, 4 and 5 day old seedlings. The experiments were done in technical and biological triplicates. The error bars represent the standard error of the mean on the biological triplicate.

Figure 41. *PollB* mutants exhibit an increased sensitivity to ciprofloxacin (CIP) and are impaired in their ability to process DNA DSBs.

(a) Representative photographs of the different *poll* mutants grown for 21 days on MS medium containing 0, 0.25 or 0.5 μ M of ciprofloxacin.

(b) Representative photographs of the first true leaves of *pollb-1* and *pollb-2* plants grown for 21 days on MS medium containing 0.25 μ M of ciprofloxacin.

(c) Histogram representing the percentage of plants with white/variegated first true leaves after 21 days of growth on medium with different concentrations of ciprofloxacin. Error bars represent the standard error of the mean of the counts made on different plates.

(d) Histogram representing the percentage of plants without any first true leaves after 21 days of growth on medium with different concentrations of ciprofloxacin. Error bars represent the standard deviation of the counts made on different plates.

(e) Histogram presenting the results of the PCR experiments to evaluate the abundance of unprocessed DSBs upon ciprofloxacin treatment in plastid DNA. Unprocessed DSBs are significantly more abundant for the *pollb-1* mutant when treated with both 0.25 and 0.5 μ M of ciprofloxacin. Error bars represent the standard error of the mean of three separate

experiments where WT signal at 0 μM of ciprofloxacin is expressed as 100 for each repetition. One asterisk indicates a significant difference with a $p < 0.05$ in the Student t-test and two asterisks, a difference with a $p < 0.01$.

(f) Histogram presenting the results of the PCR experiments to evaluate the abundance of unprocessed DSBs in mitochondrial DNA upon treatment with ciprofloxacin. Unprocessed DSBs are significantly more abundant for the *pollb-1* mutant when treated with 0.5 μM of ciprofloxacin. Error bars represent the standard error of the mean of three separate experiments where WT signal at 0 μM of ciprofloxacin is expressed as 100 for each repetition. Asterisk indicates a significant difference with a $p < 0.05$ in the Student t-test.

Figure 42. A *pollb-1/why1why3* triple mutant yields a pale-green dwarf phenotype and gives rise to abundant DNA rearrangements in plastids.

(a) Representative photographs of four week old plants for the genotypes used in the cross and the triple mutant progeny compared to WT.

(b) Representative PCR reactions to evaluate the abundance of DNA rearrangements dependant on microhomology in plastids of the plants presented in (a). The plants were grown for 21 days on MS medium containing 0 or 0.25 μM of ciprofloxacin. Two DNA samples were tested for each condition. PCR reactions were done using the primers indicated on the right of the gels.

(c) Schematic representation of DNA rearrangements evaluated by the PCR approach. Inward-facing primers placed far apart will yield a PCR product only if a deletion has occurred in the region that separates the primers. Outward-facing primers will yield a PCR product only if the template DNA is duplicated or circularized. Primers are represented by black half-arrows.

(d) Schematic representation of the DNA rearrangements amplified by PCR. The upper panel shows the position of the primers used to test the plastid genome of Arabidopsis. The lower panel gives a graphical representation of the rearranged DNA molecules. The sequences in the white boxes are the microhomologous sequences that were used to prime the MMBIR process. The original position of the microhomologous sequences in WT ptDNA is indicated.

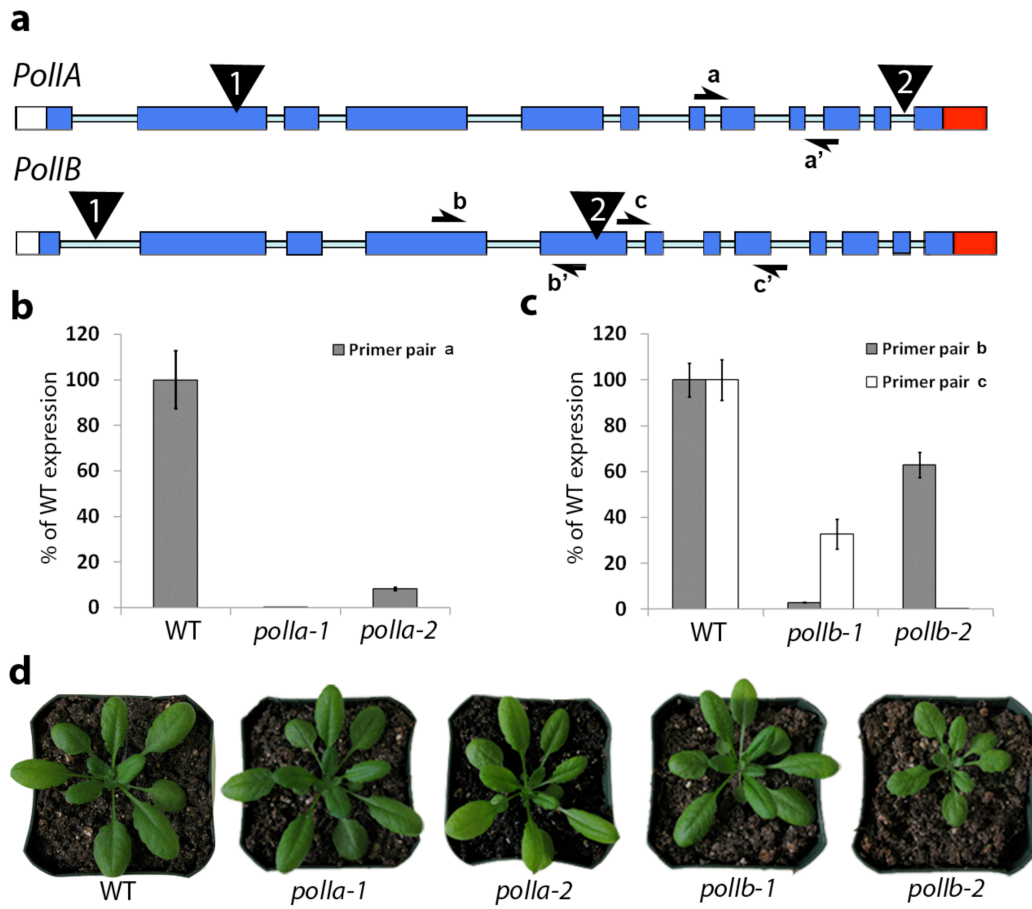


Figure 39. Isolation of *PollA* and *PollB* mutant lines.

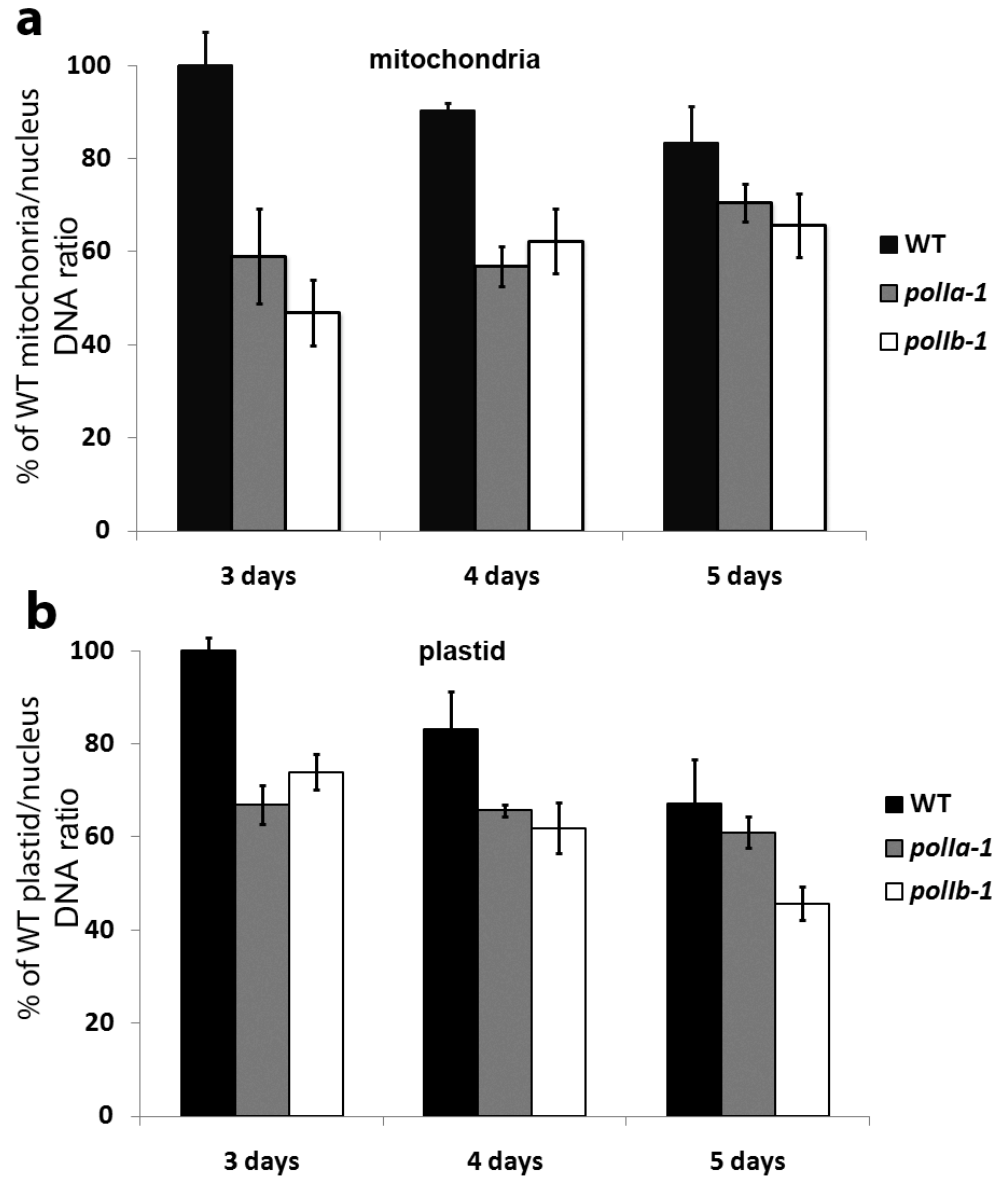


Figure 40. Organelle DNA is less abundant in seedlings of *polla-1* and *pollb-1* mutants.

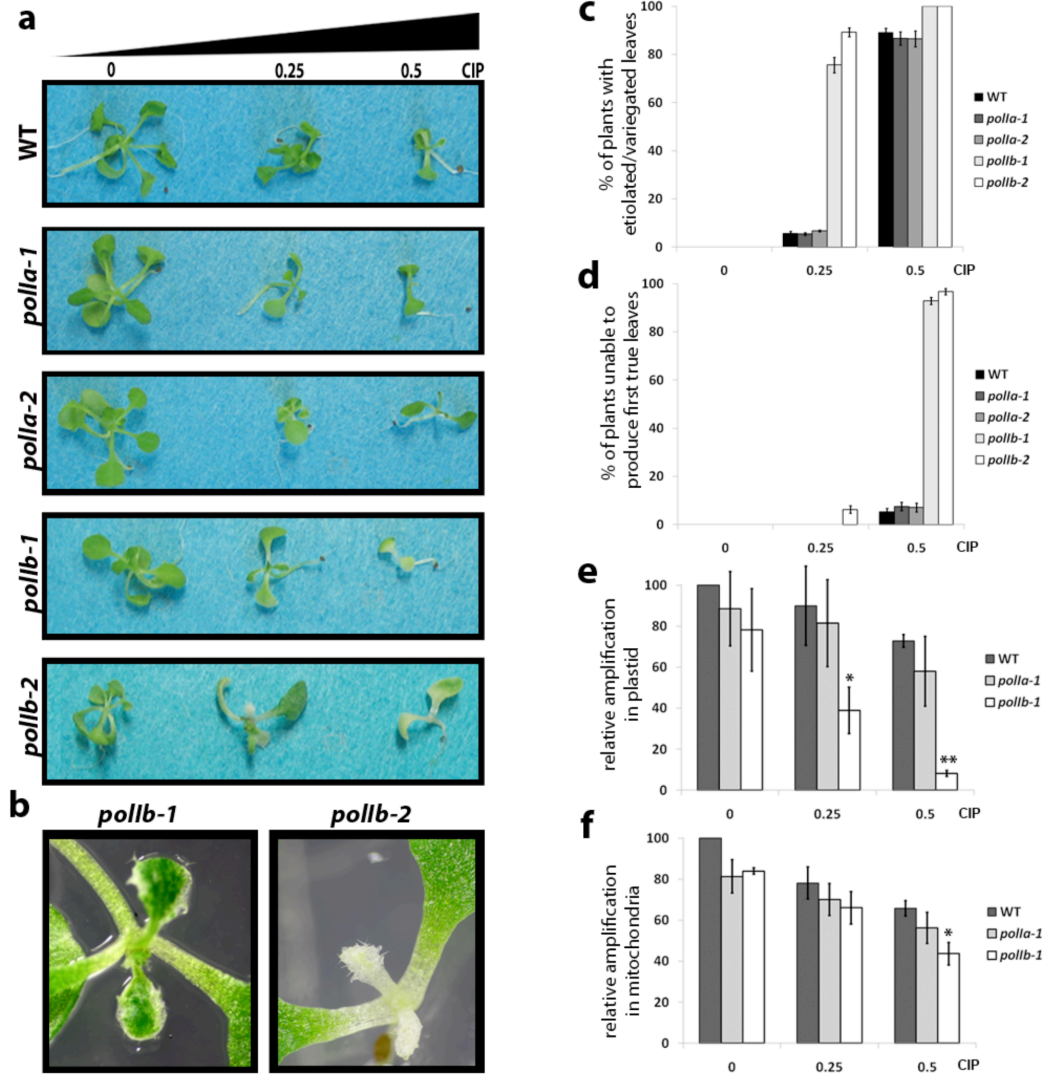


Figure 41. *PollB* mutants exhibit an increased sensitivity to ciprofloxacin (CIP) and are impaired in their ability to process DNA DSBs.

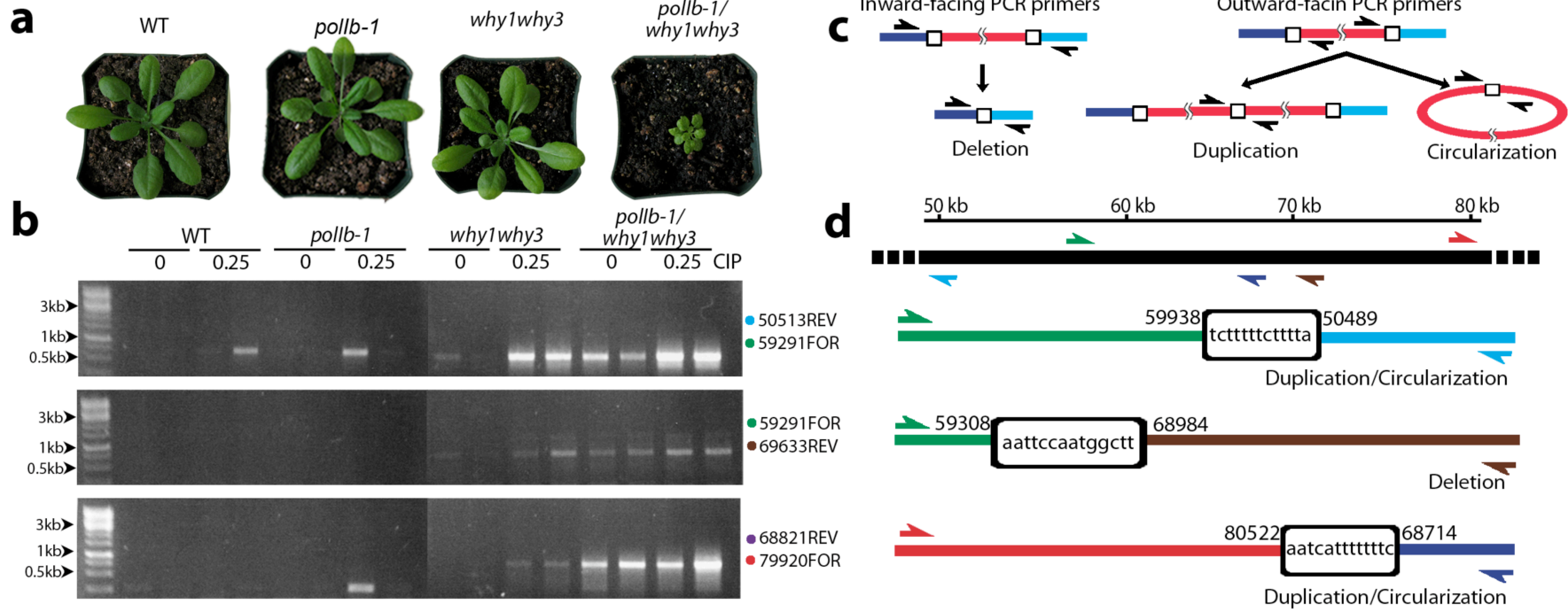


Figure 42. A *pollb-1/why1why3* triple mutant yields a pale-green dwarf phenotype and gives rise to abundant DNA rearrangements in plastids.

4.10 Supporting Information

Figure 43. Organelle DNA is less abundant in seedlings of *polla-2*.

(a) Histogram presenting the results of the qPCR analysis of mitochondrial genome copy number relative to nuclear genome copy number for 3, 4 and 5 day old seedlings. The experiments were done in technical and biological triplicates. The error bars represent the standard error of the mean on the biological triplicate. The results for the *polla-1* line (Fig. 2) are also presented.

(b) Histogram presenting the results of the qPCR analysis of plastid genome copy number relative to nuclear genome copy number for 3, 4 and 5 day old seedlings. The experiments were done in technical and biological triplicates. The error bars represent the standard error of the mean on the biological triplicate. The results for the *polla-1* line (Fig. 2) are also presented.

Figure 44. Ciprofloxacin does not generate DSB in the nuclear genome.

Histogram presenting the relative amplification of a long nuclear DNA fragment and a short nuclear DNA fragment for WT and *pollb-1* plants. Error bars represent the standard error of the mean of three separate experiments where WT signal at 0 μ M of ciprofloxacin is expressed as 100 for each repetition.

Figure 45. Crosses between *Poll* and *Whirly* mutants.

Representative photographs of WT and mutant plants from the indicated genotypes.

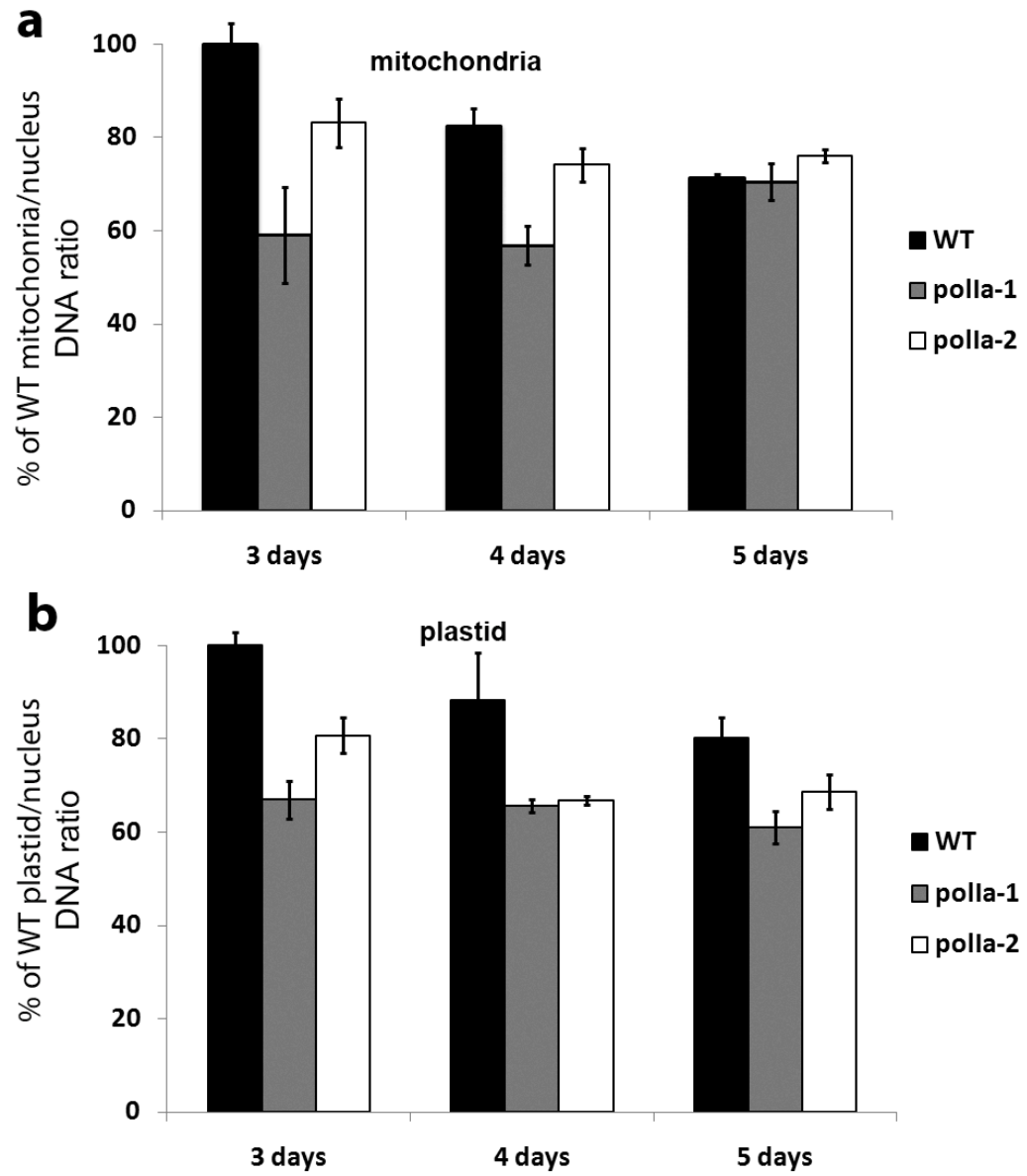


Figure 43. Organelle DNA is less abundant in seedlings of *polla-2*.

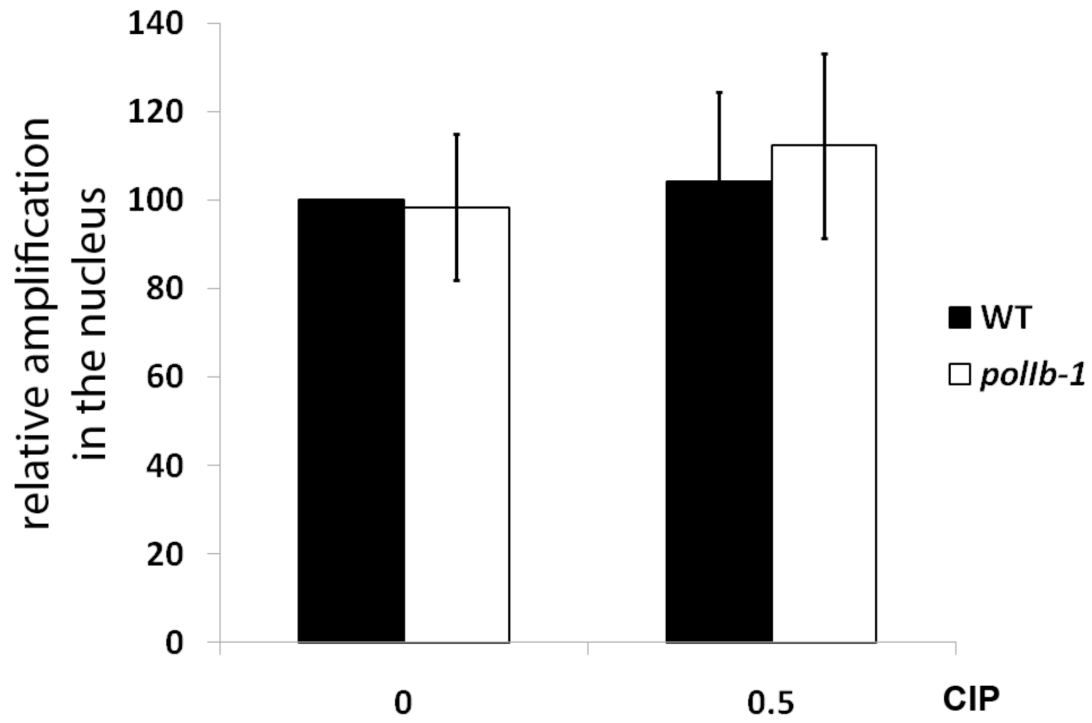


Figure 44. Ciprofloxacin does not generate DSB in the nuclear genome.

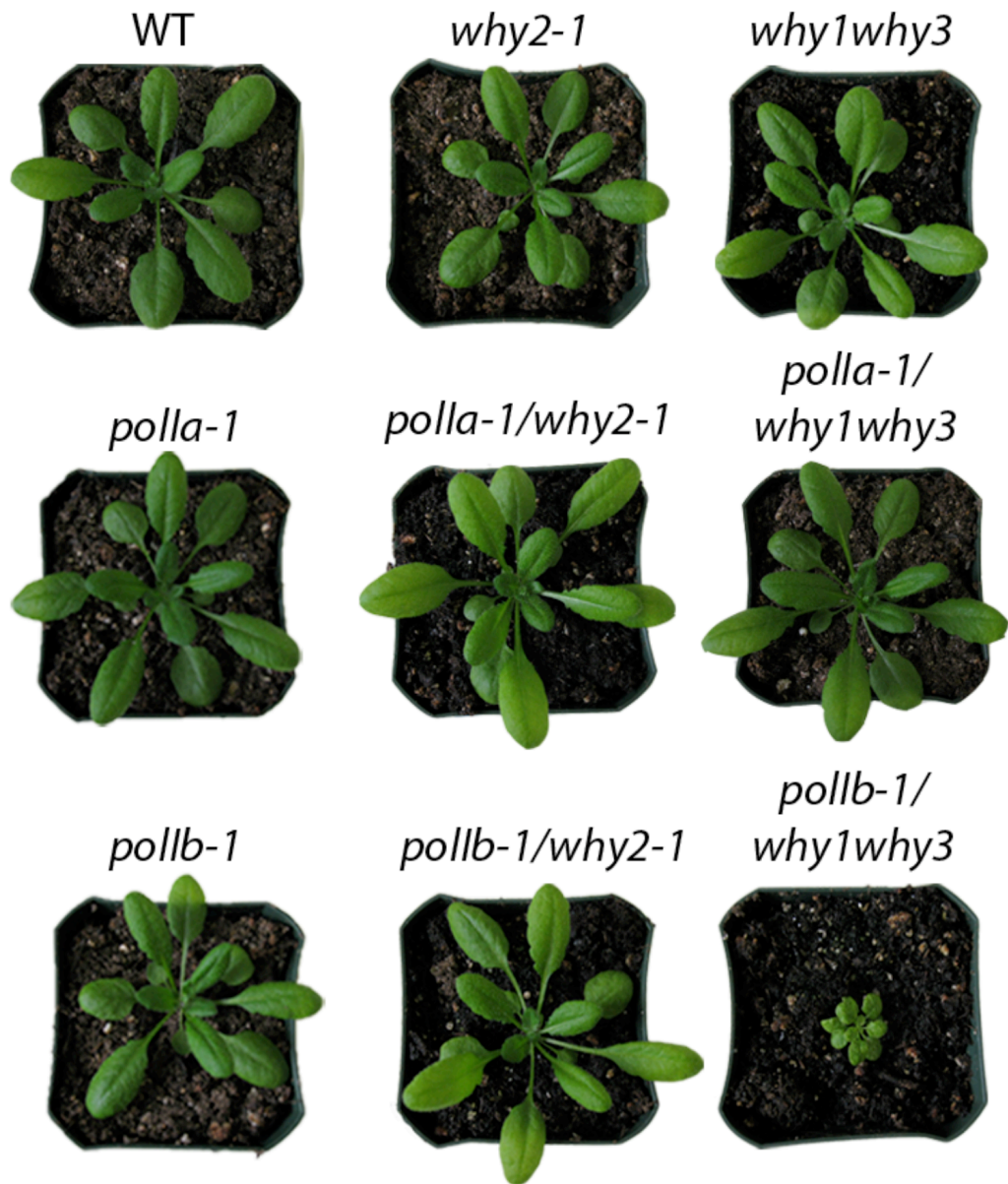


Figure 45. Crosses between *Poll* and *Whirly* mutants.

Chapitre 5. Discussion

5.1 Survol

Le but de cette étude était l'identification et la caractérisation de facteurs impliqués dans le métabolisme de l'ADN des organelles de plantes. Nous avons caractérisé deux familles de protéines spécifiques aux plantes chez l'organisme modèle *Arabidopsis thaliana* et présenté ces résultats par articles au fil des Chapitres 2, 3 et 4. Nous avons en effet découvert un rôle dans le maintien des génomes des organelles pour les protéines Whirly et les polymérases de l'ADN de type I. Considérant le peu d'information existant quant au métabolisme de l'ADN des mitochondries et des plastides chez les plantes, l'identification de facteurs impliqués dans ces mécanismes revêt une grande importance pour notre compréhension de la biologie végétale. La présente étude a permis d'identifier deux de ces facteurs, des protéines impliquées dans la réparation des dommages à l'ADN et dans sa réplication. L'identification de ces facteurs ouvre maintenant la porte à des études plus poussées en fournissant un nouveau modèle et de nouveaux outils. Une compréhension approfondie des processus liés à la réplication et la réparation de l'ADN constitue un enjeu important dans plusieurs domaines de l'agriculture.

En attendant, plusieurs questions intéressantes sont soulevées par les résultats présentés dans cet ouvrage. Dans ce chapitre, il sera traité de la réplication et du maintien des génomes des organelles de plantes en fonction des nouveaux résultats que nous avons obtenus. Ensuite, la portée du rôle des protéines Whirly et des polymérases de l'ADN de type I chez les plantes sera abordée. Finalement, l'importance de la stabilité des génomes des organelles pour l'évolution sera discutée et considérée comme source potentielle de variabilité pour le développement de nouveaux cultivars de plantes.

5.2 La réplication et le maintien des génomes des organelles de plante

Étant donné l'importance primordiale des plantes comme source d'alimentation, il est surprenant de constater à quel point certains aspects de la biologie des végétaux restent méconnus. Par exemple, la réplication de l'ADN des organelles de plante fait encore aujourd'hui l'objet d'un débat. Alors que plusieurs ouvrages font mention de génomes circulaires dans des organelles (13, 17), les résultats de plusieurs études contredisent cette affirmation (24, 25, 94). Cependant très peu de facteurs impliqués dans la réplication et le maintien de l'ADN des organelles ont été identifiés à ce jour, ce qui limite la portée des modèles proposés ((148) et revue dans (63)). Dans cette section, il est discuté du paradoxe que représente la réplication de l'ADN dans les organelles et des formes de réarrangements de cet ADN qu'on y retrouve. La nouvelle forme de réarrangement de l'ADN que nous avons identifiée au Chapitre 2 est ensuite considérée pour son apport potentiel à la réplication et au maintien du génome. Finalement, nos résultats viennent supporter un modèle où les molécules d'ADN linéaires des organelles se répliquent grâce la réplication dépendante de la recombinaison.

Les mécanismes qui permettent de maintenir les génomes des plastides et des mitochondries doivent être efficaces si l'on considère la faible accumulation de mutations silencieuses dans l'ADN mitochondrial et plastidique, comparativement à l'accumulation plus importante de telles mutations dans le noyau ((40) et revue dans (149)). De plus, on s'attendrait à ce que la reproduction asexuée des organelles et la présence dans ces compartiments de molécules hautement oxydatives nuisent à la stabilité de ces génomes. Une option pour la plante serait de préserver un petit nombre d'organelles loin des différentes sources de dommages, tels que ceux qui se produisent dans les cellules du mésophylle où se déroule la photosynthèse. En effet, il a récemment été proposé que la préservation des génomes des organelles à l'intérieur de cellules quiescentes pourrait expliquer la grande stabilité de ces génomes à travers le temps (150). Par ailleurs, il existe de nombreuses preuves démontrant l'existence de mécanismes de recombinaison actifs chez les organelles ((91, 93) et revue dans (63)). En effet, puisque le plastide et la

mitochondrie sont tous deux polyploïdes, la recombinaison de leur ADN offre la possibilité de corriger les dommages à la double hélice et les erreurs de réplication grâce à la conversion génique (voir section 1.2.2.5.2). Les résultats présentés ici suggèrent aussi que la recombinaison occupe une place importante dans le métabolisme de l'ADN des organelles.

L'importance de la recombinaison chez les organelles a également été suggérée par une étude chez le tabac qui visait à transformer l'ADN du plastide pour qu'il contienne deux gènes situés sur un même locus mais chacun rattaché à un marqueur de sélection différent (41). Le but était de forcer le plastide à maintenir deux génomes différents, un état que l'on nomme hétéroplasmie, grâce aux marqueurs de résistance aux antibiotiques. Cette étude rapporte qu'avant même la régénération complète de plantes transformées, on retrouve un seul génome (homoplasmie) qui est le produit d'un événement de recombinaison produisant une molécule d'ADN unique contenant les deux marqueurs de sélection. L'étude de Khakhlova *et al.* révèle deux aspects très importants du maintien de l'ADN_{pt}. D'abord, la présence d'une activité de recombinaison capable de remodeler rapidement l'ADN pour s'adapter à de nouvelles conditions (la présence d'antibiotiques) et ensuite la conversion de toutes les copies du plastome pour qu'il ne reste qu'une seule séquence d'ADN. Les auteurs de cette étude suggèrent que ce dernier phénomène est responsable de la grande stabilité de l'ADN_{pt} (41) et qu'il en va peut-être de même pour l'ADN_{mt}. Il est donc postulé qu'une nouvelle mutation se retrouvant en une seule copie du génome de l'organelle, contre des centaines de copies non mutées sera convertie très rapidement sans avoir vraiment aucune chance de s'établir.

En plus du rôle de correction des mutations, la recombinaison homologue est aussi le mécanisme privilégié dans tous les organismes pour réparer les bris double brin de l'ADN. Il est néanmoins difficile d'estimer l'ampleur de ce phénomène puisque ces réparations ne laissent aucune trace et qu'aucun système rapporteur pour ce phénomène n'existe. Chez la mitochondrie par contre, il existe un mécanisme assez bien caractérisé qui utilise de courtes séquences d'ADN (>100 pb et <1000 pb) pour effectuer la recombinaison ((52, 91) et revue dans (12, 63)). Ce phénomène crée de nouvelles molécules d'ADN qui peuvent avoir un effet considérable sur le phénotype de la plante (66). Ces molécules réarrangées

se retrouvent dans le génome mitochondrial en très faible quantité en temps normal, mais leur nombre peut être augmenté par certaines mutations (voir section 1.2.2.5.3). La mutation du gène *MSH1* impliqué dans la réparation de mésappariements entraîne en effet l'augmentation du nombre de molécules réarrangées, ce qui cause un phénotype semblable à la variégation étudiée au Chapitre 2 (52). Lorsque *RECA3*, qui code pour un homologue de la protéine RecA bactérienne, est muté, on observe aussi une augmentation du nombre de réarrangements causés par la recombinaison, bien que dans ce cas-ci aucun phénotype n'y soit associé. Les figures 31 et 32 montrent que cette même recombinaison est augmentée en présence de deux stress exogènes spécifiques à l'ADN des organelles. Ce mécanisme n'existe cependant pas dans le plastide où l'on ne retrouve pas de courtes séquences répétées à la base du processus. Cette différence entre les deux organelles pourrait bien être la raison pour laquelle l'ADNmt accumule moins de mutations que l'ADNpt ((40) et revue dans (149)).

Les résultats présentés dans les Chapitres 2 et 3 suggèrent l'existence d'un mécanisme de type MMBIR (voir section 1.2.2.5.2) à l'intérieur des organelles de plante. Cette découverte porte à croire qu'on retrouve également le mécanisme du BIR à l'intérieur de ces compartiments. Ce mécanisme est utilisé dans d'autres organismes pour réparer les bris double brin dans l'ADN lorsqu'une seule extrémité de la double hélice de l'ADN est disponible (revue dans (63, 76, 77)). Bien que cette voie de réparation n'implique pas nécessairement de réarrangements de l'ADN, elle y est plus susceptible que la recombinaison homologue classique (revue dans (63, 77)). En effet, le BIR peut utiliser différentes longueurs de séquences homologues, et plus cette séquence est courte, plus les chances d'un mauvais appariement augmentent, comme c'est le cas avec les microhomologies. Le BIR est cependant plus conservateur que les réparations par recombinaison non homologue comme le NHEJ et le MMEJ (voir section 1.2.2.5.1) récemment découverts dans le plastide d'*Arabidopsis* (64). Ce type de réparation peut causer de graves réarrangements et ne peut donc être une voie privilégiée, ce qui expliquerait pourquoi ces réparations n'ont jamais été observées chez la plante en l'absence d'un stress exogène (72). Ces réarrangements n'ont été détectés dans notre étude que lors du traitement à la ciprofloxacine (voir Tableau X) et, par un autre groupe,

lors de l'induction de coupures à l'ADN par une endonucléase (64). Au contraire, notre étude révèle que des réarrangements de type MMBIR sont présents, bien qu'en nombres très réduits, dans le plastome d'*Arabidopsis* et du maïs en l'absence de stress, ce qui indique que le MMBIR et par extension le BIR sont utilisés en temps normal.

Chez la levure, le BIR peut également être utilisé pour réparer un chromosome brisé produits par une réplication incomplète ((80) et revue dans (76)). Si l'on prend pour acquis qu'il existe des molécules d'ADN linéaires à l'intérieur des organelles de plantes ((17) et revue dans (11, 22)), alors leurs extrémités pourraient servir de substrats à la voie du BIR comme c'est le cas chez la levure. En effet, les modèles de réplication de l'ADN des organelles qui postulent que la majorité des génomes se retrouvent sous forme linéaire font appel à la réplication dépendante de la recombinaison (RDR) ((17) et revue dans (11, 22)). Un tel phénomène impliquerait l'utilisation de l'extrémité d'une molécule d'ADN qui trouverait une région d'homologie sur une autre molécule et amorcerait ainsi une réaction de réplication (voir Figure 6). Ce mécanisme pourrait expliquer les formes d'ADN branchées qui sont observées dans le plastide par microscopie (voir Figure 3) (25, 94). De plus, une étude récente a démontré chez la levure que le BIR était indépendant du complexe de reconnaissance des Ori (151). Chez la plante, cette indépendance envers les origines de réplication expliquerait pourquoi la réplication du plastome ne semble pas être affectée par les mutations des deux Ori présentes dans l'ADN_{pt} du tabac (27, 28). De plus, chez *C. reinhardtii*, on a détecté une forme de RDR indépendante des Ori lorsque les algues étaient traitées avec un inhibiteur de gyrase qui bloque la réplication dépendante des Ori (152). Il est possible que cette voie ne soit nulle autre que celle du BIR qui prend en charge la réplication de l'ADN du plastide en situation de stress.

Du point de vue de l'ADN, une réaction de RDR est identique à une réaction de BIR; et donc, bien que nous traitons les deux concepts comme des phénomènes distincts, il se peut qu'il s'agisse d'une seule et même voie métabolique. La présence d'un tel mécanisme dans la mitochondrie en l'absence de stress exogène reste cependant à prouver, mais les résultats présentés au Chapitre 3 le suggèrent fortement. L'étude plus approfondie des facteurs identifiés dans cette thèse, tels les Poll (voir Chapitre 4), et surtout l'identification future de nouveaux facteurs, permettront d'évaluer l'importance

des différents phénomènes de recombinaison pour le maintien de l'ADN des organelles de plante à travers le temps. Un système rapporteur de recombinaison qui serait introduit dans les organelles de plante pourrait également répondre à plusieurs questions, mais serait techniquement difficile à réaliser.

5.3 Le rôle des protéines Whirly dans la stabilité des génomes

Le point de départ de notre étude était l'identification du rôle des protéines Whirly dans les organelles de plantes. Les résultats des Chapitres 2 et 3 montrent leur implication dans la réparation des dommages à l'ADN. Il est discuté dans cette section de la portée du rôle de cette famille de protéines d'abord dans le plastide, et ensuite dans la mitochondrie, pour terminer avec un modèle de l'action des Whirly dans la réparation des bris double brin de l'ADN et d'un rôle de protection de l'ADNsb.

Plusieurs rôles nucléaires ont été proposés pour les protéines Whirly (125, 129, 130). Cependant, malgré le fait que ces protéines semblent principalement localisées dans les organelles en temps normal ((126, 127) et nos résultats non publiés), peu d'études se sont intéressées à leur rôle dans ces compartiments. Les résultats présentés au Chapitre 2 montrent que la mutation des gènes *Whirly* du plastide cause une instabilité dans le plastome d'*Arabidopsis* et du maïs. La conservation de ce rôle dans une espèce monocotylédone et une espèce dicotylédone suggère que celui-ci est partagé par toutes les protéines Whirly retrouvées chez les plantes à fleurs. Il semble néanmoins y avoir une certaine variation au niveau de l'importance de la présence des protéines Whirly à l'intérieur du plastide entre les différentes espèces. En effet, la mutation du gène du maïs *ZmWhy1* est très délétère pour la plante qui meure en bas âge (133). Par contre, chez *Arabidopsis*, la double mutation des gènes *AtWhy1* et *AtWhy3* ne provoque un phénotype visible que chez environ 5% des plantes. Cette différence entre les deux espèces pourrait s'expliquer par une fonction additionnelle pour *ZmWhy1*. Il semble effectivement que cette protéine soit également impliquée dans la maturation des ARNm du plastide chez le maïs (133), ce qui ne semble pas être le cas des protéines d'*Arabidopsis* (voir Figure 18). La

famille de protéines Whirly pourrait donc s'acquitter de plusieurs tâches différentes à l'intérieur des organelles, comme c'est le cas dans le noyau.

Du côté des protéines Whirly de la mitochondrie, les seuls résultats disponibles à ce jour concernent la protéine AtWhy2 d'*Arabidopsis*. Lors d'une étude précédente, nous avons démontré que cette protéine s'associe à l'ADNmt de façon non spécifique *in vivo* (132). Aucune expérience ne nous permettait cependant de voir une différence entre une lignée mutante pour AtWhy2 et une plante de type sauvage. Les résultats du Chapitre 3 montrent toutefois qu'en présence d'un stress spécifique à l'ADN, les plantes mutantes accumulent plus de molécules d'ADNmt réarrangées que les plantes sauvages. Ce résultat indique que le rôle de maintien de la stabilité du génome des protéines ptWhirly se retrouverait aussi chez les mtWhirly. Par contre, dans la mitochondrie, il semble que les réarrangements de type MMBIR soient plus difficilement détectables, ou simplement moins nombreux que chez le plastide, même en présence de stress. Ceci suggère que d'autres voies de réparation de l'ADN n'impliquant pas les protéines Whirly existent chez la mitochondrie et non chez le plastide. L'exemple parfait d'une telle autre voie est la recombinaison homologue dépendante de courtes séquences répétées discutée précédemment et qui se retrouve seulement dans la mitochondrie et n'est pas affectée par la mutation de AtWhy2 (voir Figures 31 et 32). Alternativement, il se peut que les mtWhirly soient aidées dans leur rôle par d'autres protéines, présentes seulement dans la mitochondrie. Nous proposons à quelques reprises dans ce travail que ces protéines pourraient être les homologues des protéines SSB des prokaryotes que l'on retrouve spécifiquement dans la mitochondrie de la plante (118).

Nos données suggèrent donc, pour les protéines Whirly, un rôle de protection contre les réarrangements de l'ADN à l'intérieur des organelles. Cette hypothèse est appuyée au Chapitre 3 par la résolution de la structure de complexes Whirly-ADNsb montrant que l'association entre les deux molécules se fait grâce à des contacts qui confèrent très peu de spécificité de séquence. Cette caractéristique est typique des protéines qui lient l'ADNsb pour le protéger contre la dégradation ou les mésappariements. En effet, nous observons que l'ADNsb est enroulé autour du tétramère de protéines Whirly et que ses bases azotées sont tournées vers la protéine et ne sont donc pas exposées au solvant. Ceci

suggère donc un modèle où les protéines Whirly se lient à l'extrémité 3' simple brin d'un bris double brin de l'ADN laissé libre suite à la résection du brin 5'. Elles empêcheraient ainsi l'extrémité réactive de faire un mauvais appariement et ainsi provoquer une recombinaison illégitime (voir Figure 26). Une fois le bon appariement trouvé, les protéines Whirly se détacheraient passivement ou bien seraient enlevées par une autre protéine. Il se pourrait également que les protéines Whirly protègent non pas l'extrémité 3' réactive mais bien l'ADNsb des bulles de réplication et de transcription où l'extrémité pourrait s'hybrider (voir Figure 26). Ce mode d'action est très similaire à celui des protéines SSB chez les bactéries. Comme les protéines bactériennes, les protéines Whirly pourraient recruter les différents acteurs de la réparation de l'ADN alors qu'elles sont elles-mêmes fixées à l'ADNsb. L'interaction avec les protéines de réparation pourrait s'effectuer par la queue C-terminale conservée des protéines Whirly comme c'est d'ailleurs le cas pour les protéines SSB (revue dans (98)).

Aussi, en l'absence de stress, la capacité des protéines Whirly à lier l'ADNsb sans spécificité de séquence (voir Figure 24A) pourrait leur permettre de protéger l'ADN vulnérable des organelles. En effet, des processus comme la réplication et la transcription créent quantité d'ADNsb susceptible aux dommages par les espèces oxydatives. Une expérience de digestion avec l'endonucléase Mung Bean couplée à l'électrophorèse en champs pulsé a d'ailleurs récemment montré que l'ADNsb était abondant dans les chloroplastes d'*Arabidopsis* (93). Les protéines Whirly qui sont présentes de façon constitutive dans ces organelles pourraient protéger cet ADN par le simple fait de s'y lier. Leur absence provoquerait donc une augmentation du nombre de dommages et donc une augmentation simultanée du nombre de réparations/réarrangements. Si ce modèle explique en partie nos résultats, il n'explique cependant pas pourquoi la recombinaison dépendante de courtes séquences homologues n'est pas augmentée dans les mitochondries en l'absence de la protéine AtWhy2 (voir Figures 31 et 32).

5.4 Le rôle des polymérases de l'ADN de type I dans la réplication et la réparation de l'ADN des organelles

L'utilisation de la ciprofloxacine comme stress à l'ADN des organelles a permis l'identification de plusieurs autres gènes possiblement impliqués dans la réparation de l'ADN des plastides et des mitochondries. Le gène *PollB* fait partie du nombre et a ainsi lancé notre étude des polymérases de l'ADN des organelles. Il est discuté ici de l'apport de notre étude à l'hypothèse proposant que les enzymes Poll soient responsables de la synthèse de l'ADN dans les organelles. Il est traité ensuite du rôle de *PollB* dans la réparation des dommages à l'ADN et du lien qui existe entre cette polymérase et les protéines Whirly du plastide. Finalement nous discutons de l'existence possible d'autres polymérases de l'ADN qui viendraient en aide aux Poll.

Avant notre étude, plusieurs groupes avaient déjà proposé que la famille des Poll était responsable de la synthèse de l'ADN dans les organelles de plantes (29-31). En effet, les caractéristiques des Poll recombinantes produites chez la bactérie correspondent parfaitement aux caractéristiques rapportées pour les enzymes extraites directement des mitochondries et des plastides de différentes plantes (137-139, 143, 144). Le Chapitre 4 apporte une démonstration de ce lien grâce à des plantes mutantes pour les gènes *PollA* et *PollB* d'*Arabidopsis*. Tout d'abord, la mutation combinée des deux gènes est létale pour la plante, ce qui indique une fonction essentielle partagée par ces gènes. Ensuite, les simples mutants contiennent moins d'ADNmt et d'ADNpt dans les jeunes pousses quelques jours après la germination. La différence des niveaux d'ADN entre les mutants et les plantes WT tend à s'amenuiser avec le temps. Le fait que les niveaux d'ADN des mutants et des plantes WT tendent à s'équilibrer avec l'âge est attendu puisque l'expression des gènes *Poll* est à son maximum dans les jeunes tissus méristématiques (29-31). C'est donc dans ces tissus que nous attendons une différence plus prononcée des niveaux d'ADN lorsqu'il y a mutation des gènes *Poll*. De plus, il a été rapporté par diverses études que le contenu en ADN des organelles diminuait à mesure que les tissus vieillissaient (25, 94, 153). Il est donc normal de voir les niveaux d'ADN des organelles de

la plante WT diminuer avec le temps et se rapprocher des niveaux des mutants dont la capacité de synthèse est diminuée.

Le fait que les deux protéines Poll d'*Arabidopsis* partagent une même fonction n'a rien de surprenant vu leur haut niveau d'identité quant à leur séquence en acides aminés (30). Un rôle pour ces enzymes dans la réparation de l'ADN pouvait également être attendu si l'on considère que l'ADN polymérase de type I de *E. coli* est impliquée à la fois dans la réplication et la réparation de l'ADN (revue dans (154)). Il est néanmoins surprenant que la fonction de réparation ne soit observée que pour PolIB. Cette fonction pourrait dépendre des quelques résidus qui différencient cette protéine de PolIA. Alternativement, il se pourrait que PolIB soit spécifiquement recrutée au site des dommages à l'ADN. En effet, il arrive souvent que la réparation de l'ADN dépende du recrutement séquentiel d'une multitude de protéines. Si PolIB a récemment acquis ce nouveau rôle, il se peut que PolIA soit encore capable de remplir cette fonction, mais moins efficacement. Cette hypothèse expliquerait pourquoi les mutants *pollb* arrivent encore à réparer leur ADN si on les traite avec une dose intermédiaire de ciprofloxacine (voir figure 41). Cela expliquerait également pourquoi PolIA n'arrive pas à compenser parfaitement la perte de PolIB dans la lignée *pollb-2* qui montre un retard de croissance (voir figure 39d).

Étant donné l'implication de PolIB dans la réparation des dommages à l'ADN, il devenait intéressant de croiser les différents mutants des gènes *Poll* avec les mutants des gènes *Whirly*, puisque ces derniers sont également impliqués dans la réparation de l'ADN. Le fait que seul le croisement *pollb-1/why1why3* génère un phénotype visible apporte plusieurs informations. D'abord, cela confirme l'implication spécifique de *PolIB* dans la réparation puisqu'une fois de plus, *PolIA* ne montre aucun effet visible. Ensuite, aucun effet n'est observé lors d'une combinaison avec la mutation du gène *AtWhy2*, le mtWhirly. Cette observation renforce deux hypothèses que nous avons proposées précédemment, soit que : 1) il existe des mécanismes de réparation de l'ADN dans la mitochondrie qu'on ne retrouve pas dans le plastide, et 2) d'autres protéines spécifiques aux mitochondries compensent pour l'absence du mtWhirly.

L'analyse du mutant *pollb-1/why1why3* a également fourni plusieurs informations très intéressantes pour notre modèle. Tout d'abord, il nous renseigne sur l'épistasie qui existe entre les trois allèles impliqués. Bien que nous confirmions que PollB et les ptWhirly soient impliqués dans la réparation de l'ADN, il semble que l'effet des mutations s'additionne et que les protéines agiraient donc dans des voies distinctes. Ensuite, grâce à la stratégie de PCR servant à mesurer le niveau de réorganisation du plastome, nous avons établi que le nombre de réarrangements de l'ADN était très élevé dans le triple mutant. En effet, la figure # montre la détection, dans les plantes *pollb-1/why1why3* non traitées, d'un nombre de molécules réarrangées comparable à celui mesuré dans le mutant *why1why3* traité avec 0.25 μ M de ciprofloxacine (voir Figure 42b). Ce résultat nous apprend que PollB n'est probablement pas la polymérase responsable de la synthèse de l'ADN lors des réactions de BIR ou de MMBIR. Il s'agirait donc d'une différence majeure avec la polymérase de type I de *E. coli* qui est responsable du BIR chez cet organisme (155). Cette fonction chez la bactérie dépend néanmoins du domaine 5'-3' exonucléase de la protéine que les Poll des plantes ne possèdent pas, ce qui pourrait expliquer cette différence de fonction. Nos résultats suggèrent plutôt que PollB serait impliqué dans les voies de réparation classique comme la recombinaison homologe. Donc, en l'absence de PollB, les voies classiques seraient bloquées ou très inefficaces, ce qui dirigerait l'ADN endommagé vers d'autres voies, tel le BIR. Cependant, en l'absence de protéines Whirly, la voie du BIR manque de précision ce qui donne lieu au MMBIR qui crée les molécules d'ADN réarrangées que nous détectons. Ceci expliquerait pourquoi la mutation simultanée des gènes *PollB* et *ptWhirly* produit des plantes qui contiennent un grand nombre de molécules d'ADNpt réarrangées, ce qui entraîne des effets graves sur le développement de la plante. La même logique devrait s'appliquer à la mitochondrie, mais l'existence de voies de réparation parallèles et/ou la présence de protéines fonctionnellement similaires aux protéines Whirly empêcheraient l'accumulation excessive de mutations.

Puisque la quantité d'ADN dans les organelles n'est que faiblement affectée lors d'une diminution de la quantité de Poll, cela suggère que d'autres polymérases de l'ADN pourraient compenser pour ces mutations. De plus, la présence de molécules d'ADN

réarrangées dans le mutant *pollb-1/why1why3* indique qu'au moins une autre polymérase peut se charger de la réparation de l'ADNpt. Un raisonnement similaire s'applique à la situation des mitochondries où l'absence de PolIB semble avoir un effet beaucoup plus modeste sur la réparation des bris double brin causés par la ciprofloxacine. Les mutants des *Poll* identifiés dans cette étude se révéleront certainement des outils précieux dans l'identification future de ces autres polymérases de l'ADN.

5.5 La stabilité des génomes et l'évolution

Le maintien d'un génome stable et sa transmission aux générations subséquentes est un enjeu primordial pour tous les organismes vivants. Une certaine variabilité au niveau de la séquence d'ADN est toutefois nécessaire à l'évolution, car un génome qui ne peut muter est un génome qui pourra difficilement évoluer. Ceci est certainement l'une des raisons pour laquelle la réplication de l'ADN demeure un processus imparfait et pour laquelle des voies de réparation de l'ADN sujettes aux erreurs comme le MMBIR et le NHEJ existent. En effet, les phénomènes du BIR et du MMBIR peuvent mener à des délétions et des duplications de larges fragments d'ADN, ce qui peut avoir des effets positifs importants pour l'organisme, autant chez les bactéries (155) que chez l'humain (81, 156, 157). Puisque ces mutations peuvent également conduire à des dérèglements graves (revue dans (82)), il est cependant important de les garder sous contrôle.

Du côté des plantes, il a été démontré que des réarrangements de l'ADN par recombinaison entre de courtes séquences répétées peuvent aider la mitochondrie à s'adapter à des nouvelles conditions environnementales (66). Ces événements de recombinaison peuvent en effet créer de nouveaux fragments d'ADN et peut-être même de nouveaux gènes en joignant deux séquences qui sont normalement séparées (65). Nous proposons que les réarrangements produits par le mécanisme du MMBIR pourrait également avoir un apport positif pour l'adaptation des plantes à de nouvelles conditions. Ce phénomène peut en effet causer la duplication ou la délétion de larges fragments d'ADN qui contiennent plusieurs gènes. Des changements au niveau du nombre de copie de certains gènes pourraient entraîner des changements correspondant au niveau de

l'expression de ces gènes. De plus, la duplication de certains gènes a pour effet de relâcher la pression sélective exercée sur les deux copies ce qui permet aux gènes de muter et éventuellement d'acquérir de nouvelles fonctions (158).

Chapitre 6. Conclusion et perspectives

Le maintien de l'ADN des génomes des organelles de plante est encore un phénomène mal compris et des voies métaboliques aussi centrales que la réplication de l'ADN font encore l'objet de débats. Notre étude a permis de caractériser cinq protéines impliquées dans le maintien de ces génomes et ainsi d'apporter un peu de lumière sur cette branche négligée de la biologie des plantes. En utilisant la plante modèle *Arabidopsis thaliana* et la génétique inverse, nous avons démontré l'implication des protéines Whirly dans la stabilité des génomes des organelles et aussi dans la réparation des dommages à l'ADN. Ces protéines semblent en effet prévenir l'utilisation de voies de réparation promptes aux erreurs, erreurs qui causent des réarrangements au niveau de la séquence d'ADN. Puisque les protéines Whirly s'associent avec l'ADN simple brin de façon peu spécifique, il est logique de penser qu'elles lient l'ADN aux extrémités laissées libres à la suite d'une cassure en attendant une réparation conservatrice. Parallèlement, les protéines Whirly pourraient lier les extrémités des génomes linéaires existants dans les organelles et les protéger, un peu à l'image des protéines liant les télomères dans le noyau. Alternativement, ces protéines pourraient tout simplement lier l'ADNsb laissé libre lors des activités métaboliques régulières des organelles et le protéger contre les dommages.

Les réarrangements de l'ADN, qui deviennent abondants dans le génome du plastide en l'absence de protéines Whirly, se retrouvent également dans le génome, en faible quantité, en temps normal. Ceci nous amène à formuler l'hypothèse que ces réarrangements peuvent être bénéfiques pour la plante lorsqu'elle a besoin de s'adapter à de nouvelles conditions environnementales. Une étude récente a en effet démontré qu'une réorganisation au niveau du génome de la mitochondrie des plantes pouvait donner lieu à une plus grande résistance à la chaleur (66). Il serait intéressant de tester si les mutants des gènes *Whirly* auraient un avantage sélectif sur une plante sauvage pour s'adapter plus rapidement à des conditions extrêmes.

Notre étude des plantes mutantes pour les gènes *Whirly* a permis de mettre au point une approche générant un stress spécifique pour l'ADN des organelles des plantes. En effet, le travail présenté ici démontre que la ciprofloxacine peut être utilisée pour induire des bris

double brin dans l'ADNmt et l'ADNpt. En très peu de temps, ce nouvel outil nous a permis de préciser le rôle des ptWhirly, de trouver un rôle dans la réparation de l'ADN pour le mtWhirly d'*Arabidopsis* et pour PolIB. Ce stress représente donc une opportunité unique d'identifier rapidement un grand nombre de gènes impliqués dans la réparation de l'ADN des organelles. Au laboratoire, plusieurs autres gènes ont déjà été identifiés par cette approche et sont en voie de caractérisation. De plus, nous pensons que la ciprofloxacine pourrait être utilisée sur des cultivars dans le but de provoquer une multitude de réarrangements au niveau des génomes des organelles et d'en sélectionner des traits intéressants pour l'agriculture.

L'identification des polymérase responsables de la synthèse de l'ADN dans les organelles représente également une avancée majeure en biologie végétale. L'implication spécifique de PolIB dans la réparation des dommages à l'ADN a permis d'enrichir notre modèle pour le maintien de l'ADN des organelles. Les mutants que nous avons identifiés seront des atouts pour les études génétiques qui seront menées au laboratoire. En somme, nous proposons un premier modèle très simple le maintien des génomes des organelles de plante qui permet d'expliquer les différents résultats rapportés par cet ouvrage. Ce modèle pourra être le point de départ de plusieurs autres études qui viendront l'enrichir en y greffant des facteurs nouvellement identifiés. Un modèle solide du maintien de ces petits génomes est essentiel à notre compréhension des plantes et pourra être utilisé pour le développement de nouveaux cultivars répondant mieux aux exigences sans cesse croissantes de l'agriculture.

Références

1. Nunes-Nesi, A. & Fwenie, A. R. (2007) Mitochondrial metabolism in *Plant Mitochondria*, ed. Logan, D. (Blackwell Publishing Ltd, Oxford), pp. 212-277.
2. Pyke, K. (2007) Plastid biogenesis and differentiation in *Cell and Molecular Biology of Plastids*, ed. Bock, R. (Springer-Verlag Berlin Heiderberg, Berlin), pp. 1-28.
3. Altmann, R. (1890) *Die Elementarorganismen und ihre Beziehungen zu den Zellen* (Viet, Leipzig).
4. Gray, M. W., Burger, G., & Lang, B. F. (1999) Mitochondrial evolution. *Science* **283**, 1476-1481.
5. Dyall, S. D., Brown, M. T., & Johnson, P. J. (2004) Ancient invasions: from endosymbionts to organelles. *Science* **304**, 253-257.
6. Adams, K. L. & Daley, D. O. (2004) Plant Mitochondrial Genome Evolution and Gene Transfer to the Nucleus in *Plant Mitochondria: From Genome to Function*, ed. Day, D. A., Millar, A.H. and Whelan, J. (Kluwer Academic Publishers), pp. 107-120.
7. Gould, S. B., Waller, R. F., & McFadden, G. I. (2008) Plastid evolution. *Annu Rev Plant Biol* **59**, 491-517.
8. Kleine, T., Maier, U. G., & Leister, D. (2009) DNA transfer from organelles to the nucleus: the idiosyncratic genetics of endosymbiosis. *Annu Rev Plant Biol* **60**, 115-138.
9. Lister, D. L., Bateman, J. M., Purton, S., & Howe, C. J. (2003) DNA transfer from chloroplast to nucleus is much rarer in *Chlamydomonas* than in tobacco. *Gene* **316**, 33-38.
10. Newton, K. J., Gabay-Laughnan, S. and De Paepe, R. (2004) Mitochondrial Mutations in Plants in *Plant Mitochondria: From Genome to Function*, ed. Day, D. A., Millar, A.H. and Whelan, J. (Kluwer Academic Publishers), pp. 121-142.
11. Brackert, S., Nielsen, B.L. and Börner, T. (1997) The mystery of the rings: structure and replication of mitochondrial genomes from higher plants. *Trends in plant science* **2**, 477-483.
12. Mackenzie, S. A. (2007) The unique biology of mitochondrial genome instability in plants in *Plant Mitochondria*, ed. Logan, D. (Blackwell Publishing), pp. 36 - 49.
13. Backert, S., Nielsen, B.L. and Borner, T. (1997) The mystery of the rings: structure and replication of mitochondrial genomes from higher plants. *Trends in plant science* **2**, 477-483.
14. Kucej, M. & Butow, R. A. (2007) Evolutionary tinkering with mitochondrial nucleoids. *Trends Cell Biol* **17**, 586-592.
15. Bock, R. (2007) Structure, function, and inheritance of plastid genomes in *Cell and Molecular Biology of Plastids*, ed. Bock, R. (Springer), pp. 29-64.
16. Palmer, J. D. (1983) Chloroplast DNA exists in two orientations. *Nature* **301**, 92-93.
17. Bendich, A. J. (2004) Circular chloroplast chromosomes: the grand illusion. *Plant Cell* **16**, 1661-1666.
18. Stegemann, S. & Bock, R. (2006) Experimental reconstruction of functional gene transfer from the tobacco plastid genome to the nucleus. *Plant Cell* **18**, 2869-2878.
19. Glaser, E. a. W., J (2007) Import of nuclear-encoded mitochondrial proteins in *Plant Mitochondria*, ed. Logan, D. (Blackwell Publishing Ltd.), pp. 97-140.
20. Balsera, M., Soll, J., & Bolter, B. (2009) Protein import machineries in endosymbiotic organelles. *Cell Mol Life Sci* **66**, 1903-1923.
21. Carrie, C., Giraud, E., & Whelan, J. (2009) Protein transport in organelles: Dual targeting of proteins to mitochondria and chloroplasts. *Febs J* **276**, 1187-1195.

22. Day, A., Madesis, P. (2007) DNA replication, recombination, and repair in plastids in *Cell and Molecular Biology of Plastids*, ed. Bock, R. (Springer), pp. 65-119.
23. Kolodner, R. D. & Tewari, K. K. (1975) Chloroplast DNA from higher plants replicates by both the Cairns and the rolling circle mechanism. *Nature* **256**, 708-711.
24. Lilly, J. W., Havey, M. J., Jackson, S. A., & Jiang, J. (2001) Cytogenomic analyses reveal the structural plasticity of the chloroplast genome in higher plants. *Plant Cell* **13**, 245-254.
25. Oldenburg, D. J. & Bendich, A. J. (2004) Most chloroplast DNA of maize seedlings in linear molecules with defined ends and branched forms. *J Mol Biol* **335**, 953-970.
26. Bendich, A. J. (1993) Reaching for the ring: the study of mitochondrial genome structure. *Curr Genet* **24**, 279-290.
27. Muhlbauer, S. K., Lossl, A., Tzekova, L., Zou, Z., & Koop, H. U. (2002) Functional analysis of plastid DNA replication origins in tobacco by targeted inactivation. *Plant J* **32**, 175-184.
28. Scharff, L. B. & Koop, H. U. (2007) Targeted inactivation of the tobacco plastome origins of replication A and B. *Plant J* **50**, 782-794.
29. Kimura, S., Uchiyama, Y., Kasai, N., Namekawa, S., Saotome, A., Ueda, T., Ando, T., Ishibashi, T., Oshige, M., Furukawa, T., Yamamoto, T., Hashimoto, J., & Sakaguchi, K. (2002) A novel DNA polymerase homologous to Escherichia coli DNA polymerase I from a higher plant, rice (*Oryza sativa* L.). *Nucleic Acids Res* **30**, 1585-1592.
30. Mori, Y., Kimura, S., Saotome, A., Kasai, N., Sakaguchi, N., Uchiyama, Y., Ishibashi, T., Yamamoto, T., Chiku, H., & Sakaguchi, K. (2005) Plastid DNA polymerases from higher plants, *Arabidopsis thaliana*. *Biochem Biophys Res Commun* **334**, 43-50.
31. Ono, Y., Sakai, A., Takechi, K., Takio, S., Takusagawa, M., & Takano, H. (2007) NtPoll-like1 and NtPoll-like2, bacterial DNA polymerase I homologs isolated from BY-2 cultured tobacco cells, encode DNA polymerases engaged in DNA replication in both plastids and mitochondria. *Plant Cell Physiol* **48**, 1679-1692.
32. Elo, A., Lyznik, A., Gonzalez, D. O., Kachman, S. D., & 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.
33. Wamboldt, Y., Mohammed, S., Elowsky, C., Wittgren, C., de Paula, W. B., & Mackenzie, S. A. (2009) Participation of leaky ribosome scanning in protein dual targeting by alternative translation initiation in higher plants. *Plant Cell* **21**, 157-167.
34. Takeuchi, R., Kimura, S., Saotome, A., & Sakaguchi, K. (2007) Biochemical properties of a plastidial DNA polymerase of rice. *Plant Mol Biol* **64**, 601-611.
35. Wall, M. K., Mitchenall, L. A., & Maxwell, A. (2004) *Arabidopsis thaliana* DNA gyrase is targeted to chloroplasts and mitochondria. *Proc Natl Acad Sci USA* **101**, 7821-7826.
36. Sugino, A., Peebles, C. L., Kreuzer, K. N., & Cozzarelli, N. R. (1977) Mechanism of action of nalidixic acid: purification of Escherichia coli nalA gene product and its relationship to DNA gyrase and a novel nicking-closing enzyme. *Proc Natl Acad Sci U S A* **74**, 4767-4771.
37. Cho, H. S., Lee, S. S., Kim, K. D., Hwang, I., Lim, J. S., Park, Y. I., & Pai, H. S. (2004) DNA gyrase is involved in chloroplast nucleoid partitioning. *Plant Cell* **16**, 2665-2682.
38. Kato, J., Nishimura, Y., Imamura, R., Niki, H., Hiraga, S., & Suzuki, H. (1990) New topoisomerase essential for chromosome segregation in *E. coli*. *Cell* **63**, 393-404.
39. Muller, H. J. (1964) The Relation of Recombination to Mutational Advance. *Mutat Res* **106**, 2-9.
40. Wolfe, K. H., Li, W. H., & Sharp, P. M. (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc Natl Acad Sci U S A* **84**, 9054-9058.
41. Khakhlova, O. & Bock, R. (2006) Elimination of deleterious mutations in plastid genomes by gene conversion. *Plant J* **46**, 85-94.

42. Race, H. L., Herrmann, R. G., & Martin, W. (1999) Why have organelles retained genomes? *Trends Genet* **15**, 364-370.
43. Spampinato, C. P., Gomez, R. L., Galles, C., & Lario, L. D. (2009) From bacteria to plants: a compendium of mismatch repair assays. *Mutat Res* **682**, 110-128.
44. Liu, P. & Demple, B. (2010) DNA repair in mammalian mitochondria: Much more than we thought? *Environ Mol Mutagen* **51**, 417-426.
45. Mason, P. A., Matheson, E. C., Hall, A. G., & Lightowers, R. N. (2003) Mismatch repair activity in mammalian mitochondria. *Nucleic Acids Res* **31**, 1052-1058.
46. de Souza-Pinto, N. C., Mason, P. A., Hashiguchi, K., Weissman, L., Tian, J., Guay, D., Lebel, M., Stevensner, T. V., Rasmussen, L. J., & Bohr, V. A. (2009) Novel DNA mismatch-repair activity involving YB-1 in human mitochondria. *DNA Repair (Amst)* **8**, 704-719.
47. Chi, N. W. & Kolodner, R. D. (1994) Purification and characterization of MSH1, a yeast mitochondrial protein that binds to DNA mismatches. *J Biol Chem* **269**, 29984-29992.
48. Mookerjee, S. A., Lyon, H. D., & Sia, E. A. (2005) Analysis of the functional domains of the mismatch repair homologue Msh1p and its role in mitochondrial genome maintenance. *Curr Genet* **47**, 84-99.
49. Pogorzala, L., Mookerjee, S., & Sia, E. A. (2009) Evidence that msh1p plays multiple roles in mitochondrial base excision repair. *Genetics* **182**, 699-709.
50. Dong, C., Whitford, R., & Langridge, P. (2002) A DNA mismatch repair gene links to the Ph2 locus in wheat. *Genome* **45**, 116-124.
51. Lloyd, A. H., Milligan, A. S., Langridge, P., & Able, J. A. (2007) TaMSH7: a cereal mismatch repair gene that affects fertility in transgenic barley (*Hordeum vulgare* L.). *BMC Plant Biol* **7**, 67.
52. Abdelnoor, R. V., Yule, R., Elo, A., Christensen, A. C., Meyer-Gauen, G., & 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.
53. Chen, J. J., Mitchell, D. L., & Britt, A. B. (1994) A Light-Dependent Pathway for the Elimination of UV-Induced Pyrimidine (6-4) Pyrimidinone Photoproducts in Arabidopsis. *Plant Cell* **6**, 1311-1317.
54. Kimura, S. & Sakaguchi, K. (2006) DNA repair in plants. *Chem Rev* **106**, 753-766.
55. Draper, C. K. & Hays, J. B. (2000) Replication of chloroplast, mitochondrial and nuclear DNA during growth of unirradiated and UVB-irradiated Arabidopsis leaves. *Plant J* **23**, 255-265.
56. Petersen, J. L. & Small, G. D. (2001) A gene required for the novel activation of a class II DNA photolyase in Chlamydomonas. *Nucleic Acids Res* **29**, 4472-4481.
57. Kunz, B. A., Anderson, H. J., Osmond, M. J., & Vonarx, E. J. (2005) Components of nucleotide excision repair and DNA damage tolerance in Arabidopsis thaliana. *Environ Mol Mutagen* **45**, 115-127.
58. Rechkunova, N. I., Mal'tseva, E. A., & Lavrik, O. I. (2008) [Nucleotide excision repair in mammalia: mechanism of a primary damage recognition]. *Mol Biol (Mosk)* **42**, 24-31.
59. Richard, D. J., Bolderson, E., & Khanna, K. K. (2009) Multiple human single-stranded DNA binding proteins function in genome maintenance: structural, biochemical and functional analysis. *Crit Rev Biochem Mol Biol* **44**, 98-116.
60. Boesch, P., Ibrahim, N., Paulus, F., Cosset, A., Tarasenko, V., & Dietrich, A. (2009) Plant mitochondria possess a short-patch base excision DNA repair pathway. *Nucleic Acids Res* **37**, 5690-5700.
61. Bensen, R. J. & Warner, H. R. (1987) Partial Purification and Characterization of Uracil-DNA Glycosylase Activity from Chloroplasts of Zea mays Seedlings. *Plant Physiol* **84**, 1102-1106.
62. Gutman, B. L. & Niyogi, K. K. (2009) Evidence for base excision repair of oxidative DNA damage in chloroplasts of Arabidopsis thaliana. *J Biol Chem* **284**, 17006-17012.

63. Maréchal, A. & Brisson, N. (2010) Recombination and the maintenance of plant organelle genome stability. *New Phytol* **186**, 299-317.
64. Kwon, T., Huq, E., & Herrin, D. L. (2010) Microhomology-mediated and nonhomologous repair of a double-strand break in the chloroplast genome of Arabidopsis. *Proc Natl Acad Sci U S A* **107**, 13954-13959.
65. Arrieta-Montiel, M. P., Shedje, V., Davila, J., Christensen, A. C., & Mackenzie, S. A. (2009) Diversity of the Arabidopsis mitochondrial genome occurs via nuclear-controlled recombination activity. *Genetics* **183**, 1261-1268.
66. Shedje, V., Davila, J., Arrieta-Montiel, M. P., Mohammed, S., & Mackenzie, S. A. (2010) Extensive rearrangement of the Arabidopsis mitochondrial genome elicits cellular conditions for thermotolerance. *Plant Physiol* **152**, 1960-1970.
67. Mahaney, B. L., Meek, K., & Lees-Miller, S. P. (2009) Repair of ionizing radiation-induced DNA double-strand breaks by non-homologous end-joining. *Biochem J* **417**, 639-650.
68. McVey, M. & Lee, S. E. (2008) MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends Genet* **24**, 529-538.
69. Gherbi, H., Gallego, M. E., Jalut, N., Lucht, J. M., Hohn, B., & White, C. I. (2001) Homologous recombination in planta is stimulated in the absence of Rad50. *EMBO Rep* **2**, 287-291.
70. West, C. E., Waterworth, W. M., Jiang, Q., & Bray, C. M. (2000) Arabidopsis DNA ligase IV is induced by gamma-irradiation and interacts with an Arabidopsis homologue of the double strand break repair protein XRCC4. *Plant J* **24**, 67-78.
71. West, C. E., Waterworth, W. M., Story, G. W., Sunderland, P. A., Jiang, Q., & Bray, C. M. (2002) Disruption of the Arabidopsis AtKu80 gene demonstrates an essential role for AtKu80 protein in efficient repair of DNA double-strand breaks in vivo. *Plant J* **31**, 517-528.
72. Kohl, S. & Bock, R. (2009) Transposition of a bacterial insertion sequence in chloroplasts. *Plant J* **58**, 423-436.
73. Maréchal, A. (2009) in *Biochimie* (Université de Montréal, Montréal).
74. Szostak, J. W., Orr-Weaver, T. L., Rothstein, R. J., & Stahl, F. W. (1983) The double-strand-break repair model for recombination. *Cell* **33**, 25-35.
75. Odom, O. W., Baek, K. H., Dani, R. N., & Herrin, D. L. (2008) Chlamydomonas chloroplasts can use short dispersed repeats and multiple pathways to repair a double-strand break in the genome. *Plant J* **53**, 842-853.
76. Llorente, B., Smith, C. E., & Symington, L. S. (2008) Break-induced replication: what is it and what is it for? *Cell Cycle* **7**, 859-864.
77. Hastings, P. J., Ira, G., & Lupski, J. R. (2009) A microhomology-mediated break-induced replication model for the origin of human copy number variation. *PLoS Genet* **5**, e1000327.
78. Lambert, S., Froget, B., & Carr, A. M. (2007) Arrested replication fork processing: interplay between checkpoints and recombination. *DNA Repair (Amst)* **6**, 1042-1061.
79. Kraus, E., Leung, W. Y., & Haber, J. E. (2001) Break-induced replication: a review and an example in budding yeast. *Proc Natl Acad Sci U S A* **98**, 8255-8262.
80. Smith, C. E., Llorente, B., & Symington, L. S. (2007) Template switching during break-induced replication. *Nature* **447**, 102-105.
81. Sudmant, P. H., Kitzman, J. O., Antonacci, F., Alkan, C., Malig, M., Tsalenko, A., Sampas, N., Bruhn, L., Shendure, J., & Eichler, E. E. Diversity of human copy number variation and multicopy genes. *Science* **330**, 641-646.
82. Zhang, F., Carvalho, C. M., & Lupski, J. R. (2009) Complex human chromosomal and genomic rearrangements. *Trends Genet* **25**, 298-307.
83. Redei, G. P. & Plurad, S. B. (1973) Hereditary structural alterations of plastids induced by a nuclear mutator gene in Arabidopsis. *Protoplasma* **77**, 361-380.

84. Martinez-Zapater, J. M., Gil, P., Capel, J., & Somerville, C. R. (1992) Mutations at the Arabidopsis CHM locus promote rearrangements of the mitochondrial genome. *Plant Cell* **4**, 889-899.
85. Christensen, A. C., Lyznik, A., Mohammed, S., Elowsky, C. G., Elo, A., Yule, R., & Mackenzie, S. A. (2005) Dual-domain, dual-targeting organellar protein presequences in Arabidopsis can use non-AUG start codons. *Plant Cell* **17**, 2805-2816.
86. Abdelnoor, R. V., Christensen, A. C., Mohammed, S., Munoz-Castillo, B., Moriyama, H., & Mackenzie, S. A. (2006) Mitochondrial genome dynamics in plants and animals: convergent gene fusions of a MutS homologue. *J Mol Evol* **63**, 165-173.
87. Jun, S. H., Kim, T. G., & Ban, C. (2006) DNA mismatch repair system. Classical and fresh roles. *Febs J* **273**, 1609-1619.
88. Cox, M. M. (2007) Regulation of bacterial RecA protein function. *Crit Rev Biochem Mol Biol* **42**, 41-63.
89. Bell, C. E. (2005) Structure and mechanism of Escherichia coli RecA ATPase. *Mol Microbiol* **58**, 358-366.
90. Khazi, F. R., Edmondson, A. C., & Nielsen, B. L. (2003) An Arabidopsis homologue of bacterial RecA that complements an E. coli recA deletion is targeted to plant mitochondria. *Mol Genet Genomics* **269**, 454-463.
91. Shedge, V., Arrieta-Montiel, M., Christensen, A. C., & Mackenzie, S. A. (2007) Plant mitochondrial recombination surveillance requires unusual RecA and MutS homologs. *Plant Cell* **19**, 1251-1264.
92. Odahara, M., Kuroiwa, H., Kuroiwa, T., & Sekine, Y. (2009) Suppression of repeat-mediated gross mitochondrial genome rearrangements by RecA in the moss *Physcomitrella patens*. *Plant Cell* **21**, 1182-1194.
93. Rowan, B. A., Oldenburg, D. J., & Bendich, A. J. (2010) RecA maintains the integrity of chloroplast DNA molecules in Arabidopsis. *J Exp Bot* **61**, 2575-2588.
94. Shaver, J. M., Oldenburg, D. J., & Bendich, A. J. (2006) Changes in chloroplast DNA during development in tobacco, *Medicago truncatula*, pea, and maize. *Planta* **224**, 72-82.
95. Chen, C. R., Malik, M., Snyder, M., & Drlica, K. (1996) DNA gyrase and topoisomerase IV on the bacterial chromosome: quinolone-induced DNA cleavage. *J Mol Biol* **258**, 627-637.
96. Vermel, M., Guermann, B., Delage, L., Grienberger, J. M., Marechal-Drouard, L., & 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.
97. Zaegel, V., Guermann, B., Le Ret, M., Andres, C., Meyer, D., Erhardt, M., Canaday, J., Gualberto, J. M., & 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.
98. Shereda, R. D., Kozlov, A. G., Lohman, T. M., Cox, M. M., & Keck, J. L. (2008) SSB as an organizer/mobilizer of genome maintenance complexes. *Crit Rev Biochem Mol Biol* **43**, 289-318.
99. Glover, B. P. & McHenry, C. S. (1998) The chi psi subunits of DNA polymerase III holoenzyme bind to single-stranded DNA-binding protein (SSB) and facilitate replication of an SSB-coated template. *J Biol Chem* **273**, 23476-23484.
100. Yuzhakov, A., Kelman, Z., & O'Donnell, M. (1999) Trading places on DNA--a three-point switch underlies primer handoff from primase to the replicative DNA polymerase. *Cell* **96**, 153-163.
101. Shereda, R. D., Bernstein, D. A., & Keck, J. L. (2007) A central role for SSB in Escherichia coli RecQ DNA helicase function. *J Biol Chem* **282**, 19247-19258.

102. Slocum, S. L., Buss, J. A., Kimura, Y., & Bianco, P. R. (2007) Characterization of the ATPase activity of the Escherichia coli RecG protein reveals that the preferred cofactor is negatively supercoiled DNA. *J Mol Biol* **367**, 647-664.
103. Cadman, C. J. & McGlynn, P. (2004) PriA helicase and SSB interact physically and functionally. *Nucleic Acids Res* **32**, 6378-6387.
104. Lu, D. & Keck, J. L. (2008) Structural basis of Escherichia coli single-stranded DNA-binding protein stimulation of exonuclease I. *Proc Natl Acad Sci U S A* **105**, 9169-9174.
105. Inoue, J., Nagae, T., Mishima, M., Ito, Y., Shibata, T., & Mikawa, T. (2010) A mechanism for SSB displacement from single-stranded DNA upon SSB-RecO interaction. *J Biol Chem*.
106. Handa, P., Acharya, N., & Varshney, U. (2001) Chimeras between single-stranded DNA-binding proteins from Escherichia coli and Mycobacterium tuberculosis reveal that their C-terminal domains interact with uracil DNA glycosylases. *J Biol Chem* **276**, 16992-16997.
107. Wobbe, C. R., Weissbach, L., Borowiec, J. A., Dean, F. B., Murakami, Y., Bullock, P., & Hurwitz, J. (1987) Replication of simian virus 40 origin-containing DNA in vitro with purified proteins. *Proc Natl Acad Sci U S A* **84**, 1834-1838.
108. Maga, G., Frouin, I., Spadari, S., & Hubscher, U. (2001) Replication protein A as a "fidelity clamp" for DNA polymerase alpha. *J Biol Chem* **276**, 18235-18242.
109. Bae, S. H., Bae, K. H., Kim, J. A., & Seo, Y. S. (2001) RPA governs endonuclease switching during processing of Okazaki fragments in eukaryotes. *Nature* **412**, 456-461.
110. Stauffer, M. E. & Chazin, W. J. (2004) Physical interaction between replication protein A and Rad51 promotes exchange on single-stranded DNA. *J Biol Chem* **279**, 25638-25645.
111. Mer, G., Bochkarev, A., Gupta, R., Bochkareva, E., Frappier, L., Ingles, C. J., Edwards, A. M., & Chazin, W. J. (2000) Structural basis for the recognition of DNA repair proteins UNG2, XPA, and RAD52 by replication factor RPA. *Cell* **103**, 449-456.
112. Coverley, D., Kenny, M. K., Lane, D. P., & Wood, R. D. (1992) A role for the human single-stranded DNA binding protein HSSB/RPA in an early stage of nucleotide excision repair. *Nucleic Acids Res* **20**, 3873-3880.
113. Stigger, E., Drissi, R., & Lee, S. H. (1998) Functional analysis of human replication protein A in nucleotide excision repair. *J Biol Chem* **273**, 9337-9343.
114. Richard, D. J., Bolderson, E., Cubeddu, L., Wadsworth, R. I., Savage, K., Sharma, G. G., Nicolette, M. L., Tsvetanov, S., McIlwraith, M. J., Pandita, R. K., Takeda, S., Hay, R. T., Gautier, J., West, S. C., Paull, T. T., Pandita, T. K., White, M. F., & Khanna, K. K. (2008) Single-stranded DNA-binding protein hSSB1 is critical for genomic stability. *Nature* **453**, 677-681.
115. Van Dyck, E., Foury, F., Stillman, B., & Brill, S. J. (1992) A single-stranded DNA binding protein required for mitochondrial DNA replication in S. cerevisiae is homologous to E. coli SSB. *Embo J* **11**, 3421-3430.
116. Maier, D., Farr, C. L., Poeck, B., Alahari, A., Vogel, M., Fischer, S., Kaguni, L. S., & Schneuwly, S. (2001) Mitochondrial single-stranded DNA-binding protein is required for mitochondrial DNA replication and development in Drosophila melanogaster. *Mol Biol Cell* **12**, 821-830.
117. Ruhanen, H., Borrie, S., Szabadkai, G., Tynismaa, H., Jones, A. W., Kang, D., Taanman, J. W., & Yasukawa, T. (2010) Mitochondrial single-stranded DNA binding protein is required for maintenance of mitochondrial DNA and 7S DNA but is not required for mitochondrial nucleoid organisation. *Biochim Biophys Acta* **1803**, 931-939.
118. Edmondson, A. C., Song, D., Alvarez, L. A., Wall, M. K., Almond, D., McClellan, D. A., Maxwell, A., & Nielsen, B. L. (2005) Characterization of a mitochondrially targeted single-stranded DNA-binding protein in Arabidopsis thaliana. *Mol Genet Genomics* **273**, 115-122.

119. Desveaux, D., Després, C., Joyeux, A., Subramaniam, R., & Brisson, N. (2000) PBF-2 is a novel single-stranded DNA binding factor implicated in PR-10a gene activation in potato. *Plant Cell* **12**, 1477-1489.
120. Matton, D. P., Prescott, G., Bertrand, C., Camirand, A., & Brisson, N. (1993) Identification of cis-acting elements involved in the regulation of the pathogenesis-related gene STH-2 in potato. *Plant Mol Biol* **22**, 279-291.
121. Després, C., Subramaniam, R., Matton, D. P., & Brisson, N. (1995) The Activation of the Potato PR-10a Gene Requires the Phosphorylation of the Nuclear Factor PBF-1. *Plant Cell* **7**, 589-598.
122. Desveaux, D., Maréchal, A., & Brisson, N. (2005) Whirly transcription factors: defense gene regulation and beyond. *Trends Plant Sci* **10**, 95-102.
123. Desveaux, D., Allard, J., Brisson, N., & Sygusch, J. (2002) A new family of plant transcription factors displays a novel ssDNA-binding surface. *Nat Struct Biol* **9**, 512-517.
124. Desveaux, D., Allard, J., Brisson, N., & Sygusch, J. (2002) Crystallization and preliminary X-ray crystallographic analysis of p24, a component of the potato nuclear factor PBF-2. *Acta Crystallogr D Biol Crystallogr* **58**, 296-298.
125. Desveaux, D., Subramaniam, R., Després, C., Mess, J. N., Lévesque, C., Fobert, P. R., Dangl, J. L., & Brisson, N. (2004) A "Whirly" transcription factor is required for salicylic acid-dependent disease resistance in Arabidopsis. *Dev Cell* **6**, 229-240.
126. Krause, K., Kilbiński, I., Mulisch, M., Rodiger, A., Schafer, A., & Krupinska, K. (2005) DNA-binding proteins of the Whirly family in Arabidopsis thaliana are targeted to the organelles. *FEBS Lett* **579**, 3707-3712.
127. Grabowski, E., Miao, Y., Mulisch, M., & Krupinska, K. (2008) Single-stranded DNA-binding protein Whirly1 in barley leaves is located in plastids and the nucleus of the same cell. *Plant Physiol* **147**, 1800-1804.
128. Krause, K. & Krupinska, K. (2009) Nuclear regulators with a second home in organelles. *Trends Plant Sci* **14**, 194-199.
129. Xiong, J. Y., Lai, C. X., Qu, Z., Yang, X. Y., Qin, X. H., & Liu, G. Q. (2009) Recruitment of AtWHY1 and AtWHY3 by a distal element upstream of the kinesin gene AtKP1 to mediate transcriptional repression. *Plant Mol Biol* **71**, 437-449.
130. Yoo, H. H., Kwon, C., Lee, M. M., & Chung, I. K. (2007) Single-stranded DNA binding factor AtWHY1 modulates telomere length homeostasis in Arabidopsis. *Plant J* **49**, 442-451.
131. Kwon, C. & Chung, I. K. (2004) Interaction of an Arabidopsis RNA-binding protein with plant single-stranded telomeric DNA modulates telomerase activity. *J Biol Chem* **279**, 12812-12818.
132. Maréchal, A., Parent, J. S., Sabar, M., Véronneau-Lafortune, F., Abou-Rached, C., & Brisson, N. (2008) Overexpression of mtDNA-associated AtWhy2 compromises mitochondrial function. *BMC Plant Biol* **8**, 42.
133. Prikryl, J., Watkins, K. P., Friso, G., van Wijk, K. J., & Barkan, A. (2008) A member of the Whirly family is a multifunctional RNA- and DNA-binding protein that is essential for chloroplast biogenesis. *Nucleic Acids Res* **36**, 5152-5165.
134. Jenkins, B. D., Kulhanek, D. J., & Barkan, A. (1997) Nuclear mutations that block group II RNA splicing in maize chloroplasts reveal several intron classes with distinct requirements for splicing factors. *Plant Cell* **9**, 283-296.
135. Keller, S. J., Biedenbach, S. A., & Meyer, R. R. (1973) Partial purification of a chloroplast DNA polymerase from *Euglena gracilis*. *Biochem Biophys Res Commun* **50**, 620-628.
136. Wang, Z. F., Yang, J., Nie, Z. Q., & Wu, M. (1991) Purification and characterization of a gamma-like DNA polymerase from *Chlamydomonas reinhardtii*. *Biochemistry* **30**, 1127-1131.

137. Sala, F., Amileni, A. R., Parisi, B., & Spadari, S. (1980) A gamma-like DNA polymerase in spinach chloroplasts. *Eur J Biochem* **112**, 211-217.
138. Spencer, D. & Whitfeld, P. R. (1969) The characteristics of spinach chloroplast DNA polymerase. *Arch Biochem Biophys* **132**, 477-488.
139. McKown, R. L. & Tewari, K. K. (1984) Purification and properties of a pea chloroplast DNA polymerase. *Proc Natl Acad Sci U S A* **81**, 2354-2358.
140. Keim, C. A. & Mosbaugh, D. W. (1991) Identification and characterization of a 3' to 5' exonuclease associated with spinach chloroplast DNA polymerase. *Biochemistry* **30**, 11109-11118.
141. Castroviejo, M., Tharaud, D., Tarrago-Litvak, L., & Litvak, S. (1979) Multiple deoxyribonucleic acid polymerases from quiescent wheat embryos. Purification and characterization of three enzymes from the soluble cytoplasm and one from purified mitochondria. *Biochem J* **181**, 183-191.
142. Echeverria, M., Robert, D., Carde, J. P., & Litvak, S. (1991) Isolation from wheat mitochondria of a membrane-associated high molecular weight complex involved in DNA synthesis. *Plant Mol Biol* **16**, 301-315.
143. Meissner, K., Heinhorst, S., Cannon, G. C., & Börner, T. (1993) Purification and characterization of a gamma-like DNA polymerase from *Chenopodium album* L. *Nucleic Acids Res* **21**, 4893-4899.
144. Heinhorst, S., Cannon, G. C., & Weissbach, A. (1990) Chloroplast and mitochondrial DNA polymerases from cultured soybean cells. *Plant Physiol* **92**, 939-945.
145. Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J. F., Guindon, S., Lefort, V., Lescot, M., Claverie, J. M., & Gascuel, O. (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* **36**, W465-469.
146. Sato, S., Nakamura, Y., Kaneko, T., Asamizu, E., & Tabata, S. (1999) Complete structure of the chloroplast genome of *Arabidopsis thaliana*. *DNA Res* **6**, 283-290.
147. Unseld, M., Marienfeld, J. R., Brandt, P., & Brennicke, A. (1997) The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. *Nature Genetics* **15**, 57-61.
148. Nielsen, B. L., Cupp, J. D., & Brammer, J. (2010) Mechanisms for maintenance, replication, and repair of the chloroplast genome in plants. *J Exp Bot* **61**, 2535-2537.
149. Lynch, M., Koskella, B., & Schaack, S. (2006) Mutation pressure and the evolution of organelle genomic architecture. *Science* **311**, 1727-1730.
150. Bendich, A. J. (2010) Mitochondrial DNA, chloroplast DNA and the origins of development in eukaryotic organisms. *Biol Direct* **5**, 42.
151. Lydeard, J. R., Lipkin-Moore, Z., Sheu, Y. J., Stillman, B., Burgers, P. M., & Haber, J. E. (2010) Break-induced replication requires all essential DNA replication factors except those specific for pre-RC assembly. *Genes Dev* **24**, 1133-1144.
152. Woelfle, M. A., Thompson, R. J., & Mosig, G. (1993) Roles of novobiocin-sensitive topoisomerases in chloroplast DNA replication in *Chlamydomonas reinhardtii*. *Nucleic Acids Res* **21**, 4231-4238.
153. Rowan, B. A., Oldenburg, D. J., & Bendich, A. J. (2004) The demise of chloroplast DNA in *Arabidopsis*. *Curr Genet* **46**, 176-181.
154. O'Donnell, M. (2006) Replisome architecture and dynamics in *Escherichia coli*. *J Biol Chem* **281**, 10653-10656.
155. Slack, A., Thornton, P. C., Magner, D. B., Rosenberg, S. M., & Hastings, P. J. (2006) On the mechanism of gene amplification induced under stress in *Escherichia coli*. *PLoS Genet* **2**, e48.
156. Coyne, J. A. & Hoekstra, H. E. (2007) Evolution of protein expression: new genes for a new diet. *Curr Biol* **17**, R1014-1016.

157. Mandel, A. L., Peyrot des Gachons, C., Plank, K. L., Alarcon, S., & Breslin, P. A. (2010) Individual Differences in AMY1 Gene Copy Number, Salivary alpha-Amylase Levels, and the Perception of Oral Starch. *PLoS One* **5**.
158. Lynch, M. & Conery, J. S. (2000) The evolutionary fate and consequences of duplicate genes. *Science* **290**, 1151-1155.