UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTE

Institut de Biologie Moléculaire des Plantes – CNRS – UPR2357

THÈSE DE DOCTORAT

présentée par :

Rokas KUBILINSKAS

soutenue le : **14 juin 2024**

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

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

MitoTALENs to explore mitochondrial DNA repair and segregation

THÈSE dirigée par :

Dr. GUALBERTO José Manuel	Directeur de thèse	CNRS, Université de Strasbourg

RAPPORTEURS :

Dr. MIREAU Hakim	Rapporteur externe	INRAE, Université Paris-Saclay
Prof. PUCHTA Holger	Rapporteur externe	Karlsruher Institut für Technologie

AUTRES MEMBRES DU JURY :

Dr. ENTELIS Nina	Examinatrice interne	CNRS, Université de Strasbourg

Acknowledgments

First of all, thank you to all the members of PhD jury Dr. Mireau Hakim, Prof. Puchta Holger and Dr. Entelis Nina, who kindly accepted to evaluate and discuss my thesis work.

A huge thank you to my supervisor, Dr. Gualberto José Manuel, for providing me with such a great opportunity to take part in this project and for guiding me throughout almost four years of my PhD. Thank you for all the help with the French administration. Thank you for your advice, patience, and help with the work of PhD thesis and throughout my entire stay in Strasbourg. Otherwise, all of this would not have been possible. Thank you as well for the hours you spent reviewing and editing this manuscript.

Thank you to all the members of lab group 312, past and present, whom I had a chance to meet: Deborah Schatz Daas, Frédérique Lotfi, Anaïs Le Blevenec, and Elody Vallet. Thank you, Deborah, for such a warm welcome to Strasbourg and for being eager to introduce me to French cuisine. Thank you, Frédérique, for your tips and guidance throughout several experiments. Thank you, Anaïs, for your help with running several wet lab experiments.

Thank you to the 'Analyse de l'expression génique' platform of the IMBP Abdelmalek Alioua and Sandrine Koechler for the training and help in preparing the NGS libraries. Thank you to the IMBP bioinformatic platform for the analysis of the NGS data. A specific thank you to Stéfanie Graindorge for the training courses on the command line and R studio; these skills were essential for my data analysis. An exceptional thank you to David Pflieger for helping to perform all the NGS analyses and to make such nice figures. It was fun and interesting. None of the analyses would be possible without your help. I am grateful to all the gardeners for taking care of my plants. I would like to extend and thank all the IMBP colleagues for their appreciation, encouragement and moral support during my research.

Huge thank you to the StrsAIR society. It gave me so many opportunities to meet various international PhD students and to explore the Alsace region with them. Joining the StrasAir committee board for two years was one of the most interesting experiences that I had. Organising events with such a great team was so much fun: Michael, Tatjana, Dersu, Iryna, Renato, Teodor, Baylar, Sven, Aleksei, Ivan, Nizami. Special thanks to Iryna and Tatjana as, without you, none of the events would have been possible. At StrasAir, I made so many great

friends. Neda buvo malonu susipažinti, deja tik kad į PhD pabaigą. Buvo truputi neramu kad Strasburge nera daugiau lietuvių studentų, ir buvo šaunu sužinoti kad ne aš vienas jį pasirinkau. Vieta tikrai įspūdinga.

Thank you to Ivan, Ira, and Andrei for a marvellous cruise trip to the fjords of Norway. I am really looking forward to our next upcoming trip to Iceland together with Guido, Filipa, and Renato. All of us, together with Tatjana, Michael, and Agata, had a nice time during lunch meet-ups, festive period celebrations, and table game nights. It was an amazing time; thank you to you all. I hope we all will manage to keep in connection further wherever life sends me next. Big thanks to Tatjana for getting me involved in trips to Basel for riddle games (Что? Где? Когда?) and many other events, where we had a great time. Renato, thank you for hosting and organising so many meet-ups and activities at your place; we had a lot of fun! All of you made life in Strasbourg much more colourful and vibrant.

Big thanks to friends abroad! Thank you, Karina and Skirmantas, for hosting me in Norwich and London during my summer visit. It allowed me to flash back UEA experience and meet my sister in London. Спасибо Маше за прекрасный тур Муник-Базель-Марсель. Поездка вышла замечательная успели столько моного увидеть за караткое время.

I would like to thank my parents !! Ačiū jums už palaikymą daryti doktorantūra ir moraline pagalba sunkiais momentais. Atsiparašau kad retai atvažiuodavau ir dėkingas už supratingumą. Bus dar vienas doktorantas šeimoje !! Ugnei ačiū už galimybe susitikti Anglijoje studijų metu, gaila kad nelabai pavykdavo suderinti kėlinių namo laiko.

ABBREVIATIONS

-/-	Homozygous mutant
+/-	Heterozygous mutant
3'-PUA	3' -α, β unsaturated aldehyde
5'-dRP	5'- deoxyribose 5-phosphate
ADP	Adenosine diphosphate
AP	Apurinic/apyrimidinic site
ATP	Adenosine triphosphate
b, bp ,kb	Base, base pair, kilo base pair
BER	Base excision repair
BIR	Break-induced replication
BSA	Bovine serum albumin
cDNA	Complementary DNA
CIP	Ciprofloxacin
CMS	Cytoplasmic male sterility
cpDNA	Chloroplast DNA
CRISPR	Clustered regularly interspaced short palindromic repeats
CTAB	Cetrimonium bromide
cyt c	Cytochrome c
D-loop	Displacement loop
DSB	Double-strand break
DSBR	Double-strand break repair
dsDNA	Double-stranded DNA
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EGMS	Environment controlled genetic male sterility
EGTA	Ethylene glycol diamine tetraacetic acid
EMS	Ethyl methanesulfonate
ETC	Electron transport chain
GFP	Green fluorescence protein
GGR	Global genomic repair
GMS	Genic male sterility

HDR	Homology directed repair
HGT	Horizontal gene transfer
HJ	Holliday junction
HR	Homologous recombination
IGT	Intracellular gene transfer
IMM	Inner mitochondrial membrane
IMS	Intermembrane space
IR	Ionizing radiation
LB	Lysogeny broth medium
LP-BER	Long-patch base excision repair
LR	Large repeat
MHMR	Microhomology-mediated repair/recombination
mitoTALEN	Mitochondria-targeted transcription activator-like effector nuclease
MMBIR	Microhomology-mediated BIR
MMEJ	Microhomology-mediated end-joining
MMR	Mismatch repair
MOPS	3-(N-morpholino)propanesulfonic acid
MS	Murashige and Skoog medium
mtDNA	Mitochondrial DNA
MTS	Mitochondria targeting sequence
NADH	Nicotinamide adenine dinucleotide
NCS	Non-chromosomal stripe
NER	Nucleotide excision repair
NGS	Next generation sequencing
NHEJ	Non-homologous end-joining
NIR	Nucleotide Incision Repair
OHPHOS	Oxidative phosphorylation
OMM	Outer mitochondrial membrane
ORF	Open reading frame
PAP	3'-phosphoadenosine 5'-phosphate
PCR	Polymerase chain reaction
qPCR	Quantitative PCR
RNS	Reactive nitrogen species

ROS	Reactive oxygen species
RT-qPCR	Reverse transcription qPCR
SAM	Sorting and assembly machinery
SDS	Sodium dodecyl sulfate
SDSA	Synthesis-dependent strand annealing
SP-BER	Short-patch base excision repair
SSA	Single-Strand Annealing
SSB	Single-stranded DNA binding protein
ssDNA	Single-stranded DNA
SSG	Single strand gap
SSGR	Single strand gap repair
SSS	Substoichiometric shifting
T-DNA	Transfer DNA
TALEN	Transcription activator-like effector nucleases
TCA	Krebs cycle or tricarboxylic acid cycle
TCR	Transcription-coupled repair
TEM	Transmission electron microscopy
TES	N-tris [hydro- xymethyl]-methyl-2-aminoethanesulphonic acid
TFAM	Mitochondrial transcription factor A
TIM	Translocase of the inner membrane
TLS	Translesion DNA Synthesis
TOM	Translocase of the outer membrane
VDAC	Voltage-dependant ion channel
ZFN	Zinc finger nucleases

Table of Contents

CHAPTER I: Introduction 9

1. The Mitochondria 10
1.1. The organisation of mitochondria population 10
1.2. Origin of mitochondria 11
1.3. Structure of mitochondria 12
1.4. Physiology of mitochondria 13

2. The mitochondria genome 14
2.1. Size and content 14
2.2. Structure of mitochondrial DNA 16
2.3. Inheritance of mitochondrial genome 18
2.4. Mitochondria packing to nucleoids 19
2.5. Replication of mitochondrial genomes 21
2.6. Cytoplasmic male sterility 22

3. DNA damage 23
3.1. Sources and types of DNA damage 23
3.1.1. UV-induced damage 23
3.1.2. DNA damage from ROS and RNS 24
3.1.3. Alkylating agents 25
3.1.4. Damage from Ionizing radiation 25
3.1.5. Replication errors 26
3.2. Mitochondrial diseases 26

4. DNA repair 27
4.1. DNA repair pathways in the nucleus and organelles 27
4.1.1. DNA repair in humans 28
4.1.2. DNA repair in plants 29
4.2. Nucleotide Excision Repair (NER) 30
4.3. Base excision repair (BER) 32
4.4. Mismatch repair (MMR) 35
4.5. DSB repair 37
4.5.1. Non-Homologous End Joining (NHEJ) 38
4.5.2. Microhomology-Mediated Repair (MHMR) 39
4.5.3. Homologous recombination (HR) 40
4.5.3.1. Homologous recombination in prokaryotes 40
4.5.3.2. Homologous recombination in the nucleus of eukaryotes 42
4.5.3.3. Homologous recombination in the plant mitochondria 43
4.5.4. Synthesis-Dependant Strand Annealing (SDSA) 44
4.5.5. Single Strand Annealing (SSA) 45
4.6. Direct repair 46
4.7. Damage tolerance 46

5. Aim of the thesis 48

CHAPTER II: MitoTALENs, a new method of mtDNA editing 50

Introduction 51

1. Genome engineering 51
1.1. Zinc-Finger Nucleases (ZFN). 51
1.2. Transcription Activator-Like Effector Nucleases (TALEN) 52

1.3. CRISPR/Cas9 52
1.4. Mitochondria Genome Engineering 53

2. Challenges of sequencing the mitochondrial genomes 54

3. Aims of this chapter 55

Results 56

1. DNA repair activation by mitoTALEN 56
1.1. Generation of T1 mitoTALEN lines in WT Col-0 background 56
1.2. qPCR analysis of T1 mitoTALEN seeds 56

2. Mapping Col-0 mitoTALEN Illumina reads to the mtDNA 57
2.1. Partial mitochondrial DNA enrichment 57
2.2. Mapping to the mtDNA 58
2.2.1. DSB repair by Homologous Recombination (HR) 58
2.2.2. DSB repair by Single-Strand Annealing (SSA) 59
2.2.3. DSB repair by Microhomology-Mediated Repair (MHMR) 59

3. Bioinformatic and qPCR analysis for the extent of mtDNA structural variation 60
3.1. Read pair analysis of Illumina alignments 60
3.2. qPCR analysis of repeats involved in DSB repair by HR 61

4. Heteroplasmy of the mtDNA and segregation of structural variants 62
4.1. Heteroplasmy of the mtDNA can lead to fast structural evolution 62
4.2. Differential efficiency between mitoTALENs introduced by Agrobacterium-mediated transformation and by crossing 63
4.3. Differential segregation of mitotypes across plant tissues 63

5. Pilot study: trying to develop new mtDNA editing techniques 64
5.1. mitoTALEN transformation design for simultaneous introduction of two DSBs 65
5.2. Segregation of 2DSB-TAL introduced cut sites in the F1 and F2 generations 65

Discussion 66

1. Designing the mitoTALEN constructs 66

2. DNA library preparation 66

3. Homologous recombination involving two pairs of repeated sequences. 67

4. NGS read pair analysis 68

5. Differences in repair efficiency linked to how mitoTALEN constructs are introduced in the plant 68

6. Pilot study: what went wrong? 69

CHAPTER III: Study of Homologous recombination pathway 70

Introduction 71

1. Single-strand DNA-binding proteins 71

2. RecA-like DNA recombinases 73

3. DNA helicases 73

4. Mismatch recognition protein MSH1 75

5. Mitochondrial exonuclease OEX1 75

6. DNA synthesis by Polymerases 76

Results 77

1. Mitochondria HR T-DNA mutant lines 77
2. Illumina sequencing of mutant lines containing mitoTALEN constructs 77
3. Bioinformatic analysis of NGS data 79
4. Identify repeats involved in the DSB repair by qPCR 81
5. Effect of *pol1* mutants on mtDNA replication 82

Discussion 83

1. Segregating T-DNA lines of HR mitochondrial mutants 83
2. Divergent roles of Rad52 and the plant ODB1 ortholog 83
3. Single-strand DNA-binding protein WHY2 84
4. Loss of HR-dependent repair 85
5. Regulation of homologous recombination 85

CHAPTER IV: Importance of ATP9 subunit for proper function of ATP synthase 87

Introduction 88

1. Oxidative phosphorylation system in plant mitochondria 88
2. ATP synthase structure and function 89
3. ATP synthase linked disorders in humans and plants 90

Results 91

1. Partial loss of ATP9 in mitoTALEN lines leads to developmental defects 91
2. Role of ATP9 in Complex V 92

Discussion 93

Conclusion & Perspectives 94

Materials & Methods 97

Materials 98

1. Bacterial Strains 98
   1.1. Agrobacterium tumefaciens GV3101 pMP90 98
   1.2. Escherichia coli TOP 10 98
2. Arabidopsis thaliana plant lines 98
3. Growth media 99
   3.1. LB medium 99

3.2. SOC medium ... 99
3.3. MS medium ... 99
4. Gateway plasmid vectors ... 99
5. Oligonucleotides ... 99
6. Informatic tools ... 100
6.1. Command line tools ... 100
6.2. R studio packages ... 101
6.3. MacVector (v18.5.1) and its features ... 102
Methods ... 103
1. Quantification and quality assessment of samples ... 103
1.1. Spectrophotometric quantification ... 103
1.1.1. Nucleic acid quantification ... 103
1.1.2. Protein quantification ... 103
1.2. Fluorometric quantification ... 103
2. DNA analysis methods ... 104
2.1. Plant DNA extraction ... 104
2.1.1. DNA extraction with Edwards buffer ... 104
2.1.2. DNA extraction with CTAB ... 104
2.1.3. DNA extraction for sequencing ... 104
2.1.3.1. Mitochondrial enrichment ... 104
2.1.3.2. DNA extraction for Illumina library preparation ... 105
2.2. DNA agarose gel electrophoresis ... 105
2.3. Polymerase chain reaction (PCR) ... 105
2.4. Quantitative polymerase chain reaction (qPCR) ... 106
2.5. Bacterial transformation ... 106
2.6. Plasmid DNA extraction ... 107
2.7. Sequencing ... 107
2.7.1. Sanger sequencing ... 107
2.7.2. Illumina sequencing ... 107
3. RNA analysis methods ... 108
3.1. RNA extraction ... 108
3.1.1. Total RNA extraction with Trizol ... 108
3.1.2. RNA extraction for cDNA preparation ... 108
3.2. cDNA preparation ... 108
3.3. RNA agarose gel electrophoresis ... 108
3.4. Northern blot ... 108
4. Protein analysis methods ... 109
4.1. Protein extraction ... 109
4.1.1. Total plant protein extraction ... 109
4.1.2. Mitochondrial protein extraction ... 110
4.2. SDS-PAGE electrophoresis ... 110
4.3. Western Blot (WB) ... 111
5. Plant Methods ... 111
5.1. Plant growth ... 111
5.2. *A. thaliana* transformation by floral dip ... 111
5.3. Crossing of *A. thaliana* plants ... 112
5.4. Seed selection and embryo extraction under confocal microscope ... 112
5.5. Transmission electron microscopy imaging ... 112
5.6. In vitro methods ... 112
5.6.1. Seeds sterilisation ... 112

5.6.2. Seedlings growth on solid medium 113

6. Bioinformatics analysis 113

6.1. Initial sequencing mapping with MacVector 113

6.2. In-depth analysis of NGS Illumina data 113

6.2.1 Identifying recombination event by read pair analysis. 113

6.2.2. Analysis of Microhomology-Mediated Recombination (MHMR) 114

Bibliography 115

Appendix 149

CHAPTER I: Introduction

1. The Mitochondria

Mitochondria are organelles found in the cytoplasm of almost all eukaryotic cells. However, their number, shape, and size differ within the chondriome and between cell types. The typical size of mitochondria ranges from ~ 0.5 to 3 μm (Wiemerslage and Lee, 2016; Logan, 2010). The first observation of mitochondria was done by German pathologist and historian Richard Altmann (Altmann, 1890). Using acid-fuchsin contrasted by picric acid amid delicate heating, he observed filaments forming from the granular structures in nearly all analysed cell types. The same staining profile is found in bacteria. He assumed granules to be elementary living units and called them “bioblasts”. The well-recognised name “mitochondria” was given later by Carl Benda, deriving it from the Greek words “mitos” (thread) and “chondros” (granule) (Benda, 1898). In plants, mitochondria were observed later, in 1904 (Meves, 1904)

1.1. The organisation of mitochondria population

Not long after the discovery of mitochondria, Lewis and Lewis noticed that mitochondria are ‘never at rest, are continually changing their position and shape’ (Lewis and Lewis, 1914). However, this was forgotten or neglected for long. Now we know that mitochondria form a dynamic structure called chondriome, a name given by French cytologists to the entire mitochondrial contents (population) of a cell (Guilliermond, 1911). Mitochondria were found to be highly mobile and can merge and divide by fusion and fission (Sukhorukov *et al.,* 2012). In yeast, the shape of mitochondria varies highly depending on the culture conditions and stage of life (Sando *et al.,* 1981; Miyakawa, 2017). In the log phase with a nutrient-rich medium, mitochondria are tubular due to interaction with microtubules, while in the stationary phase or anaerobic conditions, they are small and dispersed in the cytoplasm. The number of mitochondria also varies significantly. Yeast cells grown in an ethanol medium carry around 30 mitochondria, while on a glucose-rich fermentation medium, just 2 or 3 (Visser *et al.,* 1995). In humans, as in yeast, mitochondria dynamics are closely linked to microtubules that form the cytoskeleton. Most often, they form a constantly reshaping tubular reticulum web spread over the cytosol (Bereiter-Hahn and Vöth, 1994; Sukhorukov *et al.,* 2012). The number of mitochondria in humans highly varies, depending on cell type and function. For example, blood cells erythrocytes have no mitochondria. In contrast, liver cells can contain 500-4000 per cell (Degli Esposti et al., 2012) and the HeLa cell line, the most frequently used immortalised human cell line in research, has 383-882 per cell (Posakony *et al.,* 1977). However, the homogeneous distribution of heteroplasmic mitochondria (which may differ in

mtDNA content) suggests that they lost their individuality and, independent of the actual number, should be considered as a single dynamic unit (Hayashi *et al.,* 1994).

However, different from animals and yeast, in plants, F-actin filaments and, more broadly, cytoskeleton were shown to be responsible for the translocation of mitochondria in plants (Doniwa *et al.,* 2007; Van Gestel *et al.,* 2002). As in other species could be found in spherical and tubular forms. The number of mitochondria in plants varies little under controlled growth conditions, yet the copy number of mtDNA might differ according to developmental stage and cell type: increase by 1.5 types during the growth of cotyledons to mature leaves (Preuten *et al.,* 2010) and have a 50-fold reduction during pollen development (Sodmergen *et al.,* 2002). The difference in the number of mitochondria can also be observed between plant species: Arabidopsis mesophyll contains 200-300 mitochondria per cell (Logan, 2006), while tobacco mesophyll has 500-600 per cell (Sheahan *et al.,* 2004). Furthermore, in plants, as in most eukaryotes, there is a maternal inheritance pattern for mitochondria transmission to the next generation (Birky Jr., 2001). However, some plant species like rapeseed (Oshima and Handa, 2012) and cucumber (Havey *et al.,* 2004) do show the paternal inheritance of mitochondria and biparental like in Pelargonium (Weihe *et al.,* 2009) or Chinese pine (Guo *et al.,* 2005). Regardless of the number of mitochondria in the cell, they have been shown to continuously undergo fusion and fission, leading to rapid mixing of their contents. Additionally, it has been observed that not all mitochondria contain DNA, and during fusion and fission, there is unequal transmission of nucleoids (Arimura *et al.,* 2004).

1.2. Origin of mitochondria

Contrary to other cellular organelles, mitochondria and chloroplasts have their own DNA. The endosymbiotic theory was first developed in 1905 for chloroplasts, assuming that they were derived from bacterial ancestors (Mereschkowsky, 1910). However, after 81 years, it was adapted by Lynn Margulis to mitochondria and remains widely accepted in biomedical sciences (Sagan, 1967; Margulis, 1981). She assumed that in an oxygen-containing atmosphere, heterotrophic anaerobic prokaryotes had to ingest aerobic prokaryotic microbe into the cytoplasm to survive. Currently, it is believed that the proto-mitochondrion ancestor was an *Alphaproteobacteria*, found to be related to the branch of the intracellular parasite *Rickettsia prowazekii* (Andersson *et al.,* 1998). *Alphaproteobacteria* is a large class of *Proteobacteria*, primarily known for photosynthetic, pathogenic and nitrogen-fixing capacities (Cevallos and Degli Esposti, 2022; Hördt *et al.,* 2020). However, it remains unclear and controversial from which exact lineage mitochondria have evolved. Controversially, to widely accepted opinion, recent phylogenetic studies suggest that mitochondria evolved from extinct

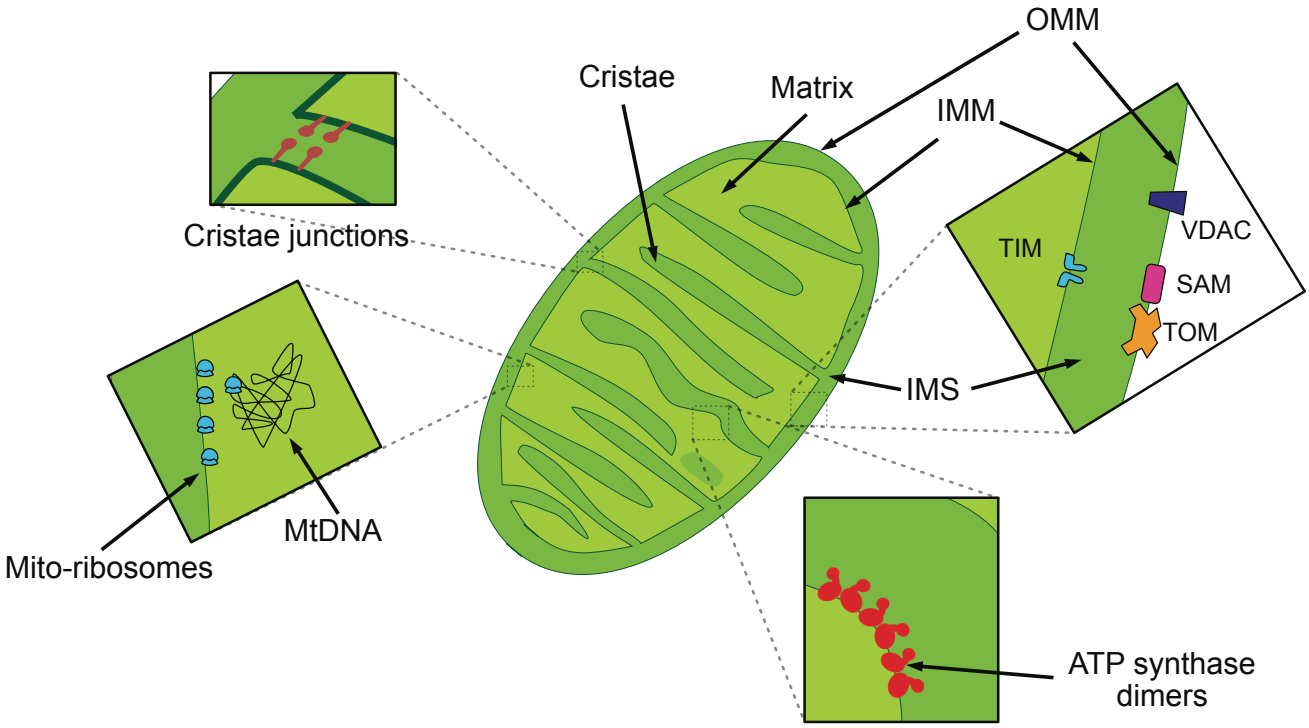


Figure I. 1. Schematic view of mitochondria structural components. The mitochondria are surrounded by two membranes that define several compartments: the intermembrane space (IMS) between the inner mitochondrial membrane (IMM) and the outer mitochondrial membrane (OMM), and the mitochondrial matrix inside the IMM. The transport of molecules across the OMM is assured by TOM (translocase of the outer membrane), SAM (sorting and assembly machinery) and VDAC (voltage-dependant ion channel), while transport across IMM is facilitated by translocase of the inner membrane (TIM). ATP synthase dimerization promotes the formation of IMM invaginations called cristae, and junctions at the neck impede the flow of proteins and metabolites with the rest of the matrix, creating a highly acidic environment needed for the electron transport chain (ETC) (Fig I. 2). Finally, mtDNA forms nucleoids, protein-DNA complexes that are most likely anchored to the IMM, together with mito-ribosomes, facilitating protein expression and complexes assembly.

proto-mitochondrion, outside the core clade of *Alphaproteobacteria* (Martijn *et al.,* 2018; Muñoz-Gómez *et al.,* 2022). However, such a hypothesis does not explain why *Alphaproteobacteria* and mitochondria share a large number of similar metabolic traits, suggesting incorrect phylogenies (Degli Esposti, 2023).

1.3. Structure of mitochondria

In 1960, there was a high controversy between two models of mitochondria structure revealed by electron microscope. The first model by Sjöstrand was fully rejected. It stated that mitochondria have a double membrane, with an internal membrane forming septa that divide the matrix into many compartments (Sjöstrand, 1956). The second model presented by Palade later evolved into what is currently accepted (Palade, 1952). According to his model, mitochondria have two aqueous sub-compartments: intermembrane space (IMS) between the outer (OMM) and the inner membrane (IMM) and mitochondrial matrix inside the IMM. However, the model was misleading as the mitochondrial matrix solvent is 3 times more viscous than water, implying being not aqueous. Furthermore, later studies discovered junctions at the neck of cristae that impede the flow of proteins and metabolites and, contrary to the Palade model, create aqueous compartments different from IMS and separated from the mitochondrial matrix (Perkins *et al.,* 1998; Perkins *et al.,* 1997; Weissig and Edeas, 2021) (Fig. I.1). Recent studies have shown that the dimerization of ATP synthase facilitates cristae formation (Blum *et al.,* 2019).

Most of the mitochondrial proteins are encoded in the nucleus and transported to mitochondria via the N-terminus targeting sequence. Thus, mitochondria have developed a transport system for the exchange of proteins as well as metabolites and ions for the proper functioning of mitochondria (Becker and Wagner, 2018). The outer membrane has a small number of channels with a broad substrate specificity. The majority could be sorted into three protein families: TOM (translocase of the outer membrane), SAM (sorting and assembly machinery) and VDAC (voltage-dependant ion channel) (Fig. I.1). TOM is responsible for the import of most mitochondria-targeted proteins (Becker *et al.,* 2005). SAM mediates the insertion of β-barrel proteins into the outer membrane(Kutik *et al.,* 2008). VDAC is a transport channel of hydrophilic ions and metabolites up to 5 kDa, driven by the concentration gradient across the outer membrane (Campo et al., 2017; Hiller et al., 2010). The transport of inner mitochondria differs in many aspects. It is a tight diffusion barrier to all ions and molecules. The potential across the mitochondrial membrane could build up to 180 mV (Hajnóczky *et al.,* 2006). To pass metabolites and proteins across the IMM, they must have precursors' targeting signals recognised by the translocase of the inner membrane (TIM)(Bauer *et al.,* 2000).

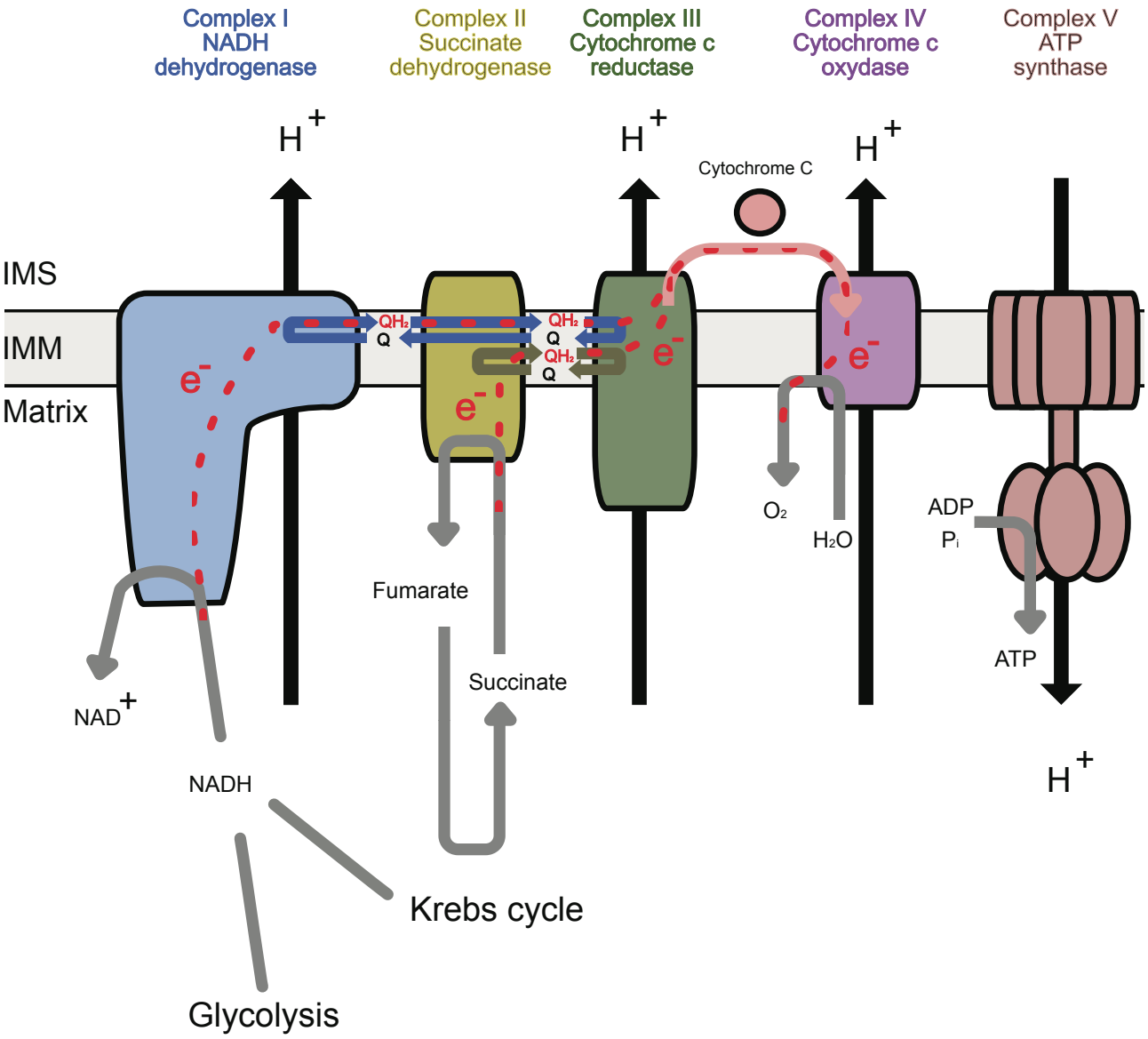


Figure I. 2. Schematic representation of electron transport chain (ETC) in mitochondria. NADH from glycolysis and the Krebs cycle feeds ETC with electrons via Complex I. Complex II acquires electrons from succinate from the Krebs cycle. Electron transfer via Complex I, III and IV is coupled with a proton pump that creates an electrochemical gradient across the inner membrane (IM). This gradient drives ATP synthase activity to produce ATP from ADP and inorganic phosphate. IMM: inner mitochondrial membrane; Q: ubiquinone; QH_2: ubiquinol. Figure prepared according to Sousa et al., 2018.

1.4. Physiology of mitochondria

The first observation of cell respiration was made in 1884 by the detection of "respiratory pigments" that were later named "cytochromes" (Kalin, 1925). In 1930, this led to the idea of a "respiratory chain" (Ernster and Schatz, 1981). Shortly after, Hans Adolf Krebs, during the analysis of ketonic acid, proposed a mechanism of the Krebs cycle (also known as citric acid or TCA cycle) (Krebs and Johnson, 1937). The first attempts to isolate mitochondria (Bensley and Hoerr, 1934) led to the first isolated, well-preserved mitochondria (Hogeboom *et al.,* 1948), significantly boosting the further biochemical and molecular study of mitochondria to identify its functions and roles in the cell. Soon after, a link was found between the aerobic oxidation of the TCA cycle intermediates and fatty acids with the formation of adenosine triphosphate (ATP) from inorganic phosphate and adenosine diphosphate (ADP) (Kennedy and Lehninger, 1949). Furthermore, the phosphorylation of ADP to ATP was found to be coupled with the oxidation of nicotinamide adenine dinucleotide (NADH) (Friedkin and Lehninger, 1949). These findings, together with the discovery of other respiratory chain compounds linked to ATP production, made mitochondria receive the nickname "the powerhouse of the cell" (Siekevitz, 1957).

To conclude, according to up-to-date knowledge, in nearly all eukaryotes, energy is primarily produced via aerobic respiration in mitochondria. The process usually starts with glycolysis in the cytosol, where glucose is metabolized. This most abundant carbohydrate can come from dietary sources or from energy-storage compounds like glycogen (in humans) or starch (in plants). Pyruvate is then transported into mitochondria and converted to acetyl-CoA, which enters the TCA cycle. As a result, NADH from glycolysis and TCA, together with succinate, feeds the Electron Transport Chain (ETC) with electrons (Sousa *et al.,* 2018) (Fig. I.2). The ETC utilizes the energy to pump protons across IMM into the IMS, establishing an electrochemical gradient that drives ATP synthase. A more detailed overview of ETC is in Chapter IV.

2. The mitochondria genome

2.1. Size and content

As previously described, mitochondria are believed to have an endosymbiotic origin, which has brought many differences across kingdoms and species over the course of evolution. The size of the mtDNA varies significantly among and within lineages. At one extreme, we have the largest sequenced mitochondrial genomes of Siberian larch (*Larix sibirica* Ledeb.) (Putintseva *et al.,* 2020) and *Silene conica*(Sloan *et al.,* 2012), with size estimates of 11.7 Mbp and 11.3 Mbp, respectively. On the other end, the smallest mtDNA is of 6 kb in size, found in apicomplexan parasites (Hikosaka *et al.,* 2010), and less than 13 kb found in some green algae (Smith *et al.,* 2010) and fungi (Pramateftaki *et al.,* 2006). There is evidence suggesting that some microbial eukaryotes might have mitochondria that lost their genome. Such mitochondria-derived organelles (mitosomes and hydrogenosomes) have been seen in several anaerobes and microaerophiles, such as *Giardia*, *Trichomonas*, *Mikrocytos*, and *Entamoeba* (Hjort *et al.,* 2010). Despite variation within different kingdoms, mtDNA size within the mammalian kingdom is quite conserved (about 15–20 kb) (Boore, 1999): 16 kb in humans (Anderson *et al.,* 1981) and rats (Laulederkind *et al.,* 2023). It is assumed that the functional replacement of mitochondrial genes by nuclear copies has effectively stopped in higher animals and thus has led to nearly constant genome size and gene content. Irrespective of size, animal mitochondrial genomes are gene-dense and highly compact. The small 16 kb human mtDNA codes for 13 proteins, 22 tRNA, and 2 rRNA genes (12S and 16S) (Anderson *et al.,* 1981). In comparison to modern *Alphaproteobacteria* like *Mesorhizobium loti* (which could be a close relative of ancient mitochondrial precursors), it has a 7 Mb genome encoding 6700 proteins, there is a great reduction in both genome size and the number of proteins encoded (Kaneko, 2000). A similar reduction in genome size can be seen in parasites: *Rickettsia prowazeckii,* which has a 1.1 Mb genome, encoding approximately 830 genes (Andersson et al., 1998), and *Mycoplasma genitalium*, with a 550kb genome, encoding around 470 genes (Su *et al.,* 1990). Studies have suggested that during evolution, many genes get relocated from ancestral organelles to the nucleus (Weeden, 1981), and the original mitochondrial orthologue can undergo recombination, mutational decay or deletion(Martin *et al.,* 1998). Yet, genome reduction in parasites and animals is different. Reduction of the genome in organelles is linked to gene transfer to the nucleus, whereas gene loss in parasites is due to the redundancy of genes that can be taken over from the host; thus, to have a copy in one's own genome is no longer needed (Gil *et al.,* 2003).

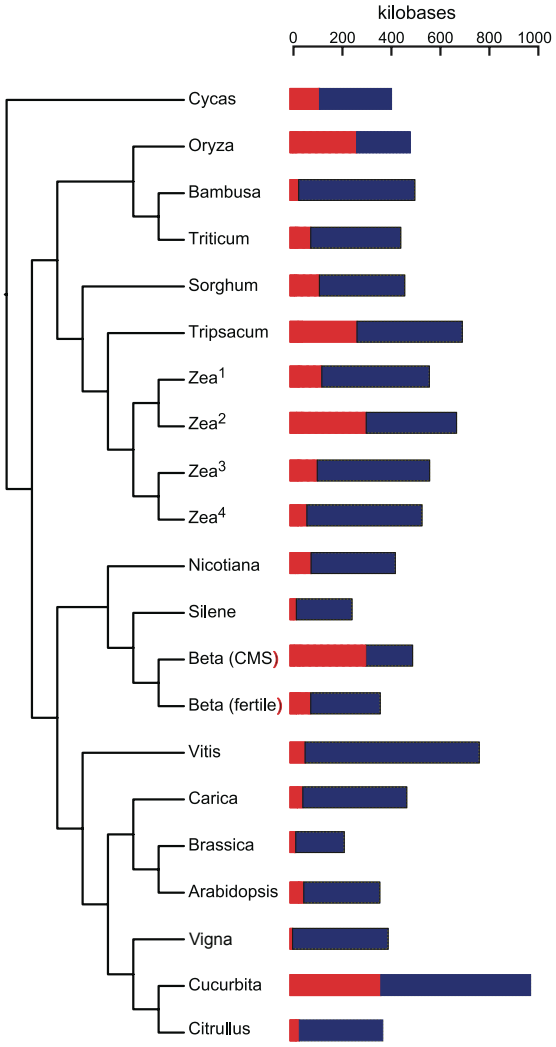


Figure I. 3. Variation of mitochondrial genome sizes across seed plant species. Indicating that in some plants, there is a high proportion of repetitive sequences (in red). In blue: non-repetitive sequences. From Alverson et al., 2011.

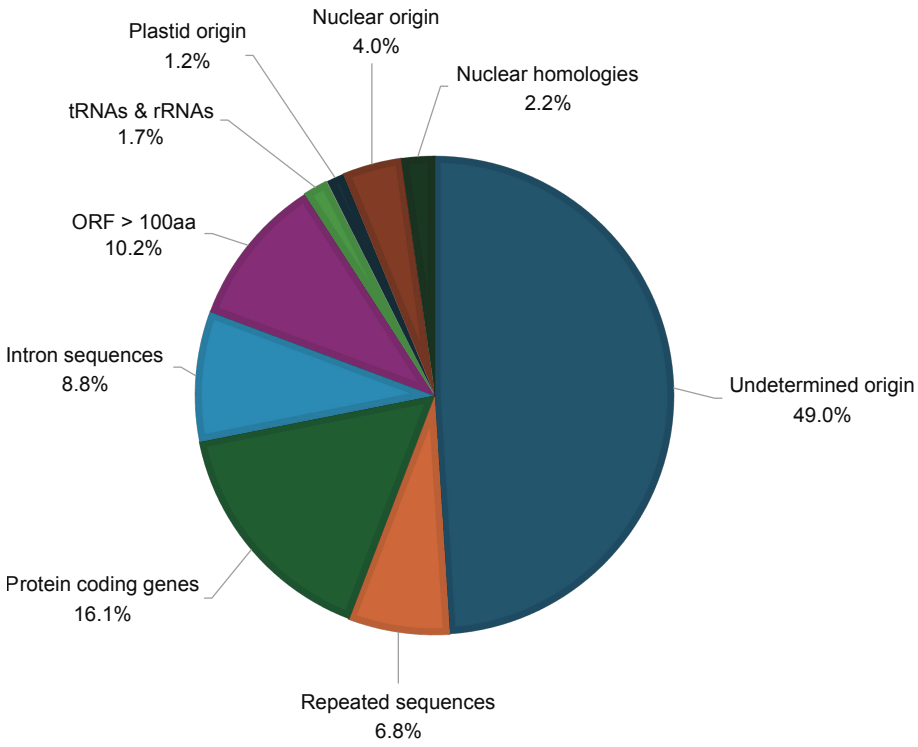


Figure I. 4. Arabidopsis mitochondrial genome composition. Mitochondria have acquired nuclear and plastid DNA via Intracellular gene transfer (IGT), while nuclear homology sequences are predicted to have been received by the nucleus from mitochondria. Adapted from Marienfeld et al., 1999.

Contrary to the animal kingdom, plant mitochondrial genomes are very large, with sizes that vary greatly among species (Fig. I.3): it can be as big as 11.7 Mbp (*Siberian larch*) and as small as 221 kb (*Brassica napus*) (Handa, 2003). However, recent observations of mitochondrial dynamics confirm that mitochondria often do not contain a full copy of the genome and can exchange genetic material between themselves (Giannakis *et al.,* 2022). Despite its large size, the mtDNA of land plants usually contains a small number of genes, between just 50 and 60 genes (Gualberto *et al.,* 2014). In this project, I will use *Arabidopsis thaliana* (ecotype Col-0). It is 367 kb in size and encodes for 32 proteins, 3 rRNA (5S,18S and 26S) and 22 tRNA (The Arabidopsis Genome, 2000). In fact, it encodes just 20 genes more than human mitochondria, while genome size is 23-fold larger. Furthermore, just a quarter of the mitochondrial genome is transcribed, and only 16.1 % encodes functional proteins (Fig. I.4) (Marienfeld *et al.,* 1999). Thus, in plants, most of the mtDNA is non-coding. The most extreme example is *Silena conica*, with an mtDNA of 11.3 Mb that is 99% noncoding and of largely unknown origin (Sloan *et al.,* 2012). Similarly, the sequence origin of more than 50 % of the Arabidopsis mitochondrial genome is undetermined (Marienfeld *et al.,* 1999). Many plants have mitochondrial genomes abundant with foreign DNA, which is suggested to have been acquired by gene transfer. It can be intracellular gene transfer (IGT) (endosymbiotic origin) or horizontal gene transfer (HGT), acquired from another organism via parasitic plant-plant, plant-pathogen interactions. However, much remains unknown about HGT between plants (Richardson and Palmer, 2006).

As described above, despite highly increased genome size in plants, they do not encode significantly more genes than animals do. Angiosperm mitogenomes most frequently encode proteins of the respiratory chain: complex I (*nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad7*, *nad9*), complex III (*cob*) complex IV (*cox1*, *cox2*, *cox3*) complex V (*atp1*, *atp4*, *atp6*, *atp8*, *atp9*), Cytochrome c maturation (ccmB, ccmC, ccmFc, ccmFn), a maturase (matR) and a gene of unknown function (mttB) (Adams and Palmer, 2003; Adams *et al.,* 2002; Knoop, 2004). The remaining mitochondrial protein genes code for some of the 17 ribosomal proteins most frequently found in the plant organelle (Fig. I.5). Small (18S) and large (26S) subunits of the ribosome are found in all characterized mitochondria. However, mitochondrial 5S rRNA has a very patchy phylogenetic distribution. It is only present in the mitochondria of land plants, a subset of green algae, red algae, and brown algae (Oudot-Le Secq *et al.,* 2001). As for transfer RNA, typically, flowering plants have 11 – 13 mt-tRNA of native plant mitochondria origin and a variable number of intracellularly and horizontally transferred tRNAs from diverse origins: plastids, bacteria, fungi, and the mitochondria of other plants (Warren and Sloan, 2020; Sinn and Barrett, 2020).

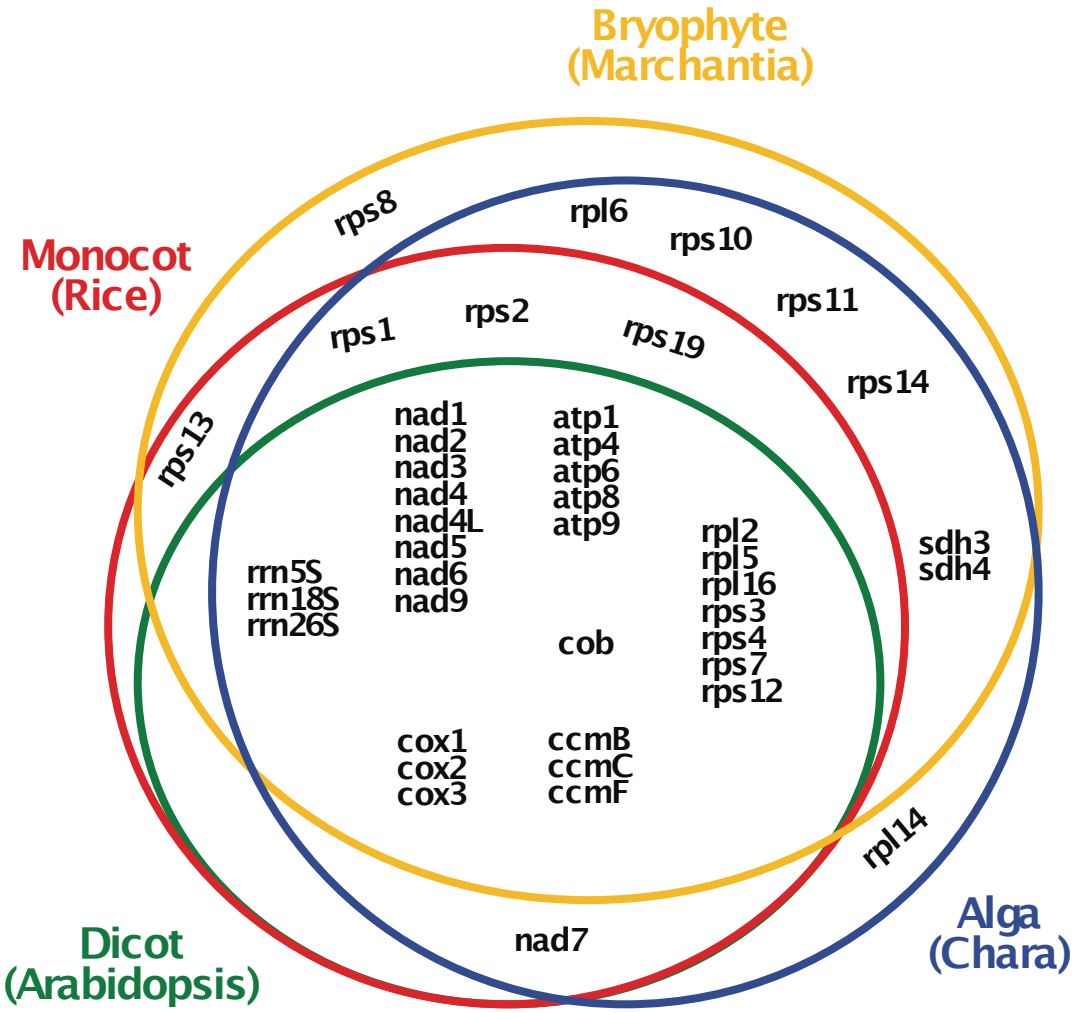


Figure I. 5. Conservation of mitochondrial genes across land plants.
Adapted from Knoop, 2004.

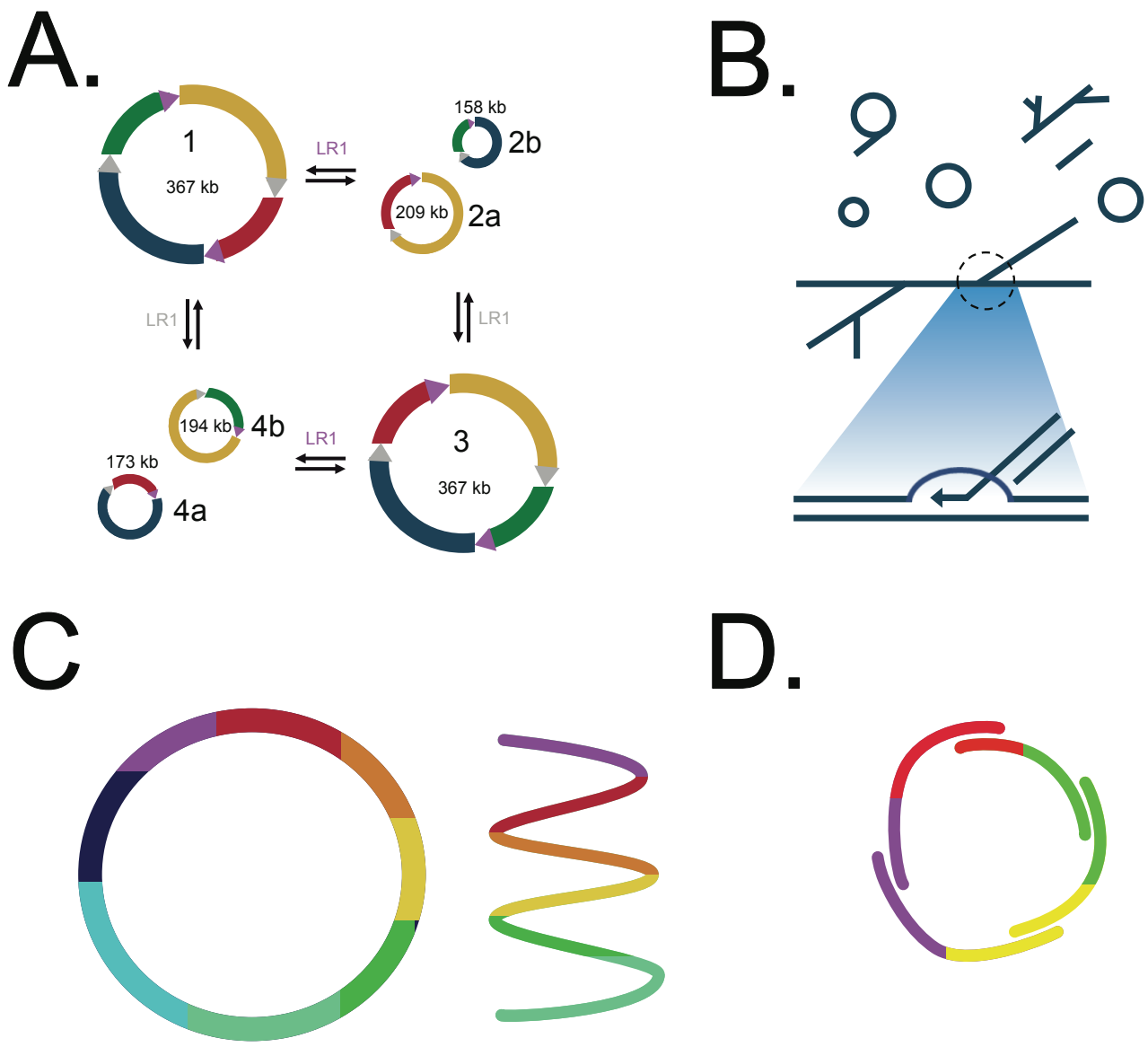


Figure I. 6. Plant mitochondrial DNA structure. A) Recombination of A. thaliana mitochondrial genome involving two sets of large repeats (LR1 and LR2). Recombination events lead to four possible mtDNA configurations: 2 alternative master circles that caries all the mtDNA sequences (1 and 3) and 2 possible configurations (2 and 4) with the mtDNA divided into sub-genomes (a and b). B) According to microscopy observation, mtDNA is found to be a mixture of branched, linear and circular molecules. Branched molecules could be intermediates of recombination. C & D) Based on bacterial studies, the mtDNA could be organised into head-to-tail concatemers (C) or circularly permuted linear molecules (D).

2.2. Structure of mitochondrial DNA

The important characteristic of plant mitogenomes is the abundance of repeated sequences (Fig. I.3) that can be involved in homologous recombination (HR), which highly impacts the structure of the genome (Alverson *et al.,* 2011). Repeats can be classified by size: large repeats (>500 bp), intermediate-size repeats (50–500 bp); and small repeats (<50 bp). In animals, it is generally accepted that HR is absent in mitochondria, as no clear cases have been observed (Rokas *et al.,* 2003). However, there is an exception: HR in mitochondria has been seen in invertebrates, such as mussels (Ladoukakis and Zouros, 2001). In yeast (Eckert-Boulet *et al.,* 2011) and bacteria (Smith, 1988), HR pathways are already well-defined. However, in plant mitochondria, the existing alternative HR pathways and the factors involved are still poorly characterized.

Early studies on plant mtDNA revealed that frequent and reciprocal recombination between large repeats is the major reason for its variable genome structure. This plasticity makes it hard to predict the actual structure of the mtDNA; thus, it is frequently represented as a single circular chromosome called a “master circle” (Lonsdale *et al.,* 1984; Palmer and Shields, 1984). The Arabidopsis mtDNA has 2 pairs of large repeats; thus, it can assume two alternative configurations of master circles and be subdivided into four sub-genomes, all these structures interconvertible by HR (Fig. I.6A). Similarly, a single large repeat in radish (*Raphanus sativus*) mtDNA can result in one master circle and two sub-genomes (Tanaka *et al.,* 2012), while three large repeats present in the mtDNA of a related ancient crop (*Eruca sativa*) theoretically result in six alternative master circles and four sub-genomes (Wang *et al.,* 2014). The existence of the plant mtDNA as a multipartite genome has been previously demonstrated. An independent small sub-genomic plasmid in the mitochondria of *Chenopodium album* has been shown to replicate independently and has no functional significance (Dorfel *et al.,* 1991). However, the sequence mutation rates of such plasmids are higher compared to the main genome (Smith *et al.,* 2014). Thus, it was assumed that periodic insertion of genes in plasmid might accelerate their evolution rates, as was shown in a maize linear mitochondrial plasmid (Warren *et al.,* 2016). Furthermore, long-read sequencing confirmed that the plant mtDNA can form big multipartite structures as well; two independent circles of *Tylosema esculentum* (Marama bean), which cannot recombine with each other, were assembled *de novo* from PacBio long reads (Li and Cullis, 2021).

The mammalian mtDNA is found as small single-circular whole genome molecules (Falkenberg and Gustafsson, 2020). In yeast, analysis of the mtDNA of multiple species revealed the presence of monogenomic linear and circular molecules, concatemeric molecules, and branched structures (Nosek *et al.,* 2006). However, gel-based approaches and electron microscopy have failed to detect genome-size circular molecules in plants. Instead, they have shown that most mtDNA is linear or forms large complex branched molecules and rosette-like structures, while most observed circular molecules were small (Fig. I.6B) (Backert *et al.,* 1997; Backert and Borner, 2000; Bendich, 1996). Recent analysis of lettuce mtDNA molecules using fluorescence microscopy has corroborated these observations (Kozik *et al.,* 2019). Thus, it is evident that the circular representation of the plant mtDNA might be convenient but artificial. Several new models have been proposed based on DNA structures of certain replicating bacteriophages, including head-to-tail concatemers (Fig. I.6C) and circularly permuted linear molecules (Fig. I.6D) (Bendich, 1993; Sloan, 2013). The first model suggests that linear mtDNA molecules can be connected end-to-end and mapped or assembled as circles, while the second model implies that linear mtDNA molecules with overlapping ends could be assembled into a circle. However, none of these models have been proven *in vivo*.

In plants, small microhomologies and intermediate-size repeats can infrequently be involved in ectopic recombination that is usually asymmetric (Woloszynska and Trojanowski, 2009; Gualberto *et al.,* 2014). These recombination events usually produce multiple configurations of the mitochondrial genome called mitotypes or sublimons. The state of the genome when multiple mitotypes co-exist is called a heteroplasmic state. Consequently, the stoichiometry of multiple mitotypes can constitute the uniform plant mtDNA during symmetric recombination, and not uniform during asymmetric recombination (Preuten *et al.,* 2010). It is not well known how a preferred orientation is selected during homologous recombination, but GC bias was observed for gene conversion. However, it was postulated that gene conversion is most powerful in the SAM and reproductive tissues during mitochondria fusion, as HR takes place only within the organelle (Broz *et al.,* 2022). Furthermore, it has been seen that heteroplasmy can be sorted out within a single generation, suggesting a tight and effective genetic bottleneck (a tight selection reducing variation). Such a tight bottleneck in plant mitochondria could be due to the physical reduction of mtDNA copies during female germline development (Zhang *et al.,* 2018). Furthermore, live imaging of germinating Arabidopsis

seedlings showed the dramatic reorganization of the chondriome, massive fusion of mitochondria and mixing of the mtDNA, followed by heterogeneous redistribution of nucleoids among the mitochondria. This showed how mixing among mitochondria is favoured while the integrity of the mtDNA remains unchanged (Paszkiewicz *et al.,* 2017).

Such asymmetric HR events and quick sorting of heteroplasmy leads to the fast evolution of the mtDNA structure. However, the plant mtDNA evolves slowly in sequence. The evolution of two genera, Brassica and Raphanus, revealed 3 to 14 events of inversion between any 2 species of 6 analysed (Palmer and Herbon, 1988). While the primary sequence differed by only 1-8 substitutions every 1000 bp, the mutation rate was 100 times slower compared to the mutation rate of the animal mtDNA. Coding sequences showed a smaller substitution rate compared to non-coding sequences, due to evolutionary pressure (Christensen, 2013). Such slow mutational rate and fast structural change of the plant mtDNA can be explained by the high-fidelity repair mechanism of HR, correcting SNP by gene conversion (Gualberto and Newton, 2017). Due to such a fast evolution of mtDNA structure, the synteny of genes might vary significantly between closely related species and even within the same species. Deletion, duplication, insertion, and inversion of sequences by homologous recombination have been seen in different accessions of *A. thaliana* (Arrieta-Montiel *et al.,* 2009; Davila *et al.,* 2011), beet(Darracq *et al.,* 2011) and maize (Allen *et al.,* 2007). However, such rearrangement could also lead to the creation of new chimeric genes. Several studies have shown that such new ORFs could lead to cytoplasmic male sterility (CMS) in crop species such as maize, rice, sunflower, beet, carrot, and cauliflower. This brings a strong genetic bottleneck for further evolution, yet brings an advantage for crop breeding programs (Chen and Liu, 2014).

2.3. Inheritance of mitochondrial genome

In most plants, mitochondria and its DNA, are maternally inherited, as in most eukaryotes (Birky Jr., 2001). However, as previously discussed, heteroplasmy is a common state of plant mitochondria, most frequently due to mutations or structural variations. Yet rarely, paternal mitochondria leakage could lead to an even more complex heteroplasmic population. In a study of natural populations of *Silene vulgaris*, paternal leakage contributed 12% to the plant's heteroplasmic state (Pearl *et al.,* 2009). However, in a heteroplasmic state, plants have a predominant mitotype, which determines the plant phenotype (Kmiec *et al.,* 2006). Furthermore, the segregation of mitotypes to the next generation is not uniform, and it has been shown that the predominant mitotype can change rapidly in just a generation, by a clonal amplification process that was called substoichiometric shifting (SSS) (Janska *et al.,* 1998; Zaegel *et al.,* 2006). Additionally, the process can be reversed, as was demonstrated in a CMS line of common beans, where fertility could be restored at low frequencies by a genomic

shifting of the CMS-associated gene *pvs-orf239* to substoichiometric levels, by SSS (Mackenzie *et al.,* 1988). A putative progenitor sequence of *pvs-orf239* was identified at substoichiometric levels in fertile plants, suggesting that the CMS gene was initially introduced by SSS (Arrieta-Montiel *et al.,* 2001). The process of SSS is not fully understood. The predominant mitotype could be selected by a genetic bottleneck linked to the preferential amplification of mitotypes and increased homologous recombination (Woloszynska, 2010). However, it never leads to full homoplasmy. Furthermore, it has been shown that the restoration of MSH1, a protein involved in mismatch repair and homologous recombination, increased the rate of heteroplasmic sorting in mitochondria (Broz *et al.,* 2022).

2.4. Mitochondria packing to nucleoids

For long the mtDNA was thought to exist naked within the mitochondrial matrix. Only during the last decades, in humans, was it shown to form compact structures called nucleoids, analogous to bacterial chromosomes (Satoh and Kuroiwa, 1991). Through super-resolution imaging it was revealed that mtDNA is coated with proteins and organized into irregular and elongated aggregates with a consistent mean size of 100 nm, across all mammals (Kukat *et al.,* 2011; Kukat *et al.,* 2015). These images also estimated an average of 1.4 mtDNA molecules per nucleoid, in human fibroblast lines. For proper mtDNA packing and maintenance, over 50 nucleoid-associated proteins interact either permanently or transiently with the mtDNA or other nucleoid-associated proteins (Lee and Han, 2017). All factors can be split into two groups: core and peripheral. In human mitochondria, the key core proteins are: mitochondrial transcription factor A (TFAM), mitochondrial single-stranded DNA-binding proteins (mtSSBs), ATP-dependent Lon protease (LONP1), Twinkle helicase, polymerase γ (POLG/POLG2). TFAM is a main nucleoid core protein with multiple roles. First, it is involved in tight compacting of the mtDNA, with a footprint of 10 bp to 30 bp (Gilkerson *et al.,* 2013). The super-resolution images of human fibroblast lines assume TFAM binds a single mtDNA molecule (TFAM:mtDNA is 1:1) in nucleoid composition (Kukat *et al.,* 2011; Kukat *et al.,* 2015). However, the ratio of TFAM:mtDNA in human mitochondria is variable depending on physiological and pathological conditions, as well as the mechanism of selection of mtDNA for replication was proposed, keeping part of mtDNA silent (Alán *et al.,* 2016; Farge *et al.,* 2014). Secondly, as the name implies, TFAM is also a transcription factor that induces a DNA U-turn and promotes transcription initiation (Hallberg and Larsson, 2011). Finally, TFAM is equally important for the initiation of mtDNA replication, and overall control of mtDNA copy number via mtDNA stabilisation and protection from turnover (C., T., Campbell *et al.,* 2012). LONP1 is involved in the degradation of TFAM and brings another layer of control of mtDNA content

Category	*A.thaliana*	AGI	Copies per mtDNA	Human homologue
DNA Exonuclease	OEX1	At3g52050	-	FEN1*
DNA Polymerase	Pol1A	At1h50840	5	POLγ*
	Pol1B	At3g20540	2	
DNA Topoisomerse I	TOPI	At4g31210	61	-
DNA Topoisomerse II	GYRA	At3g10690	981 (491)	-
	GYRB	At5g04130	676 (338)	-
DNA Helicase	TWINKLE	At1g30680	192 (32)	TWINKLE
ssDNA-binding proteins	SSB1	At4g11060	493 (123)	mtSSB
	SSB2	At3g18580	657 (164)	
	OSB1	At1g47720	4	
	OSB3	At5g44785	171	
	OSB4	At1g31010	180	
	WHY2	At1g71260	974 (41)	
DNA-binding proteins	ODB1	At1g71310	673 (61)	RAD52?
	SWIB5	At1g31760	141	
DNA Recombinases	RECA2	At2g19490	462	RAD51*
	RECA3	At3g10140	-	
MutS-like	MSH1	At3g24320	2 (1)	-
DNA Ligase	LIG1	At1g08130	42	LIG*
RNA Polymerase	RPOTm	At1g68990	5	POLRMT
	RPOTmp	At5g15700	31	
HMG-box protein	-	-	-	TFAM
ATPase family AAA domain-containing protein	SBA1	At3g03060	-	ATAD3
	SBA2	At5g16930	-	
	SBA3	At2g18330	-	
	SBA4	At4g36580	-	
Mitochondrial transcription termination factor	SHOT1	At3g60400	-	mTERF18

Table I. 1. Plant mitochondria nucleoid factors. Data is adapted from Arabidopsis mitochondrial proteomic analysis (Fuchs et al., 2020). For proteins known or assumed to form complexes, their number is shown in brackets. * not homologs, but perform the same function.

(Iborra *et al.,* 2004). The outer layer is composed of transiently recruited proteins to perform RNA processing and translation (Ambro et al., 2012). There are additional factors, such as prohibitin, that stabilise TFAM (Gilkerson et al., 2013), or ADAT3, that play an important role in the association of nucleoids to ribosomes and to the inner mitochondrial membrane (Gilkerson *et al.,* 2013; Bogenhagen *et al.,* 2008).

The full composition of plant mitochondrial nucleoids has not been determined, yet an extensive set of factors involved in mitochondrial mtDNA transcription, replication and maintenance has been identified (Table. I.1) (Chevigny et al., 2020; Fuchs et al., 2020). In early experiments, it was already seen that the nucleoids of organelles in plants are like those in animals: a complex mix of DNA and proteins. Mitochondrial nucleoids are initially bulky but progressively get smaller during plant differentiation and elongation (Kuroiwa *et al.,* 1992). Nucleoids isolated from tobacco Bright Yellow-2 (BY-2) cells resist RNase A activity, yet after treatment with DNase I, proteinase K, and high levels of salt, RNA becomes sensitive (Nemoto *et al.,* 1988). Several factors have been shown to directly disrupt nucleoids. An Arabidopsis mutant of single-strand DNA binding protein WHIRLY2 (WHY2) was shown to break mitochondrial nucleoids and affect seed germination (Golin *et al.,* 2020). Furthermore, homologs of human mitochondrial transcription termination factor (mTERF) and ATAD3 have been found in Arabidopsis: SUPPRESSOR OF HOT1-4 1 (SHOT1) and SHOT1 BINDING ATPASES (SBAs), respectively (Kim *et al.,* 2021). In humans, mTERF arrests mitochondrial RNA polymerase (mtRNAP) progression by binding downstream of the 3′ end of the 16S rRNA gene (Kruse *et al.,* 1989). In Arabidopsis, the *shot1* and *sba* mutations lead to disrupted nucleoids, a decrease in complex I accumulation, and heat tolerance (Kim *et al.,* 2021). No other factors have been linked to changes in nucleoid formation, so far. However, many other factors involved in mtDNA binding, transcription and maintenance have been identified. Arabidopsis SSB1 and SSB2 orthologs of bacterial SSB and of human mtSSB have been identified (Elo *et al.,* 2003; Edmondson *et al.,* 2005), together with other plant-specific ssDNA binding protein families: OSB (Zaegel *et al.,* 2006), WHIRLY (Cappadocia *et al.,* 2010) and ODB (Janicka *et al.,* 2012). Functional homologs of bacterial and yeast proteins involved in DNA repair by HR have also been identified in Arabidopsis mitochondria, such as RecA-like recombinase RECA2 and RECA3 (Miller-Messmer *et al.,* 2012) and RADA ortholog of bacterial RadA (Chevigny *et al.,* 2022). Two functional homologs of human POLγ have been found in Arabidopsis - POL1A and POL1B, both functionally related to bacterial polymerase I (Parent *et al.,* 2011). A MutS-like homolog, MSH1 (Xu *et al.,* 2011), involved in mismatch repair,

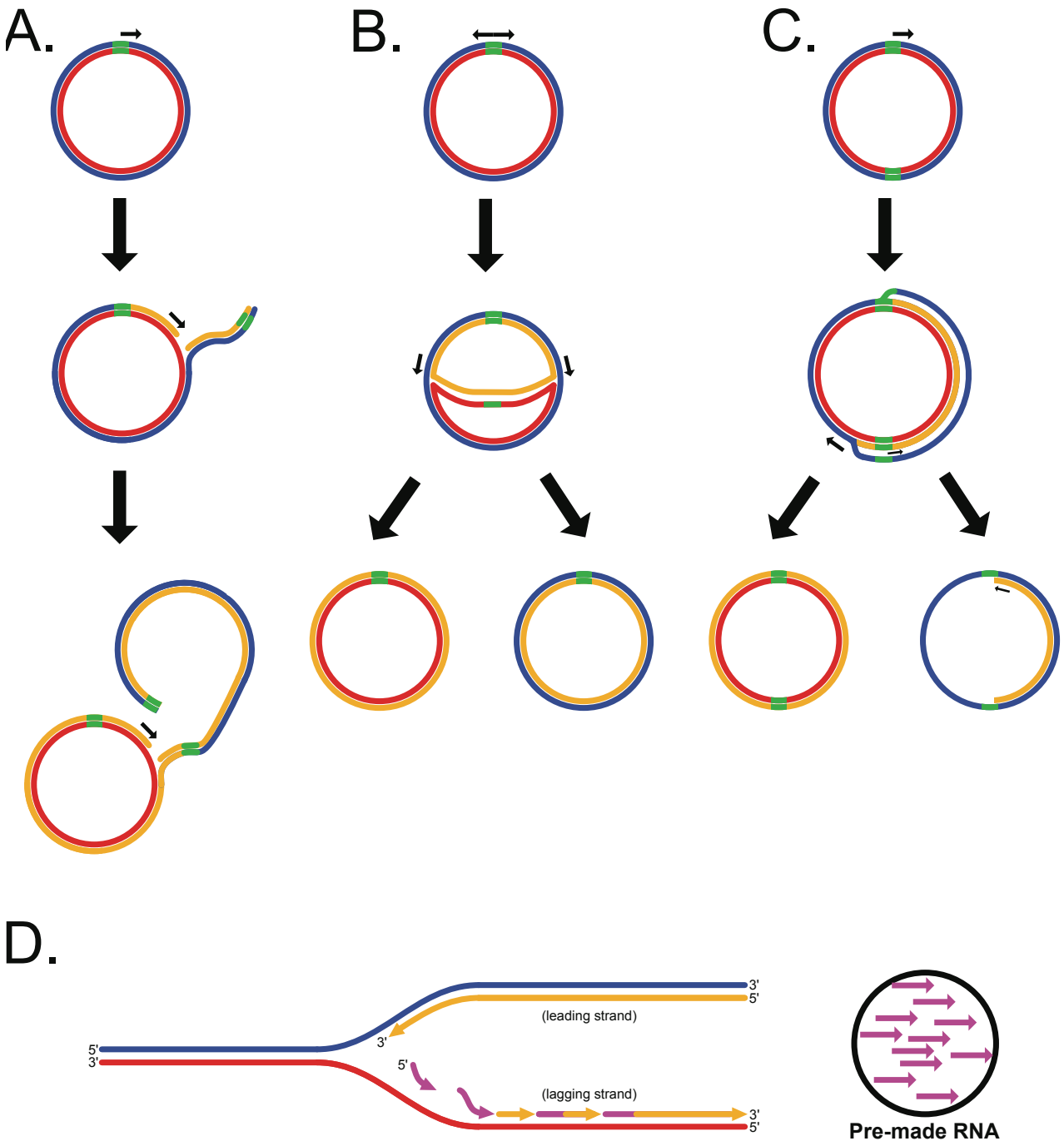


Figure I. 7. Models of animal mtDNA replication. Blue and Red represent the parental strands, and the newly synthesised strand is in yellow. The replication initiation site is shown in green. A) Rolling circle replication. The replication is initiated by a nick in one of the DNA strands, and synthesis progresses along the circular molecule, displacing the nicked strand. Upon replication reaches the initial site, the displaced strand is nicked and ligated to form a new single-stranded circular molecule. However, synthesis may continue, creating a linear concatemeric molecule, which is later converted into multiple single-stranded circular copies of the parent molecule. In both cases, DNA polymerase later synthesises the second DNA strand, forming a double-stranded DNA. B) Theta replication model. Bi-direction replication is initiated at the origin of replication. When two replication forks meet each other, two double-stranded circular molecules are separated. C) Strand displacement replication model. Replication of DNA is initiated by the displacement of one of the DNA strands by an RNA primer. Upon 2/3 of the genome is synthesised, the DNA synthesis in the opposite direction is triggered at the exposed second initiation site. When the replication fork reaches the initial start site, the parental strand is displaced as a single-stranded circular DNA molecule and converted into a double-stranded molecule by DNA polymerase. D) RITOLS/bootlace replication model. This model has a leading strand that replicates normally and a lagging strand whose replication involves hybridization ('threading') of free pre-synthesized RNA molecules onto the template DNA, starting from the 3' end. Once the gaps are filled, the RNA primers are removed, likely by RNase H, 5'-3' exonuclease, or Flap endonuclease. Adapted from Morley et al., 2019.

is also probably a part of the mitochondrial nucleoid. Finally, plant mitochondrial nucleoids most likely also include the replicative DNA helicase TWINKLE (human ortholog) (Diray-Arce *et al.,* 2013), topoisomerases type II gyrase and type I TOP1, to relieve DNA supercoiling (Cho et al., 2004; Wall et al., 2004), and phage-type DNA-directed RNA polymerase RPOTm and RPOTmp for mtDNA transcription (Hedtke *et al.,* 1997; Yadav *et al.,* 2019). The relative abundance of some proteins mentioned above has been identified by mass spectrometry (Fuchs *et al.,* 2020). It was found that ssDNA binding factors, together with gyrases, are the most abundant proteins in mitochondria. However, a homolog of human TFAM, involved in DNA packing, was not found in plants (Gilkerson *et al.,* 2013).

2.5. Replication of mitochondrial genomes

Even though it was shown that animal mtDNA is simpler in structure than mitogenomes of yeast and plants, and much has been revealed about its nucleoid composition, much about its replication still remains a mystery. Several models (Fig. I.7) have been proposed for mtDNA replication in animals. The rolling circle model (Fig. I.7A) was proposed for mtDNA replication of *Caenorhabditis elegans* (Lewis *et al.,* 2015), similar to the DNA replication in certain bacteriophages (Bernstein and Bernstein, 1973) and some eukaryotic viruses (Rampersad and Tennant, 2018). The replication of mitochondria in *Drosophila melanogaster* (Jõers and Jacobs, 2013) was assumed to undergo theta replication (Fig. I.7B), which is shared by all invertebrates and also a common type of replication in bacteria (Lilly and Camps, 2015). A strand-displacement model (Fig. I.7C) (SDM) was proposed for the replication of mouse mtDNA (Robberson *et al.,* 1972) and was widely accepted to be the sole mechanism for all mammalian mitochondria till 2000 (Holt *et al.,* 2000). The discovery of replication intermediates in humans and mice contradicted SDM and assumed a more complex mechanism. It was determined that the replication of vertebrate mitochondrial DNA has transient ribonucleotide incorporation throughout the lagging strand (Yasukawa *et al.,* 2006), which was not present in previous models, and thus, the bootlace/RITOLS model (Fig. I.7D) was proposed (Reyes *et al.,* 2013).

As described previously, the structure of plant mitochondrial genomes is much more complex than in animals. Therefore, it might require a different model of replication. It was assumed that the plant mitochondrial genome undergoes recombination-dependant replication (RDR), similar to that of bacteriophage T4 (Oldenburg and Bendich, 1996; Backert and Borner, 2000). Furthermore, as previously described, the plant mtDNA can be found in linear molecules, head-to-tail concatemers, branched, and rosette-like structures, suggesting that free DNA ends prime RDR (Oldenburg and Bendich, 1996). Break-induced replication

(BIR), or break-induced RDR, has been seen in many eukaryotes, including human nuclear DNA (Sakofsky and Malkova, 2017). The heteroplasmy previously discussed in section 2.2 may have resulted from the replication of new mitotypes initiated by BIR (Jackman *et al.,* 2020). However, considering the variable mtDNA structure and the many nucleoid replication proteins with similarities to bacteriophages (e.g., Twinkle) or of clear bacterial origin (e.g., DNA gyrase), the plant mtDNA replication system might be more mosaic and redundant. Rolling cycle replication has been observed in *Chenopodium album*, suggesting it could be present in other plant species as well (Backert *et al.,* 1996; Backert and Borner, 2000). Finally, mimicking the DNA replisome of T7 bacteriophage (Lee and Richardson, 2011), a coordinated leading and lagging-strand DNA synthesis model of mtDNA could be proposed in plants, directed by the plant Twinkle-POL1A/POL1B enzyme replisome (Brieba, 2019). A similar system was found in animals composed of DNA Polγ, Twinkle, and SSB1 protein (Korhonen *et al.,* 2004).

2.6. Cytoplasmic male sterility

Sterile male plants are plants that are incapable of producing pollen, while pistils are functional. It was first documented in 1763 by Joseph Gottlieb Kölreuter (Mayr, 1986). When such a phenotype is caused by mutations or rearrangements in the mitochondrial DNA (mtDNA), it is called "cytoplasmic male sterility" (CMS). Today, over 610 plant species have been reported with CMS, caused by mitochondrial genes, and genic male sterility (GMS) caused by nuclear genes alone (Kaul, 1988; Vedel *et al.,* 1994). Furthermore, external conditions can also affect pollen production, as shown in rice and wheat, via what is called environmental genic male sterility (EGMS) (Meng *et al.,* 2016; Zhou *et al.,* 2012). In fact, the formation of plant microspores in anthers requires cooperation between gametophytic cells (microspores) and sporophytic cells (anther wall), particularly the tapetum that directly supplies developing male gametophytes with nutrients (Levings, 1993; Guo and Liu, 2012). Thus, sterility can be divided into two types: gametophytic, which involve defective microgametogenesis of haploid pollen, and sporophytic, which affects diploid maternal tissue like the tapetum during sporogenesis, leading to pollen abortion and sterility.

The CMS gene is usually a mitochondrial chimeric gene produced via recombination, resulting in a gain of function (Xu *et al.,* 2022; Yamagishi and Bhat, 2014). It is maternally inherited as it is part of the mtDNA. However, CMS can be rescued by nuclear genes, which are defined as restorers of fertility (Rf), that suppress or downregulate the CMS genes (Jindal *et al.,* 2020; Schnable, 1998). Several Rf genes have been identified to code for PPR proteins that affect the processing or stability of the CMS transcripts (Gao *et al.,* 2022). More than 10 genes coding for ETC subunits have been found to be involved in the formation of CMS genes

(Chen and Liu, 2014). Most frequently, they originate from *cox1*, *atp8*, and *atp6*. Furthermore, most often the CMS genes encode a transmembrane proteins: *orf79* (fusion of N-terminal *cox1* and SUO, sequence of unknown origin) in rice (Z., Wang *et al.,* 2006), *orf138*, *orf125*, *orf224*, *orf222* (originated from *atp8*) in *Brassica* (Singh and Brown, 1991; Uyttewaal *et al.,* 2009), *atp6* derivatives in maize (Dewey *et al.,* 1991), pepper (Kim *et al.,* 2007) and sugar beet (Yamamoto *et al.,* 2005). More about CMS linked to ATP synthase is developed in Chapter 4.

Four different models have been proposed on how CMS genes induce male sterility (Chen and Liu, 2014). (i) Cytotoxicity model: CMS-encoded genes directly lead to the death of sporophytic or gametophytic cells in the anthers. The model is based on mitochondrial dysfunction. However, no proof of cytotoxicity in CMS lines has been shown previously. (ii) Energy deficiency model: cellular respiration is altered. CMS proteins can replace respective parts of ETC complexes, making them defective, resulting in a reduction of proton gradient and ATP production. ATP deficit has been recorded in growing anthers of CMS rice (Wang *et al.,* 2013). (iii) Asynchronous programmed cell death: tapetum cell death during seed maturation. Loss of tapetum impedes the nourishment of pollen and halts its development. Programmed cell death is assumed to be activated by increased production of reactive oxygen species (ROS) due to the release of cytochrome C from ETC. (iv) Retrograde regulation model: the CMS gene regulates the expression of nuclear genes. It has been shown how the CMS gene in carrots can reduce the expression of MADS-box proteins involved in the specification of flower organs, resulting in their deformation (Linke *et al.,* 2003).

3. DNA damage

3.1. Sources and types of DNA damage

3.1.1. UV-induced damage

Sun is needed for plants to perform photosynthesis and for humans for vitamin D production, which is required for proper calcium and phosphate assimilation (Adams and Hewison, 2010). Although solar exposure is necessary for both, UV radiation can induce DNA damage. Prolonged human exposure to the sun can lead to premature ageing (photoaging), skin cancer, and blinding eye diseases (Fisher *et al.,* 2002). UV can induce the photochemical generation of reactive oxygen species (ROS) causing direct deleterious chemical modifications to cellular components (DNA, proteins, and lipids), leading to damage to skin connective tissue. UV radiation is classified as a “complete carcinogen” as it is a mutagen that can initiate and promote tumour formation. Shorter-wavelength UV photons, particularly UV-B and UV-C, most commonly cleave internal 5–6 double bonds of pyrimidines, forming bulky

photolesions of pyrimidine, thymine dimers or pyrimidine–pyrimidone photoproducts (Sarasin, 1999; Sklar *et al.*, 2012; Britt *et al.*, 1993). The last can further photoisomerize into a third type of photoproduct called Dewar valence isomer (Batista *et al.*, 2009). Less frequently, UV can also lead to single-strand gaps (SSG) (Lankinen *et al.*, 1996) or, in rare cases, double-strand breaks (DSB) via the introduction of two single-strand breaks or UV-induced replication arrest (Limoli *et al.*, 2002). For plants, UV light is intrinsic to their survival by photosynthesis, and thus, they developed phytochrome and cryptochrome photoreceptors. To enhance UV tolerance, in response, plants can biosynthesize sunscreen flavonoids (Jenkins, 2017; Stracke *et al.*, 2010). However, as in humans, the production of photoproducts by UV damage has been detected in the plant nucleus, mitochondria and chloroplast (Molinier *et al.*, 2008; Takahashi *et al.*, 2011a). It is the most common exogenous factor affecting plant mutation rates (Quiroz *et al.*, 2023).

3.1.2. DNA damage from ROS and RNS

Reactive Oxygen Species (ROS) or Reactive Nitrogen Species (RNS) are by-products of multiple physiological activities in diverse subcellular sites. Most are produced due to electron leak from ETC: 2% of consumed oxygen in mitochondria is turned into ROS (Turrens, 2003). Their production might be increased by exogenous sources like UV light, ozone, ionizing radiations, metals, pesticides, air pollutants, or pharmaceutical drugs (Limón-Pacheco and Gonsebatt, 2009). The most abundant ROS found in living cells are superoxide ($O_2^{\bullet -}$), hydroxyl radicals (HO•) and peroxide (H_2O_2). However, there is a biological paradox, as ROS species are mitigated and controlled by various antioxidants and act in ROS-signaling pathways required for the preservation of cellular homeostasis (D'Autréaux and Toledano, 2007; Bardaweel *et al.*, 2018). Overproduction of ROS leads to deleterious effects such as DNA damage.

In mitochondria, RNS can be created from L-arginine, nitroprussides, nitrosothiols or any other organic nitrates (Steffen *et al.*, 2001; Chen *et al.*, 2002). The superoxide anion ($O_2^{\bullet -}$) can react with nitric oxide (NO•) and produce peroxynitrite ($ONOO^-$). Furthermore, the oxidation of nitric oxide forms the powerful nitrosating agent nitrous anhydride N_2O_3. It can deaminate primary amines on DNA bases, converting adenosine, cytosine, 5-methyl-cytosine and guanine to hypoxanthine, uracil, thymine and xanthine, respectively (Burney *et al.*, 1999). In general, RNS was shown to inhibit respiration, modulate ATP-dependent K^+ channels and regulate the maturation of the organelle (Boveris *et al.*, 2000; Almeida *et al.*, 1999).

Release of both ROS and RNS can oxidize nucleotide bases, most frequently forming 8-oxoguanine lesions, apurinic/apyrimidinic sites (AP) or promoting mispairing outside of normal Watson-Crick (Schulz *et al.,* 2000). The most characteristic transversion is guanine (G) to thymine (T). ROS and RNS also can introduce single-strand gaps (SSG), and more closely based SSG, AP sites can form DSB (Hegde *et al.,* 2012; Lindahl, 1993).

3.1.3. Alkylating agents

The alkylation of DNA and formation of bulky adducts can be promoted by endogenous metabolic products as well as chemical agents, such as methyl methanesulfonate (MMS) and Ethyl methanesulfonate (EMS). MMS and EMS are widely used mutagenic chemical agents in plants, causing DNA methylation and ethylation, respectively, and both can lead to the formation of intrastrand and interstrand cross-links or bulky adducts (De Silva *et al.,* 2000; Todd *et al.,* 1981; Iwamoto *et al.,* 2004). Previous studies have shown that during the metabolism of organic compounds, reactive compounds are formed that can react with DNA, creating hypermodified nucleotides called “adducts”. S-adenosylmethionine produced from amino acid biosynthesis was shown to nonenzymatically methylate the DNA (Rydberg and Lindahl, 1982). Reactive methylating species were also shown to rise from lipid peroxidation reactions (Vaca *et al.,* 1988). The 7MeG methylation has been recorded in the genomic DNA of many various organisms (Xiao and Samson, 1993; Shuker and Farmer, 1992; Tan *et al.,* 1990). 7MeG can be depurated to the AP (apurinic) site by N-methylpurine-DNA glycosylase (MPG), which in turn leads to DNA breaks (Barbara, 2003). However, in general, the accumulation of adducts in the mtDNA is higher than in the nuclear genome (Thomas and Roy, 2001; Jung *et al.,* 2009; Balansky *et al.,* 1996). Furthermore, alkylating agents like cyclophosphamides are used in chemotherapies (Sabel, 2009). It has been shown that a single alkyl group causes SSG, while two alkyl groups can lead to DNA cross-links and DSB.

3.1.4. Damage from Ionizing radiation

Ionizing radiation (IR) is another agent with mutagenic potential. In hospitals, it is used for diagnostic purposes (in X-rays and computed tomography scans) and for the treatment of cancer (Martinez-López and Hande, 2020). IR is often used as a mutagenic agent for genetic analyses in crops. For such purposes, commonly used agents are γ-rays and carbon ion beams (Yoshihara et al., 2013), X-rays (Kus et al., 2018), and fast neutrons (Li *et al.,* 2017). Resulting mutants have single-base substitutions, deletions, insertions, inversions, translocations, and tandem duplications. IR directly generates DSBs and leads to DNA fragmentation, indirectly damaging DNA via ROS species generated by water hydrolysis (Nagata *et al.,* 1999; Kim *et al.,* 2019). Studies showed that relaxed chromatin, during DNA

replication, is more sensitive to IR, and DSB distribution was found to be influenced by packaging (Lavelle and Foray, 2014). Similarly, radiometric agents like bleomycin have been shown to mimic radiation and induce DSBs, SSGs, and oxidative damage (Furukawa *et al.,* 2015).

3.1.5. Replication errors

In addition to all the factors mentioned above that cause DNA damage, one endogenous factor arises directly from DNA replication, which is always present. During DNA replication between two meiotic divisions, the Spo11 complex introduces DSB for recombination between chromosomes (Neale *et al.,* 2005). Furthermore, the precision of DNA replication and editing enzymes is not ideal, resulting in 10^{-8} errors per base or 30 de-novo mutations per human genome (Campbell et al., 2012) The incorporation of the wrong base could be just a matter of polymerase fidelity, as there should be a balance between nucleotide selectivity, concentration, and proofreading. Thus, the fidelity of various genome regions might differ, and slight changes in balance may result in errors (Ganai and Johansson, 2016). In humans, mismatches are identified by MutS homologues (MSHs) after replication, after which MutLα (MLH1 and PMS1 in plants) induce SSGs (Josephs and Marszalek, 2018; Ortega *et al.,* 2021). If not repaired, it has been shown that SSGs during replication cause replication fork collapse and the introduction of DSBs (Kuzminov, 2001). To conclude, inefficient polymerase activity can result in mismatches, SSGs and DSBs. Furthermore, it has been shown that SSG can rise from the erroneous activity of DNA topoisomerase 1 (TOP1) during replication (Wang, 2002).

3.2. Mitochondrial diseases

My thesis evolves around mtDNA repair and segregation. Thus, I will review the consequences of DNA damage in mitochondria. In plants, it has been shown that the mutation rate is slow in most species and that structural rearrangement may lead to CMS, discussed in section 2.6, or to spontaneous growth-defect mutants, like maize non-chromosomal stripe (NCS) mutants. As CMS, NCS plants have a pleiotropic type of phenotype due to heteroplasmy, because of deletions in mitochondrial genes like *cox2* (Lauer *et al.,* 1990), *nad4*, *nad7* (Marienfeld and Newton, 1994) or *rps3/rpl16* (Newton *et al.,* 1996).

Although, as discussed in sections 2.2 and 4.4.2, there is no HR in animal mitochondria and sequence mutations are 100 times more frequent. The discovery of mitochondrial diseases dates to almost 40 years ago, with the first identification of mutations in mtDNA genes leading to Leber Hereditary Optic Neuropathy (Wallace *et al.,* 1988) and Kearns-Sayre syndrome (Zeviani *et al.,* 1988). According to a review paper on mitochondrial diseases from

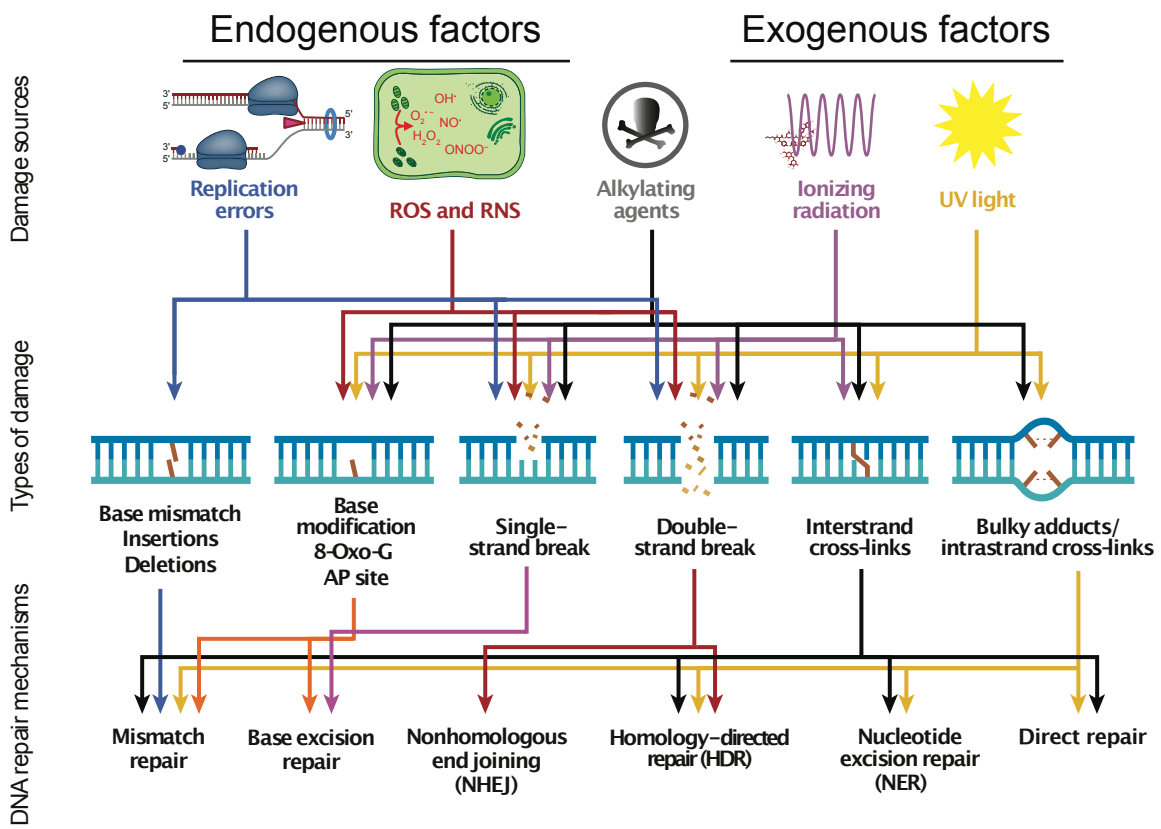


Figure I. 8. Endogenous and exogenous sources of DNA damage causing different types of DNA modification and pathways responsible for their repair. Adapted from Quiroz et al., 2023.

2020 (Stenton and Prokisch, 2020), disease genes (338) can be subdivided according to the functional roles: (i) respiratory chain proteins (102 genes); (ii) maintenance of mitochondrial genes and their expression (102 genes); (iii) mitochondrial dynamics and homeostasis (43 genes); (iv) metabolism of substrates (40 genes); (v) metabolism of cofactors (41 genes); (vi) metabolism of toxic compound (10 genes). As male sterility genes in plants, human mitochondrial disease genes can be further divided into two groups: those encoded in mitochondria and those encoded in the nucleus. However, as previously mentioned, human mitochondria are relatively small and encode just 37 genes. Mutations in 36 of them have been linked to diseases.

As discussed in section 2.2, humans naturally just have a single variant of the mtDNA; however, due to mutations, the mtDNA may become heteroplasmic. Thus, it has been shown that there is a biochemical threshold level at which a mutation can lead to clinical symptoms (Filograna *et al.,* 2021). It could be perceived as a natural defence barrier, giving the organism time to repair or segregate out the mutation. Otherwise, it may be transmitted to the next generation at different heteroplasmy levels (Wei and Chinnery, 2020), which might be lower or higher than the parental one. Furthermore, both *de novo* and inherited mtDNA mutations may have different threshold levels, determining the disease onset and the severity of the symptoms. Mutation A3243G in the human mtDNA affects tRNA leucine. Patients with 10 – 30% of their mtDNA with such mutation were shown to develop diabetes and potentially autism (van den Ouweland *et al.,* 1992; Pons *et al.,* 2004), while patients with the same mtDNA mutation at 50 – 90% heteroplasmy level, develop mitochondrial encephalomyopathy, lactic acidosis and stroke-like episodes (Picard *et al.,* 2014). The heteroplasmy of mutant mtDNA could also differ in various parts of the human body, leading to different detrimental health consequences (Wang *et al.,* 2023). Typically, organs with higher energy demand, such as the central nervous system, skeletal muscles, kidneys, and liver, are more affected (Huisman *et al.,* 2022). All this highly resembles why some plant mtDNA mutants may have just CMS phenotype, while others have stronger NCS-like phenotypes.

4. DNA repair

4.1. DNA repair pathways in the nucleus and organelles

Section 3.1 discusses different types of DNA damage, each of which may require a different repair pathway (Fig. I.8) to avoid the detrimental health consequences discussed in section 3.2. In this section, I will first review the different DNA repair mechanisms present in different organellar compartments and compare them with those found in plants. Secondly, I will review the mechanism of repair of each repair pathway.

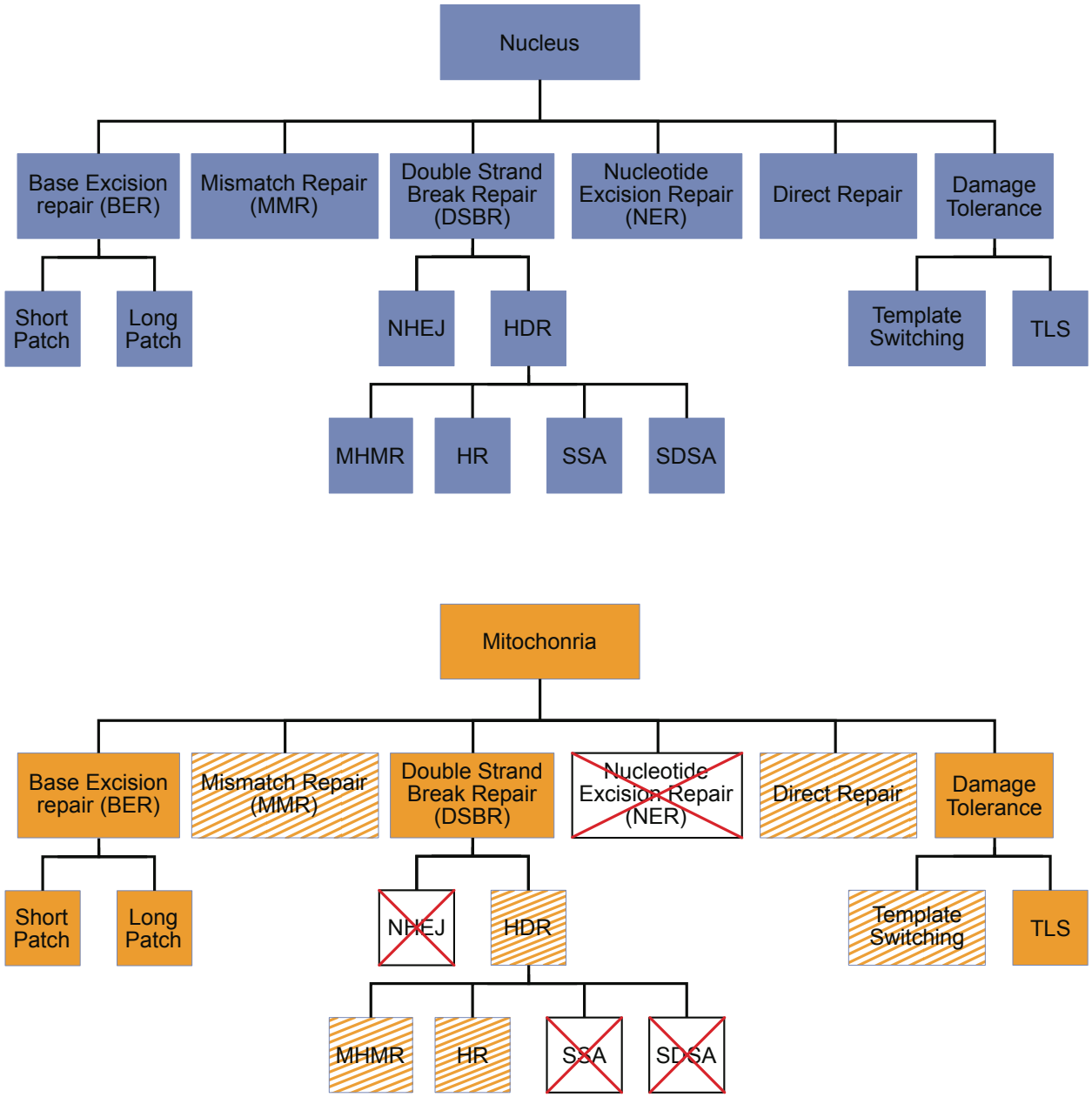


Figure I. 9. DNA repair pathways in the nucleus and mitochondria of human cells. Established pathways are shown in full colour; pathways that need further confirmation are shown with a hatched background; pathways that have not been identified or are absent are shown in a white background. Abbreviations: HDR, homology-directed repair; NHEJ, non-homologous end joining; MHMR, microhomology-mediated repair; HR, homologous recombination; SSA, single strand annealing; SDSA, synthesis-dependent strand annealing; TLS, translesion synthesis. Adapted from Boesch et al., 2011.

There are several repair pathways for each DNA damage mentioned in section 3.1 (Fig. I.8). DNA intrastrand cross-links and bulky adducts are most frequently repaired by direct repair (DR) or nucleotide excision repair (NER), and less frequently by mismatch repair (MMR) and homologous recombination (HR), if they turn to DSBs due to a stalled or collapsed replication fork (Enoiu *et al.,* 2012). However, there is evidence that lesions caused by alkylating agents can be recognised by MMR but not repaired. MMR replaces the base opposite the cisplatin adduct, yet keeps the source of mismatch intact, which re-start this process and leads to a so-called futile cycle (Rocha *et al.,* 2018). Furthermore, interstrand cross-links (ICL) can be repaired by NER, structure-specific endonucleases, translesion DNA synthesis (TLS), homologous recombination (HR) (Hashimoto *et al.,* 2016) or MMR (Kato *et al.,* 2017). However, the major target of MMR is the repair of mismatches introduced by inefficient polymerase activity, and only a few oxidised bases could be repaired in such a way (Bridge *et al.,* 2014). The major repair pathway for base modifications and SSG is base excision repair (BER) (Bordin *et al.,* 2021). However, SSG repair (SSGR) slightly differs from classical BER, in that it does not require abnormal base removal by a DNA glycosylase and AP site elimination by a variety of lyase and endonuclease activities (Caldecott, 2008). SSGs are recognised by PARP or indirectly by APE1 during BER. Finally, the most toxic DSBs are repaired either by non-homologous end joining (NHEJ) or homologous directed repair (HDR) (De Falco and De Felice, 2021).

4.1.1. DNA repair in humans

Repair pathways in mitochondria and nuclei are different (Fig. I.9). For example, in humans, mtDNA molecules harbouring abasic sites or DSBs are directly degraded rather than repaired. This is possibly the consequence of the very high ploidy of the human mtDNA. It was shown that TFAM promotes strand cleavage at AP sites (Xu *et al.,* 2019). Naturally, SSGs in both the nucleus and mitochondria are repaired by BER (Krokan and Bjoras, 2013). However, it was shown, *in situ*, that BER factors OGG1, UDG, APE1, and pol γ can be inhibited by TFAM (Canugovi *et al.,* 2010). Persistent SSGs during replication cause the collapse of replication forks and the introduction of DSBs (Kuzminov, 2001). In humans, mitochondria DNA containing DSB is selectively degraded by components of the replication machinery (Moretton *et al.,* 2017; Peeva *et al.,* 2018). Although it is generally accepted that there is no HDR in human mitochondria, several studies claimed microhomology-mediated repair and HR in rat tissues (Tadi *et al.,* 2016; Dahal *et al.,* 2018). Yet, till now, there is only one case recorded in humans where mtDNA is detected of paternal and maternal mtDNA hybrids in the muscle tissue (Kraytsberg *et al.,* 2004). Meanwhile, in the nucleus, DSBs can cause the loss of

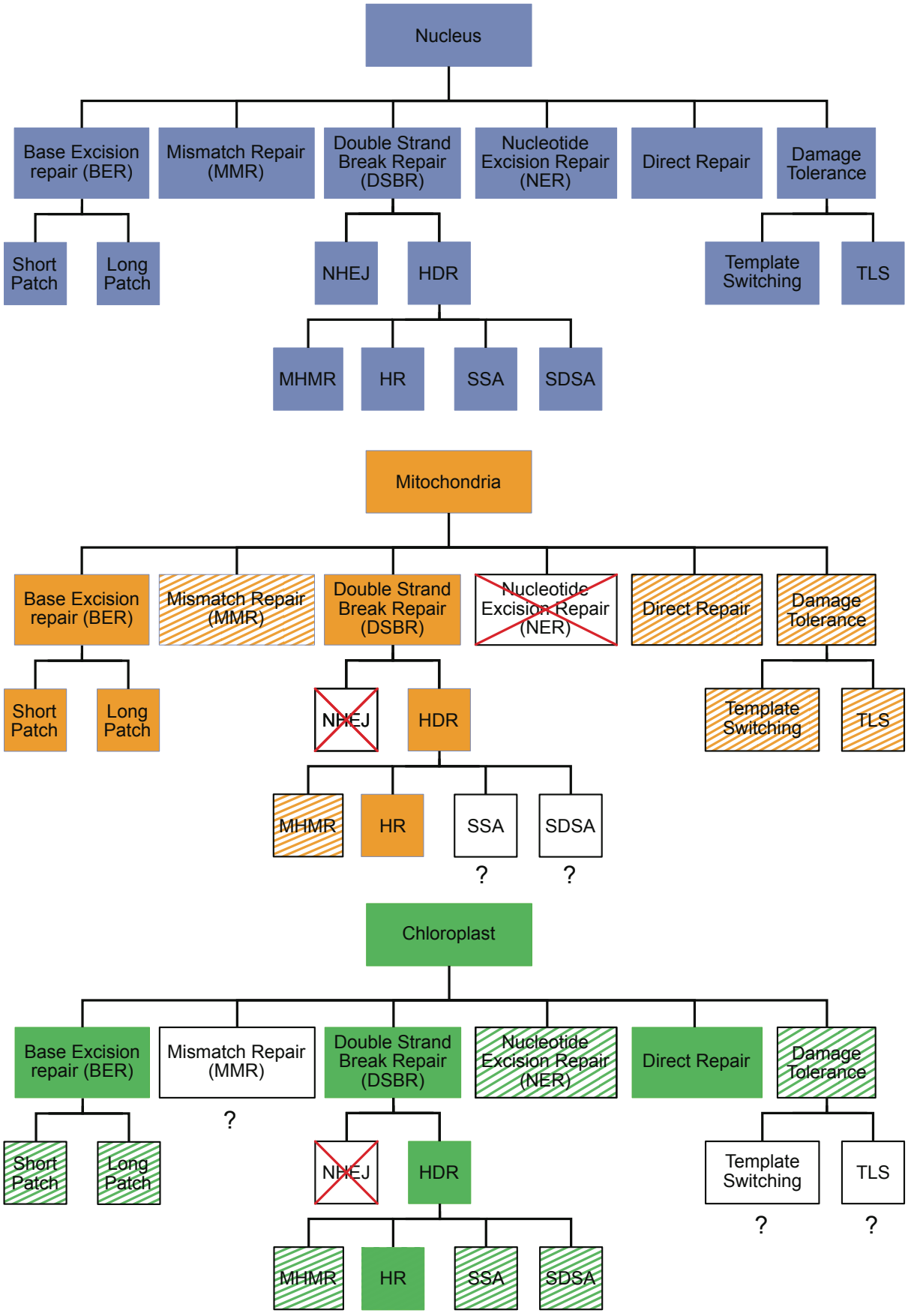


Figure I. 10. DNA repair pathways in the nucleus, mitochondria, and chloroplast of plant cells. Established pathways are shown in full colour; pathways that need further confirmation are shown with a hatched background; pathways that have not been identified or are absent are shown in a white background. Abbreviations: HDR, homology-directed repair; NHEJ, non-homologous end joining; MHMR, microhomology-mediated repair; HR, homologous recombination; SSA, single strand annealing; SDSA, synthesis-dependent strand annealing; TLS, translesion synthesis. Adapted from Boesch et al., 2011.

chromosome fragments and genomic translocations, which are repaired by HDR or NHEJ pathways (Cannan and Pederson, 2016). The BER (Prakash and Doublié, 2015), and MMR (Wen *et al.,* 2021) pathways have been detected in both mitochondria and nucleus. Yet, many components of MMR in mitochondria remain unknown. NER is the only repair pathway not identified in mitochondria (Cline, 2012). The existence of direct repair in human mitochondria remains under question. A methyltransferase with a similar molecular weight to its nuclear counterpart was found in Chinese hamster mitochondria. However, it has not been identified in human mitochondria, and the claimed mitochondrial localization remains questionable (LeDoux *et al.,* 1992; Rasmussen and Rasmussen, 2005). Furthermore, photolyases resolving thymidine and pyrimidine dimers have been found in vertebrates such as fish, amphibians, birds, and reptiles, yet not in mammals, including humans (Ramírez-Gamboa *et al.,* 2022). There is some inconclusive evidence that in plant and human mitochondria, short-range inversions might occur during replication stalling on the leading strand, which may cause a template change to the lagging strand, according to the model (Tremblay-Belzile *et al.,* 2015). Template switching could explain some deletions in mtDNA reported in patients (Oliveira *et al.,* 2020). Similarly, template switching nicely explains the clustering of mutations in the human genome (Löytynoja and Goldman, 2017), and resemble alternative polyadenylation profiles (Balázs *et al.,* 2019). Furthermore, in humans, trans-lesion synthesis was proposed to bypass the replication stall, performed by REV3 (Polζ) polymerase in mitochondria. In the nucleus, DNA polymerases ζ and κ, protect XPV cells against UV cytotoxicity, via TLS and independently of nucleotide excision repair (Ziv *et al.,* 2009).

4.1.2. DNA repair in plants

As in humans, different DNA repair pathways are found in plant nuclei, mitochondria, and chloroplast (Fig. I.10). The BER pathway has been well-defined in plants' nuclei and mitochondria. Although the activity of BER in the chloroplast was detected, many crucial details are still missing (Grin *et al.,* 2023). The MMR pathway has been extensively elucidated in plant nuclei, however, just a single MutS-like protein (MSH1) was found to be targeted to organelles (Gonzalez and Spampinato, 2020; R., Abdelnoor et al., 2006). NHEJ has been seen to be the major repair pathway in eukaryotic cell nuclei in somatic cells and has been well-defined for T-DNA insertion in rice (Saika *et al.,* 2014). Ku70 and Ku80 were found to be key for NHEJ and to suppress HR (Nishizawa-Yokoi *et al.,* 2012). NHEJ is either completely absent in the plant organelles or very rare (Kwon *et al.,* 2010; Cappadocia *et al.,* 2010). In plant mitochondria, HDRs, or more commonly HR, is the primary repair pathway (Cappadocia *et al.,* 2010). In *ku80* Arabidopsis mutants, deficient in NHEJ, it was found that MHMR (also called

back-up or alternative NHEJ, B/alt-NHEJ) was more often observed in the repair by end-joining (Jia *et al.,* 2013). However, MHMR has been shown to result in gross genome rearrangements and instability in mitochondria and chloroplast (Gualberto and Newton, 2017; Marechal *et al.,* 2009). In somatic tissues, single-strand annealing (SSA) and synthesis-dependant strand annealing (SDSA) are responsible for the majority of HDR outcomes (Puchta, 2004; Osman *et al.,* 2011). On the contrary, HR is the major HDR pathway in mitochondria, while SSA and SDSA are assumed to take place, yet such an assumption has not been validated (Chevigny et al., 2020). Repair fingerprints of HR, SSA and SDSA have been reported in the Chloroplast of *Chlamydomonas,* yet little is known about the factors involved (Odom *et al.,* 2008). Most proteins involved in NER have not been found in plant organelles. Only a homolog of bacterial TRCF, involved in transcription-coupled NER, was found in the chloroplast, which might also take part in BER, as human ortholog CSA (Kamenisch *et al.,* 2010; Gualberto and Newton, 2017). Furthermore, direct repair has been observed in the mitochondria of S*accharomyces cerevisiae* (Yasui *et al.,* 1992) and in the chloroplast of *Chlamydomonas reinhardtii* (Small, 1987). However, photolyase activity was not detected in the chloroplast and mitochondria of Arabidopsis and rice (Chen et al., 1996; Takahashi et al., 2011). Finally, little is known about DNA damage tolerance in plant organelles. Recent studies have identified Arabidopsis polymerase PrimPol, supposedly involved in nuclear, mitochondrial and chloroplast translesion DNA synthesis (TLS), which *in vitro* was shown able to bypass DNA lesions like thymine glycol and abasic sites (García-Medel *et al.,* 2021).

4.2. Nucleotide Excision Repair (NER)

Nucleotide Excision Repair is the major repair pathway that deals with UV-induced photoproducts and other bulky DNA lesions (Nouspikel, 2009). It is present in all three domains of life: eukaryotes, bacteria, and archaea (Farnell, 2011). NER was first discovered in *E. coli* and described as a “dark repair” (Harm, 1968) that works together with “light repair”, performed by photolyases (Rupert *et al.,* 1958). The major mode of action, named “dual incision“, was revealed in 1980 in bacteria and in 1990 in human cells (Huang *et al.,* 1992; Sancar and Rupp, 1983). In most eukaryotic species, including humans, it was determined that XPA, RPA, XPC, TFIIH, XPF and XPG are essential components of the NER machinery (Aboussekhra *et al.,* 1995). NER forms two main pathways: global genomic repair (GGR) and transcription-coupled repair (TCR) (Hanawalt, 2002).

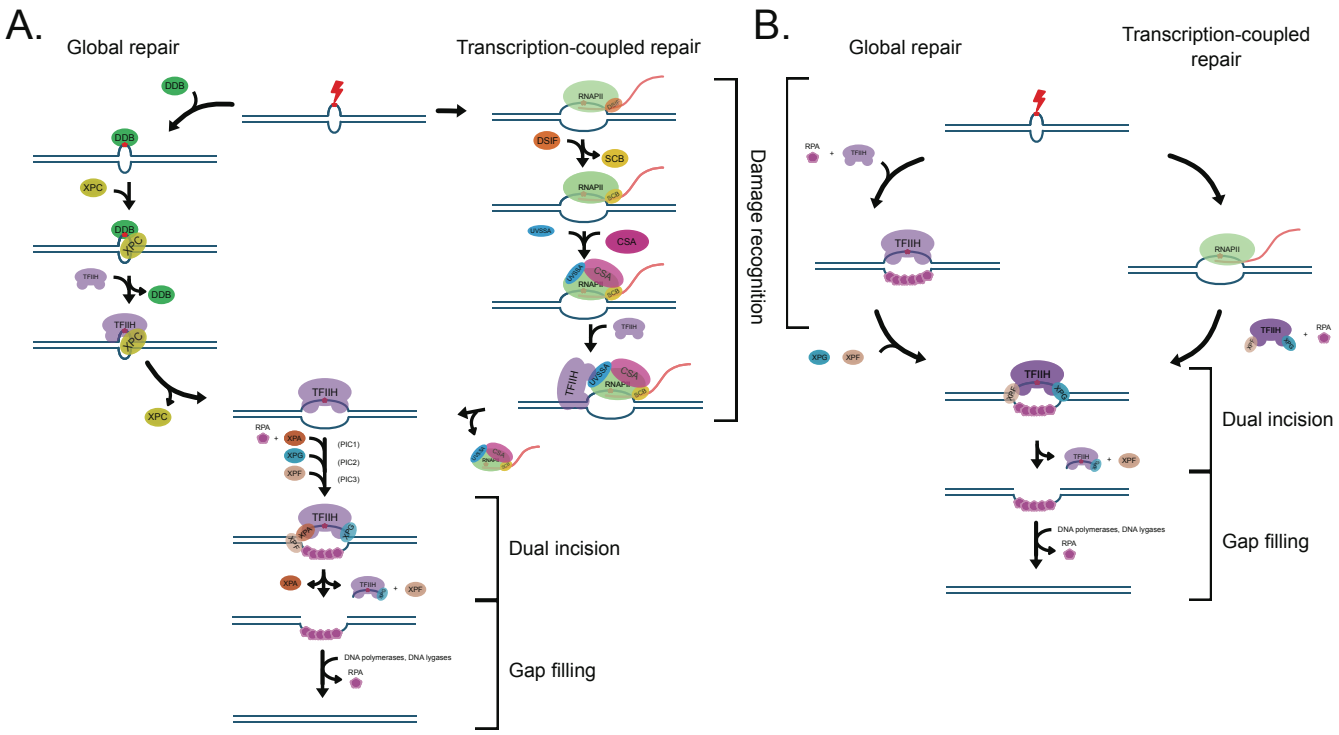


Figure I. 11. Simplified mechanism of Nucleotide Excision Repair (NER). A) Schematic representation of Global genomic repair NER (GGR) and transcription-coupled repair (TCR) in eukaryotes. Sub-pathways differ in damage recognition. In GGR, DDB2 recognise DNA lesions and recruits XPC. Upon binding of XPC to the DNA lesion, DDB is removed, and TFIIH is recruited. Ubiquitination of XPC by RNF111 leads to its eviction. In TCR, upon stall of RNA polymerase II, its factor DSIF is replaced by CSB, which promotes the recruitment of UVSSA and CSA. Ubiquitination of complex promotes recruitment of TFIIH by UVSSA. The binding of TFIIH to DNA lesions evicts the RNAPII-CSB-CSA-UVSSA complex. From here on, both subpathways GGR and TCR share the mechanism of dual incision and gap filling. TFIIH sequentially recruits the XPA, XPG, RPA and XPF, forming pre-incision complexes 1-3 (PIC1-3). XPF and XPG perform 5′ and 3′ incisions on the damaged strand. The damaged fragment is removed in a complex with TFIIH and XPG, and the gap is filled by DNA polymerase and by DNA ligase. B) The proposed NER pathways to utilise the 'minimal essential set' of excision repair factors. In GGR, the damaged site can be recognised just by the RPA-TFIIH complex, followed by subsequent recruitment of XPF and XPG nucleases. In TCR, DNA damage can be recognised by a 'transcription bubble' that initiates the assembly of the RPA-TFIIH-XPF-XPG complex. Figure adapted from (Kose et al., 2024; Nieto Moreno et al., 2023).

The first step of NER differs between GGR and TCR (Fig. I.11). GGR deals with damage across the whole genome, while TCR only repairs damage on the template strands of transcribed regions. In GGR, the DDB complex binds to the DNA damage site and recruits XPC (Fei *et al.,* 2011). Upon binding to DNA, the CRL4 subunit of DDB ubiquitinates surrounding proteins such as DDB, histones and XPC. Ubiquitination of XPC promotes damaged DNA handover from DDB to XPC, and ubiquitination of histones leads to decompaction of the damaged nucleosome (Luijsterburg *et al.,* 2012; Ribeiro-Silva *et al.,* 2020). Upon loading, XPC recruits TFIIH to the lesion (Okuda *et al.,* 2017). Ubiquitination of XPC by RNF111 leads to its eviction(van Cuijk *et al.,* 2015).

The first step of TCR involves the stall of RNA polymerase II (RNAPII) due to DNA damage, such as a bulky adduct. Upon stop of transcription, polymerase factor DSIF is replaced by CSB (Tantin *et al.,* 1997), which promotes the recruitment of UVSSA and CSA (Fischer *et al.,* 2011; van der Weegen *et al.,* 2020). Like DDB, CSA forms a complex with CRL4 and ubiquitinates TCR components. Together, they form an unstable complex, that mediates UVSSA recruitment of TFIIH to DNA damage site. Upon TFIIH binding to the site, the rest of the RNAPII-CSB-CSA-UVSSA complex is released (Nieto Moreno *et al.,* 2023).

After TFIIH loading to the DNA damage site, TCR and GGR pathways are the same. The TFIIH contains a helicase subunit that can unwind DNA double strands, allowing the binding of XPA and RPA (Coin *et al.,* 2007). TFIIH-XPA-RPA form pre-incision complex 1 (PIC1). After complex formation, endonuclease XPG is recruited, forming pre-incision complex 2 (PIC2). Recruitment of final XPF endonuclease forms pre-incision complex 3 (PIC3). XPF and XPG perform the 5′ and 3′ incisions, respectively (Mu *et al.,* 1996). The removed DNA is newly synthesized using the undamaged DNA strand as a template, and gaps are ligated (Paul-Konietzko *et al.,* 2015; Staresincic *et al.,* 2009). The best example of NER importance is that any defect causes Xeroderma Pigmentosum (XP), a rare UV hypersensitivity syndrome. These patients tend to develop skin changes, including pigmentary abnormalities, capillary telangiectasias and atrophy (Daya-Grosjean, 2008; DiGiovanna and Kraemer, 2012). Moreover, they often develop skin cancer before the second decade of life.

The homologs of almost all human NER proteins have been found in Arabidopsis, except for XPA, which verifies DNA damage and stabilizes the DNA as it is repaired (Kunz *et al.,* 2005). Recent phylogenetic analyses across multiple species, together with excision repair analysis under UV-irradiation, proposed a ‘minimal essential set’ of excision repair factors proteins, although it showed reduced repair efficiency (Fig. I.11B) (Kose *et al.,* 2024). In GGR, the damaged site can be recognised just by the RPA-TFIIH complex, followed by subsequent

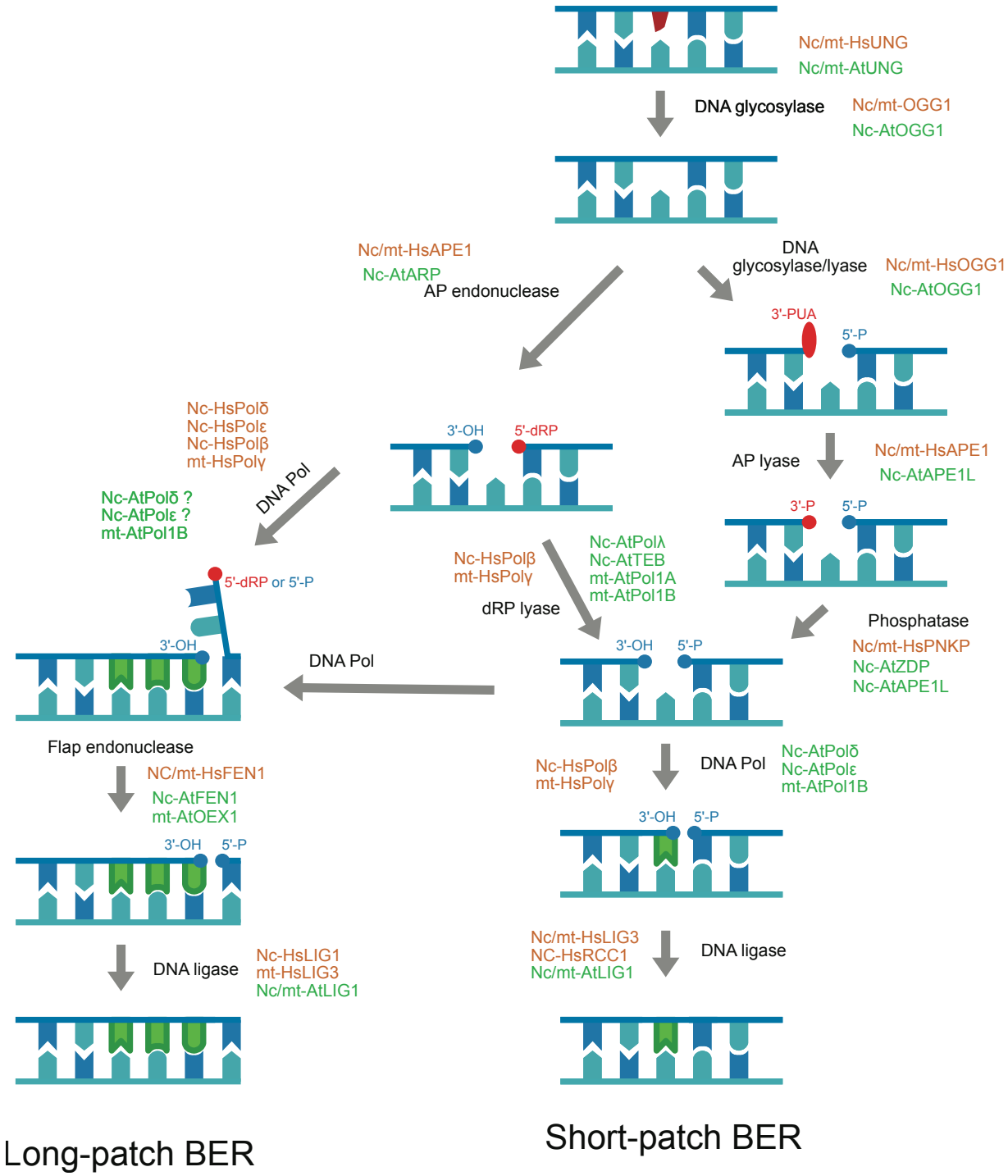


Figure I. 12. Mechanism of Base excision repair. DNA glycosylase first recognises and transforms DNA lesions into apurinic sites (AP). AP site can be incised either by AP lyase or AP endonuclease, creating 3'PUA and 5'-dRP, respectively. The 3' end is further processed by lyase and phosphatase to transform it into 3'P and 3'OH, respectively. The 5'-dRP end can be converted to 5'-P by dRP lyase activity. According to how the gap is filled by specific DNA polymerases, the BER pathway can be divided into two sub-pathways: short-patch (SP), which incorporates only one missing nucleotide, or long-patch (LP), which inserts 2 to 13 nucleotides. After DNA elongation in SP, the ends can be ligated by DNA ligase, while in LP, 5' ssDNA overhangs first need to be removed by a flap endonuclease. The corresponding proteins involved in each enzymatic step in humans (brown) and in Arabidopsis (green) cells are indicated. Abbreviation: Nc, nuclear; mt, mitochondrial.

recruitment of XPF and XPG nucleases. While in TCR, DNA damage can be recognised by a 'transcription bubble' and initiate the assembly of the RPA-TFIIH-XPF-XPG complex. As was said previously, the NER pathway is not found in organelles. However, homologues of bacterial NER have been found in Arabidopsis chloroplast: Mft (also homologs of human CSB) and UvrD/MutU/RecL (UvrD homologues to XPB domain of TFIIH) (Hays, 2002; Maddukuri *et al.,* 2007).

4.3. Base excision repair (BER)

Base excision repair is the main pathway for correcting non-bulky single-base lesions in DNA (Fig. I.12). It is initiated by DNA glycosylases, which recognize and excise damaged bases, such as those affected by oxidation, alkylation, and deamination. It is the only repair pathway for single-strand gaps (SSG). The discovery of the BER pathway started in 1974 with the detection of uracil-DNA glycosylase (Ung) in *Escherichia coli* (Lindahl, 1974). Tomas Lindahl highlighted that the produced AP site would require a DNA endonuclease, polymerase, and a ligase for its repair – the key steps of BER. The BER pathway is highly conserved in archaea, protozoa, fungi, animals, and plants (Nischwitz *et al.,* 2023). However, the evolutionary conservation of glycosylases is limited to the enzymatic core domain. The Mammalian DNA glycosylases tend to have amino- and/or carboxy-terminal extensions (Hegde *et al.,* 2010). Five structural superfamilies of DNA glycosylases have been identified, according to the conserved distinctive protein fold motifs (Dalhus *et al.,* 2009). Members of uracil DNA glycosylases (UDG) that remove uracil from DNA are found in Arabidopsis mitochondria and nuclei (Boesch *et al.,* 2009). It is the most extensively studied and the most widely distributed superfamily: eubacteria, archaea, yeast, animals, and plants (Schormann *et al.,* 2014). The AAG superfamily is composed of monofunctional glycosylases that remove alkylated purines and ethenopurines, found in humans (HsAAG) and Arabidopsis (AtAAG) (Santerre and Britt, 1994). The HhH-GPD superfamily is the most heterogenous, having a different substrate specificity: 3-methyladenine DNA glycosylases remove alkylated purines (Dalhus *et al.,* 2009); Endonucleases III remove oxidatively damaged pyrimidines and have AP lyase activity (AtNTH1 and AtNTH2 are targeted to nucleus and chloroplast (Gutman and Niyogi, 2009)); 8-oxoguanine DNA glycosylases remove 8-oxoG oxidation product of purines (AtOGG1 found in the nucleus (Dany and Tissier, 2001)); thymine-DNA mismatch glycosylases target T:G mispairs (AtMBD4 found in the nucleus (Nota et al., 2015a)); DEMETER-LIKE endonucleases are plant-specific 5-meC DNA glycosylases/lyases involved in active DNA demethylation via BER (Ortega-Galisteo *et al.,* 2008). The H2TH superfamily proteins contain a zinc finger DNA binding domain and are bifunctional DNA glycosylases that

cleave the sugar-phosphate backbone (Huffman *et al.,* 2005; Sugahara *et al.,* 2000). They could be grouped into two families: *E. coli* formamidopyrimidine (FPG) DNA glycosylases, and Endonucleases VIII (NEI). In eukaryotes, only FPG is found in the plant and fungi clades. It was found that AtFPG might be involved in mitochondrial BER (Córdoba-Cañero *et al.,* 2014). The last superfamily, HEAT-like repeat (HLR), is mostly found in prokaryotes and not present in plants (Mullins *et al.,* 2015).

After one of the glycosylases mentioned above removes the damaged base, an AP site (apurinic or apyrimidinic site) is created that can be excised by AP lyase and AP endonuclease, producing 3′-PUA or 3′-P and 5′-dRP, respectively (Fig. I.12). These need to be processed to conventional 3′-OH and 5′-P termini, to allow DNA polymerization and ligation. Blocked 3' ends arise from bifunctional DNA glycosylases/AP lyases, such as FPG or OGG in humans and Arabidopsis. The 3′-PUA can be processed by bifunctional APE1 endonuclease, which has 3′-phosphodiesterase activity and can transform 3′-PUA into 3′-P in humans (Suh *et al.,* 1997), and by the Arabidopsis homolog APE1L (Joldybayeva *et al.,* 2014). Contrarily to plant APE1L, the human APE1 has weak 3' phosphatase activity. Thus, phosphatase activity is performed by polynucleotide kinase/3′-phosphatase (PNKP) (Jilani *et al.,* 1999). Orthologs of PNKP in maize (Betti *et al.,* 2001) and Arabidopsis (Petrucco *et al.,* 2002) have also been identified.

On the other hand, AP endonucleases like APE1 can excise AP sites, generating 5′- dRP. Furthermore, APE1, specifically, has been demonstrated to participate in Nucleotide Incision Repair (NIR) in both humans and Arabidopsis. NIR is a glycosylase-independent repair pathway that was proposed to be promoted by APE1 exonucleases(Akishev *et al.,* 2016; Ischenko and Saparbaev, 2002). In both cases, BER and NIR create 5′-dRP ends that can be further processed by dRP lyase to 5′-P ends. In mammals, activity is associated with DNA Polymerase β and λ in the nucleus (Srivastava et al., 1998), and with DNA Polymerase γ (Prasad *et al.,* 2009) in mitochondria. Pol λ orthologs have been identified as the sole orthologs of the Polymerase X-family in plants. The dRP lyase activity of Pol λ has been shown in the nucleus of Arabidopsis and rice (Uchiyama *et al.,* 2004). In Arabidopsis mitochondria, organellar polymerases Pol1A and Pol1B were proven to have 5′-dRP lyase activity (Trasviña-Arenas *et al.,* 2018).

The BER pathway can be divided into two sub-pathways by gap-filling: short-patch (SP), which incorporates only one missing nucleotide, or long-patch (LP), which inserts 2 to 13 nucleotides irrespective of the gap size. In humans, during SP-BER, the incorporation of bases and ligation is facilitated by DNA Pol β and DNA Ligase III (LIG3) (Kubota *et al.,* 1996).

It was believed that plants had just LP-BER, as no LIG3 orthologs were identified in plants, until the discovery that the LIG1 plant ortholog could catalyse the ligation step of both SP-BER and LP-BER (Córdoba-Cañero *et al.,* 2011). However, the gap filling in SP-BER still remains unknown. Gap filling in LP-BER is facilitated by Pol δ and Pol ε in humans, and orthologs of both have been identified in Arabidopsis (Ronceret *et al.,* 2005). In human mitochondria, DNA synthesis during gap filling in both BER pathways is facilitated by Pol γ (Longley *et al.,* 1998), while in plant mitochondria, this function is performed by Pol1B (Trasviña-Arenas *et al.,* 2018).

In humans, the equilibrium between SP-BER and LP-BER is assumed to be affected by the nature of the lesion introduced by the DNA glycosylase, which may vary depending on the phase of the cell cycle (Fortini and Dogliotti, 2007). Conversely, in plants, it was demonstrated that BER is highly dependent on the nature of the 5'-end of the DNA gap, as the SP-BER sub-pathway is unavailable in the absence of dRP lyases (Córdoba-Cañero et al., 2011). Finally, to ligate the synthesised gap in LP-BER, overhang DNA must be cleaved by a DNA endonuclease. In human cell nuclei and mitochondria, this function is performed by Flap Endonuclease 1 (FEN1) (Roldán-Arjona *et al.,* 2019). In Arabidopsis, the FEN1 ortholog is targeted only to the nucleus (Kimura *et al.,* 2000), while in mitochondria, such function could be performed by AtOEX1, according to *in vitro* studies (Schatz-Daas, 2021).

Several additional proteins have been linked to the BER pathway, with no specific catalytic activity in repair: Proliferating Cell Nuclear Antigen (PCNA), Scaffolding Protein X- Ray Cross Complementation Group 1 (XRCC1) and Poly (ADP-Ribose) Polymerases (PARP). As previously noted, BER is the only SSG repair pathway. However, differently from classical BER, no DNA glycosylases or AP nucleases are required. It has been shown that PARP1 is required for the detection of SSG and the recruitment of BER repair proteins to the damaged site (Caldecott *et al.,* 1996). In mice, PARP1-and PARP2-deficient cells display a significant delay in resealing of DNA strand breaks (Caldecott *et al.,* 1996). Similarly, Arabidopsis PARP mutants *Atparp* and *Atparp2* show an increased sensitivity to DNA- damaging agents (Boltz *et al.,* 2014). In eukaryotes, PCNA is an accessory factor required for efficient DNA synthesis by Pol δ or Pol ε, in both LP-BER (Stucki *et al.,* 1998) and SP-BER (Fan, 2004). Arabidopsis orthologs of PCNA have been identified (Strzalka *et al.,* 2009), yet no data has been reported on their involvement in BER. The XRCC1 complex has been shown to play a major role in SSGR and BER, in mammalian cells (Caldecott, 2003), and bind most of the proteins involved in the pathway described above. XRCC1 knock-out mice are embryonic lethal (Thompson and West, 2000). In contrast, Arabidopsis *xrcc1* mutants develop normally, although they show radiosensitivity (Charbonnel *et al.,* 2010). AtXRCC1 was

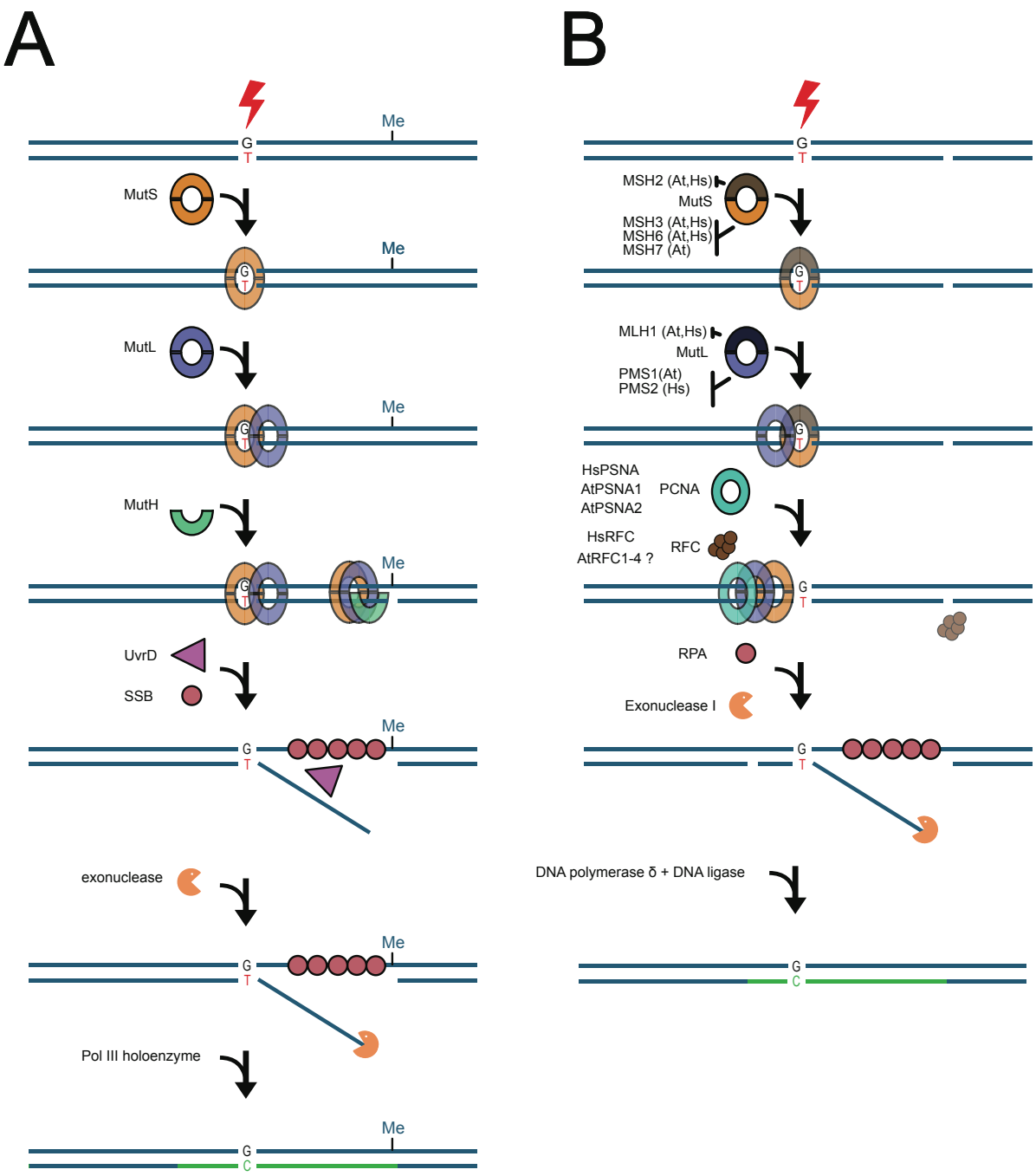


Figure I. 13. Mechanism of mismatch repair (MMR) in prokaryotes (A) and eukaryotes (B). In prokaryotes, MutS homodimer recognises the mismatched base, and MutL homodimer stabilises the complex. Activated MutH endonuclease incises the unmethylated GATC site to create an entry point for UvrD helicase. UrvD unwinds the non-parental strand, which is excised by exonuclease. ssDNA is stabilised by SSB. The gap is filled by DNA polymerase III and sealed with DNA ligase. In eukaryotes, MSH heterodimer recognises the mismatched base and subsequently recruits MutL heterodimer, PCNA and RFC. The activated PMS domain of MutL incises the non-parental DNA next to the mismatch. Strand discrimination remains unclear, but evidence suggests that it is nick-directed. The mismatch-containing strand is excised by exonuclease and ssDNA stabilised by RPA. The gap is filled by DNA polymerase δ and sealed with DNA ligase. Respective MMR proteins involved in humans (Hs) and Arabidopsis (At) are indicated. Adapted from (Kadyrova and Kadyrov, 2016; Matson and Robertson, 2006).

shown to stimulate uracil BER *in vitro* (Córdoba-Cañero *et al.,* 2009) and is required for efficient DNA ligation (Nota *et al.,* 2015). In mammalian cells, the BER pathway was also shown to be involved in transcription, yet little is known about it (Chakraborty *et al.,* 2021).

4.4. Mismatch repair (MMR)

Mismatch repair is a highly conserved mechanism of DNA repair from prokaryotes to higher eukaryotes, for repair of mispaired bases, insertion-deletions, loops, elimination of unnecessary heteroduplexes, psoralen-induced inter-strand crosslinks (ICLs), and oxidative DNA damage (Groothuizen and Sixma, 2016; Manova and Gruszka, 2015). MMR was first discovered in 1964 to explain the repair of brominated nucleotides in E. coli (Witkin and Sicurella, 1964) and as a mechanism of gene conversion in yeast (Holliday, 1964). The genetic machinery underlying MMR was established when the *hexA Pseudomonas* mutant was found to be defective in gene conversion and was linked to the previously established "Treffers Mutator" (Treffers *et al.,* 1954; Tiraby and Fox, 1973), "Siegel Mutator" (MutS) of *Escherichia coli* (Siegel and Bryson, 1967), *Salmonella* LT7 Mutator (MutL), Hill Mutator (MutH) and MutU (UvrD) (Miyake, 1960; Hill, 1970; Nevers and Spatz, 1975). Mutations in each of them resulted in increased mutation rates.

Four main enzymes (Mut)S, MutL, MutH and UvrD direct the recognition and removal of mismatches in prokaryotes (Fig. I.13A) (Glickman and Radman, 1980; Matson and Robertson, 2006). MutS homodimer recognizes base-base mismatches (Su and Modrich, 1986). Upon binding to the mismatch site, MutH and MutL homodimers are recruited to the complex (Galio *et al.,* 1999; Guarné, 2012). MutL, upon binding, activates MutH endonuclease activity that recognizes and specifically incises the unmethylated daughter strand at the closest hemimethylated dGATC site, either 5' or 3' to the mismatch (Ban, 1998; Lee *et al.,* 2005). Loading of the UvrD helicase at the cut site is stimulated by MutL. It unwinds the DNA towards the mismatch (Mechanic *et al.,* 2000), allowing the excision of the non-parental DNA strand beyond the mismatch, followed by ssDNA stabilisation and protection by SSB proteins. The non-parental DNA is excised next to the mismatch by ExoI or ExoX (3'-5' exonuclease), or by ExoVII or RecJ (5'-3' exonucleases), depending on the position of the strand break relative to the mismatch, and released from the parental strand (Burdett *et al.,* 2001). Finally, DNA polymerase III, SSB, and DNA ligase carry out DNA synthesis to seal the gap (Lahue *et al.,* 1989).

Eukaryotic MMR is similar to the bacterial one in substrate specificity, bi-directionality, and nick-directed strand specificity (Fig. I.13B). However, there are notable differences. First, in eukaryotes, MutS and MutL heterodimers are formed rather than homodimers. MutS was

shown to have seven orthologs in eukaryotes, MSH1 to MSH7. Different heterodimers of MutS were shown to have different specificities: MSH2–MSH6 primarily targets base–base mismatches and small insertion/deletion loops, while MSH2-MSH3 is involved in the removal of large insertion/deletion loops (2–12 nucleotides) (Marti *et al.,* 2002). MSH4–MSH5 is involved in heteroduplex stabilization during meiotic recombination (Hollingsworth *et al.,* 1995). The plant-specific heterodimer MSH2– MSH7 targets base–base mismatches with different specificities (Wu, 2003). It is believed that MSH7 might have evolved from MSH6.

In eukaryotes, MutL has been shown to have five orthologs: MLH1-3 and PMS1-2. Similar to MutS, MutL forms heterodimers in eukaryotes: MLH1-PMS2 in humans (Nicolaides *et al.,* 1998), MLH1-PMS1 in Arabidopsis (Alou *et al.,* 2004), and MLH1-MLH3, which is involved in meiotic crossovers in both Arabidopsis and humans (Jackson et al., 2006; Cannavo et al., 2005). Furthermore, in eukaryotes, MutH homologs are absent. Instead, endonuclease activity is performed in humans by PMS2 (Kadyrov *et al.,* 2006), and most likely by PMS1 in plants. Although the endonuclease activity of PMS1 in plants has not been demonstrated, the protein is highly similar to PMS1 from yeast (Smith *et al.,* 2015; Li *et al.,* 2009).

Several reviews suggest that strand specificity in eukaryotic MMR is ensured by leading- and lagging-strand synthesis, in so-called nick-directed MMR (Kadyrova and Kadyrov, 2016; Fishel, 2015; Fukui, 2010). Misincorporated ribonucleotides during S-phase are removed by RNase H2, leaving a strand scission that serves as a signal for strand discrimination, as shown in yeast *S. cerevisiae* (Lujan *et al.,* 2013; Ghodgaonkar *et al.,* 2013). However, it remains unclear how MMR ensures communication between two physically distant DNA sites, namely the location of nicked DNA and the loci of mismatch. As has been shown, the 5'-excision reaction only requires MSH (MSH2-MSH6), EXOI, and RPA (Dzantiev *et al.,* 2004). In contrast, proliferating cellular nuclear antigen (PCNA) and its clamp loader replication factor C (RFC) are essential for 3'-excision, involving 5'-3' exonuclease EXOI. This is because both 3' and 5' excisions utilise the same EXOI exonuclease. The exact mechanism remains to be understood. Involvement of PCNA in MMR has also been confirmed for Arabidopsis orthologs AtPCNA1 and AtPCNA2 (Xue *et al.,* 2015; Gopalakrishnan Nair and Chung, 2014).

Little is known about the MMR in human mitochondria. Yet in 2003, MMR activity in rat liver mitochondrial extracts was reported (Mason, 2003), but no mitochondrial MSH or MLH have been identified in humans. However, some reports showed that depleting Y-box binding protein1 YB-1 in mice mitochondria leads to a decrease in MMR activity and an increase in mtDNA mutations, suggesting the involvement of alternative factors (Lyabin *et al.,* 2014). In

Arabidopsis, MSH1 (MutS ortholog) is the only known MMR factor identified in both mitochondria and chloroplasts. MSH1 contains a homing endonuclease domain capable of incising DNA, akin to the role of MutH in bacteria (Abdelnoor et al., 2006). However, no activity was recorded *in vitro*, suggesting it may require interaction with other factors (Fukui *et al.,* 2018). Loss of MSH1 leads to altered organellar genome stability, which was linked to short repeated sequences, suggesting a role in the surveillance of ectopic recombination (Wu *et al.,* 2020). Disruption of *MSH1* leads to a significant increase in the point mutations and small indels. Recent findings suggested that MSH1 has a high binding affinity to D-loops, whose dismantling by endonuclease activity prevents recombination between short repeats (Penafiel-Ayala *et al.,* 2023). These findings highlight the multifaceted role of MSH1 in the maintenance of organellar genome stability and its importance in preventing aberrant recombination events that can lead to genomic instability.

4.5. DSB repair

DSBs are the most harmful lesions, which can lead to genomic instability and cell death. They can arise naturally during meiosis or cellular metabolism, but more frequently they are caused by external factors such as ultraviolet (UV) radiation, ionizing radiation and genotoxic agents. The repair of DSBs is particularly difficult, as erroneous re-joining can result in the loss or amplification of chromosomal material, or error-free DNA rearrangement that can lead to the translocation of chromosomal segments (Nikiforova *et al.,* 2000). In humans, such errors can lead to tumorigenesis via changes in gene functions. However, DSB repair pathways generate diversity, which is specifically important in several important processes that rely on the creation and repair of DSBs. These include lymphocyte development and meiosis (Ranjha *et al.,* 2018). One notable process involving DSBs is V(D)J recombination, which occurs during B and T lymphocyte differentiation, allowing the production of a wide variety of antigen receptors from a limited number of gene segments (Arya and Bassing, 2017). Furthermore, in plants, most genome engineering techniques are based on DSB repair: Agrobacterium-mediated transformation and genome editing by CRISPR/cas9 or by TALEN (Gehrke *et al.,* 2022; Arimura, 2021) (more about it in Chapter 2).

There are two major repair pathways for repairing DSBs: error-prone non-homologous end-joining (NHEJ) and homology-directed repair (HDR), specifically homologous recombination (HR). NHEJ is the main repair pathway in mitotically replicating cells. In mitotic cells, HR is regulated to avoid the generation of reciprocal crossovers, with most recombinations resulting in non-crossovers via synthesis-dependent strand annealing (SDSA) (Andersen and Sekelsky, 2010). In contrast, during meiosis, which aims to ensure accurate

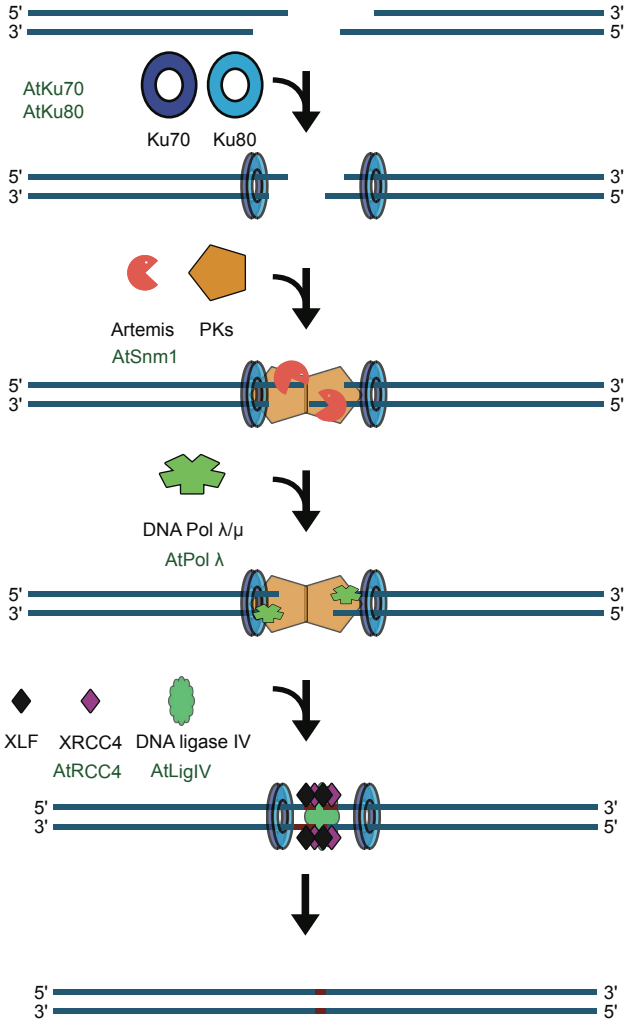


Figure I. 14. Mechanism of non-homologous end joining. Ku80 and Ku70 bind to the broken end of the double-strand break (DSB). The Ku dimer recruits protein kinase catalytic subunit (DNA-PKcs) and endonuclease Artemis. Phosphorylated Artemis processes 5' overhangs into blunt-ends and a 4 nt 3′-overhang. DNA polymerase fills the ends, and the DSB is sealed by DNA ligase. Respective proteins in humans (in black) and Arabidopsis (in green) are indicated. Adapted from Weterings and Chen, 2008.

chromosome segregation and promote genetic variability, crossover products (CO) via HR are expected to occur. In human mitochondria, DSBs in the mtDNA most frequently result in the degradation of the damaged genome copy (Moretton *et al.,* 2017). On the contrary, in plant mitochondria, HR is the major repair pathway (Chevigny et al., 2020). Furthermore, nuclear DNA is large and homologous template DNA for HR is difficult to find, while plant mtDNA is small and rich in repeated sequences, as described in section 2.2. Besides HR, there are 3 alternative HDR pathways: SDSA, Single Strand Annealing (SSA) and Microhomology-Mediated Repair (MHMR).

4.5.1. Non-Homologous End Joining (NHEJ)

The NHEJ is the major repair pathway for DSB repair in plant and mammalian cell nuclei, requiring little or no DNA sequence homology (Fig. I.14) (Weterings and Chen, 2008). First, irrespective of the broken DNA end conformation, Ku80 and Ku70 bind to them and protect them from further DNA resection (Mimitou and Symington, 2010; Mimori and Hardin, 1986). In humans, 53BP51 was shown to inhibit HR and act along Ku proteins (Bunting *et al.,* 2010). However, no orthologs have been found in plants. Ku heterodimer recruits DNA- dependent protein kinase catalytic subunit (DNA-PKcs) and endonuclease Artemis. Interaction between DNA-PKcs/Ku70/Ku80 holoenzymes forms a synaptic complex that maintains broken DNA ends in proximity, providing a platform for end-processing and ligation (Spagnolo *et al.,* 2006). It has long been thought that DNA-PKcs are absent in plants in general. However, recent phylogenetic studies have revealed them to be conserved in green plants, yet they are absent in angiosperms, suggesting that NHEJ mechanism and DNA-repair mechanism in general may not be entirely conserved in all plant lineages (Kumar, 2023). Artemis, phosphorylated by DNA-PKcs gets activated and processes the DNA ends (Goodarzi *et al.,* 2006). Artemis 5′-endonuclease activity produces a blunt end from 5' overhangs, while 3'-endonuclease activity leaves behind a 4 nt 3′-overhang (Ma *et al.,* 2005). Its ortholog in plants, SNM1, has also been shown to enhance NHEJ (Johnson *et al.,* 2011). In humans, the end-filling process is carried out by Pol X family polymerases, including polymerase λ/μ (Bertocci *et al.,* 2006), and a DSB is sealed by DNA ligase IV–XRCC4–XLF complex (Gu *et al.,* 2007). XRCC4 and XLF were shown to form long filaments along DNA ends to hold them together (Ropars *et al.,* 2011). The same process of gap-filling was observed in Arabidopsis; however, XLF does not have orthologs in plants (West *et al.,* 2000; Roy *et al.,* 2013). Regarding mitochondria, recent studies showed that there is no NHEJ in mitochondria, neither in humans (Tadi *et al.,* 2016), nor in plants (Chevigny et al., 2020).

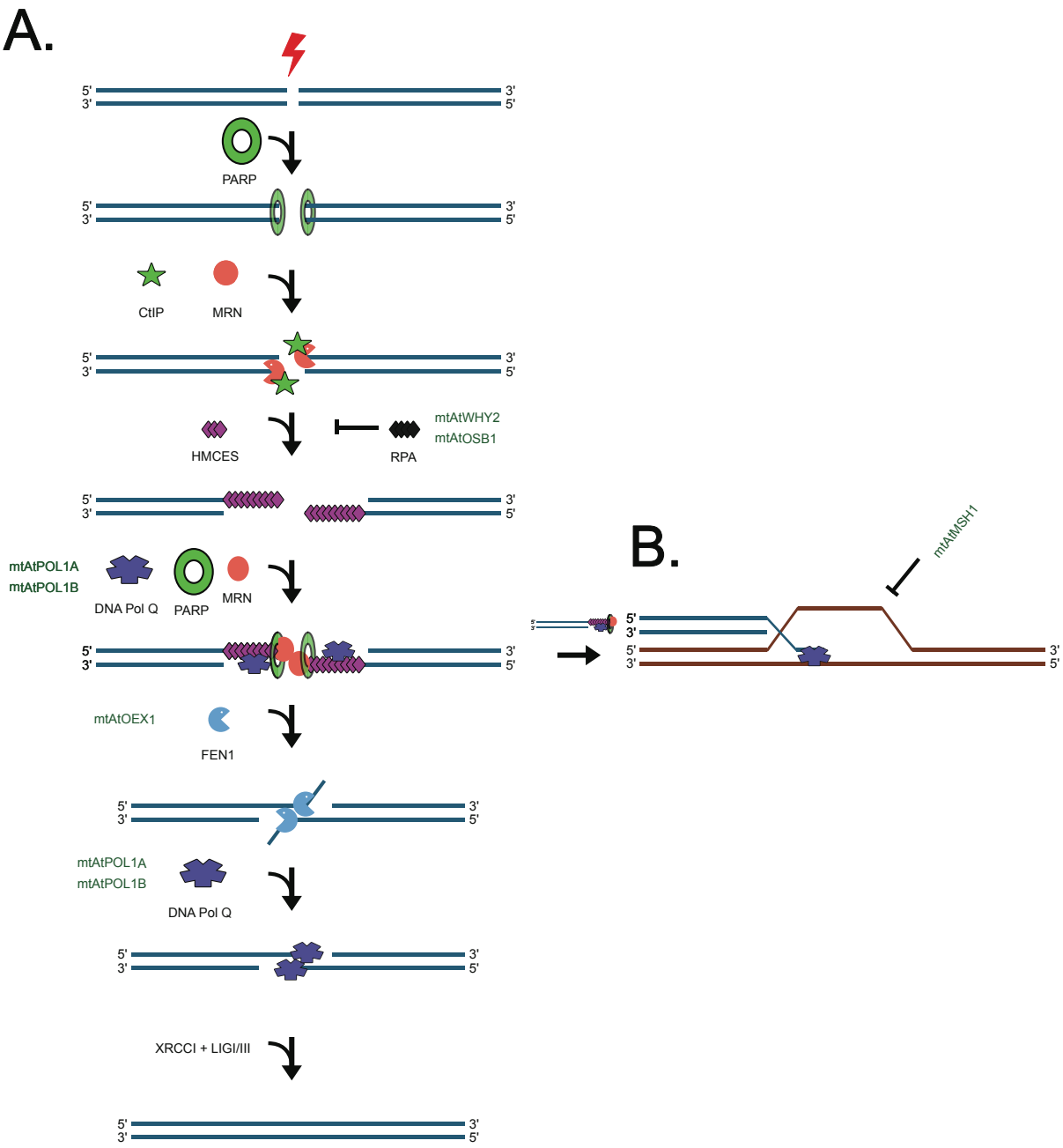


Figure I. 15. Mechanism of microhomology-mediated repair (MHMR) in eukaryotes. The pathway is determined by the binding of PARP1 to the DSB. PARP1 facilitates recruitment of MRN/CtIP complex for the recession of the DSB extremities and production of 3'-end overhangs. The ends can be protected from MHMR by RPA, or facilitated by HMCES. PolQ was shown to remove RPA, facilitating microhomology base-pairing and inhibiting HR. Homology base pairing can occur either between two ends of DSB (A), or 3'end overhang can anneal to stalled replication fork (B), to initiate break-induced replication (BIR). Non-homologous tails are removed by FLAP endonuclease FEN1, followed by DNA synthesis facilitated by Polymerase Q and end-joining by DNA ligase I/III in complex with XRCC1. The functional homologs found in Arabidopsis mitochondria are shown in green. Figure adapted from Kent et al., 2015.

4.5.2. Microhomology-Mediated Repair (MHMR)

The Microhomology-Mediated Repair or End-Joining (MHMR or MMEJ), also known as alternative MHEJ (alt-NHEJ), is a highly mutagenic DSB repair pathway driven by microhomology (MH) regions (2–20 bp) (Fig. I.15). MHMR starts with competition between Ku70-Ku80 of NHEJ and poly(ADP-ribose) polymerase-1 (PARP-1) in the nucleus (M., Wang *et al.*, 2006) and mitochondria (Tadi *et al.*, 2016). PARP-1 modifies DNA by ADP-ribosylating terminal phosphate residues (Taipakova *et al.*, 2020). Orthologs are also found in Arabidopsis (AtPARP1, AtPARP2). PARP1 helps to recruit the MRN complex (MRE11, RAD50, and NBS1) to the extremity of the 3' single-strand region (Cannavo and Cejka, 2014). Stimulated by Sae2/CtIP, MRE11 starts 3′- 5′ exonucleolytic digestion. MRN was also shown to be present in mammalian mitochondria (Dahal *et al.*, 2018). Orthologs of all MRN complexes were also found in Arabidopsis (AtMRE11, AtRAD50, and AtNBS1) (Waterworth *et al.*, 2007; Šamanić *et al.*, 2016; Gallego and White, 2001), but their mitochondrial localisation has not been demonstrated. Mutation in each protein leads to chromosomal rearrangements, which is consistent with the assumption that the MRN complex takes part in HR. Furthermore, for rearrangements involving microhomologies distant from the DSB, further resection is performed by EXO1, both in nuclei and mitochondria (Garcia et al., 2011; Dahal et al., 2018). Following resection, the ssDNA is stabilized and protected by RPA in the nucleus and by SSB protein in mitochondria. ssDNA covered in embryonic stem cells-specific protein HMCES has been shown to promote MHMR in humans (Shukla *et al.*, 2020).

Microhomology alignment is facilitated by PARP-1, MRN, and POLQ (Seki, 2003; Jäger *et al.*, 2010). POLQ has two domains: the N-terminal helicase domain and the C-terminal polymerase domain. Both were shown to be essential in MHMR, as a mutation in the helicase domain greatly reduced MMEJ (Beagan *et al.*, 2017). This is especially essential in microhomology-mediated break-induced replication (MMBIR) (Fig. I.15B). It has been proposed that during replication stalk, any single-strand tail can anneal to any ssDNA via MH, eventually re-establishing processive replication (Hastings *et al.*, 2009). A correlation was shown of PolQ activity and increased mutational rate in human mitochondria (Wisnovsky *et al.*, 2018). Similar observations have been made for Arabidopsis *tebichi*/*polq* mutants. In Arabidopsis mitochondria, MMEJ was shown to be performed by AtPolIA and AtPolIB (Lv *et al.*, 2022). Furthermore, there evidence for increased MHMR in mutants of ssDNA binding proteins AtOSBs and AtWHIRLY2, which stabilise ssDNA and protect it from MH-mediated annealing (García-Medel *et al.*, 2019). Recessed DSB ends can also interact with each other

via the MH region, similar to the SSA recombinase-independent mechanism (see section 4.5.5). In humans, it has been shown that the helicase domain of PolQ can act independently from its ATPase activity. It is required for both the suppression of HR by binding RAD51 and inhibiting its assembly along ssDNA (Ceccaldi *et al.,* 2015). RAD51 is the functional eukaryotic ortholog of bacterial RecA. It is an ATPase that forms nucleoprotein filaments on ssDNA, and is involved in locating and invading homologous DNA sequences during HR (Bonilla *et al.,* 2020), (see section 4.5.3). PolQ also facilitates the removal of RPA from ssDNA, to allow MMEJ and inhibit HR (Mateos-Gomez *et al.,* 2017). After the alignment of the MH to the complementary short sequence, non-homologous tails are removed either by the endonuclease activity of the ERCC1/XPF complex or, more likely, by flap-endonuclease FEN1 (Bennardo *et al.,* 2008). In MHMR, the gap is then filled by PolQ through the process of strand displacement and DNA synthesis, and any displaced 5′ ssDNA flaps are removed by FEN1 (Kent *et al.,* 2015; Zou and Maitra, 2008). In plant mitochondria, such function might be performed by AtOEX1 (Schatz-Daas, 2021). The MHMR is completed by joining of the two ends by ligase. In humans, it can be done by DNA Ligase I, III and IV; the major was shown to be Ligase III in complex with XRCC1 (Della-Maria *et al.,* 2011; Simsek *et al.,* 2011; Grawunder *et al.,* 1998). In plants, Ligase III has no homologs, and final ligation can be finished by ligase IV (West et al., 2000), and by LIG1 in mitochondria (Córdoba-Cañero et al., 2011).

4.5.3. Homologous recombination (HR)

Homologous Recombination (HR) is a predominant DNA repair mechanism in prokaryotes and plant mitochondria. It also occurs to a lesser extent in the nucleus of eukaryotes, although its activity is largely limited by the cell cycle, particularly when sister chromatids become available. The mechanism of HR is highly conserved, yet factors may differ across organisms. Due to the endosymbiotic origin of mitochondria, its HR machinery is partially derived from a prokaryotic ancestor (Palmer and Herbon, 1988). Thus, I will first review HR in prokaryotes, before discussing the factors identified in Arabidopsis mitochondria.

4.5.3.1. Homologous recombination in prokaryotes

In *E. coli*, HR starts with the recession of 5′-ends at the DSB, leaving a free 3'-OH ssDNA extremity (Fig. I.16). There are two distinct pathways: RecBCD and RecF. RecBCD is a complex formed by the recognition domain RecC and the two helicases RecD and RecB, of which RecB possesses endonuclease activity (Fig. I.16A). At the DSB ends, RecD (5'- 3' helicase) starts fast unwinding, followed by slow unwinding at 3' end by RecB (3'- 5' helicase). RecC forms a tunnel structure that screens DNA for crossover hot-spot

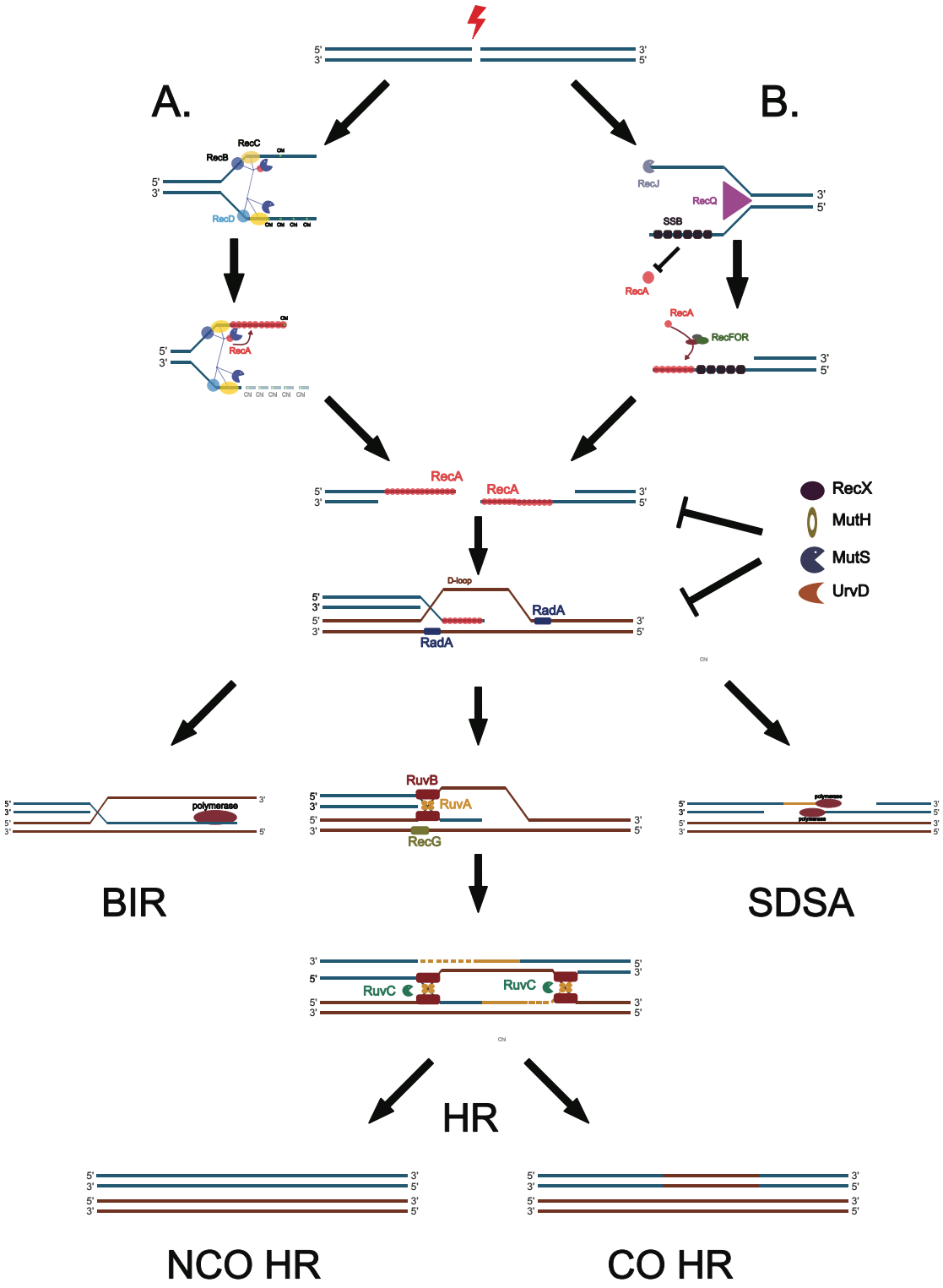


Figure I. 16. Mechanism of DSB repair by homologous recombination in bacteria. DSB repair starts with DNA ends recession. A) The RecBCD recession pathway starts with DNA unwinding facilitated by RecB and RecD helicases. RecC screens the DNA for a crossover hotspot instigator sequence (Chi), and promotes incision by RecB that also mediates the loading of RecA on the newly formed 3'-overhang. RecA protects 3-ends but not 5' ends from recession, resulting in 3'-overhangs. B) The RecF pathways involve the unwinding of DNA by the RecQ helicase, and recession of 5'-ends by the RecJ exonuclease. The resulting 3' overhang is stabilised and protected from RecA binding by SSB. The RecFOR complex manages the substitution of SSB by RecA. In both pathways, RecA forms a nucleofilament on ssDNA and screens dsDNA for homology sequences. Upon homology, RecA facilitates ssDNA invasion into dsDNA to form a D-loop. Upon D-loop formation, HR alternative pathways become available: BIR and SDSA. In BIR, DNA synthesis of the invaded strand might proceed until it covers the entire chromosome. In SDSA, after initial extension, the D-loop is dismantled, and DNA synthesis and ligation are finished on dsDNA. In HR, the proper ssDNA annealing and further extension are promoted by branch migration, mediated by helicases RadA and RecG. Recruitment of the second strand of invading DNA forms a four-stranded DNA structure called a Holliday Junction (HJ) that is stabilised by RuvAB. After synthesis of the missing DNA sequences, HJ are resolved by the RuvC endonuclease, to form either crossover (CO) or non-crossover (NCO) products. Adapted from Chevigny et al., 2020.

instigator' (Chi), of eight nucleotides long (5'-GCTGGTGG-3') (Amundsen and Smith, 2019; Lao and Forsdyke, 2000). The Chi signal detected by RecC is transduced to RecD to stop unwinding and to RecB to make a cut 4-6 nt 3' from the chi sequence and loads RecA on the newly generated 3' end. The addition of RecA stops the degradation of 3'-ends but does not stop the degradation of the 5′ ends, which leads to the formation of a 3′ ssDNA tail coated with RecA (Wiktor *et al.,* 2018). It is important to note that RecBCD is called an 'exonuclease' because it requires a free DNA-end for its activity, but it makes endonucleolytic scissions.

In the absence of RecBCD, the RecF pathway can be alternatively employed for DSB recession (Fig. I.16B). The RecF pathway is primarily involved in repair of SSG via HR (Morimatsu and Kowalczykowski, 2003). In this case, the RecQ 3'-5' helicase performs the RecBCD helicase function of unwinding the DNA (Umezu *et al.,* 1990). The 5'-end of unwound DNA becomes available for degradation by the RecJ 5′-3′ exonuclease, which produces a 3′ ssDNA overhang that is then protected by SSB proteins (Corrette-Bennett and Lovett, 1995). SSB inhibit RecA binding to ssDNA. Replacement of SSBs by RecA requires the RecFOR complex, which is recruited to the 5′-end of an ssDNA gap and promotes the binding of RecA onto the ssDNA (Morimatsu and Kowalczykowski, 2003). RecF recognizes SSB bound to ssDNA, and RecOR promotes displacement of SSB, allowing subsequent binding of RecA to ssDNA.

In both RecBCD and RecF pathways, RecA bound to ssDNA forms a presynaptic nucleoprotein filament complex that captures dsDNA and screens for homology that can be as few as 6 nucleotides (Ragunathan *et al.,* 2012; Lee and Wang, 2009). Upon finding a homology, RecA initiates strand invasion into dsDNA and promotes the formation of the heteroduplex, the so-called D-loop. However, D-loop formation can be inhibited by dismantling RecA form ssDNA by UrvD helicase (Veaute *et al.,* 2005), or by direct inhibition of RecA recombinase and coprotease activities by RecX (Stohl *et al.,* 2003).

After the establishment of the D-loop, several alternative pathways can be taken (Mawer and Leach, 2014): proceed with HR, form a replication fork by Break-Induced Replication (BIR) or switch to the Synthesis-Dependant Strand Annealing (SDSA) pathway, discussed in section 4.5.4. In the case of BIR, DNA synthesis of the invaded strand might proceed until it covers the entire chromosome. BIR was initially proposed as a process by which recombinant bacteriophages assimilate parental DNA (Meselson and Weigle, 1961). Later, studies in humans showed that BIR is the core mechanism for telomere maintenance in the absence of telomerase in tumours, to avoid cell death (Dunham *et al.,* 2000).

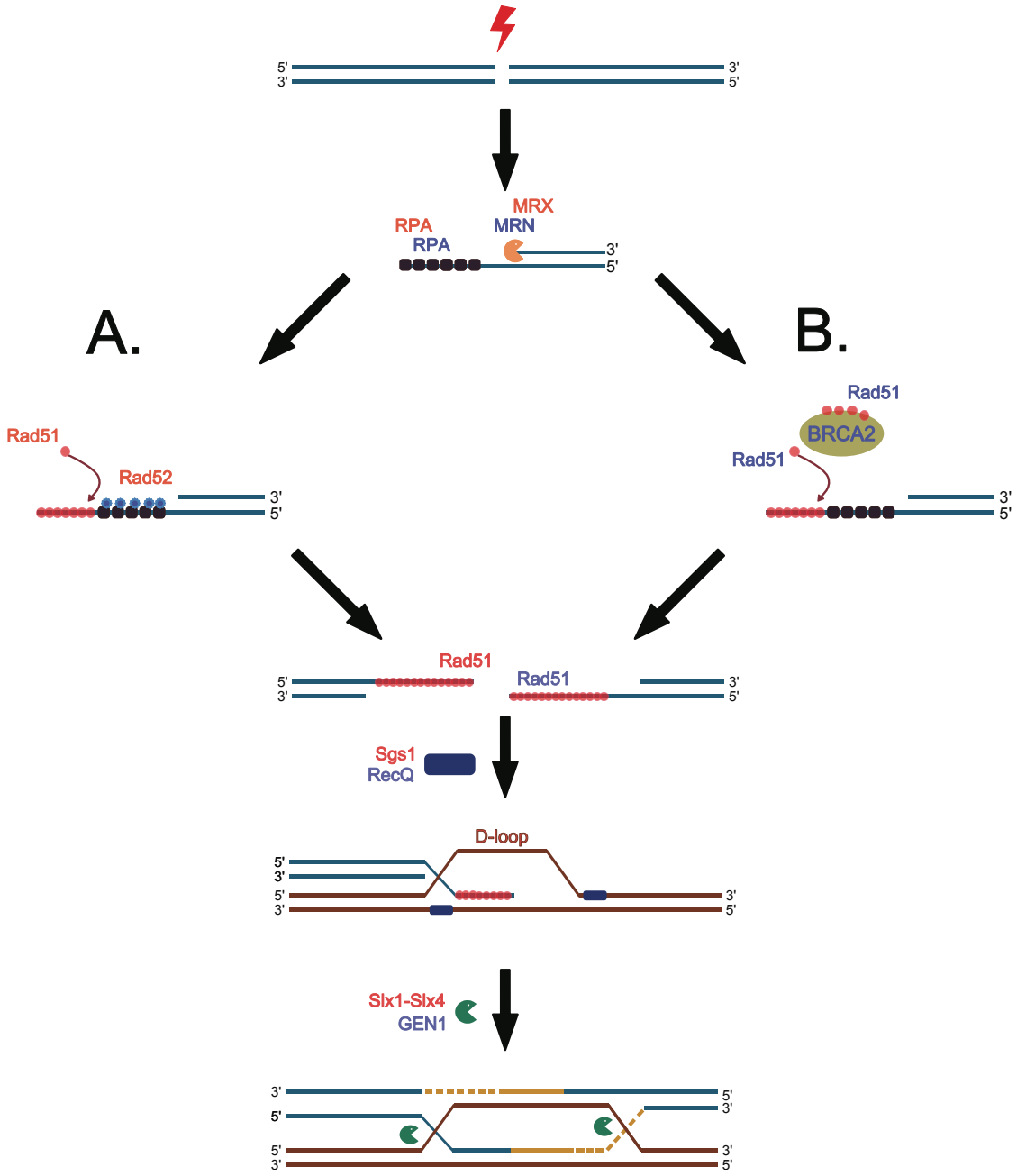


Figure I. 17. Simplified model of DSB repair by homologous recombination in yeast and human cells. DSB break repair starts with DNA recession by the MRN/MRX complex to form 3'-overhangs that are stabilised by RPA. In yeast (A), Rad51 loading on ssDNA is facilitated by prior Rad52 binding to RPA. In humans (B), the replacement of RPA by Rad51 is promoted by BRCA2. Rad51 forms a nucleofilament on ssDNA and screens dsDNA for homology sequences. Upon homology, Rad51 facilitates ssDNA invasion into dsDNA to form a D-loop. The proper ssDNA annealing and further extension are promoted by helicase Sqs1/RecQ. Recruitment of the second strand of invading DNA forms a four-stranded DNA structure called a Holliday Junction (HJ). After the synthesis of missing DNA sequences, HJs are resolved by Slx1-Slx4/Gen1 endonuclease to form either crossover (CO) or non-crossover (NCO) gene conversion products. Yeast factors are indicated in red, and human factors in blue. Figure adapted from Eckert-Boulet et al., 2011 and Ratanaphan, 2012.

During HR, the formed D loop must be extended for proper recombination of homologous sequences; otherwise, it will be aborted (Mawer and Leach, 2014). Two partially redundant ATP-dependant helicases were found in *E. coli* responsible for D-loop maturation: RadA/Sms (RecA paralog) and RecG. RecA can load RadA hexamers on the acceptor dsDNA at both ends of the D-loop (Marie *et al.,* 2017). The helicase activity of RadA extends the D-loop in both directions over long distances. Unlike RadA, RecG is recruited by SSB and acts as a monomer, to unwind receptor dsDNA for invasion and recombination of ssDNA (Whitby *et al.,* 1994). Furthermore, RecG allows the recruitment of the second strand of invading DNA, converting the triple-strand D-loop into a four-stranded DNA structure called a Holliday Junction (HJ) (Mawer and Leach, 2014). The HJ is stabilised by the RuvABC complex (West, 1997). The RuvA tetramer binds specifically to the HJ, stabilizing its structure and recruiting RuvB to both sides of the junction, which can translocate the HJ in both directions. Finally, the HJ is resolved by the RuvC endonuclease, allowing the separation of two DNA molecules. The outcome of this resolution can result in either crossover (CO) or non-crossover (NCO) products, depending on how RuvC cleaves the junction.

4.5.3.2. Homologous recombination in the nucleus of eukaryotes

The mechanism of HR in eukaryotes is more complex than in prokaryotes, but while the overall process is highly conserved, the specific proteins involved may vary (Fig. I.17). As described for MMR, in humans, DSBs are initially recessed by the MRN complex (MRE11, RAD50, and NBS1) to the 3' single-strand overhangs (Cannavo and Cejka, 2014). In yeast, the process is promoted by the MRX (Garcia et al., 2011). Stimulated by Sae2/CtIP, MRN/MRX subunit MRE11 starts 3′-5′ exonucleolytic digestion. Orthologs of all MRE complexes were also found in Arabidopsis (AtMRE11, AtRAD50, and AtNBS1) (Waterworth *et al.,* 2007; Šamanić *et al.,* 2016; Gallego and White, 2001). Mutations in any one of the corresponding genes lead to chromosomal rearrangements. However, for longer homology regions, further recession is performed by EXO1 (Nimonkar *et al.,* 2011). In eukaryotes, Rad51 (ortholog of bacterial RecA) screens for homology regions and mediates strand invasion. However, in humans, BRCA2 inhibits Rad51 ATPase activity and enhances Rad51 binding to ssDNA and nucleation (Carreira and Kowalczykowski, 2011; Ratanaphan, 2012). In yeast, it was shown that RPA is first bound by Rad52, which facilitates RPA replacement by Rad51. Furthermore, Rad52, similarly to Rad51, has the activity to promote the annealing of complementary ssDNA strands (Mortensen *et al.,* 1996). However, differently from yeast, several studies suggest that human Rad52 has lost the capacity to mediate RAD51 assembly on RPA–ssDNA complexes and suggest that second-end capture is rather facilitated by strand invasion than annealing (McIlwraith and West, 2008; Jensen *et al.,* 2010). Rad54 in all eukaryotes acts as a motor

Category	*E.coli*	*A.thaliana*	AGI	Targeting
DNA Exonuclease	RecBCD	OEX1 ?	At3g52050	Mt
DNA Plymerase	PolI	Pol1B	At3g20540	Mt + Cp
ssDNA-binding proteins / Recombination mediators	SSB	SSB1	At4g11060	Mt
		SSB2	At3g18580	Mt
		OSB1	At1g47720	Mt
		OSB3	At5g44785	Mt + Cp
		OSB4	At1g31010	Mt
		WHY2	At1g71260	Mt
		ODB1	At1g71310	Mt
DNA Recombinases	RecA	RECA2	At2g19490	Mt + Cp
		RECA3	At3g10140	Mt
DNA Helicases / DNA Translocases	RuvAB	-	-	-
	RecG	RECG1	At2g01440	Mt + Cp
	RadA	RADA	At5g50340	Mt + Cp
Recombination regulators	RecX	RECX	At3g13226	Mt + Cp
	UvrD	-	-	-
	MutH	MSH1	At3g24320	Mt + Cp
	MutS			
DNA Ligase	LigA	LIG1	At1g08130	Mt + Cp
Holliday junction resolvase	RuvC	-	-	-

Table I. 2. Comparison of bacteria and plant mitochondria factors involved in homologous recombination. Abbreviations: Mt, mitochondria; Cp, Chloroplast. Adapted from Gualberto and Newton, 2017.

protein that translocate along dsDNA and facilitates strand invasion (Swagemakers *et al.,* 1998; Mazin *et al.,* 2010). Yeast Sgs1 and human RecQ, orthologs of bacterial RecQ, ensure branch migration for perfect annealing and capture of the second strand (Lu and Davis, 2021; Ui *et al.,* 2007). Holiday junction is resolved by Slx1-Slx4 or GEN1, in yeast and humans, respectively (Rass *et al.,* 2010; Fricke and Brill, 2003). In Arabidopsis, two paralogs of GEN1 (AtGEN1 and AtSEND1) were shown to have HJ resolvase activity(Bauknecht and Kobbe, 2014). Furthermore, in humans, it was shown that Topoisomerase I can also catalyse HJ resolution (Hede *et al.,* 2007).

4.5.3.3. Homologous recombination in the plant mitochondria

Due to the endosymbiotic origin of mitochondria, many factors involved in HR in plants are orthologous to bacterial counterparts or have evolved their own proteins with functional similarity (Table. I.2) (Gualberto and Newton, 2017). However, no orthologs of eukaryotic MRN/MRX or of bacterial RecBCD complexes have been identified in plant mitochondria. Thus, the mechanism of DSB recession in plant mitochondria remains unknown. The produced ssDNA overhangs can be bound by numerous ssDNA binding proteins: SSBs, OSBs, ODB1, WHY2. Mutants deficient in SSBs and of OSBs showed an increased accumulation of recombination products (Qian *et al.,* 2022; Zaegel *et al.,* 2006), leading to major genome rearrangements, while loss of WHY2 and of OSB1 was shown to result in an accumulation of MHMR products (Cappadocia *et al.,* 2010; Janicka *et al.,* 2012). Mitochondrial SSB1 and SSB2 are orthologs of bacterial SSB. The difference between the two SSBs is unknown, yet it has been shown that SSB1 might act upstream OSB1 during DNA replication and downstream OSB1 during HR (Qian *et al.,* 2022). Furthermore, SSB1 and SSB2 do not seem to interfere with MHMR (García-Medel *et al.,* 2019).

The displacement of single-stranded DNA proteins for recombinase loading might be facilitated by WHY2 and ODB1. ODB1 is an apparent ortholog of yeast Rad52 and was confirmed to facilitate the hybridization of complementary strands *in vitro* (Janicka *et al.,* 2012). WHY2 and ODB1 are co-purified together and may act in the same way or be part of different pathways (Janicka et al., 2012; Gualberto et al., 2014). In Arabidopsis, there are two orthologs of bacterial RecA targeted to mitochondria, RECA2 and RECA3 (Miller-Messmer *et al.,* 2012; Shedge *et al.,* 2007). It is assumed that they form two independent pathways of branch migration. RECA2 is essential, and its mutation is lethal at the seedling stage. It is likely to participate in HR functions essential for proper mtDNA replication and segregation. In contrast, RECA3 mutants are viable, although they accumulate products of the ectopic recombination involving small repeats, correlating with phenotypes of reduced growth and deformed leaves

that worsen in following generations. RECA3 might be involved in HR surveillance by promoting the SDSA pathway during mtDNA repair. Notably, RECA3 evolved recently in the flowering plant lineage and is absent in mosses. It is characterized by the absence of an important acidic C-terminus that is found in RECA2 and the bacterial orthologs. In bacteria, this C-terminus has been shown to interact with SSB and with the RecX recombination suppressor (Drees *et al.,* 2004). An ortholog of RECX is present in all plants, targeted to both mitochondria and chloroplast, and in *Physcomitrium* it was shown to interact with both chloroplast RECA1 and mitochondrial RECA2 (Odahara and Sekine, 2018). Thus, in flowering plants such as Arabidopsis, the inhibition of HR by RECX might be evaded by RECA3, but not by RECA2 (Gualberto and Newton, 2017). Furthermore, RADA (the ortholog of bacterial RadA) was detected to interact with RECA2, but not RECA3, further supporting the hypothesis that there are two alternative plant mitochondrial HR pathways defined by RECA2 and RECA3 (Chevigny *et al.,* 2022).

Arabidopsis orthologs of two partially redundant bacterial helicases have been found in plant mitochondria: RECG1 and RADA. Both can complement their orthologs in *E.coli* and the corresponding mutants have severe effects on mtDNA stability (Wallet *et al.,* 2015; Chevigny *et al.,* 2022). However, *recG1* plants have normal development, whereas *radA* mutants are severely affected in development. Unlike in bacteria, RECG1 and RADA are not epistatic to each other, as there was no additive effect of double mutant *recG1 radA* compared to the single radA mutant. Finally, no orthologs of bacterial RuvABC have been found in plant organelles for HJ resolution. However, the ortholog of alternative bacterial resolvase YqgF has been found in plant mitochondria that may be involved in HJ, although it has not been experimentally tested yet (Chevigny et al*.,* 2020). Alternatively, as in humans, HJs could be resolved by branch migration factors and topoisomerase I, or RECG1 might resolve HJs by moving them to replication forks, as proposed for bacterial RecG (Wardrope *et al.,* 2009). More about factors involved in the HR of Arabidopsis mitochondria are discussed in Chapter III.

4.5.4. Synthesis-Dependant Strand Annealing (SDSA)

The SDSA pathway of HR does not involve the formation of a double holiday junction and exclusively produces NCO products (McMahill *et al.,* 2007) (Fig. I.16). This HR pathway mechanism was first proposed for *Drosophila melanogaster,* where crossover events were rare, if detected at all, and no need for Holliday junctions was identified (Nassif *et al.,* 1994). Later studies revealed that Drosophila maintains genome stability by downregulating HR via Rad51 removal by Srs2, which also unwinds D-loops, thus promoting SDSA (Meir *et al.,*

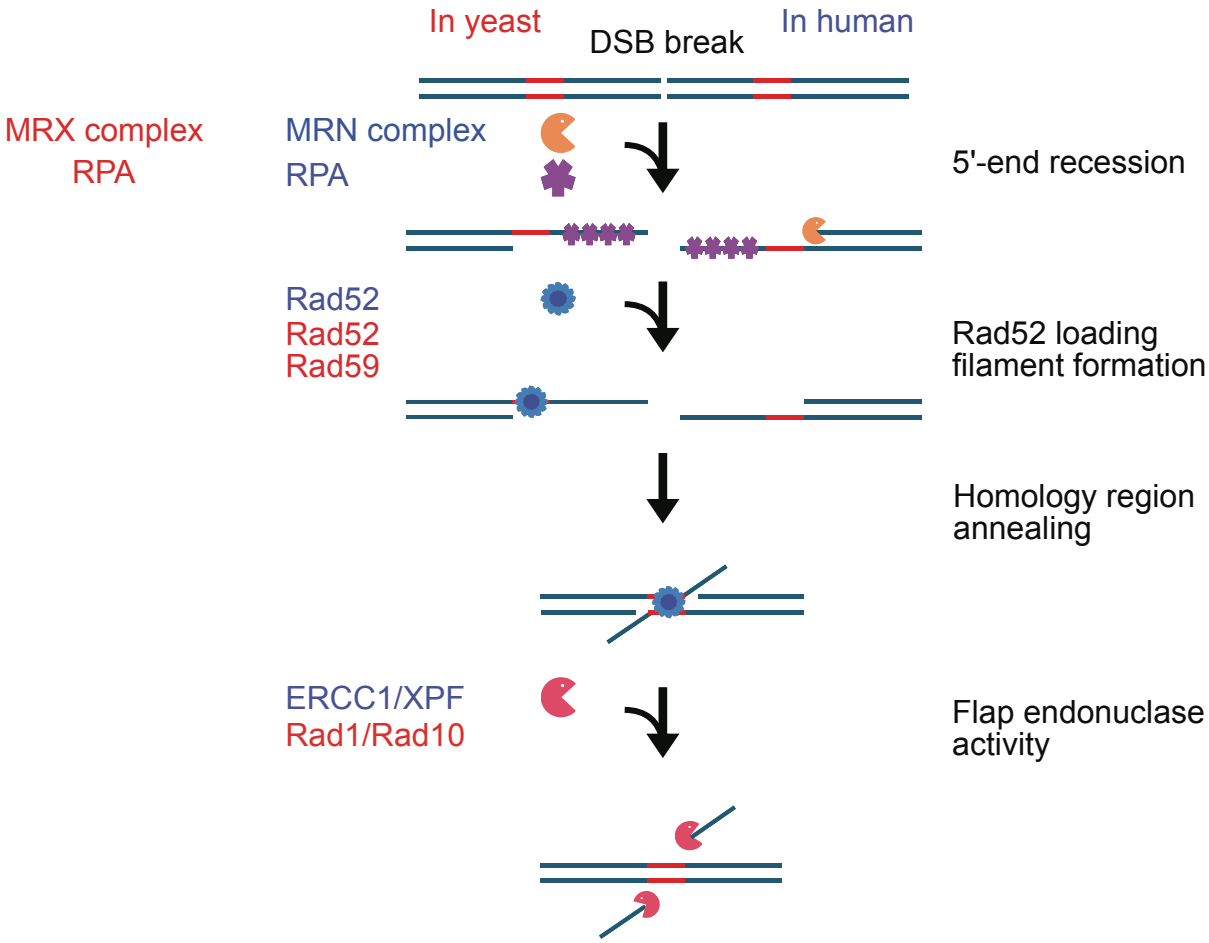


Figure I. 18. Mechanism of Single-Strand Annealing repair, in yeast and humans. It shows the key steps of repair and proteins involved in humans (blue) and in yeast (red) cells. Adapted from Bhargava et al., 2016; Haber, 2006; Hanamshet et al., 2016; Motycka et al., 2004.

2023). Furthermore, the loss of any of the helicases BLM, HELQ, or FANCM impairs gap repair and SDSA (Thomas *et al.,* 2022).

In the Arabidopsis nucleus, loss of RAD51 results in a dramatic deficiency in SDSA, suggesting its requirement for the pathway (Roth *et al.,* 2012). Furthermore, overexpression of RAD52-like ODB1 targeted to the nucleus leads to increased intra-chromosomal recombination and SDSA, but contrary to what is described for yeast Rad52, it repressed the SSA pathway (Samach *et al.,* 2018). Similarly to Srs2, MSH1 might also promote SDSA, as it has been shown to dismantle D-loops via endonuclease activity (Penafiel-Ayala et al., 2023). In the SDSA model, RAD51-coated single-stranded DNA invades a homologous template and forms a D-loop. DNA synthesis extends the 3′ end of the single-stranded DNA (Fig. I.16). However, unlike in conventional HR, the D-loop is then disassembled prior to strand annealing of the second ssDNA strand and formation of a HJ.

4.5.5. Single Strand Annealing (SSA)

The SSA pathway, like MHMR, are two HDR pathways that are RecA/Rad51 independent. In humans and yeast, Rad52 was shown to be the key player in SSA, by promoting the annealing of ssDNA regions of homology (Fig. I.18) (Shinohara *et al.,* 1998; Jensen *et al.,* 2010). When complementary regions are small, yeast rad59 has been shown to assist Rad52 with annealing (Wu *et al.,* 2006). After annealing, 3' overhangs are removed by the endonuclease activity of complexes such as ERCC1-XPF (in humans) and Rad10- Rad1 (in yeast). The remaining ends are then joined by DNA ligase (Bhargava *et al.,* 2016; Motycka *et al.,* 2004; Hanamshet *et al.,* 2016; Haber, 2006).

In plant mitochondria, as mentioned before, ODB1 and RECA2-RECA3 most likely perform the roles of Rad52 and Rad51, respectively. The endonuclease activity needed to process 3'overhangs could be carried out by OEX1, according to *in vitro* studies (Deborah, 2021). Loss of Rad52 results in defective SSA in yeast and animals (Stark *et al.,* 2004; Symington, 2002). On the contrary, overexpression of RAD52/ODB1 in the nucleus repressed the SSA pathway (Samach *et al.,* 2018). In the bacteria *Deinococcus radiodurans*, it was also shown that SSA is a RecA-independent repair pathway. Furthermore, the same study showed that DdrB, which is the *D. radiodurans* SSB, promotes SSA by an unknown mechanism (Ithurbide et al., 20 5). Thus, plants may have developed distinct SSA mechanism than those from bacteria, yeast and animals.

4.6. Direct repair

Some types of DNA damage can be directly repaired. For instance, UV-induced pyrimidine dimer can be directly repaired by photoreactivation, a process that restores the structure of the DNA helix without excising the modified nucleotides, catalysed by photolyases. Photolyases use blue light photons to reduce a flavin cofactor and revert bridging between the two pyrimidines (Brettel and Byrdin, 2010). However, photoreactivation in higher plant organelles remains controversial. While a photolyase has been detected in the mitochondria of Arabidopsis and rice, little or no repair of cyclobutyl pyrimidine dimers was observed (Chen *et al.,* 1996; Takahashi *et al.,* 2011b). Photolyases are absent in humans, but they have Fe(II)/α-ketoglutarate-dependent dioxygenases (ALKBH) and Methyl Guanine Methyl Transferase (MGMT) that can repair a variety of alkylation adducts. However, ALKBH and MGMT have not been identified in plants (Nay and O'Connor, 2013).

4.7. Damage tolerance

As a last resort, when DNA damages cannot be repaired through conventional mechanisms, cells have developed several mechanisms to cope with the damage. These can be translation synthesis (TLS) or strand displacement synthesis. TLS is conserved from bacteria to mammals and executed by an array of specialized DNA polymerases, most of which belong to the Y family, along with the B family enzyme Pol ζ. The Y-family polymerases possess a replicative polymerase domain with a large active site, capable of accommodating bulky adducts, and have fingers and thumb domains that detect the fit of the nucleotides. However, these domains are typically shorter and less efficient compared to those of replicative polymerases. Consequently, Y-family polymerases are error-prone, with incorporation error rates of up to 1 in 10 (McCulloch and Kunkel, 2008). Many have specific features to bypass the lesion. For instance, DNA polymerase η acts as a molecular splint and constrains the DNA lesion to maintain base pairing despite the distortion (Johnson *et al.,* 1999). However, due to their low fidelity, these polymerases must be tightly regulated. In prokaryotes, their concentration is regulated by the SOS response (Friedberg *et al.,* 1995). In vertebrates, Proliferating Cell Nuclear Antigen (PCNA) is a key regulator, and its ubiquitylation recruits polymerase to stalled replication forks (Moldovan *et al.,* 2007).

Pol ζ is responsible for most mutagenesis in yeast and is unique to eukaryotes. It lacks proofreading exonuclease activity and, *in vivo*, is responsible for most insertions from mismatched termini and those containing abnormal structures (Prakash *et al.,* 2005). In the Arabidopsis nucleus, DNA polymerase ζ (REV7) and DNA polymerase η (POLH) bypass UV- induced damage by TLS, preventing replication arrest and allowing damaged cells to survive and grow (Sakamoto, 2019). In mitochondria, both PolI1A and PolI1B can perform high-fidelity bypass of AP lesions, moderate 3′-5′ exonuclease activity and show strong evidence of strand displacement activity (Baruch-Torres and Brieba, 2017). During strand displacement, the 3' ssDNA is rejected from the stalled replication fork and is reannealed to a new complementary dsDNA, to resume synthesis (Plaza-G.A. *et al.,* 2023).

5. Aim of the thesis

Advancements in genome engineering constantly provide new tools for further and more sophisticated research in the field of genetics and, more broadly, in the field of biology. Currently, there are three major gene editing techniques: zinc finger nucleases (ZFNs) (Urnov *et al.,* 2010), transcription activator-like effector nucleases (TALENs)(Christian *et al.,* 2010), and clustered regularly interspaced short palindromic repeats (CRISPR) (Gilbert *et al.,* 2013). In this study, I exploited the possibility of mitoTALEN technology for the research of mtDNA repair and segregation in the plant model *Arabidopsis thaliana*. Specifically, I used the MitoTALENs to explore the role of several factors known to be involved in mtDNA Homologous Recombination. To pursue this aim, previously studied mutants, including *odb1* (Samach *et al.,* 2011; Janicka *et al.,* 2012), *why2*(Krause *et al.,* 2005), *osb1osb4* (Zaegel *et al.,* 2006), *radA* (Chevigny *et al.,* 2022), *msh1*(Virdi *et al.,* 2015), *pol1A and pol1B* (Parent *et al.,* 2011), *recA3* (Davila *et al.,* 2011; Miller-Messmer *et al.,* 2012) and *recG1* (Wallet *et al.,* 2015) were transformed with mitoTALEN constructs to induce DSB in the mtDNA , allowing for subsequent analysis of the mitochondrial mtDNA repair and segregation processes.

My aim was to analyse how a DSB is repaired in mutants of key proteins involved in HR, compared to the WT. First, my project started with the transformation of Arabidopsis HR mutants with the mitoTALEN constructs that were prepared and kindly provided by Prof. Arimura (University of Tokyo). To analyse the resulting mutants, I had to optimise mtDNA enrichment from individual plants for subsequent Illumina sequencing. This allowed me to generate a dataset of Illumina sequencing data for 90 individual T1 plants.

Several available NGS data analysis tools were used to analyse the data, and a few new ones had to be developed. NGS analysis was divided into three main stages. First, a rapid analysis was conducted with MacVector mapping tools. This involved generating a coverage plot of the reads mapping to the mitochondrial genome, to determine the fingerprints of DSB repair and manually evaluate which repeated sequences were active in the repair pathway taken. Second, a dedicated Bioinformatic pipeline was developed to analyse read- pairs mapping more than 1000 pb apart on the mtDNA. This approach aimed to identify active repeats and estimate the relative frequency of their involvement in HR-dependent repair. Finally, an NGS analysis pipeline was developed to identify MHMR. This comprehensive approach enabled to distinguish between three distinct DSB repair pathways: HR, SSA, and MHMR.

NGS sequencing is an expensive tool for analysing multiple samples. Thus, to enhance the confidence in the NGS-based findings, a qPCR approach was developed to

increase the number of individual mutants analysed. This qPCR approach was based on the DNA repair fingerprints identified in the NGS dataset. All of this allowed me to assess what repair pathways are taken in the absence of a specific protein and deduce or re-confirm its possible function in mtDNA repair.

As an unexpected and serendipitous result, a link between a mitoTALEN-induced phenotype and the integrity of the *atp9* gene was established. It was observed that several plants transformed with our mitoTALEN constructs exhibited a phenotype characterized by reduced growth and wrinkled leaves. This phenotype was correlated to a deficiency in the *atp9* gene copy number and transcript abundance. Thus, my present work led to the identification of a viable Arabidopsis line with a deficiency in Oxidative Phosphorylation (OXPHOS) due to reduced *atp9* expression. This line can be further exploited for future research on retrograde signalling pathways.

Finally, I conducted a pilot study to develop a novel genome engineering technique. The approach involved directing DNA repair by simultaneously introducing two DSBs in the mtDNA. These DSBs were strategically positioned to create relevant single-stranded DNA (ssDNA) overhangs that contribute to the general pool of recombination intermediates. The aim was to leverage the natural mitochondrial repair mechanisms and stimulate the formation of recombination intermediates that could potentially lead to long-range modifications of the mtDNA. Such an approach of inducing genetic variability in mitochondria could be of potential biotechnological interest.

CHAPTER II: MitoTALENs, a new method of mtDNA editing

Introduction

1. Genome engineering

Genome editing is a type of genetic engineering in which DNA is inserted, deleted, modified or replaced in a site-specific manner. It potentially allows the specific inactivation of a gene, the introduction of specific mutations, correct mutations, and even the insertion of a new gene. The technology is experiencing rapid growth as new tools continue to emerge. Zinc-Finger Nucleases (ZFN), Transcription Activator-Like Effector Nucleases (TALEN) and CRISPR/Cas-based technologies, whose specificity can be altered relatively easily, become the most attractive tools for genome engineering and keep evolving fast. Meganucleases were initially discovered to target a single sequence over 12 bp long. Despite the finding that mutation in a few amino acids could alter nucleotide specify, the intimate connection between the recognition and cleavage elements makes them challenging to use in research, and currently it is hard to find recent publications on using meganucleases for genome editing (Ashworth *et al.*, 2006; Paques and Duchateau, 2007).

1.1. Zinc-Finger Nucleases (ZFN).

Zinc-finger nucleases are fusion proteins of a zinc-finger DNA-binding domain and the Fok1 endonuclease DNA-cleavage domain. Fok1 is a type IIS restriction endonuclease, cleaving DNA at a defined distance from its non-palindromic asymmetric recognition site. Zinc-finger is composed of a set of Cys2His2 (3 to 6) that recognise 9-18 bp, respectively (Kim *et al.*, 1996). However, modulating the specificity of a zinc-finger had limited success, as zinc-finger specificity can change in an assembly with another zinc-finger. The most successful method of changing zinc-finger specificity was generating a pool of tree zinc-finger assemblies and screening them for desired recognition using a bacterial two-hybrid system. Upon DNA binding, a selectable marker is expressed (Maeder et al., 2008; Ramirez et al., 2008). ZFNs have been widely used for plant genome engineering in Arabidopsis, tobacco, maize, etc (Cai et al., 2009; Marton et al., 2010; Shukla et al., 2009). However, ZFNs typically have poor activity and high toxicity due to dependency on neighboring fingers (Cornu *et al.*, 2008). Multiple platforms have been created to study ZFN dependency and to produce Arabidopsis mutants: “Sangamo Bioscience”(Doyon *et al.*, 2008), “OPEN”(Maeder et al., 2008), and “CoDA”(Sander *et al.*, 2011).

1.2. Transcription Activator-Like Effector Nucleases (TALEN)

Interest in TALE domains appeared just two decades ago. It was discovered in the *Xanthomonas* plant pathogen, showing that specific repeat-variable residues (RVD) bind the corresponding nucleotide in the target site. In nature, TAL effectors bind to the host genes' promoter and activate the transcription of genes according to the bacteria's needs. It was shown that nucleotide binding specificity is given by positions 12 and 13 of each repeat (Boch et al., 2009; Moscou and Bogdanove, 2009). This breakthrough quickly led to the creation of TALEN technology for DNA editing, fusing TALE DNA-binding motifs with the Fok1 nuclease domain (Christian et al., 2010; Li et al., 2011a; Mahfouz et al., 2011). Furthermore, the TALE domain was found to be less toxic and more accessible to engineering compared to ZFN. However, due to the highly repetitive sequences, the constructions are challenging to create by PCR. The problem was overcome with Golden Gate technology, which uses Type IIS restriction endonucleases, creating unique four bp overhangs next to their specific recognition site (Engler *et al.,* 2009). This way, RVD flanked with unique cohesive ends could be designed for ordered single-reaction assembly into an array, forming a specific TALE-binding domain (Cermak et al., 2011; Li et al., 2011b; Morbitzer et al., 2011). In a 2017 review paper, more than 50 genes had already been targeted by that time across multiple plant species, including Arabidopsis, barley, *Brachypodium*, maize, tobacco, rice, soybean, tomato, and wheat (Malzahn *et al.,* 2017).

1.3. CRISPR/Cas9

Two years later, following the discovery of TALENs, another genome editing tool, CRISPR-Cas9, was introduced. Initially, CRISPR (Clustered regularly interspaced palindromic repeats) was found to be an adaptive immune system in bacteria and archaea, serving as a defence mechanism against invading viruses. Cas9 proteins paired with transcribed viral RNA fragments make DSB in an invading viral DNA (Makarova *et al.,* 2017; Wiedenheft *et al.,* 2012). The CRISPR-Cas9 system has two main elements: a guide RNA (sgRNA) that brings target specificity and a Cas9 protein. Cas9 is an endonuclease containing RuvC and HNH domains that cut the target and complementary strands, respectively (Jinek *et al.,* 2012). In 2012, Cas9 of *Streptococcus pyogenes* was paired with a synthetic single guide RNA (gRNA) to make double-strand breaks (DSB) *in vitro* in *Escherichia coli* (Jinek *et al.,* 2012). CRISPR-Cas9 system outperforms previously presented ZNF and TALEN systems, due to its high specificity, low production cost, simple design and ability to edit multiple genes simultaneously (Rao *et al.,* 2022). Thus, currently, it is the first choice when considering a gene editing tool for nuclear sequences.

1.4. Mitochondria Genome Engineering

Till now, no reliable method has been established for plant mitochondria transformation. The similarities between the plastid genome and the mitochondrial one suggested that the mitochondrial genome might be amenable to transformation. However, currently, a reliable method of mitochondrial transformation has only been described in yeast (Fox et al., 1988; Johnston et al., 1988; Zhou et al., 2010) and green algae (Bonnefoy et al., 2007; Randolph-Anderson et al., 1993; Remacle et al., 2006), by biolistic transformation. There are still no reported gene editing systems for transforming mitochondria in plants.

Although plant mitochondria remains out of scope for transformation, recent advancements in genetics suggest that adapted tools may be developed soon. Currently, the previously presented, frequently used CRISPR/Cas9 system remains under development for use in mitochondria, as there is no reliable method for guide RNA delivery to mitochondria, apart from a few reports. Multiple studies in humans showed the successful delivery of Cas9 protein to mitochondria, by the addition of an N-terminal targeting sequence. However, few studies detected mitochondrial localisation of guide RNAs, and showed the system to have a global effect on mtDNA copy number rather than specific mtDNA editing. The main difficulty is the delivery of long guide RNA to mitochondria, as known RNA import determinants (MLS) have shown no success (Nikitchina *et al.,* 2024). Similarly, a recent report targeting the *mtatp9* CMS gene by CRISPR/Cas9 in Tobacco showed controversial success. They produced male- sterile mutants with aborted stamen, yet only 30 % of the mtDNA was modified (Chang *et al.,* 2023). Furthermore, there is evidence of an RNA editing approach using *trans*-cleaving hammerhead ribozymes on mitochondrial genes in Arabidopsis, although with partial success. The hammerhead ribozymes were used as RNA cargo sequences attached to the PKTLS mitochondria import shuttle, which allowed the specific knockdown of mitochondrial RNAs. Although there was no evidence of RNA import, a partial reduction of the target mitochondrial gene transcript was observed, triggering a complex nuclear retrograde response (Niazi *et al.,* 2019).

Recent advancements in using TALE-based tools brought significant progress towards targeted mtDNA modification. The assembly of TALE repeats ensures recognition of the target site, and the construct can be deployed from nuclear genes and targeted to mitochondria by an N-terminal targeting sequence. Most certainly, TALEN is a better option than ZNF, as the TALE domain has a higher specificity and is not toxic, and most importantly, it does not require guide RNA as the CRISPR/Cas9 system. Both mitoZNF (ZNF targeted to mitochondria)

(Gammage *et al.,* 2014; Minczuk *et al.,* 2008) and mitoTALEN (TALEN targeted to mitochondria) (Bacman *et al.,* 2013; Hashimoto *et al.,* 2015) demonstrated efficacy in modifying the mitochondrial genome, introducing point mutations or large-scale deletions, as well as induce selective degradation of mutated mitotypes in cybrid cells and mouse models. Both mitoZNF and mitoTALEN were shown to be compatible with the fusion to double-stranded DNA deaminase (DddA)-derived cytosine base editors (DdCBEs), to edit DNA without the introduction of DSBs, in mice and human cells. This method has proven to be more reliable for repairing point mutation, comparing to the use of FOK1 endonuclease, which cannot induce or revert a specific mutation. Furthermore, the repair of introduced DSBs in the mammalian mtDNA is not as efficient as in the nucleus (Cho et al., 2022; Gammage et al., 2018; Lee et al., 2021; Silva-Pinheiro et al., 2022).

Recently, several publications presented the successful use of mitoTALEN to induce specific DSBs in plant mitogenomes (Arimura et al., 2020). The approach has proven most effective in validating predicted CMS-responsible genes. Indeed, multiple genes potentially responsible for CMS have been described previously, in variable plant species, yet none had been proved by gene knockout untill recently. Thus, the knockout of genes *orf79* in rice and *orf125* in rapeseed, using MitoTALENs, was shown to restore male fertility (Kazama *et al.,* 2019). A mitoTALEN was also used to produce stable knockout lines of the mitochondrial *nad9* gene (Forner et al., 2023). Forner's group also showed that mitoTALENs can be used to introduce point mutations in the mtDNA, using (TALEN) gene-drive mutagenesis (GDM), or TALEN-GDM in *Nicotiana tabacum*. As previously shown, repair mechanisms in plant mitochondria are error-prone and are based on homology recombination using an intact genome copy as a template. Since recombination is linked to DNA polymerisation, intercalating agents like ethidium bromide (EtBr) can enhance nucleotide misincorporation, with a preference for the TALEN binding site, thereby reducing its binding capacity. In Arabidopsis, TALE with bound adenine deaminase (TALED) attained more precise point mutation editing, achieving homoplasmic heritable mutations within a single plant generation (Zhou *et al.,* 2023).

2. Challenges of sequencing the mitochondrial genomes

Sequencing mitochondrial DNA is a challenging process in some species more than in others, because in total DNA extracts the amount of mtDNA is rather small compared to nuclear and chloroplast DNA. Although plant mitochondrial genomes (200 – 11,000 kbp) are much bigger than the mammalian mtDNA (16.5 kbp), it's much smaller than nuclear DNA (*A. thaliana* nuclear genome 157 Mb)(Morley and Nielsen, 2017). Furthermore, in green

tissues of *A. thaliana,* there are 80 to 120 chloroplasts per mesophyll cell (cpDNA size 155 kb, 10-500 genomes per 1 chloroplast) in comparison to 300 mitochondria (mtDNA size 368 kb, 1 genome per 3 mitochondria)(Crumpton-Taylor *et al.,* 2012; Welchen *et al.,* 2014). However, the number of plastid genome copies varies greatly during mesophyll cell development, from 80 plastome copies in meristematic cells to 2600-3300 copies in mature diploid mesophyll cells (Greiner *et al.,* 2020). On average, qPCR analysis of cpDNA in a total DNA extract has a 5 Cp value difference compared to the mtDNA, meaning about 64 times higher copy number of the cpDNA as compared to the mtDNA. Thus, Illumina 150 bp pair-end sequencing of a total DNA extract of diploid *A. thaliana* will yield approximately 9 nuclear and 27 plastid reads for every mitochondrial read. However, in this project, individual 5-6 week-old plants have been sequenced, and it is known that in Arabidopsis the nuclear DNA undergoes endoreplication – an increase in gene content without cell division. It has been shown that at 3 weeks, Arabidopsis leaf cells may have equal amounts of nuclei with 2, 4, 8 and 16 copies of the genome (del Pozo and Ramirez-Parra, 2015; Hendrix *et al.,* 2018). Consequently, if we assume that plants would have, on average, 8 copies of the nuclear genome due to endoreplication, sequencing should yield about 36 nuclear and 27 plastid reads for every mitochondrial read. It would require 157,000 reads to reach a 1X mtDNA coverage, from a total DNA extract. This is for the relatively small nuclear genome of Arabidopsis. In species with much larger nuclear genomes (ex. hexaploid wheat, 16.9 Gb) the problem is exacerbated. It can be overcome either by increasing coverage by sequencing more Illumina reads, or by enrichment of the mtDNA by purification of mitochondria previous to DNA extraction, by differential centrifugation followed (or not) by density gradient purification.

3. Aims of this chapter

The primary aim of this chapter of my thesis was to exploit mitoTALENs to analyse mitochondrial genome repair, in Arabidopsis. I present the multiple potential repair pathways observed in WT plants (accession Col-0) transformed with two different mitoTALEN constructs, along with the methods for DNA repair analysis through NGS sequencing and qPCR. Additionally, I try to understand how different mitotypes resulting from mitoTALEN-induced genome editing are maintained and segregated.

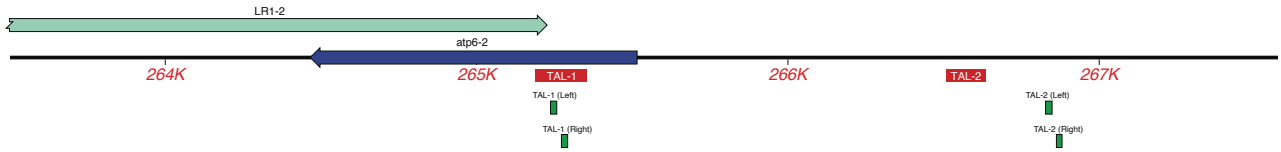


Figure II.1. mitoTALEN binding sites to mitochondrial genome of A. thaliana Col-0. A DSB is made between Left and Right TALEN proteins, via FokI dimerization. Indicated in green are the two binding sites (Left and Right) for the two mitoTALEN constructs: TAL-1 and TAL-2. The regions in red indicate qPCR amplicons used to check the copy number of the target sites.

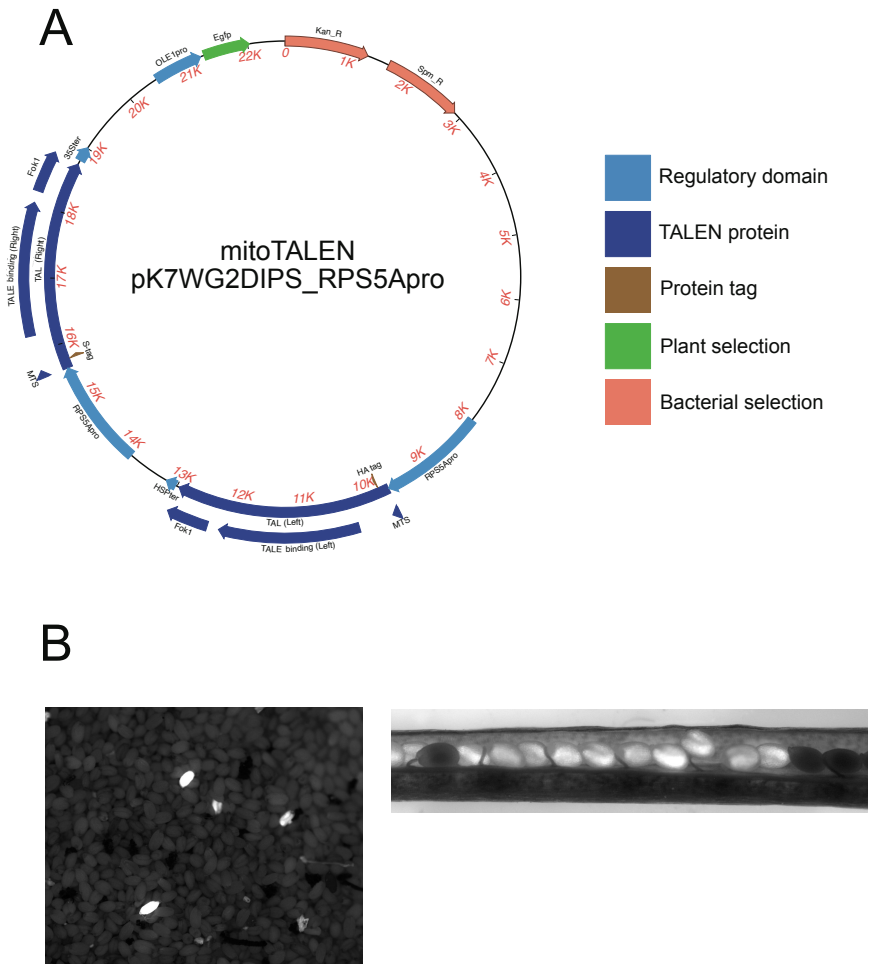


Figure II.2. (A) mitoTALEN constructs in binary vectors. Different TALEN domains are indicated: FokI endonuclease, TALE binding domain, and mitochondria targeting sequence (MTS). The genes are under control of the constitutive RPS5 promoter. The T-DNA constructions contain the marker eGFP gene expressed under the OLE1 oleosin promoter on the seed coat. This allows the rapid selection of fluorescent T1 transformed seeds and the assumption of the heterozygous or homozygous state of mitoTALENs in T2 generation. **(B) Florescent stereo microscope (Zeiss Axiozoom) image of T1 transformed seeds and of developing T2 seeds in silique from segregating mitoTALEN plant.**

Results

1. DNA repair activation by mitoTALEN

To study DNA repair in mitochondria, two mitoTALENs were designed to make a specific DSB in non-essential regions of the mtDNA, allowing exclusive study of repair (Fig. II.1). TAL-1 targets a copy of the *atp6* gene (*atp6-2*), whose loss should not have any significant impact on mitochondrial function, as there is another functional copy of *atp6* (*atp6-1*) in the genome of Arabidopsis Col-0. Furthermore, the target site is surrounded by multiple repeats that could be used for homology-based repair pathways. The TAL-2 construct targets a non-coding region further downstream of the second copy of Large repeat 1 (LR1-2). This target site does not have annotated repeats in proximity; thus, it is a good tool for studying DSB repair by Microhomology-Mediated Recombination (MHMR).

These two constructions were assembled and kindly provided by Prof. Arimura (Tokyo University). In my project, I used the previously published mitoTALEN constructs that were used to show that the two copies of *atp6* are redundant (Arimura *et al.,* 2020). Thus, TAL-1 and TAL-2 correspond to the constructions annotated as "ATP6-2 and ATP6-2UP" in (Arimura *et al.,* 2020), respectively. Prof. Arimura kindly provided the constructs cloned in binary vector pK7WG2 (Fig. II.2A), with Spectinomycin selection for bacterial transformation and GFP under expression of the Oleosin *OLE1* promoter for selection of transformed seeds.

1.1. Generation of T1 mitoTALEN lines in WT Col-0 background

The mitoTALEN constructs were amplified in *E. coli* TOP10 cells and transformed into Agrobacterium for plant transformation by floral dip. Seeds from transformed plants were selected by GFP fluorescence expressed in the seed coat, visible under fluorescent Stereo Microscope (Zeiss Axiozoom) (Fig. II.2B). Selected seeds were either used for DNA extraction by the CTAB protocol or sown for growth. Seeds selected to be grown *in* vitro, in petri dishes, were sterilised beforehand.

1.2. qPCR analysis of T1 mitoTALEN seeds

DNA extracted from T1 seedlings was used for qPCR analysis to identify changes in the copy number of the MitoTALEN target sites. DNA copy number next to the cut sites was normalised to a distant locus of the mtDNA, upstream repeat G (Fig. II.1 & Fig. II.3), and compared to WT Col-0. There was a near complete loss of the mitoTALEN target sites in seedlings, suggesting that DSBs were repaired early in development. To identify when a DSB is repaired, DNA from T1 seeds was extracted and used for the same qPCR analysis (Fig. II.4). The DNA sequence at the target site was lost in mature seeds as in seedlings, suggesting that



Figure II.3. qPCR analysis of mitoTALEN cut sites in TAL-1 and TAL-2 mutant T1 seedlings (n=45). The relative copy number of the mtDNA near the TAL-1 and TAL-2 cut sites is normalized to the copy number of a distant mtDNA sequence (cox2 gene).



Figure II.4. qPCR analysis of mitoTALEN cut sites in TAL-1 and TAL-2 mutant T1 mature seeds (n=150). The relative copy number of the mtDNA near the TAL-1 and TAL-2 cut sites is normalized to the copy number of a distant mtDNA sequence.

	in 40 mL Beckman JA25.5 tubes				in 2 mL Eppendorf	
no.	1st step	2nd step	3rd step	4th step	5th step	7th step
1	5 min 2,500 g	xxx	xxx	20 min 15,000 g	xxx	20 min 15,000 g
2	10 min 1,000 g	xxx	xxx	20 min 15,000 g	xxx	20 min 15,000 g
3	5 min 1,000 g	5 min 2,500 g	5 min 4,000 g	20 min 15,000 g	10 min 1,000 g	20 min 15,000 g
4	10 min 1,000 g	10 min 2,500 g	xxx	20 min 15,000 g	xxx	20 min 15,000 g

Table II.1. Optimisation of differential centrifugation protocol for mitochondria enrichment. The whole plant rosette is ground with mortar and pestle, and the crude mitochondrial pellet is purified by differential centrifugation. After the 4th step, the tissue pellet is transferred from 40 mL polycarbonate tubes to 2 mL Eppendorf.

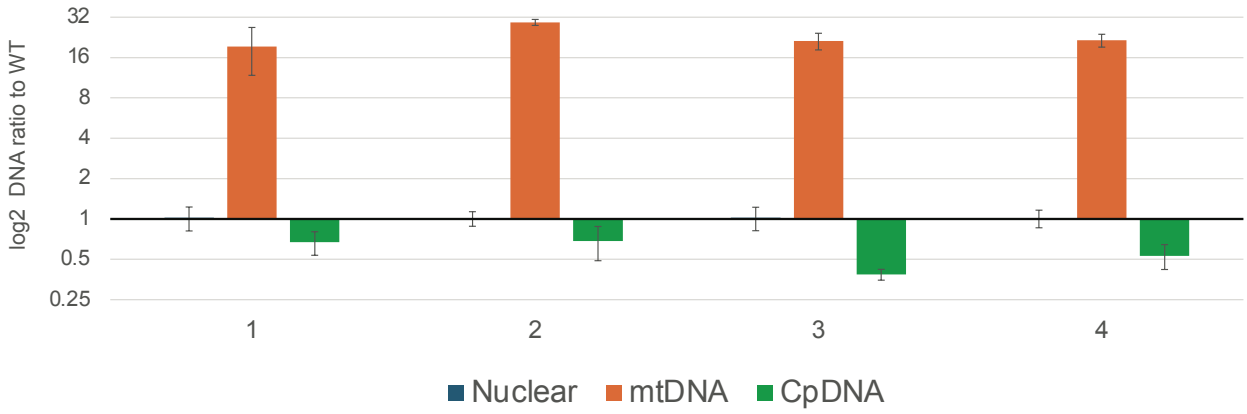


Figure II. 5. qPCR analysis of mtDNA enrichment optimization. Four protocols were tested, and all achieved an mtDNA enrichment of about 20-fold. Results were normalised to nuclear DNA.

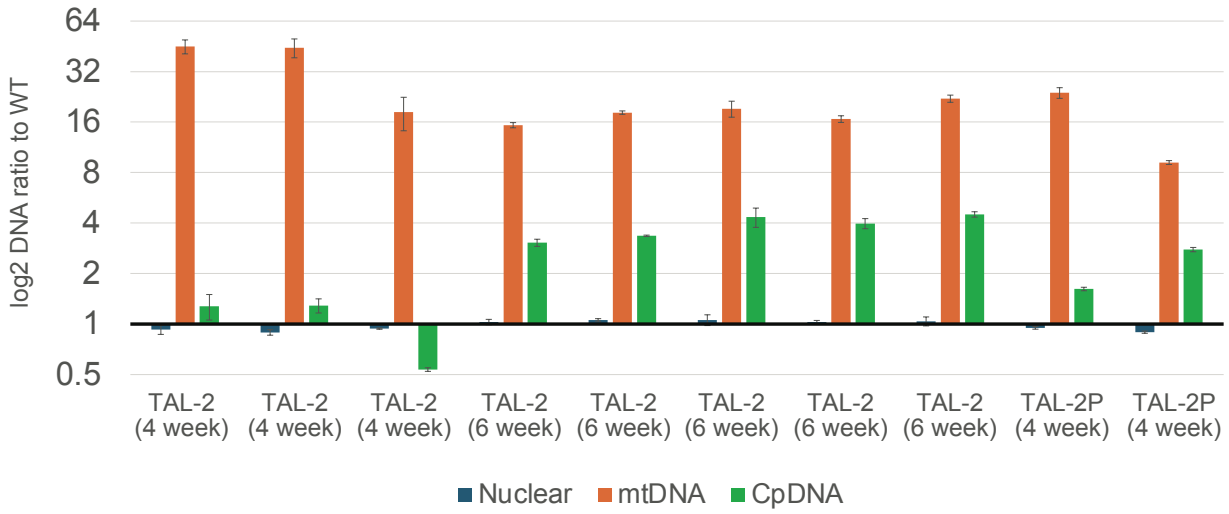


Figure II.6. qPCR analysis of mtDNA enrichment for multiple plants using no.4 differential centrifugation protocol (Fig II.5 & Table 1). Results were normalised to the nuclear DNA. There was increased chloroplast DNA contamination in 6-week old plants as compared to 4-week old plants.

the mitoTALENs are already very active early in seed development. Furthermore, both TAL-1 and TAL-2 cut sites were affected in plants transformed with either of the constructs, suggesting that repair of the DSBs resulted in loss of large regions of the mtDNA.

2. Mapping Col-0 mitoTALEN Illumina reads to the mtDNA

The primary approach of the study was to sequence the mtDNA of mitoTALEN lines and, according to the resulting DNA fingerprints next to the cut site, deduce how the mtDNA was repaired. To do so, I performed Illumina DNA sequencing on mitoTALEN samples partially enriched in mitochondria.

2.1. Partial mitochondrial DNA enrichment

To reduce sequencing costs and increase the coverage of sequenced mtDNA, we chose to extract DNA from plant extracts after partial mitochondrial enrichment. To do so, I adapted and optimised a routine in-house protocol for *A. thaliana* mitochondria purification. Because purifying mitochondria from plants using Percoll density gradients requires a substantial amount of plant material, we opted for just partial mitochondria enrichment by omitting the Percoll gradient purification step. In addition, performing purification on Percoll gradients was not an option, as we decided to study individual T1 plants, which yielded too little material. There were several reasons why I selected to sequence T1 individual plants. First, T1 plants are the first generation after mitoTALEN activity, and different T1 plants could have employed different repair pathways to repair DSBs. Secondly, individual plants needed to be studied, as minor mitotypes would have been challenging to detect in a pool of plants. Pooling plants would dilute minor alternative outcomes of DSB repair. Finally, sequencing the segregating T2 progeny would introduce a bias, due to the maternal inheritance of the mtDNA.

Partial mitochondria enrichment was achieved by differential centrifugation, starting from homogenised whole rosettes extracts of four- to six-week-old T1 plants. Several centrifugation conditions were tested (Table II.1). To assess the mtDNA enrichment, qPCR relative quantification of the mtDNA, cpDNA and nuclear DNA was performed, and the copy number of the organellar genomes normalised to the nuclear DNA. With all centrifugation conditions tested, an mtDNA enrichment of about 22-fold was achieved (Fig. II.5). Condition n° 4 was used for mtDNA enrichment in all samples of the experiments described in Sections 2 and 3. However, the achieved enrichment significantly varied from sample to sample (5-fold to 25-fold). On average, all samples had a 15-fold enrichment of the mtDNA (Appendix 2). Furthermore, due to the many samples that were processed in the experiment, some samples were prepared from plants that were up to two weeks older, and showed poorer mtDNA enrichment, with an increase of up to 4 fold in cpDNA (Fig. II.6). Further detailed discussion about the phenotypes developed by some mitoTALEN plants can be found in Chapter IV. In

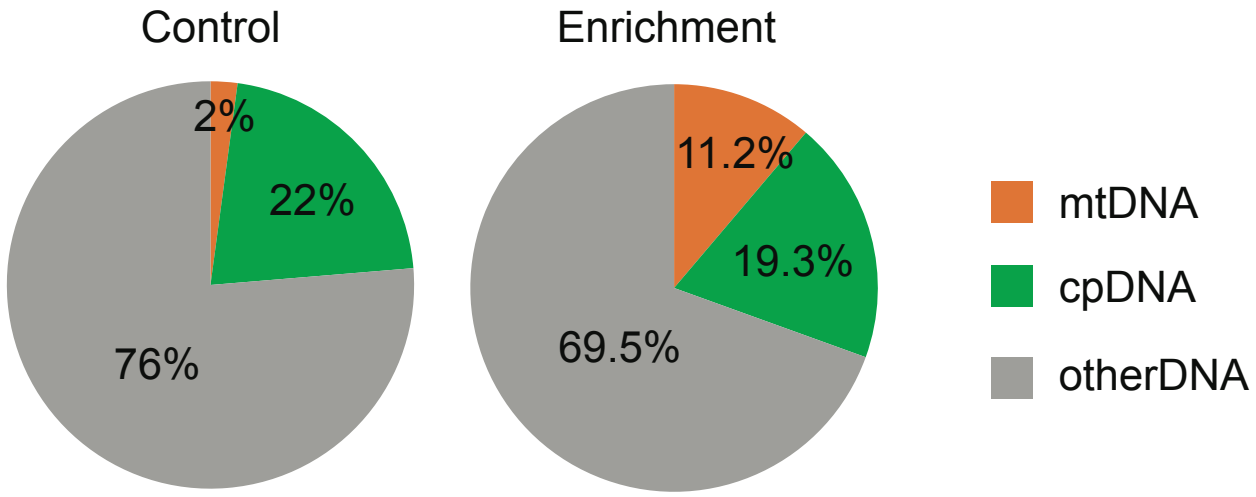


Figure II.7. NGS mapping analysis of mtDNA enrichment, according to the proportion of Illumina reads mapping to the mtDNA and to the cpDNA, of libraries prepared from mitochondria enriched fractions following the no.4 differential centrifugation protocol, compared to libraries prepared from total genomic DNA.

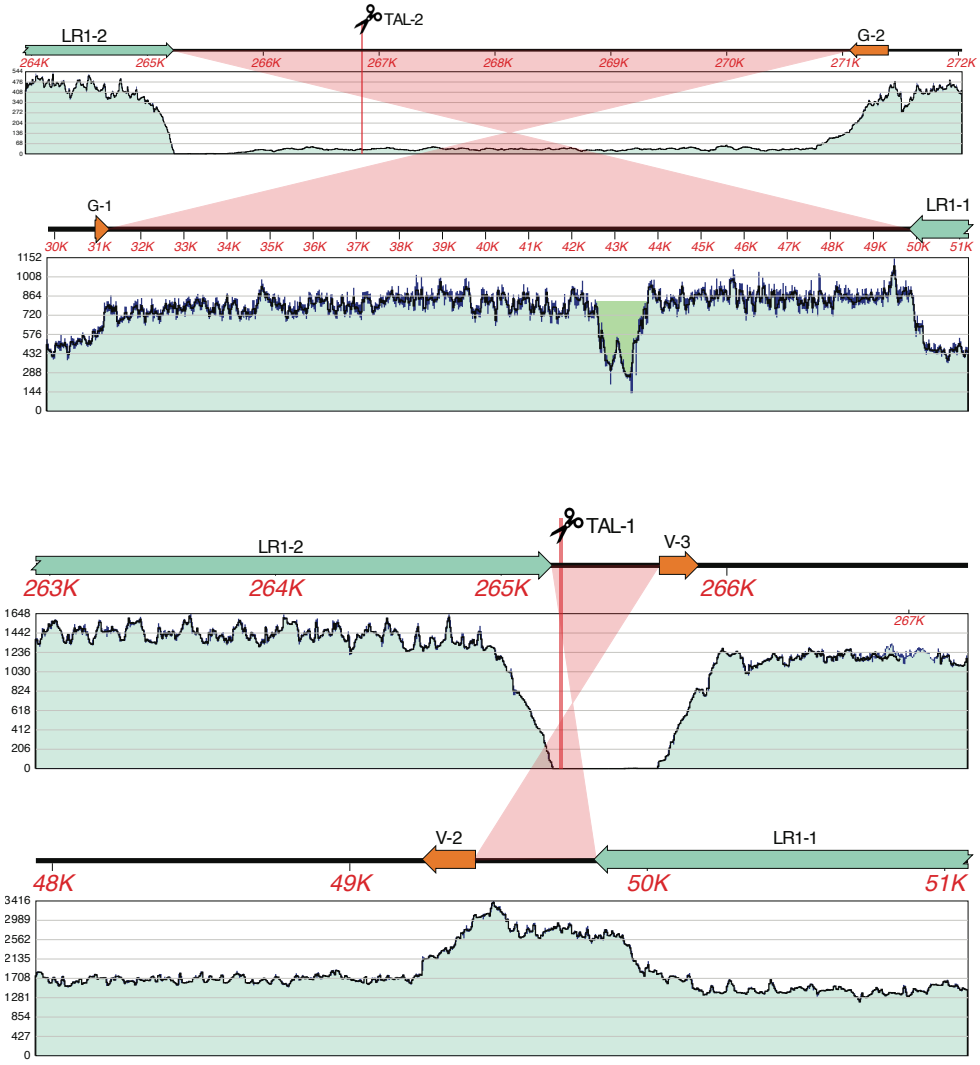


Figure II.8. Coverage plots of Col-0 TAL-2 sequencing reads (upper panel) and Col-0 TAL-1 sequencing reads (lower panel) mapped to the Col-0 mitochondrial genome. Produced by the Bowtie2 aligner implemented in Macvector Assembler. The red boxes represent the TALEN cut site region being replaced by the region below, located far apart in the mtDNA (see genome coordinates). These are the regions between large repeat LR1 and small repeat G (upper panel, for TAL-2) and between LR1 and small repeat V (lower panel, for TAL-1). The area in green represents reads removed by mapping to the chloroplast genome, corresponding to chloroplast sequence insertions in the mtDNA.

theory, with 15-fold mtDNA enrichment and 4-fold cpDNA enrichment, sequencing should result in that for every sequenced mtDNA reads, there should be approximately 2.5 nuclear and 1.8 plastid reads.

All samples discussed in Chapter II and Chapter III and presented in Appendix 2 have been sequenced by Illumina pair-end (2x150 bases) either on a Miseq or on a NextSeq 2000 sequencer. The libraries were prepared following Nextera™ DNA Flex Library Prep protocol (Illumina, San Diego, California, United States). The IBMP platform 'Gene expression analysis' managed the sequencing, with samples being sequenced either in-house on a MiSeq, or by sending them to the Genomax sequencing platform, for NextSeq 2000 sequencing. Two NextSeq 2000 sequencing runs were performed, multiplexing 50 and 60 libraries, respectively. The average library size was 320 bp, achieving an average of 900,000 reads mapping to the Arabidopsis mtDNA. On average, each library yielded 11.3% of mitochondrial reads and 19.6% of chloroplast reads (Fig. II.7). Consequently, there were approximately 6 nuclear and 2 plastid reads for every mitochondrial read. A more detailed library analysis can be found in Appendix 2.

2.2. Mapping to the mtDNA

After receiving sequencing reads, they were first analysed using the MacVector Assembler package application, utilizing the inbuilt bowtie2 mapping tool. First, all reads were mapped to the chloroplast genome to remove them from downstream analysis. This step is particularly important because several sequences of cpDNA origin are present in the mtDNA of Arabidopsis. Secondly, unmapped reads were mapped to the mitochondrial genome. From the coverage profiles, it was possible to infer the various mechanisms used to repair the mitoTALEN-induced DSBs, in WT Col-0 plants.

2.2.1. DSB repair by Homologous Recombination (HR)

Previous studies showed that homologous recombination is the major repair pathway in plant mitochondria. Coverage plots of reads mapping to the mtDNA showed that most Col-0 sequenced plants had a significant coverage drop at the mitoTALEN cut sites, flanked by copy two of Large repeat 1 (LR1-2, 6.6 kb) at the left border and either repeat G (335 bp) (in both TAL-1 and TAL-2 lines) or repeat V (170 bp) (in TAL-1 lines) at the right border (Fig. II.8). Simultaneous with the loss of these regions, a distinct region of the genome exhibited a significant increase in coverage, located distant from the TAL-1/TAL-2 cut sites. This region was also flanked by the same repeats (G or V at the left border and LR1-1 at the right border). Considering the equivalent amplitude of coverage changes along with the involvement of the same pairs of repeats, it suggested that the region with increased coverage was duplicated, replacing the region containing the cut site that was lost. This duplication likely occurred

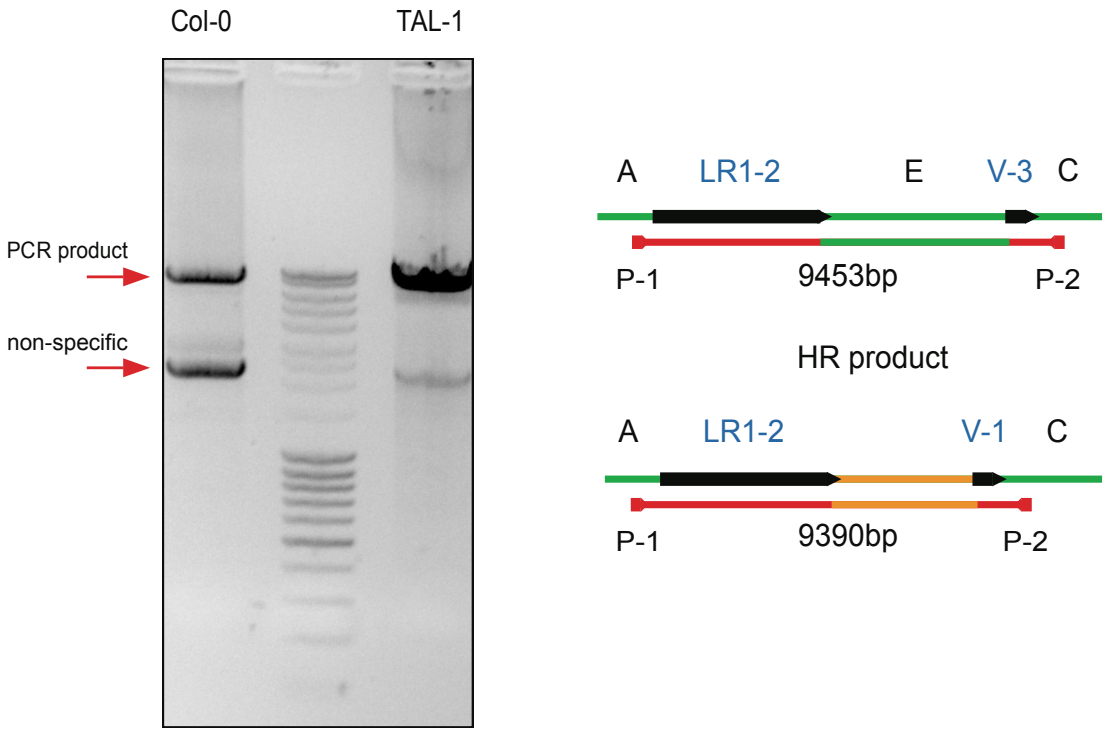


Figure II. 9. Two homologous recombination events involving Large repeat 1 (LR1) and repeat V, respectively, result in the replacement of region E by another region located in a distant locus of the genome, which becomes duplicated. Electrophoresis analysis on 0.7 % agarose gel (on the left) of PCR products and schematics (on the right) of PCR amplification of region in yellow and red, using the same P-1 and P-2 primers, yet producing different products in TAL-1 sample compared to WT Col-0. The model was confirmed by Sanger sequence of the PCR fragments (Fig II.10).

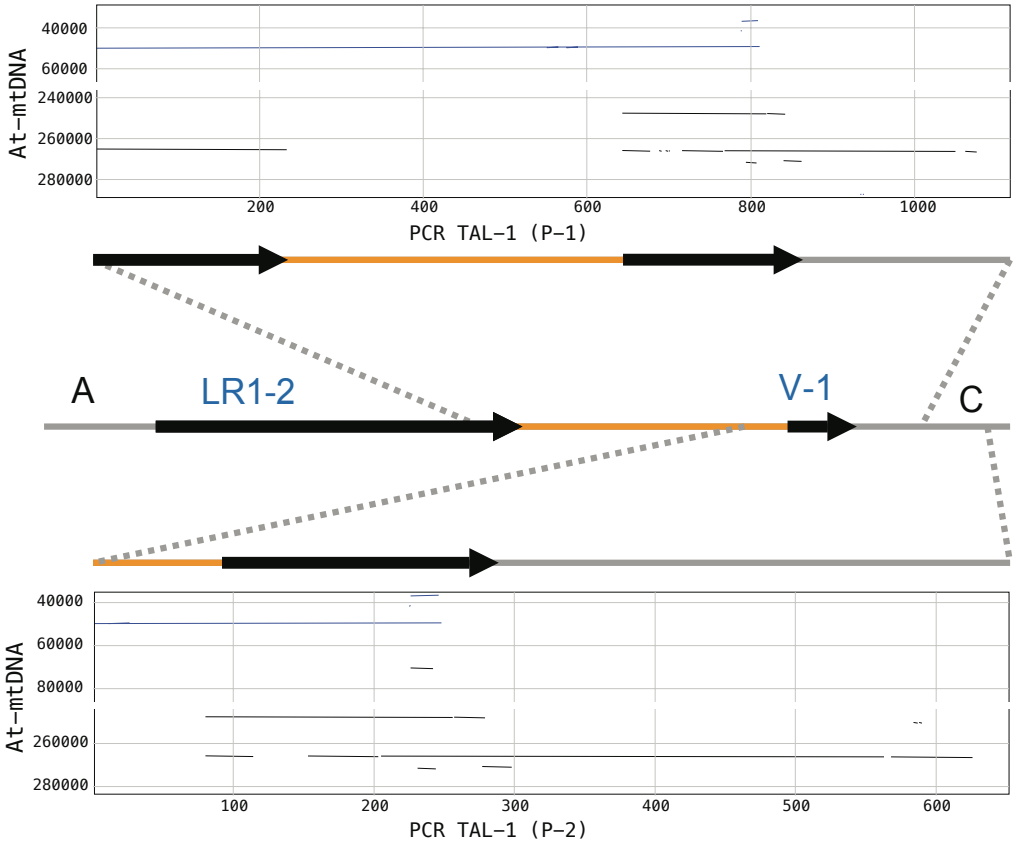


Figure II.10. Sanger sequencing of a PCR fragment resulting from homologous recombination involving repeats LR1 and V in a Col-0 TAL-1 Arabidopsis line. A dot plot alignment of the sequenced PCR fragment compared to the WT Col-0 mtDNA was created with the MacVector Pustel DNA Matrix function.

through a process of homologous recombination involving the two pairs of repeats. To prove this hypothesis, the region flanked by the repeats was amplified by long-range PCR, and the resulting fragment was purified from gel and sequenced. As expected, agarose gel electrophoresis was not resolutive enough to reveal the expected change in size, of just 63 bp for a fragment as big as 9453 bp (Fig.II.9). However, Sanger sequencing confirmed that the amplified fragment in TAL-1 had the duplicated region between repeat LR1 and repeat V (Fig.II.10). The same analysis could not be performed for recombination involving repeat LR1 and repeat G, as the resulting fragment would be too large for PCR amplification (26 kb). The putative new mitotypes resulting from this process were manually constructed and used as a reference (Fig.II.11). No significant changes in coverage were found when reads were mapped to the new mitochondrial genome, further supporting that our interpretation is correct (Fig.II.12). From these observations we could propose a model of how the DSBs were repaired, by recombination involving two pairs of repeats; This model is presented in the discussion (Fig II. 40)

2.2.2. DSB repair by Single-Strand Annealing (SSA)

Analysis of NGS sequencing data revealed coverage fingerprint for DSB repair by SSA, via the pair of direct repeats Y-2 and Y-3 that are located just 1.4 kb apart in the mtDNA. In one of the sequenced Col-0 TAL-1 plants, there was a drop in coverage flanked by the Y repeats, while the rest of mtDNA was in stochiometric balance (Fig. II. 13). I hypothesised that the region between the repeats was lost via SSA. To show it, first, I created the proposed mitotype to use it as reference to re-map the mitochondrial reads (Fig. II. 13). No changes in coverage were observed in the resulting genome, supporting that the hypothesis is correct. Secondly, I amplified the cut site region by PCR, and as expected the resulting PCR fragment was reduced in size in the mitoTALEN mutant compared to the WT plant (Fig. II. 14). Sequencing of the PCR fragments confirmed that in the mitoTALEN mutant, the region comprised between the two copies of the Y repeat is lost (Fig. II. 15).

2.2.3. DSB repair by Microhomology-Mediated Repair (MHMR)

In one of the sequenced Col-0 TAL-1 mutant lines, there was a drop in sequence coverage around the DSB, with no link to any annotated repeated sequence (Fig. II. 16). Closer analysis of the reads at these sites revealed microhomologies that most likely were used in microhomology-mediated repair (Fig. II. 17). For a comprehensive analysis of MHMR events, a pipeline for analysis of soft-clips in NGS data was developed, with the assistance of David Pflieger from the IBMP Bioinformatic platform. Some reads incompletely align to the reference genome, and an unaligned end of a read is referred to as a soft-clip. There are different reasons why the end of the read can lose a perfect match. Errors are rare and would

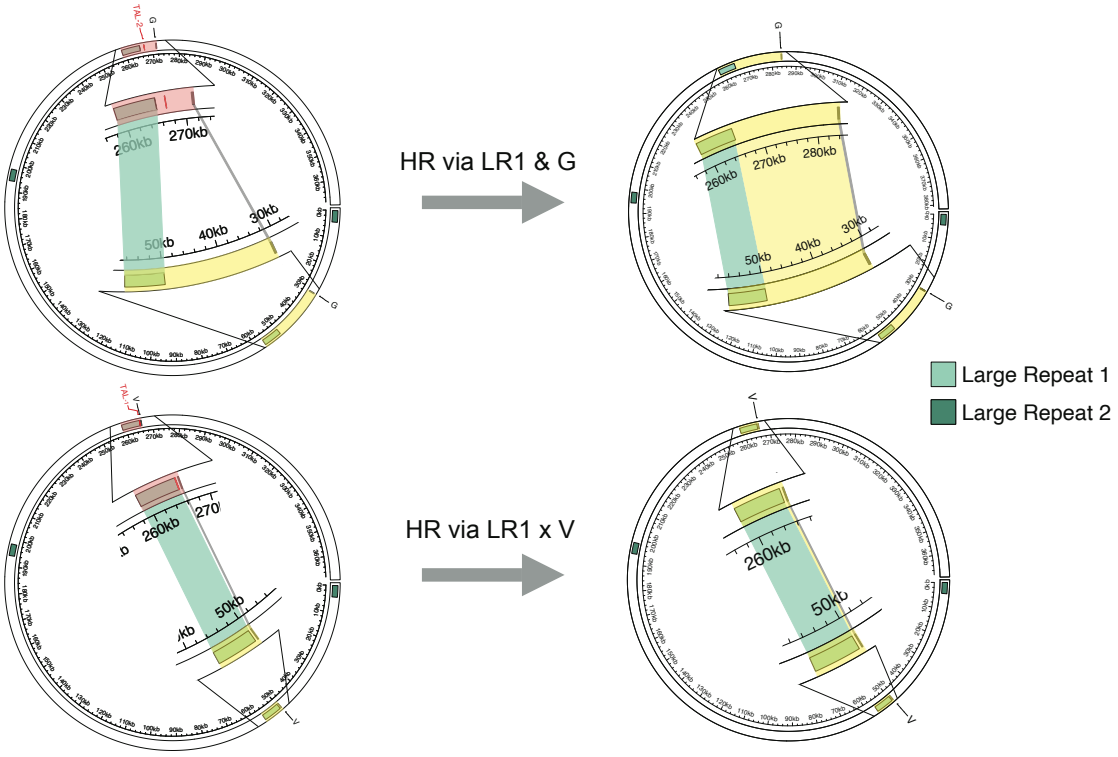


Figure II.11. Construction of proposed mitotypes in Col-0 after homologous recombination involving two pairs of repeats: Large repeat LR1 and small repeat G (upper panel) or Large repeat LR1 and small repeat V (lower panel). After repair by homologous recombination, the region in red containing the target site for the mitoTALEN is replaced by the region in yellow.

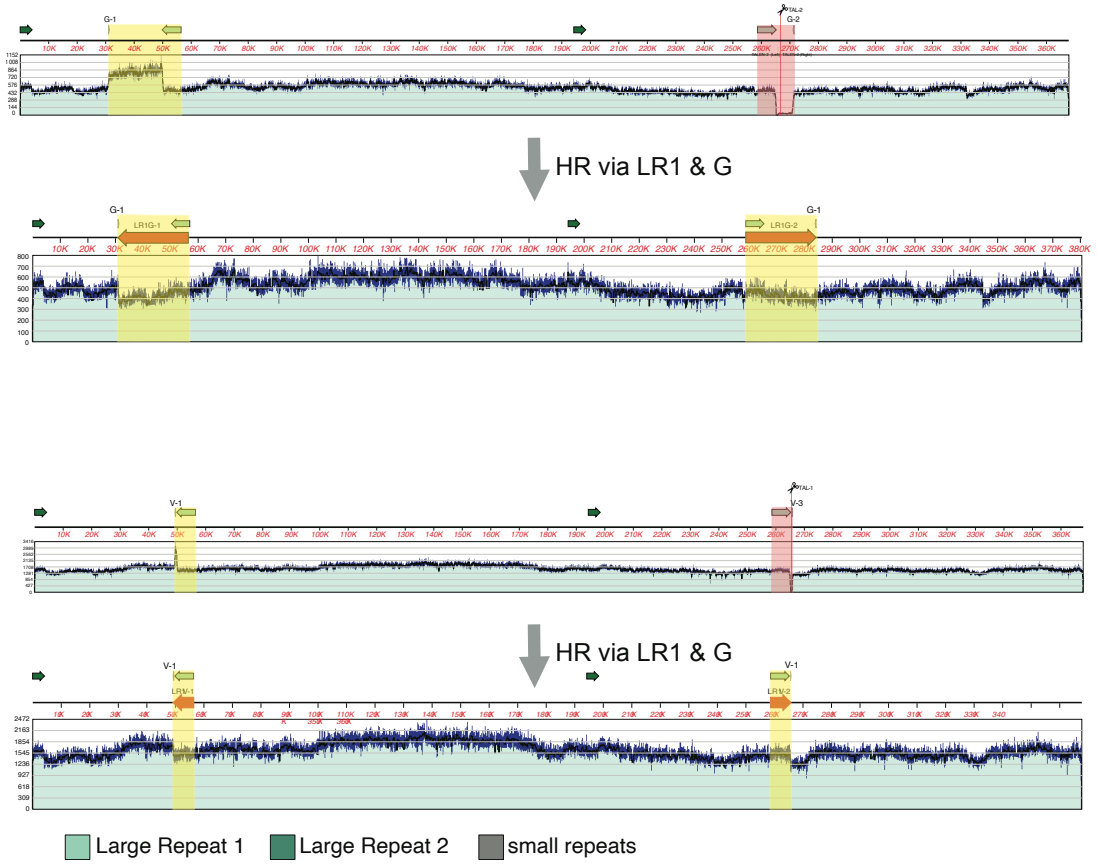


Figure II.12. Mapping of mitoTALEN lines to proposed new mitotypes that result from DSB repair via homologous recombination. Coverage plots of Col-0 TAL-2 (upper panel) and Col-0 TAL-1 (lower panel) Illumina sequencing reads, mapped to the WT Col-0 mtDNA and to the new mitotypes (as represented in Fig. II.11). The regions in red containing the target sites for the mitoTALEN are replaced by the region in yellow. Produced by Bowtie2 aligner, implemented in the MacVector Assembler package.

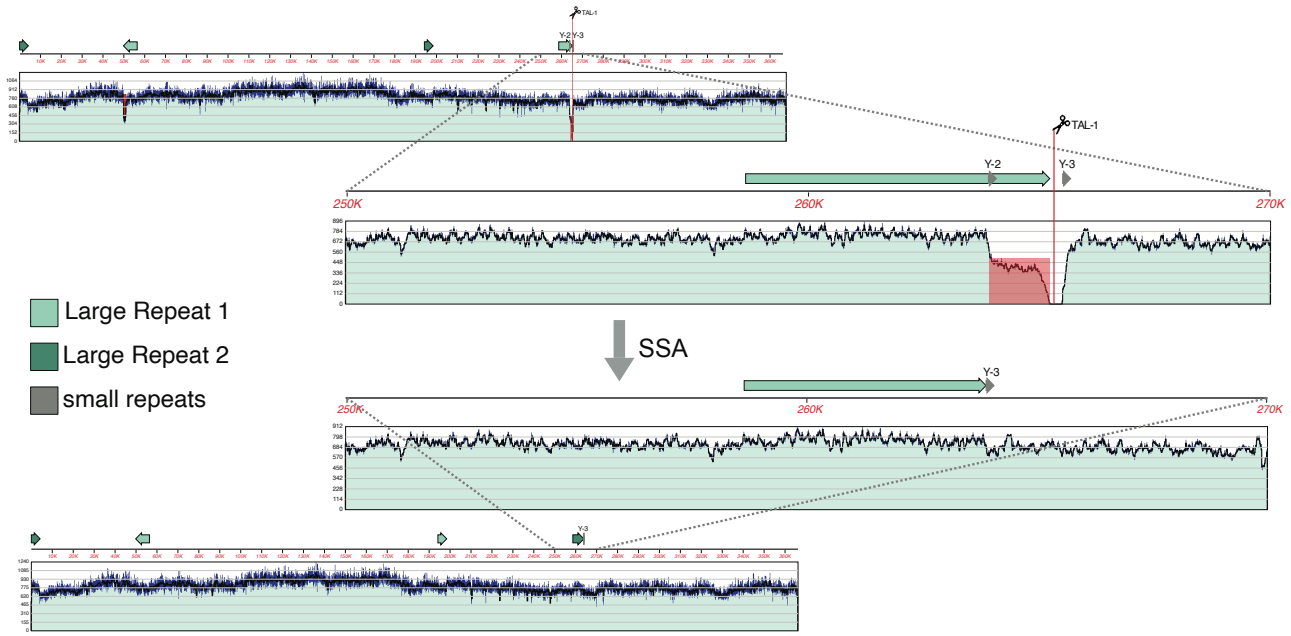


Figure II.13. Mapping of mitoTALEN lines to proposed new mitotype that results from DSB repair by Single-Strand Annealing (SSA). Coverage plot of Col-0 TAL-1 sequencing reads, mapped to the WT Col-0 mtDNA (upper panel) and to the proposed new mitotype, respectively (resulting in deletion of the region between Y repeats). The region in red between repeats Y is lost via SSA repair. The drop in coverage (in red, around 265 kb region) is due to multi-mapping of reads, resulting from an equal distribution of reads between duplicated mapping sites, although one of the sites is lost by SSA. Produced by Bowtie2 aligner implemented in the MacVector Assembler package.

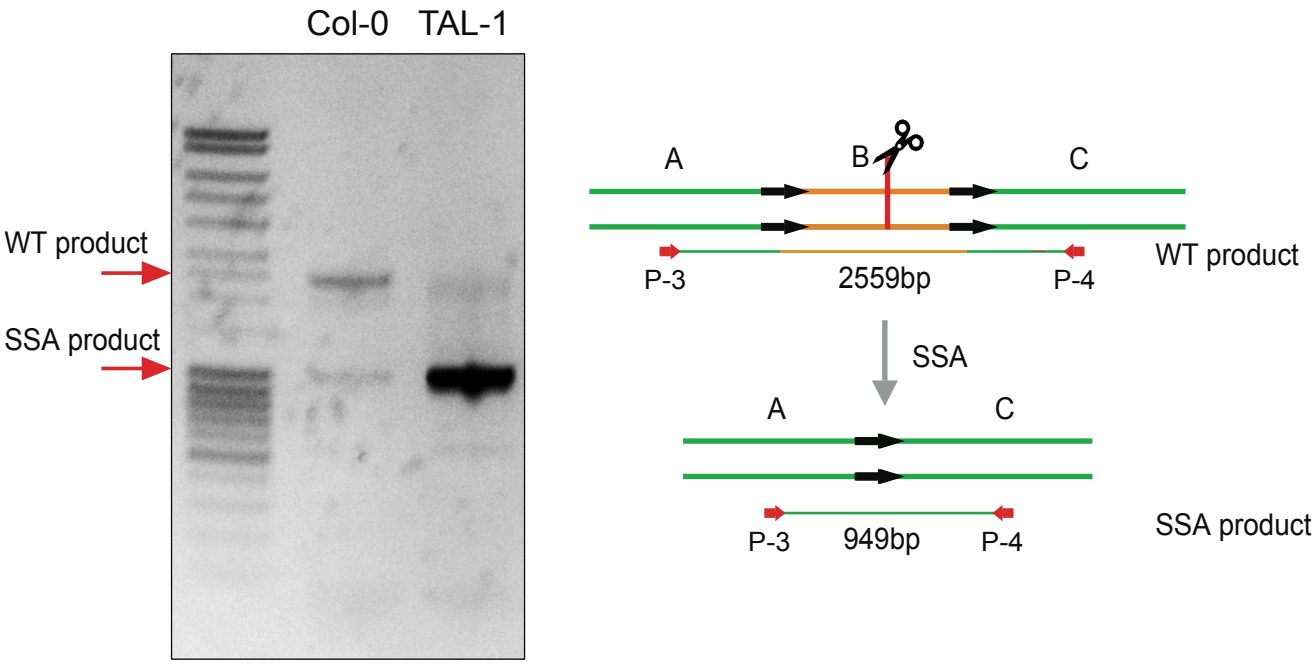


Figure II.14. Single-Strand Annealing involving repeat Y results in deletion of region B, containing the TAL-1 induced DSB. 0.7 % agarose gel electrophoresis analysis (on the left) and schematics (on the right) of PCR amplification of the corresponding regions, using P-3 and P□4 primers, from TAL-1 line and from WT Col-0. Confirmed by the Sanger sequence (Fig II.15).

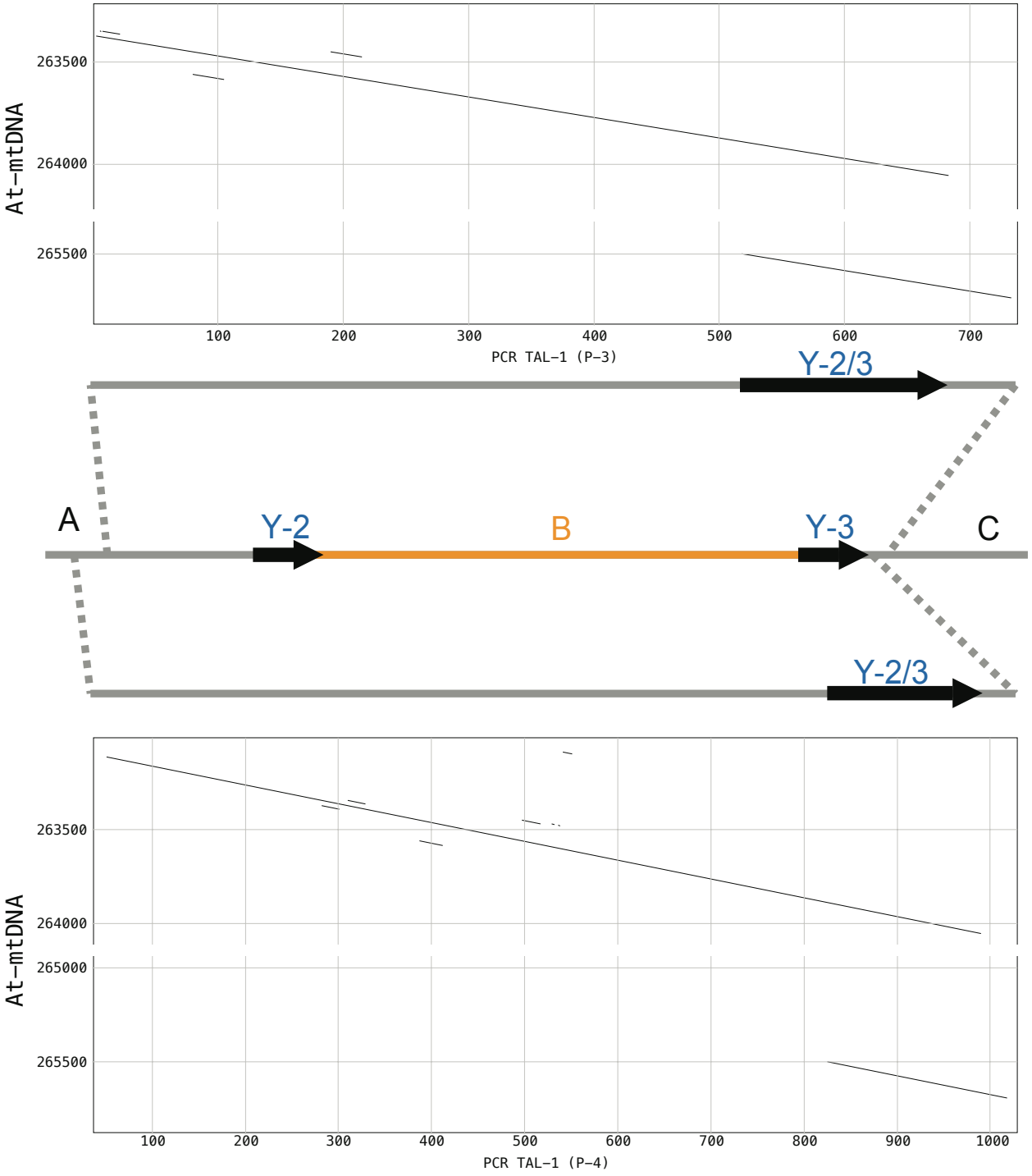


Figure II.15. Sanger sequence of PCR fragment resulting from Single-Strand Annealing between repeats Y, in Col-0 TAL-1 Arabidopsis line. A dot-plot alignment was created by the MacVector Pustel DNA Matrix function between the sequence of the PCR fragment and the WT Col-0 mtDNA.

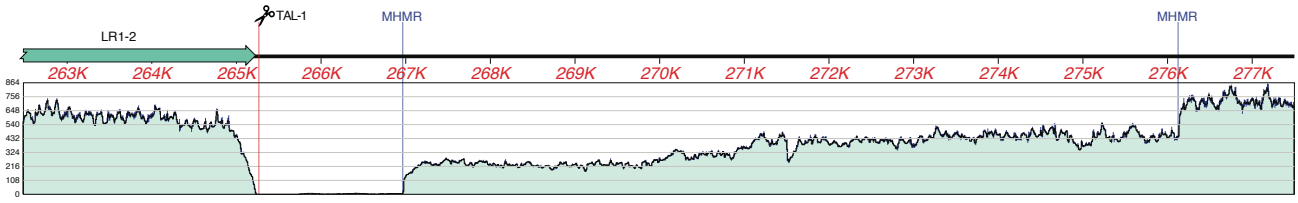


Figure II.16. Coverage plot of Col-0 TAL-1 sequencing reads mapped to the mtDNA, from a T1 transformant with evidence of repair by microhomology-mediated recombination (MHMR). The red line and scissors indicate the Tal-1 target site. The blue lines indicate the microhomology regions that undergo MHMR. Produced by Bowtie2 aligner from MacVector.

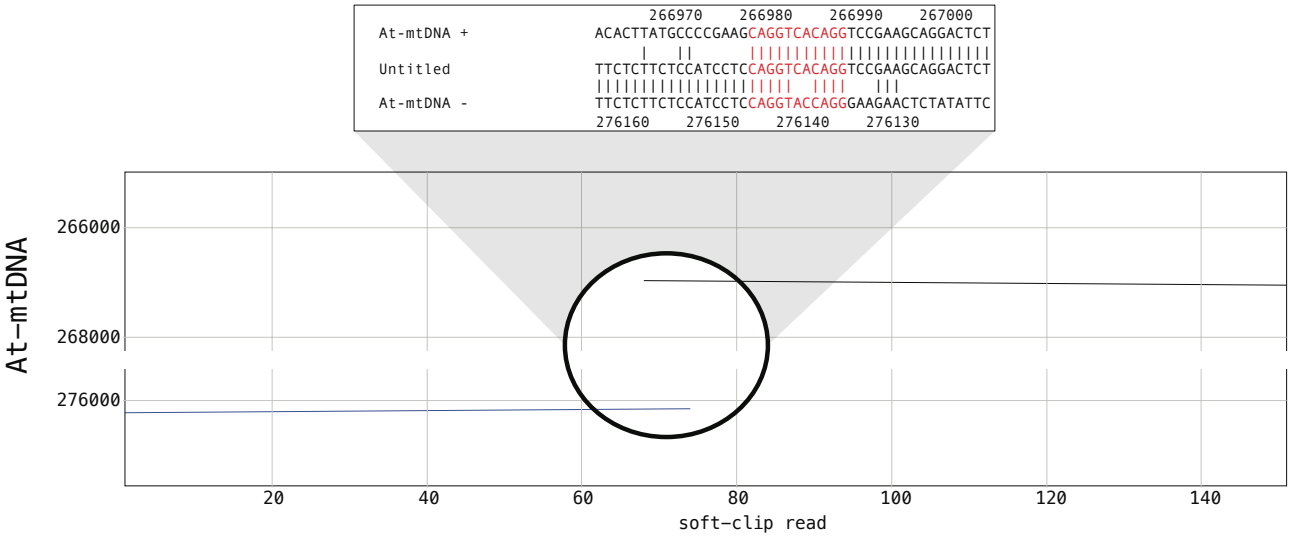


Figure II.17. Alignment of a Soft-clip read to two different mtDNA loci that share a common microhomology site (11 bp with two mispairings). Base-pairing of the read bridging the two sites of the mtDNA is shown in red. The dot-plot alignment was generated with the MacVector Pustel DNA Matrix function.

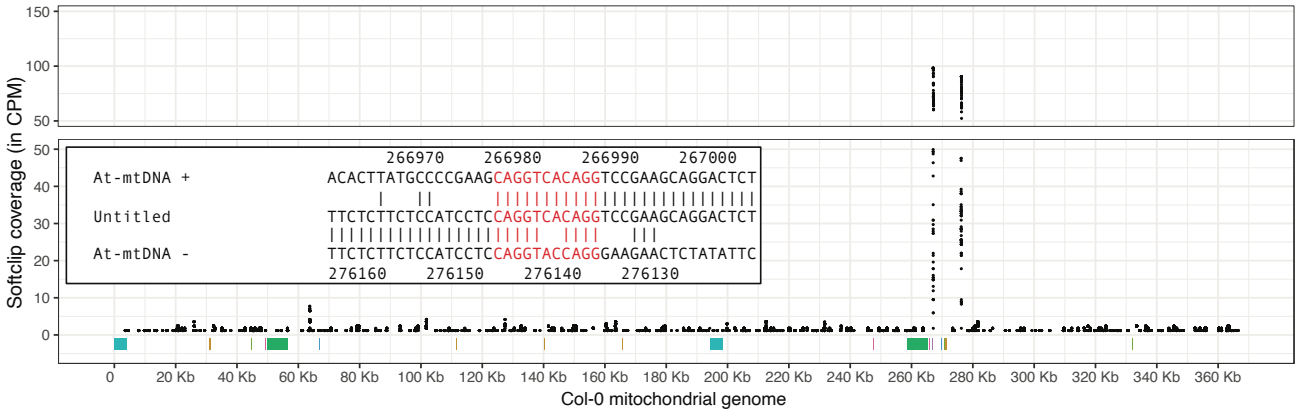


Figure II.18. Coverage plot of Col-0 TAL-1 illumina reads with soft-clips > 20 bp, mapped to the WT Col-0 mtDNA. The two peaks of high coverage correspond to the sites of MHMR. In red, the base pairing of a read bridging the two sites of the mtDNA.

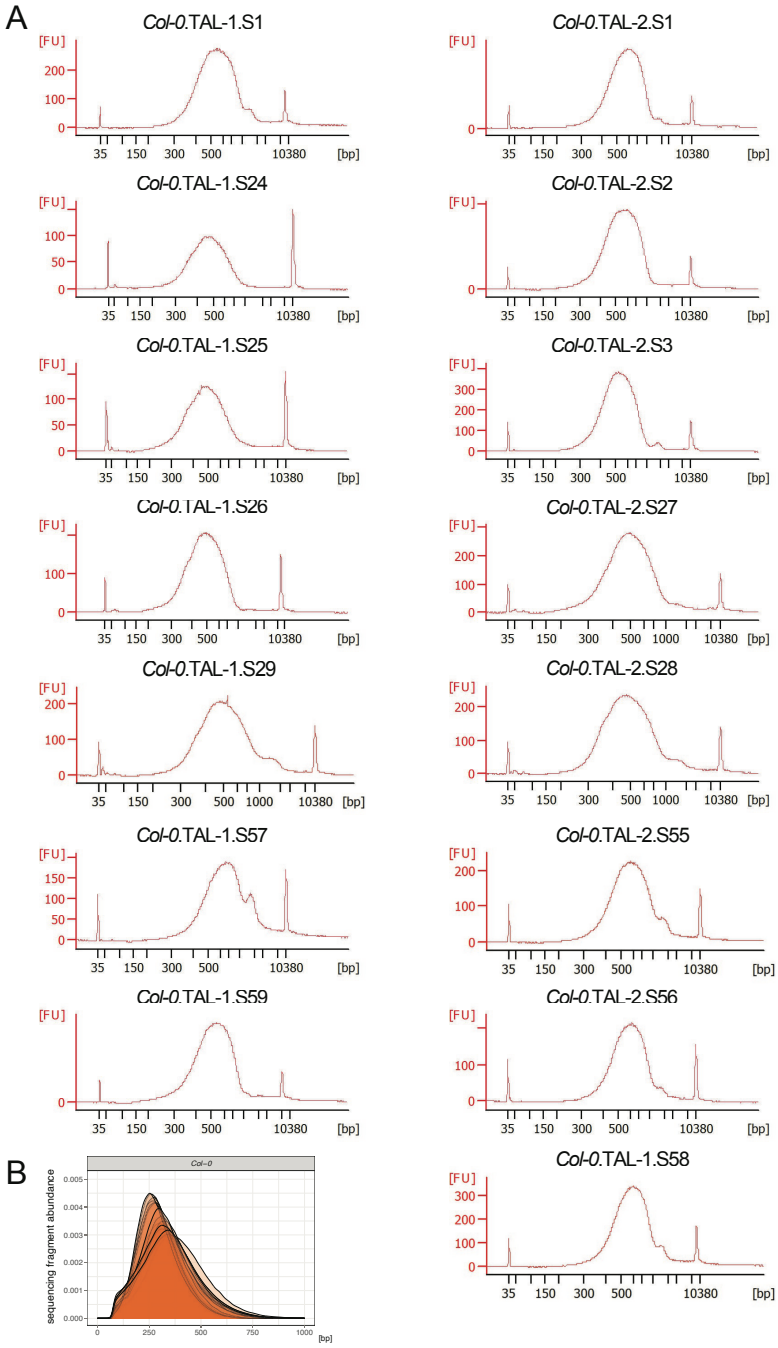


Figure II.19. Analysis of Illumina libraries prepared from T1 Col-0 mitoTALEN samples. (A) Distribution of DNA fragment sizes of Illumina libraries produced by BioAnalyser. (B) Estimated distribution of sequenced fragment sizes of Illumina libraries. The distribution of read pairs mapping closer than 1000 bp is assumed to represent the fragment size distribution.

change just one or two bp, as Illumina sequencing is highly accurate with the probability of inaccurate base call 1 in 1000. Secondly, some of the mtDNA sequences are also found in the nucleus, yet not 100% identical. Thus, a loss in alignment could be due to sequence differences found in the borders of such regions, where part of the read may be found in both compartments and another part of the read just in one, different from the mtDNA template. Finally, some reads can discontinuously map to two different regions of the template genome, due to genome reorganisation at these loci. As microhomologies are small homology regions that can be spanned by a read, MHMR produces reads mapping to both sites of the microhomology. Thus, in MHMR cases, soft-clip segments of reads mapped to the WT mtDNA can map to a distant mtDNA locus. Re-mapping of these soft-clip segments to the mtDNA results in their clustering across the genome. The resulting peaks are indicative of sites of ectopic recombination involving either microhomologies or very small repeats (smaller than the read size, which in our case was 150 bp) (Fig. II. 18). To distinguish MHMR from HR involving small repeats, small repeats annotated in the reference mtDNA were masked in the analysis. Such analysis provides evidence for the presence of MHMR and helps identify the corresponding recombination sites.

3. Bioinformatic and qPCR analysis for the extent of mtDNA structural variation

Two approaches were designed for a more detailed analysis of repeats involved in HR-dependent repair. First, to better understand which repeats are activated by the HR machinery following the introduction of the MitoTALEN constructs, we developed an NGS data analysis method according to the position of the read-pair. Second, a qPCR analysis approach was designed to estimate the frequency of the different repair pathways activated among T1 generation plants.

3.1. Read pair analysis of Illumina alignments

Previous results showed that minor mitotypes can coexist with the major mutant mitotype, resulting from alternative repair pathways. Thus, an NGS analysis pipeline was designed, to identify all homology-based recombination events within the mtDNA, based on the position and distance of mapped Illumina read-pairs. According to the Illumina library preparation kit, the sequenced fragments are below 1000 bp in size (the average library size was 500 to 600 bp, according to BioAnalyzer) (Fig. II. 19) (Appendix 2). However, the average size of the sequenced fragments, calculated from the NGS data a posteriori. was much smaller – an average of 300 bp. Thus, read-pairs mapped further than 1000 bp apart were considered indicative of mtDNA structural variants, resulting from homologous recombination. Filtering the reads pairs according to insert size on the mtDNA genome would result in the clustering of

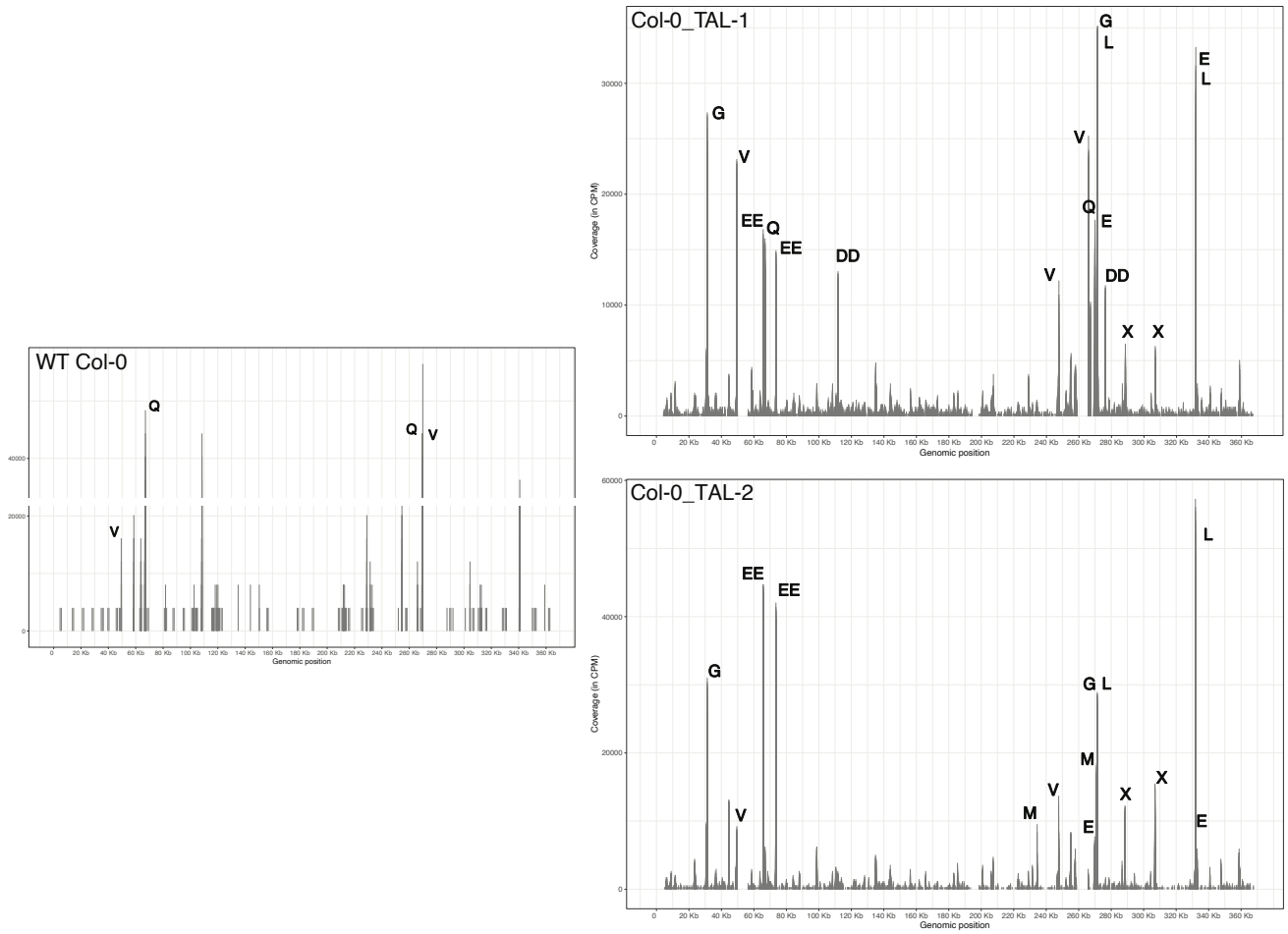


Figure II.20. Coverage plots of Illumina read pairs >1000 bp apart, mapped to the Col-0 mtDNA. From WT Col-0 plants and from Col-0 transformed with either TAL-1 or TAL-2. The peaks indicate structural rearrangements of the mtDNA linked to HR involving the annotated repeats. Coverage values are normalized in CPM (counts per million).

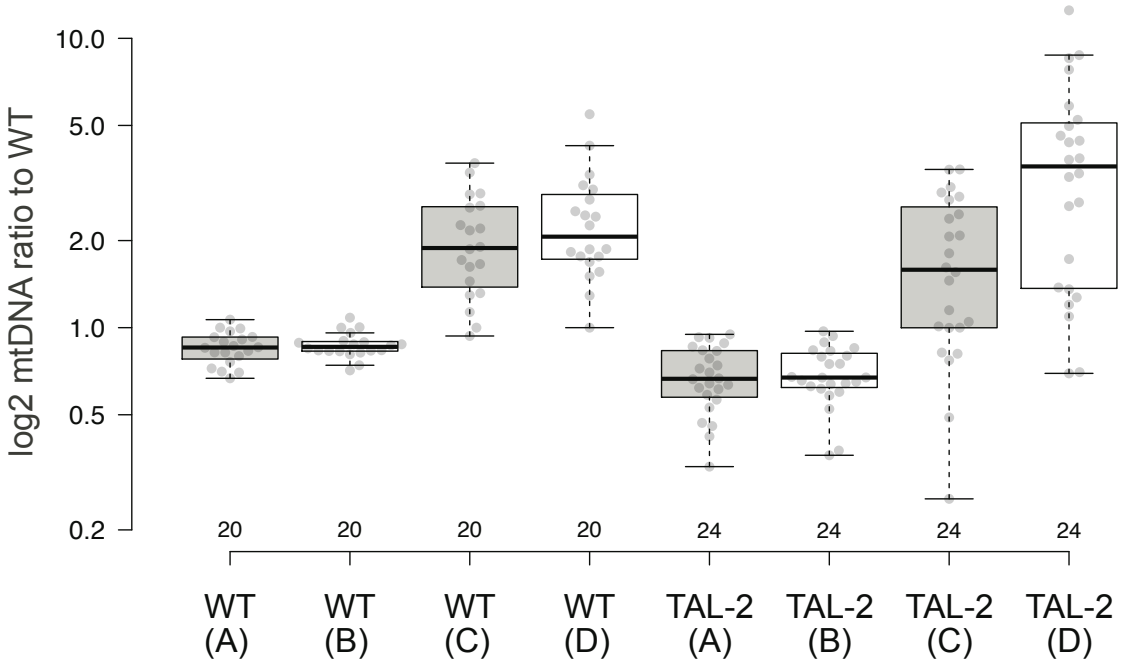


Figure II.21. qPCR analysis of recombination involving repeat EE, in WT Col-0 plants compared to T1 TAL-2 plants. Products A and B correspond to the WT parental sequences, and products C and D are the reciprocal recombination products. Results indicate a 2 fold increase in copy number of recombination product D, in TAL-2 plants as compared to WT ($p = 0.03$).

reads around the repeats (Fig. II. 20). From active repeats in proximity to the mitoTALEN cut sites, it was confirmed that repeats V, Q, G and L are frequently involved in the repair of the DSB. However, repeats distant from the cut sites, such as repeats X, R and repeats EE, also recombine more frequently compared to WT, indicating an apparent general increase in recombination activity, probably induced by the DNA damage. To confirm the activation of global recombination, qPCR analysis was done to quantify recombination involving repeat EE, in Col-0 TAL-2 mutant plants. Results showed a two-fold increase for one of the reciprocal recombination products resulting from recombination involving EE, confirming a slight increase in global recombination following mitoTALEN-induced DNA damage (Fig. II.21).

The frequency of recombination was assessed for each annotated repeat in the reference mtDNA, to determine which repeats are preferentially used as sites for homology-based repair. The number of recombined pairs for each repeat was calculated as a ratio of the total possible recombination event at that specific locus. This ratio provides insight into the relative frequency of recombination events occurring at each repeat site. A more detailed description of the calculations performed can be found in METHODS, section 6.2.2. The frequency of recombination events was represented as a heatmap link between the repeats, in a circular representation of the coverage plot (Fig. II.22) (Appendix 3). Analysis of repeats involved in HR revealed that recombination involving repeats LR1 and G was the most frequent. Yet, in most cases, it was not the only site of recombination. Analysis confirmed that repeats active in repair are not always the same, and depend on the location of the DSB. Therefore, the frequency of recombination involving repeat V is higher than for repeat G in TAL-1 plants, while in TAL-2 plants recombination involving repeat V is not available (Fig. II.23). This analysis re-confirmed previous observations that, apparently, there are mechanisms that select the HR sites in proximity to the DSB for DNA repair.

3.2. qPCR analysis of repeats involved in DSB repair by HR

To assess the frequency of DSB repair by HR involving specific repeats, in a large number of individual T1 plants, the relative copy number of the sequences upstream and downstream the repeats were determined by qPCR (Fig. II. 24). This allowed for the quantification of changes in sequence abundance, reflecting the extent of DSB repair at these repeat sites. According to the Illumina sequencing results, several repeats were identified as most frequently involved in DSB repair, and were chosen for this qPCR analysis. These included repeats Y, V, Q, G and L. Repeats Q, E, and H are too close to each other for individual analysis and, in the case of repeats H and E, they overlap. Therefore, they are collectively analyzed under the designation "Repeat Q." Recombination involving repeats Y, corresponding to events of repair by SSA, was assessed by qPCR quantifying the region between repeats Y and V. Several probes were designed for sites between repeats V and Q,

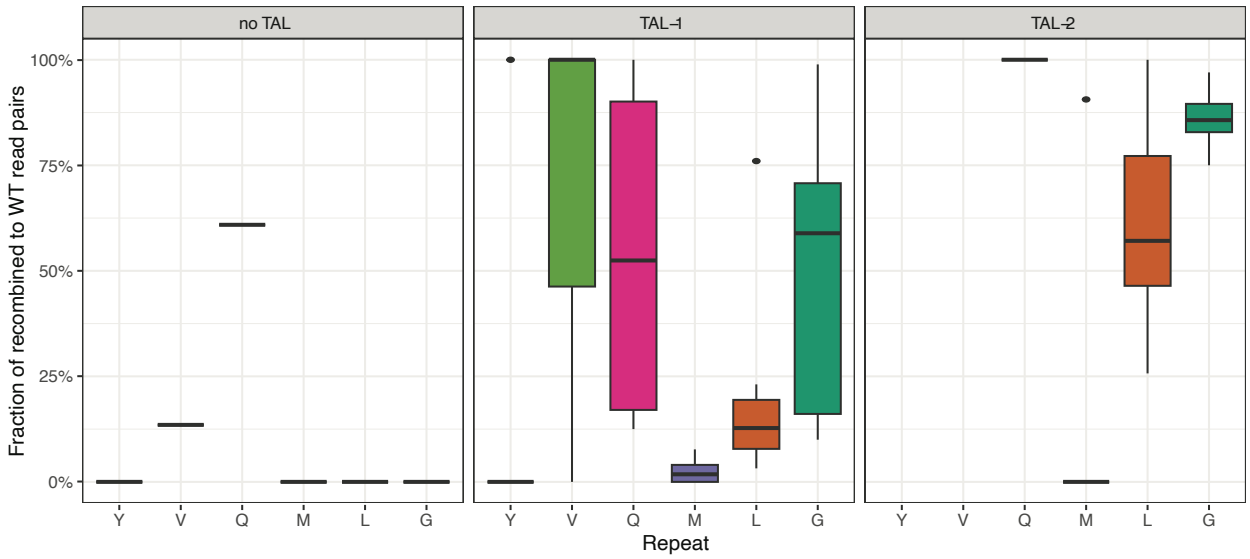


Figure II. 23. Recombination frequency, calculated from sequence data, for repeats that map next to DSB in TAL-1 (n=8) and TAL-2 (n=7) plants, compared to Col-0 (repeats Y, V, Q, M, L, and G). Analysis indicates that HR is more frequent for repeats that are closer to DSB.



Figure II.24. Probes selected for qPCR analysis of repair pathways (SSA and MHMR) and of repeats involved in HR repair of the DSB. Change in mtDNA copy number along both sides of repeat indicates a recombination event.

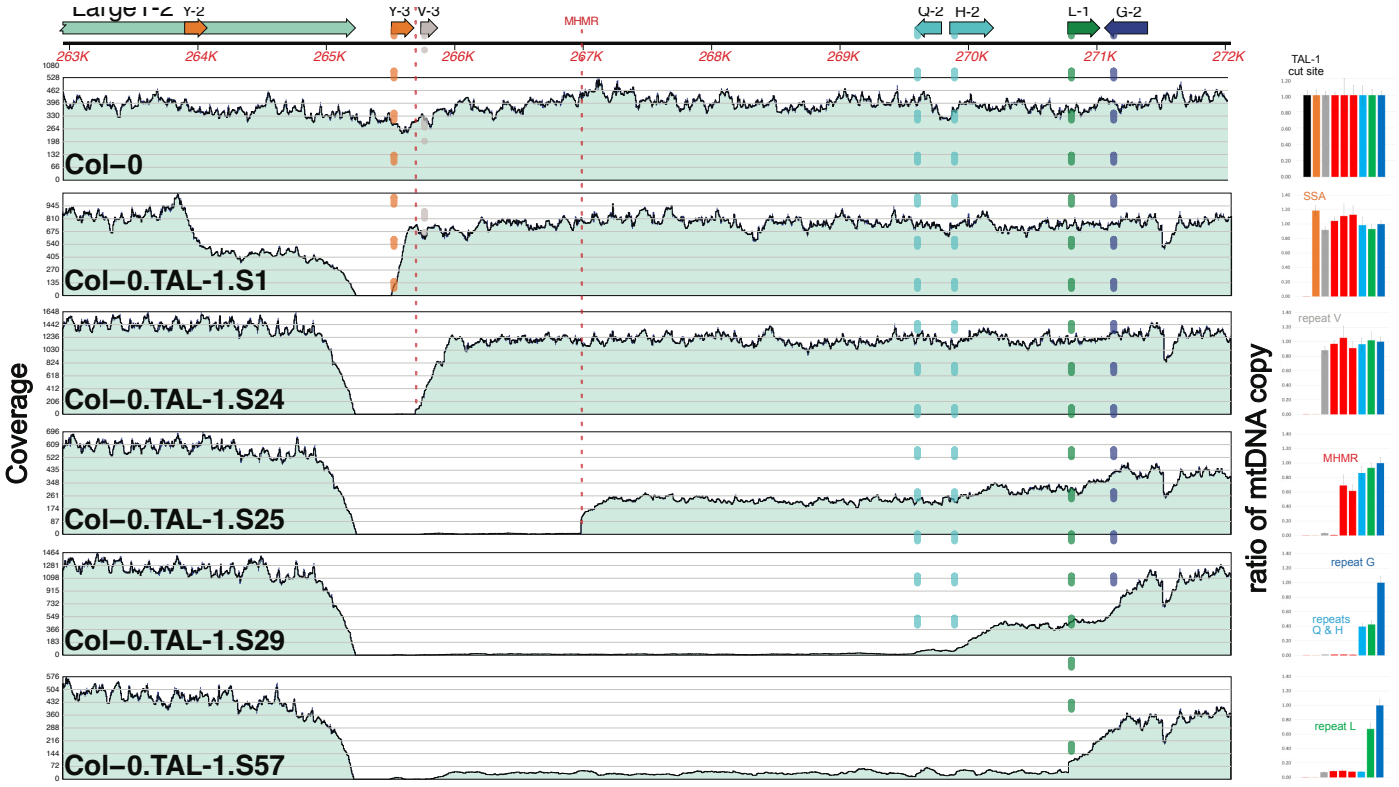


Figure II.25. Validation of the qPCR analysis on samples that had been previously analysed by Illumina sequence. (A) Coverage plots, with color-coded dotted lines indicating the repeats or microhomologies linked to a corresponding drop in coverage. (B) qPCR analysis of the relative copy number of the mtDNA sequence probes indicated in Fig II.25).

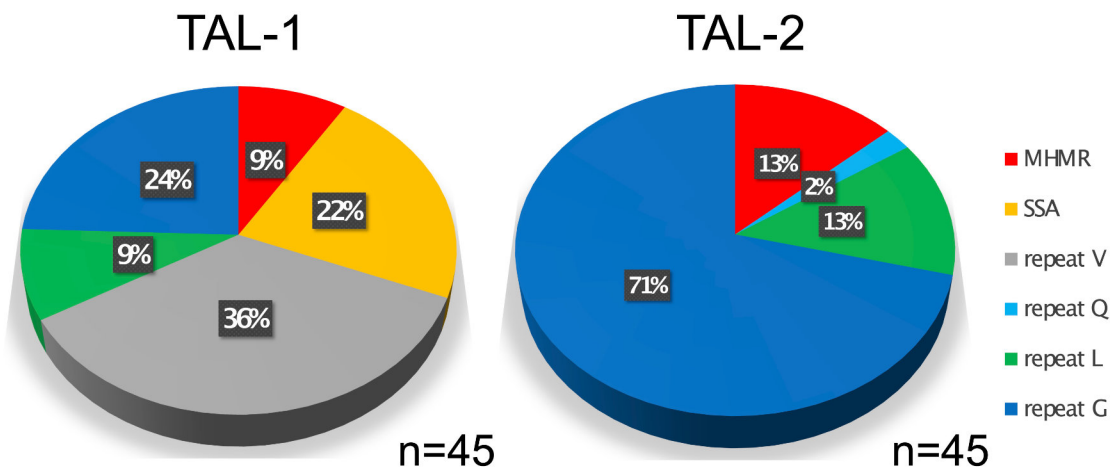


Figure II.26. Results of the qPCR analysis for 45 individual T1 plants transformed with either TAL-1 or TAL-2. Represented as the percentage of plants with a new major mitotype resulting from repair following SSA, MHMR, or HR-dependent repair involving large repeat LR1 and one of the small repeats indicated (V, Q, L or G).

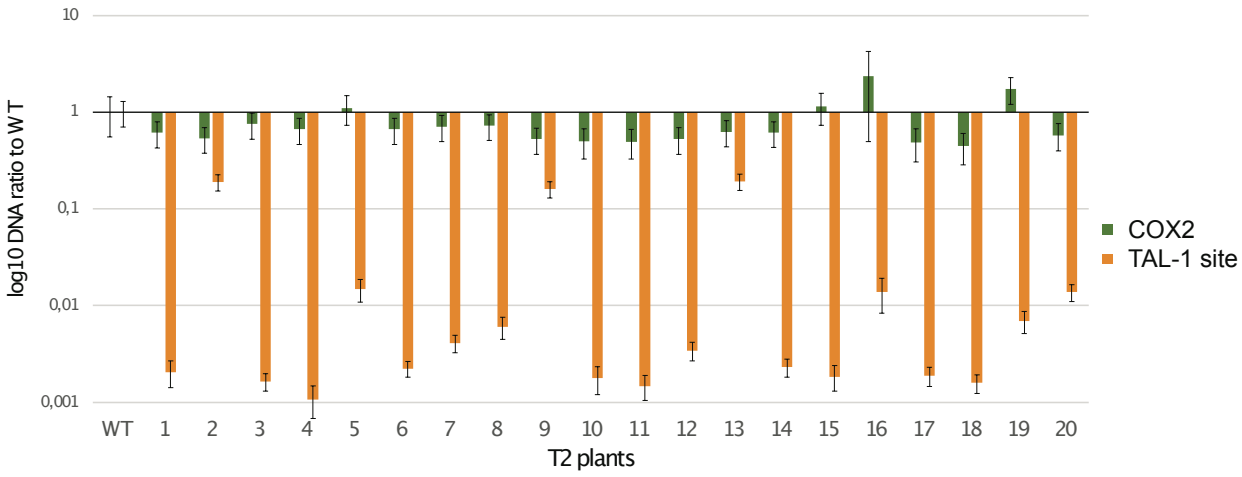


Figure II.27. qPCR analysis of mtDNA copy number at TAL-1 cut site compared to distant COX2 sequence, in T2 progeny of a heteroplasmic T1 plant.

as indicative for MHMR, as there are no repeated sequences in between. Additionally, qPCR probes for the TAL-1 and TAL-2 cut sites were added to the analysis. The validity of our qPCR- based approach was confirmed on the same DNA samples previously analysed by Illumina sequencing (Fig. II. 25). Using this qPCR-based approach we have analysed a total of 45 WT T1 plants transformed with TAL-1 and 45 transformed with TAL-2. This study re-confirmed that HR mobilising repeated sequences is the major repair pathway (60 % in TAL- 1 and 73% in TAL-2), with a preference for repeats closer to the cut site: in TAL-1 plants 36 % mobilising repeat V and only 24 % repeat G, while in TAL-2 plants 71 % mobilised repat G for repair) (Fig. II. 26). Recombination involving repeat L might play a role in the repair of the mitoTALEN-induced DSB, but might also be linked to the segregation of alternative mitogenomes, as discussed later in Chapter III. Yet, it remains unclear why repair by recombination involving repeat G is preferred to recombination involving other repeats closer to the cut site, like Q, E, or H. That might be linked to repeat size and the relative location of the other repeat pair. Overall, our analysis shows that the DSB repair mechanism is highly dependent on the availability of repeats in proximity of the cut site.

4. Heteroplasmy of the mtDNA and segregation of structural variants

4.1. Heteroplasmy of the mtDNA can lead to fast structural evolution

From sequencing results, it has become clear that heteroplasmy of the mtDNA might increase due to the creation of alternative new mitotypes during DSB repair. Although, as previously stated, most mitoTALEN mutants are nearly homoplasmic for the loss of mitotypes containing the mitoTALEN cut site, they could have generated multiple mutant mitotypes, depending on how the DSB was repaired. Furthermore, several mitotypes may make a major contribution to total mtDNA, as observed in sample Col-0.TAL-1.S29 in Fig. II.25.

Even though it was stated multiple times that mitoTALENs are highly efficient in producing homoplasmic mutants, some could still retain the WT mitotype. In order to test the segregation of the co-exiting mitotypes, we followed the T2 generation from a TAL-1 T1 mutant that was heteroplasmic for the mitoTALEN cut site and that still retained significant copies of the WT mtDNA. In T2 segregating plants that no longer contained the TAL-1 construction we tested by qPCR for the presence of the cut site in the mtDNA. Results were normalized against mitochondrial cox2 gene as a proxy to the total mtDNA copy number. We observed that in certain T2 plants (3 out of 20) the mtDNA could revert to the WT mitotype, while the remaining plants mainly retained the variants resulting from repair, which no longer retained the TAL-1 cut site (Fig. II. 27). This result is in agreement with published observations, that there is rapid

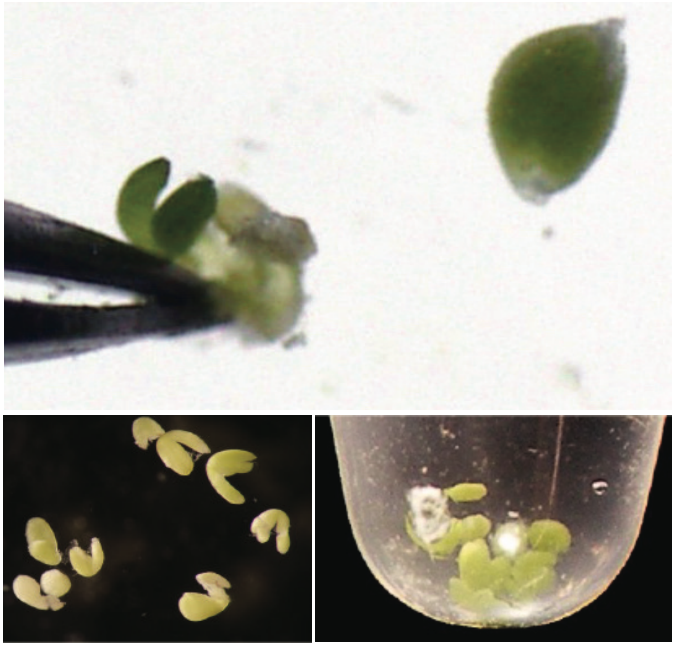


Figure II.28. Extraction of embryos from developing seeds, after pollination of WT Col-0 plants with pollen from a homozygous TAL-1 plant. Top panel left shows an extracted embryo, and on the right a 2-week-old developing seed, The lower panel shows images of the extracted embryos, and in the Eppendorf tube before DNA extraction for mtDNA analysis by qPCR.

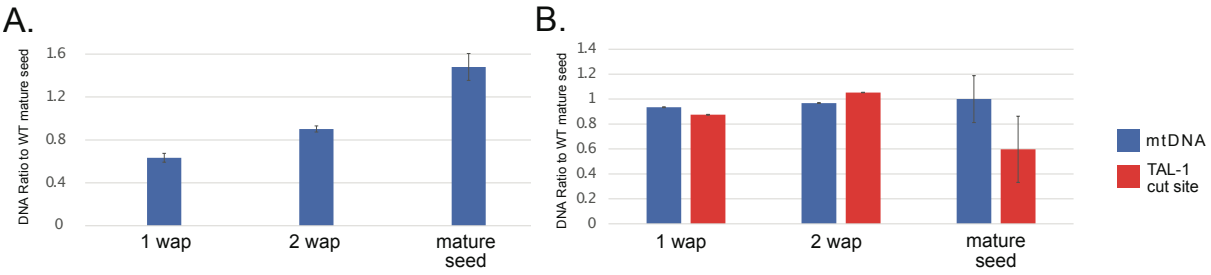


Figure II.29. qPCR analysis of TALEN activity during seed development when the TAL-1 construct was introduced through the pollen allele. (A) The mtDNA copy number (COX2 gene) was normalised to nuclear DNA, reflecting an increase in mtDNA during seed development. (B) The Copy number of the TAL-1 cut site, normalised to the COX2 copy number, shows a minor effect of mitoTALEN only in the mature seed. (wap = week after pollination).

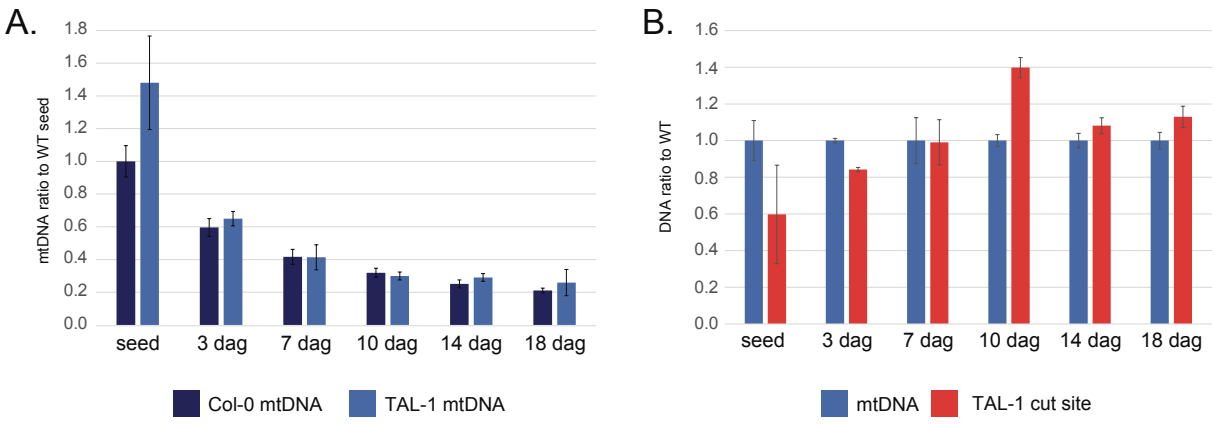


Figure II.30. qPCR analysis of TALEN activity during seedling development. A) mtDNA copy number (COX2 gene) normalised to nuclear DNA, reflecting a decrease in mtDNA during seedling development. B) TAL-1 cut site DNA copy number normalised to COX2, showing just a minor effect of mitoTALEN in mature seed. (dag – days after germination).

sorting out of co-existing mitotypes from heteroplasmic mitochondria (Cree *et al.,* 2008; Gualberto and Newton, 2017).

4.2. Differential efficiency between mitoTALENs introduced by Agrobacterium-mediated transformation and by crossing

During the project, I tested alternative methods to introduce mitoTALEN constructions to plants' nuclear genomes. Either by *de novo* Agrobacterium-mediated transformation or by the cross with a homozygous mitoTALEN line (TAL-1$^{+/+}$) as pollen donor. We observed significant differences in mitoTALEN efficiency. While agrobacterium T1 transformants were nearly homoplasmic mutants early in development, with a mtDNA that had almost completely lost the mitoTALEN target site, F1 plants from the cross Col-0 x TAL-1$^{+/+}$ revealed minor or no trace of mitoTALEN activity (Fig. II. 28 & 29). In mature seeds a minor decrease in the mtDNA sequence corresponding to the TAL-1 site was detected, with increased total mtDNA (Fig. II. 29). However, no change in copy number at TAL-1 cut site was detected at the seedlings stage, suggesting possible mtDNA sorting (Fig. II. 30). Significant loss of the mitoTALEN cut site in F1 generation was observed just in mature plants, producing high heteroplasmy (Fig. II. 31).

We considered the hypothesis that the TAL-1 construction was silenced and not expressed when transmitted through pollen. This could be because both left and right mitoTALENs are under the same RPS5 promoter, which also exists in the Arabidopsis genome. To check whether there was gene silencing in the F1 generation, RT-qPCR relative quantification was done, testing for the FokI sequence specific to the mitoTALEN transcripts. The mitoTALEN transcript was detected in both F1 and F2 mutant lines resulting from the cross with TAL-1$^{+/+}$, at similar levels as those observed in T1 transformants (Fig. II. 32). However, there was no reduction in the abundance of the mitochondrial *atp-2* gene transcript that contains the TAL-1 targeting site, while *atp6-2* was strongly reduced in the two T1 plants tested, consistent with previous mtDNA analysis. To analyse further mitoTALEN expression, a western blot was performed on mutant lines, testing for the presence of the HA-tag that is fused to the Left mitoTALEN (see Fig. II.2A). Yet, no mitoTALEN protein could be detected (expected size about 120 kDa), neither in the T1 lines produced via Agrobacterium nor in the F1 progeny of a Col-0 x TAL-1$^{+/+}$ cross (Fig. II. 33). Faint bands of about the expected size were also present in the WT control.

4.3. Differential segregation of mitotypes across plant tissues

Heteroplasmy of the mtDNA has been noticed not just across mitoTALEN plants and resulting progeny, but also in different tissues across the same plant. We noticed that several MitoTALEN plants developed chimeric phenotypes, with just part of the plant affected, having

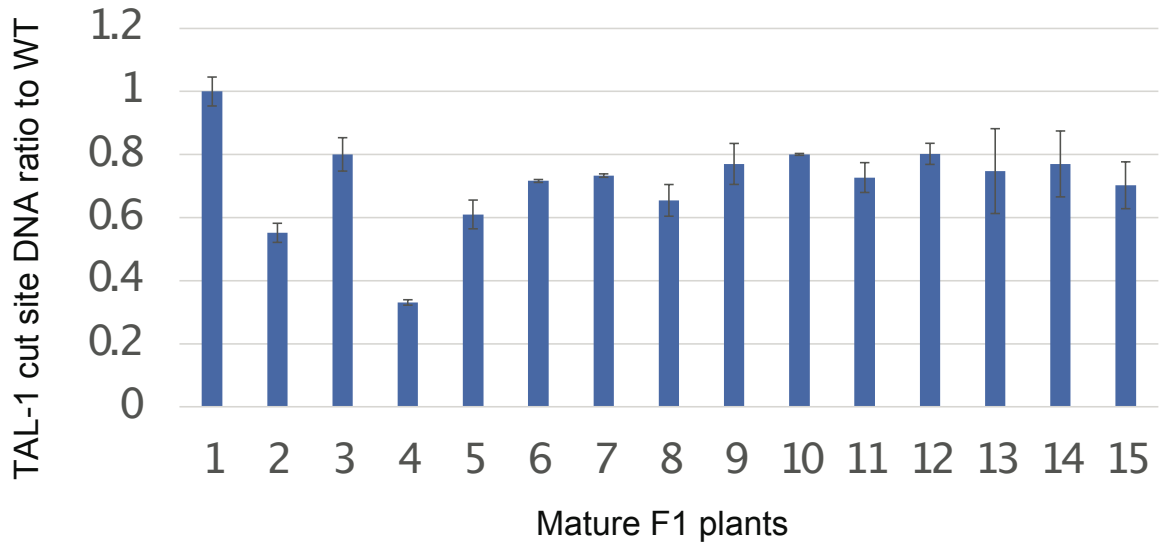


Figure II.31. qPCR analysis of mtDNA copy number at TAL-1 cut site compared to distant COX2 site, in the progeny of WT Col-0 crossed with homozygous TAL-1 mutant.

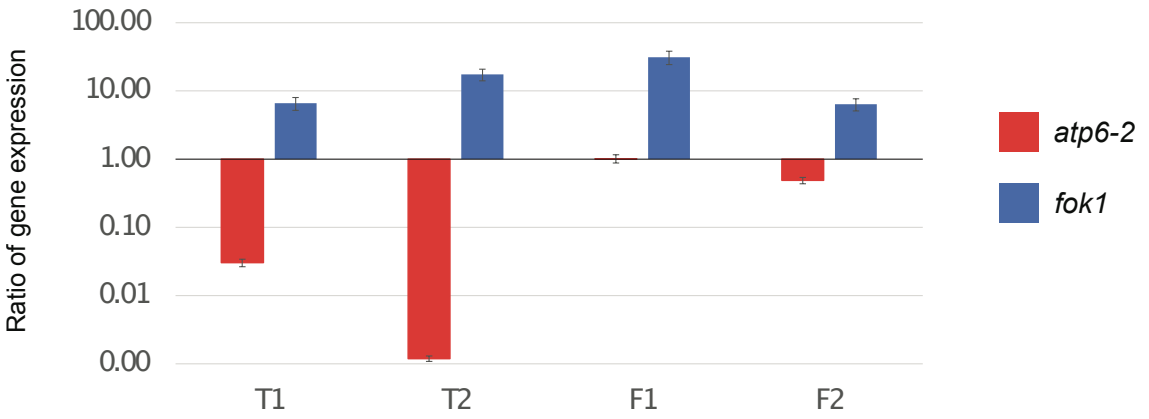


Figure II.32. RT-qPCR analysis of atp6-2 (containing TAL-1 targeting site) and fok1 (TAL-1 transcripts) gene expression in TAL-1 transformants (T1 and T2 plants) and in the progeny of Col-0 crossed with homozygous TAL-1 mutant (F1 and F2 plants). Results confirmed the expression of the mitoTALEN (fok1) transcripts in all mutants, yet a reduction of atp6-2 was only observed in the T1 and T2 mutants, where the TAL-1 construct was introduced via agrotransformation.

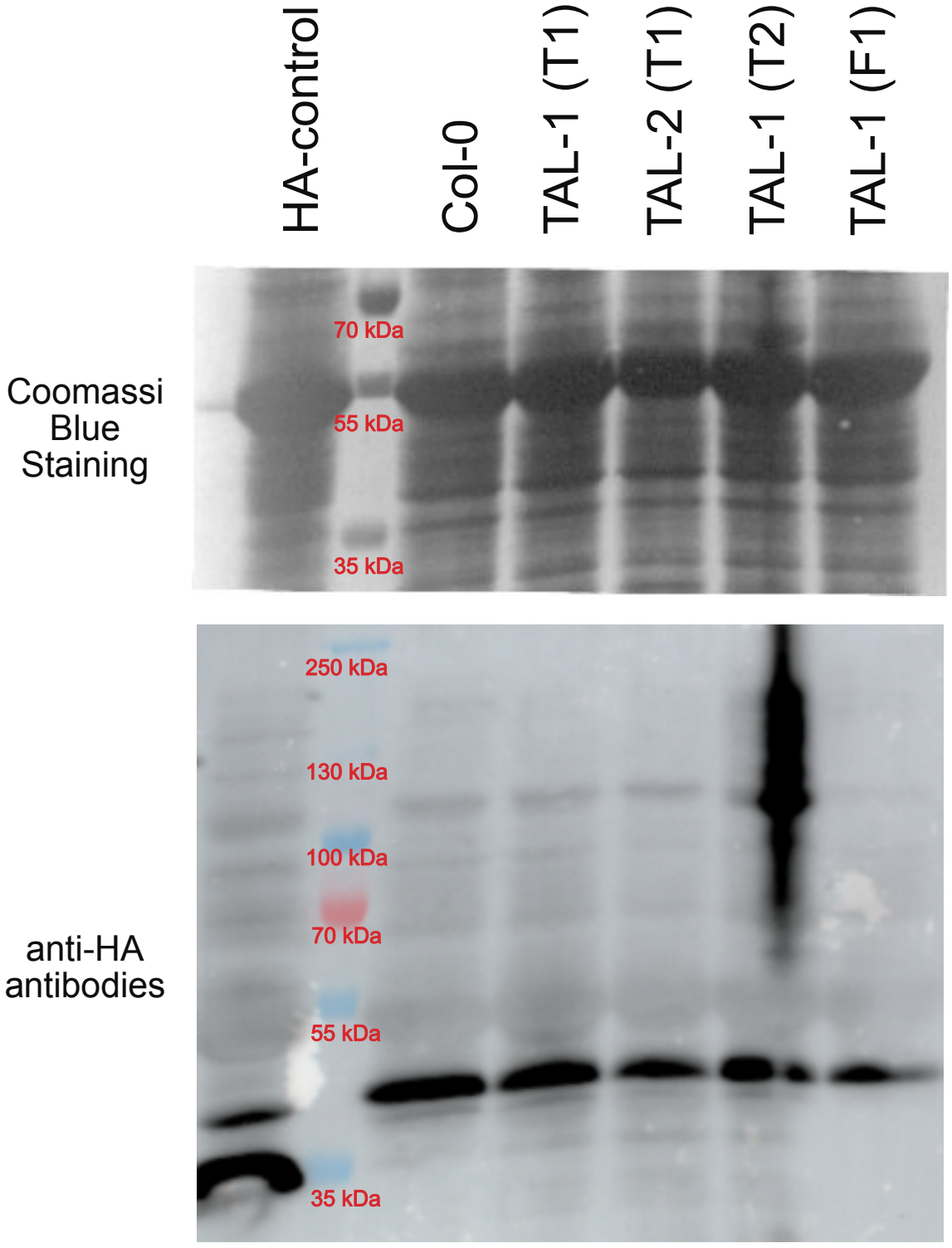


Figure II.33. Western Blot analysis of TAL-1 (T1 and T2) and TAL-2 (T1) transformants and progeny of Col-0 crossed with homozygous TAL-1 mutant (F1). Coomassie Blue Staining confirms equivalent protein loading (top image). Immunodetection using an antibody against the HA-tag (bottom image) failed to detect the TALEN proteins, of an expected molecular weight of 114 kDa. Yet, an HA-containing protein used as a positive control was detected.

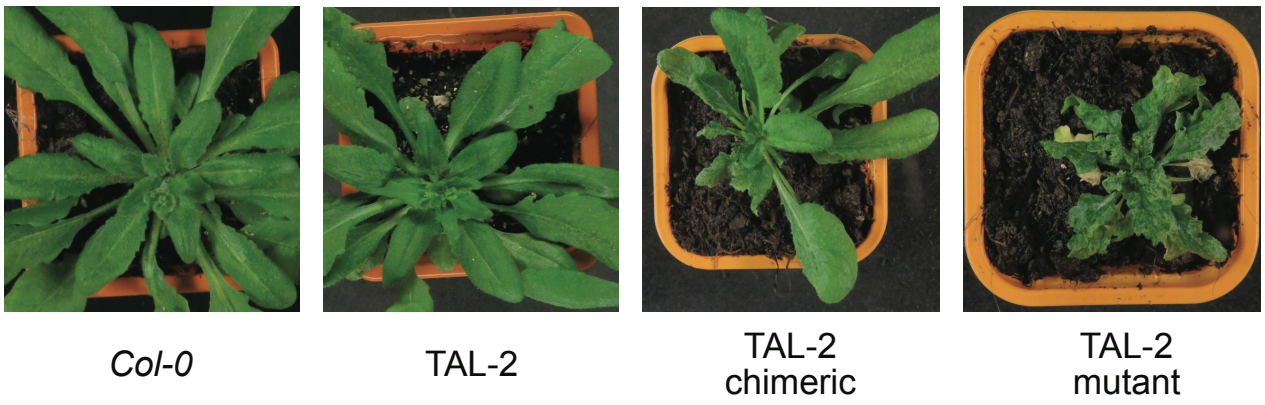


Figure II.34. Pictures of 4-weeks old Col-0 and of TAL-2 Col-0 mutant plants depicting pleiotropic developmental phenotypes: WT looking, chimeric (just part of the plant is affected) and the whole plant is severely affected.

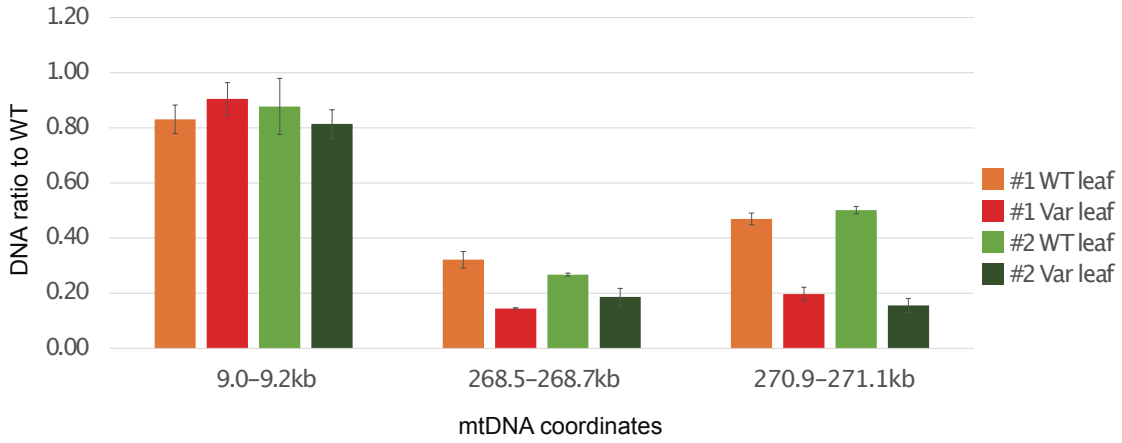


Figure II.35. qPCR analysis of TALEN activity in two independent chimeric plants, comparing an affected leaf (small and curly phenotype) to a WT-looking leaf. The analysis targeted a region distant from the cut site (9.0-9.2 kb locus of the mtDNA) and next to the TALEN cut site (268.5-268.7 kb & 270.9-271.1 kb). Results showed a more pronounced decrease in copy number at the TALEN cut site in the affected leaves compared to the WT-looking ones.

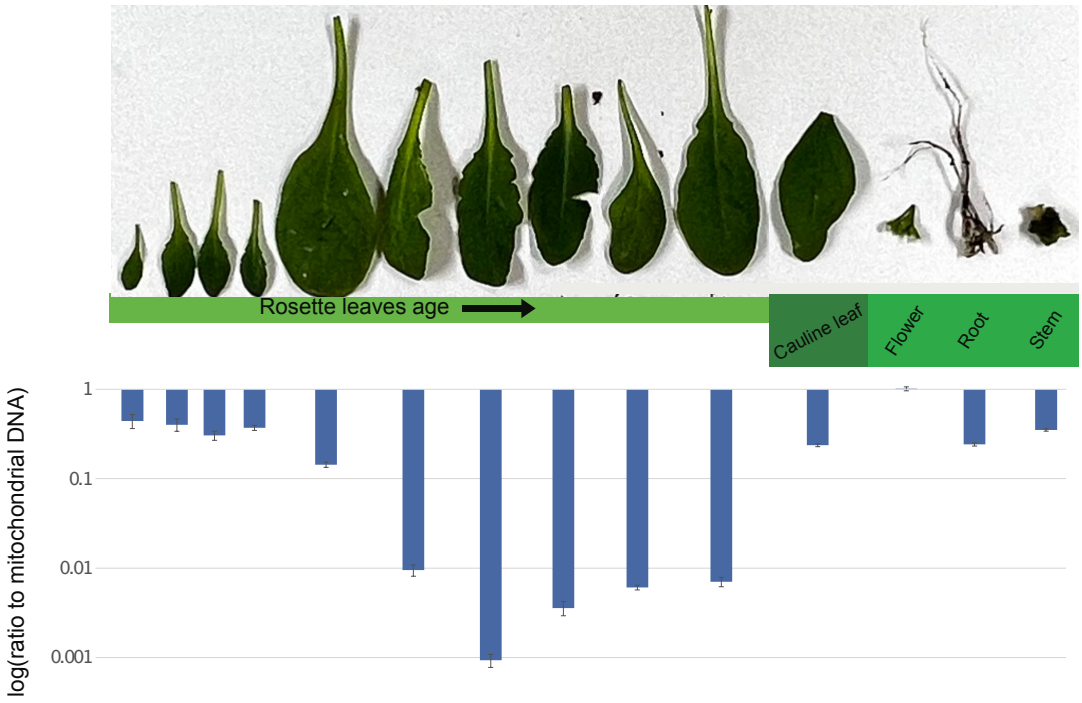


Figure II.36. Analysis of TALEN activity in different tissues dissected from F1 mutant plant. The results of the qPCR relative quantification of the TAL-1 cut site in the plant tissues shown above.

growth defect phenotypes of reduced growth and crumpled leaves, while the remaining parts of the plant developed normally (Fig. II.34). Comparison of the mtDNA present in leaves from the same plant with and without developmental phenotype revealed differences in how DSB was repaired. The qPCR analysis showed that a region not linked to the locus targeted by the MitoTALEN is not affected in copy number, while two regions closer to the cut site were significantly reduced in copy number, and more in the affected leaves than in the normal-looking leaves of the same plants (Fig. II. 35). These results were consistent with the hypothesis that the phenotypes result from the outcomes the DSB repair has on the mtDNA structure and on mitochondrial gene expression. More detailed studies and discussion on the origin of the developmental phenotype observed in several MitoTALEN lines is presented in Chapter IV.

To further study the mtDNA heteroplasmy across plant tissue type, DNA was extracted from different plant organs and leaves, at different developmental stages of F1 mitoTALEN plant resulting from Col-0 X TAL-1(+/+) crosses. Relative copy number of the mtDNA at the mitoTALEN cut site was analysed by qPCR (Fig. II. 36). Analysis revealed that there was a more significant loss of mitoTALEN cut site mtDNA in older leaves of F1 plant and no detectable change in the flower part, compared to Col-0.

5. Pilot study: trying to develop new mtDNA editing techniques

In the previous sections, I have shown that mitoTALEN is a highly efficient tool for introducing a specific DSB in the mtDNA. However, I also showed that sites of homology recombination to repair DSB vary from plant to plant. Although I showed that there is some preference for repeats closer to the DSB for the repair, there are no available strategy to choose the pathway of repair to direct the production of a desired new mitotype. In this section I present a pilot study to artificially link two not-associated distant loci by the production of exposed DNA regions for DNA repair. After DSB, the recession produces exposed single-strand DNA overhangs that can act as intermediates for homology-based repair. Thus, the introduction of intermediates from two distant loci might facilitate their connection via homology that, under normal circumstances, does not occur. Such a tool would be of great interest for mtDNA engineering, as it would provide a method to artificially produce new mitotypes and generate mtDNA genetic variability.

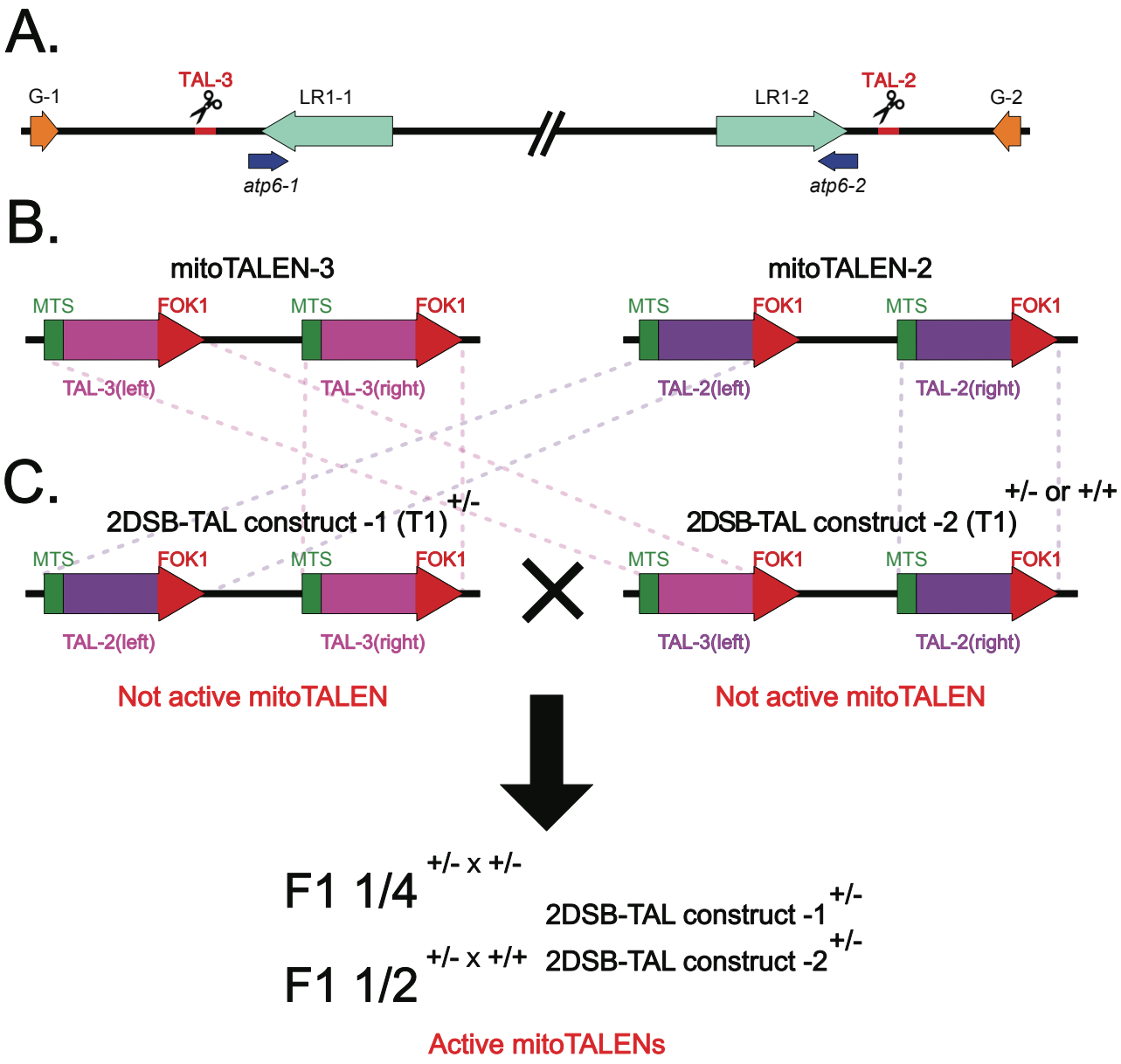


Figure II.37. Schematics of introducing two DSBs using two 2DSB-TAL constructs. A) Binding sites to the mitochondrial genome for mitoTALEN constructs: TAL-2 and TAL-3. B) The structure of mitoTALEN constructs indicating two copies of protein left and right and their composition: mitochondria targeting sequence (MTS), TALE binding domain, and FokI endonuclease. C) Schematics of two 2DSB-TAL constructs each containing just one copy of TAL-2 and TAL-3, which are reconstituted in active form by cross in F1 generation. Indicating the segregation ratios of active forms of mitoTALENs in F1 generation upon the cross of two heterozygous constructs (Fig II. 38) and heterozygous x homozygous of the the second cross experiment (Fig II. 39).

5.1. mitoTALEN transformation design for simultaneous introduction of two DSBs

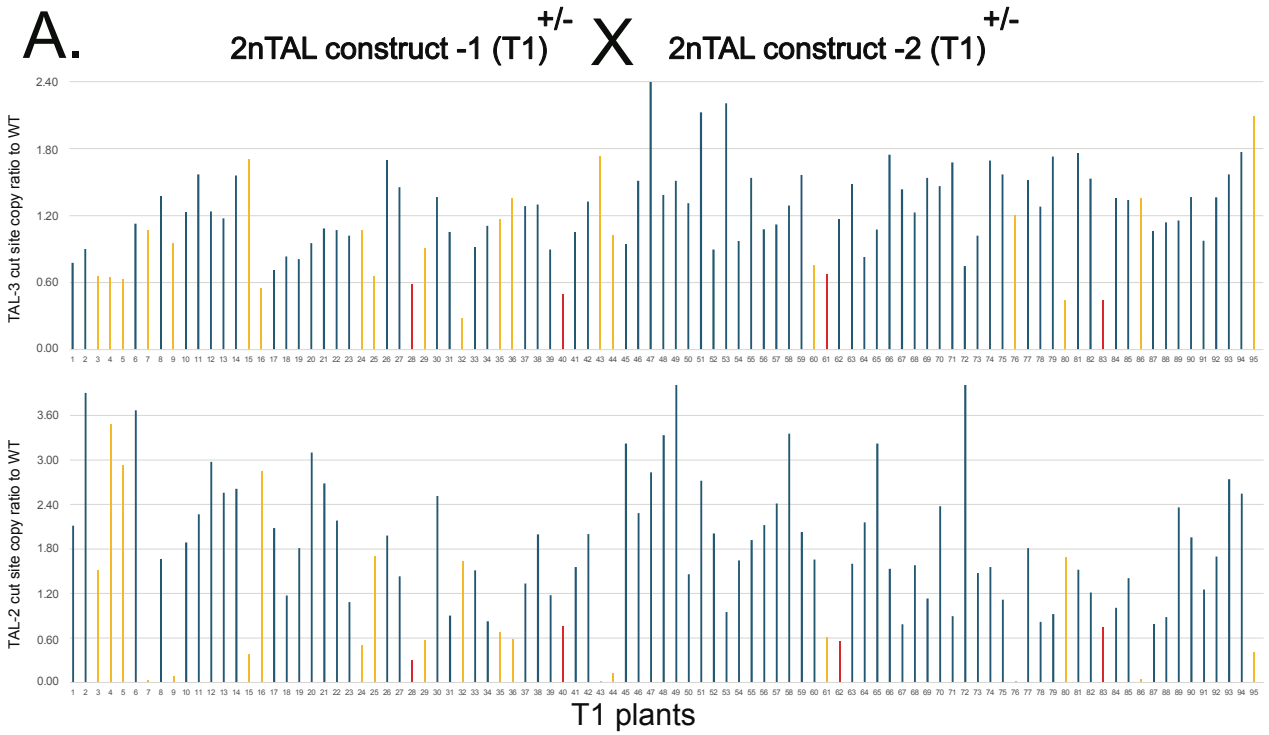
To achieve the co-direction of two DSBs, they must be introduced into the mtDNA simultaneously. This can be done by two mitoTALEN constructs (each consisting of two proteins, Left and Right) split into two vectors. Each contains just one part of both constructs. Therefore, the cross of plants, each containing one of the vectors, would allow simultaneous expression of the two active mitoTALENs in the F1 generation.

For the study, I used two constructs previously published, each targeting a non-coding sequence downstream LR1 copy 2 (LR1-2, used in previous Chapters) and upstream LR1 copy 1 (LR1-1) (Arimura *et al.,* 2020). Thus, TAL-2 and TAL-3 correspond to the constructions annotated as "ATP6-2UP" and "ATP6-1UP" in (Arimura *et al.,* 2020), respectively (Fig. II. 37A). Prof. Arimura kindly provided the constructs, cloned in binary vector pK7WG2 (2DSB-TAL construct-1 and 2DSB-TAL construct-2) (Fig. II. 37B), each vector containing just one part of each TAL-2 and TAL-3 construct (Left-2 and Right-3 in one vector and Left-3 and Right-2 in the second vector). The vectors coded spectinomycin resistance for bacterial transformation and GFP under expression of the oleosin *OLE1* promoter for selection of transformed seeds.

5.2. Segregation of 2DSB-TAL introduced cut sites in the F1 and F2 generations

To reconstitute active mitoTALENs, T1 transformants, each containing a different 2DSB-TAL construct, have been generated (Fig. II. 37C). In the T2 generation the genetic state of the plant (homozygous/ heterozygous) was assessed according to the abundance of seeds positive for GFP expression (Fig. II. 2). We adopted this strategy because the TALE constructs are too similar and repetitive to be distinguishable by PCR. I could obtain a line homozygous for construct 2 (2DSB-TALconstruct-2+/+), but did not manage to segregate plants homozygous for construct-1. Thus, crossing experiments were done between heterozygous 2DSB-TAL construct 1+/- and heterozygous 2DSB-TAL construct 2+/-. To increase the number of F1 plants containing both constructs, crossing experiments were also done between heterozygous 2DSB-TAL construct 1+/- and homozygous 2DSB-TAL construct 2+/+ (Fig. II. 37C). The resulting F1 plants were screened for mitoTALEN activity at TAL-2 and TAL-3 cut sites, by qPCR quantification of the targeting sequences (Fig. II. 38 & 39).

On the first plant crossing experiment, 2DSB-TAL construct $1^{-/+}$ x 2DSB-TAL construct $2^{-/+}$, 24/95 plants had a 40% reduction in mtDNA copy at one of the cut sites (Fig. II. 38A). This aligned with Mendelian genetics, as we expected 1/4 of progeny -/+ for both constructs. However, just 4 plants had a 40% reduction of both sites, and there were no plants with complete absence of both cut sites. Some mutant plants were pale in colour, and some had reduced growth, but no link of phenotype with mtDNA genotype was found, suggesting side



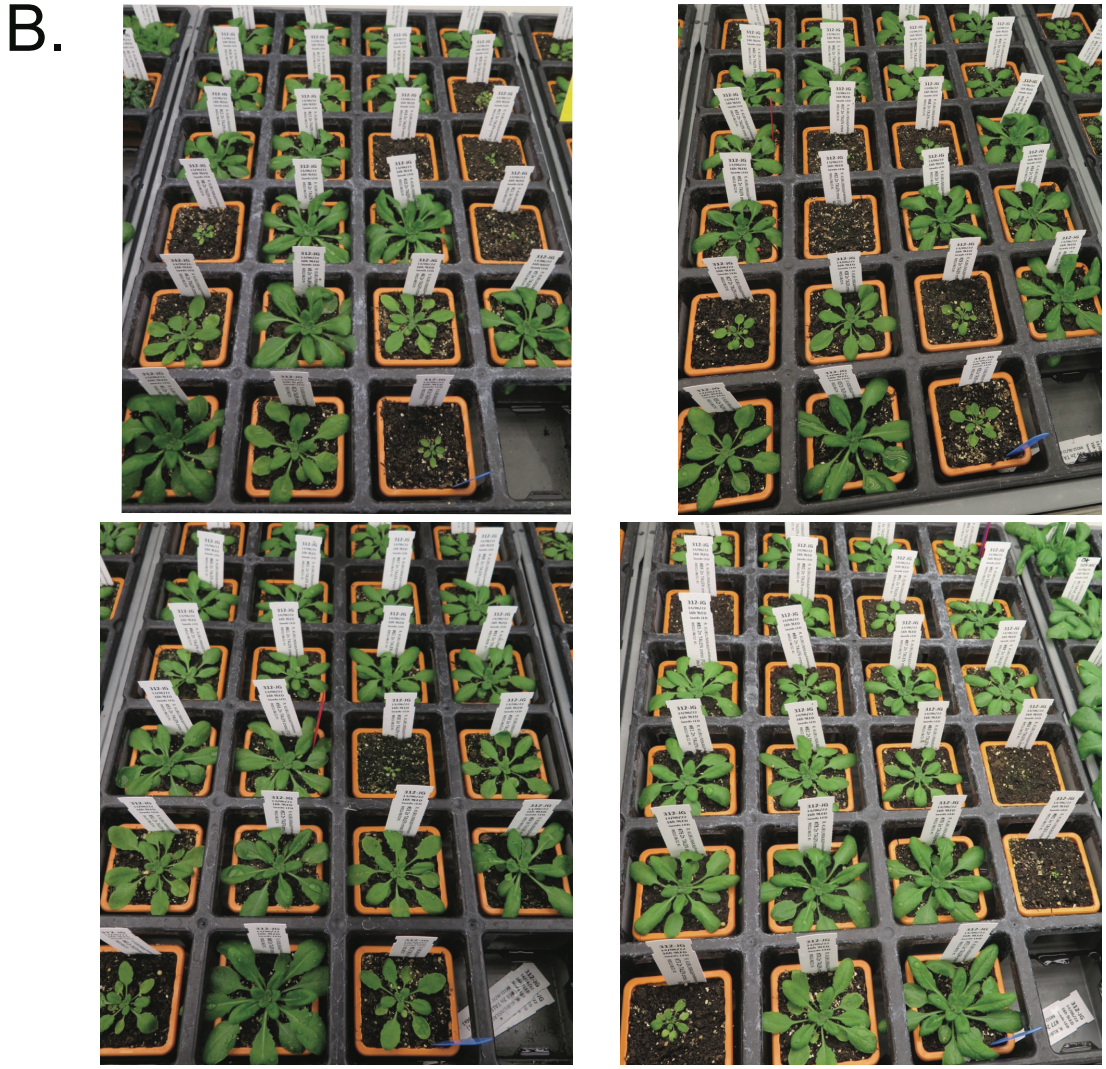


Figure II.38. Screening the F1 progeny of 2DSB-TAL construct-1$^{-/+}$ X 2DSB-TAL construct-2$^{-/+}$. A) Pictures of four-week-old F1 progeny of 2DSB-TAL construct-1$^{-/+}$ X 2DSB-TAL construct-2$^{-/+}$. B) qPCR analysis of mtDNA copy number at TAL-2 and TAL-3 cut sites compared to distant COX2 site, in F1 plants shown in A.

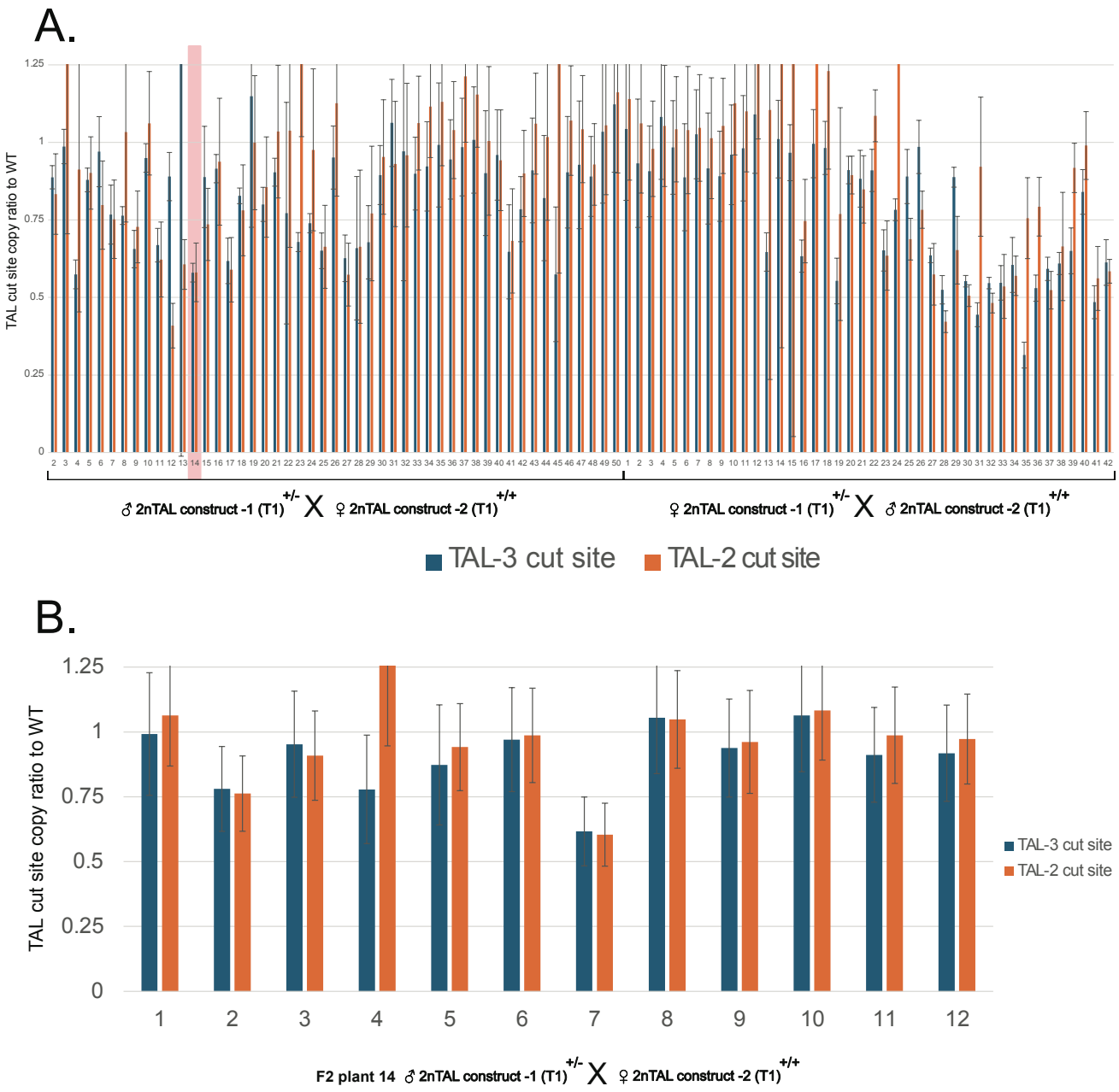


Figure II.39. Screening the F1 progeny of 2DSB-TAL construct-1$^{-/+}$ X 2DSB-TAL construct-2$^{-/-}$. A) qPCR analysis of mtDNA copy number at TAL-2 and TAL-3 cut sites compared to distant COX2 site, in F1 produced by pollen of heterozygous 2DSB-TAL construct-1 crossed with homozygous 2DSB-TAL construct-1, and reverse. B) qPCR analysis of mtDNA copy number at TAL-2 and TAL-3 cut site compared to distant COX2 site, in F2 progeny of plant 14 indicated in red in A.

effects of the T-DNA insertion sites (Fig. II. 38B). For the second crossing experiment, 2DSB-TAL construct $1^{-/+}$ x 2DSB-TAL construct $2^{+/+}$, one of the parents was homozygous for the construct. However, still only 24/95 plants had a 40% reduction in mtDNA copy in one of the cut sites (Fig. II. 39A). Which was not following Mendelian segregation, as it was expected half of the plants to have both constructs. No homoplastic mutants were found for either cut site deletion. In the hope of segregating homoplasmic mutants during SSS, I looked at the progeny of plant 14, which had a 40 % copy number reduction at both TAL-2 and TAL-3 cut sites. Nevertheless, only one out of 12 F2 plants had a reduction of 40 % at both TAL-2 and TAL-3 cut sites, and there were reversions to WT for other F2 plants (Fig. II. 39B). At this stage of the project, we decided that the results obtained from this pilot study were not sufficiently promising to justify further investment in time and effort.

Discussion

1. Designing the mitoTALEN constructs

The Prof. Arimura group provided us with both the TAL-1 and TAL-2 constructs, assembled in the same binary Gateway destination vector. Genotyping of the constructs was difficult, as the only differences between constructs were in the highly repetitive TALEN genes. Designing constructs with different selection markers would have made genotyping easier. In the end, the only reliable way of selecting plants containing the constructs was either to confirm the presence of the Fok1 domain or to analyse the mitoTALEN cut site in the mtDNA by qPCR.

2. DNA library preparation

Compared to a library prepared from total plant DNA, there was a 5-fold final enrichment of mtDNA reads in NGS sequencing runs, from mitochondria-enriched samples. The enrichment protocol could have been further optimised by determining the best plant age for the experiments. While too young plants may yield insufficient material for mitochondria enrichment, those that are too old have increased nuclear endoreplication and higher levels of chloroplast DNA contamination (Cookson et al., 2006; Greiner et al., 2020). Therefore a compromise had to be reached. Furthermore, optimising the sample collection time might have improved mtDNA enrichment. Starch content in chloroplasts typically increases during the day and decreases at night (Graf *et al.,* 2010). Thus, sedimentation and breakage of chloroplasts, which are highly dependent on starch granules content, might differ for centrifugation steps performed during different times of the day. Nevertheless, despite the potential for optimizing mtDNA enrichment, we ultimately achieved sufficient mtDNA coverage from all libraries for all subsequent analyses.

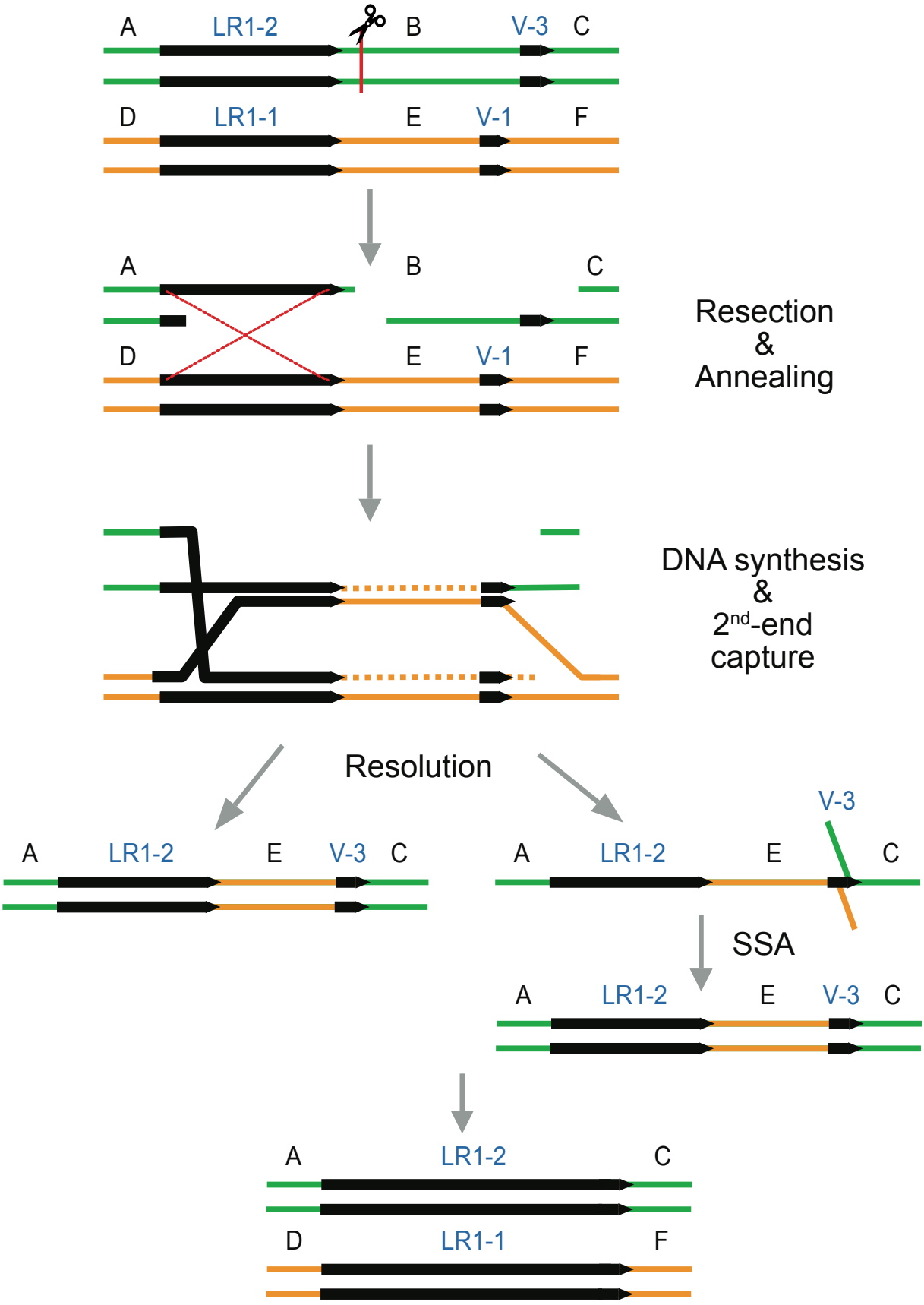


Figure II.40. Proposed mechanism of HR-dependent repair involving two pairs of repeated sequences, in this case, the large repeat LR1 and the small repeated sequence V. The mtDNA regions represented in green and orange are far apart in the genome, yet share the same repeats.

Illumina sequencing was selected over other NGS methods due to its high sequencing coverage at a moderate price. However, 2x150 bp reads might not be the best choice for soft-clip analysis, as bigger reads would result in more reads detected with the soft-clip software, resulting in higher confidence of findings. Furthermore, a library preparation kit yielding larger fragment sizes would have highly improved the detection of homologous recombination sites in NGS read-pair analysis. Fewer fragments can span larger repeats compared to small repeats. Thus, increasing library size could enhance the detection of HR involving larger repeats. Furthermore, it is important to note that Illumina sequencers tend to preferentially sequence shorter fragments; therefore, the library sizes determined by BioAnalyser significantly differed from the average sizes of the fragments that were actually sequenced (Ribarska *et al.,* 2022). In addition, short fragments lead to a loss of sequence coverage due to overlapping paired-end reads.

With foresight and in an ideal scenario, we would have used different libraries for different purposes: single reads of 300 bp for detecting MHMR events, and long reads (PacBio, nanopore) or mate-pair libraries for analysing repair by recombination involving larger repeats. However, these latter approaches would necessitate larger DNA quantities and significantly higher costs for library preparation, thus making it unfeasible to analyse individual T1 plants. In the end, a compromise had to be made.

3. Homologous recombination involving two pairs of repeated sequences.

Three different repair pathways were found in WT plants transformed with the mitoTALEN constructs: HR, SSA and MHMR. Repair of the DSB by HR is predominant, creating a new major mitotype. According to the repair fingerprint in WT plants, deduced from NGS coverage profiles and PCR analyses, the hypothesis was made that DSBs are most frequently repaired by homologous recombination directed by two pairs of repeats. The repair pathway involved the large repeat LR1 (6.5 kb), located on one side (left) of the cut site and one of the small repeats in proximity (170-340 bp) located on the other side (right). The proposed mechanism suggests that one of the homology sites is used during strand capture, while the second homology region locks and stabilises the extension by 2nd-end capture, resulting in duplication of the template DNA at the cut site (Fig II. 40).This mechanism suggests how new large repeats can be created, becoming predominant in plant mitochondrial genomes, and greatly contributing to the large content of non-coding DNA and the big size of plant mtDNA genomes. Indeed, repair options that lead to the loss of large fragments, replaced by smaller fragments, increase the likelihood of losing essential sequences. There may be a preference for replacing smaller fragments with larger ones, which, over evolutionary

timescales, contribute to increased repeat and genome sizes. Furthermore, the resulting mtDNA after DSB repair establishes a new equilibrium of mtDNA copy numbers across the genome, suggesting the existence of control over DNA replication of resulting mitotypes and homologous recombination.

4. NGS read pair analysis

Coverage analysis of reads mapped to the mitochondrial genome revealed structural changes that result in major changes in the equilibrium between alternative mtDNA conformations. Recombination events that do not disrupt mtDNA equilibrium are hidden and may not be readily apparent in coverage plots. Thus, to detect all recombination events, read-pairs further apart than 1000 bp were analysed separately. The Illumina library preparation kit utilized in this project allows the preparation of sequencing fragments much smaller than 1000 bp; thus, a read pair with a higher insert size was interpreted as representing structural variants likely resulting from homology recombination. Several factors need to be taken into account in the analysis. First, some repeats are larger than the average size of library fragments, limiting recombination event detection as fragments may be too small to span the repeat. Secondly, detecting HR for small repeats is not limited by fragment size, and sequencers have a higher preference for smaller fragments. To conclude, quantitative NGS analysis of recombination involving repeats should considered with a pinch of salt, as it relies on many variables, which could be overcome by the preparation of more homogeneous libraries of bigger fragment size. For example, nanopore sequencing could be better for the analysis of recombination frequency via discontinuous mapping of long reads.

5. Differences in repair efficiency linked to how mitoTALEN constructs are introduced in the plant

With the initial objective of determining the timing of DSB repair following the introduction of the mitoTALEN construct, WT Col-0 plants were crossed with homozygous TAL-1 mutants as the pollen donor. An unexpected difference in DSB repair efficiency was noticed, compared to what had been observed in primary T1 transformants. While plants transformed with agrobacterium were almost completely homoplasmic and showed a highly reduced copy number of the mitoTALEN target sites, no apparent mtDNA damage was observed in F1 generation when the mitoTALEN construct was introduced by crossing. The major difference between the two approaches is that in Agrobacterium-mediated transformation developing ovules are the targets of agrobacterium (Ye et al., 1999), while in crosses the pollen alleles are introduced in mature ovules. It could be assumed that mitoTALEN was active in flowers just upon construct introduction by agrobacterium, and we would then expect mitoTALEN to be restored in the F2 generation, but that was not the case.

To explain these results, we considered the possibility that the mitoTALEN construct might get silenced, as it contains multiple identical promoters. However, RT-qPCR indicated that mitoTALEN transcripts are present in adult plants at levels similar to those observed in T1 transformed plants. Still, no mitoTALEN protein could be detected in any of the TAL lines. This could be due to the HA-tag being hidden by the mitochondrial targeting sequence (the HA-tag is found between MTS and TALEN sequences), or just that the proteins are expressed at very low levels, despite the strong constitutive RPS5 promoter.

Analysis of different tissues and leaf ages of F1 plants revealed differential mitoTALEN activity in leaves, having DSB repair detectable just in older leaves. There might be differential protein expression under the RPS5 promoter depending on the tissue, not described in the literature. Alternatively, it is also possible that in young tissues, the damaged mtDNA is preferentially discarded, while it is repaired in older tissues, depending on the differential development expression of the mitochondrial repair factors.

6. Pilot study: what went wrong?

There are several reasons why the introduction of two DSBs did not produce homoplasmic mutants. The first and most significant issue is that, according to our results, ethe fficiency of mitoTALENs drops significantly during the cross of mitoTALEN lines, which I did not know at the start of the pilot study but later demonstrated with the results of section 4.2. Secondly, as I presented in the result section 2.2.1, the DSB produced by TAL-2 in WT plants is usually processed via homology recombination utilising two sets of repeats, duplicating the distant region at the cut site. The issue is that the duplicated region in the pilot study is compromised by TAL-3. Thus, either the resulting mutant is non-viable, or it is repaired by a mechanism that I did not detect. It is possible that the mitoTALEN constructs were silenced in the pilot study, as already in a heterozygous situation, when both constructs are present there are already 4 copies of the same promoter. This might trigger transcriptional gene silencing. But this didn't seem to be the case, as we could detect the transcripts. Furthermore, many resulting F1 mutants had only a single target site affected. This could be due to different efficiencies of the two mitoTALENs, or to the genome environment of the DSB. Finally, some studies previously have shown that TALENs may not necessarily require two different proteins. According to that study, dimerization of the FokI endonuclease domain can occur between two same TALEN proteins, with only one TALEN binding to the DNA (Forner *et al.,* 2022).

CHAPTER III: Study of Homologous recombination pathway

Introduction

To study the role of key proteins in plant mitochondrial HR pathways, the mitoTALEN constructs have been transformed into the corresponding mutant. Below, I briefly describe such factors. The corresponding specific mutant lines will be described in results section 1.

1. Single-strand DNA-binding proteins

Single-strand DNA binding proteins play essential roles in DNA replication and repair. Arabidopsis have at least two major types of ssDNA binding proteins in their organelles. The first one is orthologs to bacterial SINGLE STRAND DNA BINDING PROTEIN (SSB). In animals and yeast mitochondria, mtSSB assumes functions equivalent to those of its bacterial counterpart. It is a homotetrameric protein closely linked to mtDNA replication either by protecting ssDNA from degradation or mediating replication activity. Reflecting its crucial role in mtDNA maintenance, mutations in the gene leads to mtDNA depletion (Backert *et al.,* 1996; Bianco, 2017; Jiang *et al.,* 2021). Orthologs in *Drosophila melanogaster* regulate mtDNA replication activity by promoting dsDNA unwinding at the replication fork, by replicative DNA helicase TWINKLE (Farr *et al.,* 2004). In Arabidopsis, two SSB proteins, SSB1 and SSB2, were found to localise in mitochondria. On the contrary to what described for animal mtSSB, plant mutants of SSBs showed increased mtDNA levels, suggesting their role in the negative regulation of mtDNA replication (Qian et al., 2022, 2021). The double mutation *ssb1 ssb2* results in severe germination and growth retardation. Furthermore, mutants were shown to have increased accumulation of mtDNA ectopic recombination products, and to play a role in mtRNA splicing and response to the phytohormone abscisic acid (Qian et al., 2022, 2021).

An additional plant ssDNA-binding protein group are ORGANELLAR ssDNA-BINDING PROTEINS (OSB). In Arabidopsis, four OSB proteins are found: two localise in mitochondria (OSB1 and OSB4), one in plastids (OSB2), and one can be found in both organelles (OSB3). Like SSB proteins, *osb1* and *osb4* mutants have distorted mtDNA stability and an increase in ectopic homologous recombination products (Zaegel *et al.,* 2007). SSB1/2 were proposed to compete with OSB1, thus re-reciprocally inhibiting each other. Furthermore, it was assumed that elevated repression of OSB1 leads to increased mtDNA replication (Qian *et al.,* 2022).

Finally, Whirly 2 (WHY2) and organellar DNA-binding protein 1 (ODB1) are two additional ssDNA-binding proteins, which are most abundant in Arabidopsis mitochondria. They form complex multimeric structures: WHY2 forms homotetramers, and, upon binding, ssDNA forms six-fold filaments of tetramers (24-mers) (Cappadocia *et al.,* 2012). ODB1 also appears to form homo-oligomeric complexes of ten or more monomers, as estimated by size-

exclusion chromatography (Janicka *et al.,* 2012). Furthermore, at least 80% of mtDNA is not covered in proteins, and ssDNA-binding proteins keep 10–15% of the genome in ssDNA configuration (Fuchs *et al.,* 2020).

Whirly proteins are plant-specific ssDNA binding proteins found in all flowering plants. Three proteins of this family are found in organelles: WHY1 and WHY3 in chloroplasts and WHY2 in mitochondria (Golin et al., 2020; Krause et al., 2005). Whirly proteins play a role in replication, transcription, RNA processing, and DNA repair and compaction (Krupinska *et al.,* 2022; Oetke *et al.,* 2022). Overexpression of the *WHY2* gene indicated a correlation to mtDNA replication efficiency (Cai *et al.,* 2015). Mutant *why2* plants are more sensitive to ciprofloxacin (CIP) treatment and show increased microhomology-mediated recombination (Cappadocia *et al.,* 2010). Furthermore, accumulation of U-turn rearrangements *via* microhomologies is observed in *why1why3* mutants, deficient in the two plastidial whirly proteins, and is suggested to result from a replication-dependent repair mechanism (Zampini *et al.,* 2015). In addition, *why2* mutant plants have mitochondria with a reduced abundance of crista and low matrix density (Golin *et al.,* 2020). Finally, Whirly proteins, might also have nuclear functions. WHY2 was proposed to be involved in regulating stress resistance toward abiotic and biotic stress. Expression of *WHY2* is closely linked to drought tolerance in tomato and tobacco (Meng et al., 2020; Zhao et al., 2018), and it was shown to play a role in pathogen defence via interaction with TGA1 (Wanamaker *et al.,* 2017).

As previously noted, WHY2 and ODB1 make up most of the mitochondrial ssDNA binding proteins, yet whether they compete or cooperate in recruiting different factors to ssDNA regions remains unknown. ODB1 shows sequence and structural similarities to the N-terminal domain of RAD52 (Janicka *et al.,* 2012). Arabidopsis has two orthologs of RAD52, each having two splicing variants: ODB1 found in mitochondria and nucleus, and ODB2 localised to chloroplast and nucleus. Both proteins are suggested to assist intrachromosomal recombination events in the nucleus (Samach *et al.,* 2011). Arabidopsis *odb1* mutants show increased mtDNA molecules generated by microhomology-mediated recombination events. It is assumed that ODB1 promotes the loading of RecA recombinase by displacement of SSB, similar to RAD52 initiating loading of RAD51 by displacing the ssDNA-binding protein replication protein A (RPA) (Janicka *et al.,* 2012; Roy *et al.,* 2009). However, the double mutant *odb1recA3* was shown to have a more severe developmental phenotype than the parents, indicating synergetic effects on mtDNA stability (Chevigny et al., 2020). Overexpression of ODB1 in the Arabidopsis nucleus leads to increased intra-chromosomal recombination and

synthesis-dependent strand-annealing (SDSA) HR repair rates. However, it repressed the single-strand annealing (SSA) pathway, contrary to what was expected from a RAD52 functional homolog (Samach *et al.,* 2018). Furthermore, ODB1 was shown to be also involved in mitochondrial splicing. The *odb1* mutants were shown to produce hydrolytic splicing variants via excision of two cis-spliced group II introns. ODB1-dependent introns, during splicing, release linear or circular molecules instead of lariats (Gualberto *et al.,* 2015).

2. RecA-like DNA recombinases

The central proteins of homologous recombination pathways are recombinases that catalyse strand exchange. The eucaryotic homolog of bacterial recombinase RecA (RAD51) from the RecA/RecBCD pathway is found in many organisms, including plants (Michel *et al.,* 2007; Rowan *et al.,* 2010). In bacteria, RecA forms nucleofilament structure on ssDNA and screens the genome for homologies by probing multiple heterologous sequences (Forget and Kowalczykowski, 2012; Ragunathan *et al.,* 2012). In Arabidopsis, there are five predicted homologues of RecA. However, two are loosely connected: AT3G32920 is a pseudogene, and DRT100 has weak similarity and can only partially restore bacterial *recA* phenotype (Pang et al., 1992; Shedge et al., 2007). Another three homologs have been shown to have a role in recombination-dependant repair similar to bacterial RecA and can be found in different organellar compartments: RECA1 and RECA3 localise to chloroplasts and mitochondria, respectively, and RECA2 localises to both (Shedge et al., 2007; Miller-Messmer et al., 2012). Mutant plants of all RecA-like genes had altered structural forms of organellar genomes, indicating their role in homologous recombination. Loss of RECA1 leads to reduced cpDNA copy numbers and the accumulation of ssDNA in the chloroplast (Rowan *et al.,* 2010). The *recA2* and *recA3* mutations in Arabidopsis result in different phenotypes: *recA2* is lethal at the early seedling stage, while *recA3* first transformants most frequently look like wild-type plants. Yet, *recA3* may develop a more severe phenotype in later generations. RECA3 seemed to be involved in a different recombination-dependent repair pathway than RECA2 (Miller-Messmer *et al.,* 2012). While RECA2 is present in all land plants, RECA3 seems to have emerged in evolution after the development of flowering plants.

3. DNA helicases

In Arabidopsis mitochondria, at least two DNA helicases are involved in homologous recombination pathways: RADA and RECG1. Both can be found in chloroplasts and mitochondria. The Arabidopsis *recG1* mutants showed no developmental phenotype (Wallet et al., 2015), while the *radA* mutation has severe effects on plant development: deformed

leaves, short roots and very small siliques containing mostly aborted seeds. Furthermore, double mutants showed the epistatic relationship of genes with no additive effect observed compared to a single *radA* mutation phenotype. Phenotypic analysis suggested that proteins act in the same pathways, and RADA activity is dominant in branch migration (Chevigny et al., 2022; Wallet et al., 2015). In addition, plants do not encode any homolog of the bacterial RuvAB helicase, which is believed to be the predominant branch migration pathway, recruiting the RuvC resolvase. Thus, mildly affected *recG1* mutants could also suggest additional pathways for the maturation of recombination intermediates in plant organelles.

RECG1 is a plant ortholog of bacterial RecG, and homologs can be found in all sequenced plant genomes, including vascular plants, mosses, and algae. Bacterial RecG is a multifunctional translocase that can guide RecA-promoted strand exchange, rescue stalled replication forks and suppress pathological DNA replication (Manosas et al., 2013; McGlynn and Lloyd, 2002; Rudolph et al., 2013, 2010a; Vincent et al., 1996). RecG can process different DNA substrates: three-strand junctions, D-loops, and R-loops (Rudolph et al., 2010; Vincent et al., 1996; Whitby and Lloyd, 1998). The loss of *recG1* in Arabidopsis reduces recombination-dependent repair of DSBs; however, *recG1* leads to mitochondrial genome instability caused by increased asymmetric ectopic homologous recombination involving intermediate-size repeats. Furthermore, the crossing of *recG1* and *recA3* mutant plants showed the synergy of the two genes, as the progeny had a much more severe developmental phenotype than their parents (Wallet *et al.,* 2015). Such an additive effect on phenotype could be due to partial functional similarities of the proteins in recovering stalled replication forks (Cox, 2007).

RADA is a plant ortholog of bacterial RadA/Sms, a DNA helicase that extends D-loops, allowing the recruitment of the fourth DNA strand to form a Holliday Junction (Cooper and Lovett, 2016; West, 1997; Whitby et al., 1994). In bacteria, RadA interacts with RecA filaments; similarly, in Arabidopsis, RADA was shown to interact with RECA2, but not with RECA3. This reconfirms the apparent existence of two independent RECA pathways; even though the double mutant *recA2 recA3* could not be obtained, the double mutant *radA recA3* was shown to be lethal(Chevigny et al., 2022; Miller-Messmer et al., 2012). In Arabidopsis, loss of *radA* leads to massive mtDNA reshuffling driven by the formation of sub-genomes by recombination involving small direct repeats, which replicate autonomously and uncoordinated from the remaining mtDNA, as also observed in *recG1* and *osb1* mutants (Chevigny *et al.,* 2022; Wallet *et al.,* 2015; Zaegel *et al.,* 2007). In addition, *radA* mutations were shown to trigger retrograde signaling that alters cell cycle progression, concluded from increased epidermal cell size, nuclear ploidy and reduced mitotic activity (Chevigny *et al.,* 2022).

4. Mismatch recognition protein MSH1

In addition to the proteins mentioned previously, disruption of the MutS HOMOLOGUE 1 gene (*MSH1*) also increases ectopic recombination involving small repeats. The MSH1 protein is also targeted to both mitochondria and chloroplasts, and the frequency of mitochondrial and plastid sequence variants in *msh1* plants is approximately 10-fold to 1,000-fold higher than in WT plants (Wu *et al.,* 2020; Xu *et al.,* 2011). MutS is a protein that was initially found in prokaryotes, involved in mismatch repair (MMR). It recognises and binds to mismatches on double-strand DNA, forming a complex with MutL that recruits the endonuclease MutH to create a lesion. Exonuclease then extends lesions to remove a mismatch, and DNA polymerase synthesises the correct DNA sequence (Schofield and Hsieh, 2003; Spampinato *et al.,* 2009). In Arabidopsis, other MMR proteins were not identified in organelles. Yet, MSH1 has a C-terminal GIY-YIG endonuclease domain that could potentially perform the MutH function (Abdelnoor et al., 2006). This led to the hypothesis that MSH1 could recognise mismatches and create DSB, initiating accurate repair by homologous recombination (Christensen, 2018, 2014). However, the GIY-YIG domain exhibited no endonuclease activity by itself *in vitro*. Thus, the model needs further inspection as endonuclease activity may require the whole protein or interaction with additional factors (Fukui *et al.,* 2018).

Extensive studies of MSH1 suggested that its primary role is recombination surveillance rather than correcting mismatches (Abdelnoor et al., 2006; Morley et al., 2019; Shedge et al., 2007; Wu et al., 2020). Arabidopsis *msh1* mutants have dramatic structural reorganisation of organelle genomes via repeat-mediated recombination (Abdelnoor et al., 2003; Davila et al., 2011; Xu et al., 2011), accompanied by phenotypic variation; they develop variegated leaves that could be traced by maternal inheritance, with some exhibiting a chlorotic phenotype due to defective chloroplasts (Xu *et al.,* 2011). The mutants undergo widespread epigenetic modifications in the nucleus via methylation (Kundariya et al., 2022; Virdi et al., 2015; Xu et al., 2012). Furthermore, it was proposed that MSH1 suppresses ectopic recombination between non-identical repeats by detecting mismatches and destabilising recombination intermediates (Xu *et al.,* 2011; Zou *et al.,* 2022).

5. Mitochondrial exonuclease OEX1

In plant organelles, the proteins involved in DNA recession at the onset of DSB repair by homologous recombination remain to be identified. In her PhD thesis, Deborah Schatz-Daas presented the OEX1 exonuclease that might be involved in DNA recession. OEX1

seems to derive from the 5'-3' exonuclease domain of bacterial DNA polymerase I. She showed that OEX1 has 5'-3' exonuclease and flap-endonuclease activities, and can also rapidly process the RNA on RNA-DNA hybrids. These activities suggested that OEX1 could be also involved not just in DNA repair by HR and base excision repair (BER), but also in the maturation of Okazaki fragments during replication and the elimination of R-loops, which compromise genome stability. Arabidopsis mutants of *oex1* present severe developmental phenotypes similar to those observed in other mutants of the plant mitochondrial HR pathway, such as *radA* mutants (Schatz-Daas, 2021).

6. DNA synthesis by Polymerases

In Arabidopsis mitochondria, two DNA polymerases, DNA Polymerase IA (Pol1IA) and DNA Polymerase IB (Pol1IB), have been identified, both dual-targeted to mitochondria and plastids. They are phylogenetically related to the eubacterial DNA polymerase I (PolI) (Shutt and Gray, 2006). Arabidopsis *pol1A* or *pol1B* mutants do not develop any remarkable developmental phenotype; just *pol1b* mutants have been reported to have reduced growth rates, but this observation remained to be confirmed. Attempts to generate double mutants proved unsuccessful, suggesting lethality and redundant functions on the replication of the organellar genomes (Parent *et al.,* 2011).

Both single mutants showed reduced mitochondrial DNA copy number, suggesting their role in mtDNA replication. However, POL1B seems to be specifically involved in DNA repair, as *pol1B* mutants are sensitive to the genotoxic agent ciprofloxacin, which introduces DSBs in the organellar genomes (Cupp and Nielsen, 2013; Parent et al., 2011). In agreement with this assumption, a cross between *pol1B* and a plastid Whirly mutant line (*why1 why3*) yielded plants with high levels of plastid DNA rearrangements and severe growth defects, indicating impairments in plastid DNA repair pathways. This correlated with an increase in repair by error-prone MHMR pathways (Parent *et al.,* 2011).

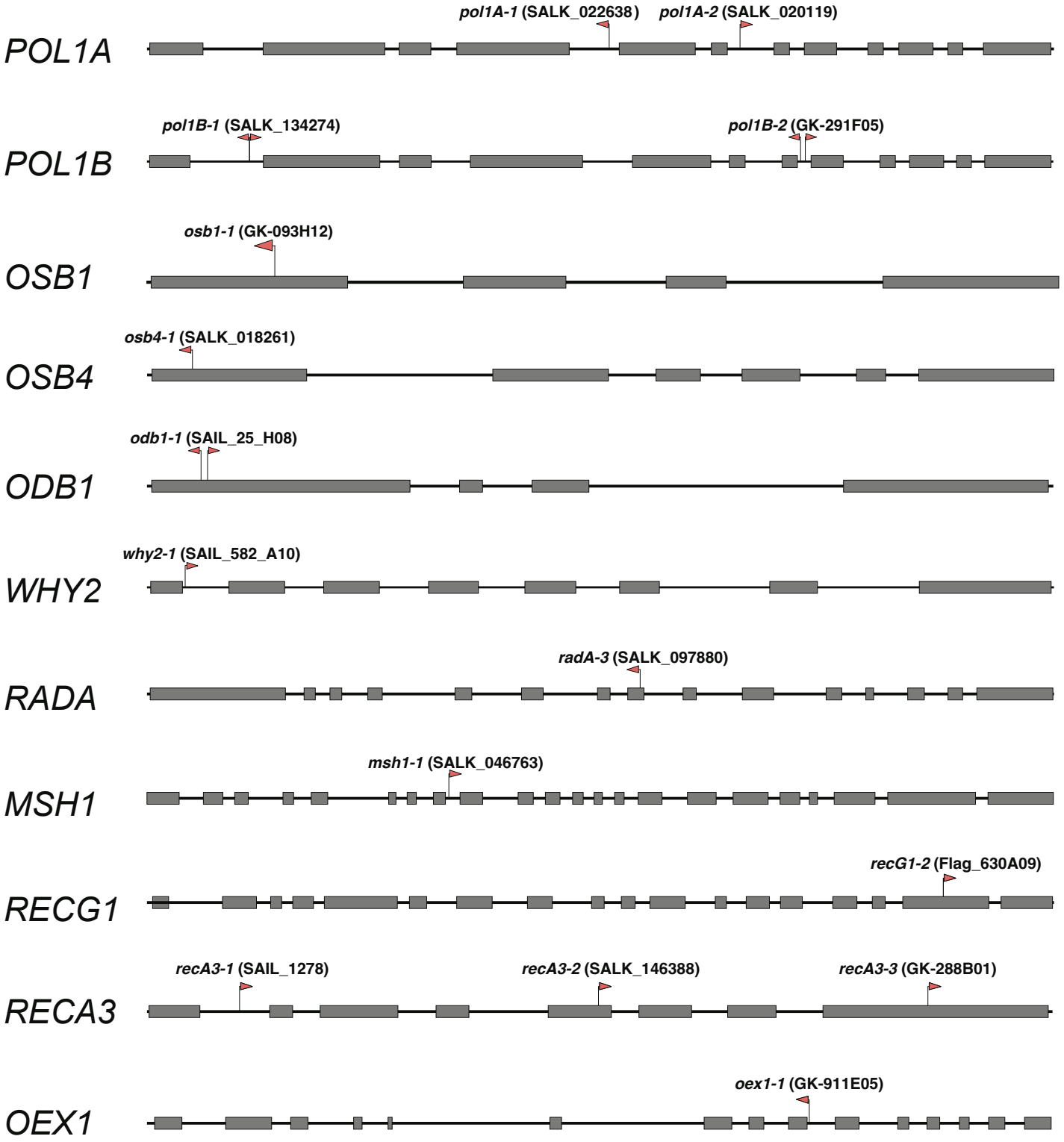


Figure III. 1 Gene models of key proteins involved in plant mitochondrial homologous recombination, which were studied in this project. Flags indicate the location and orientation (left-border) of T-DNA insertions in the mutant lines, with names and codes from the stock-centers specified.

Results

1. Mitochondria HR T-DNA mutant lines

A. thaliana accession Columbia 0 (Col-0) was used as the model organism in our studies, to study homologous recombination repair in organelles. The T-DNA insertion mutant lines described previously have been retrieved from the Nottingham Arabidopsis Stock Centre (*odb1-1*: SAIL_25_H08, *why2-1*: SAIL_582_A10, *osb1-1*: GK-093H12, *osb4-1*: SALK_018261, *radA-3*: SALK*_097880, msh1:* SALK_046763*, pol1B-1:* SALK_134274*, pol1B-2:* GK-291F05*, pol1A-1:* SALK_022638, *pol1A-2*: SALK_020119, *recA*3-1: SAIL_1278, *recA*3-2: SALK_146388, *recA*3-3: GK-288B01, *oex*1-1: GK-911E05). An *osb1-1 osb4-1* double mutant was obtained via crossing of the corresponding lines. *recG1-2*: Flag_630A09 line, in the accession Wassilewskija (Ws) background, was retrieved from the Versailles Arabidopsis Stock Centre, and subsequently the mutation was introgressed into Col-0. The position of all T-DNA insertions is presented in Fig. III.1. As described in the Methods section, mutants were genotyped by PCR, using primers flanking the insertion and a primer specific for the left border of the T-DNA (Appendix 1). Some of the mutants had deleterious developmental phenotypes (*recA3, osb1 osb4, oex1, msh1, recG1* and *radA*), yet all of them remained viable until maturity (Fig. III.2). This suggests the presence of alternative repair pathways for given genes, as complete loss of mitochondrial HR is apparently lethal, as indicated by the lethality of the *recA2 recA3* mutation. Mutants *recA3* and *osb1 osb4* displayed pleiotropic phenotypes, most likely linked to the creation and differential segregation of alternative major mitotypes resulting from recombination. Plants *recA3-1* and *recA3-3* rarely developed any visible phenotype, while *recA3-2* plants frequently developed deleterious phenotypes. This might be because *recA3-1* and *recA3-3* have T-DNA insertions in non-coding sequences (3'-UTR and first exon, respectively) and might be leaky mutants, while *recA3-2* has the T-DNA in the fifth exon and is most probably a true knock-out. Furthermore, *oex1, radA* and *osb1 osb4* mutant plants have little fertility and yielded many aborted or unviable seeds (Fig.III.3).

2. Illumina sequencing of mutant lines containing mitoTALEN constructs

To study how a DSB is repaired in mutants lacking specific factors involved in the mitochondrial HR pathway, the corresponding mutant lines were transformed with the TAL1 and TAL2 constructions, and the DNA from T1 plants was sequenced, as described in Chapter II for WT plants. However, the transformation of *radA* and *oex1* was difficult as homozygous mutants mostly give aborted seeds. Thus, heterozygous mutants *radA* and *oex1* were transformed. Several homozygous r*adA* plants containing the TAL constructs could be

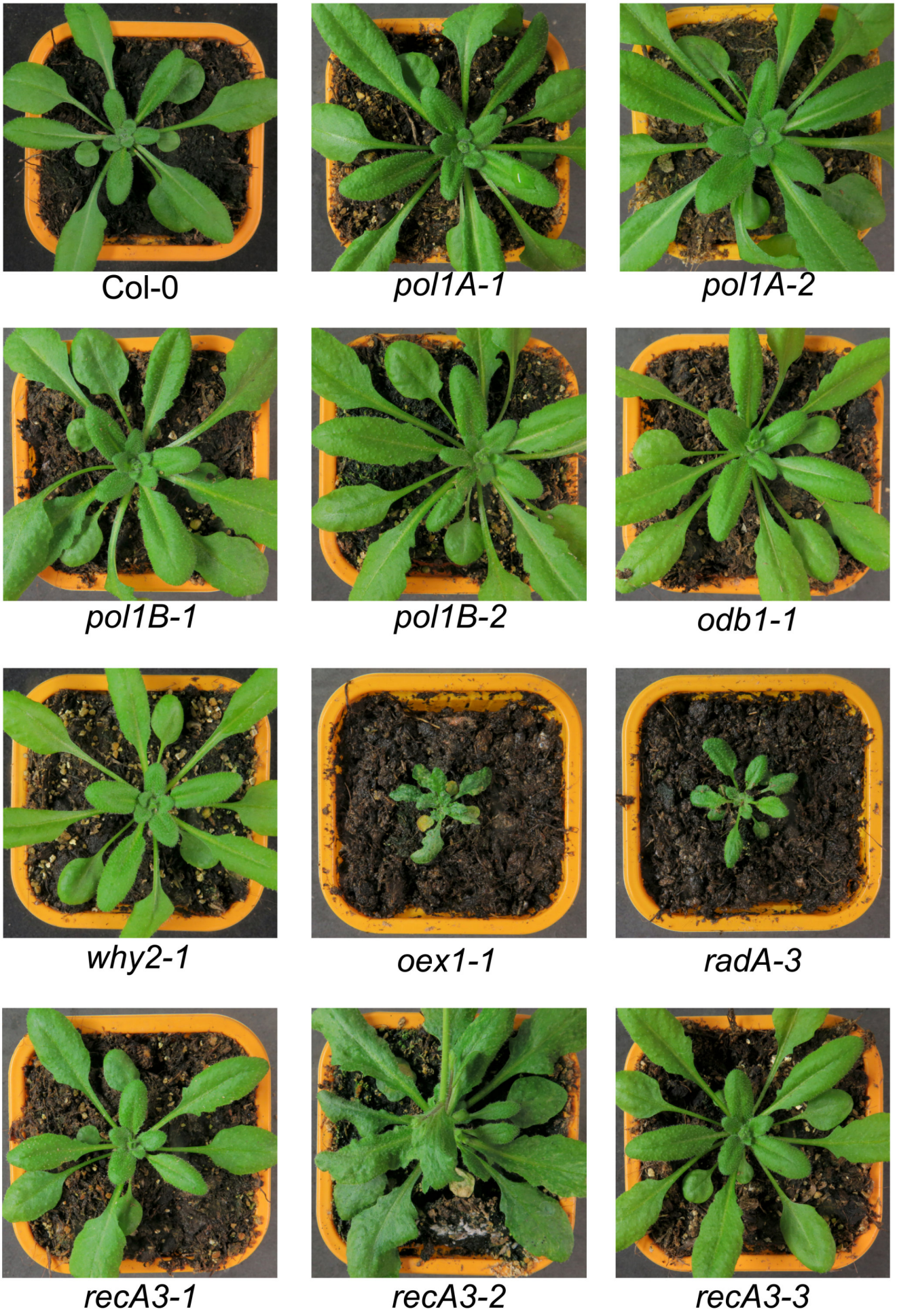


Figure III. 2 Pictures of 5-week-old Arabidopsis plants from the indicated mutant lines.

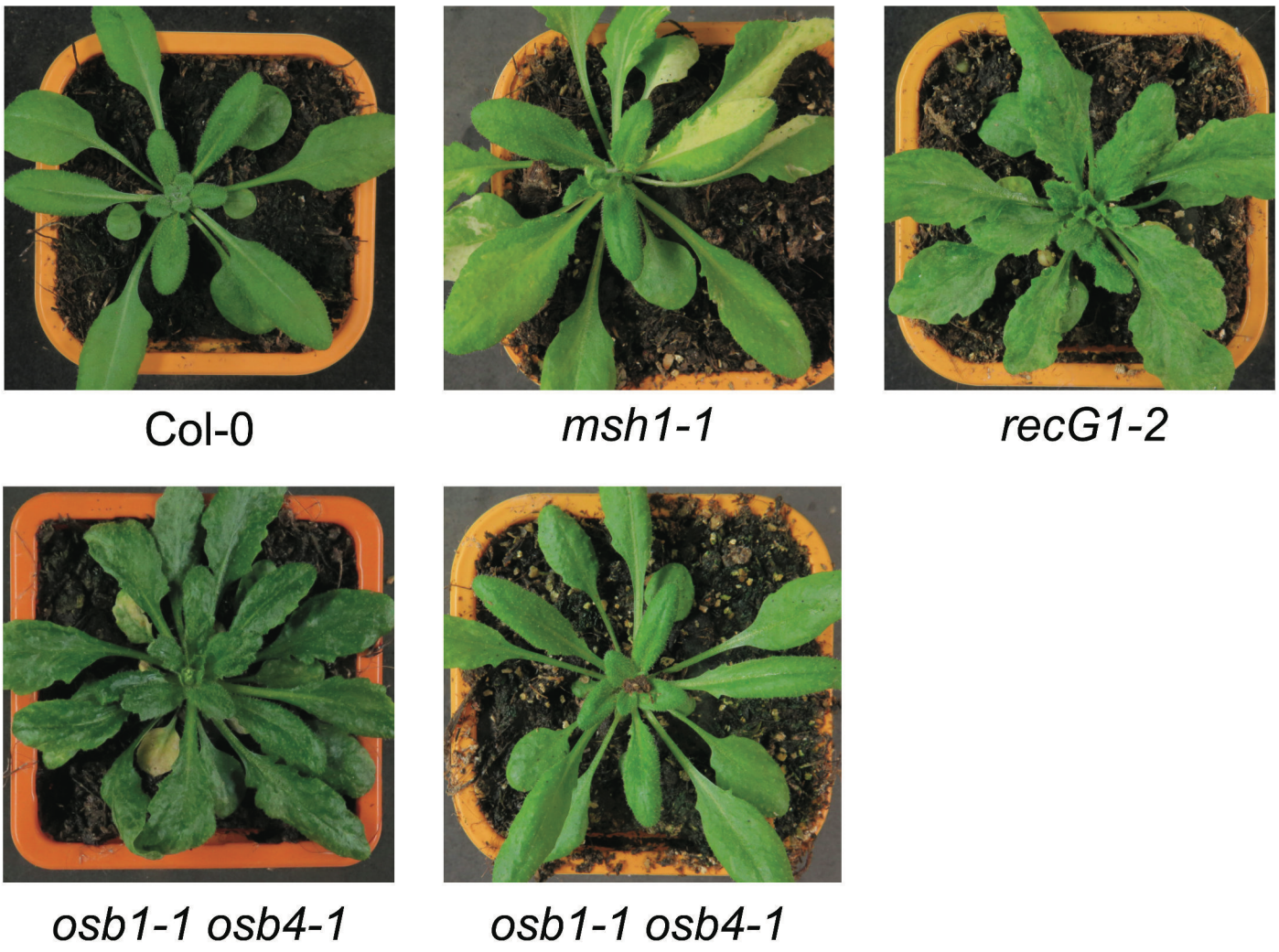


Figure III. 2 Pictures of 5-week-old Arabidopsis plants from the indicated mutant lines. (continue)

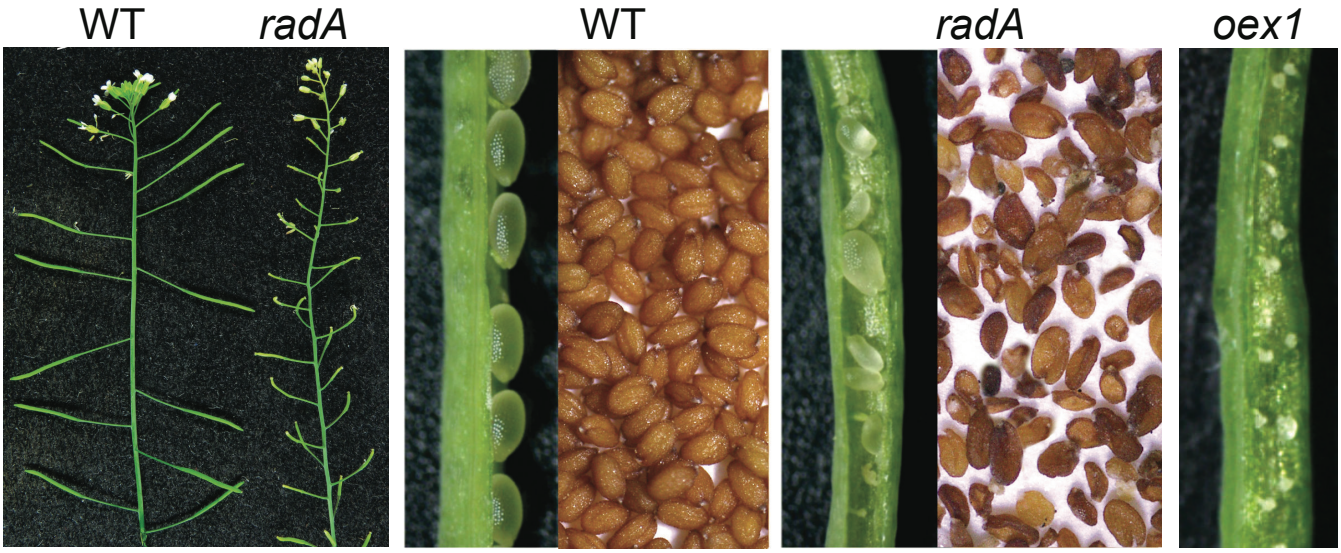


Figure III. 3 Small siliques, containing mostly aborted seeds, in radA and oex1 mutant plants

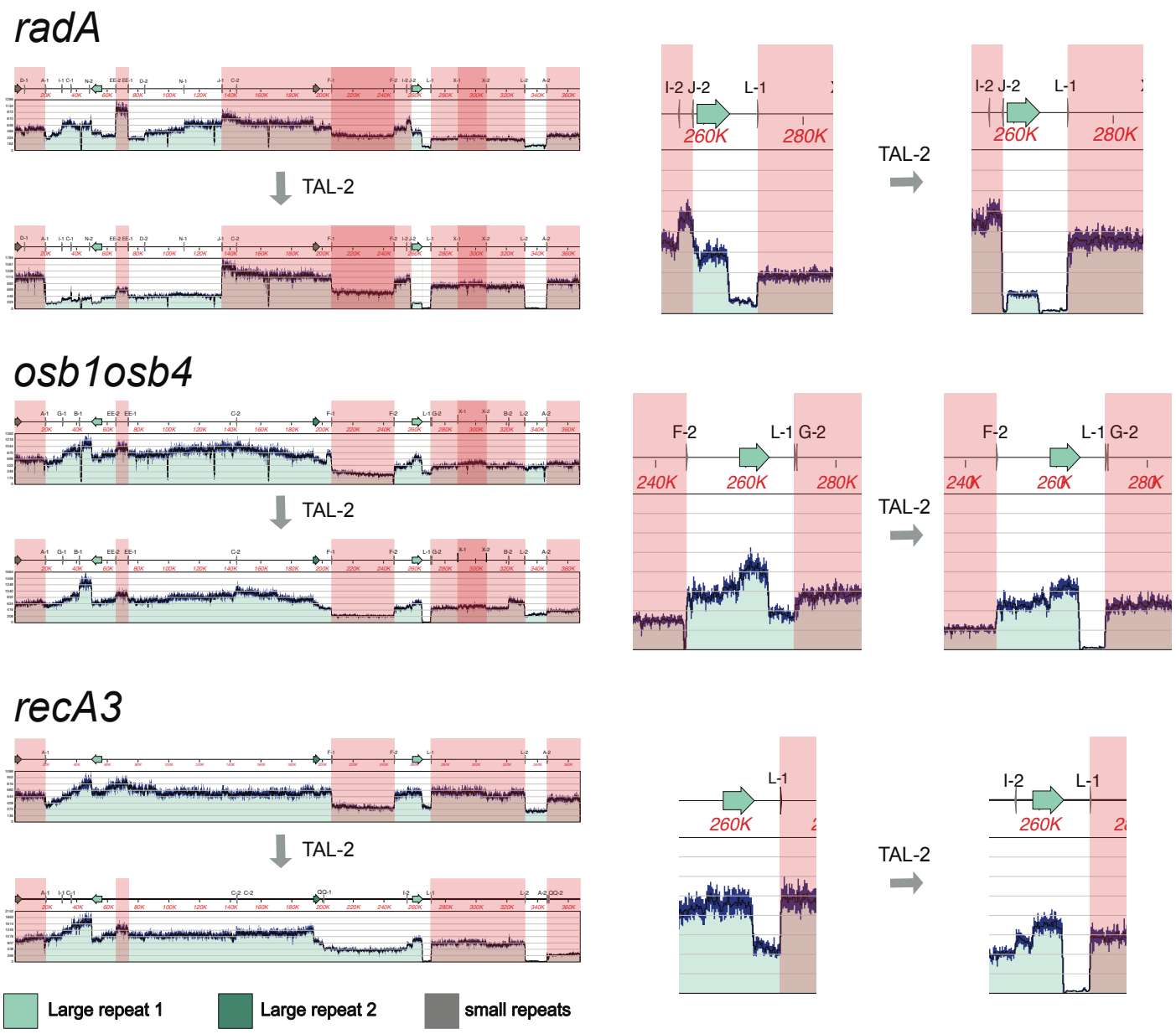


Figure III. 4 Coverage of Illumina sequencing reads from Arabidopsis mutants radA, osb1osb4 and recA3, before and after introduction of the TAL-2 construction. Reads were mapped to the WT Col-0 mtDNA with the Bowtie2 aligner implemented in the MacVector Assembler application. Red boxes indicate regions flanked by direct repeats supposedly forming independently replicating sub-genomes.

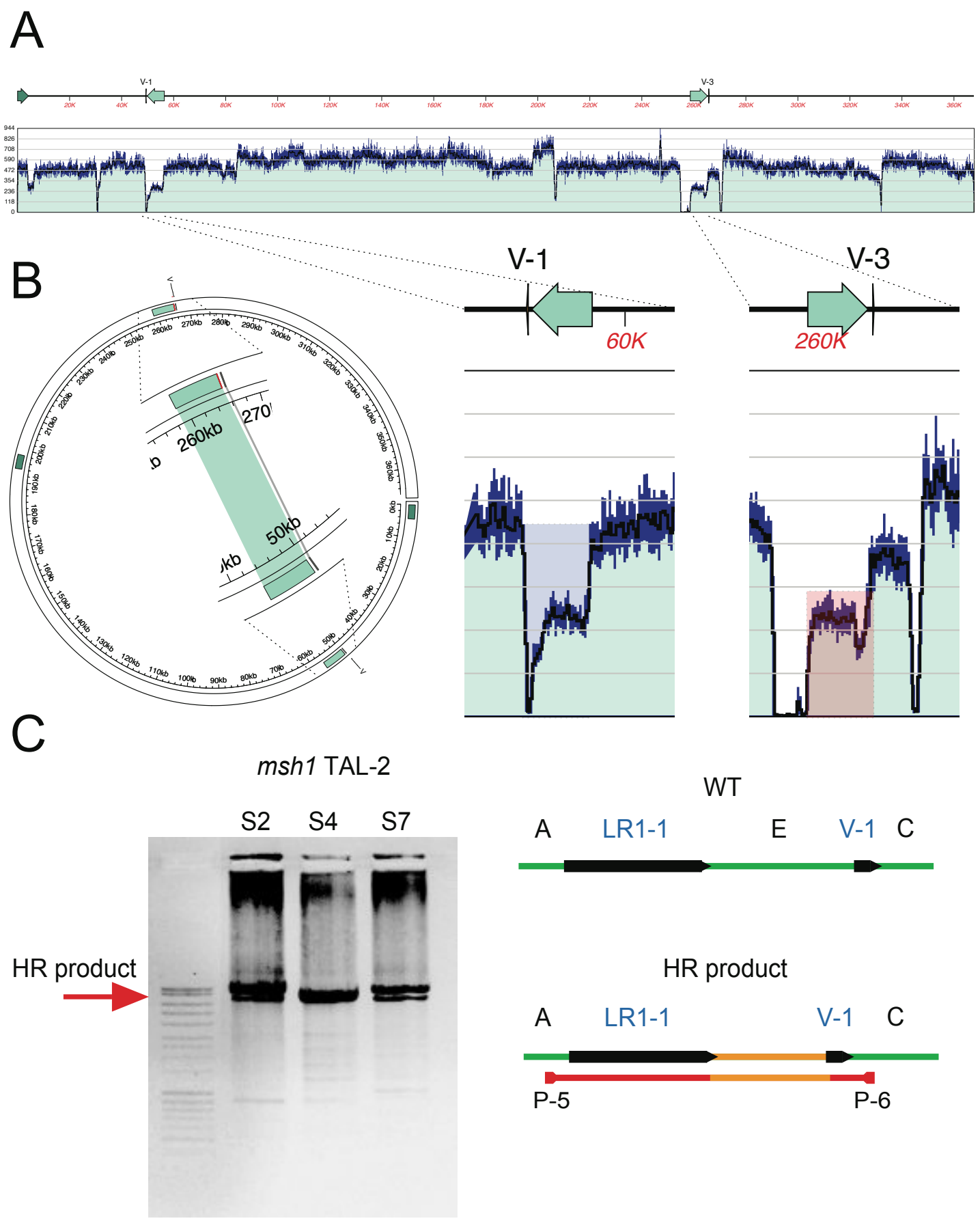


Figure III. 5 Structural rearrangements of the msh1 mitochondrial genome compared to WT Col-0. (A) Coverage map of Illumina sequencing reads from msh1 mapped to the Col-0 mtDNA. (B) Circular representation of the Col-0 mtDNA indicating homologous recombination involving the two pairs of repeats LR1 and G. A red box indicates the region comprising LR1-2 and V-3 that by recombination replaced the region in the blue box, comprising repeats LR1-1 and V-1. (C) Electrophoresis analysis on 0.7 % agarose gel (left) of PCR products and schematics (right) of PCR amplification of the region in blue, using P-5 and P-6 primers. Sanger sequence of the PCR products confirmed the model.

obtained this way, as T1 transformants. However, no *oex1* homozygous mutants were obtained containing the mitoTALENs. To produce *oex1* $^{-/-}$ and additiona *radA*$^{-/-}$ plants with mitoTALEN, pollen of *radA*$^{-/+}$ TAL-1/2 or *oex1*$^{-/+}$ TAL-1/2 was used as a mitoTALEN donor to introgress into respective heterozygous mutants, and finally screening for homozygous F1 plants with TALEN construct.

As for WT plants, the DNA of the selected plants was enriched for mtDNA by partial purification of mitochondria. Libraries were prepared as described in Material and Methods and analysed by Illumina sequencing. NGS data was first analysed using MacVector Assembler tools. As described previously, chloroplast reads were first removed by mapping to its genome, and the remaining were mapped to the Col-0 mitochondrial genome.

All repair pathways have been identified from changes in coverage, as mentioned in Chapter II, yet at different frequencies according to the mutants (Appendix 2&3). Initial analysis revealed no difference in DSB repair between Col-0, *pol1A*, *pol1B* and *oex1* mutants, suggesting that the predominant HR-dependent repair pathway is not significantly compromised in these mutants. Unlike others, *oex1* also shows more complex genome rearrangement that is not linked to DNA repair. Furthermore, it's important to note that *oex1* mitoTALEN plants were obtained through crossing, rather than directly by *Agrobacterium* transformation. Thus, several samples showed strong heteroplasmy, with WT mtDNA, probably because of lower efficiency of mitoTALENs when introduced by crossing, as discussed in Chapter II.5. Conversely, in *odb1* and *why2,* there was an increased preference for repair following alternative repair pathways, namely SSA and MHMR. In contrast, and consistent with plant phenotypes, Arabidopsis mutants *recA3, osb1osb4, oex1, msh1, recG1* and *radA* had a strong structural reorganisation of the mtDNA via repeated sequences that directly or indirectly affected plant development. In *recA3, osb1 osb4 and radA*, there was no evidence of the HR-dependent repair involving repeats V, G, or Q, which was previously shown in Chapter II to be predominant in WT plants (Fig. III.4). However, these regions indicated in red in the figure, flanked by direct repeats, seems to form episomes that pre-existed TALEN activity. Upon DSB, the damaged DNA was apparently segregated-out, by reconstituting viable genomes from autonomously replicating episomes. Thus, the molecules containing a DSB were removed, without involvement of repair pathways.

Furthermore, the mapping of *msh1* sequences to the WT mtDNA revealed that already in the mutant line, before transformation with the mitoTALEN constructs, there was a recombination event similar to the one between Large repeat 1 and repeat V, described in Chapter II, but in opposite orientation, resulting in the replacement of the *atp6-1* gene by *atp6-2* (Fig. III.5A). To prove this hypothesis, the region comprising the *atp6-1* gene was amplified by PCR and sequenced. Amplification from WT total DNA was not successful,

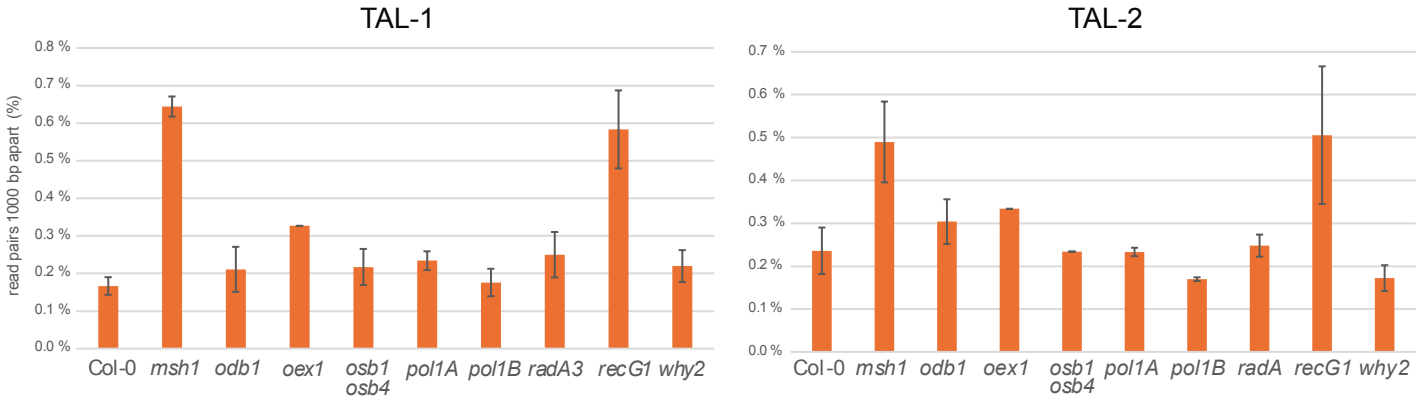


Figure III. 6 Proportion of Illumina read pairs more than 1000 bp apart, indicating structural variants of the mitochondrial genome, in the mutants transformed with either TAL-1 or TAL-2 constructs. Calculated relative to the total number of reads that have been mapped to the WT Col-0 mitochondrial genome.

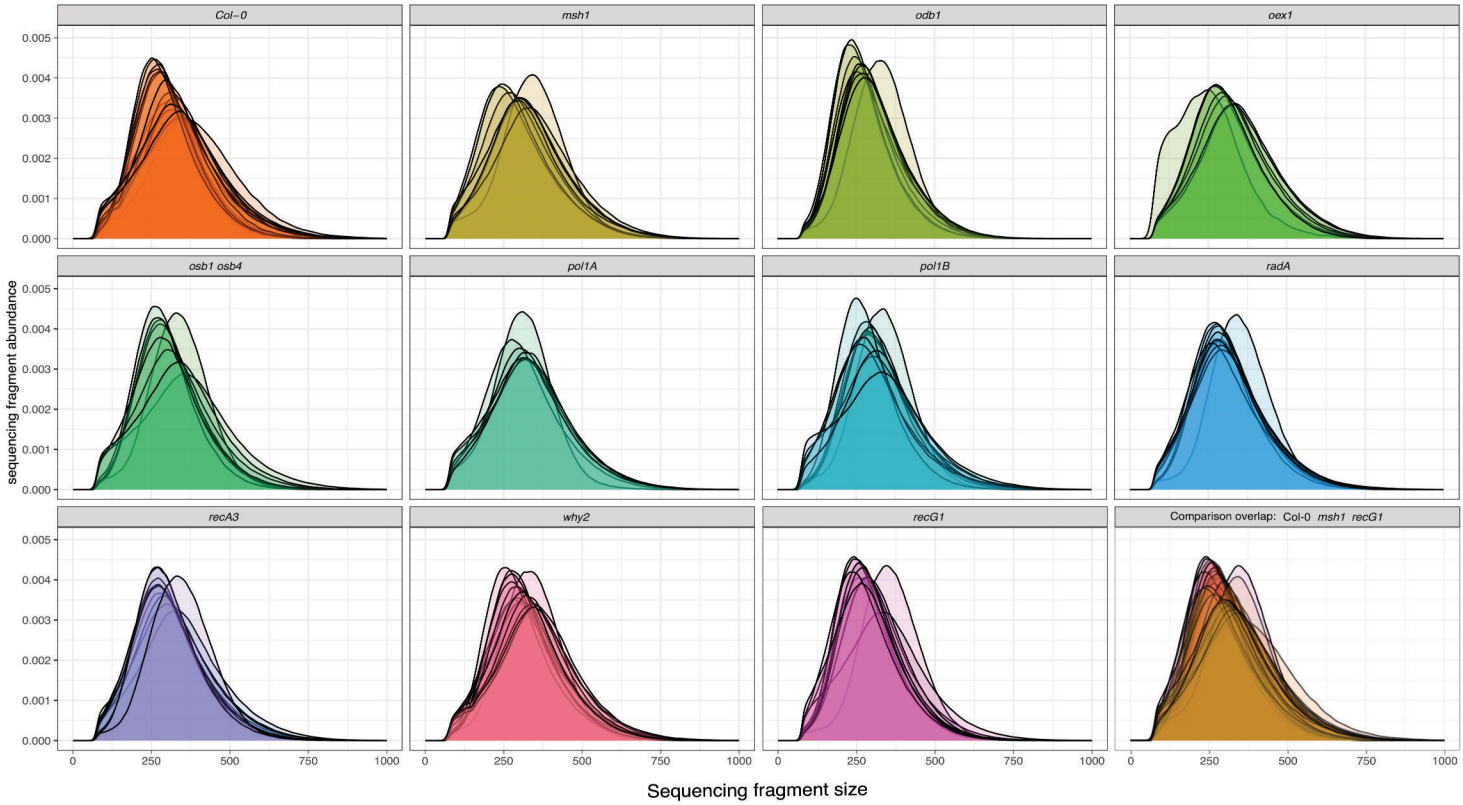


Figure III. 7 Size distribution of fragments sequenced by Illumina, from different T-DNA insertion mutants transformed with the mitoTALEN constructs. Fragment sizes are assumed to be below 1000 bp, according to read pair insert size, as structural variants of the mtDNA result in a greater increase of distance between read pairs.

confirming that such configuration of the mtDNA is virtually absent in WT plants. However, the amplification from *msh1* worked (Fig. III.5B), and sequencing of the amplified fragment confirmed the replacement of *atp6-1* by the *atp6-2,* the region comprised between Large repeat LR1-2 and intermediate size repeat V-3, replaced by the region between LR1-1 and V- 1 (Fig. III.5C).

3. Bioinformatic analysis of NGS data

To examine DSB repair and track genome structural variants more closely, NGS data was analysed based on read-pairs insert sizes and read soft-clips, with developed scripts described in Chapter II. These scripts allowed us to assess the relative importance of HR-dependent repair involving small repeats and repair by MHMR, as described previously. Analysis of rea-pairs allowed us to determine the frequencies of recombination of individual repeats, and they are represented in circos plots as links between the repeats (Appendix 3). For a more systematic analysis, samples were combined according to genotype and the mitoTALEN construct (TAL-1 or TAL-2). The abundance of read-pairs further apart than 1000 bp was two times higher in *msh1* and *recG1* as compared to WT, while in other mutants it was much less significant (Fig. III.6). This suggested that in *msh1* and *recG1* mutants, homologous recombination is significantly increased, resulting in a highly altered structure of the mtDNA. However, results should be interpreted critically, as the detection of structural variants is highly dependent on library size, and it varies from sample to sample (Fig. III.7). Advantageously, library size variation in *msh1* and *recG1* samples was not higher than in other mutants. Thus, it should not be the sole contributor to such a significant increase in the number of homology recombination events found.

To identify which intermediate-size repeats are most frequently involved in HR-dependent repair in the different mutants, all read-pairs that were not further apart than 1000 bp were removed from mapping to the Col-0 mtDNA. The remaining pairs are indicative of HR involving the corresponding repeat in proximity (Fig. III.8). The analysis reveals that repeats V, Q, H, L and G, located next to the cut site, are particularly active. Furthermore, in *msh1* and *recG1* mutants, the number of significantly active HR sites is higher (26-43 and 24-29, respectively) compared to WT (20-22), resulting in increased mtDNA structural variation. To compare how DSBs are repaired in different mutants, the ratio of reads in recombined and WT configurations was calculated (Fig. III.9). In *pol1A, pol1B* and *why2* mutants, as in WT, most recombined forms were formed via recombination involving repeats V and G for mutants transformed with TAL-1, and repeat G for mutants transformed with TAL-2. However, in other mutants (*oex1, radA, osb1osb4, recG1, msh1* and *recA3*) recombination involving the pair of

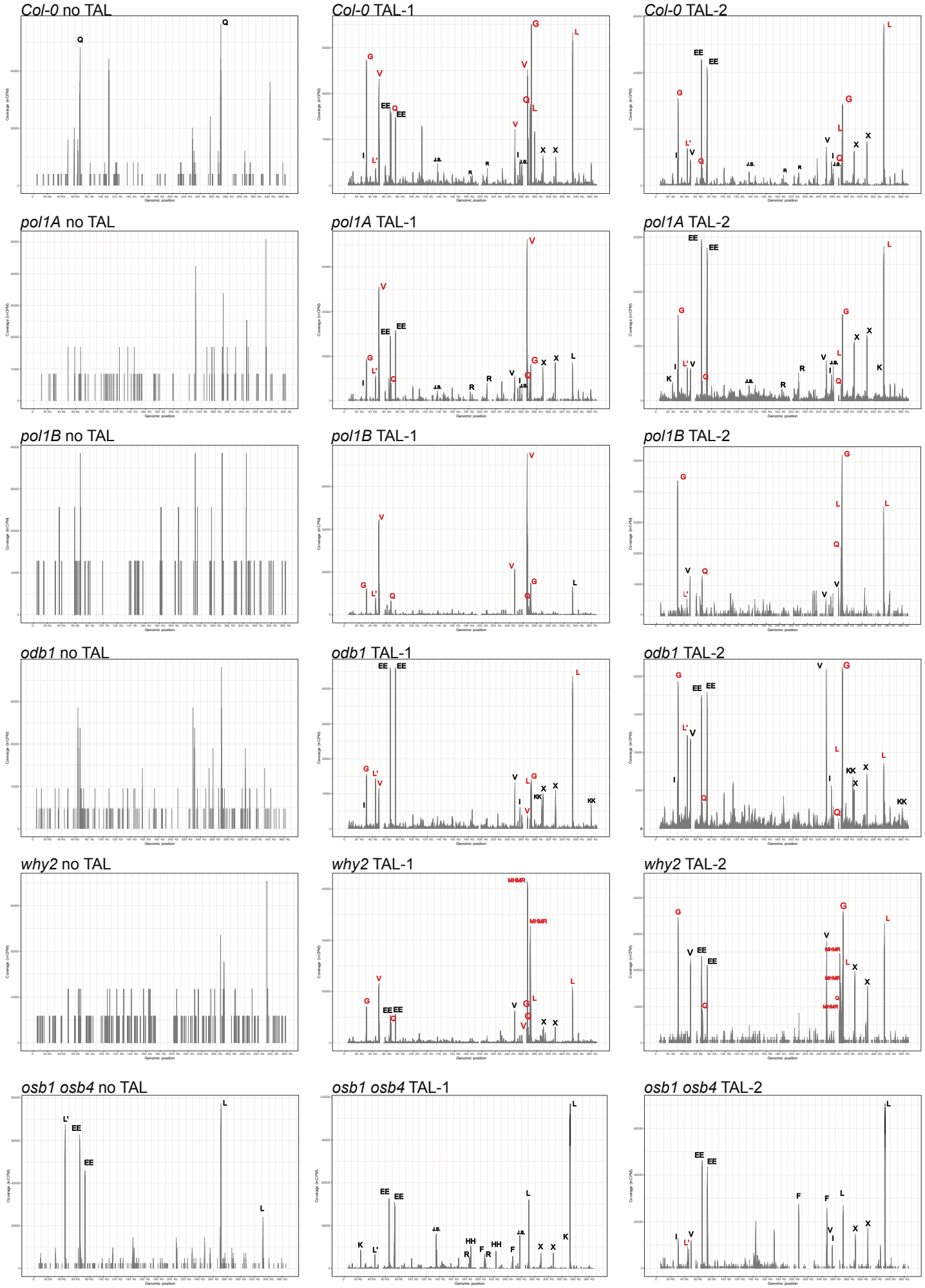


Figure III. 8 Combined coverage maps of read pairs mapping more than 1000 bp apart to the Col-0 mtDNA, for T-DNA mutant lines both before and after transformation with either the TAL-1 or TAL-2 mitoTALEN constructs. Peaks on the maps indicate clustering of reads next to annotated repeats of the mtDNA. Annotations in red indicate peaks corresponding to repeats next to the mitoTALEN cut site that are assumed to play major roles in the HR-dependent repair of the DSB.

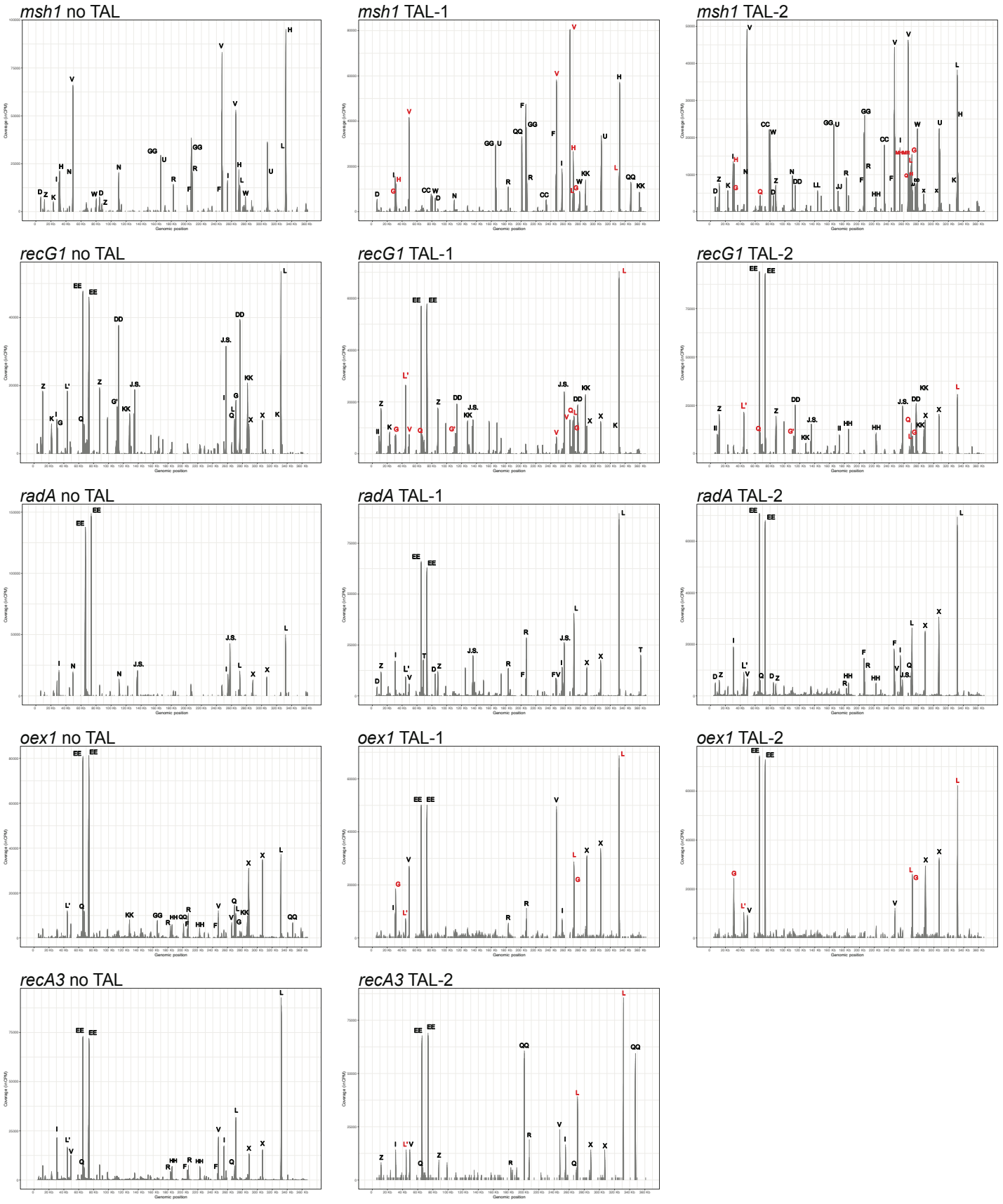


Figure III. 8 (continue) Combined coverage maps of read pairs mapping more than 1000 bp apart to the Col-0 mtDNA, for T-DNA mutant lines both before and after transformation with either the TAL-1 or TAL-2 mitoTALEN constructs. Peaks on the maps indicate clustering of reads next to annotated repeats of the mtDNA. Annotations in red indicate peaks corresponding to repeats next to the mitoTALEN cut site that are assumed to play major roles in the HR-dependent repair of the DSB.

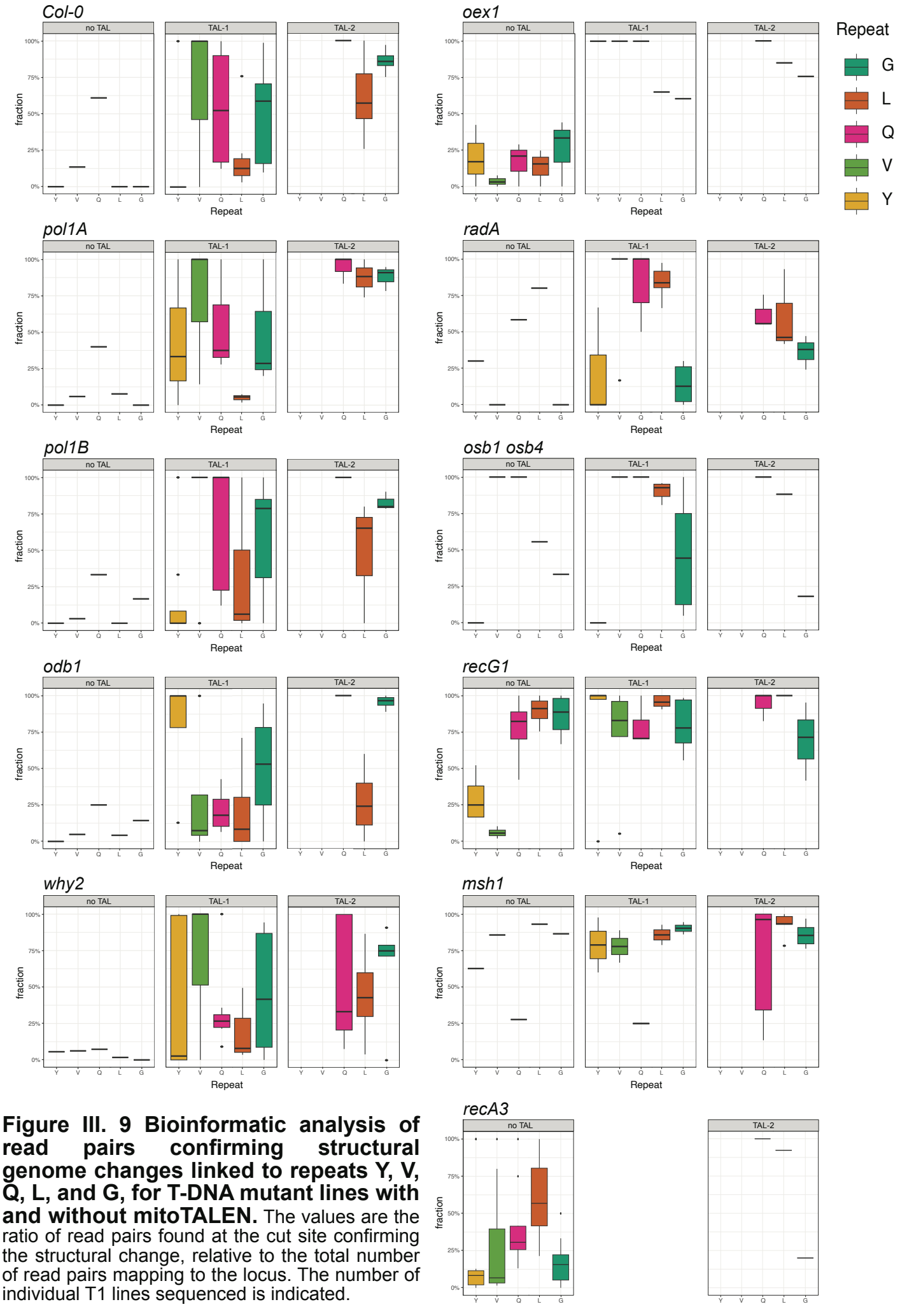


Figure III. 9 Bioinformatic analysis of read pairs confirming structural genome changes linked to repeats Y, V, Q, L, and G, for T-DNA mutant lines with and without mitoTALEN. The values are the ratio of read pairs found at the cut site confirming the structural change, relative to the total number of read pairs mapping to the locus. The number of individual T1 lines sequenced is indicated.

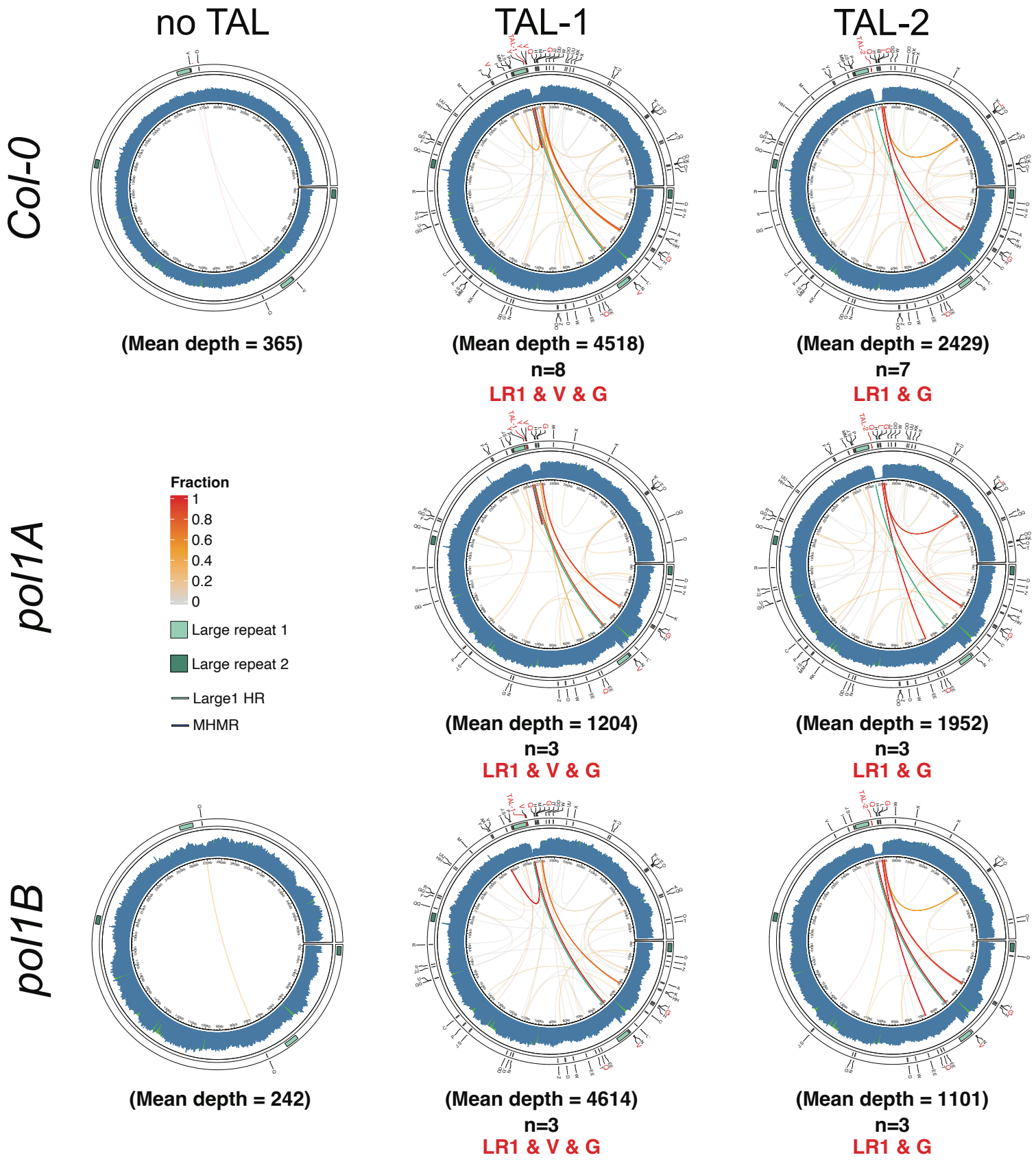


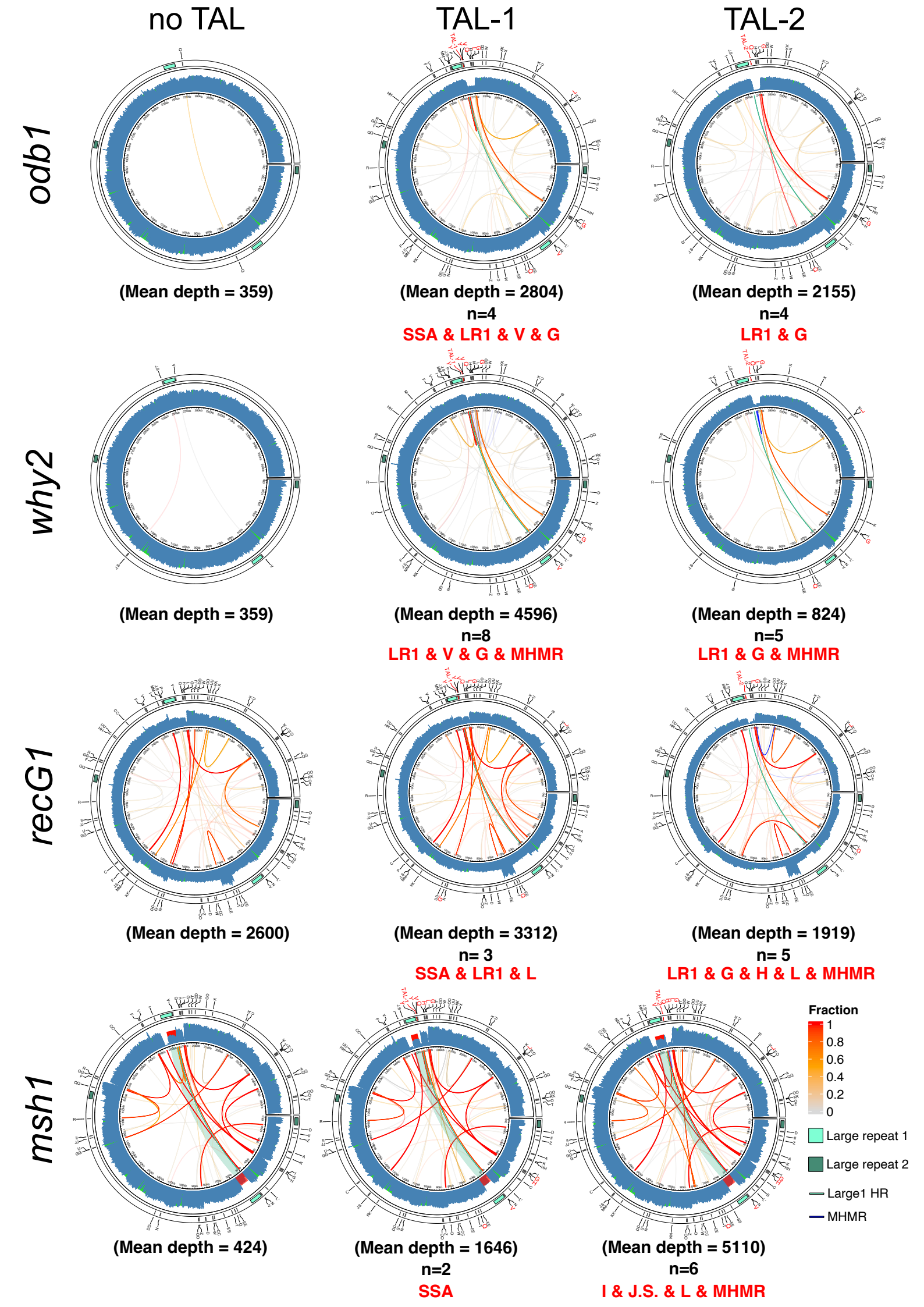
Figure III. 10 Circular representation of coverage maps of Illumina sequencing reads, mapped to the WT Col-0 mtDNA using the BWA aligner. Drops in coverage indicated in green represent reads removed from the analysis by mapping to the chloroplast genome, corresponding to regions of the mtDNA that are chloroplast sequence insertions. Links represent recombination between the indicated repeats, colored according to the ratio of read pairs confirming the recombination, relative to the total number of read pairs mapping to the locus. Annotations in red indicate the mitoTALEN cut site and the repeats in proximity involved in the DSB repair, resulting in the reconstitution of a viable genome.

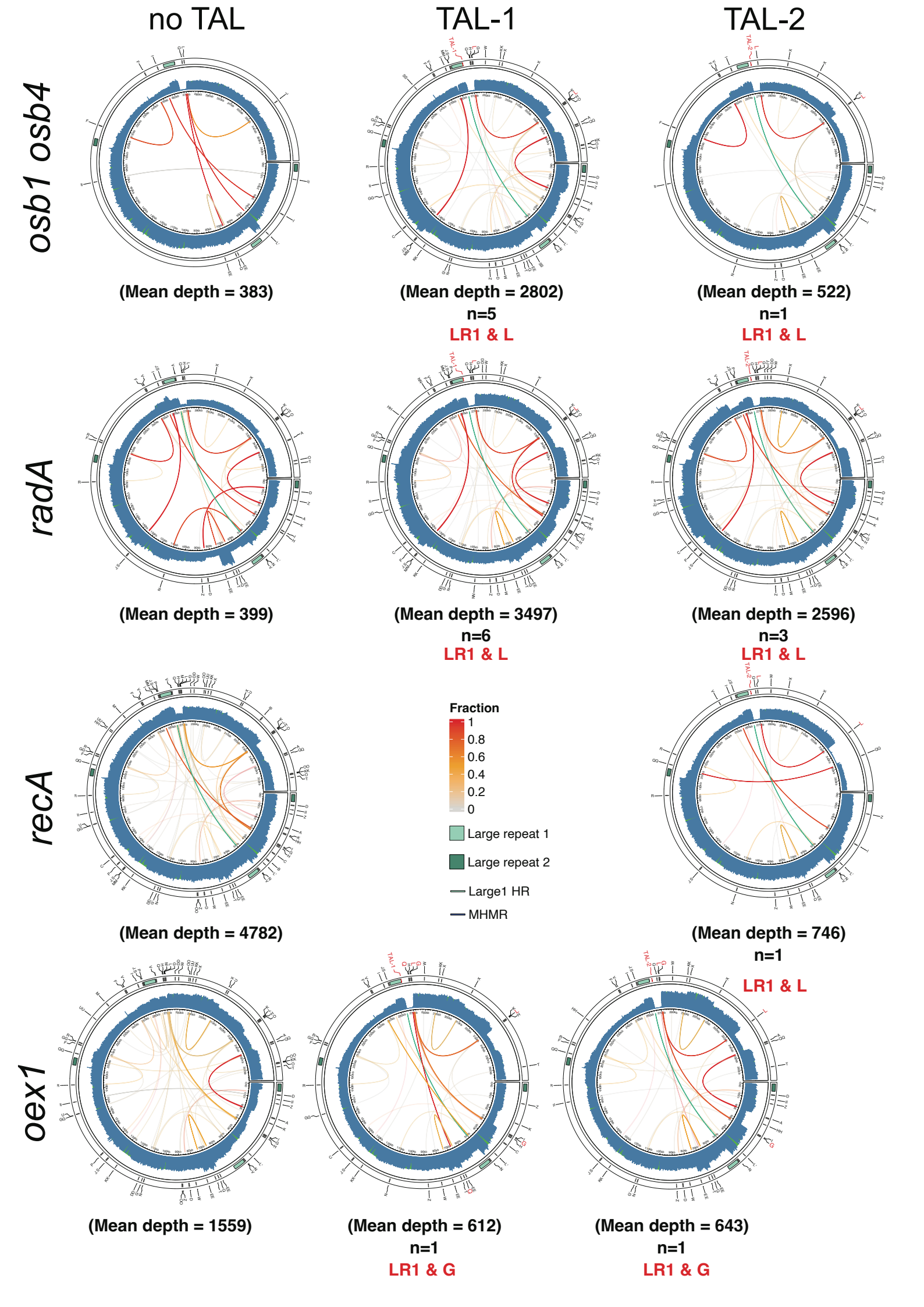
repeats L had higher frequency compared to recombination involving repeat G. Repeats L, unlike other repeats in the analysis, are a pair of direct repeats previously shown to be involved in the formation of a sub-genomic episome that can independently replicate (Wallet *et al.,* 2015). Consistently, with the previous analysis, *recG1* and *msh1* show increased recombination of all repeats considered in the analysis.

Finally, in *odb1* mutants transformed with TAL-1, the majority of new mitotypes result from DSB repaired via SSA involving the Y repeats. However, it needs to be taken into account that not all recombination events lead to a stochiometric change in mtDNA regions visible in coverage profiles. Thus, for better representation, reads mapped to the mtDNA can be shown in a circos plot with links representing the read-pairs confirming recombination events involving specific repeats, with respective recombination ratios indicated by heatmaps (Fig. III.10).

All repeats identified in the mtDNA could be grouped by size and orientation. Recombination events linked to corresponding repeats were sorted according to repeat size and compared in different mutants with and without the TALEN construct (Fig. III.11). Within a given experiment, repeats most frequently engaged in recombination were 200–300 bp in size, approximately the same size as most sequenced fragments (Fig. III.7). However, there was an increased preference for intermediate-size repeats in mutants *odb1, why2, oex1, msh1* and *recG1,* and for larger repeats in *pol1B* containing the TAL-2 construct. It could also be seen as there is a decrease in ectopic recombination involving small repeats in *pol1B* mutants (Fig. III.11).

Furthermore, the frequency of recombination events could be sorted according to the orientation of the repeat pairs (Fig. III.12). Repeats can be in direct or inverted orientation, facing the same direction or facing opposite directions, respectively. Mutants that had increased recombination involving direct repeats are the same that had increased DSB repair involving repeat L, and that underwent a significant structural reorganisation of the mtDNA: o*ex1*, *osb1osb4*, *radA* and *recG1*. However, although *msh1* also showed important mtDNA reorganisation, the ratio between repeats engaged in recombination, whether in direct or inverted orientation, is similar to that of WT. Recombination involving direct repeats is indicative of the formation of circular sub-genomes that might behave as episomes, capable of independent replication. However, the orientation of repeats can change due to genome reorganization. Therefore, conclusions from the analysis should be taken with caution.


no TAL
TAL-1
TAL-2
odb1
(Mean depth = 359)
(Mean depth = 2804)
n=4
SSA & LR1 & V & G
(Mean depth = 2155)
n=4
LR1 & G
why2
(Mean depth = 359)
(Mean depth = 4596)
n=8
LR1 & V & G & MHMR
(Mean depth = 824)
n=5
LR1 & G & MHMR
recG1
(Mean depth = 2600)
(Mean depth = 3312)
n= 3
SSA & LR1 & L
(Mean depth = 1919)
n= 5
LR1 & G & H & L & MHMR
msh1
(Mean depth = 424)
(Mean depth = 1646)
n=2
SSA
(Mean depth = 5110)
n=6
I & J.S. & L & MHMR
Fraction
1
0.8
0.6
0.4
0.2
0
Large repeat 1
Large repeat 2
Large1 HR
MHMR


no TAL
TAL-1
TAL-2
osb1 osb4
(Mean depth = 383)
(Mean depth = 2802)
n=5
LR1 & L
(Mean depth = 522)
n=1
LR1 & L
radA
(Mean depth = 399)
(Mean depth = 3497)
n=6
LR1 & L
(Mean depth = 2596)
n=3
LR1 & L
recA
(Mean depth = 4782)
Fraction
1
0.8
0.6
0.4
0.2
0
Large repeat 1
Large repeat 2
Large1 HR
MHMR
(Mean depth = 746)
n=1
LR1 & L
oex1
(Mean depth = 1559)
(Mean depth = 612)
n=1
LR1 & G
(Mean depth = 643)
n=1
LR1 & G

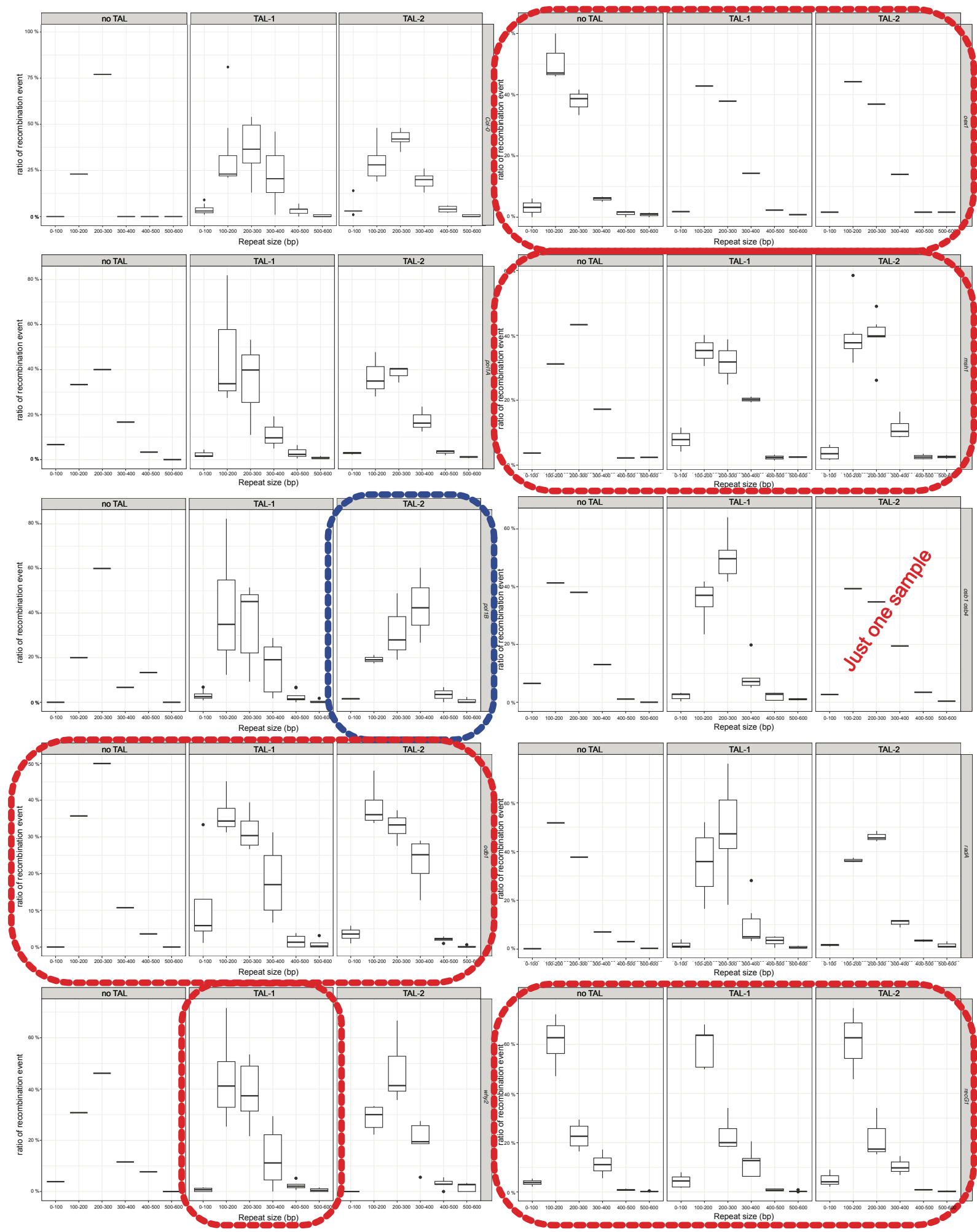


Figure III. 11 Bioinformatic analysis of read pairs confirming structural genome variants associated with repeats, grouped according to repeat size, for T-DNA mutant lines with and without mitoTALEN. Mutant lines indicated in red show increased structural recombination next to small repeats (100-300 bp); in blue preference for larger repeats (>300 bp).

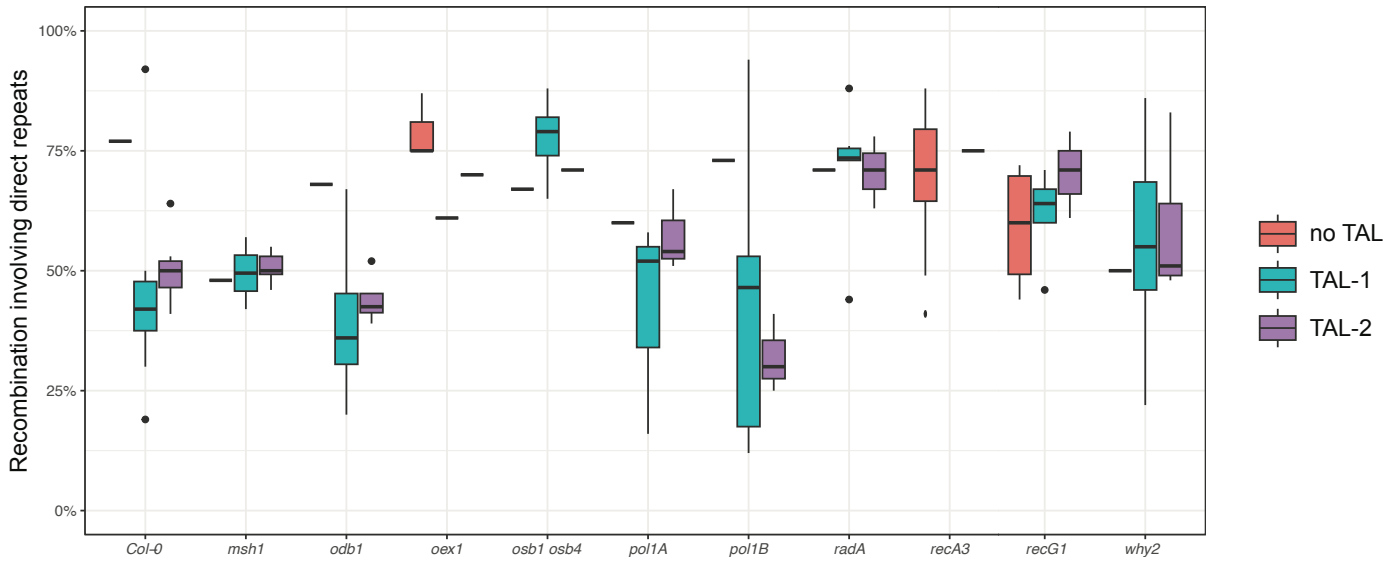


Figure III. 12 Bioinformatic analysis of read pairs confirming structural genome changes associated to recombination involving repeats, grouped according to repeat orientation in the mtDNA. For T-DNA mutant lines with and without mitoTALEN. Values above 50% indicate increased structural variation involving direct repeats, which is assumed to be linked to the formation of autonomously replicating sub-genomes.

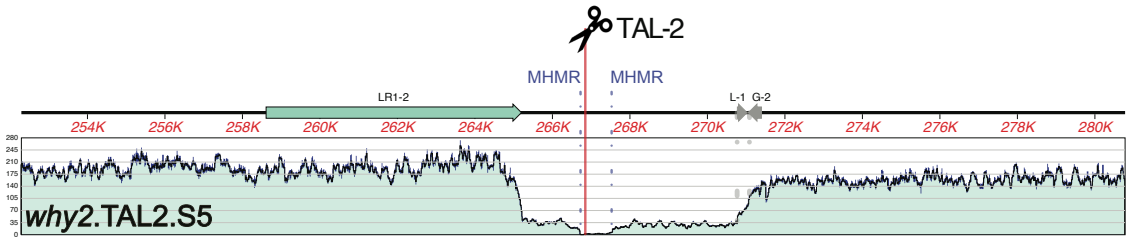


Figure III. 13 Coverage map of Illumina sequencing reads from Arabidopsis why2.TAL2.S5 mutant plant, mapped to the Col-0 mtDNA using the Bowtie aligner of the MacVector application. Blue dotted lines indicate the microhomology sites creating a minor mitotype via microhomology-mediated recombination, and small repeats involved in HR creating the major mitotype are indicated in grey.

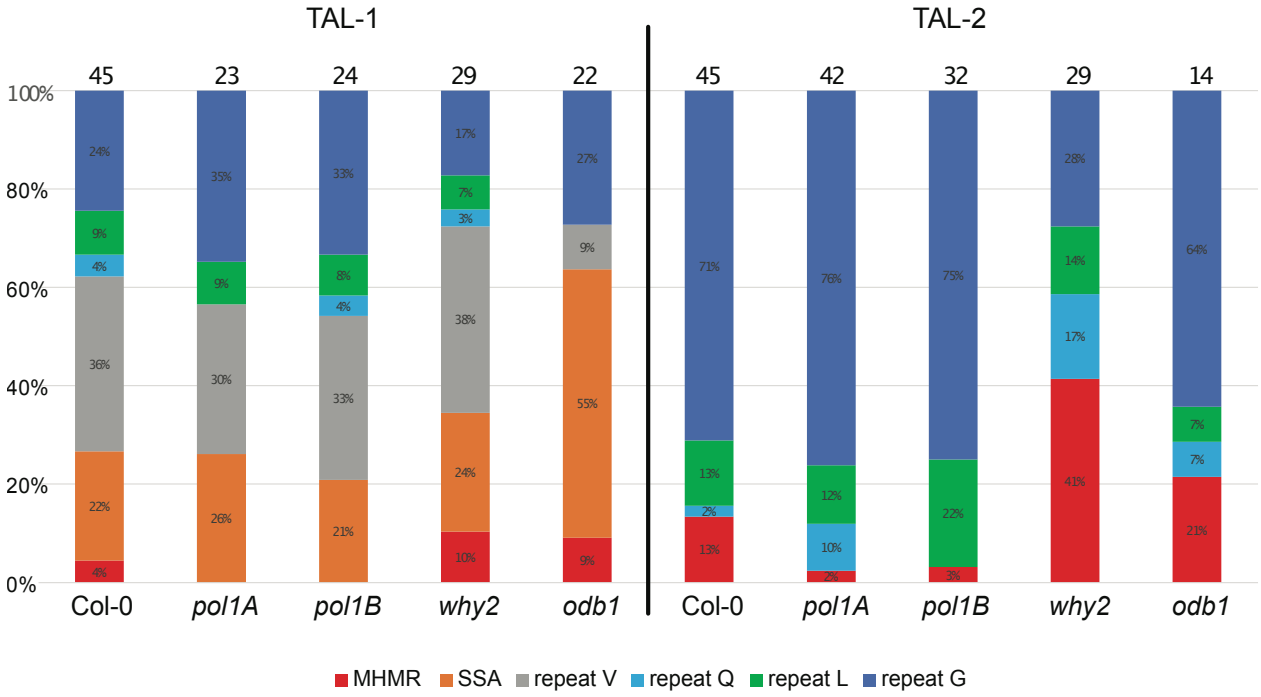


Figure III. 14 Frequencies of the different repair mechanisms involved in the repair of the DBS introduced by MitoTALENs, in the mutant lines transformed with the mitoTALEN constructs (MHMR, SSA, and HR involving repeats V, Q, L, or G). Assessed by qPCR analysis, as described in the Material and Methods and Chapter II sections. The number of independent T1 plants analyzed is indicated.

Chi-square test of independence

Genotype	N value	Total repair	SSA	MHMR
*pol1A*_TAL-1	23	0.642	0.694	0.153
*pol1B*_TAL-1	24	0.704	0.885	0.144
*why2*_TAL-1	29	0.970	0.827	0.793
*odb1*_TAL-1	22	0.011**	0.001**	0.975
*pol1A*_TAL-2	42	0.550	X	0.052*
*pol1B*_TAL-2	32	0.504	X	0.114
*why2*_TAL-2	29	0.000**	X	0.000**
*odb1*_TAL-2	14	0.947	X	0.407

Figure III. 15 Chi-square test of independence, done to distinguish significant changes of DSB repair in the mutants compared to WT, analyzed in Figure III. 14. Asterisks (*) indicate a significant difference.

As described in Chapter II, for WT plants containing the TAL-1 or TAL-2 constructs, analysis of soft-clip reads was used to assess the frequency of DSB repair involving microhomologies. The analysis revealed a significant increase in Microhomology-Mediated Repair (MHMR) in *why2* mutants (see Appendix 5). Four out of eight *why2* TAL-1 plants and all five *why2* TAL-2 plants sequenced contained mitotypes with DSBs repaired via MHMR. However, not all these mutants had MHMR as a major repair pathway, some having major mitotype created similarly as in WT, via recombination at Large repeat LR1 and intermediate size repeats V or G (Fig. III.13).

4. Identify repeats involved in the DSB repair by qPCR

The NGS analysis allowed us to identify the DNA repair fingerprints and repeats active in DSB repair, in the various mutant plants transformed with the mitoTALENs. Yet, analysing more plants for statistical significance by Illumina sequencing was impractical, due to the high cost and workload involved. Thus, we used the qPCR-based analysis approach described in Chapter II, to increase the number of individual T1 plants analysed and boost the confidence in repair fingerprints and determine the frequency of the alternative repair pathways mobilized to repair the DSBs in the different mutant plants. This analysis allowed us to compare which repair pathways are preferred in different mutant plants: *pol1A, pol1B, odb1* and *why2,* compared to WT Col-0 (Fig. III.14). Individual qPCR analyses can be found in Appendix 6. However, mutants such as *osb1osb4, radA, oex1, recG1,* and *msh1* exhibited too complex mtDNA structural changes and were not amenable to the qPCR analysis. As described in Chapter II, plants transformed with TAL-1 were optimal for analysis of repair by SSA involving repeats V, while in TAL-2 plants there are no repeats capable of undergoing SSA for the repair. Plants transformed with TAL-2 were better for analysis of MHMR, because the detection of MHMR in TAL-1 mutants is impaired by the presence of too many repeats around the target sequence.

The analysis revealed that loss of POL1A and POL1B does not lead to any significant change in how DSBs are repaired (Fig. III.14 and 15). Conversely, the loss of ODB1 resulted in a greater than 2-fold increase in the preference for SSA, in TAL-1 mutants, confirming the previous results from illumina sequencing. Also supporting the observations from illumina sequencing, the qPCR analysis revealed that loss of WHY2 leads to a greater than 3-fold increase in the preference for MHMR, in TAL-2 mutants. χ^2 test of independence supported the statistical significance of these observations.

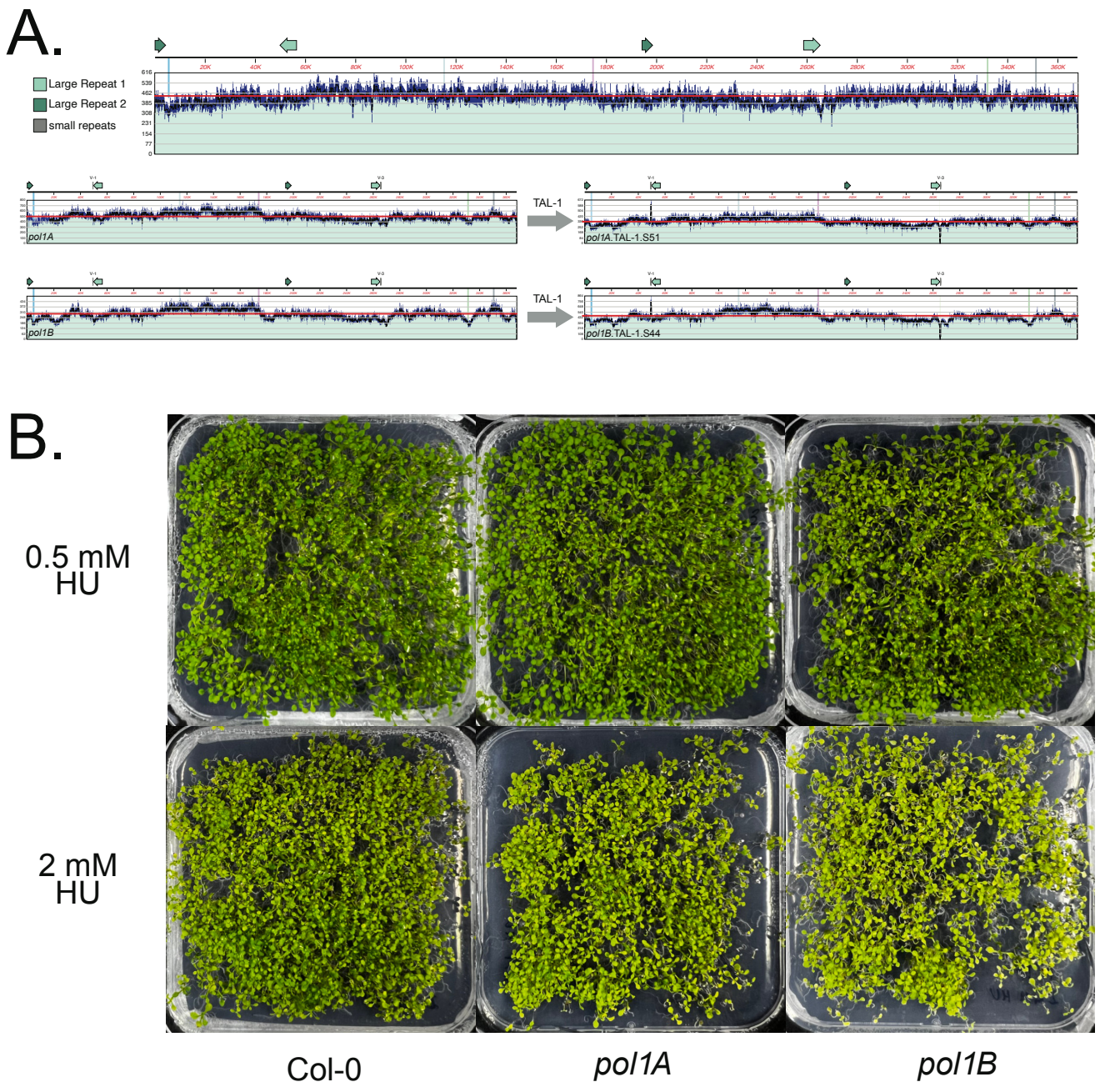


Figure III. 16 Analysis of pol1A and pol1B Arabidopsis mutants effects on mtDNA replication. (A) Coverage map of Illumina sequencing reads from Arabidopsis Col-0 and pol1A and pol1B mutants, with and without mitoTALEN constructions. Mapped to the Col-0 mtDNA using the Bowtie aligner of the MacVector application. The red line indicates the variation of mtDNA copy (read coverage) across the mitochondrial genome in the mutants. Colour-coded peaks indicate the region taken for qPCR analysis in C. (B) Images of Arabidopsis Col-0, pol1A and pol1B 10-day seedlings grown in vitro, on MS medium containing 0.5 mM or 2 mM hydroxyurea. (C) qPCR analysis of mtDNA copy number of the variable region indicated in A for seedlings in B.

5. Effect of *pol1* mutants on mtDNA replication

Although NGS sequencing results showed that *pol1A* and *pol1B* mutations do not affect how DSBs are repaired as compared to WT, we observed that the mtDNA coverage profiles for *pol1A* and *pol1B* plants were slightly distorted, with ups and downs that suggested putative sites of replication initiation (Fig. III.16A). To further explore the impact of POL1A and POL1B on replication and repair, plants were treated with hydroxyurea (HU), which inhibits ribonucleotide reductase and depletes deoxyribonucleotide triphosphates (dNTPs) required for DNA synthesis (Żabka *et al.,* 2013). This might result in replication fork collapse, requiring the BIR repair pathway to resume replication. Following HU treatment, the relative copy number of different mtDNA regions were quantified by qPCR.

Results showed that *pol1B* plants were more sensitive to HU, being more chlorotic than WT and *pol1A* mutants (Fig. III.16B). Consistently, *pol1B* had a greater reduction in mtDNA copy number. This observation aligns with previously published results suggesting that POL1B is preferentially involved in mtDNA repair. However, our analysis did not reveal significant differences in the relative copy numbers of different mtDNA regions between *pol1A* and *pol1B*, indicating that both DNA polymerases are involved in mtDNA replication.

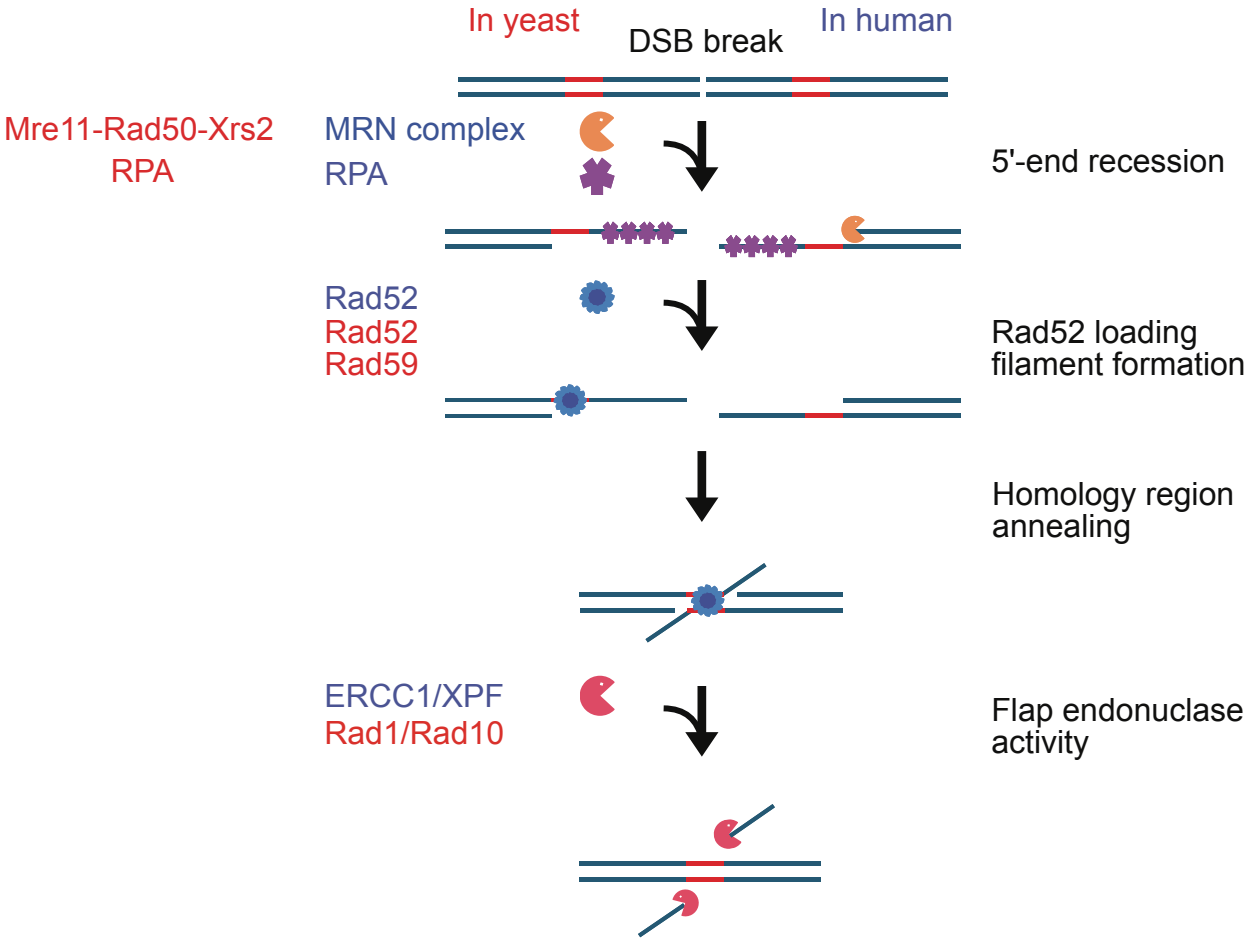


Figure III. 17 Mechanism of Single-Strand Annealing repair, in yeast and humans. Adapted from Bhargava et al., 2016; Haber, 2006; Hanamshet et al., 2016; Motycka et al., 2004.

Discussion

1. Segregating T-DNA lines of HR mitochondrial mutants

Previous to transformation with the TAL-1 or TAL-2 constructs, the different mutant lines were Illumina sequenced, to establish the status of the mtDNA before TALEN-induced DSB and repair. The *odb1, why2, pol1A, and pol1B* mutants did not show any discernible changes in the mtDNA as compared to WT. This consistent with the absence of any visible developmental phenotype. However, *oex1*, *radA*, *osb1osb4*, *msh1*, *recA* and *recG1* mutants showed major structural reorganisation of the mtDNA, already in the absence of mitoTALEN mtDNA damage. Previous studies have shown that the *oex1, radA, recA3* and *osb1* mutations have detrimental effects on mtDNA stability, with subsequent mutant generations more severely affected than the first homozygous plants (Chevigny et al., 2022; Shedge et al., 2007; Zaegel et al., 2006). For example, the *osb1* mutants were shown to form a stable new mitotype after the second homozygote generation, a mtDNA sorting that is no longer reversible by backcrossing (Zaegel et al., 2006). Thus, mutants that affect the mtDNA stability might have different mtDNA configurations, as shown by previous research. Therefore, we opted to study the T1 generation to avoid parental bias of inherited mtDNA and mitotype selection mechanisms that can happen over generations. It's important to note that previous studies done on the *msh1 (*SALK_046763) line are different from studies done in this project, as the mtDNA of our line was missing the *atp6-1* gene region (Fig. III.5), a mitotype segregation that apparently occurred over successive generations of propagation of the mutant.

2. Divergent roles of Rad52 and the plant ODB1 ortholog

In yeast, Rad52 has been shown to be a key player in homologous recombination and DSB repair. After DSB and 5′-strand recession, 3'-overhangs are stabilised and protected from degradation by replication protein A (RPA). However, the formed complex is inert, and to proceed with HR, RPA must be replaced by Rad52, which facilitates DNA invasion by Rad51 to form a D-loop (Sung, 1997; Sung and Klein, 2006). In humans, the role of Rad52 is performed by BRCA2 (Esashi *et al.,* 2007). Similarly, in yeast and humans, Rad52 plays a role in SSA (Fig. III.17). It promotes the annealing of ssDNA regions of homology, captures ERCC1-XPF (in humans) and Rad10-Rad1 (in yeast) to form a complex and promote endonuclease activity to remove flaps. The endonuclease activity in plants could be performed by OEX1 (Schatz-Daas, 2021). The remaining ends are joined by DNA ligase (Bhargava et al., 2016; Haber, 2006; Hanamshet et al., 2016; Motycka et al., 2004). In yeast, a close relationship was proposed between replication, recombination and SSA, with replication and recombination activities preventing Rad52-dependent SSA (Onaka *et al.,* 2020).

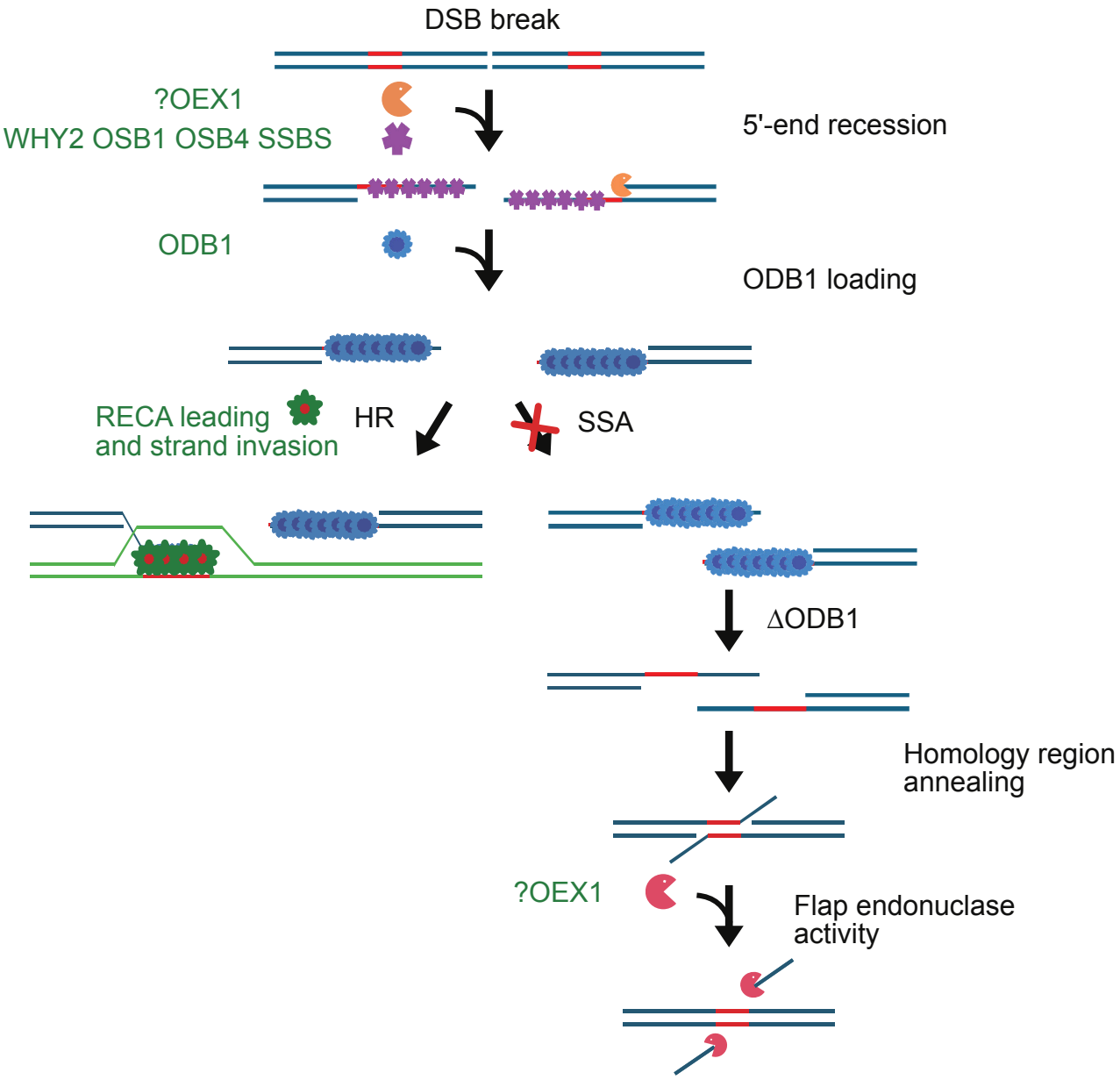


Figure III. 18 Proposed role of ODB1 in the mechanism of Single-Strand Annealing repair in Arabidopsis mitochondria.

As previously discussed, ODB1 and RECAs (RECA2 and RECA3) are considered to be plant mitochondrial functional orthologs of yeast Rad52 and Rad51, respectively. The exonuclease activity needed to proceed with SSA could be performed by OEX1, as shown by *in vitro* studies (Schatz-Daas, 2021). However, our study indicates that loss of ODB1 leads to increased SSA upon DSB, contrary to what was found in yeast and animals for RAD52/ BRCA2, where *rad52* mutations lead to defective SSA (Stark *et al.,* 2004; Symington, 2002). Similar results have been seen by overexpression of ODB1 targeted to the nucleus (Samach *et al.,* 2018): mutant plants had increased intra-chromosomal recombination and SDSA, while the SSA pathway was repressed. The repression of SSA could be the result of increased HR, as SSA and HR competes for Rad52. In the nucleus, it has been shown that an increase in HR employs most Rad52, leaving the SSA pathway deficient. On the contrary, no decrease in HR was noticed in *odb1* TAL-2 mutants, suggesting that the increase in SSA in *odb1* TAL-1 was not caused by changes in HR efficiency. This suggests that ODB1 is not a limiting factor between SSA and HR in mitochondria, as when targeted to the nucleus (Samach *et al.,* 2018). Furthermore, this suggests that there are alternative pathways independent of ODB1, as *odb1* TAL-2 mutants do not affect HR or plant ODB1 functions differently than its orthologs in eukaryotes.

We propose a model where, in the presence of ODB1, homologous ssDNA regions are hidden by the binding of high-affinity and abundant single-strand DNA binding proteins, thereby suppressing SSA. However, in the absence of ODB1, homology base pairing becomes accessible (Fig. III.18). Furthermore, previous studies showed that *odb1* mutants have increased DSB repair by MHMR (Janicka *et al.,* 2012). Our present study also indicates a 1.6-time increase in MHMR for *odb1* TAL-2 mutants, although the sample size was insufficient to achieve statistical significance.

3. Single-strand DNA-binding protein WHY2

Our studies, consistent with published results, show an increased preference for MHMR in *why2* plants (Cappadocia *et al.,* 2010). Recent studies on *why2* showed that the protein is closely involved in the structural organization of the mtDNA, particularly in the packing into nucleoids (Golin *et al.,* 2020). We proposed a repair mechanism akin to ODB1, whereby homology sites are hidden by the WHY2 protein (Fig. III.19). However, the functional differences between these two abundant, ssDNA-binding proteins from plant mitochondria remain unknown.

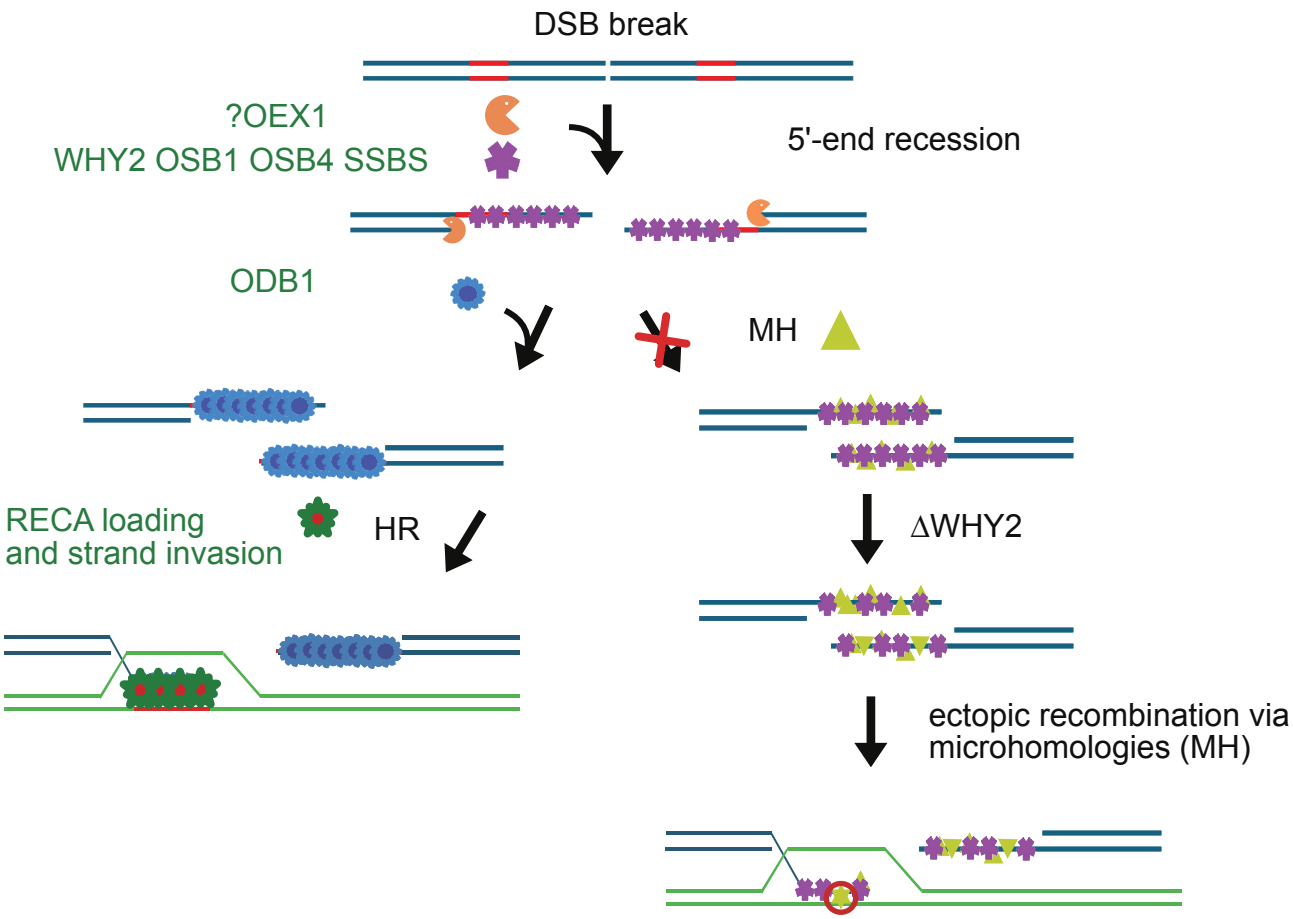


Figure III. 19 Proposed role of WHY2 in the mechanism of Microhomology Mediated Repair in Arabidopsis mitochondria.

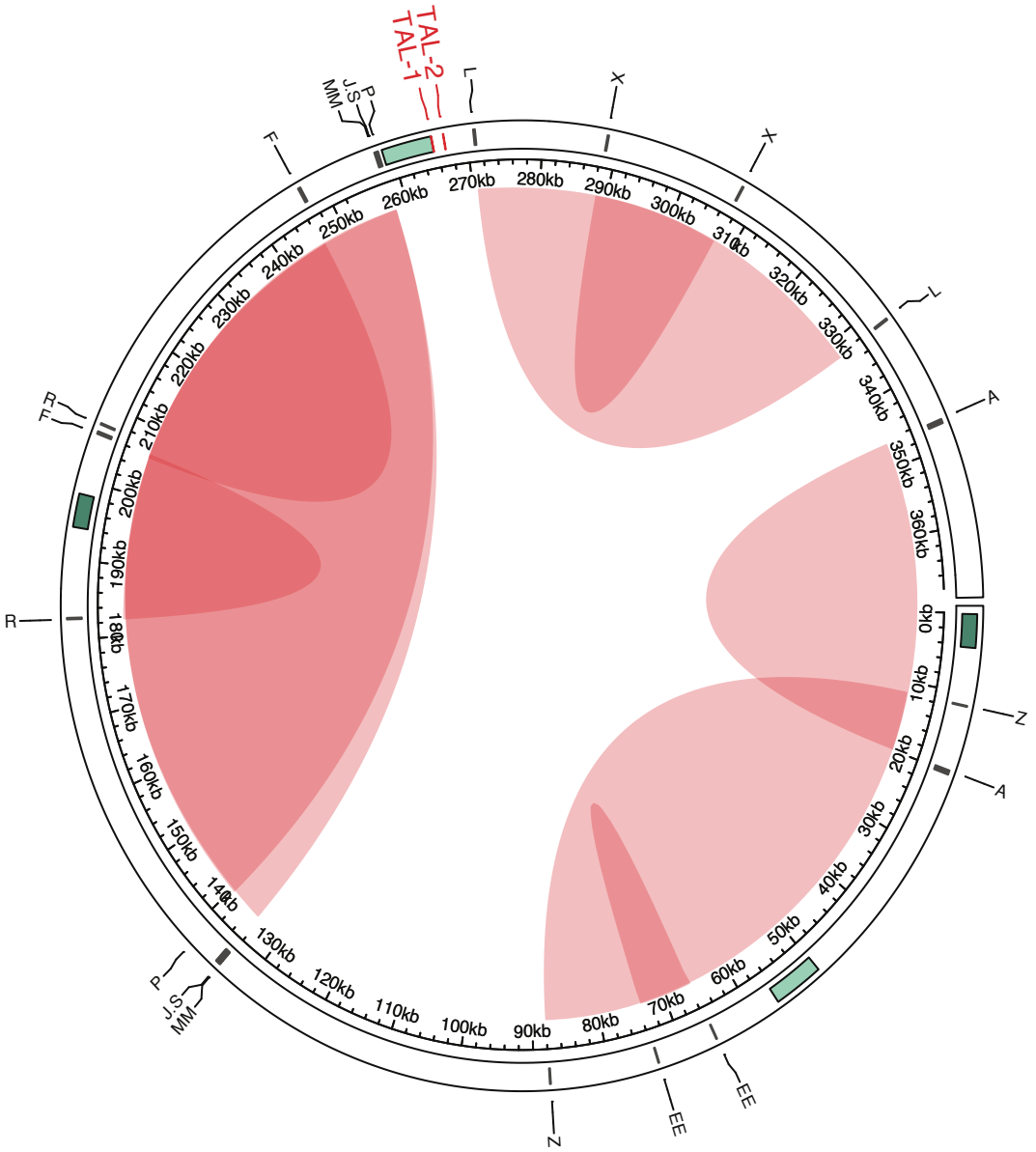


Figure III. 20 Circular representation of the Arabidopsis Col-0 mitochondrial genome, indicating in red sub-genomes that might be formed via Homologous Recombination involving direct repeats.

4. Loss of HR-dependent repair

Analysis of Illumina sequences of mitoTALEN mutants of *radA, osb1osb4,* and *recA3* revealed no fingerprints of the HR-dependent repair that is predominately engaged in WT plants, as seen in Chapter II. This suggests that these mutant lines might have lost the ability to repair DSB by homology-based recombination pathways. However, these mutants are viable, although they do show pronounced developmental phenotypes and global mtDNA reorganisation.

We proposed that in these plants an alternative process allows survival when the mtDNA integrity is compromised. This process would involve the segregation of a viable mitochondrial genome, reconstructed from minor pre-existing sub-genomes created before the introduction of the mitoTALEN constructs (Fig. III.20). These sub-genomes seem to be created via recombination involving direct repeats, leading to the formation of episomes that can segregate independently. The formation of such episomes was demonstrated in the study of *recG1* mutants, in which small episomes are formed from recombination between direct repeats EE (Wallet *et al.,* 2015). These sub-stoichiometric alternative mtDNA configurations (mitotypes) constitute a sink of genomic variability, contributing to the fast evolution of the plant mtDNA, by a clonal amplification process that can occur as rapidly as over a single generation (Janska *et al.,* 1998; Small *et al.,* 1989; Woloszynska and Trojanowski, 2009; Zaegel *et al.,* 2006). Mitochondrial instability was also shown to lead to cell cycle arrest, resulting in an even greater mitochondrial sorting mechanism (Chevigny *et al.,* 2022). Genomes created from pre-existing mitotypes in *radA, osb1osb4,* and *recA* are similar, yet not the same. Whether heteroplasmic sorting differences are stochastic natural variations or linked to protein functions, remains unknown. Previous studies have shown that sorting of the heteroplasmic genome favours GC base pairs. Thus, there is some bias in the sorting process which might accelerate the removal of minor mitotypes (Broz *et al.,* 2022).

5. Regulation of homologous recombination

In some of the mutants studied, the repair or the mitoTALEN-induced DSB remains puzzling. Arabidopsis *recG1, oex1* and *msh1* mutants showed similar repair fingerprints for sub-genomes segregation, as seen in Chapter III.3.4. However, DSB repair by HR, as seen for WT plants in Chapter II, can also be detected in these mutants. The number of read-pairs corresponding to structural variants is 2-3 times higher in *recG1* and *msh1* mutants compared to WT, while in other mutants, such an increase is relatively small. This suggests that RECG1, OEX1 and MSH1 have important roles in the surveillance of HR. Furthermore, more HR sites were detected in *recG1* and *msh1*, as compared to other analysed mitoTALEN mutants.

However, it remains challenging to distinguish between an increase in HR and the amplification of sub-genomes created via HR.

In the *msh1* line we used, the cut site region for TAL-1 was moved to the atp6-1 region before the introduction of the mitoTALEN constructs via recombination involving Large repeat LR1 and intermediate size repeat V. Thus, the study of repair in *msh1* was highly biased towards segregation that occurred before construct introduction. Finally, it's important to note that in this *msh1* line, recombination involving repeats LR1 and repeats V differs from that described in Chapter II. In this *msh1* line, the *atp6-1* copy was replaced by the *atp6-2* copy, while in WT MitoTALENs, it is the *atp6-1* gene copy that was duplicated and replaced the *atp6-2* copy. However, it remains unknown if the orientation of recombination is controlled and, if so, by what mechanism. In *msh1, the atp6-1* copy might have been lost during heteroplasmic sorting. Furthermore, In WT mitoTALEN, HR is directed by the gene loss; thus only one recombination outcome is possible. Yet in *msh1*, both respective recombination outcomes contain TAL-1 cut site and are compromised by mitoTALEN.

CHAPTER IV: Importance of ATP9 subunit for proper function of ATP synthase

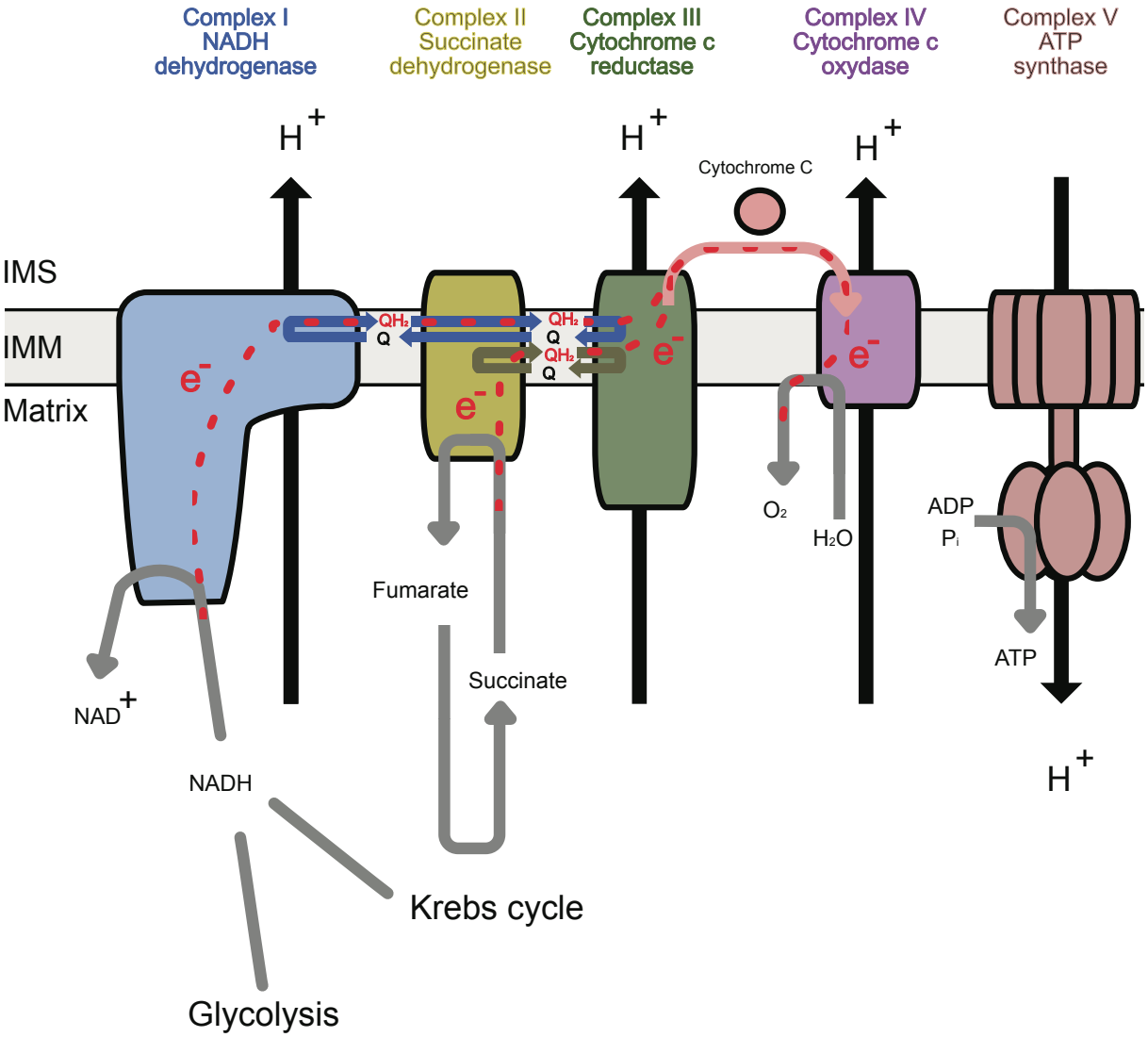


Figure IV. 1. Schematic representation of electron transport chain (ETC) in mitochondria. NADH from glycolysis and the Krebs cycle feeds ETC with electrons via Complex I. Complex II acquires electrons from succinate from the Krebs cycle. Electron transfer via Complex I, III and IV is coupled with a proton pump that creates an electrochemical gradient across the inner membrane (IM). This gradient drives ATP synthase activity to produce ATP from ADP and inorganic phosphate. IMM: inner mitochondrial membrane; Q: ubiquinone; QH_2: ubiquinol. Figure prepared according to Sousa et al., 2018.

Introduction

As previously described, mitochondria is a powerhouse in nearly all eukaryotes. Despite differences between various lineages, energy is primarily produced via aerobic respiration. This catabolic pathway involves three main steps: glycolysis of nutrients, tricarboxylic acid (TCA) cycle and the oxidative phosphorylation (OXPHOS) system (Roger *et al.,* 2017). In this Chapter, I will briefly review the OXPHOS system in plants and focus on the ATP synthase (Complex V).

1. Oxidative phosphorylation system in plant mitochondria

In plants, the OXPHOS system is composed of more than 100 proteins assembled into five functional complexes, with only about 18 of these proteins encoded by the mitochondrial genome. OXPHOS can be divided into two parts: the electron transfer chain (ETC) and ATP production (Meyer *et al.,* 2019). The ETC reduces organic compounds from TCA onto molecular oxygen through several complexes (Fig. IV.1): NADH dehydrogenase (complex I), succinate dehydrogenase (complex II), cytochrome c reductase (complex III) and cytochrome c oxidase (complex IV). The backbone process of ETC is highly conserved in most eukaryotic cells. First, complex I and complex II transfer electrons from NADH and succinate, respectively, to ubiquinone, a mobile electron transporter within the inner mitochondrial membrane. Next, ubiquinol, the reduced form of ubiquinone, transfers electrons to complex III. Then, cytochrome c, a small hemoprotein located in the intermembrane space, mediates electron transfer from complex III to complex IV, which ultimately reduces oxygen to water. ETC is coupled with the formation of an electrochemical proton gradient across the inner mitochondrial membrane, with protons pumped into the intermembrane space by complexes I, III and IV. Finally, the proton gradient drives ATP synthase, the final complex in the OXPHOS system, to produce ATP.

Compared to fungal and animal systems, in plants OXPHOS is tightly linked to photosynthesis and is sensitive to day/night periods (van Dongen *et al.,* 2011). Furthermore, Complexes I, II and IV have additional subunits (Eubel et al., 2003; Millar et al., 2004; Sunderhaus et al., 2006). During the day, large amounts of reducing compounds are formed by photosynthesis that can indirectly affect mitochondria, putting them at risk of becoming 'over-reduced'. Chloroplasts perform photorespiration, driven by ribulose bisphosphate carboxylase/oxygenase (RubisCO), forming vast amounts of NADH, becoming the primary substrate of the ETC (Bauwe *et al.,* 2010). However, this would require substantial NADH re-oxidation capacities in mitochondria. To maintain the redox state, allosteric inhibition of

anaerobic respiration in the mitochondria was shown in potato illuminated leaves (Bykova *et al.,* 2005). Furthermore, the OXPHOS system in plants has developed side pathways to mediate the redox state without affecting the electrochemical proton gradient. This function is performed by alternative NAD(P)H dehydrogenases and alternative oxidase (AOX), by removing electrons from ETC via direct transfer of electrons from ubiquinol to oxygen, also playing an important role in abiotic stress responses (Selinski et al., 2018; Vanlerberghe et al., 2020, 2016).

2. ATP synthase structure and function

The primary focus of this Chapter is ATP synthase (complex V). It is composed of 15 subunits organised into two domains, F_0 and F_1. The membrane-bound peripheral stalk domain F_0 and the soluble central stalk domain F_1 are both assembled separately and joined on the mitochondrial membrane (Boyer, 1997; He et al., 2018). ATP synthase has been studied extensively in yeast and humans. However, the exact composition of plant ATP synthase remains elusive (Zancani *et al.,* 2020). The F_1 domain is a stator composed of 3 copies of α (ATP1) and β (ATP2) subunits, containing a catalytic binding site, and a central stalk formed by γ (ATP3), δ (ATP16) and ε (ATP15) subunits. The F_0 domain is composed of a motor formed by a (ATP6), e (ATP21), g (ATP20), f (ATP17), 8 (ATP8), and a ring of c (ATP9) subunits (number of ATP9 subunits in Arabidopsis unknown), and a peripheral stalk composed of b (ATP4), d (ATP7) and OSCP. Subunits h and e found in yeast and humans have not been identified in plants (Zancani *et al.,* 2020).

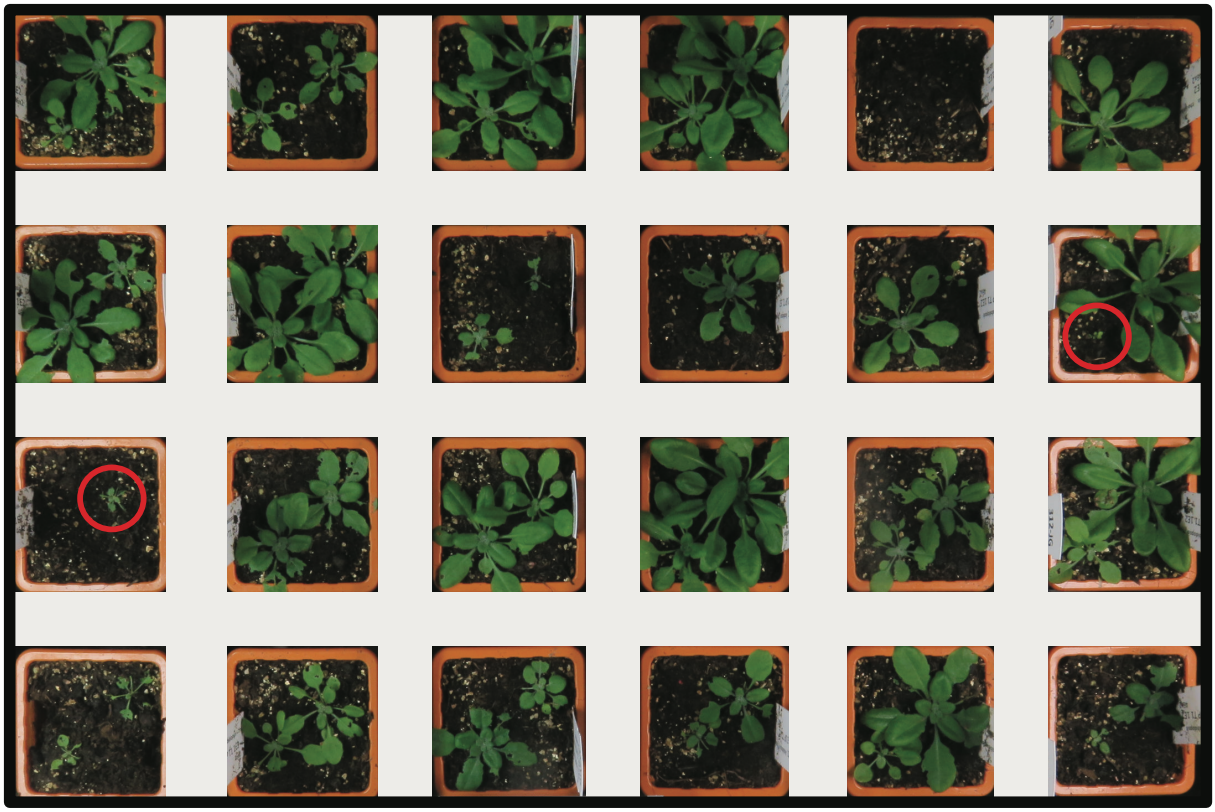
The mechanism of ATP production is highly conserved in eukaryotic cells (Zancani *et al.,* 2020). The ring of c subunits together with γε creates a rotary shaft linked via γ against the $\alpha_3\beta_3$ stator (Stock *et al.,* 1999). The F_0 domain forms a proton translocation channel, and proton movement driven by an electrochemical gradient powers the rotation of c-ring and γ within the $\alpha_3\beta_3$ catalytic sub-complex (Guo et al., 2017; Kubo and Takada, 2022). The full 360° rotation of the catalytic sub-unit produces 3 ATP molecules and comprises three steps followed by major conformation changes: catalytic site empty, ADP bound and ATP bound (Noji *et al.,* 2017). Reverse catalytic activity of $\alpha_3\beta_3$ complex to perform ATP hydrolysis is also possible (Nakano *et al.,* 2023). Furthermore, Complex V was shown to form dimers that impose curvatures on the inner mitochondrial membrane. Arrays of ATP synthase dimers are formed at the tip of the cristae (Blum *et al.,* 2019).

3. ATP synthase linked disorders in humans and plants

Mutations of the *atp6* gene (subunit a of the F_0 domain) in humans have been shown to reduce the mitochondria's ability to produce ATP and lead to neuropathy, ataxia, and retinitis pigmentosa (NARP). The symptoms of the disorder tend to manifest in childhood or early adulthood and are characterised by sensory neuropathy, ataxia, vision loss due to retinal deterioration, cognitive decline, seizures, hearing loss, and cardiac conduction defects (Blanco-Grau *et al.,* 2013; Thorburn *et al.,* 1993). Mutation in the *atp6* gene can also cause maternally inherited Leigh's syndrome (MILS), which is characterised by high mortality from respiratory or cardiac failure by the age of 3, and symptoms typically include psychomotor retardation, cerebellar ataxia, peripheral neuropathy, and hypertrophic cardiomyopathy (Carrozzo et al., 2004; Thorburn et al., 1993).

In plants, increased ATP synthase expression is tightly linked with tolerance to abiotic stress (temperature, salt), while a decrease in expression occurs during oxidative stress (Zancani *et al.,* 2020). Although total loss of complex V was shown to be lethal (Robison *et al.,* 2009; Moore *et al.,* 2003), a dysfunctional complex can be tolerated in many vegetative organs. However, proper ATP synthase activity is crucial for pollen production, due to its high energy requirements. Thus, dysfunction in ATP synthase leads to CMS (Cytoplasmic Male Sterility) (Chase, 2007; Carlsson *et al.,* 2008). Multiple studies highlighted the importance of RNA editing for the correct expression of ATP synthase subunits ATP1 (He *et al.,* 2018; Hammani *et al.,* 2011), ATP6 (Hu *et al.,* 2013; Wang *et al.,* 2009), ATP8 (Das *et al.,* 2010) and ATP9 (Wei *et al.,* 2011; Yang *et al.,* 2007). Mutations and premature stop codons in the transcripts of these genes lead to a drastic reduction in ATP synthase assembly or to a change in the stability and efficient functioning of the complex. Ultimately, all this leads to CMS in Arabidopsis (Li *et al.,* 2010) and multiple agricultural species such as maize (Li *et al.,* 2019), rice (Chakraborty *et al.,* 2015), wheat (Wang *et al.,* 2015) and soybean (Wei *et al.,* 2011). As previously shown in Chapters II and III, the plant mtDNA undergoes multiple genome rearrangements via HR, some of which can result in the production of new chimeric genes. The creation of pseudogenes and open reading frames (ORFs) involving genes coding for ATP synthase subunits ATP1 (Yang *et al.,* 2009; Yu *et al.,* 2010), ATP6(Satoh *et al.,* 2004), and ATP9 (Reddemann and Horn, 2018; Szklarczyk *et al.,* 2014) were also shown to cause CMS.

mitoTALEN-1



mitoTALEN-2

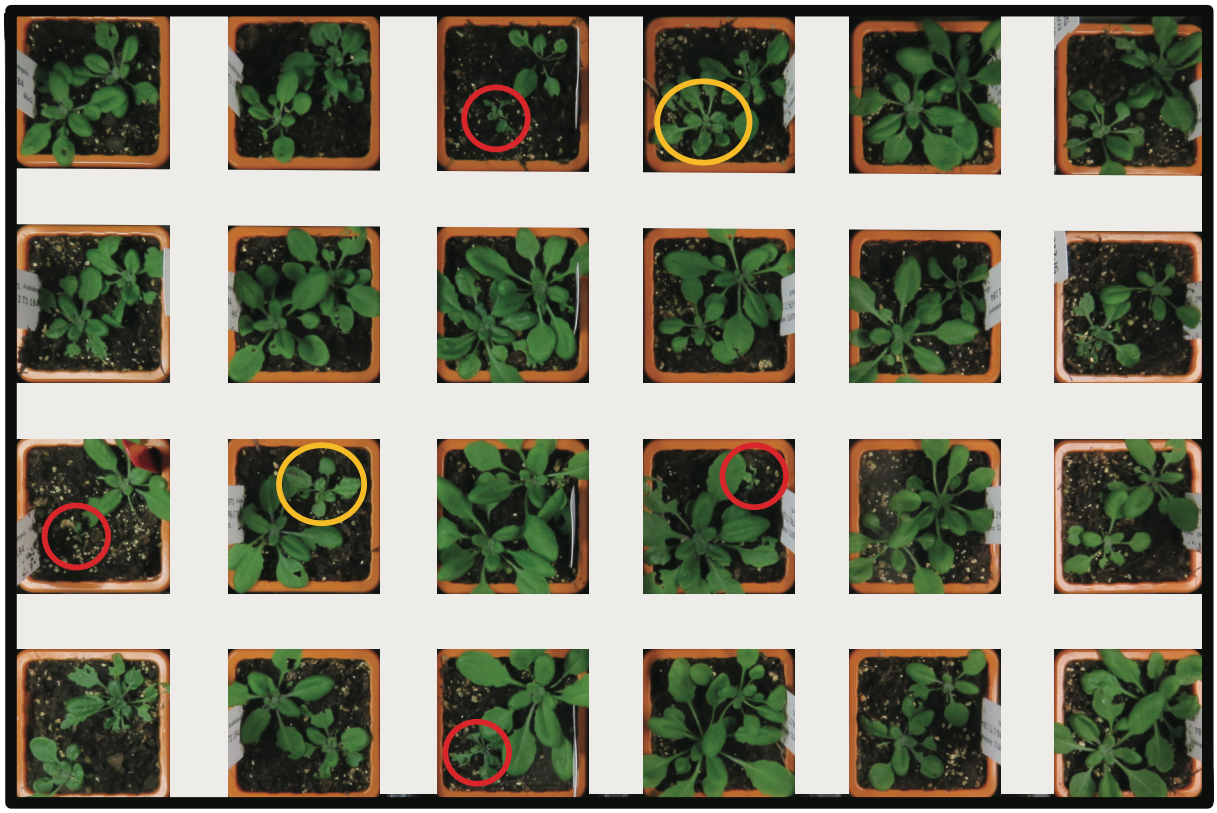


Figure IV. 2. Pictures of three-week-old T1 transformants of Col-0 Arabidopsis containing the TAL-1 (top) and TAL-2 (bottom) constructs. Red circles identify plants with strong developmental defects. Yellow circles identify plants with a chimeric phenotype: only some parts of the plant show defects in leaf morphology.

Results

1. Partial loss of ATP9 in mitoTALEN lines leads to developmental defects

After the introduction of mitoTALEN in WT Col-0 plants, we noticed that the development of some plants was abnormal: leaves were curly and variegated, sometimes reduced in size, and, depending on phenotype severity, affected plants failed to develop flowering stems and flower organs. Furthermore, the developmental phenotype was pleiotropic - not all T1 transformants exhibited the phenotype, and among those that did it was with variable severity (Fig. IV.2). When grown *in vitro* on MS medium, about a third of T1 transformants showed the phenotype (28% and 35% of TAL-1 and TAL-2 transformants, respectively) (Fig. IV.3 & 4). Of those phenotypical T1 plants, most were chimeric, with parts of the plant affected, while other parts of the plant developed normally. Furthermore, the T2 progeny of phenotypic TAL-2 plants showed strange mendelian segregation of the phenotype, as if it was for a recessive mutation, with 11 out of 48 plants exhibiting the phenotype (Fig. IV.5). However, mitoTALENs are under constitutive promoter and are expressed in the heterozygous state. Thus, such segregation indicates that the phenotype most likely originates from mitochondria.

Plants grown on soil or *in vitro* showed that TAL-2 transformants developed phenotypes more frequently than TAL-1 transformants (Fig. IV.2 & 4). In Chapter II we investigated the repair mechanisms of DSBs induced by mitoTALEN. Our findings showed that in TAL-2 transformants, the DSB is most frequently repaired via recombination involving repeats LR1 and G. However, recombination at repeat G leads to the loss of the C-terminus of the *atp9* gene and potentially creates a fusion between the 5' of *atp9* and *orf315* (Fig. IV.6). To test whether the observed deleterious phenotypes in TAL-2 plants are linked to a loss of *atp9* and the creation of a chimeric gene, we compared the copy number of the 3' terminus of the *atp9* gene, between phenotypic and WT-looking TAL-2 transformants. The qPCR results confirmed a more severe decrease of the 3' terminus of *atp9* in phenotypic TAL-2 plants (Fig. IV.7A). Similarly, in chimeric plants, variegated leaves exhibited a greater decrease in the copy number of the 3' of *atp9,* compared to leaves of the same plant that looked WT (Fig. IV.7B). Interestingly, variegated leaves were found to have increased mtDNA copy number, as evaluated by relative qPCR quantification of mtDNA regions that are far apart and not linked to the TAL-2 target region (Fig. IV.8). To further analyse whether a decrease in *atp9* gene expression correlates with the defective development phenotypes, we performed RT-qPCR and Northen blot analysis. Total RNA was extracted from TAL-1 and TAL-2 transformants, with and without phenotypic defects. Both RT-qPCR (Fig. IV.9) and Northern

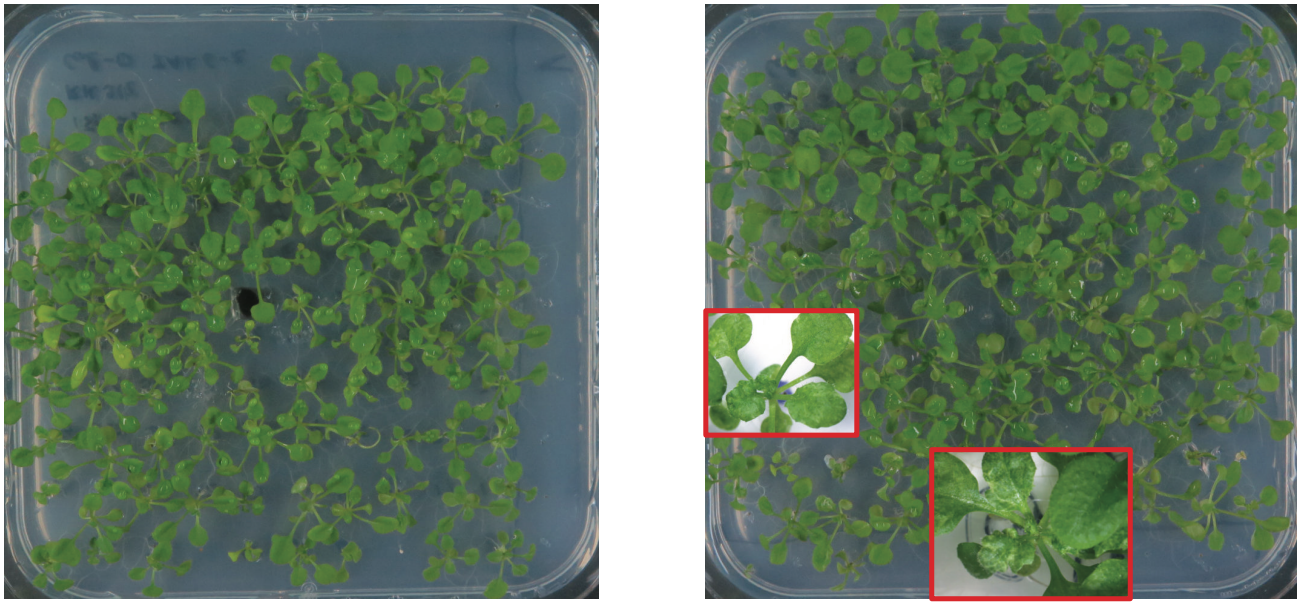


Figure IV. 3. Pictures of two-week-old T1 transformants of Col-0 Arabidopsis containing the TAL-1 (left) and TAL-2 (right) constructs, grown in vitro. Seedlings displaying the typical phenotype are enlarged within the red boxes.

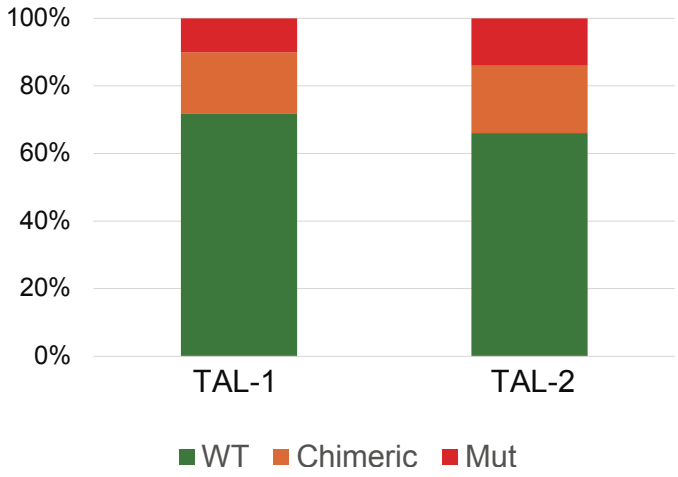


Figure IV. 4. The frequency of plants displaying defective development among T1 mitoTALEN transformants grown in vitro (Fig. IV. 3).



Figure IV. 5. Pictures of four-week-old T2 TAL-2 transformants segregating from a phenotypic chimeric T1 transformant. Red circles identify plants with strong developmental defects (11 out of 24).

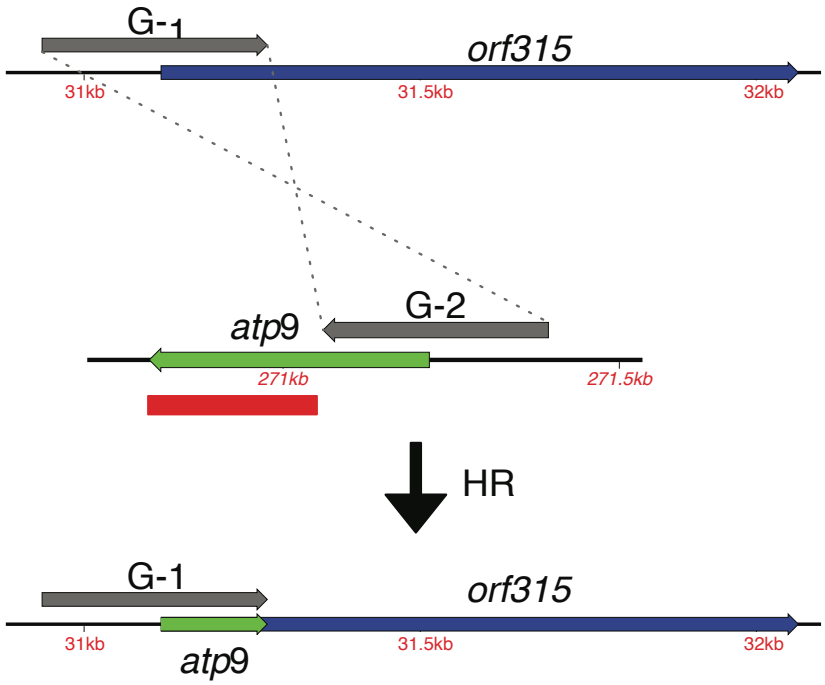


Figure IV. 6. Schematic representation of recombination involving repeat G. The 3' end of atp9 is lost due to recombination; however, as it is part of repeat G, a truncated version of atp9 is fused in frame to part of orf315, resulting in a chimeric gene that can be transcribed by the promoter located upstream atp9. The region in red region of atp9 is lost due to the repair of DSB.

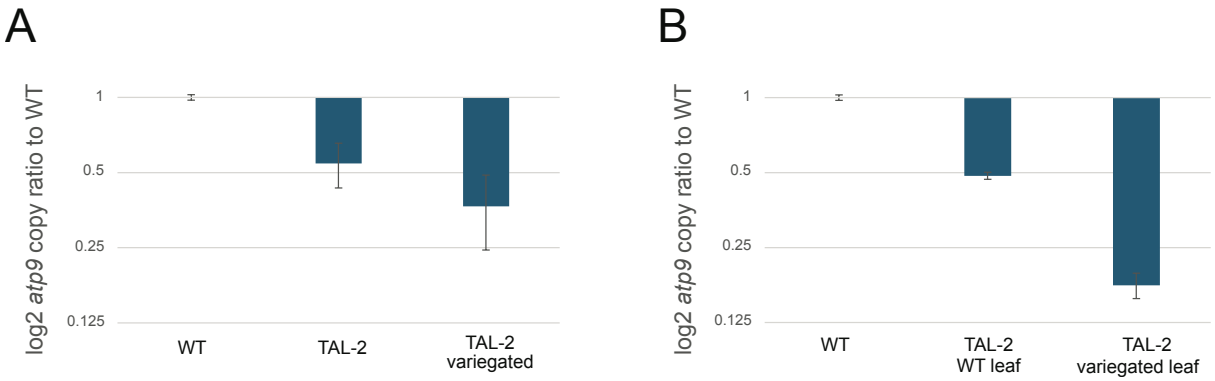


Figure IV. 7. qPCR quantification of the 3' end of atp9. (A) Comparison in T1 TAL-2 transformants with (variegated) and without defective development. (B) Comparison between variegated and wild-type-looking leaves from the same chimeric TAL-2 plant.

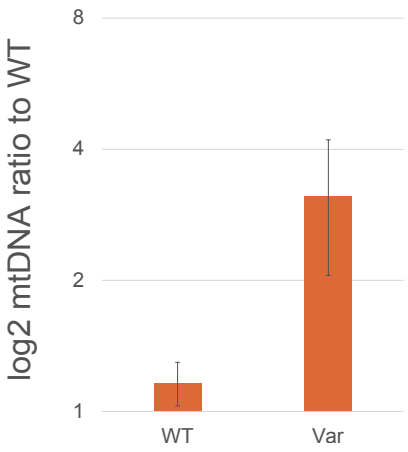


Figure IV. 8. qPCR analysis of mtDNA copy number in leaves with and without defective developmental.

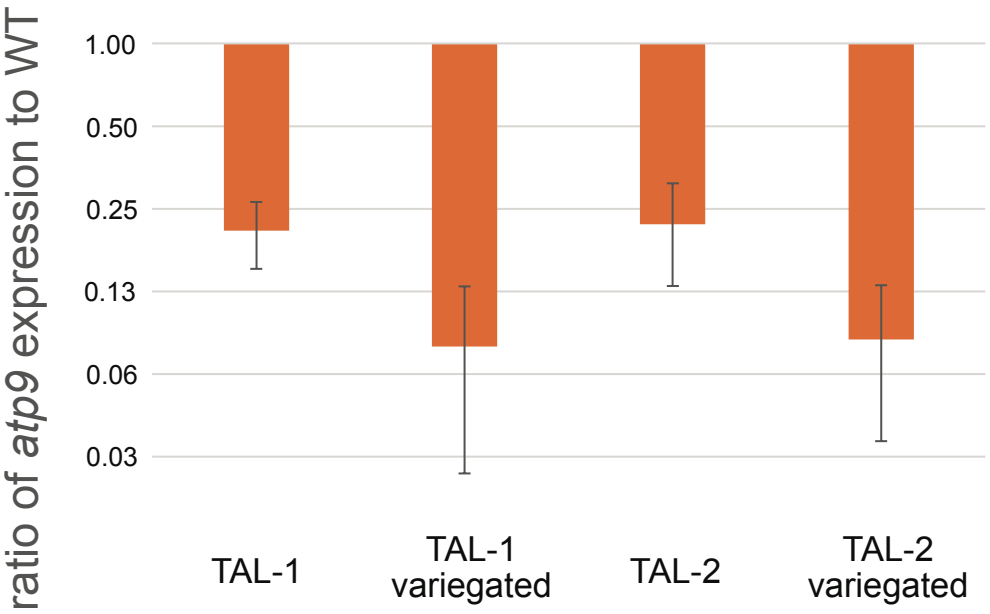


Figure IV. 9. RT-qPCR analysis of atp9 expression in TAL-1 and Tal-2 transformants with and without defective development.

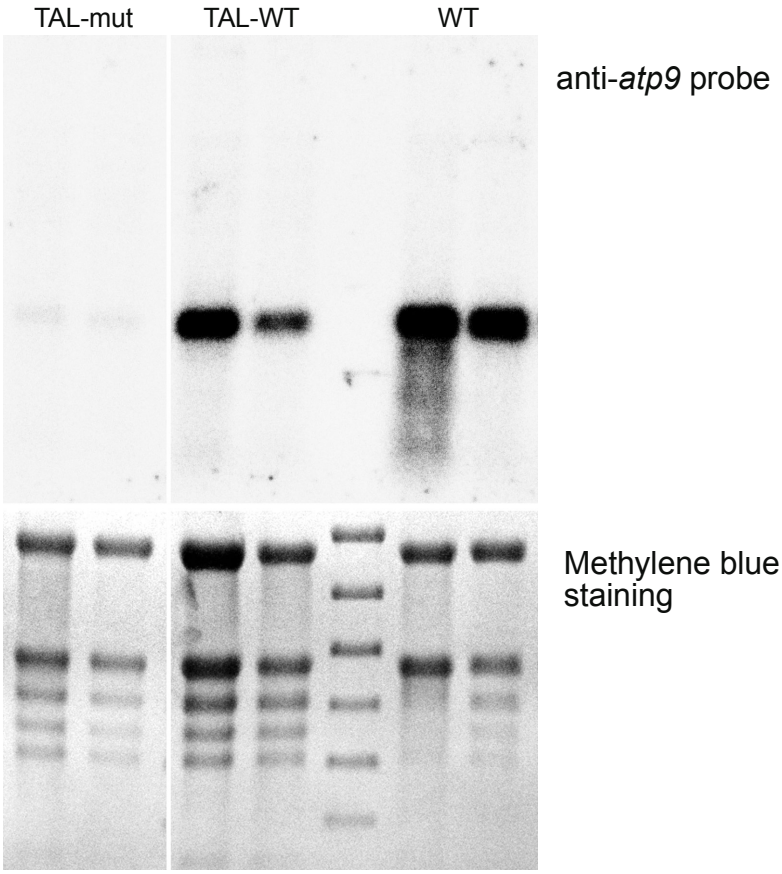


Figure IV. 10. Northern blot analysis of RNA extracts from TAL-2 transformants with and without development defect compared to WT, using a radio-labelled probe against the *atp9* gene transcript (Appendix 1). The equivalent loading of total plant RNA was assessed by methylene blue staining of the membrane after transfer.

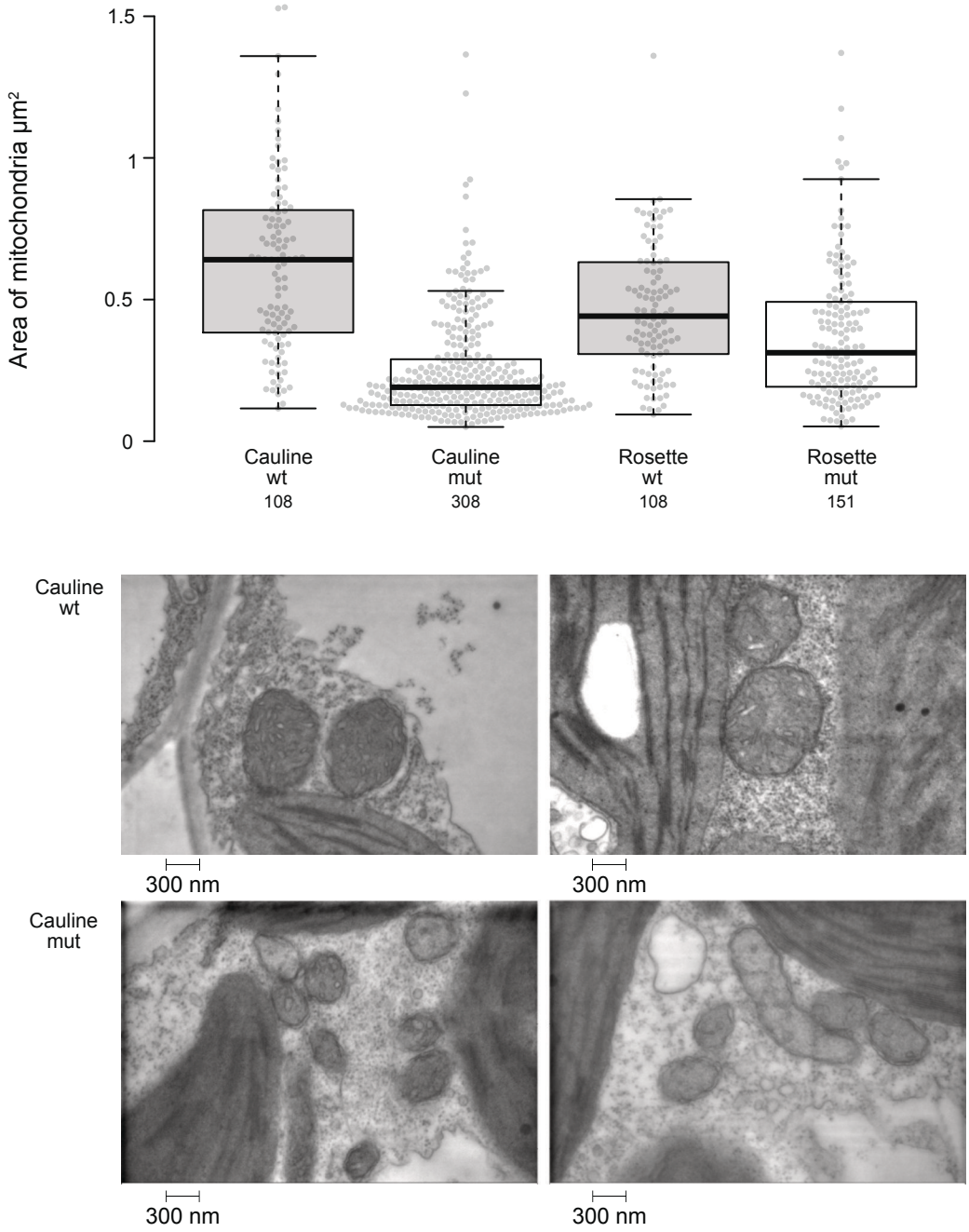


Figure IV. 11. Analysis of mitochondrial TEM images from variegated and WT-looking rosette and cauline leaves of TAL-2 transformants. Boxplots represent the area of mitochondria in TEM images, serving as a proxy for mitochondrial size. Representative images are shown below.

blot analysis (Fig. IV.10), with a probe specific to the *atp9* gene, revealed a decrease in the *atp9* transcript levels in TAL-1 and TAL-2 plants with defective development, compared to transformants with no phenotype and WT Col-0 plants. However, no fusion transcript of *atp9* and *orf315* was detected in the Northern blots (Fig. IV.6 & 10). These findings collectively confirm that the developmental defects in mitoTALEN plants arise from a loss of the *atp9* gene and a decrease in its expression, rather than the expression of a chimeric *atp9-org315* gene

2. Role of ATP9 in Complex V

To further study the outcomes of the decrease in *atp9* transcripts in mitoTALEN transformants, we intend to check how the partial loss of ATP9 protein affects the assembly of the ATP synthase complex. For this, we plan to perform blue-native gel electrophoresis analysis of mitochondrial protein extracts, from TAL-2 transformants with and without the defected phenotypes. The analysis will involve Coomassie staining to visualize the respiratory complexes and western blot with antibodies specific to complex V. We anticipate observing a decrease in the amount and/or size of Complex V in protein extracts from phenotypic TAL-2 transformants. Because of time constraints, this could not be done before the defence of this thesis.

Previous studies have shown that the clustering of ATP synthase complexes is responsible for the formation of cristae in mitochondria. With the help of Mathieu Erhardt from the Bioimaging platform of IBMP, we obtained transmission electron microscope (TEM) images of mitochondria in variegated and WT-looking cauline and rosette leaves from chimeric TAL-2 transformants. We observed a significant decrease in the size of mitochondria in both cauline and rosette variegated leaves, compared to mitochondria from WT leaves (Fig. IV.11). Additionally, we observed numerous mitochondria with reduced electron density, suggesting a reduced innermembrane network. Both analyses support our conclusion that defects in the development of plants transformed with our mitoTALEN constructs are due to a defect in ATP synthase activity, because of reduced *atp9* expression, leading to profound effects in mitochondria biogenesis.

Discussion

In this Chapter, I present the possible origin of the pleiotropic phenotype developed by several plants transformed with our mitoTALEN constructs, as seen in the experiments described in Chapters II and III. I identified a correlation between the phenotype and a decreased expression of subunit 9 of ATP synthase. The segregation of mitotypes, as discussed in Chapter III, could give clues about the segregation of the phenotype. There appears to be a critical threshold at which loss of ATP synthase leads to an energy shortage in cells, triggering various secondary effects that can cause defects in organ morphology and delayed development. While repair often involves recombination mediated by repeat G, leading to the truncation of *atp9*, a functional selection among the different outcomes of repair that coexist within the cell maintains a sufficient copy number of the gene essential for survival.

Previous studies showed that the loss of ATP synthase in Arabidopsis is lethal (Moore *et al.,* 2003). Our findings are consistent with previous results on Arabidopsis RNA editing mutants with defective ATP synthase, which displayed slow growth and delayed development, and some were seedling lethal (Robison *et al.,* 2009; Hammani *et al.,* 2011). However, our presented line is the first *atp9* gene mutant described, produced via HR in Arabidopsis. Rearrangements involving the *atp9* gene were shown to cause CMS in other species, such like in carrot (Szklarczyk *et al.,* 2014), cauliflower(Dieterich *et al.,* 2003) and sunflower (Makarenko *et al.,* 2018). However, Arabidopsis is a much better model organism for genetic studies. Thus, our mitoTALEN-produced *atp9*-deficient lines open the way for future studies on ATP synthase deficiency, induced retrograde responses, as well as the possible production and study of Arabidopsis CMS lines.

Finally, these results highlight a limitation of mitoTALENs in editing plant mitochondrial genomes. Due to the predominant repair mechanisms for DSB involving repeated sequences distant from the TALEN targeting site, there is always a risk of inadvertently affecting essential sequences located far apart.

	TAL-1	TAL-2
pol1A	3	3
pol1B	8	3
recG1	5	3
oex1 (cross)	1	4
recA3	0	1
radA	4 (+2 cross)	3 (cross)
osb1 osb4	5	1
why2	8	5
odb1	4	4
msh1	2	6
Col-0 (WT)	8	7

Table Conclusions 1. A number of sequenced mitoTALEN mutant plants in different backgrounds. In red, the number of plants produced by crossing heterozygous mutants.

Conclusion & Perspectives

The major aim of my PhD thesis was to utilize mitoTALENs to explore DSB repair in the mitochondria of *Arabidopsis thaliana,* and to acquire a deeper understanding of the roles of key proteins involved in mtDNA maintenance and homologous recombination (HR). Thus, my PhD project started with the generation of mitoTALEN lines, first in the Col-0 WT background and subsequently in mutants deficient in factors of the mitochondrial HR pathways.

At first, I analysed DSB repair in WT plants and devised a pipeline to analyse mtDNA repair in mitoTALEN mutants (discussed in Chapter II). Prior to more detailed analysis of the mitoTALEN lines, I first confirmed that the mitoTALENs are a high efficient tools of mtDNA editing, resulting in complete loss of the cut site early during seed development. Later inspection revealed the creation of mtDNA heteroplasmy, with co-existence of multiple mutant mitotypes, via variable repair pathways. Two major approaches were used to analyse DSB repair outcomes in mutant plants. Firstly, mtDNA enrichment was done on mitoTALEN mutant plants, and DNA was analysed by Illumina paired-end (150 bp) sequencing. With the help of the IBMP bioinformatic platform, I developed two dedicated NGS data analysis pipelines: one to identify HR sites and their relative recombination frequencies, and another to detect MHMR recombination sites. Secondly, I developed a qPCR analysis approach to confirm repair pathways with greater confidence, on a larger pool of T1 mutant plants, thus disclosing the frequency of the alternative repair pathways. In WT plants, three different DSB repair pathways were identified: Homologous Recombination, Single-Strand Annealing, and Microhomology-mediated repair.

The DSB in WT plants was most frequently repaired by HR, yet multiple recombination sites, defined by different pairs of repeated sequences, could be involved in the repair process, leading to the generation of a diverse collection of new mitotypes. Although multiple repeats could be involved, a distinct pattern of preference was observed. I identified that the DSB repair often resulted in the loss of the regions containing the cut site, accompanied by a corresponding increase in copy number of a distal region flanked by the same set of sequences (repeats). I proposed a model where the region containing the DSB is replaced by the duplication of distant regions by homologous recombination involving two sets of repeated sequences, via the break-induced replication (BIR) pathway. Furthermore, I show that DSB repair shows a general preference for HR sites in proximity: TAL-1 mutants preferred recombination involving repeats LR1 and V, while TAL-2 mutants favored recombination involving repeats LR1 and G. However, such selection of an HR site in proximity to DSB repair was not strict, as in TAL-2 mutants, there are multiple recombination sites closer than G that

could be used. It could be assumed that besides the proximity factor, the stability and organisation of resulting recombination outcomes may also influence the selection of HR sites for DSB repair.

In Chapter II, I also observed that repair of mitoTALEN-induced DSBs generates a heteroplasmy of mitotypes between tissues of the same plant and across generations. I showed that heteroplasmic mutant mtDNA can revert to WT in the next generation, if they have maintained the WT mitogenome. In Chapter IV, I show that heteroplasmy of the mtDNA leads to a pleiotropic phenotype observed in certain mitoTALEN mutants. I identified that the phenotype results from the loss of the *atp9* gene, due to recombination involving repeat G during DSB repair. I observed a correlation between reduced levels of the *atp9* mtDNA sequence and of its transcript in phenotypic tissue. This finding suggests that there a critical threshold at which loss of ATP synthase leads to an energy deficit in cells, triggering various secondary effects that cause defects in organ morphology and delayed development. Moreover, different levels of heteroplasmy among different tissues of the same plant leads to a chimeric phenotype, where some leaves have defects in morpholog while others look normal.

Furthermore, analysis of mitoTALEN progeny resulting from a cross between Col-0 TAL-1($^{+/+}$) revealed a more significant loss of the TALEN cut site in older leaves of F1 plants, and no detectable change in the flower parts, compared to Col-0. This may be either because of differential TALEN expression in different plant tissues or because the damaged mtDNA is preferentially discarded in young tissues while being repaired in older tissues, depending on the differential expression and availability of mitochondrial repair factors. However, further studies into these observations were out of the scope of my PhD study. Also, the observation that the TALEN construct loses its efficiency when introduced by cross rather than by agrotransformation suggests that we still don't understand when and where the mtDNA is repaired. This also deserves a more detailed analysis that was beyond the scope of my thesis.

I finish Chapter II with a pilot study in which I hypothesise that mtDNA DSB repair could be directed by another DSB, via the introduction of HR intermediates from two distant genomic loci to the overall repair pool. This might facilitate their linkage via homology, which does not occur under normal circumstances. However, due to the inefficiency of the mitoTALENs introduced by crossing, the study was stopped as it was deemed unfeasible.

The second part of my PhD concerned the study of how a DSB is repaired in the background of mutants deficient in factors of the mitochondrial HR pathways, and deduce the function of the respective factors (discussed in Chapter III). The following lines have been produced with the mitoTALEN constructs: *odb1, why2*, *osb1 osb4*, *radA, msh1, pol1A, pol1B,*

*oex*1 *and recG1* (Table.1). Despite multiple attempts to transform *recA3* plants with the mitoTALEN constructs, just a single *recA3* mutant containing the TAL-2 construction was successfully obtained. This suggests that mtDNA DSBs in *recA3* might be highly toxic, resulting in non-viable seed. From NGS analysis, it became apparent that when HR activity is limited in the mutant plants, different repair pathways can be recruited to reconstitute a stable genome via Single-Strand Annealing (SSA), as observed in *odb1* mutant plants or by microhomology-mediated recombination (MHMR), as observed in *why2* mutants. However, in *pol1A and pol1B* mutants, there was no difference in the repair of DSB compared to WT, but I did observe that mtDNA coverage profiles were slightly distorted, with ups and downs that suggested putative sites of replication initiation. In both *pol1A and pol1B* mutants, I observed a reduction in mtDNA copy number, yet only *pol1B* plants were more sensitive to replicative stress induced by hydroxyurea treatment. This indicated that both DNA polymerases are involved in DNA replication, but that only POL1B is preferentially involved in mtDNA repair. Mutants of *oex1*, *radA*, *osb1osb4*, *msh1*, *recA3* and *recG1* showed major structural reorganisation of the mtDNA, already in the absence of mitoTALEN-induced mtDNA damage. Upon DSB introduction, no repair fingerprints identified in Chapter II were observed in *radA*, *osb1osb4* and *recA3* mutants. We concluded that in these mutants, a viable mitogenome is reconstituted from autonomous replication of pre-existing sub-genomes created by recombination between direct repeats. Thus, in these mutants, the DSB is segregated out rather than repaired, suggesting that these proteins are essential for HR-dependent repair in plant mitochondria. Inversely, *recG1* and *msh1* mutants displayed repair fingerprints, and read-pairs analysis suggested that an increase in ectopic recombination was much greater than in other mutants. All this suggested that MSH1 and RECG1 are involved in the surveillance of HR, possibly by suppressing events involving hemeologous sequences.

In Chapter IV, I show that the mitoTALEN line we generated could be the first *atp9* gene mutant described described so far, produced via HR in Arabidopsis. Previous studies in other species have shown that deficiency in *atp9* gene expression might lead to CMS. However, Arabidopsis is a much better model organism for genetic studies. Thus, our mitoTALEN-produced *atp9*-deficient lines open the way for future studies into ATP synthase deficiency-induced retrograde responses, and the potential generation and study of Arabidopsis CMS lines.

Finally, my studies highlight the limitations of mitoTALENs in editing plant mitochondrial genomes. Given that HR is the major repair pathway, introducing DSBs into the mtDNA always carries the risk of inadvertently impacting essential sequences located at distant sites. However, this problem for mtDNA editing might be mitigated if performed in the *why2* mutant background, in which DSB repair by MHMR is increased, thereby reducing the risk of affecting regions located far apart.

Materials & Methods

Materials

1. Bacterial Strains

1.1. Agrobacterium tumefaciens GV3101 pMP90

Agrobacterium tumefaciens is a soil phytopathogen that naturally infects plant wound sites, causing crown gall disease via the delivery of transferred T-DNA to the host plant via a bacterial type IV secretion system. The strain used is a derivative of C58C1, employed for stable transformation of *A. thaliana* by floral dip transformation (Bechtold and Pelletier, 1998). It contains rifampicin resistant gene (*rif*) in nuclear genes as a screening label. GV3101 carries a disarmed *pMP90* Ti plasmid that contains *vir* genes for the transfer of foreign binary vector T-DNA. Additionally, pMP90 carries a gentamicin resistance gene.

1.2. Escherichia coli TOP 10

E. coli is a gram-negative bacteria commonly found in the lower intestine of warm-blooded organisms. Chemically competent TOP10 strain cells were used for standard cloning and plasmid amplification. The mutations recA1 and endA1 benefit the stability of inserted DNA and the extraction of high-purity plasmid DNA.

Genotype of TOP10*: F^- mcrA (mrr-hsdRMS-mcrBC) 80lacZ M15 lacX74 recA1 ara139 (ara-leu)7697 galU galK rpsL (StrR) endA1 nupG.*

2. Arabidopsis thaliana plant lines

A. thaliana accession Columbia-0 (Col-0) was used as a model organism to study homologous recombination. This accession is generally used because most mutant lines available are in the Col-0 genetic background, including the lines deficient in mitochondrial HR factors that were used in my studies. T-DNA lines were obtained from the Nottingham Arabidopsis Stock Centre (*odb1-1*: SAIL_25_H08, *why2-1*: SAIL_582_A10, *osb1-1*: GK-093H12, *osb4-1*: SALK_018261, *radA-3*: SALK_*097880, msh1:* SALK_046763*, pol1B-1:* SALK_134274*, pol1B-2:* GK-291F05*, pol1A-1:* SALK_022638, *pol1A-2*: SALK_020119, *recA*3-1: SAIL_1278, *recA*3-2: SALK_146388, *recA*3-3: GK-288B01, *oex1-1*: GK-911E05). *osb1-1 osb4*-1 double mutant was obtained via crossing of the corresponding lines*. recG1-2*: Flag_630A09 line in accession Wassilewskija (Ws) background was retrieved from the Versailles Arabidopsis Stock Centre and was subsequently introgressed into Col-0.

Name	Concentration
Gentamicin	50 µg/ml
Rifampicin	50 µg/ml
Kanamycin	50 µg/ml
Streptomyscin	50 µg/ml

Table MM.1. Antibiotics used in this project for Bacterial selection.

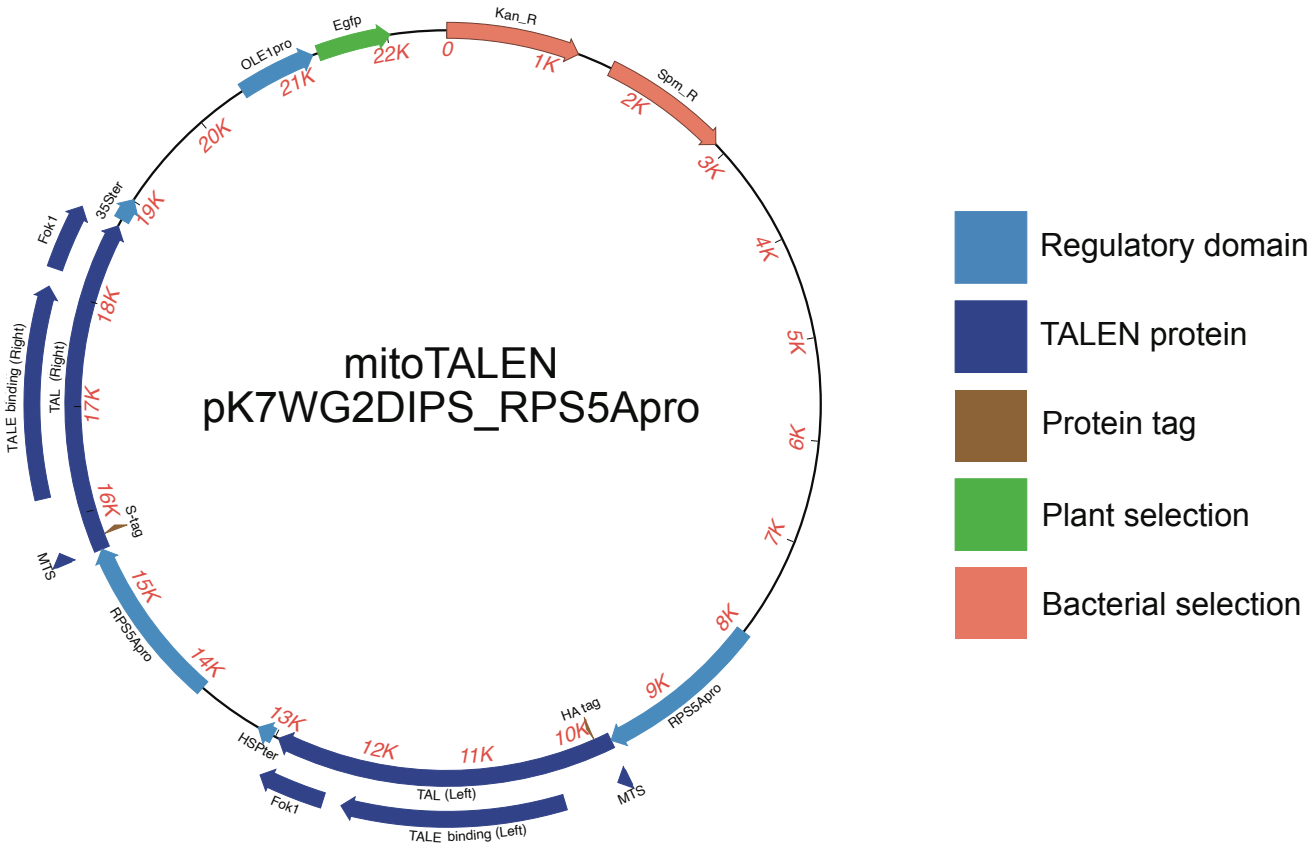


Figure MM.1. mitoTALEN constructs in binary vectors. Different domains are indicated: FokI endonuclease, TALE binding domain, and mitochondria targeting sequence (MTS). The TALEN genes are under the control of the constitutive RPS5 promoter. The T-DNA constructions contain the marker eGFP gene expressed under the OLE1 oleosin promoter on the seed coat.

3. Growth media

3.1. LB medium

Lysogeny broth medium (1 % tryptone (w/v); 0.5 % yeast extract (w/v); 0.5 % NaCl (w/v); pH 7.4) was used for bacterial growth. For bacterial selection on plates, the medium was complemented with suitable antibiotics (Table MM.1) and solidified with 1.5 % agar.

3.2. SOC medium

Super optimal medium with catabolic repressor (2 % tryptone (w/v); 0.5 % yeast extract (w/v); 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 20 mM Glucose; pH 7.0) was used for primary growth in the recovery step of bacterial competent cells transformation.

3.3. MS medium

Murashige and Skoog media (micro and macro elements M0255 (DUCHEFA BIOCHEMIE B.V., Haarlem, Netherlands); 1 % sucrose (w/v); 1 % agar (w/v); pH 5.8) was used for *in vitro* growth of *A. thaliana* plants on petri dish.

4. Gateway plasmid vectors

Gateway vectors are designed for cloning experiments based on att recombination sites using recombination enzyme mix (BP and LR clonases). Genes flanked by matching att sites can be transferred from donor to destination vectors (Hartley *et al.,* 2000). All gateway vectors contain the toxic *ccdB* gene for negative selection. For this project, destination vectors pK7WG2DIPS_RPS5Apro with TALEN constructs were provided by Prof. Arimura team (University of Tokyo). Constructs were made via the multi-gateway system, where two TALEN genes, preceded by N-terminal mitochondrial targeting sequences, were combined and expressed under the constitutive *RPS*5A promoter. A marker *GFP* gene is expressed under the *OLE1* oleosin promoter. These components were integrated into the final binary vector derived from pK7WG2 (Fig. MM. 1).

5. Oligonucleotides

The sequences of all primers used in this project are provided in the appendix 1 section.

6. Informatic tools

6.1. Command line tools

6.1.1. Samtools (v1.18)

A set of bioinformatic tools written in C language for manipulating next-generation sequencing data. It was used to convert the Sequence Alignment Map (SAM) into the Binary Alignment Map (BAM) format, filter mapped reads, and extract them in FASTA or FASTQ formats.

6.1.2. Burrows-Wheeler Aligner (BWA) (v0.7.17-r1188)

BWA (Li and Durbin, 2009) is a software package for mapping low-divergent sequences against a large reference genome. The BWA-MEM algorithm supports long-reads and split alignment. It is recommended for high-quality queries of 70-100 bp Illumina reads due to its faster and more accurate performance. The software works by seeding alignments with maximal exact matches (MEMs) and then extends the seeds with the affine-gap Smith-Waterman algorithm (SW).

6.1.3. Bowtie2 (v2.5.2)

Bowtie2 (Langmead and Salzberg, 2012) is an ultrafast and memory-efficient mapping tool designed for long reference sequences. It is usually used as the first step in pipelines for comparative genomics. Bowtie2 is considered more accurate than BWA, however, more split reads and soft-clips were found with BWA. Thus Bowtie2 was selectively used for remapping the soft-clipped parts during the analysis of microhomologies.

6.1.4. GATK (v4.5.0.0)

The Genome Analysis Toolkit (GATK) is a set of tools designed for efficient and robust analysis of next-generation DNA sequencers (NGS). The MarkDuplicates (Picard) sub-package was used to identify and remove duplicate reads.

6.1.5. SE-MEI

SE-MEI is a tool package deposited in GitHub by Devon Ryan: https://github.com/dpryan79/SE-MEI. It is a repository containing multiple scripts for finding mobile element insertion (MEI) sites from single-end datasets. The C script named “extractSoftclipped” was used to extract the soft-clipped portions of reads for subsequent mapping to the mtDNA.

6.1.6. Subread (v2.0.6)

It is a package of software programs for the analysis of NGS reads. The tool called “featureCounts” was used to count the number of reads mapped to the regions flanking repeats. It is a software designed for counting reads to genomic features such as genes, exons, promoters and genomic bins (an artificial separation of the genome to equal size regions).

6.1.7. Snakemake (v2.0.6)

This is a package that was used to organise scripts in specific workflows. Thus, the script can be easily executed for multiple data sets and adapted to the data source.

6.2. R studio packages

RStudio is an environment for R language programming. It is a language widely used for statistical computing and graphics representation.

6.2.1. Data management packages

Packages data.table and dplyr were used for fast organisation and management of big data frames.

6.2.2. ggplot2

The ggplot2 is a data visualization package that creates graphics, based on “The Grammar of Graphics”. This package was used to create pie and bar charts, as well as coverage plots. The following extension packages were additionally used for data visualisation: ggraph, igraph, ggrepel, ggpubr, grid, patchwork.

6.2.3. circlize

The Circlize is a package for visualizing huge amounts of data in circular representation. It was used to create plots representing the links between recombining repeats. The ComplexHeatmap package allowed the creation of links with colour according to the number of events found.

6.2.4. Rsamtools

It is a package allowing the use of samtools tools (presented in Methods 5.3.1) in the Rstudio environment.

6.2.5. Viridis

It is a companion package for data visualisation. It provides different series of colour maps that can be used within the figures produced by ggplot and circlize.

6.2.6. Cli

It is a package that creates a command-line interface in the Rstudio environment.

6.2.7. rtracklayer

Rtracklayer tools allows to create, import and manipulate annotation tracks from genome browsers in various formats (GFF, BED).

6.3. MacVector (v18.5.1) and its features

It is a commercial sequence analysis application developed for the Macintosh operating system. During PhD, it has been used for quick DNA analysis and data visualisation.

6.3.1. Primer design feature

The primer design feature has been used to design all the primers for genotyping and PCR and qPCR experiments. MacVector uses Primer3 algorithm to propose the best primers according to given PCR conditions.

6.3.2. Dot plot feature

MacVector use Pustell DNA Matrix function to create a dot plot according to the alignment using ClustalW function of two given sequences. I was used for the analysis of reads confirming the MHMR, to confirm that they map to two distinct locations.

6.3.3. Protein Analysis toolbox

It contains multiple algorithms that can analyse the composition of protein sequences. In the current project, it was used to predict the Molecular weight of the TALEN protein in western blot.

6.3.4. Macvector Assembler

Macvector Assembler is a Macvector extension that uses various algorithms to assemble reads into De Novo contigs or to align them against reference genomes to produce a coverage map. The bowtie aligner algorithm (also described in section 5.3.2) of MacVector has been used for the initial analysis of Illumina sequencing reads by producing a read coverage map over the mitochondrial genome.

Methods

1. Quantification and quality assessment of samples

1.1. Spectrophotometric quantification

1.1.1. Nucleic acid quantification

For simple and rapid quantification of DNA and RNA, the *NanoDrop 2000* (Thermo Fisher Scientific, Waltham, Massachusetts, USA) spectrophotometer was used. Both nucleic acids are quantified by their absorbance at 260 nm (A260), and the machine cannot determine from the sample whether it is DNA or RNA, leading to potential overestimation of concentration readings. Furthermore, sample degradation is not considered. The ratios A260/A280 and A260/A230 allowed us to estimate the contamination by proteins, phenols or other contaminants that absorb strongly near 230 nm or 280 nm wavelengths. Samples are considered virtually pure when ratio A260/A280 is ~1.8 (DNA) or ~ 2 (RNA), and ratio A260/A230 is in the range of 2.0 – 2.2.

1.1.2. Protein quantification

The Bradford test was used to determine protein concentrations (Bradford, 1976). Bradford reagent (Bio-Rad, Hercules, California, USA) is a dye that binds proteins unspecifically, based on the change of absorbance of Coomassie G-250 from 450 nm to 595 nm in the presence of proteins. The quantification is carried out using the Protein Assay Dye Reagent Concentrate (BioRad®), diluted one-sixth in water, and protein concentration is estimated from absorbance at 595 nm (Shimadzu™ UV 1800 Spectrophotometer, Thermo Fisher Scientific, Waltham, Massachusetts, USA) and compared to a standard curve established with known concentrations of BSA. The method holds significant limitations, as it is very sensitive to contaminants, in particular to detergents that are often used during protein extraction.

1.2. Fluorometric quantification

For sequencing purposes, highly accurate quantification of nucleic acids is required. The Qubit® 3.0 Fluorometer (Invitrogen, Waltham, Massachusetts, USA) was used in all step of DNA library preparations using high sensitivity (HS) reagents. This technique uses fluorometric dyes specific to DNA or RNA and is more sensitive and reliable than NanoDrop. Qubit was also used to quantify RNA during cDNA preparation, using broad range (BR) reagents.

2. DNA analysis methods

2.1. Plant DNA extraction

2.1.1. DNA extraction with Edwards buffer

Crude DNA extraction from leaves by the Edwards protocol was done for genotyping purposes (Edwards *et al.,* 1991). Leaves were frozen in liquid nitrogen and ground with metal beads (∅ 3 mm) using TissueLyser (QIAGEN, Germantown, Maryland, USA) for 2 minutes at 25 Hz. The ground tissue was resuspended in 400 µL of Edwards buffer (200 mM Tris-Cl, 250 mM NaCl, 25 mM EDTA, 0.5 % SDS (v/v), pH 7.5). After centrifugation for 15 min at 3,200 g (in plates) or 2 min at 15,000 g (in Eppendorf) the supernatant was recovered, and nucleic acids precipitated with one volume of isopropanol. After incubation at room temperature for 5 min, nucleic acids were sedimented by centrifugation for 20 min at 3,200 g (in plates) or for 5 min at 15000 g (in Eppendorf tubes). The pellet was washed with 70% ethanol and resuspended in TE (10 mM Tris-Cl, 0.1 mM EDTA, pH 8.0) or autoclaved milliQ water.

2.1.2. DNA extraction with CTAB

When purer plant DNA was required, as for qPCR analysis, extraction was done using cetyltrimethylammonium bromide (CTAB). Leaves or seedlings were frozen in liquid nitrogen and ground with metal beads using TissueLyser (QIAGEN, Germantown, Maryland, USA) for 2 minutes at 25 Hz. The ground tissue was resuspended in 400 µL CTAB buffer (2 % (w/v) CTAB, 1.4 M NaCl, 20 mM EDTA, 100 mM Tris-HCl, pH 8.0) preheated to 65 °C and supplemented with 0.2 % (v/v) β-mercaptoethanol and 50 mg/mL of RNase A. After incubation at 65 °C for 1 hour, 400 µL of chloroform was added, and after agitation, the aqueous and organic phases were separated by centrifugation for 10 min at 12,000 g. The DNA in the aqueous phase was precipitated with an equal volume of isopropanol. For more quantitative purification or when the sample amount was small, it was incubated at -20 °C overnight before centrifugation for 10 min at 4 °C 15,000 g. The DNA pellet was washed twice with 70% ethanol and resuspended in TE (10 mM Tris-Cl, 0.1 mM EDTA, pH 8.0) or autoclaved milliQ water.

2.1.3. DNA extraction for sequencing

2.1.3.1. Mitochondrial enrichment

Mitochondrial enrichment was done for mtDNA sequencing of individual plants to increase coverage and allow cost-effective multiplexing of multiple samples. Full rosettes of four- to six-week-old T1 first transformants, grown on soil, were taken for mtDNA enrichment by differential centrifugation as described previously (Weber-Lotfi *et al.,* 2023). All steps were

carried out in the cold room, and centrifugations were at 4 °C. The full rosettes were ground using mortar and pestle with 20 mL cold grinding buffer (0.3 M sucrose, 60 mM TES, 10 mM EDTA, 10 mM KH_2PO_4, 25 mM tetrasodium pyrophosphate, 1 mM glycine, 1 % (w/v) polyvinylpyrrolidone-40, 1 % (w/v) BSA, 50 mM sodium ascorbate, pH 8.0), supplemented with 20 mM cysteine before use. The mix was filtered through Miracloth (Merck®) and a 20 µm nylon mesh, and centrifuged at 1,000 g for 10 min to remove most nuclei, chloroplasts and thylakoid membranes. The supernatant was transferred to a new tube and centrifuged at 2,500 g for 10 min. The crude mitochondrial pellet was obtained by subsequent centrifugation of supernatant at 15,000 g for 20 min, and washed with 1.5 mL wash buffer (0.3 M sucrose, 10 mM TES, 2 mM EDTA, 10 mM KH_2PO_4, pH 7.5) in a 2 mL Eppendorf tube. The pellet was frozen in liquid nitrogen and stored at -80 °C for later DNA extraction.

2.1.3.2. DNA extraction for Illumina library preparation

The DNA from enriched mitochondria was extracted using QIAprep Spin Miniprep Kit (QIAGEN, Germantown, Maryland, USA), incubating the sample in 180 µL ATL buffer with 20 µL RNase A at 37 °C for 30 min before addition of Proteinase K. As a modification to established kit protocol, if the sample contained too much tissue material that could block the columns, the sample was centrifuged at max speed before addition of ethanol to remove undissolved tissue.

2.2. DNA agarose gel electrophoresis

Agarose gel electrophoresis is a method to separate nucleic acids in a matrix of agarose. Different agarose concentrations (0.7 % up to 3 % (w/v)) were used according to the size of nucleic acids needed to be separated. The gel was prepared in Tris-acetate-EDTA (TAE) buffer (20 mM Tris-acetate, 0.5 mm EDTA, pH 8.0). For nucleic acid detection, 0.5 µL/mL of ethidium bromide (10 µg/mL stock solution) is added to the gel. It is an intercalator that binds to DNA, making it visible under UV light. Samples are supplemented with 0.2 volume of Loading buffer (10 mM Tris-HCl, 60 % glycerol (v/v), 60 mM EDTA, 0.03 % bromophenol blue (w/v), pH 7.6). Samples are migrated in TEA running buffer at 100 V for 20-60 min depending on the separation needed. The size of DNA fragments is estimated by comparison to size markers of known size run in parallel (80 bp to 10,000 bp, MassRuler DNA ladder, Thermo Fisher Scientific, Waltham, Massachusetts, USA).

2.3. Polymerase chain reaction (PCR)

PCR is a method to *in vitro* amplify a specific DNA sequence from small amounts of template, using heat-resistant DNA polymerases. A pair of oligonucleotides flanking the target site act as primers for initiating polymerase activity, creating a copy of the selected sequence. PCR involves a repeated set of reactions, which constitute a PCR cycle. Each cycle starts

	Genotyping			Long range PCR*				
Step	Number of cycles	Phire Hot Start		Number of cycles	Phusion high fidelity		Platinum™ SuperFi II	
		Temperature	Time		Temperature	Time	Temperature	Time
Initial denaturation	1	98℃	30 sec	1	98℃	1 min	98℃	1 min
Denaturation	30	98℃	10 sec	35	98℃	1 min	98℃	1 min
Annealing		55℃	15 sec		64℃	1 min	64℃	1 min
Extension		72℃	15 sec		72℃	10 min	68℃	10 min
Final extension	1	72℃	2 min	1	72℃	10 min	68℃	10 min

Table MM.2. Conditions of Polymerase Chain Reaction. *Long range PCR with Phusion polymerase was successful only in amplifying large SSA products (949 bp and 2559 bp) in Chapter II, but not the product of HR involving repeats LR1 and V (9453 bp and 9390 bp). HR product amplification was successful only with Platinum SuperFi II.

Step	Number of cycles	Temperature	Time
Initial denaturation	1	95 ℃	7 min
Denaturation	40	95 ℃	10 sec
Annealing		58 ℃	15 sec
Extending		72 ℃	20 sec
Melting curve	1	95 ℃	5 sec
		55 ℃	1 min
		55–95 ℃	0.11 ℃/ sec
		40 ℃	30 sec

Table MM.3. Conditions of quantitative Polymerase Chain Reaction

with high-temperature denaturation of double-stranded DNA, releasing the single-strand DNA templates to which the designed primers can base-pair at their target sites. After the annealing of primers, DNA elongation is catalysed by the DNA polymerase. These three steps (denaturation, annealing and elongation) constitute a single PCR cycle that is repeated 30-35 times, resulting in the exponential amplification of the target DNA fragment.

During my experiments, three different DNA polymerases were used due to their differences in fidelity, processivity and stability. Phire Hot Start II DNA polymerase (Thermo Fisher Scientific, Waltham, Massachusetts, USA) was used for genotyping, due to its processivity and stability, less sensitive to inhibitors present in crude DNA preparations than the classical *Taq* DNA polymerase. For amplification of long DNA fragments, higher polymerase fidelity and fragment stability were required: Phusion High Fidelity DNA polymerase (Thermo Fisher Scientific, Waltham, Massachusetts, USA) or Platinum SuperFi II DNA polymerase (Thermo Fisher Scientific, Waltham, Massachusetts, USA). Specific PCR conditions were used for different polymerases (Table MM.2).

2.4. Quantitative polymerase chain reaction (qPCR)

qPCR is a method for the relative quantification of nucleic acid sequences. The principle of qPCR and PCR is the same, with the difference that the total DNA is quantified after each cycle by measuring the fluorescence of a DNA binding dye. The qPCR conditions can be found in Table MM.3. From the amplification curves, a detection threshold value is determined (Cp: crossing point), which is directly proportional to the initial amount of DNA. In the current projects, 384 well microwell plates were used to perform qPCR reactions in the LightCycler 480 II (ROCHE, Basel, Switzerland). Each reaction was made in a total volume of 6 µL containing 0.5 volumes of SYBR Green I Master mix (ROCHE, Basel, Switzerland), 1.5 mM of each primer and about 2 ng of total plant DNA. The 2nd derivative maximum method was used to determine Cp values, roughly corresponding to the cycle where the increase in amplification rate is maximum. Each reaction was conducted in triplicate, and the results were normalised to reference genes and represented as the difference to the control sample. Primer efficiency was either arbitrarily assumed to be 1.9 or determined using LinRegPCR software (http://LinRegPCR.nl).

2.5. Bacterial transformation

Bacterial transformation by heat shock was performed in current projects. The bacterial membrane of competent cells, suspended in a calcium-rich media, upon heat shock, becomes permeable to DNA molecules. 50 µL of frozen (-80 °C) competent bacteria thawed on ice were incubated with 0.1-100 ng of plasmid DNA for 30 min. Heat shock was carried out on a water bath (45 sec at 42 °C for *E. coli* ; 5 min at 37 °C for *A. tumefaciens)* and put back on ice for

2 min. 1 mL of SOC media was added to the transformation reaction and incubated for 1 h (at 37 °C for *E.coli* ; at 28 °C for *A. tumefaciens)* allowing resistance genes to be expressed and bacteria to proliferate. Finally, cells were plated on Petri dishes containing LB agar media supplemented with appropriate antibiotics.

2.6. Plasmid DNA extraction

To purify amplified plasmid DNA from *E. coli*, 2 mL of culture grown overnight was pelleted by centrifugation and DNA extracted using the Nucleospin Plasmid EasyPure kit (Macherey-Nagel, Düren, Germany).

2.7. Sequencing

2.7.1. Sanger sequencing

DNA Sanger sequencing was performed by the “Gene expression analysis” platform of the IBMP (Mr. A. Alioua and Ms. S. Koechler) by providing them with 15 µL of sample (600 ng for plasmid DNA; 100 ng for PCR products) supplemented with 20 pmoles of sequencing primer. The method relies on DNA extension chain-termination by dideoxynucleotides incorporation, which lacks the 3'-OH group required for phosphodiester bond formation (Sanger *et al.,* 1977). Resulting DNA fragments are separated by capillary electrophoresis on the ABI PRISM 3100 Genetic Analyser (Applied Biosystem, Foster City, California).

2.7.2. Illumina sequencing

DNA for sequencing (section 2.1.3.2.) was first cleaned using the NucleoMag kit (*Macherey-Nagel, Düren, Germany*) and quantified using the Qubit® 3.0 Fluorometer (Invitrogen, Waltham, Massachusetts, USA). A total of 50 ng of the sample was used for DNA library preparation following the Nextera™ DNA Flex Library Prep protocol (Illumina, San Diego, California, United States). Library preparation starts with DNA tagmentation, where a Bead-Linked Transposome fragments the DNA and adds adaptor sequences. Afterwards, DNA fragments are amplified using primers complementary to the adaptor sequence. Primers for amplification also contain and add to the amplified fragments the index sequences that allow tracing the fragments to a specific sample in the library pool.

Libraries were analysed on Agilent 2100 Bioanalyzer (Agilent, Santa Clara, California, USA) and quantified with the Qubit® 3.0 Fluorometer (Invitrogen, Waltham, Massachusetts, USA). Equal moles of each library were pooled and SPRIselect beads (BeckmanCoulter, Brea, California, USA) were used to clean library pools from adaptors. Libraries were pair-end sequenced (2 x 150 reads) on a NextSeq 2000 or on a MiSeq sequencing systems. MiSeq sequencing was performed by Ms. S. Koechler of the “Gene expression analysis” platform of

2 µg RNA (extracted from kit) 1 µL of 50 µM oligo-dT 1 µL of 200 ng/µL random primers 1 µL of 10 mM dNTP 13 µL H_20
5 min incubation at 65 ºC
5 min incubation on ice
4 µL of 5X Superscript IV (invitrogen) 1 µL of 100 mM DTT 1 µL of 40 U/µL RNaseOUT (invitrogen) 1 µL of 200 U/µL Superscript IV (invitrogen)
5 min incubation at 25 ºC
60 min incubation at 50 ºC
15 min incubation at 70 ºC

Table MM.4. Reverse-transcription protocol for production of cDNA from pure DNA free RNA samples.

the IBMP, and NextSeq 2000 sequencing at the Genomax sequencing platform (www.france-genomique.org).

3. RNA analysis methods

3.1. RNA extraction

Young *A. thaliana* leaves (about 100 mg) were frozen in liquid nitrogen and ground with metal beads on a TissueLyser (QIAGEN, Germantown, Maryland, USA) for 2 minutes at 25 Hz.

3.1.1. Total RNA extraction with Trizol

Ground tissue is resuspended in 800 µL TRIzol™ reagent (Life Technologies, Carlsbad, California, USA). After adding 200 µL of chloroform, the suspension is incubated on ice for 5 min, with occasional agitation, and centrifuged for 10 min at 16,000 g. The aqueous phase containing the RNA is collected and precipitated with 0.8 volumes of Isopropanol after the addition of 1/10 volume of 3M Sodium Acetate. After centrifugation at 16,000g for 20 mins, the resulting pellet was washed with 70% EtOH and resuspended in autoclaved MilliQ water.

3.1.2. RNA extraction for cDNA preparation

For cDNA preparation, RNA was extracted using the NucleoSpin RNA Plant isolation kit *(Macherey-Nagel, Düren, Germany).*

3.2. cDNA preparation

Complementary DNA (cDNA) was synthesised using RNA-dependant DNA polymerase Superscript IV (Invitrogen, Waltham, Massachusetts, USA). The levels of cDNA are proportional to transcribed RNA. Thus, cDNA quantification by qPCR (RT-qPCR) allows us to determine relative gene expression levels. Results need to be standardized against the RT-qPCR amplification of reference genes whose expression is constant on the tissues analysed. The steps of cDNA synthesis are presented in Table MM.4.

3.3. RNA agarose gel electrophoresis

The method for RNA analysis on agarose gels is as for DNA, described in Method section 2.2., but with additional care to avoid degradation by RNases. All buffers need to be prepared fresh and filtrated.

3.4. Northern blot

The Northern blot is a method used to analyse RNA by size and detect target RNA molecules by hybridization with complementary DNA or RNA probes. The experiment starts with denaturing RNA gel electrophoresis. The gel is prepared 1 hour before run (1.2 % agarose

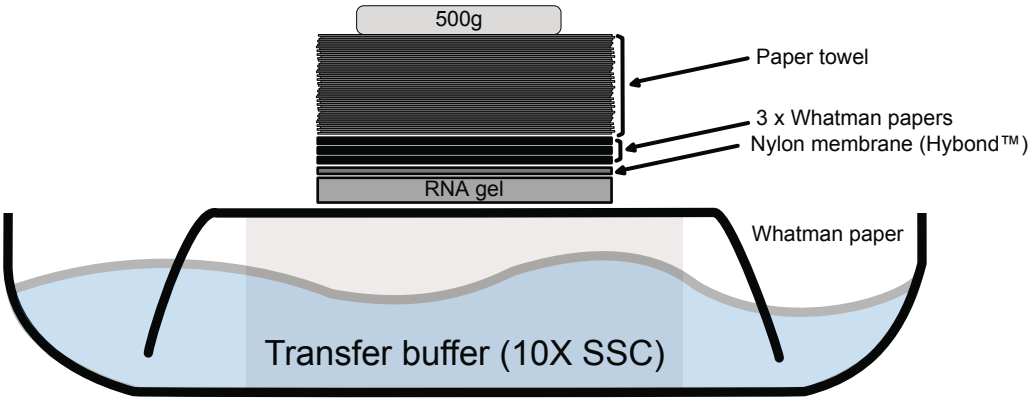


Figure MM.2. Schematics of a "sandwich" set up for Northern blot transfer of RNA from agarose gel onto a positively charged Nylon membrane (Hybond™-N^+) via the capillary method.

(w/v), 10 % (v/v) formaldehyde, 1X MOPS buffer (0.2 M MOPS, 5 mM EDTA, 50 mM sodium acetate, pH 7.0) and let to solidify under the hood. Samples were prepared by mixing 1 volume of RNA (3-5 µg) with 2 volumes of RNA sample buffer (1.3X MOPS, 8% formaldehyde, 65 % (v/v) formamide) and incubated for 10 min at 65 °C to denature the RNA. Before loading, the samples were supplemented with 3 µL of loading buffer (10 mM Tris-HCl, 60 % glycerol (v/v), 60 mM EDTA, 0.03 % bromophenol blue (w/v), pH 7.6). The samples and size marker were run in 1X MOPS buffer at 35 V until the dye front reached the bottom of the gel.

After gel electrophoresis, the RNA is transferred by capillarity from the gel to a positively charged nylon membrane (Hybond™-N+, Thermo Fisher Scientific, Waltham, Massachusetts, USA). The scheme of the transfer “sandwich” is shown (Fig. MM. 2). The transfer is done in 10X SSC buffer (3 M NaCl, 0.3 M sodium citrate, pH 7.2). The RNA gets transferred to the membrane by capillarity overnight. The RNA is then bound to the membrane by UV-crosslinking (120 mJ/cm2), using a UVC5000 Ultraviolet Crosslinker-230V (Hoefer, San Francisco, California, USA). To control the transfer, the RNA bound to the membrane is visualised by methylene blue staining (0.5M sodium acetate pH 5.2, methylene blue 0.04%) for 15 mins, followed by several washes with MilliQ water.

Specific RNA molecules are detected by hybridization with a labelled complementary DNA probe. It is prepared from corresponding purified PCR fragments labelled with [α-^{32}P]dCTP by random priming, using the Ambion™ DECAprimeTM DNA labelling kit (Thermo Fisher Scientific, Waltham, Massachusetts, USA). The northern blot membrane was pré-incubated for 1 h at 65 °C in Church buffer (7% SDS, 0.5 M sodium phosphate, 1 mM EDTA, pH 7.0) prior to hybridization. Hybridization with the probe was done overnight, in 5 mL of Church buffer at 65 °C. After extensive washes with 2x SSC (0.1%), the radioactive bands were visualised with an Amersham Typhoon™ phosphor imager (Cytiva, Washington, D.C., USA).

4. Protein analysis methods

4.1. Protein extraction

4.1.1. Total plant protein extraction

An Arabidopsis rosette leaf was ground in 1.5 mL Eppendorf with extraction buffer (50 mM Tris-HCl, 300 mM NaCl, 1% (v/v) TritonX-100, 1 mM DTT, Complete Protease Inhibitor cocktail (Roche), pH 8.0) using a plastic pestle. After centrifugation at 4°C for 10 mins at 21,000 g the supernatant is collected in a new tube and frozen in liquid nitrogen for storage at -80 °C. Extracts were used for Western blot detection of the mitoTALEN HA-tag.

Speed	Time
3200 g	5 min
3200 g	5 min
6000 g	5 min
18000 g	10 min

Table MM.5. Centrifugation steps taken for mitochondria purification from full rosette extracts before loading final pellet onto Percoll gradient.

4.1.2. Mitochondrial protein extraction

Five 6 weeks-old full rosettes were homogenized in 45 mL grinding buffer (0.3 M sucrose, 5 mM tetrasodium pyrophosphate, 2 mM EDTA, 10 mM potassium dihydrogenphospahte, polyvinylpyrrolidone (PVP-40), 1 % (w/v) BSA, 20 mM sodium L-ascorbate, 5 mM cysteine, 1 mM DTT, Complete Protease Inhibitor cocktail (Roche), pH 7.2) using a Waring blender. The suspension was filtered through two layers of nylon cloth (∅ 96 µm). To remove tissue debris and most of the chloroplasts and thylakoids, the suspension was submitted to differential centrifugation at 4 °C (Table MM.5). To purify mitochondria, the final pellet was resuspended in 200 mL of wash buffer (0.3 M sucrose, 20 mM MOPS, 1mM EGTA, pH 7.2) and loaded on a discontinuous Percoll® gradient and centrifuged at 21,000 g for 20 min, with no break. The gradient was composed of three layers of different Percoll solutions, 45 %, 29 % and 18 %, in gradient buffer (0.6 M sucrose, 20 mM MOPS, pH 7.2). Mitochondria were recovered from the 29-18 % interphase. After two washes with washing buffer and centrifugation for 20 mins at 21,000 g, the purified mitochondria pellet was frozen in nitrogen and conserved at -80 °C. Prepared extracts will be analysed using blue native gel to check the integrity of the ATPase complex, yet due to time constraints, this could not be done before the defense of this thesis.

4.2. SDS-PAGE electrophoresis

SDS-PAGE gel is a method to separate denatured proteins according to their size in an electric field (Laemmli, 1970). SDS denatures the proteins and gives them a uniform negative charge. Proteins are separated in the gel matrix formed by polymerised polyacrylamide/ bis-acrylamide (37.5:1). The top stacking gel contains 5% (w/v) acrylamide in 125 mM Tris-Cl pH 6.8, 0.1 % (v/v) SDS. It allows proteins to concentrate on top of the gel and uniformly enter the separation gel, whose concentration determines the protein sizes that are best resolved. The best resolution was achieved with gradient gels (7%-17 % acrylamide in 375 mM Tris-Cl pH 8.8, 0.1 % (v/v) SDS). Protein samples before loading are denatured at 96 °C for 10 mins with 4X sample buffer (8 % SDS (p/v), 40 % glycerol, 20 % 2-mercaptoethanol, 0.008 % bromophenol blue, 0.250 M Tris-HCl, pH 6.8). A protein ladder is run in parallel (Prestained protein marker page ruler plus, Thermo Fisher Scientific, Waltham, Massachusetts, USA). The migration was done in running buffer (25 mM Tris, 192 mM glycine, 0.1 % (w/v) SDS) with a constant current of 20-25 mA until the dye front reached the bottom of the gel. Proteins in the gel were visualised by Coomasie Blue staining (10 % (v/v) acetic acid, 40 % (v/v) methanol, 0.25 % (w/v) Coomassie Brilliant Blue R250). Afterwards, the gel was washed with destaining solution (20% (v/v) ethanol, 7% (v/v) acetic acid).

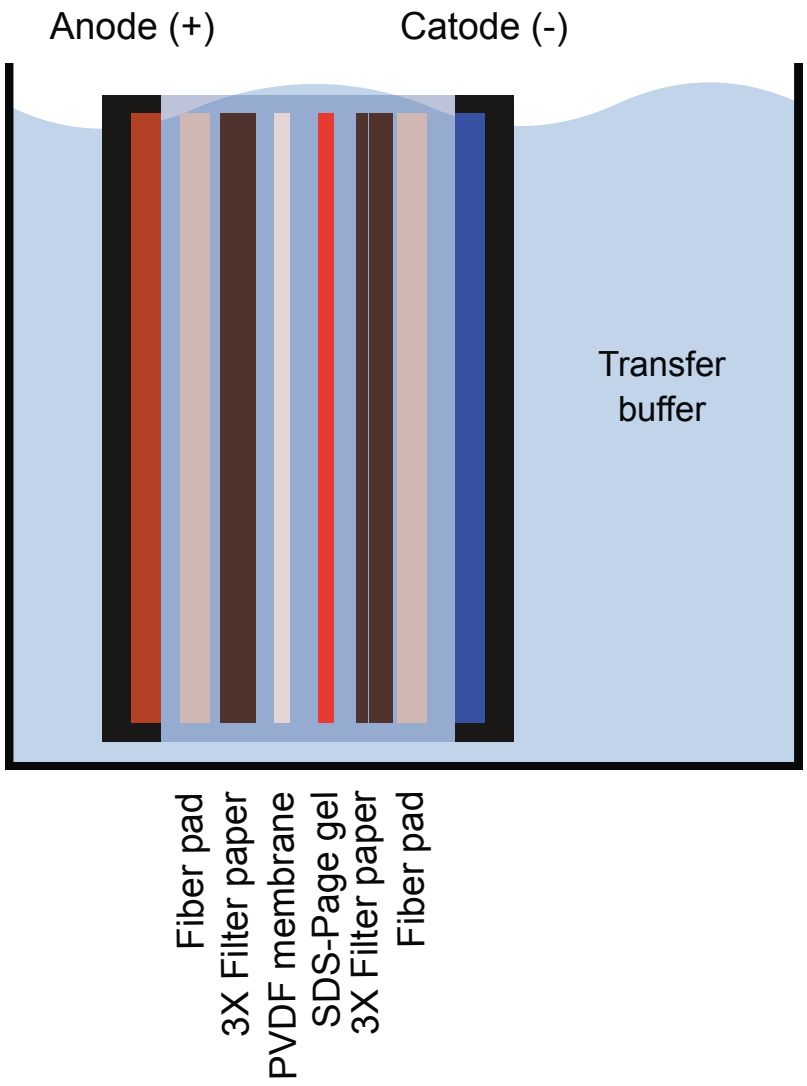


Figure MM.3. Schematics of a "sandwich" set up for Western blot transfer of Proteins from an acrylamide gel onto a PDVF membrane (Immobilon®-P, Thermo Fisher Scientific), under an electric field.

4.3. Western Blot (WB)

Western blot is a technique to identify specific proteins, fractionated by SDS-PAGE, by immunodetection with specific antibodies. In the current project, following SDS-PAGE analysis, the proteins in the gel are transferred to a PVDF membrane (Immobilon®-P, Thermo Fisher Scientific, Waltham, Massachusetts, USA) using the wet transfer method. The method involves preparing a “sandwich” with the gel in contact with the membrane (Fig. MM. 3) and equilibrating it in the transfer tank with transfer buffer (25 mM Tris, 192 mM Glycine, 0.1 % (w/v) SDS, 20 % (v/v) ethanol). The membrane is oriented next to the cathode, making the SDS-negatively charged proteins migrate from the gel onto the membrane in the electric field. The transfer was done in the cold room, under a constant voltage of 100 V for 90 min. After transfer, the membrane is incubated in a blocking solution (5 % skimmed milk in TBS-T (20 mM Tris-HCl, 150 mM NaCl, 0.1 % (v/v) Tween®20, pH 7.4)) for 1 hour. Several washes of the membrane with TBS-T are followed by overnight incubation in primary antibodies diluted in 1 % skimmed milk-TBS-T (1:1000 up to 1:10,000 dilution, according to the antibody). After several washes with TBS-T, the membrane is incubated in secondary antibodies solution for 1-2 hours (1:10,000 anti-rabbit or anti-mouse IgG, coupled to peroxidase). The revelation is carried out using ECL ("Enhanced Chemical Luminescence") Western Blotting detection systems, such as the Amersham kit or the Lumi-Light Western blotting system (Roche), depending on the desired detection threshold/background noise. Chemiluminescence was recorded using a FUSION FX imaging system (Vilber, Paris, France).

5. Plant Methods

5.1. Plant growth

Plants were grown either on soil or *in vitro,* in plates of 0.5x MS media (M0255, Duchefa Farma®, Haarlem, Netherlands) supplemented with 1 % sucrose and 0.8 % agar. Plants grown on soil were under a long-day photoperiod of 16 hours light (21 °C) and 8 hours dark (18 °C), or a 12-hour light-dark cycle. Plants on soil were exposed to light from LED lamps (Philips LED tube SA1 Master). Plants cultivated *in vitro* were exposed to light from fluorescent lamps (Osram BioluxT5).

5.2. *A. thaliana* transformation by floral dip

Plants were taken for transformation when they started bolting. For optimisation of transformation efficiency, all pre-formed siliques were cut. The transformation was mediated by *Agrobacterium tumefaciens* carrying the mitoTALEN constructions on the pK7WG2DIPS_RPS5Apro binary vector. 2 mL of preculture was used to start 250 mL cultures in LB medium, supplemented with antibiotics and grown overnight at 28 °C. Upon culture up

to an OD_{600} of 1.5-1.8, cells were sedimented and resuspended in an equal volume of infiltration medium (¼ MS, 5 % (w/v) Sucrose, 0.03 % (v/v) Silwet). Afterwards, flower organs were dipped in the bacterial suspension for 10 seconds, and plants were put under a dark dome to keep humidity and out of light for 2 days.

5.3. Crossing of *A. thaliana* plants

All mature siliques and open flowers of the amps used for crossing are removed. The emasculation of the pollen recipient flowers was performed one day before the crossing, under a stereo microscope and using very sharp tweezers (Dumont Biology n°5, super fine tip). Before pollination, the recipient pistil should exhibit developed stigmata with a rough, sticky surface. Using forceps, the anthers of pollen donors are brought into contact with the stigmata for pollination.

5.4. Seed selection and embryo extraction under confocal microscope

Plants transformed with the mitoTALEN constructs contained the marker *GFP* gene, expressed under the OLE1 oleosin promoter on the seed coat. T1 selection was done by screening for green fluorescence on seeds, using a fluorescent stereo microscope (Zeiss Axiozoom). Similarly, F1 transformants upon crossing were confirmed to have the mitoTALEN construct, and embryos were squashed out from the developing seed (3-14 days after germination) for DNA extraction and qPCR analysis (see Chapter II). We observed that the fluorescence was coming not from the seed coat, but mostly from the embryo endosperm.

5.5. Transmission electron microscopy imaging

Transmission electron microscopy (TEM) is a technique widely used in biology to observe the internal structure of tissues and cells. During the imaging, a beam of high-energy electrons is targeted to the samples and the transmitted through sample electrons create the electron density images. During my project, with the help of Mathieu Erhardt from the Bioimaging platform of IBMP, I obtained TEM images of mitochondria from cauline and rosette leaves of 6-week chimeric TAL-2 transformants.

5.6. In vitro methods

5.6.1. Seeds sterilisation

To grow plants *in vitro* on sucrose-containing media without overwhelming growth of bacteria and fungi, the seeds must be surface sterilised, to remove fungi and bacteria. This was done by incubation in 70 % ethanol, 0.05 % TritonX-100 for 15 minutes. Afterwards, TritonX-100 is washed out by 2 washes with 70 % ethanol, followed by a final wash with 90 % ethanol. The seeds were air-dried for 2 hours in a fume hood to remove all traces of ethanol before sowing.

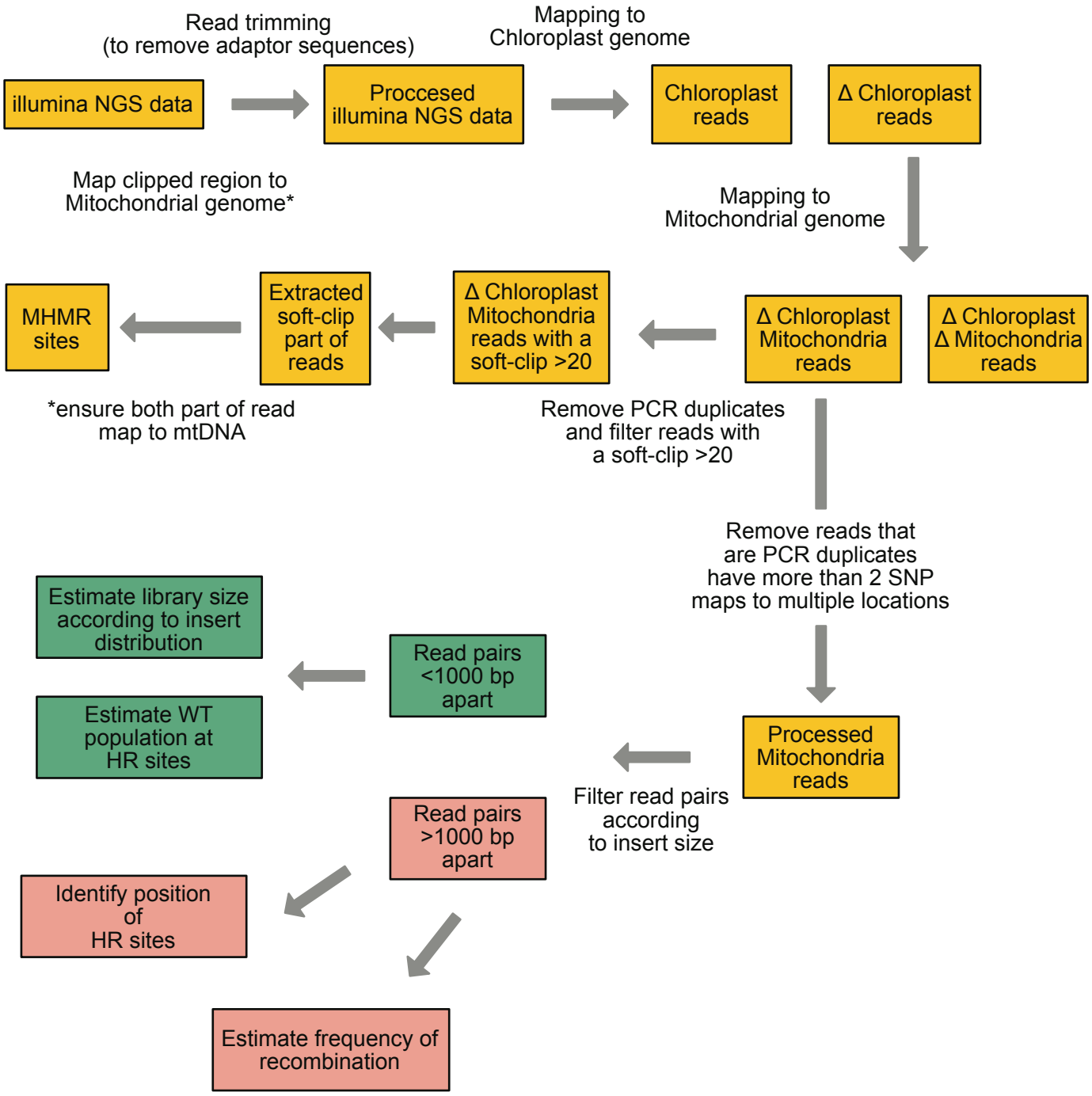


Figure MM.4. Schematics of the pipeline used for NGS data analysis in this project. In yellow are the basic steps and MHMR analysis; in green are the analyses of the WT fragment population; and in red are the analyses of read pairs that confirm structural variation.

5.6.2. Seedlings growth on solid medium

Sterilised seeds are sown on Petri dish containing 0.5x MS media (M0255, Duchefa Farma®, Haarlem, Netherlands) supplemented with 1 % sucrose and solidified with 0.8 % agar. MS medium was supplemented with hydroxyurea 0.5 mM - 2 mM in the experiment for analysis of DNA replication.

6. Bioinformatics analysis

6.1. Initial sequencing mapping with MacVector

Illumina sequencing reads were first analysed with MacVector inbuild bowtie2 assembler. Reads mapping to chloroplast DNA (cpDNA) sequences were first removed from analysis before mapping the remaining reads to the mtDNA. Total coverage maps were retrieved from coverage representations produced by the MacVector application.

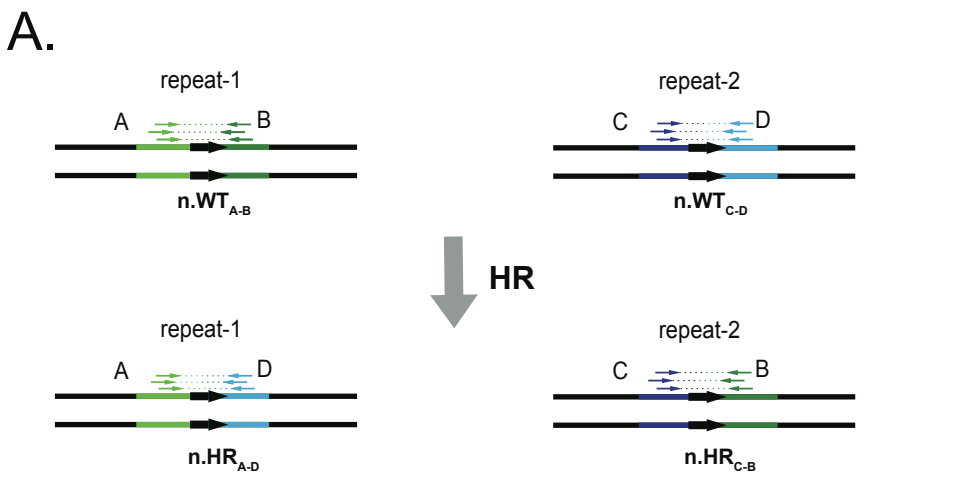
6.2. In-depth analysis of NGS Illumina data

Command line tools were used for more detailed analysis. As with MacVector, chloroplast reads were first removed before mapping to the mtDNA. Afterwards, the remaining reads were used according to developed NGS analysis pipelines of HR or MHMR (Fig MM. 4).

6.2.1 Identifying recombination event by read pair analysis.

The produced alignments were filtered to remove PCR duplicates, reads that map to multiple locations, had a mapping quality below 1, a soft-clip larger than half the size of the read and more than 2 SNP. Afterwards, two sets were produced: read pairs with insert size smaller than 1000 bp and read pairs with insert size larger than 1000 bp. Reads pairs that are 1000 bp apart resulted from structural variation, while reads below 1000 bp represent WT mitotype.

Further analysis was done with Rstudio software. The set off read pairs with insert size below 1000 bp were used to calculate the Illumina library fragment size distribution, as well as the average library size. It was assumed that most pairs with insert size below 1000 bp represent true fragment size, as most mtDNA modifications alter the whole genome and are not local. As Illumina fragments obtained by tagmentation are not larger than 1000 bp, pairs with an insert above 1000 bp represent mtDNA modification events, mainly by recombination. Windows of recombination events are set at both sides of each repeat (Appendix 4). Read pairs with an insert size above 1000 bp, identified in two distinct unique windows for the same repeat but for different copies, represent a potential read pair indicative of a recombination event (Fig. MM. 5A). Similarly, if a read pair is identified in two different windows for the same repeat and the same copy, it serves as a representative pair for the WT mitotype and suggests a possible recombination event involving the repeat. The efficiency of recombination involving



B.

Calculations

$c.WT = MIN\{n.WT_{A-B}; n.WT_{C-D}\}$

$HR = n.HR_{A-D} + n.HR_{C-B}$

$$\text{ratio of } HR_{repeat} = \frac{HR}{c.WT + HR}$$

Dictionary

n.WT - number of pairs with a read on both sides of the repeat

n.HR - number of pairs with a read on both sides of the repeat after HR

c.WT - "capacity" to undergo recombination

HR - read pairs confirming recombination

Figure MM.5. Schematic representation of Homologous Recombination event detection using read pair data. A) Color-coded sequences and reads represent windows that are analysed. When both reads of the pair are found around the same copy of the repeat-1 or repeat-2, they are in WT configuration. When one of the reads of the pair is found in the window of repeat-1 and the other in the window of repeat-2 it represents a Homologous recombination event. B) Calculations that were taken to determine the frequency of recombination for specific repeats.

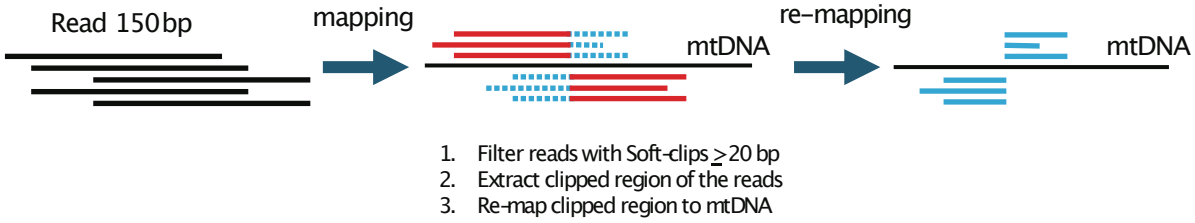


Figure MM.6. Schematic representation of microhomology-mediated recombination event detection using soft-clips. Reads that map to the mitochondrial genome and contain soft clips are filtered (1), and clipped parts (2) of reads are re-mapped to the mitochondrial genome (3). When both parts of the reads map to different loci of the mtDNA usually corresponds to a recombination event involving a microhomology region.

a pair of repeated sequences was reported as the ratio of how many reads were found in windows at two different loci compared to the total reads at the site of recombination (Fig. MM. 5B). Total reads at recombination sites are the sum of pairs confirming recombination and reads pairs found in a single locus at both sides of a same repeat. For comparison purposes, all read counts were normalised to "counts per million", initially mapped before filtering. All analyses are represented in Circos plots with the circlize R package. The opacity of several links was manually corrected according to the change in coverage. Scripts used for the analysis are available on GitHub: https://github.com/rokaskub.

6.2.2. Analysis of Microhomology-Mediated Recombination (MHMR)

To analyse the repair of cut sites by MHMR, reads containing soft-clips were analysed more closely. Regions with small repeats, smaller than half the read size, were excluded from mapping to eliminate soft-clips related to these repeats. Only reads containing soft-clips larger than 20 bp were selected from read alignments to the mtDNA. The soft-clip part of the read was then remapped to the mtDNA using bowtie2, to confirm that both parts of the read originated from mitochondria (Fig. MM. 6). Re-mapped reads were represented as a coverage map normalised to counts per million, initially mapped before filtering. The resulting peaks in the coverage map represent MHMR sites. Soft-clip reads at these sites were further confirmed by aligning them to the mtDNA using the Pustell DNA matrix dot plot function of MacVector, to identify the multiple mapping sites and the microhomology sequence involved. The scripts used for the analysis are available on GitHub: https://github.com/rokaskub.

Bibliography

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

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 U S A*, **100**, 5968–5973.

Aboussekhra, A., Biggerstaff, M., Shivji, M.K.K., et al. (1995) Mammalian DNA nucleotide excision repair reconstituted with purified protein components. *Cell*, **80**, 859–868.

Adams, J.S. and Hewison, M. (2010) Update in Vitamin D. *J Clin Endocrinol Metab*, **95**, 471–478.

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

Adams, K.L., Qiu, Y.L., Stoutemyer, M. and Palmer, J.D. (2002) Punctuated evolution of mitochondrial gene content: high and variable rates of mitochondrial gene loss and transfer to the nucleus during angiosperm evolution. *Proc Natl Acad Sci U S A*, **99**, 9905–12.

Akishev, Z., Taipakova, S., Joldybayeva, B., Zutterling, C., Smekenov, I., Ishchenko, A.A., Zharkov, D.O., Bissenbaev, A.K. and Saparbaev, M. (2016) The major Arabidopsis thaliana apurinic/apyrimidinic endonuclease, ARP is involved in the plant nucleotide incision repair pathway. *DNA Repair (Amst)*, **48**, 30–42.

Alán, L., Špaček, T. and Ježek, P. (2016) Delaunay algorithm and principal component analysis for 3D visualization of mitochondrial DNA nucleoids by Biplane FPALM/dSTORM. *European Biophysics Journal*, **45**, 443–461.

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

Almeida, A., Bolaños, J.P. and Medina, J.M. (1999) Nitric oxide mediates brain mitochondrial maturation immediately after birth. *FEBS Lett*, **452**, 290–294.

Alou, A.H., Azaiez, A., Jean, M. and Belzile, F.J. (2004) Involvement of the Arabidopsis thaliana AtPMS1 gene in somatic repeat instability. *Plant Mol Biol*, **56**, 339–349.

Altmann, R. (1890) *The elementary organisms and their relationships to the cells*, Leipzig: Veit.

Alverson, A.J., Zhuo, S., Rice, D.W., Sloan, D.B. and Palmer, J.D. (2011) The Mitochondrial Genome of the Legume Vigna radiata and the Analysis of Recombination across Short Mitochondrial Repeats. *PLoS One*, **6**, e16404.

Ambro, L., Pevala, V., Bauer, J. and Kutejová, E. (2012) The influence of ATP-dependent proteases on a variety of nucleoid-associated processes. *J Struct Biol*, **179**, 181–192.

Amundsen, S.K. and Smith, G.R. (2019) The RecB helicase-nuclease tether mediates Chi hotspot control of RecBCD enzyme. *Nucleic Acids Res*, **47**, 197–209.

Andersen, S.L. and Sekelsky, J. (2010) Meiotic *versus* mitotic recombination: Two different routes for double-strand break repair. *BioEssays*, **32**, 1058–1066.

Anderson, S., Bankier, A.T., Barrell, B.G., et al. (1981) Sequence and organization of the human mitochondrial genome. *Nature*, **290**, 457–465.

Andersson, S.G.E., Zomorodipour, A., Andersson, J.O., et al. (1998) The genome sequence of Rickettsia prowazekii and the origin of mitochondria. *Nature*, **396**, 133–140.

Arimura, S. (2021) Effects of mitoTALENs-Directed Double-Strand Breaks on Plant Mitochondrial Genomes. *Genes (Basel)*, **12**, 153.

Arimura, S., Yamamoto, J., Aida, G.P., Nakazono, M. and Tsutsumi, N. (2004) Frequent fusion and fission of plant mitochondria with unequal nucleoid distribution. *Proceedings of the National Academy of Sciences*, **101**, 7805–7808.

Arimura, S.I., Ayabe, H., Sugaya, H., et al. (2020) Targeted gene disruption of ATP synthases 6-1 and 6-2 in the mitochondrial genome of Arabidopsis thaliana by mitoTALENs. *Plant J*, **104**, 1459–1471.

Arrieta-Montiel, M., Lyznik, A., Woloszynska, M., Janska, H., Tohme, J. and Mackenzie, S. (2001) Tracing evolutionary and developmental implications of mitochondrial stoichiometric shifting in the common bean. *Genetics*, **158**, 851–864.

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.

Arya, R. and Bassing, C.H. (2017) V(D)J Recombination Exploits DNA Damage Responses to Promote Immunity. *Trends in Genetics*, **33**, 479–489.

Ashworth, J., Havranek, J.J., Duarte, C.M., Sussman, D., Monnat, R.J., Stoddard, B.L. and Baker, D. (2006) Computational redesign of endonuclease DNA binding and cleavage specificity. *Nature*, **441**, 656–659.

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

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

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

Bacman, S.R., Williams, S.L., Pinto, M., Peralta, S. and Moraes, C.T. (2013) Specific elimination of mutant mitochondrial genomes in patient-derived cells by mitoTALENs. *Nat Med*, **19**, 1111–1113.

Balansky, R., Izzotti, A., Scatolini, L., D'Agostini, F. and Flora, S. De (1996) Induction by carcinogens and chemoprevention by N-acetylcysteine of adducts to mitochondrial DNA in rat organs. *Cancer Res*, **56**, 1642–7.

Balázs, Z., Tombácz, D., Csabai, Z., Moldován, N., Snyder, M. and Boldogkői, Z. (2019) Template-switching artifacts resemble alternative polyadenylation. *BMC Genomics*, **20**, 824.

Ban, C. (1998) Structural basis for MutH activation in E.coli mismatch repair and relationship of MutH to restriction endonucleases. *EMBO J*, **17**, 1526–1534.

Barbara, T. (2003) Imidazole Ring-Opened DNA Purines and Their Biological Significance. *BMB Rep*, **36**, 12–19.

Bardaweel, S.K., Gul, M., Alzweiri, M., Ishaqat, A., ALSalamat, H.A. and Bashatwah, R.M. (2018) Reactive Oxygen Species: the Dual Role in Physiological and Pathological Conditions of the Human Body. *Eurasian J Med*, **50**, 193–201.

Baruch-Torres, N. and Brieba, L.G. (2017) Plant organellar DNA polymerases are replicative and translesion DNA synthesis polymerases. *Nucleic Acids Res*, **45**, 10751–10763.

Batista, L.F.Z., Kaina, B., Meneghini, R. and Menck, C.F.M. (2009) How DNA lesions are turned into powerful killing structures: Insights from UV-induced apoptosis. *Mutation Research/Reviews in Mutation Research*, **681**, 197–208.

Bauer, M.F., Hofmann, S., Neupert, W. and Brunner, M. (2000) Protein translocation into mitochondria: the role of TIM complexes. *Trends Cell Biol*, **10**, 25–31.

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

Bauwe, H., Hagemann, M. and Fernie, A.R. (2010) Photorespiration: players, partners and origin. *Trends Plant Sci*, **15**, 330–336.

Beagan, K., Armstrong, R.L., Witsell, A., Roy, U., Renedo, N., Baker, A.E., Schärer, O.D. and McVey, M. (2017) Drosophila DNA polymerase theta utilizes both helicase-like and polymerase domains during microhomology-mediated end joining and interstrand crosslink repair. *PLoS Genet*, **13**, e1006813.

Bechtold, N. and Pelletier, G. (1998) In Planta AgrobacteriumMediated Transformation of Adult Arabidopsis thaliana Plants by Vacuum Infiltration. In *Arabidopsis Protocols*. Totowa, NJ: Humana Press, pp. 259–266.

Becker, L., Bannwarth, M., Meisinger, C., et al. (2005) Preprotein Translocase of the Outer Mitochondrial Membrane: Reconstituted Tom40 Forms a Characteristic TOM Pore. *J Mol Biol*, **353**, 1011–1020.

Becker, T. and Wagner, R. (2018) Mitochondrial Outer Membrane Channels: Emerging Diversity in Transport Processes. *BioEssays*, **40**.

Benda, C. (1898) On spermatogenesis of vertebrates and higher invertebrates. *Archiv für Anatomie und Physiologie*, **73**, 393–398.

Bendich, A.J. (1993) Reaching for the ring: the study of mitochondrial genome structure. *Curr. Genet.*, **24**, 279–290.

Bendich, A.J. (1996) Structural analysis of mitochondrial DNA molecules from fungi and plants using moving pictures and pulsed-field gel electrophoresis. *J. Mol. Biol.*, **255**, 564–588.

Bennardo, N., Cheng, A., Huang, N. and Stark, J.M. (2008) Alternative-NHEJ Is a Mechanistically Distinct Pathway of Mammalian Chromosome Break Repair. *PLoS Genet*, **4**, e1000110.

Bereiter-Hahn, J. and Vöth, M. (1994) Dynamics of mitochondria in living cells: Shape changes, dislocations, fusion, and fission of mitochondria. *Microsc Res Tech*, **27**, 198–219.

Bernstein, H. and Bernstein, C. (1973) Circular and branched circular concatenates as possible intermediates in bacteriophage T4 DNA replication. *J Mol Biol*, **77**, 355–361.

Bertocci, B., Smet, A. De, Weill, J.-C. and Reynaud, C.-A. (2006) Nonoverlapping Functions of DNA Polymerases Mu, Lambda, and Terminal Deoxynucleotidyltransferase during Immunoglobulin V(D)J Recombination In Vivo. *Immunity*, **25**, 31–41.

Betti, M., Petrucco, S., Bolchi, A., Dieci, G. and Ottonello, S. (2001) A Plant 3′-Phosphoesterase Involved in the Repair of DNA Strand Breaks Generated by Oxidative Damage. *Journal of Biological Chemistry*, **276**, 18038–18045.

Bhargava, R., Onyango, D.O. and Stark, J.M. (2016) Regulation of Single-Strand Annealing and its Role in Genome Maintenance. *Trends in Genetics*, **32**, 566–575.

Bianco, P.R. (2017) The tale of SSB. *Prog Biophys Mol Biol*, **127**, 111–118.

Birky Jr., C.W. (2001) The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms, and models. *Annu. Rev. Genet.*, **35**, 125–148.

Blanco-Grau, A., Bonaventura-Ibars, I., Coll-Cantí, J., Melià, M.J., Martinez, R., Martínez-Gallo, M., Andreu, A.L., Pinós, T. and García-Arumí, E. (2013) Identification and biochemical characterization of the novel mutation m.8839G>C in the mitochondrial ATP6 gene associated with NARP syndrome. *Genes Brain Behav*, **12**, 812–820.

Blum, T.B., Hahn, A., Meier, T., Davies, K.M. and Kühlbrandt, W. (2019) Dimers of mitochondrial ATP synthase induce membrane curvature and self-assemble into rows. *Proceedings of the National Academy of Sciences*, **116**, 4250–4255.

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A. and Bonas, U. (2009) Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors. *Science (1979)*, **326**, 1509–1512.

Boesch, P., Ibrahim, N., Paulus, F., Cosset, A., Tarasenko, V. and Dietrich, A. (2009) Plant mitochondria possess a short-patch base excision DNA repair pathway. *Nucleic Acids Res*, **37**, 5690–5700.

Bogenhagen, D.F., Rousseau, D. and Burke, S. (2008) The Layered Structure of Human Mitochondrial DNA Nucleoids. *Journal of Biological Chemistry*, **283**, 3665–3675.

Boltz, K.A., Jasti, M., Townley, J.M. and Shippen, D.E. (2014) Analysis of Poly(ADP-Ribose) Polymerases in Arabidopsis Telomere Biology. *PLoS One*, **9**, e88872.

Bonilla, B., Hengel, S.R., Grundy, M.K. and Bernstein, K.A. (2020) RAD51 Gene Family Structure and Function. *Annu Rev Genet*, **54**, 25–46.

Bonnefoy, N., Remacle, C. and Fox, T.D. (2007) Genetic transformation of Saccharomyces cerevisiae and Chlamydomonas reinhardtii mitochondria. *Methods Cell Biol.*, **80**, 525–548.

Boore, J.L. (1999) Animal mitochondrial genomes. *Nucleic Acids Res*, **27**, 1767–1780.

Bordin, D.L., Lirussi, L. and Nilsen, H. (2021) Cellular response to endogenous DNA damage: DNA base modifications in gene expression regulation. *DNA Repair (Amst)*, **99**, 103051.

Boveris, A., Costa, L.E., Poderoso, J.J., Carreras, M.C. and Cadenas, E. (2000) Regulation of Mitochondrial Respiration by Oxygen and Nitric Oxide. *Ann N Y Acad Sci*, **899**, 121–135.

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.*, **72**, 248–254.

Brettel, K. and Byrdin, M. (2010) Reaction mechanisms of DNA photolyase. *Curr Opin Struct Biol*, **20**, 693–701.

Bridge, G., Rashid, S. and Martin, S. (2014) DNA Mismatch Repair and Oxidative DNA Damage: Implications for Cancer Biology and Treatment. *Cancers (Basel)*, **6**, 1597–1614.

Brieba (2019) Structure–Function Analysis Reveals the Singularity of Plant Mitochondrial DNA Replication Components: A Mosaic and Redundant System. *Plants*, **8**, 533.

Britt, A.B., Chen, J.J., Wykoff, D. and Mitchell, D. (1993) A UV-sensitive mutant of Arabidopsis defective in the repair of pyrimidine-pyrimidinone(6-4) dimers. *Science (1979)*, **261**, 1571–1574.

Broz, A.K., Keene, A., Fernandes Gyorfy, M., Hodous, M., Johnston, I.G. and Sloan, D.B. (2022) Sorting of mitochondrial and plastid heteroplasmy in *Arabidopsis* is extremely rapid and depends on MSH1 activity. *Proceedings of the National Academy of Sciences*, **119**.

Bunting, S.F., Callén, E., Wong, N., et al. (2010) 53BP1 Inhibits Homologous Recombination in Brca1-Deficient Cells by Blocking Resection of DNA Breaks. *Cell*, **141**, 243–254.

Burdett, V., Baitinger, C., Viswanathan, M., Lovett, S.T. and Modrich, P. (2001) In vivo requirement for RecJ, ExoVII, ExoI, and ExoX in methyl-directed mismatch repair. *Proceedings of the National Academy of Sciences*, **98**, 6765–6770.

Burney, S., Caulfield, J.L., Niles, J.C., Wishnok, J.S. and Tannenbaum, S.R. (1999) The chemistry of DNA damage from nitric oxide and peroxynitrite. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, **424**, 37–49.

Bykova, N. V., Keerberg, O., Pärnik, T., Bauwe, H. and Gardeström, P. (2005) Interaction between photorespiration and respiration in transgenic potato plants with antisense reduction in glycine decarboxylase. *Planta*, **222**, 130–140.

Cai, C.Q., Doyon, Y., Ainley, W.M., et al. (2009) Targeted transgene integration in plant cells using designed zinc finger nucleases. *Plant Mol Biol*, **69**, 699–709.

Cai, Q., Guo, L., Shen, Z.-R., Wang, D.-Y., Zhang, Q. and Sodmergen (2015) Elevation of Pollen Mitochondrial DNA Copy Number by WHIRLY2: Altered Respiration and Pollen Tube Growth in Arabidopsis. *Plant Physiol*, **169**, 660–673.

Caldecott, K.W. (2008) Single-strand break repair and genetic disease. *Nat Rev Genet*, **9**, 619–631.

Caldecott, K.W. (2003) XRCC1 and DNA strand break repair. *DNA Repair (Amst)*, **2**, 955–969.

Caldecott, K.W., Aoufouchi, S., Johnson, P. and Shall, S. (1996) XRCC1 Polypeptide Interacts with DNA Polymerase and Possibly Poly (ADP-Ribose) Polymerase, and DNA Ligase III Is a Novel Molecular "Nick-Sensor" In Vitro. *Nucleic Acids Res*, **24**, 4387–4394.

Campbell, C.D., Chong, J.X., Malig, M., et al. (2012) Estimating the human mutation rate using autozygosity in a founder population. *Nat Genet*, **44**, 1277–1281.

Campbell, C.T., Kolesar, J.E. and Kaufman, B.A. (2012) Mitochondrial transcription factor A regulates mitochondrial transcription initiation, DNA packaging, and genome copy number. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, **1819**, 921–929.

Campo, M.L., Peixoto, P.M. and Martínez-Caballero, S. (2017) Revisiting trends on mitochondrial mega-channels for the import of proteins and nucleic acids. *J Bioenerg Biomembr*, **49**, 75–99.

Cannan, W.J. and Pederson, D.S. (2016) Mechanisms and Consequences of Double-Strand DNA Break Formation in Chromatin. *J Cell Physiol*, **231**, 3–14.
Cannavo, E. and Cejka, P. (2014) Sae2 promotes dsDNA endonuclease activity within Mre11–Rad50–Xrs2 to resect DNA breaks. *Nature*, **514**, 122–125.
Cannavo, E., Marra, G., Sabates-Bellver, J., Menigatti, M., Lipkin, S.M., Fischer, F., Cejka, P. and Jiricny, J. (2005) Expression of the MutL Homologue hMLH3 in Human Cells and its Role in DNA Mismatch Repair. *Cancer Res*, **65**, 10759–10766.
Canugovi, C., Maynard, S., Bayne, A.-C. V., Sykora, P., Tian, J., Souza-Pinto, N.C. de, Croteau, D.L. and Bohr, V.A. (2010) The mitochondrial transcription factor A functions in mitochondrial base excision repair. *DNA Repair (Amst)*, **9**, 1080–1089.
Cappadocia, L., Marechal, 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.
Cappadocia, L., Parent, J.S., Zampini, E., Lepage, E., Sygusch, J. and Brisson, N. (2012) A conserved lysine residue of plant Whirly proteins is necessary for higher order protein assembly and protection against DNA damage. *Nucleic Acids Res.*, **40(1):258-69**, 258–269.
Carlsson, J., Leino, M., Sohlberg, J., Sundström, J.F. and Glimelius, K. (2008) Mitochondrial regulation of flower development. *Mitochondrion*, **8**, 74–86.
Carreira, A. and Kowalczykowski, S.C. (2011) Two classes of BRC repeats in BRCA2 promote RAD51 nucleoprotein filament function by distinct mechanisms. *Proceedings of the National Academy of Sciences*, **108**, 10448–10453.
Carrozzo, R., Rizza, T., Stringaro, A., Pierini, R., Mormone, E., Santorelli, F.M., Malorni, W. and Matarrese, P. (2004) Maternally-inherited Leigh syndrome-related mutations bolster mitochondrial-mediated apoptosis. *J Neurochem*, **90**, 490–501.
Ceccaldi, R., Liu, J.C., Amunugama, R., et al. (2015) Homologous-recombination-deficient tumours are dependent on Polθ-mediated repair. *Nature*, **518**, 258–262.
Cermak, T., Doyle, E.L., Christian, M., et al. (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res*, **39**, e82–e82.
Cevallos, M.A. and Degli Esposti, M. (2022) New Alphaproteobacteria Thrive in the Depths of the Ocean with Oxygen Gradient. *Microorganisms*, **10**, 455.
Chakraborty, A., Mitra, J., Bhattacharyya, J., et al. (2015) Transgenic expression of an unedited mitochondrial orfB gene product from wild abortive (WA) cytoplasm of rice (Oryza sativa L.) generates male sterility in fertile rice lines. *Planta*, **241**, 1463–1479.
Chakraborty, A., Tapryal, N., Islam, A., Mitra, S. and Hazra, T. (2021) Transcription coupled base excision repair in mammalian cells: So little is known and so much to uncover. *DNA Repair (Amst)*, **107**, 103204.
Chang, Y., Liu, B., Jiang, Y., Cao, D., Liu, Y. and Li, Y. (2023) Induce male sterility by CRISPR/Cas9-mediated mitochondrial genome editing in tobacco. *Funct Integr Genomics*, **23**, 205.
Charbonnel, C., Gallego, M.E. and White, C.I. (2010) Xrcc1-dependent and Ku-dependent DNA double-strand break repair kinetics in Arabidopsis plants. *The Plant Journal*, **64**, 280–290.
Chase, C.D. (2007) Cytoplasmic male sterility: a window to the world of plant mitochondrial–nuclear interactions. *Trends in Genetics*, **23**, 81–90.
Chen, J.J., Jiang, C.Z. and Britt, A.B. (1996) Little or No Repair of Cyclobutyl Pyrimidine Dimers Is Observed in the Organellar Genomes of the Young Arabidopsis Seedling. *Plant Physiol*, **111**, 19–25.
Chen, L. and Liu, Y.-G. (2014) Male Sterility and Fertility Restoration in Crops. *Annu Rev Plant Biol*, **65**, 579–606.
Chen, Z., Zhang, J. and Stamler, J.S. (2002) Identification of the enzymatic mechanism of nitroglycerin bioactivation. *Proceedings of the National Academy of Sciences*, **99**, 8306–8311.

Chevigny, N, Schatz-Daas, D., Lotfi, F. and Gualberto, J.M. (2020) DNA Repair and the Stability of the Plant Mitochondrial Genome. *Int J Mol Sci*, **21**.
Chevigny, N., Weber-Lotfi, F., Blevenec, A. Le, et al. (2022) RADA-dependent branch migration has a predominant role in plant mitochondria and its defect leads to mtDNA instability and cell cycle arrest. *PLoS Genet*, **18**, e1010202.
Cho, H.S., Lee, S.S., Kim, K.D., Hwang, I., Lim, J.-S., Park, Y.-I. and Pai, H.-S. (2004) DNA Gyrase Is Involved in Chloroplast Nucleoid Partitioning. *Plant Cell*, **16**, 2665–2682.
Cho, S.-I., Lee, S., Mok, Y.G., Lim, K., Lee, J., Lee, J.M., Chung, E. and Kim, J.-S. (2022) Targeted A-to-G base editing in human mitochondrial DNA with programmable deaminases. *Cell*, **185**, 1764-1776.e12.
Christensen, A.C. (2014) Genes and junk in plant mitochondria-repair mechanisms and selection. *Genome Biol. Evol.*, **6**, 1448–1453.
Christensen, A.C. (2018) Mitochondrial DNA Repair and Genome Evolution. In *Annual Plant Reviews online*. pp. 11–32.
Christensen, A.C. (2013) Plant mitochondrial genome evolution can be explained by DNA repair mechanisms. *Genome Biol. Evol.*, **5**, 1079–1086.
Christian, M., Cermak, T., Doyle, E.L., Schmidt, C., Zhang, F., Hummel, A., Bogdanove, A.J. and Voytas, D.F. (2010) Targeting DNA Double-Strand Breaks with TAL Effector Nucleases. *Genetics*, **186**, 757–761.
Cline, S.D. (2012) Mitochondrial DNA damage and its consequences for mitochondrial gene expression. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, **1819**, 979–991.
Coin, F., Oksenych, V. and Egly, J.-M. (2007) Distinct Roles for the XPB/p52 and XPD/p44 Subcomplexes of TFIIH in Damaged DNA Opening during Nucleotide Excision Repair. *Mol Cell*, **26**, 245–256.
Cookson, S.J., Radziejwoski, A. and Granier, C. (2006) Cell and leaf size plasticity in Arabidopsis: what is the role of endoreduplication? *Plant Cell Environ*, **29**, 1273–1283.
Cooper, D.L. and Lovett, S.T. (2016) Recombinational branch migration by the RadA/Sms paralog of RecA in Escherichia coli. *Elife*, **5**.
Córdoba-Cañero, D., Morales-Ruiz, T., Roldán-Arjona, T. and Ariza, R.R. (2009) Single-nucleotide and long-patch base excision repair of DNA damage in plants. *The Plant Journal*, **60**, 716–728.
Córdoba-Cañero, D., Roldán-Arjona, T. and Ariza, R.R. (2011) Arabidopsis ARP endonuclease functions in a branched base excision DNA repair pathway completed by LIG1. *The Plant Journal*, **68**, 693–702.
Córdoba-Cañero, D., Roldán-Arjona, T. and Ariza, R.R. (2014) Arabidopsis ZDP DNA 3'-phosphatase and ARP endonuclease function in 8-oxoG repair initiated by FPG and OGG1 DNA glycosylases. *The Plant Journal*, **79**, 824–834.
Cornu, T.I., Thibodeau-Beganny, S., Guhl, E., Alwin, S., Eichtinger, M., Joung, J. and Cathomen, T. (2008) DNA-binding Specificity Is a Major Determinant of the Activity and Toxicity of Zinc-finger Nucleases. *Molecular Therapy*, **16**, 352–358.
Corrette-Bennett, S.E. and Lovett, S.T. (1995) Enhancement of RecA Strand-transfer Activity by the RecJ Exonuclease of Escherichia coli. *Journal of Biological Chemistry*, **270**, 6881–6885.
Cox, M.M. (2007) Motoring along with the bacterial RecA protein. *Nat. Rev. Mol. Cell Biol.*, **8**, 127–138.
Cree, L.M., Samuels, D.C., Sousa Lopes, S.C. de, Rajasimha, H.K., Wonnapinij, P., Mann, J.R., Dahl, H.H. and Chinnery, P.F. (2008) A reduction of mitochondrial DNA molecules during embryogenesis explains the rapid segregation of genotypes. *Nat. Genet.*, **40**, 249–254.
Crumpton-Taylor, M., Grandison, S., Png, K.M.Y., Bushby, A.J. and Smith, A.M. (2012) Control of Starch Granule Numbers in Arabidopsis Chloroplasts . *Plant Physiol*, **158**, 905–916.
Cuijk, L. van, Belle, G.J. van, Turkyilmaz, Y., et al. (2015) SUMO and ubiquitin-dependent XPC exchange drives nucleotide excision repair. *Nat Commun*, **6**, 7499.

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

Dahal, S., Dubey, S. and Raghavan, S.C. (2018) Homologous recombination-mediated repair of DNA double-strand breaks operates in mammalian mitochondria. *Cellular and Molecular Life Sciences*, **75**, 1641–1655.

Dalhus, B., Laerdahl, J.K., Backe, P.H. and Bjørås, M. (2009) DNA base repair – recognition and initiation of catalysis. *FEMS Microbiol Rev*, **33**, 1044–1078.

Dany, A.L. and Tissier, A. (2001) A functional OGG1 homologue from Arabidopsis thaliana. *Molecular Genetics and Genomics*, **265**, 293–301.

Darracq, A., Varre, J.S., Marechal-Drouard, L., et al. (2011) Structural and content diversity of mitochondrial genome in beet: a comparative genomic analysis. *Genome Biol. Evol.*, **3**, 723–736.

Das, S., Sen, S., Chakraborty, A., Chakraborti, P., Maiti, M.K., Basu, A., Basu, D. and Sen, S.K. (2010) An unedited 1.1 kb mitochondrial orfB gene transcript in the Wild Abortive Cytoplasmic Male Sterility (WA-CMS) system of Oryza sativa L. subsp. indica. *BMC Plant Biol*, **10**, 39.

D'Autréaux, B. and Toledano, M.B. (2007) ROS as signalling molecules: mechanisms that generate specificity in ROS homeostasis. *Nat Rev Mol Cell Biol*, **8**, 813–824.

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

Daya-Grosjean, L. (2008) Xeroderma Pigmentosum and Skin Cancer. In *Molecular Mechanisms of Xeroderma Pigmentosum*. New York, NY: Springer New York, pp. 19–27.

Degli Esposti, M. (2023) The bacterial origin of mitochondria: Incorrect phylogenies and the importance of metabolic traits. In pp. 1–35.

Della-Maria, J., Zhou, Y., Tsai, M.-S., Kuhnlein, J., Carney, J.P., Paull, T.T. and Tomkinson, A.E. (2011) Human Mre11/Human Rad50/Nbs1 and DNA Ligase IIIα/XRCC1 Protein Complexes Act Together in an Alternative Nonhomologous End Joining Pathway. *Journal of Biological Chemistry*, **286**, 33845–33853.

Dewey, R.E., Timothy, D.H. and Levings, C.S. (1991) Chimeric mitochondrial genes expressed in the C male-sterile cytoplasm of maize. *Curr Genet*, **20**, 475–482.

Dieterich, J.-H., Braun, H.-P. and Schmitz, U.K. (2003) Alloplasmic male sterility in Brassica napus (CMS 'Tournefortii-Stiewe') is associated with a special gene arrangement around a novel atp9 gene. *Molecular Genetics and Genomics*, **269**, 723–731.

DiGiovanna, J.J. and Kraemer, K.H. (2012) Shining a Light on Xeroderma Pigmentosum. *Journal of Investigative Dermatology*, **132**, 785–796.

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

Dongen, J.T. van, Gupta, K.J., Ramírez-Aguilar, S.J., Araújo, W.L., Nunes-Nesi, A. and Fernie, A.R. (2011) Regulation of respiration in plants: A role for alternative metabolic pathways. *J Plant Physiol*, **168**, 1434–1443.

Doniwa, Y., Arimura, S. and Tsutsumi, N. (2007) Mitochondria use actin filaments as rails for fast translocation in Arabidopsis and tobacco cells. *Plant Biotechnology*, **24**, 441–447.

Dorfel, P., Weihe, A., Dolferus, R. and Borner, T. (1991) DNA sequence of a mitochondrial plasmid fromChenopodium album. *Plant Mol Biol*, **17**, 155–156.

Doyon, Y., McCammon, J.M., Miller, J.C., et al. (2008) Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. *Nat Biotechnol*, **26**, 702–708.

Drees, J.C., Lusetti, S.L. and Cox, M.M. (2004) Inhibition of RecA protein by the Escherichia coli RecX protein: modulation by the RecA C terminus and filament functional state. *J. Biol. Chem.*, **279**, 52991–52997.

Dunham, M.A., Neumann, A.A., Fasching, C.L. and Reddel, R.R. (2000) Telomere maintenance by recombination in human cells. *Nat Genet*, **26**, 447–450.

Dzantiev, L., Constantin, N., Genschel, J., Iyer, R.R., Burgers, P.M. and Modrich, P. (2004) A Defined Human System That Supports Bidirectional Mismatch-Provoked Excision. *Mol Cell*, **15**, 31–41.

Eckert-Boulet, N., Rothstein, R. and Lisby, M. (2011) Cell Biology of Homologous Recombination in Yeast. In pp. 523–536.

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. *Molecular Genetics and Genomics*, **273**, 115–122.

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

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[W]. *Plant Cell*, **15**, 1619–1631.

Engler, C., Gruetzner, R., Kandzia, R. and Marillonnet, S. (2009) Golden Gate Shuffling: A One-Pot DNA Shuffling Method Based on Type IIs Restriction Enzymes. *PLoS One*, **4**, e5553.

Enoiu, M., Jiricny, J. and Schärer, O.D. (2012) Repair of cisplatin-induced DNA interstrand crosslinks by a replication-independent pathway involving transcription-coupled repair and translesion synthesis. *Nucleic Acids Res*, **40**, 8953–8964.

Ernster, L. and Schatz, G. (1981) Mitochondria: a historical review. *J Cell Biol*, **91**, 227s–255s.

Esashi, F., Galkin, V.E., Yu, X., Egelman, E.H. and West, S.C. (2007) Stabilization of RAD51 nucleoprotein filaments by the C-terminal region of BRCA2. *Nat Struct Mol Biol*, **14**, 468–474.

Eubel, H., Jänsch, L. and Braun, H.-P. (2003) New Insights into the Respiratory Chain of Plant Mitochondria. Supercomplexes and a Unique Composition of Complex II. *Plant Physiol*, **133**, 274–286.

Falco, M. De and Felice, M. De (2021) Take a Break to Repair: A Dip in the World of Double-Strand Break Repair Mechanisms Pointing the Gaze on Archaea. *Int J Mol Sci*, **22**, 13296.

Falkenberg, M. and Gustafsson, C.M. (2020) Mammalian mitochondrial DNA replication and mechanisms of deletion formation. *Crit Rev Biochem Mol Biol*, **55**, 509–524.

Fan, J. (2004) XRCC1 co-localizes and physically interacts with PCNA. *Nucleic Acids Res*, **32**, 2193–2201.

Farge, G., Mehmedovic, M., Baclayon, M., van den Wildenberg, S.M.J.L., Roos, W.H., Gustafsson, C.M., Wuite, G.J.L. and Falkenberg, M. (2014) In Vitro-Reconstituted Nucleoids Can Block Mitochondrial DNA Replication and Transcription. *Cell Rep*, **8**, 66–74.

Farnell, D.A. (2011) Nucleotide Excision Repair in the Three Domains of Life. *Western Undergraduate Research Journal: Health and Natural Sciences*, **2**.

Farr, C.L., Matsushima, Y., Lagina 3rd, A.T., Luo, N. and Kaguni, L.S. (2004) Physiological and biochemical defects in functional interactions of mitochondrial DNA polymerase and DNA-binding mutants of single-stranded DNA-binding protein. *J. Biol. Chem.*, **279**, 17047–17053.

Fei, J., Kaczmarek, N., Luch, A., Glas, A., Carell, T. and Naegeli, H. (2011) Regulation of Nucleotide Excision Repair by UV-DDB: Prioritization of Damage Recognition to Internucleosomal DNA. *PLoS Biol*, **9**, e1001183.

Filograna, R., Mennuni, M., Alsina, D. and Larsson, N. (2021) Mitochondrial DNA copy number in human disease: the more the better? *FEBS Lett*, **595**, 976–1002.

Fischer, E.S., Scrima, A., Böhm, K., et al. (2011) The Molecular Basis of CRL4DDB2/CSA Ubiquitin Ligase Architecture, Targeting, and Activation. *Cell*, **147**, 1024–1039.

Fishel, R. (2015) Mismatch Repair. *Journal of Biological Chemistry*, **290**, 26395–26403.

Fisher, G.J., Kang, S., Varani, J., Bata-Csorgo, Z., Wan, Y., Datta, S. and Voorhees, J.J. (2002) Mechanisms of Photoaging and Chronological Skin Aging. *Arch Dermatol*, **138**.

Forget, A.L. and Kowalczykowski, S.C. (2012) Single-molecule imaging of DNA pairing by RecA reveals a three-dimensional homology search. *Nature*, **482**, 423–427.
Forner, J., Kleinschmidt, D., Meyer, E.H., Fischer, A., Morbitzer, R., Lahaye, T., Schöttler, M.A. and Bock, R. (2022) Targeted introduction of heritable point mutations into the plant mitochondrial genome. *Nat Plants*, **8**, 245–256.
Forner, J., Kleinschmidt, D., Meyer, E.H., Gremmels, J., Morbitzer, R., Lahaye, T., Schöttler, M.A. and Bock, R. (2023) Targeted knockout of a conserved plant mitochondrial gene by genome editing. *Nat Plants*, **9**, 1818–1831.
Fortini, P. and Dogliotti, E. (2007) Base damage and single-strand break repair: Mechanisms and functional significance of short- and long-patch repair subpathways. *DNA Repair (Amst)*, **6**, 398–409.
Fox, T.D., Sanford, J.C. and McMullin, T.W. (1988) Plasmids can stably transform yeast mitochondria lacking endogenous mtDNA. *Proc. Natl. Acad. Sci. U. S. A.*, **85**, 7288–7292.
Fricke, W.M. and Brill, S.J. (2003) Slx1—Slx4 is a second structure-specific endonuclease functionally redundant with Sgs1—Top3. *Genes Dev*, **17**, 1768–1778.
Friedberg, E.C., Walker, G.C. and Siede, W. (1995) DNA Repair and Mutagenesis . In *Am. Soc. for Microbiology*. Washington: DC.
Friedkin, Morris. and Lehninger, A.L. (1949) Oxidation-coupled incorporation of inorganic radiophosphate into phospholipide and nucleic acid in a cell-free system. *Journal of Biological Chemistry*, **177**, 775–788.
Fuchs, P., Rugen, N., Carrie, C., et al. (2020) Single organelle function and organization as estimated from Arabidopsis mitochondrial proteomics. *The Plant Journal*, **101**, 420–441.
Fukui, K. (2010) DNA Mismatch Repair in Eukaryotes and Bacteria. *J Nucleic Acids*, **2010**, 1–16.
Fukui, K., Harada, A., Wakamatsu, T., Minobe, A., Ohshita, K., Ashiuchi, M. and Yano, T. (2018) The GIY-YIG endonuclease domain of Arabidopsis MutS homolog 1 specifically binds to branched DNA structures. *FEBS Lett*, **592**, 4066–4077.
Furukawa, T., Angelis, K.J. and Britt, A.B. (2015) Arabidopsis DNA polymerase lambda mutant is mildly sensitive to DNA double strand breaks but defective in integration of a transgene. *Front Plant Sci*, **6**.
Galio, L., Bouquet, C. and Brooks, P. (1999) ATP hydrolysis-dependent formation of a dynamic ternary nucleoprotein complex with MutS and MutL. *Nucleic Acids Res*, **27**, 2325–2331.
Gallego, M.E. and White, C.I. (2001) RAD50 function is essential for telomere maintenance in Arabidopsis. *Proceedings of the National Academy of Sciences*, **98**, 1711–1716.
Gammage, P.A., Rorbach, J., Vincent, A.I., Rebar, E.J. and Minczuk, M. (2014) Mitochondrially targeted ZFNs for selective degradation of pathogenic mitochondrial genomes bearing large-scale deletions or point mutations. *EMBO Mol Med*, **6**, 458–466.
Gammage, P.A., Viscomi, C., Simard, M.-L., et al. (2018) Genome editing in mitochondria corrects a pathogenic mtDNA mutation in vivo. *Nat Med*, **24**, 1691–1695.
Ganai, R.A. and Johansson, E. (2016) DNA Replication—A Matter of Fidelity. *Mol Cell*, **62**, 745–755.
Gao, B., Ren, G., Wen, T., Li, H., Zhang, X. and Lin, Z. (2022) A super PPR cluster for restoring fertility revealed by genetic mapping, homocap-seq and de novo assembly in cotton. *Theoretical and Applied Genetics*, **135**, 637–652.
Garcia, V., Phelps, S.E.L., Gray, S. and Neale, M.J. (2011) Bidirectional resection of DNA double-strand breaks by Mre11 and Exo1. *Nature*, **479**, 241–244.
García-Medel, P.L., Baruch-Torres, N., Peralta-Castro, A., Trasviña-Arenas, C.H., Torres-Larios, A. and Brieba, L.G. (2019) Plant organellar DNA polymerases repair double-stranded breaks by microhomology-mediated end-joining. *Nucleic Acids Res*, **47**, 3028–3044.
García-Medel, P.L., Peralta-Castro, A., Baruch-Torres, N., Fuentes-Pascacio, A., Pedroza-García, J.A., Cruz-Ramirez, A. and Brieba, L.G. (2021) Arabidopsis thaliana PrimPol is a primase and lesion bypass DNA polymerase with the biochemical

characteristics to cope with DNA damage in the nucleus, mitochondria, and chloroplast. *Sci Rep*, **11**, 20582.

Gehrke, F., Schindele, A. and Puchta, H. (2022) Nonhomologous end joining as key to CRISPR/Cas-mediated plant chromosome engineering. *Plant Physiol*, **188**, 1769–1779.

Gestel, K. Van, Köhler, R.H. and Verbelen, J. (2002) Plant mitochondria move on F-actin, but their positioning in the cortical cytoplasm depends on both F-actin and microtubules. *J Exp Bot*, **53**, 659–667.

Ghodgaonkar, M.M., Lazzaro, F., Olivera-Pimentel, M., et al. (2013) Ribonucleotides Misincorporated into DNA Act as Strand-Discrimination Signals in Eukaryotic Mismatch Repair. *Mol Cell*, **50**, 323–332.

Giannakis, K., Chustecki, J.M. and Johnston, I.G. (2022) Exchange on dynamic encounter networks allows plant mitochondria to collect complete sets of mitochondrial DNA products despite their incomplete genomes. *Quantitative Plant Biology*, **3**, e18.

Gil, R., Silva, F.J., Zientz, E., et al. (2003) The genome sequence of *Blochmannia floridanus* : Comparative analysis of reduced genomes. *Proceedings of the National Academy of Sciences*, **100**, 9388–9393.

Gilbert, L.A., Larson, M.H., Morsut, L., et al. (2013) CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell*, **154**, 442–451.

Gilkerson, R., Bravo, L., Garcia, I., Gaytan, N., Herrera, A., Maldonado, A. and Quintanilla, B. (2013) The mitochondrial nucleoid: integrating mitochondrial DNA into cellular homeostasis. *Cold Spring Harb. Perspect. Biol.*, **5**, a011080

Glickman, B.W. and Radman, M. (1980) Escherichia coli mutator mutants deficient in methylation-instructed DNA mismatch correction. *Proceedings of the National Academy of Sciences*, **77**, 1063–1067.

Golin, S., Negroni, Y.L., Bennewitz, B., et al. (2020) WHIRLY2 plays a key role in mitochondria morphology, dynamics, and functionality in Arabidopsis thaliana. *Plant Direct*, **4**.

Gonzalez, V. and Spampinato, C.P. (2020) The mismatch repair protein MSH6 regulates somatic recombination in Arabidopsis thaliana. *DNA Repair (Amst)*, **87**, 102789.

Goodarzi, A.A., Yu, Y., Riballo, E., et al. (2006) DNA-PK autophosphorylation facilitates Artemis endonuclease activity. *EMBO J*, **25**, 3880–3889.

Gopalakrishnan Nair, P.M. and Chung, I.-M. (2014) Cell Cycle and Mismatch Repair Genes as Potential Biomarkers in Arabidopsis thaliana Seedlings Exposed to Silver Nanoparticles. *Bull Environ Contam Toxicol*, **92**, 719–725.

Graf, A., Schlereth, A., Stitt, M. and Smith, A.M. (2010) Circadian control of carbohydrate availability for growth in Arabidopsis plants at night. *Proceedings of the National Academy of Sciences*, **107**, 9458–9463.

Grawunder, U., Zimmer, D., Fugmann, S., Schwarz, K. and Lieber, M.R. (1998) DNA Ligase IV Is Essential for V(D)J Recombination and DNA Double-Strand Break Repair in Human Precursor Lymphocytes. *Mol Cell*, **2**, 477–484.

Greiner, S., Golczyk, H., Malinova, I., Pellizzer, T., Bock, R., Börner, T. and Herrmann, R.G. (2020) Chloroplast nucleoids are highly dynamic in ploidy, number, and structure during angiosperm leaf development. *The Plant Journal*, **102**, 730–746.

Grin, I.R., Petrova, D. V., Endutkin, A. V., Ma, C., Yu, B., Li, H. and Zharkov, D.O. (2023) Base Excision DNA Repair in Plants: Arabidopsis and Beyond. *Int J Mol Sci*, **24**, 14746.

Groothuizen, F.S. and Sixma, T.K. (2016) The conserved molecular machinery in DNA mismatch repair enzyme structures. *DNA Repair (Amst)*, **38**, 14–23.

Gu, J., Lu, H., Tippin, B., Shimazaki, N., Goodman, M.F. and Lieber, M.R. (2007) XRCC4:DNA ligase IV can ligate incompatible DNA ends and can ligate across gaps. *EMBO J*, **26**, 1010–1023.

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

Gualberto, J.M. and Newton, K.J. (2017) Plant Mitochondrial Genomes: Dynamics and Mechanisms of Mutation. *Annu Rev Plant Biol*, **68**, 225–252

Gualberto, J.M., Ret, M. Le, Beator, B. and Kuhn, K. (2015) The RAD52-like protein ODB1 is required for the efficient excision of two mitochondrial introns spliced via first-step hydrolysis. *Nucleic Acids Res.*, **43**, 6500–6510.

Guarné, A. (2012) The Functions of MutL in Mismatch Repair. In pp. 41–70.

Guilliermond, M.A.A. (1911) Sur les mitochondries des cellules vegetales. *C. R. Acad. Sci.* , **153**, 183–185.

Guo, F., Hu, S.-Y., Yuan, Z., Zee, S.-Y. and Han, Y. (2005) Paternal cytoplasmic transmission in Chinese pine (Pinus tabulaeformis). *Protoplasma*, **225**, 5–14.

Guo, H., Bueler, S.A. and Rubinstein, J.L. (2017) Atomic model for the dimeric F $_{O}$ region of mitochondrial ATP synthase. *Science (1979)*, **358**, 936–940.

Guo, J. and Liu, Y. (2012) Molecular Control of Male Reproductive Development and Pollen Fertility in Rice ^{F}. *J Integr Plant Biol*, **54**, 967–978.

Gutman, B.L. and Niyogi, K.K. (2009) Evidence for Base Excision Repair of Oxidative DNA Damage in Chloroplasts of Arabidopsis thaliana. *Journal of Biological Chemistry*, **284**, 17006–17012.

Haber, J.E. (2006) Transpositions and translocations induced by site-specific double-strand breaks in budding yeast. *DNA Repair (Amst)*, **5**, 998–1009.

Hajnóczky, G., Csordás, G., Das, S., Garcia-Perez, C., Saotome, M., Sinha Roy, S. and Yi, M. (2006) Mitochondrial calcium signalling and cell death: Approaches for assessing the role of mitochondrial Ca2+ uptake in apoptosis. *Cell Calcium*, **40**, 553–560.

Hallberg, B.M. and Larsson, N.-G. (2011) TFAM forces mtDNA to make a U-turn. *Nat Struct Mol Biol*, **18**, 1179–1181.

Hammani, K., Gobert, A., Hleibieh, K., Choulier, L., Small, I. and Giegé, P. (2011) An *Arabidopsis* Dual-Localized Pentatricopeptide Repeat Protein Interacts with Nuclear Proteins Involved in Gene Expression Regulation. *Plant Cell*, **23**, 730–740.

Hanamshet, K., Mazina, O. and Mazin, A. (2016) Reappearance from Obscurity: Mammalian Rad52 in Homologous Recombination. *Genes (Basel)*, **7**, 63.

Hanawalt, P.C. (2002) Subpathways of nucleotide excision repair and their regulation. *Oncogene*, **21**, 8949–8956.

Handa, H. (2003) The complete nucleotide sequence and RNA editing content of the mitochondrial genome of rapeseed (Brassica napus L.): comparative analysis of the mitochondrial genomes of rapeseed and Arabidopsis thaliana. *Nucleic Acids Res.*, **31**, 5907–5916.

Harm, W. (1968) Dark repair of photorepairable uv lesions in escherichia coli. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, **6**, 25–35.

Hartley, J.L., Temple, G.F. and Brasch, M.A. (2000) DNA cloning using in vitro site-specific recombination. *Genome Res.*, **10**, 1788–1795.

Hashimoto, M., Bacman, S.R., Peralta, S., Falk, M.J., Chomyn, A., Chan, D.C., Williams, S.L. and Moraes, C.T. (2015) MitoTALEN: A General Approach to Reduce Mutant mtDNA Loads and Restore Oxidative Phosphorylation Function in Mitochondrial Diseases. *Molecular Therapy*, **23**, 1592–1599.

Hashimoto, S., Anai, H. and Hanada, K. (2016) Mechanisms of interstrand DNA crosslink repair and human disorders. *Genes and Environment*, **38**, 9.

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.

Havey, M.J., Park, Y.H. and Bartoszewski, G. (2004) The Psm locus controls paternal sorting of the cucumber mitochondrial genome. *J. Hered.*, **95**, 492–497.

Hayashi, J., Takemitsu, M., Goto, Y. and Nonaka, I. (1994) Human mitochondria and mitochondrial genome function as a single dynamic cellular unit. *J Cell Biol*, **125**, 43–50.

Hays, J.B. (2002) Arabidopsis thaliana, a versatile model system for study of eukaryotic genome-maintenance functions. *DNA Repair (Amst)*, **1**, 579–600.

He, J., Ford, H.C., Carroll, J., Douglas, C., Gonzales, E., Ding, S., Fearnley, I.M. and Walker, J.E. (2018) Assembly of the membrane domain of ATP synthase in human mitochondria. *Proceedings of the National Academy of Sciences*, **115**, 2988–2993.

Hede, M.S., Petersen, R.L., Frøhlich, R.F., Krüger, D., Andersen, F.F., Andersen, A.H. and Knudsen, B.R. (2007) Resolution of Holliday Junction Substrates by Human Topoisomerase I. *J Mol Biol*, **365**, 1076–1092.

Hedtke, B., Borner, T. and Weihe, A. (1997) Mitochondrial and chloroplast phage-type RNA polymerases in Arabidopsis. *Science (1979)*, **277**, 809–811.

Hegde, M.L., Hazra, T.K. and Mitra, S. (2010) Functions of disordered regions in mammalian early base excision repair proteins. *Cellular and Molecular Life Sciences*, **67**, 3573–3587.

Hegde, M.L., Izumi, T. and Mitra, S. (2012) Oxidized Base Damage and Single-Strand Break Repair in Mammalian Genomes. In pp. 123–153.

Hendrix, S., Keunen, E., Mertens, A.I.G., Beemster, G.T.S., Vangronsveld, J. and Cuypers, A. (2018) Cell cycle regulation in different leaves of Arabidopsis thaliana plants grown under control and cadmium-exposed conditions. *Environ Exp Bot*, **155**, 441–452.

Hikosaka, K., Watanabe, Y. i., Tsuji, N., et al. (2010) Divergence of the Mitochondrial Genome Structure in the Apicomplexan Parasites, Babesia and Theileria. *Mol Biol Evol*, **27**, 1107–1116.

Hill, R.F. (1970) Location of genes controlling excision repair of UV damage and mutator activity in Escherichia coli WP2. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, **9**, 341–344.

Hiller, S., Abramson, J., Mannella, C., Wagner, G. and Zeth, K. (2010) The 3D structures of VDAC represent a native conformation. *Trends Biochem Sci*, **35**, 514–521.

Hjort, K., Goldberg, A. V., Tsaousis, A.D., Hirt, R.P. and Embley, T.M. (2010) Diversity and reductive evolution of mitochondria among microbial eukaryotes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **365**, 713–727.

Hogeboom, G.H., Schneider, W.C. and Pallade, G.E. (1948) Cytochemical studies of mammalian tissues; isolation of intact mitochondria from rat liver; some biochemical properties of mitochondria and submicroscopic particulate material. *J Biol Chem*, **172**, 619–35.

Holliday, R. (1964) A mechanism for gene conversion in fungi. *Genet Res*, **5**, 282–304.

Hollingsworth, N.M., Ponte, L. and Halsey, C. (1995) MSH5, a novel MutS homolog, facilitates meiotic reciprocal recombination between homologs in Saccharomyces cerevisiae but not mismatch repair. *Genes Dev*, **9**, 1728–1739.

Holt, I.J., Lorimer, H.E. and Jacobs, H.T. (2000) Coupled Leading- and Lagging-Strand Synthesis of Mammalian Mitochondrial DNA. *Cell*, **100**, 515–524.

Hördt, A., López, M.G., Meier-Kolthoff, J.P., et al. (2020) Analysis of 1,000+ Type-Strain Genomes Substantially Improves Taxonomic Classification of Alphaproteobacteria. *Front Microbiol*, **11**.

Hu, J., Yi, R., Zhang, H. and Ding, Y. (2013) Nucleo-cytoplasmic interactions affect RNA editing of cox2, atp6 and atp9 in alloplasmic male-sterile rice (Oryza sativa L.) lines. *Mitochondrion*, **13**, 87–95.

Huang, J.C., Svoboda, D.L., Reardon, J.T. and Sancar, A. (1992) Human nucleotide excision nuclease removes thymine dimers from DNA by incising the 22nd phosphodiester bond 5' and the 6th phosphodiester bond 3' to the photodimer. *Proceedings of the National Academy of Sciences*, **89**, 3664–3668.

Huffman, J.L., Sundheim, O. and Tainer, J.A. (2005) DNA base damage recognition and removal: New twists and grooves. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, **577**, 55–76.

Huisman, T.A.G.M., Kralik, S.F., Desai, N.K., Serrallach, B.L. and Orman, G. (2022) Neuroimaging of primary mitochondrial disorders in children: A review. *Journal of Neuroimaging*, **32**, 191–200.

Iborra, F.J., Kimura, H. and Cook, P.R. (2004) The functional organization of mitochondrial genomes in human cells. *BMC Biol*, **2**, 9.

Ischenko, A.A. and Saparbaev, M.K. (2002) Alternative nucleotide incision repair pathway for oxidative DNA damage. *Nature*, **415**, 183–187.

Ithurbide, S., Bentchikou, E., Coste, G., Bost, B., Servant, P. and Sommer, S. (2015) Single Strand Annealing Plays a Major Role in RecA-Independent Recombination

between Repeated Sequences in the Radioresistant Deinococcus radiodurans Bacterium. *PLoS Genet*, **11**, e1005636.
Iwamoto, T., Hiraku, Y., Oikawa, S., Mizutani, H., Kojima, M. and Kawanishi, S. (2004) DNA intrastrand cross-link at the 5′-GA-3′ sequence formed by busulfan and its role in the cytotoxic effect. *Cancer Sci*, **95**, 454–458.
Jackman, S.D., Coombe, L., Warren, R.L., et al. (2020) Complete Mitochondrial Genome of a Gymnosperm, Sitka Spruce (Picea sitchensis), Indicates a Complex Physical Structure. *Genome Biol Evol*, **12**, 1174–1179.
Jackson, N., Sanchez-Moran, E., Buckling, E., Armstrong, S.J., Jones, G.H. and Franklin, F.C.H. (2006) Reduced meiotic crossovers and delayed prophase I progression in AtMLH3-deficient Arabidopsis. *EMBO J*, **25**, 1315–1323.
Jäger, R., Gisslinger, H., Passamonti, F., et al. (2010) Deletions of the transcription factor Ikaros in myeloproliferative neoplasms. *Leukemia*, **24**, 1290–1298.
Janicka, S., Kuhn, K., Ret, M. Le, Bonnard, G., Imbault, P., Augustyniak, H. and Gualberto, J.M. (2012) A RAD52-like single-stranded DNA binding protein affects mitochondrial DNA repair by recombination. *Plant J*, **72**, 423–435.
Janska, H., Sarria, R., Woloszynska, M., Arrieta-Montiel, M. and Mackenzie, S.A. (1998) Stoichiometric shifts in the common bean mitochondrial genome leading to male sterility and spontaneous reversion to fertility. *Plant Cell*, **10**, 1163–1180.
Jenkins, G.I. (2017) Photomorphogenic responses to ultraviolet-B light. *Plant Cell Environ*, **40**, 2544–2557.
Jensen, R.B., Carreira, A. and Kowalczykowski, S.C. (2010) Purified human BRCA2 stimulates RAD51-mediated recombination. *Nature*, **467**, 678–683.
Jia, Q., Dulk-Ras, A. den, Shen, H., Hooykaas, P.J.J. and Pater, S. de (2013) Poly(ADP-ribose)polymerases are involved in microhomology mediated back-up non-homologous end joining in Arabidopsis thaliana. *Plant Mol Biol*, **82**, 339–351.
Jiang, M., Xie, X., Zhu, X., et al. (2021) The mitochondrial single-stranded DNA binding protein is essential for initiation of mtDNA replication. *Sci Adv*, **7**.
Jilani, A., Ramotar, D., Slack, C., Ong, C., Yang, X.M., Scherer, S.W. and Lasko, D.D. (1999) Molecular Cloning of the Human Gene, PNKP, Encoding a Polynucleotide Kinase 3′-Phosphatase and Evidence for Its Role in Repair of DNA Strand Breaks Caused by Oxidative Damage. *Journal of Biological Chemistry*, **274**, 24176–24186.
Jindal, S.K., Dhaliwal, M.S. and Meena, O.P. (2020) Molecular advancements in male sterility systems of *Capsicum* : A review. *Plant Breeding*, **139**, 42–64.
Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A. and Charpentier, E. (2012) A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science (1979)*, **337**, 816–821.
Jõers, P. and Jacobs, H.T. (2013) Analysis of Replication Intermediates Indicates That Drosophila melanogaster Mitochondrial DNA Replicates by a Strand-Coupled Theta Mechanism. *PLoS One*, **8**, e53249.
Johnson, R.A., Hellens, R.P. and Love, D.R. (2011) A transient assay for recombination demonstrates that Arabidopsis SNM1 and XRCC3 enhance non-homologous recombination. *Genetics and Molecular Research*, **10**, 2104–2132.
Johnson, R.E., Prakash, S. and Prakash, L. (1999) Efficient Bypass of a Thymine-Thymine Dimer by Yeast DNA Polymerase, Polη. *Science (1979)*, **283**, 1001–1004.
Johnston, S.A., Anziano, P.Q., Shark, K., Sanford, J.C. and Butow, R.A. (1988) Mitochondrial transformation in yeast by bombardment with microprojectiles. *Science (1979)*, **240**, 1538–1541.
Joldybayeva, B., Prorok, P., Grin, I.R., Zharkov, D.O., Ishenko, A.A., Tudek, B., Bissenbaev, A.K. and Saparbaev, M. (2014) Cloning and Characterization of a Wheat Homologue of Apurinic/Apyrimidinic Endonuclease Ape1L. *PLoS One*, **9**, e92963.
Josephs, E.A. and Marszalek, P.E. (2018) Endonuclease-independent DNA mismatch repair processes on the lagging strand. *DNA Repair (Amst)*, **68**, 41–49.
Jung, D., Cho, Y., Collins, L.B., Swenberg, J.A. and Giulio, R.T. Di (2009) Effects of benzo[a]pyrene on mitochondrial and nuclear DNA damage in Atlantic killifish (Fundulus

heteroclitus) from a creosote-contaminated and reference site. *Aquatic Toxicology*, **95**, 44–51.
Kadyrov, F.A., Dzantiev, L., Constantin, N. and Modrich, P. (2006) Endonucleolytic Function of MutLα in Human Mismatch Repair. *Cell*, **126**, 297–308.
Kadyrova, L.Y. and Kadyrov, F.A. (2016) Endonuclease activities of MutLα and its homologs in DNA mismatch repair. *DNA Repair (Amst)*, **38**, 42–49.
Kamenisch, Y., Fousteri, M., Knoch, J., et al. (2010) Proteins of nucleotide and base excision repair pathways interact in mitochondria to protect from loss of subcutaneous fat, a hallmark of aging. *Journal of Experimental Medicine*, **207**, 379–390.
Kaneko, T. (2000) Complete Genome Structure of the Nitrogen-fixing Symbiotic Bacterium Mesorhizobium loti. *DNA Research*, **7**, 331–338.
Kato, N., Kawasoe, Y., Williams, H., et al. (2017) Sensing and Processing of DNA Interstrand Crosslinks by the Mismatch Repair Pathway. *Cell Rep*, **21**, 1375–1385.
Kaul, M.L.H. (1988) *Male Sterility in Higher Plants*, New York: Springer-Verlag .
Kazama, T., Okuno, M., Watari, Y., et al. (2019) Curing cytoplasmic male sterility via TALEN-mediated mitochondrial genome editing. *Nat Plants*, **5**, 722–730.
Kennedy, E.P. and Lehninger, A.L. (1949) Oxidation of fatty acids and tricarboxylic acid cycle intermediates by isolated rat liver mitochondria. *Journal of Biological Chemistry*, **179**, 957–972.
Kent, T., Chandramouly, G., McDevitt, S.M., Ozdemir, A.Y. and Pomerantz, R.T. (2015) Mechanism of microhomology-mediated end-joining promoted by human DNA polymerase θ. *Nat Struct Mol Biol*, **22**, 230–237.
Kim, D.H., Kang, J.G. and Kim, B.-D. (2007) Isolation and characterization of the cytoplasmic male sterility-associated orf456 gene of chili pepper (Capsicum annuum L.). *Plant Mol Biol*, **63**, 519–532.
Kim, J.-H., Ryu, T.H., Lee, S.S., Lee, S. and Chung, B.Y. (2019) Ionizing radiation manifesting DNA damage response in plants: An overview of DNA damage signaling and repair mechanisms in plants. *Plant Science*, **278**, 44–53.
Kim, M., Schulz, V., Brings, L., Schoeller, T., Kühn, K. and Vierling, E. (2021) mTERF18 and ATAD3 are required for mitochondrial nucleoid structure and their disruption confers heat tolerance in *Arabidopsis thaliana*. *New Phytologist*, **232**, 2026–2042.
Kim, Y.G., Cha, J. and Chandrasegaran, S. (1996) Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proceedings of the National Academy of Sciences*, **93**, 1156–1160.
Kimura, S., Ueda, T., Hatanaka, M., Takenouchi, M., Hashimoto, J. and Sakaguchi, K. (2000) Plant homologue of flap endonuclease-1: molecular cloning, characterization, and evidence of expression in meristematic tissues. *Plant Mol Biol*, **42**, 415–427.
Kmiec, B., Woloszynska, M. and Janska, H. (2006) Heteroplasmy as a common state of mitochondrial genetic information in plants and animals. *Curr. Genet.*, **50**, 149–159.
Knoop, V. (2004) The mitochondrial DNA of land plants: peculiarities in phylogenetic perspective. *Curr Genet*, **46**.
Korhonen, J.A., Pham, X.H., Pellegrini, M. and Falkenberg, M. (2004) Reconstitution of a minimal mtDNA replisome in vitro. *EMBO J*, **23**, 2423–2429.
Kose, C., Cao, X., Dewey, E.B., Malkoç, M., Adebali, O., Sekelsky, J., Lindsey-Boltz, L.A. and Sancar, A. (2024) Cross-species investigation into the requirement of XPA for nucleotide excision repair. *Nucleic Acids Res*, **52**, 677–689.
Kozik, A., Rowan, B.A., Lavelle, D., Berke, L., Schranz, M.E., Michelmore, R.W. and Christensen, A.C. (2019) The alternative reality of plant mitochondrial DNA: One ring does not rule them all. *PLoS Genet*, **15**, e1008373.
Krause, K., Kilbienski, I., Mulisch, M., Rodiger, A., Schafer, 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.
Kraytsberg, Y., Schwartz, M., Brown, T.A., Ebralidse, K., Kunz, W.S., Clayton, D.A., Vissing, J. and Khrapko, K. (2004) Recombination of Human Mitochondrial DNA. *Science (1979)*, **304**, 981–981.

Krebs, H.A. and Johnson, W.A. (1937) Metabolism of ketonic acids in animal tissues. *Biochemical Journal*, **31**, 645–660.
Krokan, H.E. and Bjoras, M. (2013) Base Excision Repair. *Cold Spring Harb Perspect Biol*, **5**, a012583–a012583.
Krupinska, K., Desel, C., Frank, S. and Hensel, G. (2022) WHIRLIES Are Multifunctional DNA-Binding Proteins With Impact on Plant Development and Stress Resistance. *Front Plant Sci*, **13**.
Kruse, B., Narasimhan, N. and Attardi, G. (1989) Termination of transcription in human mitochondria: Identification and purification of a DNA binding protein factor that promotes termination. *Cell*, **58**, 391–397.
Kubo, S. and Takada, S. (2022) Rotational Mechanism of FO Motor in the F-Type ATP Synthase Driven by the Proton Motive Force. *Front Microbiol*, **13**.
Kubota, Y., Nash, R.A., Klungland, A., Schär, P., Barnes, D.E. and Lindahl, T. (1996) Reconstitution of DNA base excision-repair with purified human proteins: interaction between DNA polymerase beta and the XRCC1 protein. *EMBO J*, **15**, 6662–6670.
Kukat, C., Davies, K.M., Wurm, C.A., et al. (2015) Cross-strand binding of TFAM to a single mtDNA molecule forms the mitochondrial nucleoid. *Proceedings of the National Academy of Sciences*, **112**, 11288–11293.
Kukat, C., Wurm, C.A., Spåhr, H., Falkenberg, M., Larsson, N.-G. and Jakobs, S. (2011) Super-resolution microscopy reveals that mammalian mitochondrial nucleoids have a uniform size and frequently contain a single copy of mtDNA. *Proceedings of the National Academy of Sciences*, **108**, 13534–13539.
Kumar, K.R.R. (2023) Lost in the bloom: DNA-PKcs in green plants. *Front Plant Sci*, **14**.
Kundariya, H., Sanchez, R., Yang, X., Hafner, A. and Mackenzie, S.A. (2022) Methylome decoding of RdDM-mediated reprogramming effects in the Arabidopsis MSH1 system. *Genome Biol*, **23**, 167.
Kunz, B.A., Anderson, H.J., Osmond, M.J. and Vonarx, E.J. (2005) Components of nucleotide excision repair and DNA damage tolerance in Arabidopsis thaliana. *Environ Mol Mutagen*, **45**, 115–127.
Kuroiwa, T., Fujie, M. and Kuroiwa, H. (1992) Studies on the behavior of mitochondrial DNA: Synthesis of mitochondrial dna occurs actively in a specific region just above the quiescent center in the root meristem of *Pelargonium zonale*. *J Cell Sci*, **101**, 483–493.
Kutik, S., Stojanovski, D., Becker, T., et al. (2008) Response: The Mitochondrial β-Signal and Protein Sorting. *Cell*, **135**, 1159–1160.
Kuzminov, A. (2001) Single-strand interruptions in replicating chromosomes cause double-strand breaks. *Proceedings of the National Academy of Sciences*, **98**, 8241–8246.
Kwon, T., Huq, E. and Herrin, D.L. (2010) Microhomology-mediated and nonhomologous repair of a double-strand break in the chloroplast genome of *Arabidopsis*. *Proceedings of the National Academy of Sciences*, **107**, 13954–13959.
Ladoukakis, E.D. and Zouros, E. (2001) Direct Evidence for Homologous Recombination in Mussel (Mytilus galloprovincialis) Mitochondrial DNA. *Mol Biol Evol*, **18**, 1168–1175.
Laemmli, U.K. (1970) Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature*, **227**, 680–685.
Lahue, R.S., Au, K.G. and Modrich, P. (1989) DNA Mismatch Correction in a Defined System. *Science (1979)*, **245**, 160–164.
Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods*, **9**, 357–359.
Lankinen, M.H., Vilpo, L.M. and Vilpo, J.A. (1996) UV- and γ-irradiation-induced DNA single-strand breaks and their repair in human blood granulocytes and lymphocytes. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, **352**, 31–38.
Lao, P.J. and Forsdyke, D.R. (2000) Crossover hot-spot instigator (Chi) sequences in Escherichia coli occupy distinct recombination/transcription islands. *Gene*, **243**, 47–57.
Lauer, M., Knudsen, C., Newton, K., Gabay-Laughnan, S. and Laughnan, J. (1990) A partially deleted mitochondrial cytochrome oxidase gene in the NCS6 abnormal growth mutant of maize. *The New Biologist. 1990 Feb;2(2):179-186.*, **2**, 179–186.

Laulederkind, S.J.F., Hayman, G.T., Wang, S., et al. (2023) The Rat Genome Database: Genetic, Genomic, and Phenotypic Data Across Multiple Species. *Curr Protoc*, **3**.
Lavelle, C. and Foray, N. (2014) Chromatin structure and radiation-induced DNA damage: From structural biology to radiobiology. *Int J Biochem Cell Biol*, **49**, 84–97.
LeDoux, S.P., Wilson, G.L., Beecham, E.J., Stevnsner, T., Wassermann, K. and Bohr, V.A. (1992) Repair of mitochondrial DNA after various types of DNA damage in Chinese hamster ovary cells. *Carcinogenesis*, **13**, 1967–1973.
Lee, C.-D. and Wang, T.-F. (2009) The N-terminal domain of Escherichia coli RecA have multiple functions in promoting homologous recombination. *J Biomed Sci*, **16**, 37.
Lee, H., Lee, S., Baek, G., Kim, A., Kang, B.-C., Seo, H. and Kim, J.-S. (2021) Mitochondrial DNA editing in mice with DddA-TALE fusion deaminases. *Nat Commun*, **12**, 1190.
Lee, J.Y., Chang, J., Joseph, N., Ghirlando, R., Rao, D.N. and Yang, W. (2005) MutH Complexed with Hemi- and Unmethylated DNAs: Coupling Base Recognition and DNA Cleavage. *Mol Cell*, **20**, 155–166.
Lee, S.-J. and Richardson, C.C. (2011) Choreography of bacteriophage T7 DNA replication. *Curr Opin Chem Biol*, **15**, 580–586.
Lee, S.R. and Han, J. (2017) Mitochondrial Nucleoid: Shield and Switch of the Mitochondrial Genome. *Oxid Med Cell Longev*, **2017**, 1–15.
Levings, C.S. (1993) Thoughts on Cytoplasmic Male Sterility in cms-T Maize. *Plant Cell*, 1285–1290.
Lewis, M.R. and Lewis, W.H. (1914) Mitochondria in Tissue Culture. *Science (1979)*, **39**, 330–333.
Lewis, S.C., Joers, P., Willcox, S., Griffith, J.D., Jacobs, H.T. and Hyman, B.C. (2015) A Rolling Circle Replication Mechanism Produces Multimeric Lariats of Mitochondrial DNA in Caenorhabditis elegans. *PLoS Genet*, **11**, e1004985.
Li, G., Jain, R., Chern, M., et al. (2017) The Sequences of 1504 Mutants in the Model Rice Variety Kitaake Facilitate Rapid Functional Genomic Studies. *Plant Cell*, **29**, 1218–1231.
Li, H. and Durbin, R. (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, **25**, 1754–1760.
Li, J. and Cullis, C. (2021) The Multipartite Mitochondrial Genome of Marama (Tylosema esculentum). *Front Plant Sci*, **12**.
Li, L., Dion, E., Richard, G., Domingue, O., Jean, M. and Belzile, F.J. (2009) The Arabidopsis DNA mismatch repair gene PMS1 restricts somatic recombination between homeologous sequences. *Plant Mol Biol*, **69**, 675–684.
Li, T., Huang, S., Jiang, W.Z., Wright, D., Spalding, M.H., Weeks, D.P. and Yang, B. (2011) TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. *Nucleic Acids Res*, **39**, 359–372.
Li, T., Huang, S., Zhao, X., Wright, D.A., Carpenter, S., Spalding, M.H., Weeks, D.P. and Yang, B. (2011) Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes. *Nucleic Acids Res*, **39**, 6315–6325.
Li, W.-Q., Zhang, X.-Q., Xia, C., Deng, Y. and Ye, D. (2010) MALE GAMETOPHYTE DEFECTIVE 1, encoding the FAd subunit of mitochondrial F1FO-ATP synthase, is essential for pollen formation in Arabidopsis thaliana. *Plant Cell Physiol*, **51**, 923–935.
Li, X., Huang, W., Yang, H., Jiang, R., Sun, F., Wang, H., Zhao, J., Xu, C. and Tan, B. (2019) EMP18 functions in mitochondrial atp6 and cox2 transcript editing and is essential to seed development in maize. *New Phytologist*, **221**, 896–907.
Lilly, J. and Camps, M. (2015) Mechanisms of Theta Plasmid Replication. *Microbiol Spectr*, **3**.
Limoli, C.L., Giedzinski, E., Bonner, W.M. and Cleaver, J.E. (2002) UV-induced replication arrest in the xeroderma pigmentosum variant leads to DNA double-strand breaks, γ-H2AX formation, and Mre11 relocalization. *Proceedings of the National Academy of Sciences*, **99**, 233–238.
Limón-Pacheco, J. and Gonsebatt, M.E. (2009) The role of antioxidants and antioxidant-related enzymes in protective responses to environmentally induced oxidative stress. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, **674**, 137–147.

Lindahl, T. (1974) An N-Glycosidase from Escherichia coli That Releases Free Uracil from DNA Containing Deaminated Cytosine Residues. *Proceedings of the National Academy of Sciences*, **71**, 3649–3653.
Lindahl, T. (1993) Instability and decay of the primary structure of DNA. *Nature*, **362**, 709–715.
Linke, B., Nothnagel, T. and Börner, T. (2003) Flower development in carrot CMS plants: mitochondria affect the expression of MADS box genes homologous to GLOBOSA and DEFICIENS. *The Plant Journal*, **34**, 27–37.
Logan, D.C. (2010) Mitochondrial fusion, division and positioning in plants. *Biochem Soc Trans*, **38**, 789–795.
Longley, M.J., Prasad, R., Srivastava, D.K., Wilson, S.H. and Copeland, W.C. (1998) Identification of 5'-deoxyribose phosphate lyase activity in human DNA polymerase gamma and its role in mitochondrial base excision repair in vitro. *Proceedings of the National Academy of Sciences*, **95**, 12244–12248.
Lonsdale, D.M., Hodge, T.P. and Fauron, C.M. (1984) The physical map and organisation of the mitochondrial genome from the fertile cytoplasm of maize. *Nucleic Acids Res*, **12**, 9249–9261.
Löytynoja, A. and Goldman, N. (2017) Short template switch events explain mutation clusters in the human genome. *Genome Res*, **27**, 1039–1049.
Lu, H. and Davis, A.J. (2021) Human RecQ Helicases in DNA Double-Strand Break Repair. *Front Cell Dev Biol*, **9**.
Luijsterburg, M.S., Lindh, M., Acs, K., Vrouwe, M.G., Pines, A., Attikum, H. van, Mullenders, L.H. and Dantuma, N.P. (2012) DDB2 promotes chromatin decondensation at UV-induced DNA damage. *Journal of Cell Biology*, **197**, 267–281.
Lujan, S.A., Williams, J.S., Clausen, A.R., Clark, A.B. and Kunkel, T.A. (2013) Ribonucleotides Are Signals for Mismatch Repair of Leading-Strand Replication Errors. *Mol Cell*, **50**, 437–443.
Lv, Q., Han, S., Wang, L., et al. (2022) TEB/POLQ plays dual roles in protecting Arabidopsis from NO-induced DNA damage. *Nucleic Acids Res*, **50**, 6820–6836.
Lyabin, D.N., Eliseeva, I.A. and Ovchinnikov, L.P. (2014) YB-1 protein: functions and regulation. *WIREs RNA*, **5**, 95–110.
Ma, Y., Schwarz, K. and Lieber, M.R. (2005) The Artemis:DNA-PKcs endonuclease cleaves DNA loops, flaps, and gaps. *DNA Repair (Amst)*, **4**, 845–851.
Mackenzie, S.A., Pring, D.R., Bassett, M.J. and Chase, C.D. (1988) Mitochondrial DNA rearrangement associated with fertility restoration and cytoplasmic reversion to fertility in cytoplasmic male sterile Phaseolus vulgaris L. *Proc Natl Acad Sci U S A*, **85**, 2714–2717.
Maddukuri, L., Dudzińska, D. and Tudek, B. (2007) Bacterial DNA repair genes and their eukaryotic homologues: 4. The role of nucleotide excision DNA repair (NER) system in mammalian cells. *Acta Biochim Pol*, **54**, 469–482.
Maeder, M.L., Thibodeau-Beganny, S., Osiak, A., et al. (2008) Rapid "Open-Source" Engineering of Customized Zinc-Finger Nucleases for Highly Efficient Gene Modification. *Mol Cell*, **31**, 294–301.
Mahfouz, M.M., Li, L., Shamimuzzaman, Md., Wibowo, A., Fang, X. and Zhu, J.-K. (2011) De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks. *Proceedings of the National Academy of Sciences*, **108**, 2623–2628.
Makarenko, M.S., Kornienko, I. V., Azarin, K. V., Usatov, A. V., Logacheva, M.D., Markin, N. V. and Gavrilova, V.A. (2018) Mitochondrial genomes organization in alloplasmic lines of sunflower (*Helianthus annuus* L.) with various types of cytoplasmic male sterility. *PeerJ*, **6**, e5266.
Makarova, K.S., Zhang, F. and Koonin, E. V. (2017) SnapShot: Class 2 CRISPR-Cas Systems. *Cell*, **168**, 328-328.e1.
Malzahn, A., Lowder, L. and Qi, Y. (2017) Plant genome editing with TALEN and CRISPR. *Cell Biosci*, **7**, 21.

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

Manova, V. and Gruszka, D. (2015) DNA damage and repair in plants – from models to crops. *Front Plant Sci*, **6**.

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

Margulis, L. (1981) *Symbiosis in Cell Evolution*, New York: W. H. Freeman.

Marie, L., Rapisarda, C., Morales, V., et al. (2017) Bacterial RadA is a DnaB-type helicase interacting with RecA to promote bidirectional D-loop extension. *Nat Commun*, **8**, 15638.

Marienfeld, J., Unseld, M. and Brennicke, A. (1999) The mitochondrial genome of Arabidopsis is composed of both native and immigrant information. *Trends Plant Sci*, **4**, 495–502.

Marienfeld, J.R. and Newton, K.J. (1994) The maize NCS2 abnormal growth mutant has a chimeric nad4-nad7 mitochondrial gene and is associated with reduced complex I function. *Genetics*, **138**, 855–863.

Marti, T.M., Kunz, C. and Fleck, O. (2002) DNA mismatch repair and mutation avoidance pathways. *J Cell Physiol*, **191**, 28–41.

Martijn, J., Vosseberg, J., Guy, L., Offre, P. and Ettema, T.J.G. (2018) Deep mitochondrial origin outside the sampled alphaproteobacteria. *Nature*, **557**, 101–105.

Martin, W., Stoebe, B., Goremykin, V., Hansmann, S., Hasegawa, M. and Kowallik, K. V. (1998) Gene transfer to the nucleus and the evolution of chloroplasts. *Nature*, **393**, 162–165.

Martinez-López, W. and Hande, M.P. (2020) Health effects of exposure to ionizing radiation. In *Advanced Security and Safeguarding in the Nuclear Power Industry*. Elsevier, pp. 81–97.

Marton, I., Zuker, A., Shklarman, E., Zeevi, V., Tovkach, A., Roffe, S., Ovadis, M., Tzfira, T. and Vainstein, A. (2010) Nontransgenic Genome Modification in Plant Cells. *Plant Physiol*, **154**, 1079–1087.

Mason, P.A. (2003) Mismatch repair activity in mammalian mitochondria. *Nucleic Acids Res*, **31**, 1052–1058.

Mateos-Gomez, P.A., Kent, T., Deng, S.K., McDevitt, S., Kashkina, E., Hoang, T.M., Pomerantz, R.T. and Sfeir, A. (2017) The helicase domain of Polθ counteracts RPA to promote alt-NHEJ. *Nat Struct Mol Biol*, **24**, 1116–1123.

Matson, S.W. and Robertson, A.B. (2006) The UvrD helicase and its modulation by the mismatch repair protein MutL. *Nucleic Acids Res*, **34**, 4089–4097.

Mawer, J.S.P. and Leach, D.R.F. (2014) Branch Migration Prevents DNA Loss during Double-Strand Break Repair. *PLoS Genet*, **10**, e1004485.

Mayr, E. (1986) Joseph Gottlieb Kolreuter's Contributions to Biology. *Osiris*, **2**, 135–176.

Mazin, A. V., Mazina, O.M., Bugreev, D. V. and Rossi, M.J. (2010) Rad54, the motor of homologous recombination. *DNA Repair (Amst)*, **9**, 286–302.

McCulloch, S.D. and Kunkel, T.A. (2008) The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. *Cell Res*, **18**, 148–161.

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

McIlwraith, M.J. and West, S.C. (2008) DNA Repair Synthesis Facilitates RAD52-Mediated Second-End Capture during DSB Repair. *Mol Cell*, **29**, 510–516.

McMahill, M.S., Sham, C.W. and Bishop, D.K. (2007) Synthesis-Dependent Strand Annealing in Meiosis. *PLoS Biol*, **5**, e299.

Mechanic, L.E., Frankel, B.A. and Matson, S.W. (2000) Escherichia coli MutL Loads DNA Helicase II onto DNA. *Journal of Biological Chemistry*, **275**, 38337–38346.

Meir, A., Raina, V.B., Rivera, C.E., Marie, L., Symington, L.S. and Greene, E.C. (2023) The separation pin distinguishes the pro– and anti–recombinogenic functions of Saccharomyces cerevisiae Srs2. *Nat Commun*, **14**, 8144.

Meng, C., Yang, M., Wang, Y., Chen, C., Sui, N., Meng, Q., Zhuang, K. and Lv, W. (2020) SlWHY2 interacts with SlRECA2 to maintain mitochondrial function under drought stress in tomato. *Plant Science*, **301**, 110674.

Meng, L., Liu, Z., Zhang, L., Hu, G. and Song, X. (2016) Cytological characterization of a thermo-sensitive cytoplasmic male-sterile wheat line having K-type cytoplasm of Aegilops kotschyi; *Breed Sci*, **66**, 752–761.

Mereschkowsky, K.S. (1910) Theorie der zwei Plasmaarten als Grundlage der Symbiogenesis, einer neuen Lehre von der Entstehung der Organismen. *Biologisches Centralblatt* , **30**.

Meselson, M. and Weigle, J.J. (1961) Chromosome breakage accompanying genetic recombination in bacteriophage. *Proceedings of the National Academy of Sciences*, **47**, 857–868.

Meves, F. (1904) Über das Vorkommen von Mitochondrien bezw. Chondromiten in Pflanzenzellen. *Ber. Dtsch. Bot. Ges.*, **22**, 284–286.

Meyer, E.H., Welchen, E. and Carrie, C. (2019) Assembly of the Complexes of the Oxidative Phosphorylation System in Land Plant Mitochondria. *Annu Rev Plant Biol*, **70**, 23–50.

Michel, B., Boubakri, H., Baharoglu, Z., LeMasson, M. and Lestini, R. (2007) Recombination proteins and rescue of arrested replication forks. *DNA Repair (Amst)*, **6**, 967–980.

Millar, A.Harvey., Eubel, H., J�nsch, L., Kruft, V., Heazlewood, J.L. and Braun, H.-P. (2004) Mitochondrial cytochrome c oxidase and succinate dehydrogenase complexes contain plant specific subunits. *Plant Mol Biol*, **56**, 77–90.

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

Mimitou, E.P. and Symington, L.S. (2010) Ku prevents Exo1 and Sgs1-dependent resection of DNA ends in the absence of a functional MRX complex or Sae2. *EMBO J*, **29**, 3358–3369.

Mimori, T. and Hardin, J.A. (1986) Mechanism of interaction between Ku protein and DNA. *Journal of Biological Chemistry*, **261**, 10375–10379.

Minczuk, M., Papworth, M.A., Miller, J.C., Murphy, M.P. and Klug, A. (2008) Development of a single-chain, quasi-dimeric zinc-finger nuclease for the selective degradation of mutated human mitochondrial DNA. *Nucleic Acids Res*, **36**, 3926–3938.

Miyakawa, I. (2017) Organization and dynamics of yeast mitochondrial nucleoids. *Proceedings of the Japan Academy, Series B*, **93**, 339–359.

Miyake, T. (1960) Mutator factor in Salmonella typhimurium. *Genetics*, **45**, 11–14.

Moldovan, G.-L., Pfander, B. and Jentsch, S. (2007) PCNA, the Maestro of the Replication Fork. *Cell*, **129**, 665–679.

Molinier, J., Lechner, E., Dumbliauskas, E. and Genschik, P. (2008) Regulation and Role of Arabidopsis CUL4-DDB1A-DDB2 in Maintaining Genome Integrity upon UV Stress. *PLoS Genet*, **4**, e1000093.

Moore, R.C., Kozyreva, O., Lebel-Hardenack, S., Siroky, J., Hobza, R., Vyskot, B. and Grant, S.R. (2003) Genetic and Functional Analysis of *DD44* , a Sex-Linked Gene From the Dioecious Plant *Silene latifolia* , Provides Clues to Early Events in Sex Chromosome Evolution. *Genetics*, **163**, 321–334.

Morbitzer, R., Elsaesser, J., Hausner, J. and Lahaye, T. (2011) Assembly of custom TALE-type DNA binding domains by modular cloning. *Nucleic Acids Res*, **39**, 5790–5799.

Moretton, A., Morel, F., Macao, B., et al. (2017) Selective mitochondrial DNA degradation following double-strand breaks. *PLoS One*, **12**, e0176795.

Morimatsu, K. and Kowalczykowski, S.C. (2003) RecFOR Proteins Load RecA Protein onto Gapped DNA to Accelerate DNA Strand Exchange. *Mol Cell*, **11**, 1337–1347.

Morley, S.A., Ahmad, N. and Nielsen, B.L. (2019) Plant Organelle Genome Replication. *Plants (Basel)*, **8**.

Morley, S.A. and Nielsen, B.L. (2017) Plant mitochondrial DNA. *Front Biosci (Landmark Ed)*, **22**, 1023–1032.

Mortensen, U.H., Bendixen, C., Sunjevaric, I. and Rothstein, R. (1996) DNA strand annealing is promoted by the yeast Rad52 protein. *Proceedings of the National Academy of Sciences*, **93**, 10729–10734.

Moscou, M.J. and Bogdanove, A.J. (2009) A Simple Cipher Governs DNA Recognition by TAL Effectors. *Science (1979)*, **326**, 1501–1501.

Motycka, T.A., Bessho, T., Post, S.M., Sung, P. and Tomkinson, A.E. (2004) Physical and Functional Interaction between the XPF/ERCC1 Endonuclease and hRad52. *Journal of Biological Chemistry*, **279**, 13634–13639.

Mu, D., Hsu, D.S. and Sancar, A. (1996) Reaction Mechanism of Human DNA Repair Excision Nuclease. *Journal of Biological Chemistry*, **271**, 8285–8294.

Mullins, E.A., Shi, R., Kotsch, L.A. and Eichman, B.F. (2015) A New Family of HEAT-Like Repeat Proteins Lacking a Critical Substrate Recognition Motif Present in Related DNA Glycosylases. *PLoS One*, **10**, e0127733.

Muñoz-Gómez, S.A., Susko, E., Williamson, K., Eme, L., Slamovits, C.H., Moreira, D., López-García, P. and Roger, A.J. (2022) Site-and-branch-heterogeneous analyses of an expanded dataset favour mitochondria as sister to known Alphaproteobacteria. *Nat Ecol Evol*, **6**, 253–262.

Nagata, T., Todoriki, S., Hayashi, T., Shibata, Y., Mori, M., Kanegae, H. and Kikuchi, S. (1999) γ-Radiation Induces Leaf Trichome Formation in Arabidopsis1. *Plant Physiol*, **120**, 113–120.

Nakano, A., Kishikawa, J., Mitsuoka, K. and Yokoyama, K. (2023) Mechanism of ATP hydrolysis dependent rotation of bacterial ATP synthase. *Nat Commun*, **14**, 4090.

Nassif, N., Penney, J., Pal, S., Engels, W.R. and Gloor, G.B. (1994) Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. *Mol Cell Biol*, **14**, 1613–1625.

Nay, S.L. and O'Connor, T.R. (2013) Direct Repair in Mammalian Cells. In *New Research Directions in DNA Repair*. InTech.

Neale, M.J., Pan, J. and Keeney, S. (2005) Endonucleolytic processing of covalent protein-linked DNA double-strand breaks. *Nature*, **436**, 1053–1057.

Nemoto, Y., Kawano, S., Nakamura, S., Mita, T., Nagata, T. and Kuroiwa, T. (1988) Studies on Plastid-Nuclei (Nucleoids) in Nicotiana tabacum L. I. Isolation of Proplastid-Nuclei from Cultured Cells and Identification of Proplastid-Nuclear Proteins. *Plant Cell Physiol*, **29**, 167–177.

Nevers, P. and Spatz, H.-C. (1975) Escherichia coli mutants uvr D and uvr E deficient in gene conversion of λ-heteroduplexes. *Mol Gen Genet*, **139**, 233–243.

Newton, K.J., Mariano, J.M., Gibson, C.M., Kuzmin, E. and Gabay-Laughnan, S. (1996) Involvement of S2 episomal sequences in the generation of NCS4 deletion mutation in maize mitochondria. *Dev Genet*, **19**, 277–286.

Niazi, A.K., Delannoy, E., Iqbal, R.K., et al. (2019) Mitochondrial Transcriptome Control and Intercompartment Cross-Talk During Plant Development. *Cells*, **8**, 583.

Nicolaides, N.C., Littman, S.J., Modrich, P., Kinzler, K.W. and Vogelstein, B. (1998) A Naturally Occurring hPMS2 Mutation Can Confer a Dominant Negative Mutator Phenotype. *Mol Cell Biol*, **18**, 1635–1641.

Nieto Moreno, N., Olthof, A.M. and Svejstrup, J.Q. (2023) Transcription-Coupled Nucleotide Excision Repair and the Transcriptional Response to UV-Induced DNA Damage. *Annu Rev Biochem*, **92**, 81–113.

Nikiforova, M.N., Stringer, J.R., Blough, R., Medvedovic, M., Fagin, J.A. and Nikiforov, Y.E. (2000) Proximity of Chromosomal Loci That Participate in Radiation-Induced Rearrangements in Human Cells. *Science (1979)*, **290**, 138–141.

Nikitchina, N., Ulashchik, E., Shmanai, V., Heckel, A.-M., Tarassov, I., Mazunin, I. and Entelis, N. (2024) Targeting of CRISPR-Cas12a crRNAs into human mitochondria. *Biochimie*, **217**, 74–85.

Nimonkar, A. V., Genschel, J., Kinoshita, E., Polaczek, P., Campbell, J.L., Wyman, C., Modrich, P. and Kowalczykowski, S.C. (2011) BLM–DNA2–RPA–MRN and EXO1–

BLM–RPA–MRN constitute two DNA end resection machineries for human DNA break repair. *Genes Dev*, **25**, 350–362.
Nischwitz, E., Schoonenberg, V.A.C., Fradera-Sola, A., Dejung, M., Vydzhak, O., Levin, M., Luke, B., Butter, F. and Scheibe, M. (2023) DNA damage repair proteins across the Tree of Life. *iScience*, **26**, 106778.
Nishizawa-Yokoi, A., Nonaka, S., Saika, H., Kwon, Y., Osakabe, K. and Toki, S. (2012) Suppression of Ku70/80 or Lig4 leads to decreased stable transformation and enhanced homologous recombination in rice. *New Phytologist*, **196**, 1048–1059.
Noji, H., Ueno, H. and McMillan, D.G.G. (2017) Catalytic robustness and torque generation of the F1-ATPase. *Biophys Rev*, **9**, 103–118.
Nosek, J., Tomaska, L., Bolotin-Fukuhara, M. and Miyakawa, I. (2006) Mitochondrial chromosome structure: an insight from analysis of complete yeast genomes. *FEMS Yeast Res*, **6**, 356–370.
Nota, F., Cambiagno, D.A., Ribone, P. and Alvarez, M.E. (2015) Expression and function of AtMBD4L, the single gene encoding the nuclear DNA glycosylase MBD4L in Arabidopsis. *Plant Science*, **235**, 122–129.
Nouspikel, T. (2009) DNA Repair in Mammalian Cells. *Cellular and Molecular Life Sciences*, **66**, 994–1009.
Odahara, M. and Sekine, Y. (2018) RECX Interacts with Mitochondrial RECA to Maintain Mitochondrial Genome Stability. *Plant Physiol*, **177**, 300–310.
Odom, O.W., Baek, K., 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. *The Plant Journal*, **53**, 842–853.
Oetke, S., Scheidig, A.J. and Krupinska, K. (2022) WHIRLY1 of Barley and Maize Share a PRAPP Motif Conferring Nucleoid Compaction. *Plant Cell Physiol*, **63**, 234–247.
Okuda, M., Nakazawa, Y., Guo, C., Ogi, T. and Nishimura, Y. (2017) Common TFIIH recruitment mechanism in global genome and transcription-coupled repair subpathways. *Nucleic Acids Res*, **45**, 13043–13055.
Oldenburg, D.J. and Bendich, A.J. (1996) Size and Structure of Replicating Mitochondrial DNA in Cultured Tobacco Cells. *Plant Cell*, **8**, 447–461.
Oliveira, M.T., Pontes, C. de B. and Ciesielski, G.L. (2020) Roles of the mitochondrial replisome in mitochondrial DNA deletion formation. *Genet Mol Biol*, **43**.
Onaka, A.T., Su, J., Katahira, Y., et al. (2020) DNA replication machinery prevents Rad52-dependent single-strand annealing that leads to gross chromosomal rearrangements at centromeres. *Commun Biol*, **3**, 202.
Ortega, J., Lee, G.S., Gu, L., Yang, W. and Li, G.-M. (2021) Mispair-bound human MutS–MutL complex triggers DNA incisions and activates mismatch repair. *Cell Res*, **31**, 542–553.
Ortega-Galisteo, A.P., Morales-Ruiz, T., Ariza, R.R. and Roldán-Arjona, T. (2008) Arabidopsis DEMETER-LIKE proteins DML2 and DML3 are required for appropriate distribution of DNA methylation marks. *Plant Mol Biol*, **67**, 671–681.
Oshima, M. and Handa, H. (2012) The identification of quantitative trait loci that control the paternal inheritance of a mitochondrial plasmid in rapeseed (Brassica napus L.). *Genes Genet. Syst.*, **87**, 19–27.
Osman, K., Higgins, J.D., Sanchez-Moran, E., Armstrong, S.J. and Franklin, F.C.H. (2011) Pathways to meiotic recombination in *Arabidopsis thaliana*. *New Phytologist*, **190**, 523–544.
Oudot-Le Secq, M.-P., Fontaine, J.-M., Rousvoal, S., Kloareg, B. and Loiseaux-de Goër, S. (2001) The Complete Sequence of a Brown Algal Mitochondrial Genome, the Ectocarpale Pylaiella littoralis (L.) Kjellm. *J Mol Evol*, **53**, 80–88.
Ouweland, J.M.W. van den, Lemkes, H.H.P.J., Ruitenbeek, W., Sandkuijl, L.A., Vijlder, M.F. de, Struyvenberg, P.A.A., Kamp, J.J.P. van de and Maassen, J.A. (1992) Mutation in mitochondrial tRNALeu(UUR) gene in a large pedigree with maternally transmitted type II diabetes mellitus and deafness. *Nat Genet*, **1**, 368–371.
Palade, G.E. (1952) The fine structure of mitochondria. *Anat Rec*, **114**, 427–451.

Palmer, J.D. and Herbon, L.A. (1988) Plant mitochondrial DNA evolved rapidly in structure, but slowly in sequence. *J Mol Evol*, **28**, 87–97.
Palmer, J.D. and Shields, C.R. (1984) Tripartite Structure of the Brassica-Campestris Mitochondrial Genome. *Nature*, **307**, 437–440.
Pang, Q., Hays, J.B. and Rajagopal, I. (1992) A plant cDNA that partially complements Escherichia coli recA mutations predicts a polypeptide not strongly homologous to RecA proteins. *Proc Natl Acad Sci U S A*, **89**, 8073–8077.
Paques, F. and Duchateau, P. (2007) Meganucleases and DNA Double-Strand Break-Induced Recombination: Perspectives for Gene Therapy. *Curr Gene Ther*, **7**, 49–66.
Parent, J.S., Lepage, E. and Brisson, N. (2011) Divergent Roles for the Two PolI-Like Organelle DNA Polymerases of Arabidopsis. *Plant Physiol*, **156**, 254–262.
Paszkiewicz, G., Gualberto, J.M., Benamar, A., Macherel, D. and Logan, D.C. (2017) Arabidopsis Seed Mitochondria Are Bioenergetically Active Immediately upon Imbibition and Specialize via Biogenesis in Preparation for Autotrophic Growth. *Plant Cell*, **29**, 109–128.
Paul-Konietzko, K., Thomale, J., Arakawa, H. and Iliakis, G. (2015) DNA Ligases I and III Support Nucleotide Excision Repair in DT40 Cells with Similar Efficiency. *Photochem Photobiol*, **91**, 1173–1180.
Pearl, S.A., Welch, M.E. and McCauley, D.E. (2009) Mitochondrial Heteroplasmy and Paternal Leakage in Natural Populations of Silene vulgaris, a Gynodioecious Plant. *Mol Biol Evol*, **26**, 537–545.
Peeva, V., Blei, D., Trombly, G., et al. (2018) Linear mitochondrial DNA is rapidly degraded by components of the replication machinery. *Nat Commun*, **9**, 1727.
Penafiel-Ayala, A., Peralta-Castro, A., Mora-Garduno, J., et al. (2023) Plant organellar MSH1 is a displacement loop specific endonuclease. *Plant Cell Physiol*.
Perkins, G.A., Renken, C.W., Song, J.Y., Frey, T.G., Young, S.J., Lamont, S., Martone, M.E., Lindsey, S. and Ellisman, M.H. (1997) Electron Tomography of Large, Multicomponent Biological Structures. *J Struct Biol*, **120**, 219–227.
Perkins, G.A., Song, J.Y., Tarsa, L., Deerinck, T.J., Ellisman, M.H. and Frey, T.G. (1998) Electron tomography of mitochondria from brown adipocytes reveals crista junctions. *J Bioenerg Biomembr*, **30**, 431–442.
Petrucco, S., Volpi, G., Bolchi, A., Rivetti, C. and Ottonello, S. (2002) A Nick-sensing DNA 3′-Repair Enzyme from Arabidopsis. *Journal of Biological Chemistry*, **277**, 23675–23683.
Picard, M., Zhang, J., Hancock, S., et al. (2014) Progressive increase in mtDNA 3243A>G heteroplasmy causes abrupt transcriptional reprogramming. *Proceedings of the National Academy of Sciences*, **111**.
Plaza-G.A., I., Lemishko, K.M., Crespo, R., Truong, T.Q., Kaguni, L.S., Cao-García, F.J., Ciesielski, G.L. and Ibarra, B. (2023) Mechanism of strand displacement DNA synthesis by the coordinated activities of human mitochondrial DNA polymerase and SSB. *Nucleic Acids Res*, **51**, 1750–1765.
Pons, R., Andreu, A.L., Checcarelli, N., et al. (2004) Mitochondrial DNA abnormalities and autistic spectrum disorders. *J Pediatr*, **144**, 81–85.
Posakony, J., England, J. and Attardi, G. (1977) Mitochondrial growth and division during the cell cycle in HeLa cells. *J Cell Biol*, **74**, 468–491.
Pozo, J.C. del and Ramirez-Parra, E. (2015) Whole genome duplications in plants: an overview from *Arabidopsis*. *J Exp Bot*, **66**, 6991–7003.
Prakash, A. and Doublié, S. (2015) Base Excision Repair in the Mitochondria. *J Cell Biochem*, **116**, 1490–1499.
Prakash, S., Johnson, R.E. and Prakash, L. (2005) Eukaryotic translesion synthesis DNA polymerases: Specificity of Structure and Function. *Annu Rev Biochem*, **74**, 317–353.
Pramateftaki, P. V., Kouvelis, V.N., Lanaridis, P. and Typas, M.A. (2006) The mitochondrial genome of the wine yeast *Hanseniaspora uvarum* : a unique genome organization among yeast/fungal counterparts. *FEMS Yeast Res*, **6**, 77–90.

Prasad, R., Longley, M.J., Sharief, F.S., Hou, E.W., Copeland, W.C. and Wilson, S.H. (2009) Human DNA polymerase possesses 5'-dRP lyase activity and functions in single-nucleotide base excision repair in vitro. *Nucleic Acids Res*, **37**, 1868–1877.

Preuten, T., Cincu, E., Fuchs, J., Zoschke, R., Liere, K. and Börner, T. (2010) Fewer genes than organelles: extremely low and variable gene copy numbers in mitochondria of somatic plant cells. *The Plant Journal*, **64**, 948–959.

Puchta, H. (2004) The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. *J Exp Bot*.

Putintseva, Y.A., Bondar, E.I., Simonov, E.P., et al. (2020) Siberian larch (Larix sibirica Ledeb.) mitochondrial genome assembled using both short and long nucleotide sequence reads is currently the largest known mitogenome. *BMC Genomics*, **21**, 654.

Qian, J., Li, M., Zheng, M. and Hsu, Y.-F. (2021) *Arabidopsis* SSB1, a Mitochondrial Single-Stranded DNA-Binding Protein, is Involved in ABA Response and Mitochondrial RNA Splicing. *Plant Cell Physiol*, **62**, 1321–1334.

Qian, J., Zheng, M., Wang, L., Song, Y., Yan, J. and Hsu, Y.F. (2022) Arabidopsis mitochondrial single-stranded DNA-binding proteins SSB1 and SSB2 are essential regulators of mtDNA replication and homologous recombination. *J Integr Plant Biol*, **64**, 1952–1965.

Quiroz, D., Lensink, M., Kliebenstein, D.J. and Monroe, J.G. (2023) Causes of Mutation Rate Variability in Plant Genomes. *Annu Rev Plant Biol*, **74**, 751–775.

Ragunathan, K., Liu, C. and Ha, T. (2012) RecA filament sliding on DNA facilitates homology search. *Elife*, **1**.

Ramirez, C.L., Foley, J.E., Wright, D.A., et al. (2008) Unexpected failure rates for modular assembly of engineered zinc fingers. *Nat Methods*, **5**, 374–375.

Ramírez-Gamboa, D., Díaz-Zamorano, A.L., Meléndez-Sánchez, E.R., et al. (2022) Photolyase Production and Current Applications: A Review. *Molecules*, **27**, 5998.

Rampersad, S. and Tennant, P. (2018) Replication and Expression Strategies of Viruses. In *Viruses*. Elsevier, pp. 55–82.

Randolph-Anderson, B.L., Boynton, J.E., Gillham, N.W., Harris, E.H., Johnson, A.M., Dorthu, M.P. and Matagne, R.F. (1993) Further characterization of the respiratory deficient dum-1 mutation of Chlamydomonas reinhardtii and its use as a recipient for mitochondrial transformation. *Mol. Gen. Genet.*, **236**, 235–244.

Ranjha, L., Howard, S.M. and Cejka, P. (2018) Main steps in DNA double-strand break repair: an introduction to homologous recombination and related processes. *Chromosoma*, **127**, 187–214.

Rao, Y., Yang, X., Pan, C., Wang, C. and Wang, K. (2022) Advance of Clustered Regularly Interspaced Short Palindromic Repeats-Cas9 System and Its Application in Crop Improvement. *Front Plant Sci*, **13**.

Rasmussen, A.K. and Rasmussen, L.J. (2005) Targeting of O6-MeG DNA methyltransferase (MGMT) to mitochondria protects against alkylation induced cell death. *Mitochondrion*, **5**, 411–417.

Rass, U., Compton, S.A., Matos, J., Singleton, M.R., Ip, S.C.Y., Blanco, M.G., Griffith, J.D. and West, S.C. (2010) Mechanism of Holliday junction resolution by the human GEN1 protein. *Genes Dev*, **24**, 1559–1569.

Ratanaphan, A. (2012) A DNA Repair BRCA1 Estrogen Receptor and Targeted Therapy in Breast Cancer. *Int J Mol Sci*, **13**, 14898–14916.

Reddemann, A. and Horn, R. (2018) Recombination Events Involving the atp9 Gene Are Associated with Male Sterility of CMS PET2 in Sunflower. *Int J Mol Sci*, **19**, 806.

Remacle, C., Cardol, P., Coosemans, N., Gaisne, M. and Bonnefoy, N. (2006) High-efficiency biolistic transformation of Chlamydomonas mitochondria can be used to insert mutations in complex I genes. *Proc. Natl. Acad. Sci. U. S. A.*, **103**, 4771–4776.

Reyes, A., Kazak, L., Wood, S.R., Yasukawa, T., Jacobs, H.T. and Holt, I.J. (2013) Mitochondrial DNA replication proceeds via a 'bootlace' mechanism involving the incorporation of processed transcripts. *Nucleic Acids Res*, **41**, 5837–5850.

Ribarska, T., Bjørnstad, P.M., Sundaram, A.Y.M. and Gilfillan, G.D. (2022) Optimization of enzymatic fragmentation is crucial to maximize genome coverage: a comparison of library preparation methods for Illumina sequencing. *BMC Genomics*, **23**, 92.

Ribeiro-Silva, C., Sabatella, M., Helfricht, A., Marteijn, J.A., Theil, A.F., Vermeulen, W. and Lans, H. (2020) Ubiquitin and TFIIH-stimulated DDB2 dissociation drives DNA damage handover in nucleotide excision repair. *Nat Commun*, **11**, 4868.

Richardson, A.O. and Palmer, J.D. (2006) Horizontal gene transfer in plants. *J Exp Bot*, **58**, 1–9.

Robberson, D.L., Kasamatsu, H. and Vinograd, J. (1972) Replication of Mitochondrial DNA. Circular Replicative Intermediates in Mouse L Cells. *Proceedings of the National Academy of Sciences*, **69**, 737–741.

Robison, M.M., Ling, X., Smid, M.P.L., Zarei, A. and Wolyn, D.J. (2009) Antisense Expression of Mitochondrial ATP Synthase Subunits OSCP (ATP5) and γ (ATP3) Alters Leaf Morphology, Metabolism and Gene Expression in Arabidopsis. *Plant Cell Physiol*, **50**, 1840–1850.

Rocha, C.R.R., Silva, M.M., Quinet, A., Cabral-Neto, J.B. and Menck, C.F.M. (2018) DNA repair pathways and cisplatin resistance: an intimate relationship. *Clinics*, **73**, e478s.

Roger, A.J., Muñoz-Gómez, S.A. and Kamikawa, R. (2017) The Origin and Diversification of Mitochondria. *Current Biology*, **27**, R1177–R1192.

Rokas, A., Ladoukakis, E. and Zouros, E. (2003) Animal mitochondrial DNA recombination revisited. *Trends Ecol Evol*, **18**, 411–417.

Roldán-Arjona, T., Ariza, R.R. and Córdoba-Cañero, D. (2019) DNA Base Excision Repair in Plants: An Unfolding Story With Familiar and Novel Characters. *Front Plant Sci*, **10**.

Ronceret, A., Guilleminot, J., Lincker, F., et al. (2005) Genetic analysis of two Arabidopsis DNA polymerase epsilon subunits during early embryogenesis. *The Plant Journal*, **44**, 223–236.

Ropars, V., Drevet, P., Legrand, P., et al. (2011) Structural characterization of filaments formed by human Xrcc4–Cernunnos/XLF complex involved in nonhomologous DNA end-joining. *Proceedings of the National Academy of Sciences*, **108**, 12663–12668.

Roth, N., Klimesch, J., Dukowic-Schulze, S., Pacher, M., Mannuss, A. and Puchta, H. (2012) The requirement for recombination factors differs considerably between different pathways of homologous double-strand break repair in somatic plant cells. *The Plant Journal*, **72**, 781–790.

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.

Roy, R., Kozlov, A.G., Lohman, T.M. and Ha, T. (2009) SSB protein diffusion on single-stranded DNA stimulates RecA filament formation. *Nature*, **461**, 1092–1097.

Roy, S., Choudhury, S.R., Sengupta, D.N. and Das, K.P. (2013) Involvement of AtPolλ in the repair of high salt- and DNA cross-linking agent-induced double strand breaks in Arabidopsis. *Plant Physiol*, **162**, 1195–1210.

Rudolph, C.J., Mahdi, A.A., Upton, A.L. and Lloyd, R.G. (2010) RecG protein and single-strand DNA exonucleases avoid cell lethality associated with PriA helicase activity in Escherichia coli. *Genetics*, **186**, 473–492.

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

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

Rupert, C.S., Goodgal, S.H. and Herriott, R.M. (1958) Photoreactivation in vitro of ultraviolet-inactivated hemophilus influenzae transforming factor. *J Gen Physiol*, **41**, 451–471.

Rydberg, B. and Lindahl, T. (1982) Nonenzymatic methylation of DNA by the intracellular methyl group donor S-adenosyl-L-methionine is a potentially mutagenic reaction. *EMBO J*, **1**, 211–216.

Sabel, M.S. (2009) Principles of Adjuvant Chemotherapy for Breast Cancer. In *Essentials of Breast Surgery*. Elsevier, pp. 247–266.

Sagan, L. (1967) On the origin of mitosing cells. *J Theor Biol*, **14**, 255–274.
Saika, H., Nishizawa-Yokoi, A. and Toki, S. (2014) The non-homologous end-joining pathway is involved in stable transformation in rice. *Front Plant Sci*, **5**.
Sakamoto, A.N. (2019) Translesion Synthesis in Plants: Ultraviolet Resistance and Beyond. *Front Plant Sci*, **10**.
Sakofsky, C.J. and Malkova, A. (2017) Break induced replication in eukaryotes: mechanisms, functions, and consequences. *Crit Rev Biochem Mol Biol*, **52**, 395–413.
Samach, A., Gurevich, V., Avivi-Ragolsky, N. and Levy, A.A. (2018) The effects of AtRad52 over-expression on homologous recombination in Arabidopsis. *Plant J*, **95**, 30–40.
Samach, A., Melamed-Bessudo, C., Avivi-Ragolski, N., Pietrokovski, S. and Levy, A.A. (2011) Identification of plant RAD52 homologs and characterization of the Arabidopsis thaliana RAD52-like genes. *Plant Cell*, **23**, 4266–4279.
Šamanić, I., Cvitanić, R., Simunić, J. and Puizina, J. (2016) Arabidopsis thaliana MRE11 is essential for activation of cell cycle arrest, transcriptional regulation and DNA repair upon the induction of double-stranded DNA breaks. *Plant Biol*, **18**, 681–694.
Sancar, A. and Rupp, W.D. (1983) A novel repair enzyme: UVRABC excision nuclease of Escherichia coli cuts a DNA strand on both sides of the damaged region. *Cell*, **33**, 249–260.
Sander, J.D., Dahlborg, E.J., Goodwin, M.J., et al. (2011) Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). *Nat Methods*, **8**, 67–69.
Sando, N., Miyakawa, I., Nishibayashi, S. and Kuroiwa, T. (1981) Arrangement of mitochondrial nucleoids during life cycle of Saccharomyces cerevisiae. *J Gen Appl Microbiol*, **27**, 511–516.
Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*, **74**, 5463–5467.
Santerre, A. and Britt, A.B. (1994) Cloning of a 3-methyladenine-DNA glycosylase from Arabidopsis thaliana. *Proceedings of the National Academy of Sciences*, **91**, 2240–2244.
Sarasin, A. (1999) The molecular pathways of ultraviolet-induced carcinogenesis. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, **428**, 5–10.
Satoh, M., Kubo, T., Nishizawa, S., Estiati, A., Itchoda, N. and Mikami, T. (2004) The cytoplasmic male-sterile type and normal type mitochondrial genomes of sugar beet share the same complement of genes of known function but differ in the content of expressed ORFs. *Molecular Genetics and Genomics*, **272**, 247–256.
Satoh, M. and Kuroiwa, T. (1991) Organization of multiple nucleoids and DNA molecules in mitochondria of a human cell. *Exp Cell Res*, **196**, 137–140.
Schatz-Daas, D. (2021) *OEX1, une exonucléase/flap-endonucléase nécessaire à la stabilité du génome mitochondrial d'Arabidopsis thaliana*. University of Strasbourg.
Schnable, P. (1998) The molecular basis of cytoplasmic male sterility and fertility restoration. *Trends Plant Sci*, **3**, 175–180.
Schofield, M.J. and Hsieh, P. (2003) DNA mismatch repair: molecular mechanisms and biological function. *Annu. Rev. Microbiol.*, **57**, 579–608.
Schormann, N., Ricciardi, R. and Chattopadhyay, D. (2014) Uracil-DNA glycosylases—Structural and functional perspectives on an essential family of DNA repair enzymes. *Protein Science*, **23**, 1667–1685.
Schulz, I., Mahler, H.-C., Boiteux, S. and Epe, B. (2000) Oxidative DNA base damage induced by singlet oxygen and photosensitization: recognition by repair endonucleases and mutagenicity. *Mutation Research/DNA Repair*, **461**, 145–156.
Seki, M. (2003) POLQ (Pol), a DNA polymerase and DNA-dependent ATPase in human cells. *Nucleic Acids Res*, **31**, 6117–6126.
Selinski, J., Scheibe, R., Day, D.A. and Whelan, J. (2018) Alternative Oxidase Is Positive for Plant Performance. *Trends Plant Sci*, **23**, 588–597.
Sheahan, M.B., Rose, R.J. and McCurdy, D.W. (2004) Organelle inheritance in plant cell division: the actin cytoskeleton is required for unbiased inheritance of chloroplasts, mitochondria and endoplasmic reticulum in dividing protoplasts. *The Plant Journal*, **37**, 379–390.

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.

Shinohara, A., Shinohara, M., Ohta, T., Matsuda, S. and Ogawa, T. (1998) Rad52 forms ring structures and co-operates with RPA in single-strand DNA annealing. *Genes to Cells*, **3**, 145–156.

Shuker, D.E. and Farmer, P.B. (1992) Relevance of urinary DNA adducts as markers of carcinogen exposure. *Chem Res Toxicol*, **5**, 450–460.

Shukla, V., Halabelian, L., Balagere, S., Samaniego-Castruita, D., Feldman, D.E., Arrowsmith, C.H., Rao, A. and Aravind, L. (2020) HMCES Functions in the Alternative End-Joining Pathway of the DNA DSB Repair during Class Switch Recombination in B Cells. *Mol Cell*, **77**, 384-394.e4.

Shukla, V.K., Doyon, Y., Miller, J.C., et al. (2009) Precise genome modification in the crop species Zea mays using zinc-finger nucleases. *Nature*, **459**, 437–441.

Shutt, T.E. and Gray, M.W. (2006) Bacteriophage origins of mitochondrial replication and transcription proteins. *Trends Genet*, **22**, 90–95.

Siegel, E.C. and Bryson, V. (1967) Mutator gene of Escherichia coli B. J. Bacteriol. *J Bacteriol*, **94**, 38–47.

Siekevitz, P. (1957) Powerhouse of the Cell. *Sci Am*, **197**, 131–144.

Silva, I.U. De, McHugh, P.J., Clingen, P.H. and Hartley, J.A. (2000) Defining the Roles of Nucleotide Excision Repair and Recombination in the Repair of DNA Interstrand Cross-Links in Mammalian Cells. *Mol Cell Biol*, **20**, 7980–7990.

Silva-Pinheiro, P., Nash, P.A., Haute, L. Van, Mutti, C.D., Turner, K. and Minczuk, M. (2022) In vivo mitochondrial base editing via adeno-associated viral delivery to mouse post-mitotic tissue. *Nat Commun*, **13**, 750.

Simsek, D., Brunet, E., Wong, S.Y.-W., et al. (2011) DNA Ligase III Promotes Alternative Nonhomologous End-Joining during Chromosomal Translocation Formation. *PLoS Genet*, **7**, e1002080.

Singh, M. and Brown, G.G. (1991) Suppression of cytoplasmic male sterility by nuclear genes alters expression of a novel mitochondrial gene region. *Plant Cell*, **3**, 1349–1362.

Sinn, B.T. and Barrett, C.F. (2020) Ancient Mitochondrial Gene Transfer between Fungi and the Orchids. *Mol Biol Evol*, **37**, 44–57.

Sjöstrand, F.S. (1956) The Ultrastructure of Cells as Revealed by the Electron Microscope. In pp. 455–533.

Sklar, L.R., Almutawa, F., Lim, H.W. and Hamzavi, I. (2012) Effects of ultraviolet radiation, visible light, and infrared radiation on erythema and pigmentation: a review. *Photochemical & Photobiological Sciences*, **12**, 54–64.

Sloan, D.B. (2013) One ring to rule them all? Genome sequencing provides new insights into the “master circle” model of plant mitochondrial DNA structure. *New Phytol.*, **200**, 978–985.

Sloan, D.B., Alverson, A.J., Chuckalovcak, J.P., Wu, M., McCauley, D.E., Palmer, J.D. and Taylor, D.R. (2012) Rapid Evolution of Enormous, Multichromosomal Genomes in Flowering Plant Mitochondria with Exceptionally High Mutation Rates. *PLoS Biol*, **10**, e1001241.

Small, G.D. (1987) Rapair systems for nuclear and chloroplast DNA in Chlamydomonas reinhardtii. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, **181**, 31–35.

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

Smith, C.E., Bowen, N., Graham, W.J., Goellner, E.M., Srivatsan, A. and Kolodner, R.D. (2015) Activation of Saccharomyces cerevisiae Mlh1-Pms1 Endonuclease in a Reconstituted Mismatch Repair System. *Journal of Biological Chemistry*, **290**, 21580–21590.

Smith, D.R., Hua, J. and Lee, R.W. (2010) Evolution of linear mitochondrial DNA in three known lineages of Polytomella. *Curr Genet*, **56**, 427–438.

Smith, D.R., Jackson, C.J. and Reyes-Prieto, A. (2014) Nucleotide substitution analyses of the glaucophyte Cyanophora suggest an ancestrally lower mutation rate in plastid vs mitochondrial DNA for the Archaeplastida. *Mol Phylogenet Evol*, **79**, 380–384.

Smith, G.R. (1988) Homologous recombination in procaryotes. *Microbiol Rev*, **52**, 1–28.

Sodmergen, Zhang, Q., Zhang, Y., Sakamoto, W. and Kuroiwa, T. (2002) Reduction in amounts of mitochondrial DNA in the sperm cells as a mechanism for maternal inheritance in Hordeum vulgare. *Planta*, **216**, 235–244.

Sousa, J.S., D'Imprima, E. and Vonck, J. (2018) Mitochondrial Respiratory Chain Complexes. In pp. 167–227.

Spagnolo, L., Rivera-Calzada, A., Pearl, L.H. and Llorca, O. (2006) Three-Dimensional Structure of the Human DNA-PKcs/Ku70/Ku80 Complex Assembled on DNA and Its Implications for DNA DSB Repair. *Mol Cell*, **22**, 511–519.

Spampinato, C.P., Gomez, R.L., Galles, C. and Lario, L.D. (2009) From bacteria to plants: a compendium of mismatch repair assays. *Mutat. Res.*, **682**, 110–128.

Srivastava, D.K., Berg, B.J. Vande, Prasad, R., Molina, J.T., Beard, W.A., Tomkinson, A.E. and Wilson, S.H. (1998) Mammalian Abasic Site Base Excision Repair. *Journal of Biological Chemistry*, **273**, 21203–21209.

Staresincic, L., Fagbemi, A.F., Enzlin, J.H., et al. (2009) Coordination of dual incision and repair synthesis in human nucleotide excision repair. *EMBO J*, **28**, 1111–1120.

Stark, J.M., Pierce, A.J., Oh, J., Pastink, A. and Jasin, M. (2004) Genetic Steps of Mammalian Homologous Repair with Distinct Mutagenic Consequences. *Mol Cell Biol*, **24**, 9305–9316.

Steffen, M., Sarkela, T.M., Gybina, A.A., Steele, T.W., Trasseth, N.J., Kuehl, D. and Giulivi, C. (2001) Metabolism of S-nitrosoglutathione in intact mitochondria. *Biochemical Journal*, **356**, 395.

Stenton, S.L. and Prokisch, H. (2020) Genetics of mitochondrial diseases: Identifying mutations to help diagnosis. *EBioMedicine*, **56**, 102784.

Stock, D., Leslie, A.G.W. and Walker, J.E. (1999) Molecular Architecture of the Rotary Motor in ATP Synthase. *Science (1979)*, **286**, 1700–1705.

Stohl, E.A., Brockman, J.P., Burkle, K.L., Morimatsu, K., Kowalczykowski, S.C. and Seifert, H.S. (2003) Escherichia coli RecX Inhibits RecA Recombinase and Coprotease Activities in Vitro and in Vivo. *Journal of Biological Chemistry*, **278**, 2278–2285.

Stracke, R., Favory, J.-J., Gruber, H., Bartelniewoehner, L., Bartels, S., Binkert, M., Funk, M., Wwisshaar, B. and Ulm, R. (2010) The Arabidopsis bZIP transcription factor HY5 regulates expression of the PFG1/MYB12 gene in response to light and ultraviolet-B radiation. *Plant Cell Environ*.

Strzalka, W., Oyama, T., Tori, K. and Morikawa, K. (2009) Crystal structures of the *Arabidopsis thaliana* proliferating cell nuclear antigen 1 and 2 proteins complexed with the human p21 C-terminal segment. *Protein Science*, **18**, 1072–1080.

Stucki, M., Pascucci, B., Parlanti, E., Fortini, P., Wilson, S.H., Hübscher, U. and Dogliotti, E. (1998) Mammalian base excision repair by DNA polymerases δ and ε. *Oncogene*, **17**, 835–843.

Su, C.J., Chavoya, A., Dallo, S.F. and Baseman, J.B. (1990) Sequence divergency of the cytadhesin gene of Mycoplasma pneumoniae. *Infect Immun*, **58**, 2669–2674.

Su, S.S. and Modrich, P. (1986) Escherichia coli mutS-encoded protein binds to mismatched DNA base pairs. *Proceedings of the National Academy of Sciences*, **83**, 5057–5061.

Sugahara, M., Mikawa, T., Kumasaka, T., Yamamoto, M., Kato, R., Fukuyama, K., Inoue, Y. and Kuramitsu, S. (2000) Crystal structure of a repair enzyme of oxidatively damaged DNA, MutM (Fpg), from an extreme thermophile, Thermus thermophilus HB8. *EMBO J*, **19**, 3857–3869.

Suh, D., Wilson, D.M. and Povirk, L.F. (1997) 3'-Phosphodiesterase activity of human apurinic/apyrimidinic endonuclease at DNA double-strand break ends. *Nucleic Acids Res*, **25**, 2495–2500.

Sukhorukov, V.M., Dikov, D., Reichert, A.S. and Meyer-Hermann, M. (2012) Emergence of the Mitochondrial Reticulum from Fission and Fusion Dynamics. *PLoS Comput Biol*, **8**, e1002745.

Sunderhaus, S., Dudkina, N. V., Jänsch, L., Klodmann, J., Heinemeyer, J., Perales, M., Zabaleta, E., Boekema, E.J. and Braun, H.-P. (2006) Carbonic Anhydrase Subunits Form a Matrix-exposed Domain Attached to the Membrane Arm of Mitochondrial Complex I in Plants. *Journal of Biological Chemistry*, **281**, 6482–6488.

Sung, P. (1997) Function of Yeast Rad52 Protein as a Mediator between Replication Protein A and the Rad51 Recombinase. *Journal of Biological Chemistry*, **272**, 28194–28197.

Sung, P. and Klein, H. (2006) Mechanism of homologous recombination: mediators and helicases take on regulatory functions. *Nat Rev Mol Cell Biol*, **7**, 739–750.

Swagemakers, S.M.A., Essers, J., Wit, J. de, Hoeijmakers, J.H.J. and Kanaar, R. (1998) The Human Rad54 Recombinational DNA Repair Protein Is a Double-stranded DNA-dependent ATPase. *Journal of Biological Chemistry*, **273**, 28292–28297.

Symington, L.S. (2002) Role of *RAD52* Epistasis Group Genes in Homologous Recombination and Double-Strand Break Repair. *Microbiology and Molecular Biology Reviews*, **66**, 630–670.

Szklarczyk, M., Szymański, M., Wójcik-Jagła, M., Simon, P.W., Weihe, A. and Börner, T. (2014) Mitochondrial atp9 genes from petaloid male-sterile and male-fertile carrots differ in their status of heteroplasmy, recombination involvement, post-transcriptional processing as well as accumulation of RNA and protein product. *Theoretical and Applied Genetics*, **127**, 1689–1701.

Tadi, S.K., Sebastian, R., Dahal, S., Babu, R.K., Choudhary, B. and Raghavan, S.C. (2016) Microhomology-mediated end joining is the principal mediator of double-strand break repair during mitochondrial DNA lesions. *Mol Biol Cell*, **27**, 223–235.

Taipakova, S., Kuanbay, A., Saint-Pierre, C., Gasparutto, D., Baiken, Y., Groisman, R., Ishchenko, A.A., Saparbaev, M. and Bissenbaev, A.K. (2020) The Arabidopsis thaliana Poly(ADP-Ribose) Polymerases 1 and 2 Modify DNA by ADP-Ribosylating Terminal Phosphate Residues. *Front Cell Dev Biol*, **8**.

Takahashi, M., Teranishi, M., Ishida, H., Kawasaki, J., Takeuchi, A., Yamaya, T., Watanabe, M., Makino, A. and Hidema, J. (2011a) Cyclobutane pyrimidine dimer (CPD) photolyase repairs ultraviolet-B-induced CPDs in rice chloroplast and mitochondrial DNA. *The Plant Journal*, **66**, 433–442.

Takahashi, M., Teranishi, M., Ishida, H., Kawasaki, J., Takeuchi, A., Yamaya, T., Watanabe, M., Makino, A. and Hidema, J. (2011b) Cyclobutane pyrimidine dimer (CPD) photolyase repairs ultraviolet-B-induced CPDs in rice chloroplast and mitochondrial DNA. *The Plant Journal*, **66**, 433–442.

Tan, B.H., Bencsath, F.A. and Gaubatz, J.W. (1990) Steady-state levels of 7-methylguanine increase in nuclear DNA of postmitotic mouse tissues during aging. *Mutation Research/DNAging*, **237**, 229–238.

Tanaka, Y., Tsuda, M., Yasumoto, K., Yamagishi, H. and Terachi, T. (2012) A complete mitochondrial genome sequence of Ogura-type male-sterile cytoplasm and its comparative analysis with that of normal cytoplasm in radish (Raphanus sativus L.). *BMC Genomics*, **13**, 352.

Tantin, D., Kansal, A. and Carey, M. (1997) Recruitment of the Putative Transcription-Repair Coupling Factor CSB/ERCC6 to RNA Polymerase II Elongation Complexes. *Mol Cell Biol*, **17**, 6803–6814.

The Arabidopsis Genome, I. (2000) Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature*, **408**, 796–815.

Thomas, A., Cox, J., Wolfe, K.B., Mingalone, C.H., Yaspan, H.R. and McVey, M. (2022) Division of Labor by the HELQ, BLM, and FANCM Helicases during Homologous Recombination Repair in Drosophila melanogaster. *Genes (Basel)*, **13**, 474.

Thomas, R. and Roy, D. (2001) Stilbene estrogen produces higher levels of mitochondrial DNA adducts than nuclear DNA adducts in the target organ of cancer (liver) of male Sprague Dawley rats. *Oncol Rep*.

Thompson, L.H. and West, M.G. (2000) XRCC1 keeps DNA from getting stranded. *Mutation Research/DNA Repair*, **459**, 1–18.

Thorburn, D.R., Rahman, J. and Rahman, S. (1993) *Mitochondrial DNA-Associated Leigh Syndrome and NARP*,.

Tiraby, J.-G. and Fox, M.S. (1973) Marker Discrimination in Transformation and Mutation of Pneumococcus. *Proceedings of the National Academy of Sciences*, **70**, 3541–3545.

Todd, P.A., Brouwer, J. and Glickman, B.W. (1981) Influence of DNA-repair deficiencies on MMS- and EMS-induced mutagenesis in Escherichia coli K-12. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, **82**, 239–250.

Trasviña-Arenas, C.H., Baruch-Torres, N., Cordoba-Andrade, F.J., Ayala-García, V.M., García-Medel, P.L., Díaz-Quezada, C., Peralta-Castro, A., Ordaz-Ortiz, J.J. and Brieba, L.G. (2018) Identification of a unique insertion in plant organellar DNA polymerases responsible for 5′-dRP lyase and strand-displacement activities: Implications for Base Excision Repair. *DNA Repair (Amst)*, **65**, 1–10.

Treffers, H.P., Spinelli, V. and Belser, N.O. (1954) A factor (or mutator gene) influencing mutation rates in Escherichia coli. *Proceedings of the National Academy of Sciences*, **40**, 1064–1071.

Tremblay-Belzile, S., Lepage, É., Zampini, É. and Brisson, N. (2015) Short-range inversions: Rethinking organelle genome stability. *BioEssays*, **37**, 1086–1094.

Turrens, J.F. (2003) Mitochondrial formation of reactive oxygen species. *J Physiol*, **552**, 335–344.

Uchiyama, Y., Kimura, S., Yamamoto, T., Ishibashi, T. and Sakaguchi, K. (2004) Plant DNA polymerase λ, a DNA repair enzyme that functions in plant meristematic and meiotic tissues. *Eur J Biochem*, **271**, 2799–2807.

Ui, A., Seki, M., Ogiwara, H., Lai, M.S., Yamamoto, K., Tada, S. and Enomoto, T. (2007) Activation of a novel pathway involving Mms1 and Rad59 in sgs1 cells. *Biochem Biophys Res Commun*, **356**, 1031–1037.

Umezu, K., Nakayama, K. and Nakayama, H. (1990) Escherichia coli RecQ protein is a DNA helicase. *Proceedings of the National Academy of Sciences*, **87**, 5363–5367.

Urnov, F.D., Rebar, E.J., Holmes, M.C., Zhang, H.S. and Gregory, P.D. (2010) Genome editing with engineered zinc finger nucleases. *Nat Rev Genet*, **11**, 636–646.

Uyttewaal, M., Arnal, N., Quadrado, M., et al. (2009) Characterization of Raphanus sativus Pentatricopeptide Repeat Proteins Encoded by the Fertility Restorer Locus for Ogura Cytoplasmic Male Sterility. *Plant Cell*, **20**, 3331–3345.

Vaca, C.E., Wilhelm, J. and Harms-Ringdahl, M. (1988) Interaction of lipid peroxidation products with DNA. A review. *Mutation Research/Reviews in Genetic Toxicology*, **195**, 137–149.

Vanlerberghe, G.C., Dahal, K., Alber, N.A. and Chadee, A. (2020) Photosynthesis, respiration and growth: A carbon and energy balancing act for alternative oxidase. *Mitochondrion*, **52**, 197–211.

Vanlerberghe, G.C., Martyn, G.D. and Dahal, K. (2016) Alternative oxidase: a respiratory electron transport chain pathway essential for maintaining photosynthetic performance during drought stress. *Physiol Plant*, **157**, 322–337.

Veaute, X., Delmas, S., Selva, M., Jeusset, J., Cam, E. Le, Matic, I., Fabre, F. and Petit, M.-A. (2005) UvrD helicase, unlike Rep helicase, dismantles RecA nucleoprotein filaments in Escherichia coli. *EMBO J*, **24**, 180–189.

Vedel, F., Paepe, R., Chétrit, P., Gutierres, S., Vitart, V. and Pla, M. (1994) Molecular basis of nuclear and cytoplasmic male sterility in higher plants. *Plant Physiology and Biochemistry* , **32**, 601–618.

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

Virdi, K.S., Laurie, J.D., Xu, Y.Z., et al. (2015) Arabidopsis MSH1 mutation alters the epigenome and produces heritable changes in plant growth. *Nat. Commun.*, **6**, 6386.

Visser, W., Spronsen, E.A. van, Nanninga, N., Pronk, J.T., Kuenen, J.G. and Dijken, J.P. van (1995) Effects of growth conditions on mitochondrial morphology inSaccharomyces cerevisiae. *Antonie Van Leeuwenhoek*, **67**, 243–253.

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.

Wallace, D.C., Singh, G., Lott, M.T., Hodge, J.A., Schurr, T.G., Lezza, A.M.S., Elsas, L.J. and Nikoskelainen, E.K. (1988) Mitochondrial DNA Mutation Associated with Leber's Hereditary Optic Neuropathy. *Science (1979)*, **242**, 1427–1430.

Wallet, C., Ret, M. Le, Bergdoll, M., Bichara, M., Dietrich, A. and Gualberto, J.M. (2015) The RECG1 DNA Translocase Is a Key Factor in Recombination Surveillance, Repair, and Segregation of the Mitochondrial DNA in Arabidopsis. *Plant Cell*, tpc.15.00680.

Wanamaker, S.A., Garza, R.M., MacWilliams, A., et al. (2017) CrY2H-seq: a massively multiplexed assay for deep-coverage interactome mapping. *Nat Methods*, **14**, 819–825.

Wang, J., Cao, M.J., Pan, G.T., Lu, Y.L. and Rong, T.Z. (2009) RNA editing of mitochondrial functional genes atp6 and cox2 in maize (Zea mays L.). *Mitochondrion*, **9**, 364–369.

Wang, J.C. (2002) Cellular roles of DNA topoisomerases: a molecular perspective. *Nat Rev Mol Cell Biol*, **3**, 430–440.

Wang, K., Gao, F., Ji, Y., Liu, Y., Dan, Z., Yang, P., Zhu, Y. and Li, S. (2013) ORFH79 impairs mitochondrial function via interaction with a subunit of electron transport chain complex III in honglian cytoplasmic male sterile rice. *New Phytologist*, **198**, 408–418.

Wang, M., Wu, Weizhong, Wu, Wenqi, Rosidi, B., Zhang, L., Wang, H. and Iliakis, G. (2006) PARP-1 and Ku compete for repair of DNA double strand breaks by distinct NHEJ pathways. *Nucleic Acids Res*, **34**, 6170–6182.

Wang, S., Zhang, G., Zhang, Y., et al. (2015) Comparative studies of mitochondrial proteomics reveal an intimate protein network of male sterility in wheat (Triticum aestivum L.). *J Exp Bot*, **66**, 6191–6203.

Wang, W., Lin, L., Zhang, Q., Yang, J., Kamili, E., Chu, J., Li, X., Yang, S. and Xu, Y. (2023) Heteroplasmy and Individual Mitogene Pools: Characteristics and Potential Roles in Ecological Studies. *Biology (Basel)*, **12**, 1452.

Wang, Y., Chu, P., Yang, Q., Chang, S., Chen, J., Hu, M. and Guan, R. (2014) Complete Mitochondrial Genome of Eruca sativa Mill. (Garden Rocket). *PLoS One*, **9**, e105748.

Wang, Z., Zou, Y., Li, X., et al. (2006) Cytoplasmic Male Sterility of Rice with Boro II Cytoplasm Is Caused by a Cytotoxic Peptide and Is Restored by Two Related PPR Motif Genes via Distinct Modes of mRNA Silencing. *Plant Cell*, **18**, 676–687.

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

Warren, J.M., Simmons, M.P., Wu, Z. and Sloan, D.B. (2016) Linear Plasmids and the Rate of Sequence Evolution in Plant Mitochondrial Genomes. *Genome Biol Evol*, **8**, 364–374.

Warren, J.M. and Sloan, D.B. (2020) Interchangeable parts: The evolutionarily dynamic tRNA population in plant mitochondria. *Mitochondrion*, **52**, 144–156.

Waterworth, W.M., Altun, C., Armstrong, S.J., Roberts, N., Dean, P.J., Young, K., Weil, C.F., Bray, C.M. and West, C.E. (2007) NBS1 is involved in DNA repair and plays a synergistic role with ATM in mediating meiotic homologous recombination in plants. *The Plant Journal*, **52**, 41–52.

Weber-Lotfi, F., Fertet, A., Kubilinskas, R., Wallet, C. and Gualberto, J.M. (2023) Mitochondrial DNA Isolation from Plants. In *Methods in Molecular Biology*. Humana Press Inc., pp. 57–75.

Weeden, N.F. (1981) Genetic and biochemical implications of the endosymbiotic origin of the chloroplast. *J Mol Evol*, **17**, 133–139.

Weegen, Y. van der, Golan-Berman, H., Mevissen, T.E.T., et al. (2020) The cooperative action of CSB, CSA, and UVSSA target TFIIH to DNA damage-stalled RNA polymerase II. *Nat Commun*, **11**, 2104.

Wei, J., Shouping, Y., Deyue, Y. and Junyi, G. (2011) A comparative study of ATPase subunit 9 (Atp9) gene between cytoplasmic male sterile line and its maintainer line in soybeans. *Afr J Biotechnol*, **10**, 10387–10392.
Wei, W. and Chinnery, P.F. (2020) Inheritance of mitochondrial DNA in humans: implications for rare and common diseases. *J Intern Med*, **287**, 634–644.
Weihe, A., Apitz, J., Pohlheim, F., Salinas-Hartwig, A. and Börner, T. (2009) Biparental inheritance of plastidial and mitochondrial DNA and hybrid variegation in Pelargonium. *Molecular Genetics and Genomics*, **282**, 587–593.
Weissig, V. and Edeas, M. (2021) Recent developments in mitochondrial medicine (Part 1). *4open*, **4**, 2.
Welchen, E., García, L., Mansilla, N. and Gonzalez, D.H. (2014) Coordination of plant mitochondrial biogenesis: keeping pace with cellular requirements. *Front Plant Sci*, **4**.
Wen, J., Wang, Y., Yuan, M., Huang, Z., Zou, Q., Pu, Y., Zhao, B. and Cai, Z. (2021) Role of mismatch repair in aging. *Int J Biol Sci*, **17**, 3923–3935.
West, C.E., Waterworth, W.M., Jiang, Q. and Bray, C.M. (2000) Arabidopsis DNA ligase IV is induced by γ-irradiation and interacts with an Arabidopsis homologue of the double strand break repair protein XRCC4. *The Plant Journal*, **24**, 67–78.
West, S.C. (1997) Processing of recombination intermediates by the RuvABC proteins. *Annu. Rev. Genet.*, **31**, 213–244.
Weterings, E. and Chen, D.J. (2008) The endless tale of non-homologous end-joining. *Cell Res*, **18**, 114–124.
Whitby, M.C. and Lloyd, R.G. (1998) Targeting Holliday junctions by the RecG branch migration protein of Escherichia coli. *J. Biol. Chem.*, **273**, 19729–19739.
Whitby, M.C., Vincent, S.D. and Lloyd, R.G. (1994) Branch migration of Holliday junctions: identification of RecG protein as a junction specific DNA helicase. *EMBO J.*, **13**, 5220–5228.
Wiedenheft, B., Sternberg, S.H. and Doudna, J.A. (2012) RNA-guided genetic silencing systems in bacteria and archaea. *Nature*, **482**, 331–338.
Wiemerslage, L. and Lee, D. (2016) Quantification of mitochondrial morphology in neurites of dopaminergic neurons using multiple parameters. *J Neurosci Methods*, **262**, 56–65.
Wiktor, J., van der Does, M., Büller, L., Sherratt, D.J. and Dekker, C. (2018) Direct observation of end resection by RecBCD during double-stranded DNA break repair in vivo. *Nucleic Acids Res*, **46**, 1821–1833.
Wisnovsky, S., Sack, T., Pagliarini, D.J., Laposa, R.R. and Kelley, S.O. (2018) DNA Polymerase θ Increases Mutational Rates in Mitochondrial DNA. *ACS Chem Biol*, **13**, 900–908.
Woloszynska, M. (2010) Heteroplasmy and stoichiometric complexity of plant mitochondrial genomes--though this be madness, yet there's method in't. *J. Exp. Bot.*, **61**, 657–671.
Woloszynska, M. and Trojanowski, D. (2009) Counting mtDNA molecules in Phaseolus vulgaris: sublimons are constantly produced by recombination via short repeats and undergo rigorous selection during substoichiometric shifting. *Plant Mol Biol*, **70**, 511–521.
Wu, S.-Y. (2003) Dissimilar mispair-recognition spectra of Arabidopsis DNA-mismatch-repair proteins MSH2{middle dot}MSH6 (MutS) and MSH2{middle dot}MSH7 (MutS). *Nucleic Acids Res*, **31**, 6027–6034.
Wu, Y., Sugiyama, T. and Kowalczykowski, S.C. (2006) DNA Annealing Mediated by Rad52 and Rad59 Proteins. *Journal of Biological Chemistry*, **281**, 15441–15449.
Wu, Z., Waneka, G., Broz, A.K., King, C.R. and Sloan, D.B. (2020) MSH1 is required for maintenance of the low mutation rates in plant mitochondrial and plastid genomes. *Proceedings of the National Academy of Sciences*, **117**, 16448–16455.
Xiao, W. and Samson, L. (1993) In vivo Evidence for Endogeneous DNA Alkylation Damage as a Source of Spontaneous Mutation in Eukaryotic Cells. *Proceedings of the National Academy of Sciences*, **90**, 2117–2121.
Xu, F., Yang, X., Zhao, N., Hu, Z., Mackenzie, S.A., Zhang, M. and Yang, J. (2022) Exploiting sterility and fertility variation in cytoplasmic male sterile vegetable crops. *Hortic Res*, **9**.

Xu, W., Boyd, R.M., Tree, M.O., Samkari, F. and Zhao, L. (2019) Mitochondrial transcription factor A promotes DNA strand cleavage at abasic sites. *Proceedings of the National Academy of Sciences*, **116**, 17792–17799.

Xu, Y.Z., Arrieta-Montiel, M.P., Virdi, K.S., et al. (2011) MutS HOMOLOG1 is a nucleoid protein that alters mitochondrial and plastid properties and plant response to high light. *Plant Cell*, **23**, 3428–3441.

Xu, Y.Z., Santamaria Rde, L., Virdi, K.S., et al. (2012) The chloroplast triggers developmental reprogramming when mutS HOMOLOG1 is suppressed in plants. *Plant Physiol*, **159**, 710–720.

Xue, C., Liang, K., Liu, Z., Wen, R. and Xiao, W. (2015) Similarities and differences between Arabidopsis PCNA1 and PCNA2 in complementing the yeast DNA damage tolerance defect. *DNA Repair (Amst)*, **28**, 28–36.

Yadav, A.K., Sahoo, P.K., Goswami, H.N. and Jain, D. (2019) Transcriptional Fidelity of Mitochondrial RNA Polymerase RpoTm from Arabidopsis thaliana. *J Mol Biol*, **431**, 4767–4783.

Yamagishi, H. and Bhat, S.R. (2014) Cytoplasmic male sterility in Brassicaceae crops. *Breed Sci*, **64**, 38–47.

Yamamoto, M.P., Kubo, T. and Mikami, T. (2005) The 5′-leader sequence of sugar beet mitochondrial atp6 encodes a novel polypeptide that is characteristic of Owen cytoplasmic male sterility. *Molecular Genetics and Genomics*, **273**, 342–349.

Yang, J., Zhang, M. and Yu, J. (2007) Alterations of RNA Editing for the Mitochondrial *ATP9* Gene in a New *orf220* -type Cytoplasmic Male-sterile Line of Stem Mustard (*Brassica juncea* var. *tumida*). *J Integr Plant Biol*, **49**, 672–677.

Yang, J.-H., Huai, Y. and Zhang, M.-F. (2009) Mitochondrial atpA gene is altered in a new orf220-type cytoplasmic male-sterile line of stem mustard (Brassica juncea). *Mol Biol Rep*, **36**, 273–280.

Yasui, A., Yajima, H., Kobayashi, T., Eker, A.P.M. and Oikawa, A. (1992) Mitochondrial DNA repair by photolyase. *Mutation Research/DNA Repair*, **273**, 231–236.

Yasukawa, T., Reyes, A., Cluett, T.J., Yang, M.-Y., Bowmaker, M., Jacobs, H.T. and Holt, I.J. (2006) Replication of vertebrate mitochondrial DNA entails transient ribonucleotide incorporation throughout the lagging strand. *EMBO J*, **25**, 5358–5371.

Yu, X., Lu, H., Lu, G., Chen, Z., Cao, J. and Hirata, Y. (2010) Analysis of genetic diversity in cytoplasmic male sterility, and association of mitochondrial genes with petaloid-type cytoplasmic male sterility in tuber mustard (Brassica juncea var. tumida Tsen et Lee). *Mol Biol Rep*, **37**, 1059–1067.

Żabka, A., Trzaskoma, P. and Maszewski, J. (2013) Dissimilar effects of β-lapachone- and hydroxyurea-induced DNA replication stress in root meristem cells of Allium cepa. *Plant Physiology and Biochemistry*, **73**, 282–293.

Zaegel, V., Guermann, B., Ret, M. Le, Andres, 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.

Zaegel, V., Guermann, B., Ret, M. Le, Andrés, C., Meyer, D., Erhardt, M., Canaday, J., Gualberto, J.M. and Imbault, P. (2007) The Plant-Specific ssDNA Binding Protein OSB1 Is Involved in the Stoichiometric Transmission of Mitochondrial DNA in *Arabidopsis* . *Plant Cell*, **18**, 3548–3563.

Zampini, E., Lepage, E., Tremblay-Belzile, S., Truche, S. and Brisson, N. (2015) Organelle DNA rearrangement mapping reveals U-turn-like inversions as a major source of genomic instability in Arabidopsis and humans. *Genome Res.*, **25**, 645–654

Zancani, M., Braidot, E., Filippi, A. and Lippe, G. (2020) Structural and functional properties of plant mitochondrial F-ATP synthase. *Mitochondrion*, **53**, 178–193.

Zeviani, M., Moraes, C.T., DiMauro, S., Nakase, H., Bonilla, E., Schon, E.A. and Rowland, L.P. (1988) Deletions of mitochondrial DNA in Kearns-Sayre syndrome. *Neurology*, **38**, 1339–1339.

Zhang, H., Burr, S.P. and Chinnery, P.F. (2018) The mitochondrial DNA genetic bottleneck: inheritance and beyond. *Essays Biochem*, **62**, 225–234.

Zhao, S.-Y., Wang, G.-D., Zhao, W.-Y., Zhang, S., Kong, F.-Y., Dong, X.-C. and Meng, Q.-W. (2018) Overexpression of tomato WHIRLY protein enhances tolerance to drought stress and resistance to Pseudomonas solanacearum in transgenic tobacco. *Biol Plant*, **62**, 55–68.

Zhou, C., Okuno, M., Nakazato, I., Tsutsumi, N. and Arimura, S. (2023) Targeted A-to-G base editing in the organellar genomes of Arabidopsis with monomeric programmable deaminases. *Plant Physiol*.

Zhou, J., Liu, L. and Chen, J. (2010) Mitochondrial DNA heteroplasmy in Candida glabrata after mitochondrial transformation. *Eukaryot. Cell*, **9**, 806–814.

Zhou, Z., Dun, X., Xia, S., et al. (2012) BnMs3 is required for tapetal differentiation and degradation, microspore separation, and pollen-wall biosynthesis in Brassica napus. *J Exp Bot*, **63**, 2041–2058.

Ziv, O., Geacintov, N., Nakajima, S., Yasui, A. and Livneh, Z. (2009) DNA polymerase ζ cooperates with polymerases κ and ι in translesion DNA synthesis across pyrimidine photodimers in cells from XPV patients. *Proceedings of the National Academy of Sciences*, **106**, 11552–11557.

Zou, G.-M. and Maitra, A. (2008) Small-molecule inhibitor of the AP endonuclease 1/REF-1 E3330 inhibits pancreatic cancer cell growth and migration. *Mol Cancer Ther*, **7**, 2012–2021.

Zou, Y., Zhu, W., Sloan, D.B. and Wu, Z. (2022) Long-read sequencing characterizes mitochondrial and plastid genome variants in Arabidopsis *msh1* mutants. *The Plant Journal*, **112**, 738–755.

Appendix

Appendix.1. List of primers used in this work

Genotyping		
Name	**Sequence**	**Function**
MSH1-1-F	AGTGGCCATTGCGTACACGGA	Genotyping MSH1 (WT copy)
MSH1-1-R	TTGCTGAACGAGATATCCCTGTA	Genotyping MSH1 (WT copy and SALK T-DNA)
ODB1-1-F	GCGGGACGATGATACAAAC	Genotyping ODB1-1 (WT copy and SAIL T-DNA)
ODB1-1-R	CGAGTTCGGATGACTGAATCTG	Genotyping ODB1-1 (WT copy)
OSB1-1-F	CAATAACAAGAAACCAGAGACC	Genotyping OSB1-1 (WT copy and GK T-DNA)
OSB1-1-R	CAACAGACTATGAGAATCGAAC	Genotyping OSB1-1 (WT copy)
OSB4-1-F	GAGAATCTAAAAAGGCGTAGTC	Genotyping OSB4-1 (WT copy and SALK T-DNA)
OSB4-1-R	GGTTACGTTCACCAACCAGT	Genotyping OSB4-1 (WT copy)
RECG1-2-F	ACTGTTACACGGGAGGATGAAG	Genotyping RECG1-2 (WT copy)
RECG1-2-R	TGTTTCTTCCCAAGCAAATCAC	Genotyping RECG1-2 (WT copy and FLAG T-DNA)
POL1A-2-F	CAGGTATCTTTCTCCTACGTC	Genotyping POL1A-2 (WT copy and SALK T-DNA)
POL1A-2-R	CAAATCGTGTACTTACAGTTGCC	Genotyping POL1A-2 (WT copy)
POL1B-1-F	CAGACTTGACTCATCGCTGAA	Genotyping POL1B-1 (WT copy and SALK T-DNA)
POL1B-1-R	AATGAGTAAATCTAATTCTCGATG	Genotyping POL1B-1 (WT copy)
POL1B-2-F	GTTAGCTTGAATGGAGATGCAATC	Genotyping POL1B-2 (WT copy)
POL1B-2-R	GAGTCCCGCTCTGCAGTTCAG	Genotyping POL1B-2 (WT copy and GK T-DNA)
RADA-3-F	CCTGTTTTCTTGGTGAGTAAG	Genotyping RADA-3 (WT copy and SALK T-DNA)
RADA-3-R	GATTCTTCACAGATCGAAGC	Genotyping RADA-3 (WT copy)
WHY2-1-F	AAATCATAAAACCTCAGAAGTCG	Genotyping WHY2-1 (WT copy)
WHY2-1-R	AGTGAAACTGGGAAGAACAGG	Genotyping WHY2-1 (WT copy and SAIL T-DNA)
tDNA-SALK	TGGTTCACGTAGTGGGCCATCG	Genotyping T-DNA-SALK mutants
tDNA-SAIL	TAGCATCTGAATTTCATAACCAATCTCGATACAC	Genotyping T-DNA-SAIL mutants
tDNA-KG	ATATTGACCATCATACTCATTGC	Genotyping T-DNA-GK mutants
tDNA-FLAG	CGGCTATTGGTAATAGGACACTGG	Genotyping T-DNA-FLAG mutants
multi-cox2-F	ACAAATCTCACTGCACTGAC	Multiplexing for copetitive PCR TALEN genotyping according to Arimura et al., 2020
multi-cox2-R	GTGGTGGGGCCTAGTAGTATTA	Multiplexing for copetitive PCR TALEN genotyping according to Arimura et al., 2020
multi-TAL-1-F	GCAGTTCCACGAATTCCGGA	Multiplexing for copetitive PCR TALEN genotyping according to Arimura et al., 2020
multi-TAL-1-R	CAGGAACGCCCAAATTCTGC	Multiplexing for copetitive PCR TALEN genotyping according to Arimura et al., 2020
multi-TAL-2-F	ACTGATCAGTTGTGTGAGGG	Multiplexing for copetitive PCR TALEN genotyping according to Arimura et al., 2020
multi-TAL-2-R	CCTCCAGACAGCTTACGTCA	Multiplexing for copetitive PCR TALEN genotyping according to Arimura et al., 2020
Fok1-F	TGAGGCACAAGCTGAAGTACGTG	Genotyping for FOK1 domain of mitoTALEN construct
Fok1-R	TGCAGTTGGTGATGTGGTTCAGC	Genotyping for FOK1 domain of mitoTALEN construct
qPCR analysis		
Name	**Sequence**	**Function**
TAL-1F	TGGGACAATCTCAAATTGGTCC	qPCR repair pathway check TAL-1 cut site appendix 6
TAL-1R	ATATGCAGAATTTGGGCGTTCC	qPCR repair pathway check TAL-1 cut site appendix 6
Yrepeat-F	CTGGCAGAACGTCGTTAGTTCC	qPCR repair pathway check Y repeat (SSA) appendix 6
Yrepeat-R	ACGTAAGCTGTCTGGAGGGAATC	qPCR repair pathway check Y repeat (SSA) appendix 6
Vrepeat-F	TGAAATTCGAGCAGCGAAC	qPCR repair pathway check V repeat appendix 6
Vrepeat-R	AAGTATGGCTTGGCGTGTTG	qPCR repair pathway check V repeat appendix 6
TAL-2F	GATCACGGGATGGAGTCAAG	qPCR repair pathway check TAL-2 cut site appendix 6
TAL-2R	TACTGATCAGTTGTGTGAGG	qPCR repair pathway check TAL-2 cut site appendix 6
Qrepeat-F	ATACATCCGCGGAGAATGTACC	qPCR repair pathway check Q repeat appendix 6
Qrepeat-R	AGGAAAGAGGATGACACAATGG	qPCR repair pathway check Q repeat appendix 6
Lrepeat-F	GAAACCTTCTTTCTTCGATCA	qPCR repair pathway check L repeat appendix 6
Lrepeat-R	GTTATGCCATTTTGGGCTTTGCTC	qPCR repair pathway check L repeat appendix 6
Grepeat-F	TAGGGTCCCATCTCTCTTCG	qPCR repair pathway check G repeat appendix 6
Grepeat-R	TGGTGTATCTGGTGCTCTCG	qPCR repair pathway check G repeat appendix 6
MHMR1-F	AGCAGCACCAACCTAGCTCTAC	qPCR repair pathway check MHMR-1 appendix 6
MHMR1-R	GCTTATTTCGATACCGCTTCTG	qPCR repair pathway check MHMR-1 appendix 6
MHMR2-F	GGTGTATGGTTCCTAGACCTGT	qPCR repair pathway check MHMR-2 appendix 6
MHMR2-R	AATACTGGGAGACTGAGGAG	qPCR repair pathway check MHMR-2 appendix 6
TAL-3F	CGATAAACGGGAATGCGTTACC	qPCR quntification of the mtDNA at TAL-3 cut site
TAL-3R	CATCGCCTTACTTATCCAAGCG	qPCR quntification of the mtDNA at TAL-3 cut site
atp9-F	GAAACCTTCTTTCTTCGATCA	qPCR quantification of at C terminus of atp9 gene
atp9-R	TTACTACAACTCTTGAGGTGGG	qPCR quantification of at C terminus of atp9 gene
COX2-F	AATAAACGTGATTGACCCAATTCT	qPCR quantification of mtDNA
COX2-R	TCCGATGAGCAGTCACTCAC	qPCR quantification of mtDNA
RRN18-F	CCTTGAGCTAGGAGCCTCTTT	qPCR quantification of mtDNA
RRN18-R	CATGCAAGTCGAACGTTGTT	qPCR quantification of mtDNA
CLp-F	TCGCACTATATGTCAACCCAAG	qPCR quantification of cpDNA
CLp-R	CTATTGGCGTTCCAAAAGTACC	qPCR quantification of cpDNA
NadH-F	TCCGGATAAACCCCAATTTA	qPCR quantification of cpDNA
NadH-R	AGAACGGGTTGAAGGAGTTG	qPCR quantification of cpDNA
ACT-F	CATCTTGGCCTCCCTCAGTA	qPCR quantification of nuclear genome
ACT-R	GAGTAAACAAGTGATGGGACTGTG	qPCR quantification of nuclear genome
UBQ-F	TCACCGGAAAGACCATCACT	qPCR quantification of nuclear genome
UBQ-R	CGGTGGGATACCCTCTTTG	qPCR quantification of nuclear genome
mtDNA-1-F	TGGTGGAAAAAGATCGTGAAATGG	Quatification of the mtDNA region coordinates 9.0 kb - 9.2 kb (phenotype variagation test)
mtDNA-1-R	GGGACTTCAACGGGTCTTCCAAG	Quatification of the mtDNA region coordinates 9.0 kb - 9.2 kb (phenotype variagation test)
mtDNA-2-F	AGCAGCACCAACCTAGCTCTAC	Quatification of the mtDNA region coordinates 268.5 kb - 268.7 kb (phenotype variagation test)

mtDNA-2-R	GCTTATTTCGATACCGCTTCTG	Quatification of the mtDNA region coordinates 268.5 kb - 268.7 kb (phenotype variagation test)
mtDNA-3-F	GAAACCTTCTTTCTTCGATCA	Quatification of the mtDNA region coordinates 270.9 kb - 271.1 kb (phenotype variagation test)
mtDNA-3-R	GTTATGCCATTTTGGGCTTTGCTC	Quatification of the mtDNA region coordinates 270.9 kb - 271.1 kb (phenotype variagation test)
mtDNA-4-F	AATAAACGTGATTGACCCAATTCT	Quatification of the mtDNA region coordinates 6.6 kb - 6.8 kb (Pol1 test upon urea treatment)
mtDNA-4-R	TCCGATGAGCAGTCACTCAC	Quatification of the mtDNA region coordinates 6.6 kb - 6.8 kb (Pol1 test upon urea treatment)
mtDNA-5-F	AACAAACCAATTACCAAATCCAC	Quatification of the mtDNA region coordinates 116.1 kb - 116.3 kb (Pol1 test upon urea treatment)
mtDNA-5-R	ATTCTTGGTGGGAATCATCAAC	Quatification of the mtDNA region coordinates 116.1 kb - 116.3 kb (Pol1 test upon urea treatment)
mtDNA-6-F	TGGTCTCTGGTTTGATGGTTGC	Quatification of the mtDNA region coordinates 165.9 kb - 166.1 kb (Pol1 test upon urea treatment)
mtDNA-6-R	AGGAATGAAACGGCTATAGCTC	Quatification of the mtDNA region coordinates 165.9 kb - 166.1 kb (Pol1 test upon urea treatment)
mtDNA-7-F	CGGATCAAAGATAGCACGAACC	Quatification of the mtDNA region coordinates 332.7 kb - 332.8 kb (Pol1 test upon urea treatment)
mtDNA-7-R	GTTTCATTCTGTGCGCTAGTG	Quatification of the mtDNA region coordinates 332.7 kb - 332.8 kb (Pol1 test upon urea treatment)
mtDNA-8-F	CGGCACTACTCTTGTTTACAGTT	Quatification of the mtDNA region coordinates 350.8 kb - 351.0 kb (Pol1 test upon urea treatment)
mtDNA-8-R	TCACAACCAGAATGAGAACTC	Quatification of the mtDNA region coordinates 350.8 kb - 351.0 kb (Pol1 test upon urea treatment)
	Long fragment PCR	
Name	**Sequence**	**Function**
PCR-1	GATAAGGAGAGGGGTCTTCCAT	Amplification Long PCR fragment for LR1 - V recombination analysis
PCR-2	CTGGCAGAACGTCGTTAGTTCC	Amplification Long PCR fragment for LR1 - V recombination analysis
PCR-3	GACGGTTGGCTAAAGACATGG	Amplification of Long PCR fragment for SSA analysis
PCR-4	GGAGAGGTGCTTTAGTCAGGA	Amplification of Long PCR fragment for SSA analysis
PCR-5	GATCAAGCCAGCTAGAGAGGCA	Amplification Long PCR fragment in MSH1 for LR1 - V recombination analysis
PCR-6	CGATAAACGGGAATGCGTTACC	Amplification Long PCR fragment in MSH1 for LR1 - V recombination analysis
	Sanger sequencing	
Name	**Sequence**	**Function**
P-1	CAAAAGGCTACATCAGACGATG	Sequencing Long PCR fragment from PCR-1 end
P-2	CTGGCAGAACGTCGTTAGTTCC	Sequencing Long PCR fragment from PCR-2 end
P-3	AAGTATGGCTTGGCGTGTTG	Sequencing Long PCR fragment from PCR-3 end
P-4	AACAAAGATGGATTTGTGAA	Sequencing Long PCR fragment from PCR-4 end
P-5	AGTGCAGCAGCCAATAATACG	Sequencing Long PCR fragment from PCR-5 end
P-6	AACAAAGATGGATTTGTGAA	Sequencing Long PCR fragment from PCR-6 end
	RT-qPCR analysis	
Name	**Sequence**	**Function**
EXP-F	GAGCTGAAGTGGCTTCAATGAC	Reference gene for RT-qPCR
EXP-R	GGTCCGACATACCCATGATCC	Reference gene for RT-qPCR
ACT-F	GGATCTGTACGGTAACATTGTGC	Reference gene for RT-qPCR
ACT-R	TGCTCATACGGTCAGCGATA	Reference gene for RT-qPCR
GADPH-F	TTGGTGACAACAGGTCAAGCA	Reference gene for RT-qPCR
GADPH-R	AAACTTGTCGCTCAATGCAATC	Reference gene for RT-qPCR
atp6-F	ctttgttcatgctgctaactctg	quatification of atp6 gene transcript by RT-qPCR
atp6-R	gccaagcatttgggactaag	quatification of atp6 gene transcript by RT-qPCR
atp9-F	CGGAGCTGCTACAATTGCTT	quatification of atp9 gene transcript by RT-qPCR
atp9-R	CGCCACAGAATGAATCAAAG	quatification of atp9 gene transcript by RT-qPCR
Fok1-F	CGAAATGCAGAGGTACGTGGAG	quatification of mitoTALEN gene transcript by RT-qPCR
Fok1-R	ACAGGAACTTGAACTCGGTCAC	quatification of mitoTALEN gene transcript by RT-qPCR
	Northern blot	
Name	**Sequence**	**Function**
Probe-atp9-F	GACGGATCAAATTGAATAATCGAAG	Preparation of probe for Northern blot
Probe-atp9-R	GAGCAAAGCCCAAAATGGCATAAC	Preparation of probe for Northern blot

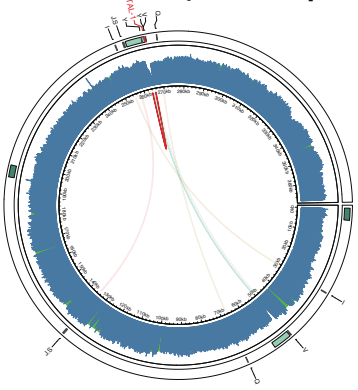
Appendix.2. Table of sequenced mitoTALEN mutants in this project:

Genotype	No.	HR tDNA	TALEN	Sample ID	mtDNA Enrichment (qPCR)	Sequencing	RAW	reads mapped to CpDNA	cpDNA (%)	reads mapped to mtDNA	mtDNA (%)	mtDNA enrichment (x times)	av.library.size (bp)	recombined_pairs (%)	Repeat activity next to cut site comments
Col-0	S1	YES	TAL-2	N1	65	Miseq NGS246	3903294	133880	3.43%	475038	12.17%	5.7	354	0.13%	Large_1, Q_repeat, G_repeat, L_repeat
Col-0	S2	YES	TAL-2	N2	14	Miseq NGS246	3596638	108092	3.01%	264746	7.36%	3.5	343	0.14%	Large_1, G_repeat
Col-0	S27	NO	TAL-2	1A7	16	Nextseq NGS257	16935524	2205548	13.02%	1992034	11.76%	5.5	303	0.17%	Large_1, M_repeat, G_repeat
Col-0	S28	NO	TAL-2	1B6	22	Nextseq NGS257	17481358	2710824	15.51%	1736968	9.94%	4.7	294	0.31%	Large_1, G_repeat, L_repeat
Col-0	S29	NO	TAL-1	1E10	60	Nextseq NGS257	17025212	3729648	21.91%	2777428	16.31%	7.6	302	0.17%	Large_1, H_repeat, G_repeat
Col-0	S3	YES	TAL-2	N3	11	Miseq NGS246	3364054	153416	4.56%	267188	7.94%	3.7	322	0.11%	Large_1, Q_repeat, G_repeat
Col-0	S57	NO	TAL-1	G5	7	Nextseq NGS313	13744446	2160912	15.72%	1282898	9.33%	4.4	354	0.18%	Large_1, V_repeat ,G_repeat, L_repeat
Col-0	S58	NO	TAL-1	1E12	17	Nextseq NGS313	21821898	4374232	20.05%	2919480	13.38%	6.3	355	0.20%	Large_1, Q_repeat, G_repeat
Col-0	S59	NO	TAL-1	1E8	10	Nextseq NGS313	16143332	2996436	18.56%	1825146	11.31%	5.3	340	0.20%	Large_1, Q_repeat, G_repeat
Col-0	S1	NO	TAL-1	1E7	7	Nextseq NGS313	19927936	3716590	18.65%	1816470	9.12%	4.3	328	0.15%	Y_repeat
Col-0	S24	NO	TAL-1	N1	44	Nextseq NGS257	15614424	1317480	8.44%	2974380	19.05%	8.9	295	0.15%	Large_1, V_repeat
Col-0	S25	NO	TAL-1	N2	44	Nextseq NGS257	15095146	674642	4.47%	1666874	11.04%	5.2	300	0.14%	Large_1, G_repeat, MHMR
Col-0	S26	NO	TAL-1	N3	17	Nextseq NGS257	14853054	668332	4.50%	1347440	9.07%	4.3	310	0.14%	Large_1, AA_repeat, V_repeat, G_repeat
Col-0	S55	NO	TAL-2	1A12	12	Nextseq NGS313	18918682	2929624	15.49%	1138466	6.02%	2.8	333	0.20%	Large_1, G_repeat, L_repeat
Col-0	S56	NO	TAL-2	1A6	8	Nextseq NGS313	18577454	2895926	15.59%	1755446	9.45%	4.4	351	0.27%	Large_1, G_repeat, L_repeat
msh1-1	msh1-1	YES	NO	msh1-1	not detected	Nextseq NGS354	8629594	647278	7.50%	1048750	12.15%	5.7	346	0.49%	Increased ectopic recombination at multiple sites, atp6-1 gene lost
msh1-1	S2	YES	TAL-2	6E5	40	Nextseq NGS313	21144218	2918360	13.80%	5601994	26.49%	12.4	290	0.39%	I_repeat, J_repeat, L_repeat, possible sub-genome segregation
msh1-1	S3	YES	TAL-2	6E4	29	Nextseq NGS313	20804242	3221562	15.49%	4432372	21.31%	10.0	307	0.41%	I_repeat, MHMR, H_repeat, L_repeat, possible sub-genome segregation
msh1-1	S5	YES	TAL-1	4D11	5	Nextseq NGS313	17056406	2951988	17.31%	1570514	9.21%	4.3	339	0.62%	Y_repeat SSA for TAL-2 cut site, heteroplasmic state, TAL-2 cut site still present
msh1-1	S6	YES	TAL-1	4D10	9	Nextseq NGS313	16827040	1151522	6.84%	2859864	17.00%	8.0	335	0.67%	Y_repeat SSA
msh1-1	S7	YES	TAL-2	4D6	5	Nextseq NGS313	13255710	2462164	18.57%	1710530	12.90%	6.0	354	0.61%	I_repeat, L_repeat, possible sub-genome segregation
msh1-1	S8	YES	TAL-2	4D5	7	Nextseq NGS313	21097718	4507840	21.37%	1545188	7.32%	3.4	322	0.49%	I_repeat, L_repeat, possible sub-genome segregation
msh1-1	S9	YES	TAL-2	4D3	4	Nextseq NGS313	11478364	2199350	19.16%	1164388	10.14%	4.8	335	0.63%	I_repeat, MHMR, H_repeat, L_repeat, G_repeat, possible sub-genome segregation
msh1-1	S4	YES	TAL-2	6E1	27	Nextseq NGS313	17640128	2865126	16.24%	3826350	21.69%	10.2	293	0.42%	I_repeat, L_repeat, G_repeat, possible sub-genome segregation
odb1-1	odb1-1	YES	NO	odb1-1	7	Nextseq NGS354	8894368	629380	7.08%	882532	9.92%	4.7	334	0.15%	odb1 mutant no TALEN - mtDNA looks like WT
odb1-1	S6	YES	TAL-1	1C5	10	Nextseq NGS257	16024650	7153098	44.64%	1628772	10.16%	4.8	269	0.14%	Y_repeat
odb1-1	S7	YES	TAL-1	1C6	15	Nextseq NGS257	13877200	4404052	31.74%	2113242	15.23%	7.1	287	0.27%	Large_1, V_repeat, G_repeat
odb1-1	S8	YES	TAL-1	1C7	9	Nextseq NGS257	13173820	7463986	56.66%	1166974	8.86%	4.2	274	0.16%	Y_repeat
odb1-1	S9	YES	TAL-1	1D6	13	Nextseq NGS257	39194154	9163360	23.38%	3826304	9.76%	4.6	301	0.27%	Large_1, G_repeat, L_repeat, has some WT mitotype in heteroplasmy
odb1-1	S47	YES	TAL-2	2A4	11	Nextseq NGS257	18322574	5312752	29.00%	1408822	7.69%	3.6	304	0.28%	Large_1, G_repeat, L prime repeat
odb1-1	S48	YES	TAL-2	2A6	11	Nextseq NGS257	16401728	7304660	44.54%	1191362	7.26%	3.4	305	0.23%	Large_1, G_repeat, L_repeat
odb1-1	S49	YES	TAL-2	2B3	19	Nextseq NGS257	17358504	6028568	34.73%	1793264	10.33%	4.8	312	0.35%	Large_1, H_repeat, G_repeat
odb1-1	S50	YES	TAL-2	2B11	12	Nextseq NGS257	19971046	5846324	29.27%	2159754	10.81%	5.1	313	0.35%	Large_1, G_repeat
oex1	oex1	YES	NO	mix	10	Nextseq NGS313	16896412	1795386	10.63%	1806330	10.69%	5.0	312	0.26%	oex sample .. No TALEN, segregation of sub-genomes
oex1	S34	YES	TAL-2	4B12	16	Nextseq NGS313	16731086	1351164	8.08%	1889154	11.29%	5.3	309	0.40%	Large_1, L_repeat, G_repeat, heteroplasmic state, WT mitotype present
oex1	S35	YES	TAL-2	4B10	16	Nextseq NGS313	11082938	1534322	13.84%	1592866	14.37%	6.7	336	0.33%	Large_1, G_repeat, L_repeat
oex1	S33	YES	TAL-1	4A3	7	Nextseq NGS313	13000594	1230292	9.46%	1513196	11.64%	5.5	358	0.33%	Large_1, G_repeat, L_repeat
oex1	S31	YES	TAL-2	4A11	60	Nextseq NGS313	19749600	1740432	8.81%	2031786	10.29%	4.8	328	0.25%	Large_1, L_repeat, heteroplasmic state, WT mitotype present or no cut DNA
oex1	S32	YES	TAL-2	4A9	10	Nextseq NGS313	13402928	1080278	8.06%	1790042	13.36%	6.3	351	0.28%	Large_1, L_repeat, G_repeat, heteroplasmic state, WT mitotype present
osb1 osb4	osb1 osb4	YES	NO	osb1 osb4	10	Nextseq NGS354	10568682	797696	7.55%	945104	8.94%	4.2	341	0.19%	osb1 osb4, segregation of sub-genomes
osb1 osb4	S20	1 YES 4 YES	TAL-1	1G7	19	Nextseq NGS257	16688850	6073368	36.39%	2086816	12.50%	5.9	312	0.19%	J_repeat, L_repeat; it seems no repair; segregation of sub-genomes
osb1 osb4	S21	1 YES 4 YES	TAL-1	1G10	16	Nextseq NGS257	15868410	3926616	24.74%	1731528	10.91%	5.1	313	0.18%	Large_1, L_repeat, B_repeat, sub-genome segregation
osb1 osb4	S22	1 YES 4 YES	TAL-1	1G11	26	Nextseq NGS257	13752670	4653414	33.84%	1411252	10.26%	4.8	292	0.18%	J_repeat, L_repeat; it seems no repair; segregation of sub-genomes
osb1 osb4	S23	1 YES 4 YES	TAL-1	1G12	20	Nextseq NGS257	21859968	4732826	21.65%	1819250	8.32%	3.9	307	0.24%	Large_1, L_repeat, sub-genome segregation
osb1 osb4	S16	1 YES 4 YES	TAL-1	OSUP2	9	Nextseq NGS313	15607458	1739748	11.15%	1914114	12.26%	5.7	366	0.30%	J_repeat, L_repeat; it seems no repair; segregation of sub-genomes
osb1 osb4	S15	1 YES 4 YES	TAL-2	R1	15	Nextseq NGS313	12893276	752232	5.83%	1310066	10.16%	4.8	326	0.23%	Large_1, L_repeat, sub-genome segregation
pol1A-2	pol1A-2	YES	NO	pol1A-2	not detected	Nextseq NGS354	13563870	659432	4.86%	1014282	7.48%	3.5	311	0.11%	pol1A-2 no TALEN slight bumps in coverage - link to replication
pol1A-2	S49	YES	TAL-1	E2	7	Nextseq NGS313	17217706	1512238	8.78%	1694266	9.84%	4.6	315	0.17%	Large_1, Q_repeat, G_repeat
pol1A-2	S50	YES	TAL-1	D11	4	Nextseq NGS313	16692588	1084874	6.50%	857590	5.14%	2.4	332	0.16%	Y_repeat
pol1A-2	S51	YES	TAL-1	D10	4	Nextseq NGS313	15695548	1050158	6.69%	762486	4.86%	2.3	338	0.18%	Large_1, V_repeat
pol1A-2	S46	YES	TAL-2	2H7	19	Nextseq NGS313	16773860	2119066	12.63%	2489676	14.84%	7.0	344	0.27%	Large_1, G_repeat, L_repeat
pol1A-2	S47	YES	TAL-2	2H3	16	Nextseq NGS313	16327210	2510440	15.38%	2039304	12.49%	5.9	329	0.17%	Large_1, G_repeat, L_repeat
pol1A-2	S48	YES	TAL-2	2G9	14	Nextseq NGS313	10864748	1222654	11.25%	991100	9.12%	4.3	346	0.14%	Large_1, G_repeat
pol1B-1	S14	YES	TAL-1	1H9	15	Nextseq NGS257	19079830	3516202	18.43%	2684662	14.07%	6.6	328	0.27%	Large_1, V_repeat, G_repeat
pol1B-1	S15	YES	TAL-1	1H10	40	Nextseq NGS257	17180784	2553746	14.86%	3543878	20.63%	9.7	330	0.21%	Large_1, V_repeat
pol1B-1	S16	YES	TAL-1	1H11	14	Nextseq NGS257	25665906	2906442	11.32%	2527412	9.85%	4.6	330	0.22%	Large_1, G_repeat, L_repeat
pol1B-1	S45	YES	TAL-1	1H8	30	Nextseq NGS313	21248396	3673166	17.29%	3597378	16.93%	7.9	308	0.16%	Large_1, M_repeat, G_repeat
pol1B-1	S43	YES	TAL-2	3F6	9	Nextseq NGS313	13895742	2023888	14.56%	1057884	7.61%	3.6	340	0.17%	Large_1, G_repeat, L_repeat
pol1B-2	pol1B-2	YES	NO	pol1B-2	8	Nextseq NGS354	9185884	459820	5.01%	597996	6.51%	3.1	339	0.22%	pol1B-2 no TALEN slight bumps in coverage - link to replication
pol1B-2	S17	YES	TAL-1	3C9	9	Nextseq NGS257	14869546	4411802	29.67%	1153376	7.76%	3.6	318	0.24%	Large_1, Q_repeat, G_repeat
pol1B-2	S18	YES	TAL-1	3D2	9	Nextseq NGS257	17919548	3861948	21.55%	1289896	7.20%	3.4	331	0.88%	Large_1, G_repeat, L_repeat
pol1B-2	S19	YES	TAL-1	3D3	23	Nextseq NGS257	11777510	3971670	33.72%	1305454	11.08%	5.2	285	0.12%	Large_1, G_repeat, L_repeat
pol1B-2	S44	YES	TAL-1	3C10	5	Nextseq NGS313	13512376	2908338	21.52%	1002924	7.42%	3.5	328	0.15%	Large_1, V_repeat
pol1B-2	S41	YES	TAL-2	2B2	7	Nextseq NGS313	15541148	2161040	13.91%	946122	6.09%	2.9	306	0.18%	Large_1, H_repeat, G_repeat
pol1B-2	S42	YES	TAL-2	2A1	12	Nextseq NGS313	11032664	1500538	13.60%	1057362	9.58%	4.5	340	0.17%	Large_1, G_repeat, L_repeat
RadA-3	radA-3	YES	NO	radA-3	64	Nextseq NGS354	9701740	161430	1.66%	980458	10.11%	4.7	354	0.33%	radA-3, segregation of sub-genomes
RadA-3	S35	YES	TAL-1	1G1	13	Nextseq NGS257	22466054	9140712	40.69%	2108034	9.38%	4.4	312	0.17%	S_repeat, L_repeat, segregation of sub-genomes
RadA-3	S36	YES	TAL-1	1G2	12	Nextseq NGS257	17187108	3839830	22.34%	2252558	13.11%	6.1	319	0.29%	S_repeat, L_repeat, segregation of sub-genomes
RadA-3	S37	YES	TAL-1	1G3	11	Nextseq NGS257	17433154	3971722	22.78%	1532234	8.79%	4.1	300	0.22%	J_repeat, L_repeat, segregation of sub-genomes
RadA-3	S38	YES	TAL-1	1G4	15	Nextseq NGS257	17634650	2900108	16.45%	1638168	9.29%	4.4	327	0.36%	Large_1, L_repeat, segregation of sub-genomes
RadA-3	S20	YES	TAL-1	X8	12	Nextseq NGS313	17679768	1337428	7.56%	2366934	13.39%	6.3	322	0.21%	Large_1, L_repeat, segregation of sub-genomes
RadA-3	S21	YES	TAL-1	X5	14	Nextseq NGS313	16463106	1193780	7.25%	1990932	12.09%	5.7	318	0.25%	Large_1, L_repeat, segregation of sub-genomes
RadA-3	S17	YES	TAL-2	3F4	14	Nextseq NGS313	14279488	1577124	11.04%	2041772	14.30%	6.7	312	0.25%	Large_1, L_repeat, segregation of sub-genomes
RadA-3	S18	YES	TAL-2	3E4	17	Nextseq NGS313	21904470	2276900	10.39%	3169878	14.47%	6.8	325	0.21%	Large_1, H_repeat, L_repeat, segregation of sub-genomes
RadA-3	S19	YES	TAL-2	3E3	22	Nextseq NGS313	18842130	1389432	7.37%	2542938	13.50%	6.3	323	0.28%	Large_1, L_repeat, segregation of sub-genomes
recA3-2	recA3-2	YES	NO	recA3-2	not detected	Nextseq NGS354	9432254	621824	6.59%	1117340	11.85%	5.6	347	0.19%	recA3 sample .. No TALEN, segregation of sub-genomes
recA3-3	S44	YES	TAL-2	2C9	14	Nextseq NGS257	21974324	6191426	28.18%	1849258	8.42%	3.9	296	0.23%	Large_1, L_repeat, segregation of sub-genomes
RecG1-2	recG1-2	YES	NO	recG1-2	13	Nextseq NGS354	9727770	1143904	11.76%	1359446	13.97%	6.6	359	0.74%	Increased ectopic recombination at multiple sites
RecG1-2	S34	YES	TAL-1	Recg1-2 Up	13	Nextseq NGS257	17136120	2425474	14.15%	1782636	10.40%	4.9	287	0.71%	Y_repeat
RecG1-2	S39	YES	TAL-1	C2	16	Nextseq NGS257	25984078	5106112	19.65%	3274394	12.60%	5.9	283	0.49%	Large_1, V_repeat
RecG1-2	S41	YES	TAL-1	C4	20	Nextseq NGS257	16888768	4662976	27.61%	2381440	14.10%	6.6	315	0.68%	Large_1, L_repeat, S_repeat, Q_repeat, MHMR
RecG1-2	S42	YES	TAL-1	C5	10	Nextseq NGS257	21566458	6218084	28.83%	1641070	7.61%	3.6	297	0.44%	Large_1, V_repeat, Y repeat
RecG1-2	S33	YES	TAL-2	RecG1-2 6-2	9	Nextseq NGS257	15440288	3897622	25.24%	1555306	10.07%	4.7	305	0.73%	Large_1, G_repeat, S_repeat
RecG1-2	S39	YES	TAL-1	C9	12	Nextseq NGS313	19835694	6167070	31.09%	1861594	9.39%	4.4	344	0.60%	Large_1, V_repeat
RecG1-2	S37	YES	TAL-2	A9	11	Nextseq NGS313	20524562	5476330	26.68%	2137158	10.41%	4.9	276	0.38%	Large_1, L_repeat, J_repeat, MHMR, E_repeat,
RecG1-2	S38	YES	TAL-2	A4	19	Nextseq NGS313	15300148	3056920	19.98%	1997942	13.06%	6.1	286	0.40%	Large_1, L_repeat, S_repeat
why2-1	why2-1	YES	NO	why2-1	10	Nextseq NGS354	9157246	786234	8.59%	1074418	11.73%	5.5	340	0.12%	why2-1 mutant no TALEN - mtDNA looks like WT
why2-1	S1	YES	TAL-1	1E5	16	Nextseq NGS257	12450072	4885588	39.24%	1498550	12.04%	5.6	313	0.17%	Large_1, H_repeat, G_repeat
why2-1	S2	YES	TAL-1	1E9	12	Nextseq NGS257	19024456	7715672	40.56%	1895254	9.96%	4.7	301	0.16%	Y_repeat
why2-1	S3	YES	TAL-1	1E10	16	Nextseq NGS257	15446130	6366230	41.22%	1537392	9.95%	4.7	316	0.20%	Large_1, MHMR
why2-1	S4	YES	TAL-1	1E12	14	Nextseq NGS257	26308408	9081574	34.52%	2685374	10.21%	4.8	334	0.23%	Large_1, G_repeat, L_repeat, MHMR
why2-1	S5	YES	TAL-1	1F4	26	Nextseq NGS257	18788796	8327192	44.32%	2455690	13.07%	6.1	315	0.26%	Large_1, V_repeat, M_repeat, G_repeat
why2-1	S4	YES	TAL-2	1A1	16	Miseq NGS246	3305946	1353654	40.95%	506928	15.33%	7.2	357	0.21%	MHMR , Large_1, G_repeat, L_repeat
why2-1	S5	YES	TAL-2	1A3	19	Miseq NGS246	3255800	1306750	40.14%	422444	12.98%	6.1	368	0.18%	MHMR , Large_1, G_repeat, L_repeat
why2-1	S6	YES	TAL-2	1A6	12	Miseq NGS246	3413636	1774464	51.98%	324142	9.50%	4.5	346	0.16%	MHMR
why2-1	S8	YES	TAL-2	1A12	16	Miseq NGS246	3673390	1600532	43.57%	517818	14.10%	6.6	331	0.12%	Large_1, Q_repeat, G_repeat
why2-1	S7	YES	TAL-2	1B9	29	Miseq NGS246	3079388	1089248	35.37%	464622	15.09%	7.1	362	0.19%	MHMR , Large_1, G_repeat, L_repeat
why2-1	S10	YES	TAL-1	6G6	20	Nextseq NGS313	18486132	2356910	12.75%	3223328	17.44%	8.2	327	0.26%	Large_1, V or AA or imperfect repeat (not clear) , L_repeat, G_repeat
why2-1	S11	YES	TAL-1	6G3	7	Nextseq NGS313	21311336	1360084	6.38%	1991774	9.35%	4.4	330	0.28%	Large_1, Y_repeat, V_repeat
why2-1	S12	YES	TAL-1	6C5	5	Nextseq NGS313	14608904	1654274	11.32%	1357900	9.30%	4.4	332	0.18%	Large_1, V_repeat

Appendix.3. Circos representation of HR events found in WT and mutant plants produced in this project

Col-0 TAL-1

Col–0.TAL-1.S1 (Mean depth = 721)

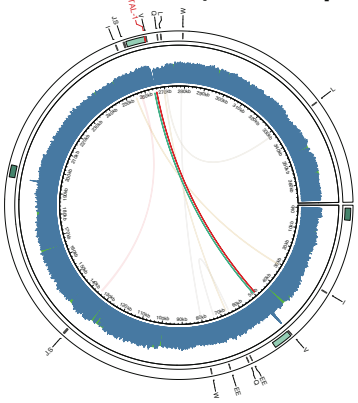


DSB Repaired by

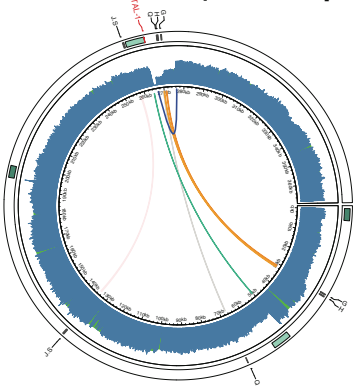
SSA via repeats Y

HR via repeats Large1 x V

Col–0.TAL-1.S24 (Mean depth = 1205)



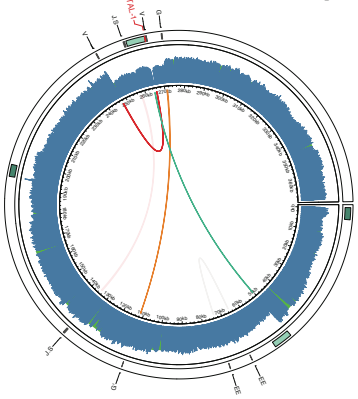
Col–0.TAL-1.S25 (Mean depth = 672)



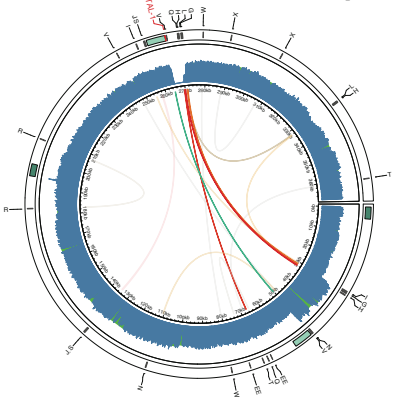
MHMR & HR via repeats Large1 x G/H

HR via repeats V & Large1

Col–0.TAL-1.S26 (Mean depth = 543)



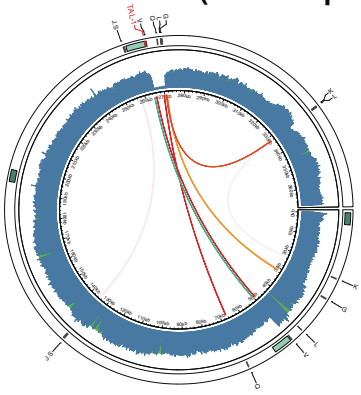
Col–0.TAL-1.S29 (Mean depth = 1123)



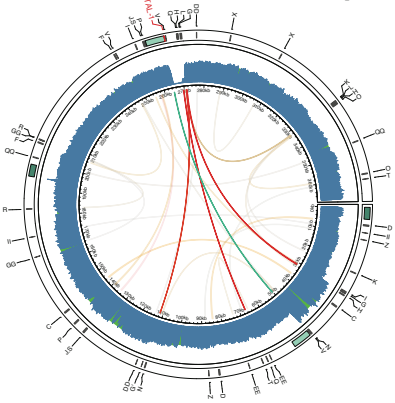
HR via repeats Large1 x G & Q

HR via repeats Large1 x V/G & L

Col–0.TAL-1.S57 (Mean depth = 506)

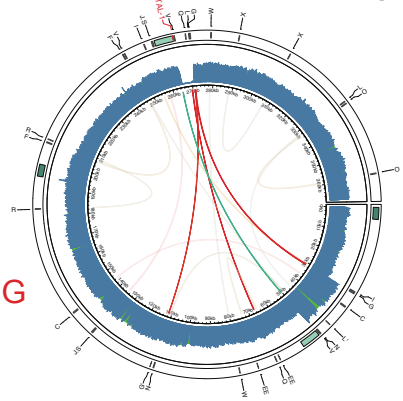


Col–0.TAL-1.S58 (Mean depth = 1163)



HR via repeats Large1 x G & Q

Col–0.TAL-1.S59 (Mean depth = 726)



HR via repeats Large1 x G

Fraction

1
0.8
0.6
0.4
0.2
0

Large1 repeat

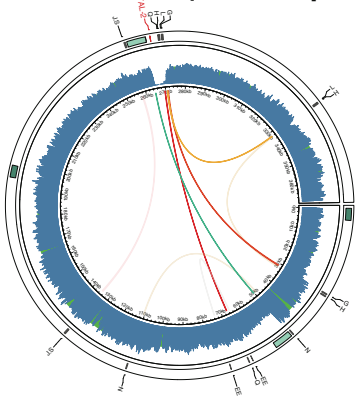
Large2 repeat

Large1 HR

MHMR

Col-0 TAL-2

Col−0.TAL-2.S1 (Mean depth = 192)

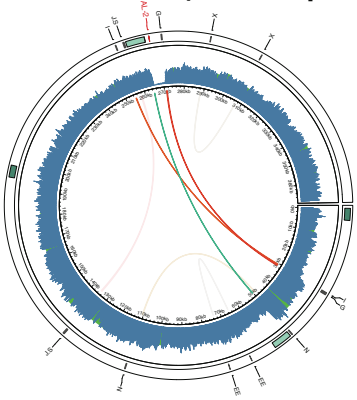


DSB Repaired by

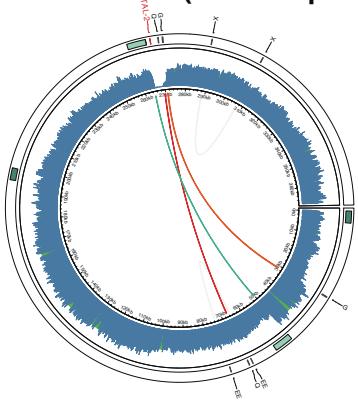
HR via repeats Large1 x G & L/Q

HR via repeats Large1 x G

Col−0.TAL-2.S2 (Mean depth = 107)



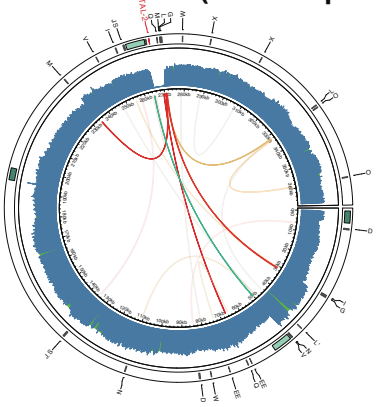
Col−0.TAL-2.S3 (Mean depth = 108)



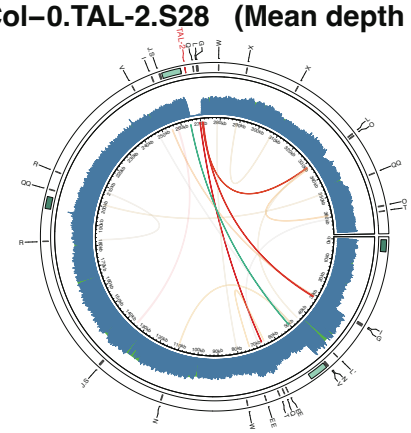
HR via repeats Large1 x G & Q

HR via repeats Large1 x G & M

Col−0.TAL-2.S27 (Mean depth = 802)



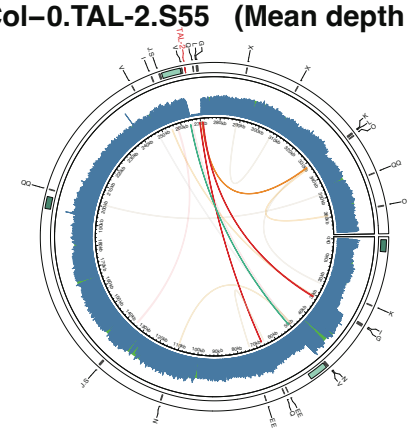
Col−0.TAL-2.S28 (Mean depth = 703)



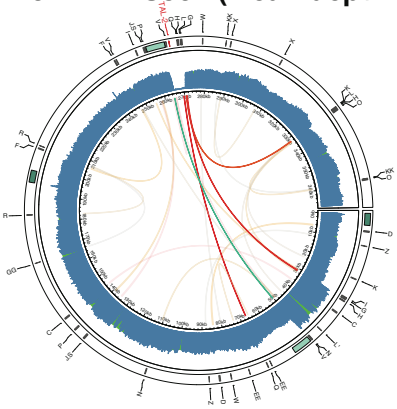
HR via repeats Large1 x G & L

HR via repeats Large1 & G & L

Col−0.TAL-2.S55 (Mean depth = 453)



Col−0.TAL-2.S56 (Mean depth = 701)



HR via repeats Large1 & G & L

Fraction

1
0.8
0.6
0.4
0.2
0

Large1 repeat

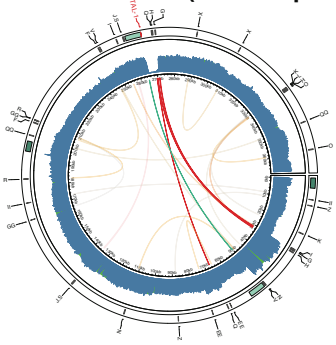
Large2 repeat

Large1 HR

MHMR

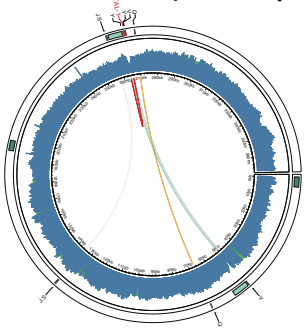
pol1A-2 TAL-1

***pol1A*.TAL-1.S49 (Mean depth = 680)**



HR
via
repeats
Large1 x H/G

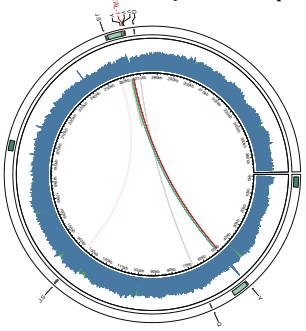
***pol1A*.TAL-1.S50 (Mean depth = 343)**



DSB Repaired by

SSA & HR
via
repeats
Y & Large1 x V

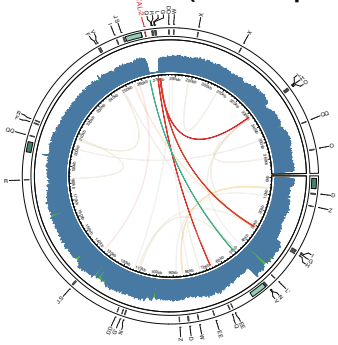
***pol1A*.TAL-1.S51 (Mean depth = 301)**



SSA & HR
via
repeats
Y & Large1 x V

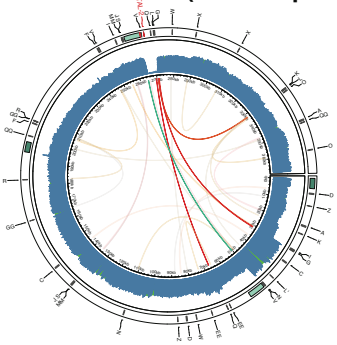
pol1A-2 TAL-2

***pol1A*.TAL-2.S46 (Mean depth = 996)**



HR
via
repeats
Large1 x G
& L

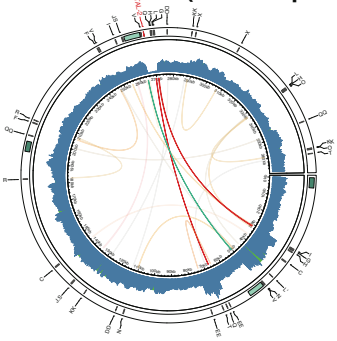
***pol1A*.TAL-2.S47 (Mean depth = 807)**



DSB Repaired by

HR
via
repeats
Large1 x G
& L

***pol1A*.TAL-2.S48 (Mean depth = 406)**



HR
via
repeats
Large1 x G

Fraction

1
0.8
0.6
0.4
0.2
0

Large1 repeat

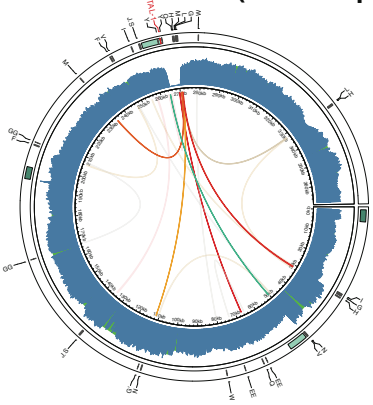
Large2 repeat

Large1 HR

MHMR

pol1B-2 TAL-1

***pol1B* -1.TAL-1.S45 (Mean depth = 1437)**

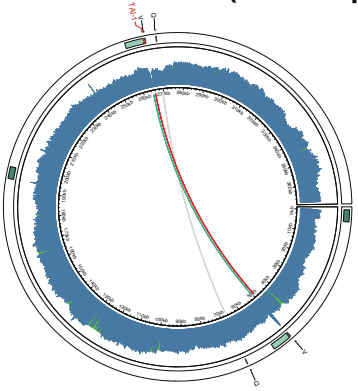


HR via repeats Large1 x G & M

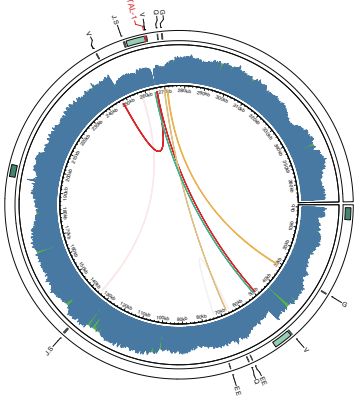
DSB Repaired by

HR via repeats Large1 & V

***pol1B* -2.TAL-1.S44 (Mean depth = 396)**



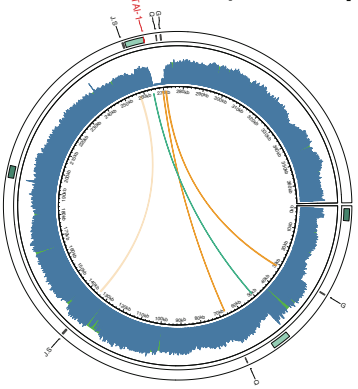
***pol1B* -1.TAL-1.S14 (Mean depth = 1088)**



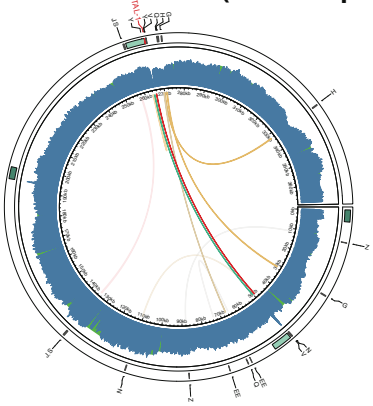
HR via repeats Large1 & V

HR via repeats Large1 x G & Q

***pol1B* -2.TAL-1.S17 (Mean depth = 467)**



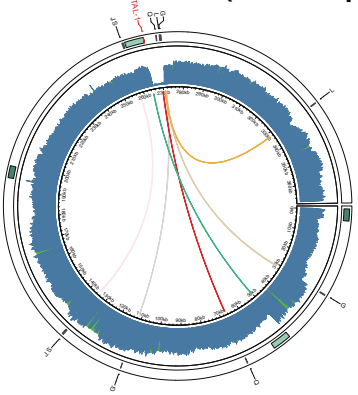
***pol1B* -1.TAL-1.S15 (Mean depth = 1439)**



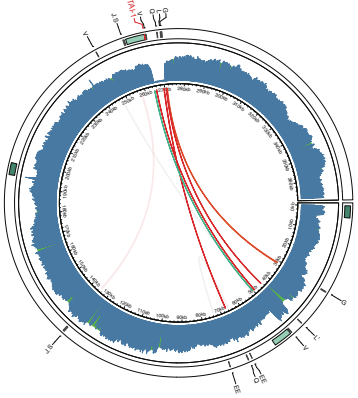
HR via repeats Large1 x V

HR via repeats Large1 x G & L

***pol1B* -2.TAL-1.S18 (Mean depth = 523)**



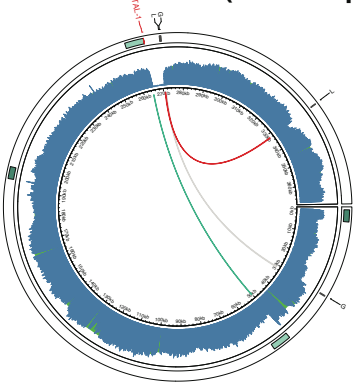
***pol1B* -1.TAL-1.S16 (Mean depth = 1026)**



HR via repeats Large1 x G/L'

HR via repeats Large1 x G & L

***pol1B* -2.TAL-1.S19 (Mean depth = 528)**



Fraction

1
0.8
0.6
0.4
0.2
0

Large1 repeat

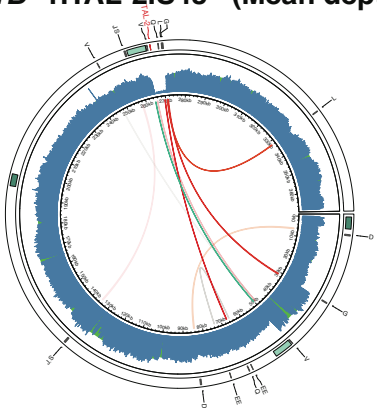
Large2 repeat

Large1 HR

MHMR

pol1B-2 TAL-2

pol1B **-1.TAL-2.S43 (Mean depth = 412)**

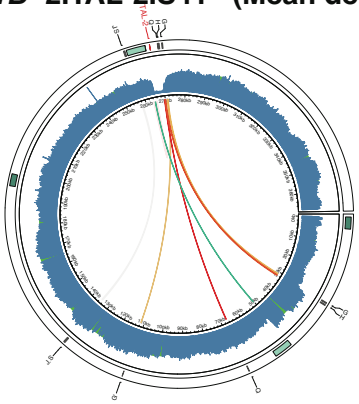


HR via repeats Large1 x G

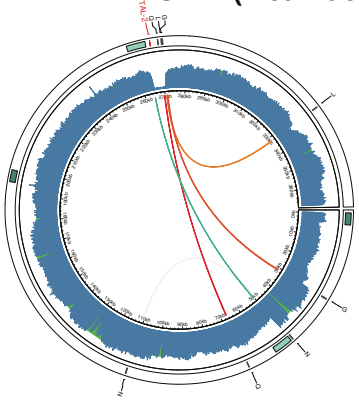
DSB Repaired by

HR via repeats Large1 x G/H

pol1B **-2.TAL-2.S41 (Mean depth = 376)**



pol1B **-2.TAL-2.S42 (Mean depth = 424)**



HR via repeats Large1 x G & L

Fraction

1
0.8
0.6
0.4
0.2
0

Large1 repeat

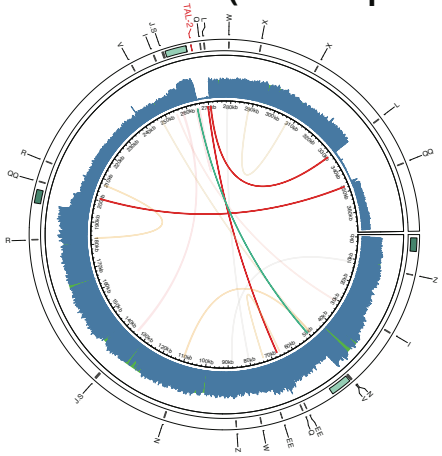
Large2 repeat

Large1 HR

MHMR

recA3 TAL-2

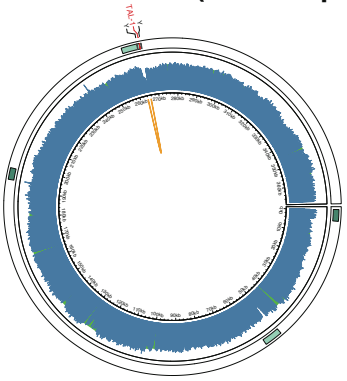
recA3**.TAL-2.S44 (Mean depth = 746)**



Segregation via repeats Large1 & L

odb1-1 TAL-1

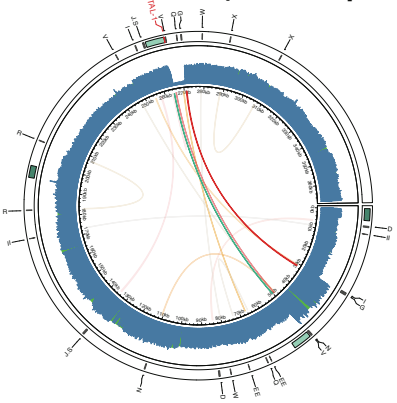
***odb*1–1.TAL-1.S6 (Mean depth = 655)**



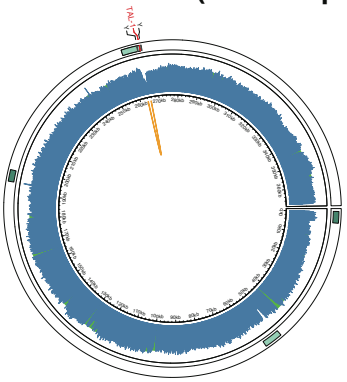
SSA via repeats Y

HR via repeats Large1 x V/G

***odb*1–1.TAL-1.S7 (Mean depth = 854)**



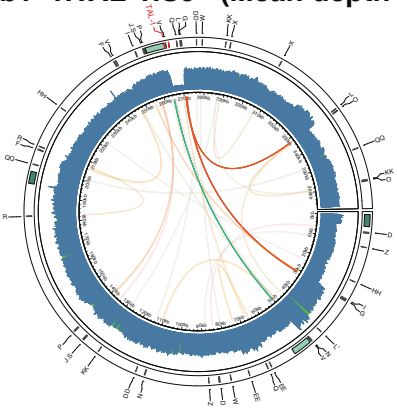
***odb*1–1.TAL-1.S8 (Mean depth = 471)**



SSA via repeats Y

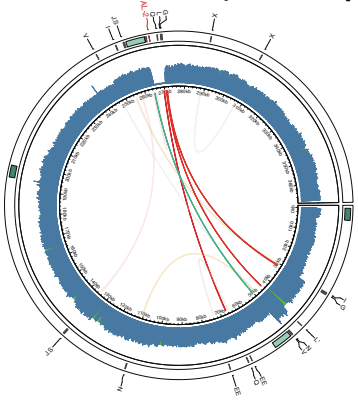
HR via repeats Large1 x G & L

***odb*1–1.TAL-1.S9 (Mean depth = 1548)**



odb1-1 TAL-2

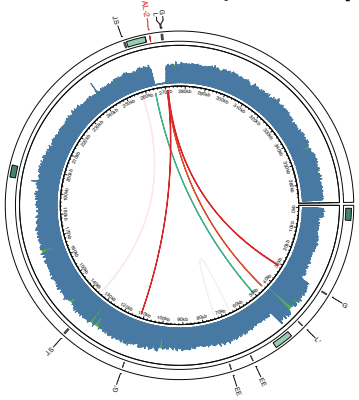
***odb*1–1.TAL-2.S47 (Mean depth = 570)**



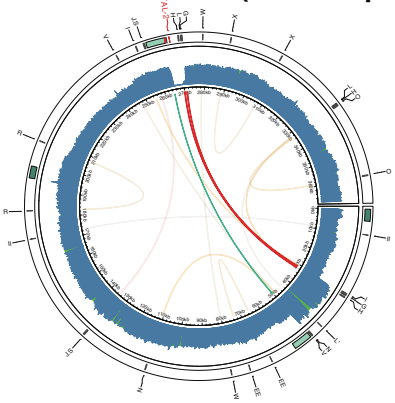
HR via repeats Large1 x G/L'

HR via repeats Large1 x G

***odb1*–1.TAL-2.S48 (Mean depth = 481)**



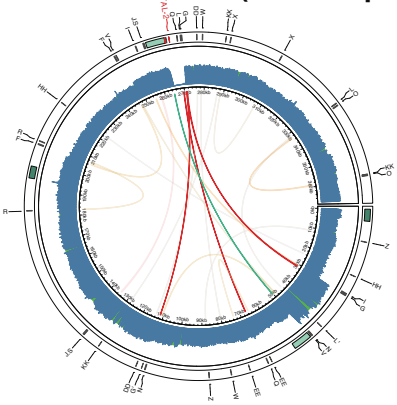
odb1–1.TAL-2.S49 (Mean depth = 727)



HR via repeats Large1 x H/G

HR via repeats Large1 x G

***odb*1–1.TAL-2.S50 (Mean depth = 871)**



Fraction

1
0.8
0.6
0.4
0.2
0

Large1 repeat

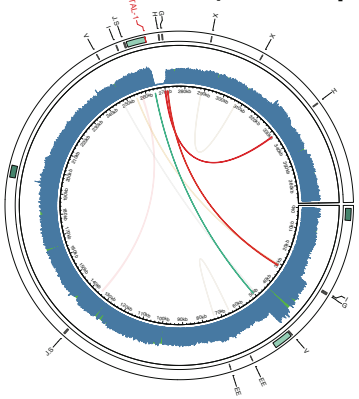
Large2 repeat

Large1 HR

MHMR

why2-1 TAL-1

***why 2*–1.TAL-1.S1 (Mean depth = 607)**

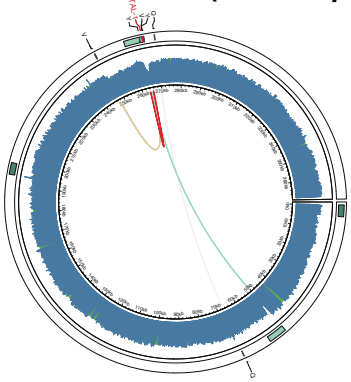


DSB Repaired by

HR via

repeats Large1 x H/G

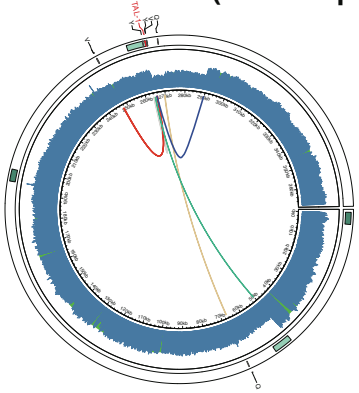
***why 2*–1.TAL-1.S2 (Mean depth = 766)**



SSA & HR via

repeats Y & Large1 & V

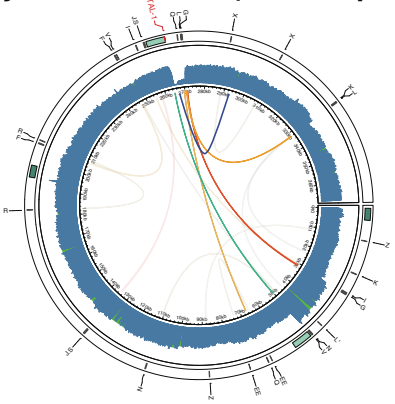
***why 2*–1.TAL-1.S3 (Mean depth = 619)**



MHMR & SSA via

repeats Y

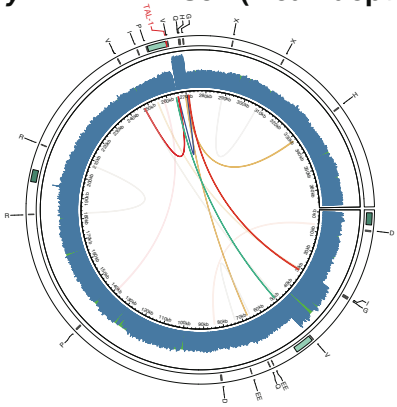
***why 2*–1.TAL-1.S4 (Mean depth = 1089)**



MHMR & HR via

repeats Large1 x G & L

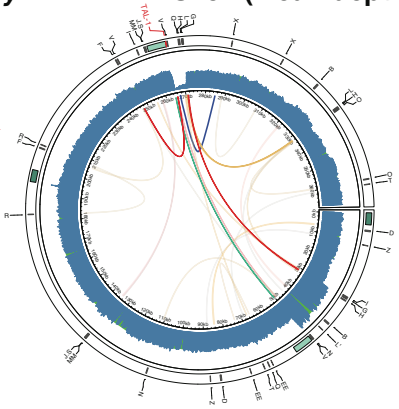
***why 2* –1.TAL-1.S5 (Mean depth = 996)**



MHMR & HR via

repeats Large1 x G & L

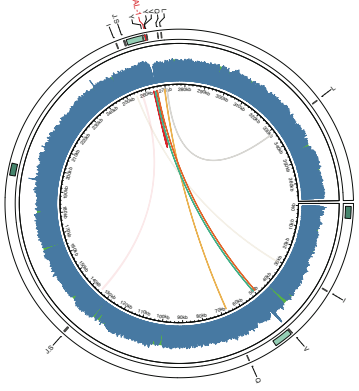
***why 2* –1.TAL-1.S10 (Mean depth = 1298)**



MHMR & HR via

repeats Large1 x G & V

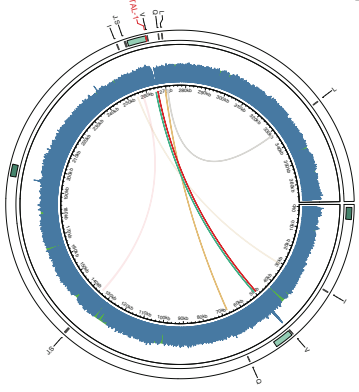
***why 2*–1.TAL-1.S11 (Mean depth = 802)**



SSA & HR via

repeats Y & Large1 x V

***why 2*–1.TAL-1.S12 (Mean depth = 543)**



HR via

repeats Large1 x V

Fraction

1
0.8
0.6
0.4
0.2
0

Large1 repeat

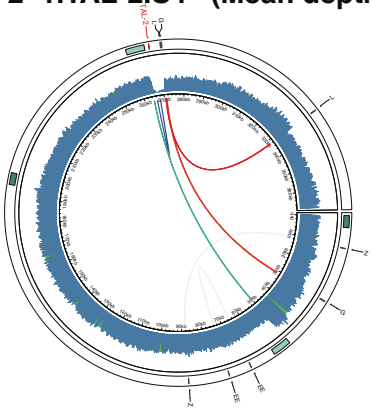
Large2 repeat

Large1 HR

MHMR

why2-1 TAL-2

why 2–1.TAL-2.S4 (Mean depth = 209)

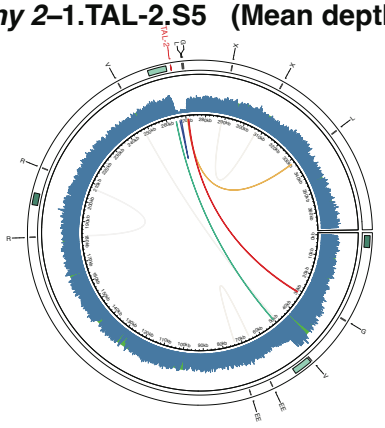


MHMR & HR via

repeats Large1 x G & L

DSB Repaired by

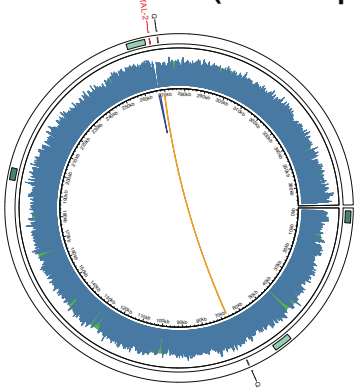
why 2–1.TAL-2.S5 (Mean depth = 171)



MHMR & HR via

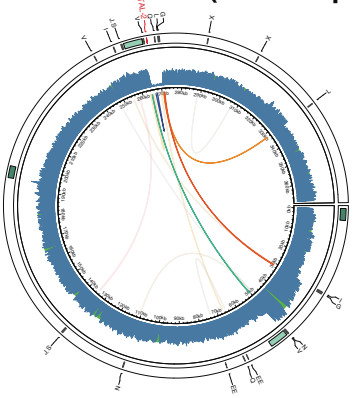
repeats Large1 & G & L

why 2–1.TAL-2.S6 (Mean depth = 131)



MHMR

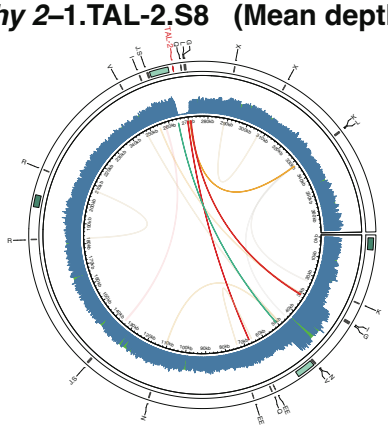
why 2–1.TAL-2.S7 (Mean depth = 187)



MHMR & HR via

repeats Large1 & G & L

why 2–1.TAL-2.S8 (Mean depth = 205)



HR via

repeats Large1 & G & L

Fraction

1
0.8
0.6
0.4
0.2
0

Large1 repeat

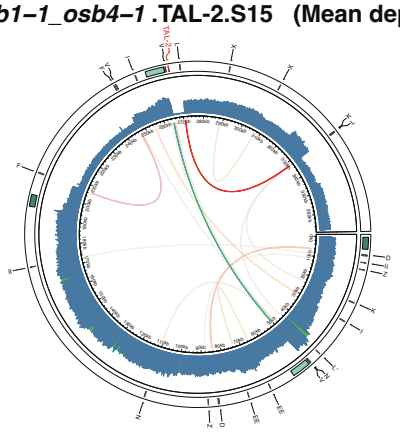
Large2 repeat

Large1 HR

MHMR

osb1-1 osb4-1 TAL-2

osb1–1_osb4–1 **.TAL-2.S15 (Mean depth = 522)**



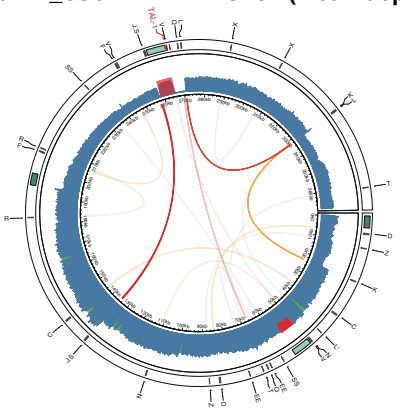
DSB Repaired by

Segregation via

repeats Large1 & L

osb1-1 osb4-1 TAL-1

osb1–1_osb4–1**.TAL-1.S16 (Mean depth = 761)**



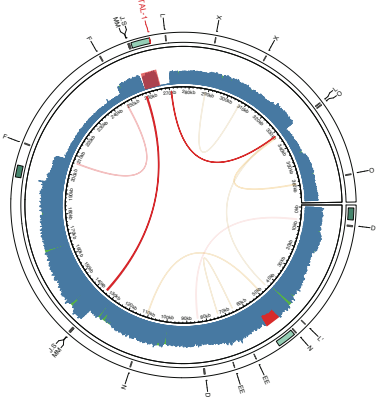
Segregation via

repeats J.S & L

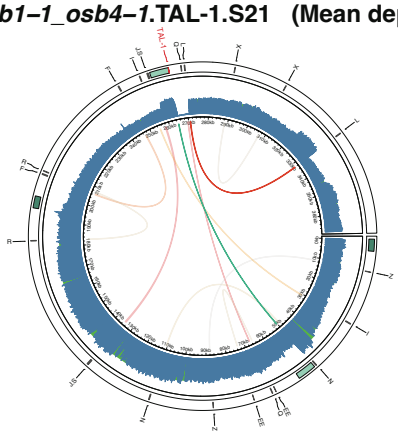
osb1–1_osb4–1**.TAL-1.S20 (Mean depth = 846)**

Segregation via

repeats J.S & L



osb1–1_osb4–1**.TAL-1.S21 (Mean depth = 852)**



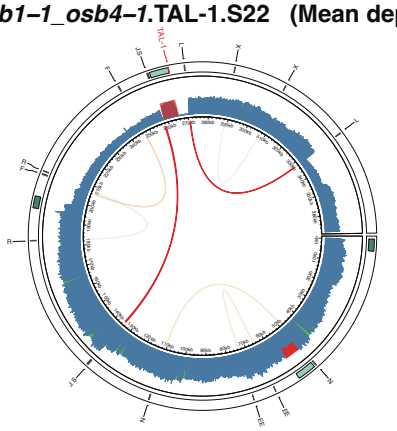
Segregation via

repeats Large1 & L

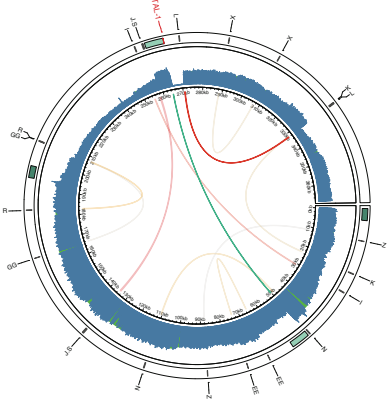
osb1–1_osb4–1**.TAL-1.S22 (Mean depth = 701)**

Segregation via

repeats J.S & L



osb1–1_osb4–1**.TAL-1.S23 (Mean depth = 569)**



Segregation via

repeats Large1 & L

Fraction

1
0.8
0.6
0.4
0.2
0

Large1 repeat

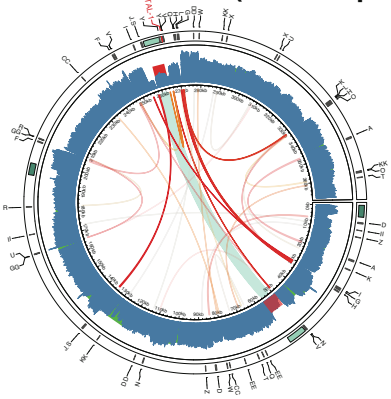
Large2 repeat

Large1 HR

MHMR

msh1 TAL-1

msh1–1 **.TAL-1.S5 (Mean depth = 636)**

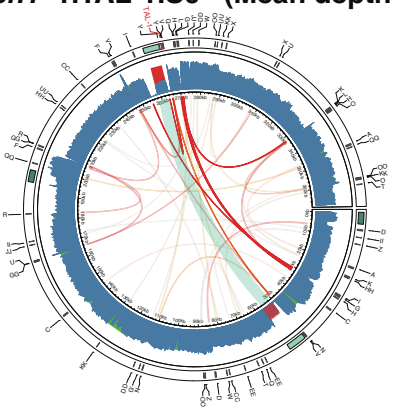


DSB Repaired by

SSA via repeats Y

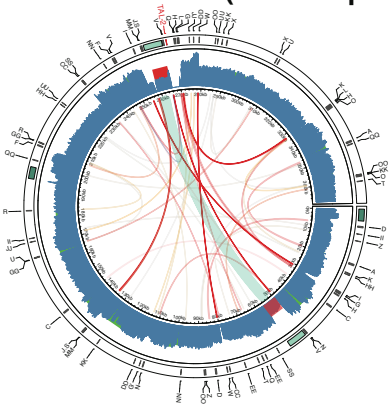
SSA via repeats Y

***msh1*–1.TAL-1.S6 (Mean depth = 1159)**



msh1 TAL-2

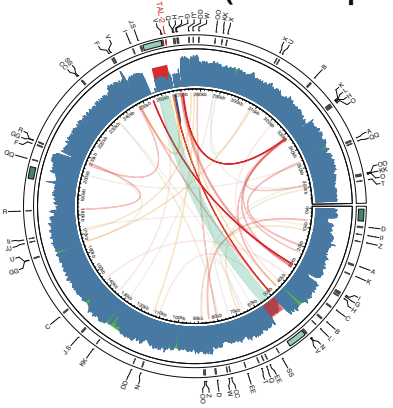
***msh1*–1.TAL-2.S2 (Mean depth = 2241)**



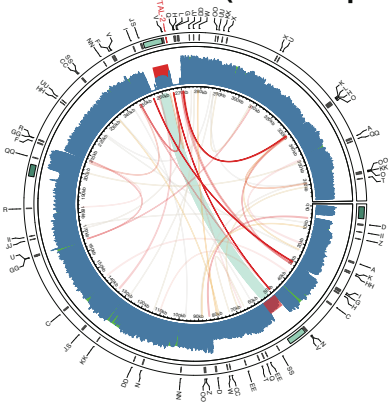
Segregation via repeats J.S/I & L

MHMR & Segregation via repeats I & L

***msh1*–1.TAL-2.S3 (Mean depth = 1777)**



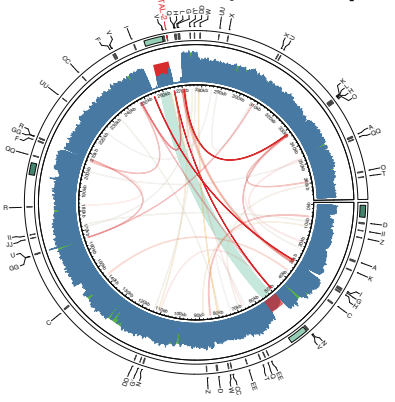
***msh1*–1.TAL-2.S4 (Mean depth = 1535)**



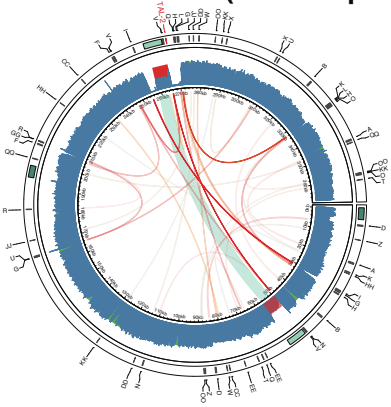
Segregation via repeats I & L

Segregation via repeats I & L

***msh1*–1.TAL-2.S7 (Mean depth = 688)**



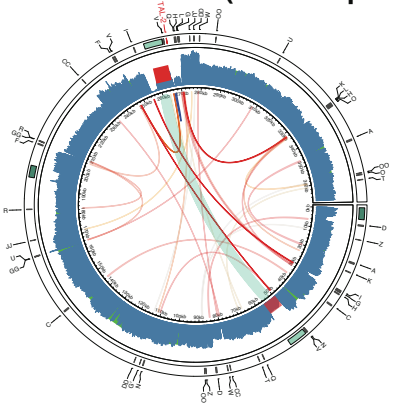
***msh1*–1.TAL-2.S8 (Mean depth = 614)**



Segregation via repeats I & L

MHMR & Segregation via repeats I & L

***msh1*–1.TAL-2.S9 (Mean depth = 471)**



Fraction

1
0.8
0.6
0.4
0.2
0

Large1 repeat

Large2 repeat

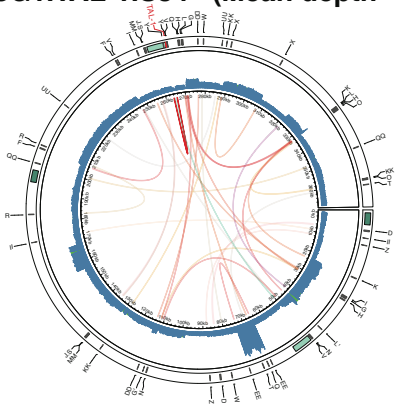
Large1 HR

MHMR

recG1 TAL-1

DSB Repaired by

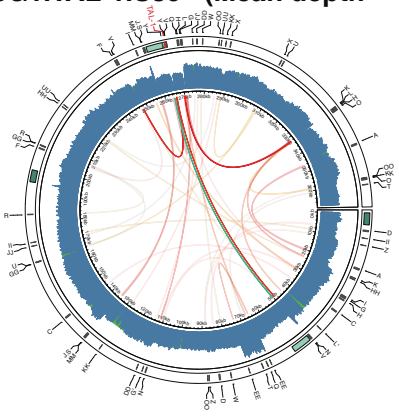
***recG1*.TAL-1.S34 (Mean depth = 726)**



SSA via

repeats Y

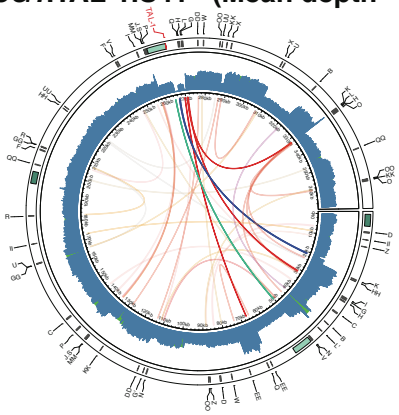
***recG1*.TAL-1.S39 (Mean depth = 1322)**



HR via

repeats Large1 x V

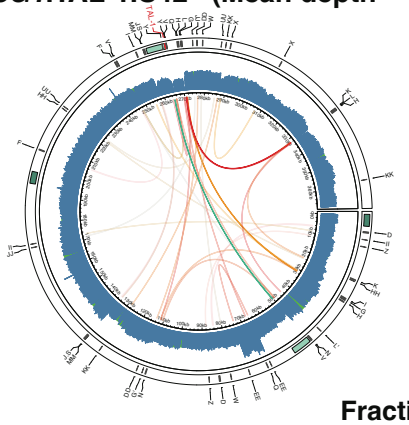
***recG1*.TAL-1.S41 (Mean depth = 968)**



MHMR & HR via

repeats Large1 x G

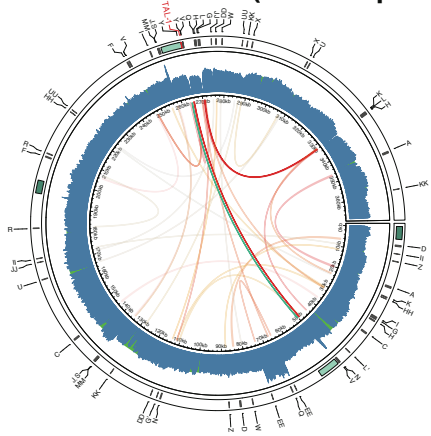
***recG1*.TAL-1.S42 (Mean depth = 664)**



SSA & HR via

repeats Y & Large1 x G

***recG1*.TAL-1.S39-C9 (Mean depth = 748)**



HR via

repeats Large1 x V

Fraction

1
0.8
0.6
0.4
0.2
0

Large1 repeat

Large2 repeat

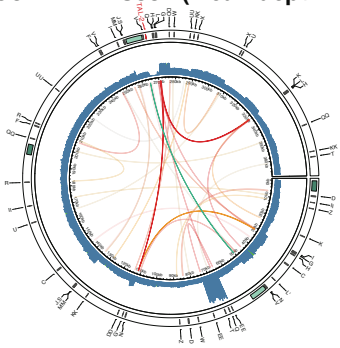
Large1 HR

MHMR

other HR

recG1 TAL-2

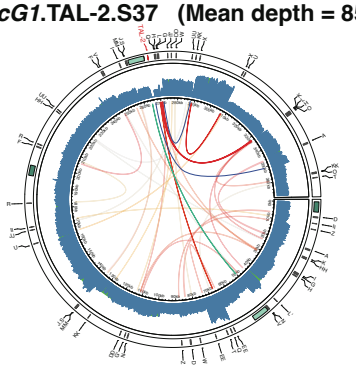
***recG1*.TAL-2.S33 (Mean depth = 633)**



HR via

repeats Large1 x H & G'

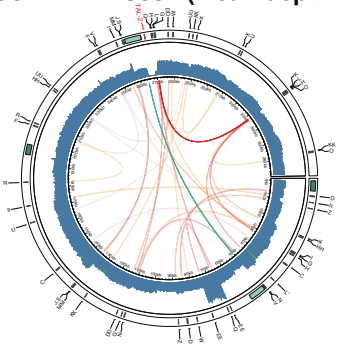
***recG1*.TAL-2.S37 (Mean depth = 853)**



MHMR & HR via

repeats Large1 & L & H

***recG1*.TAL-2.S38 (Mean depth = 792)**

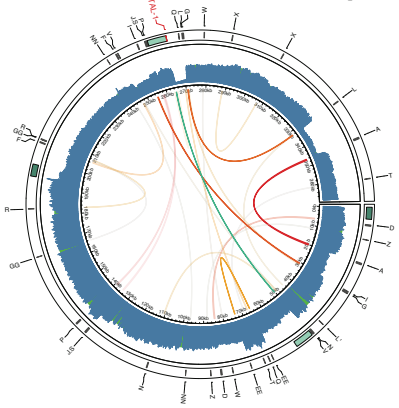


MHMR & HR via

repeats Large1 & L

redA-3 TAL-1

***radA–3*.TAL-1.S20 (Mean depth = 953)(F1)**



DSB Repaired by

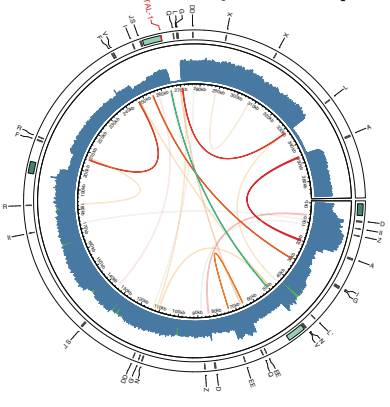
Segregation via

repeats I & L & Large1

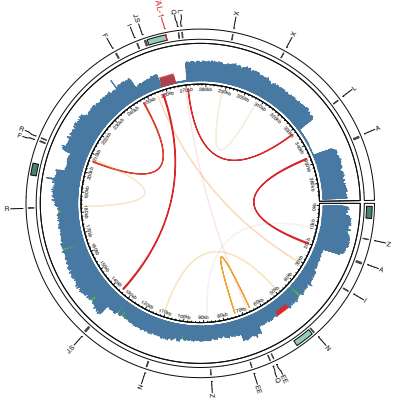
Segregation via

repeats I & L & Large1

***radA–3*.TAL-1.S21 (Mean depth = 801) (F1)**



***radA–3*.TAL-1.S35 (Mean depth = 736)**



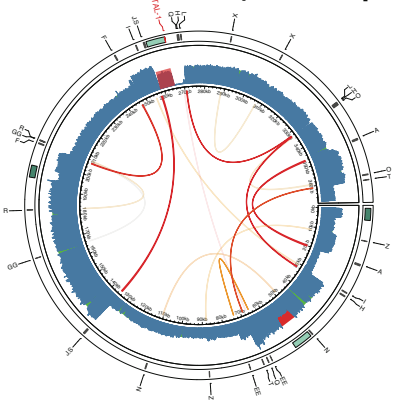
Segregation via

repeats J.S & L

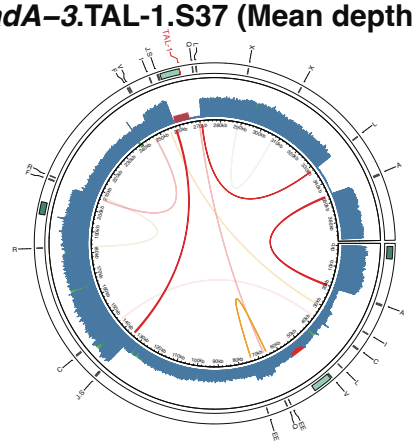
Segregation via

repeats J.S & L

***radA–3*.TAL-1.S36 (Mean depth = 913)**



***radA–3*.TAL-1.S37 (Mean depth = 615)**



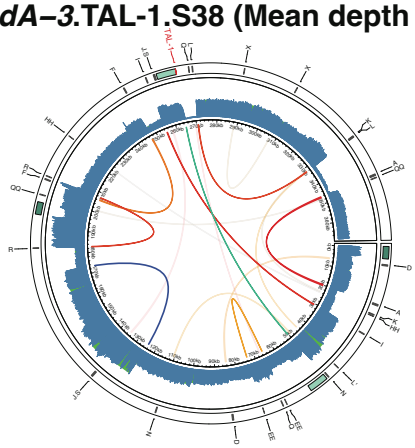
Segregation via

repeats J.S & L

Segregation via

repeats I & L & Large1

***radA–3*.TAL-1.S38 (Mean depth = 666)**



Fraction

1
0.8
0.6
0.4
0.2
0

Large1 repeat

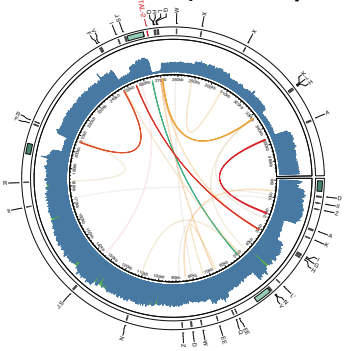
Large2 repeat

Large1 HR

MHMR

redA-3 TAL-2

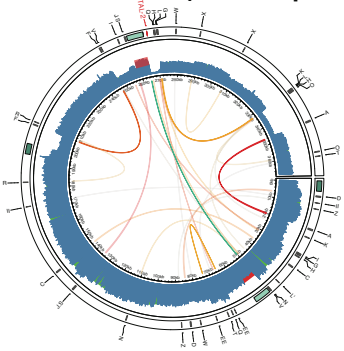
***radA–3*.TAL-2.S17(Mean depth = 819)(F1)**



Segregation via

repeats I & L & Large1

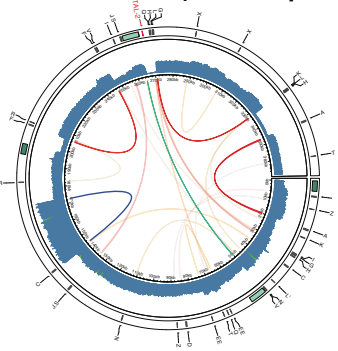
***radA*–3.TAL-2.S18(Mean depth = 1270)(F1)**



Segregation via

repeats L & Large1

***radA–3*.TAL-2.S19(Mean depth = 1022)(F1)**



Segregation via

repeats L & Large1

oex1$^{-/+}$ T1 x *oex1*$^{-/+}$

F1

TAL-1

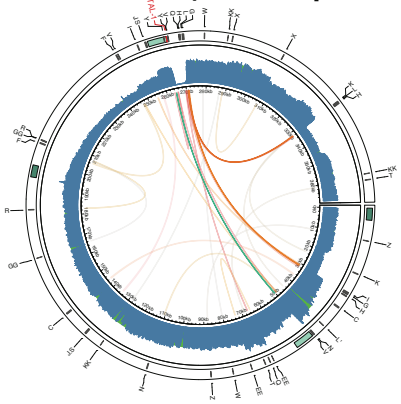
TAL-2

DSB Repaired by

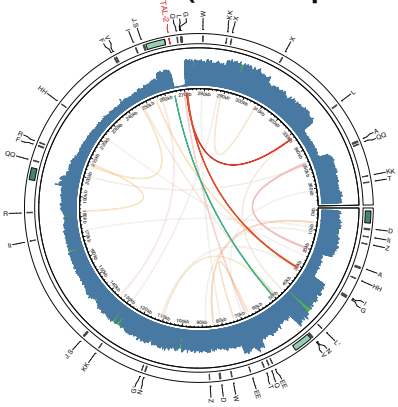
HR via repeats Large1 x G & L

HR via repeats Large1 x G & L

***oex1*.TAL-1.S33 (Mean depth = 612)**



***oex1*.TAL-2.S35 (Mean depth = 643)**



Fraction

1
0.8
0.6
0.4
0.2
0

Large1 repeat

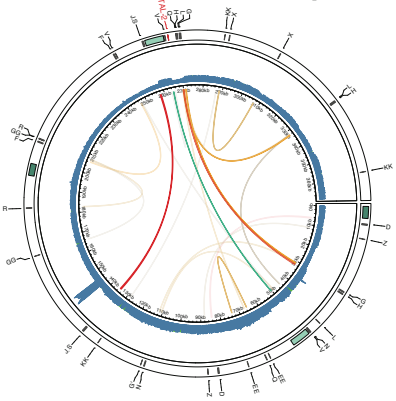
Large2 repeat

Large1 HR

MHMR

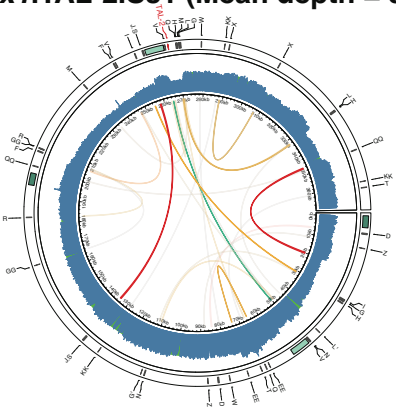
Heteroplasmic TAL-2

***oex1*.TAL-2.S34 (Mean depth = 761)**

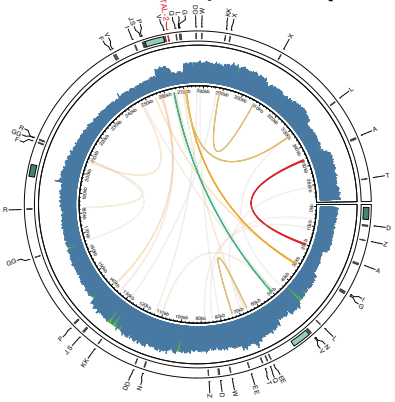


It might be there is no mitotype with a cut site due to low mitoTALEN activity when introduced by cross or it is heteroplasmic

***oex1*.TAL-2.S31 (Mean depth = 819)**



***oex1*.TAL-2.S32 (Mean depth = 721)**

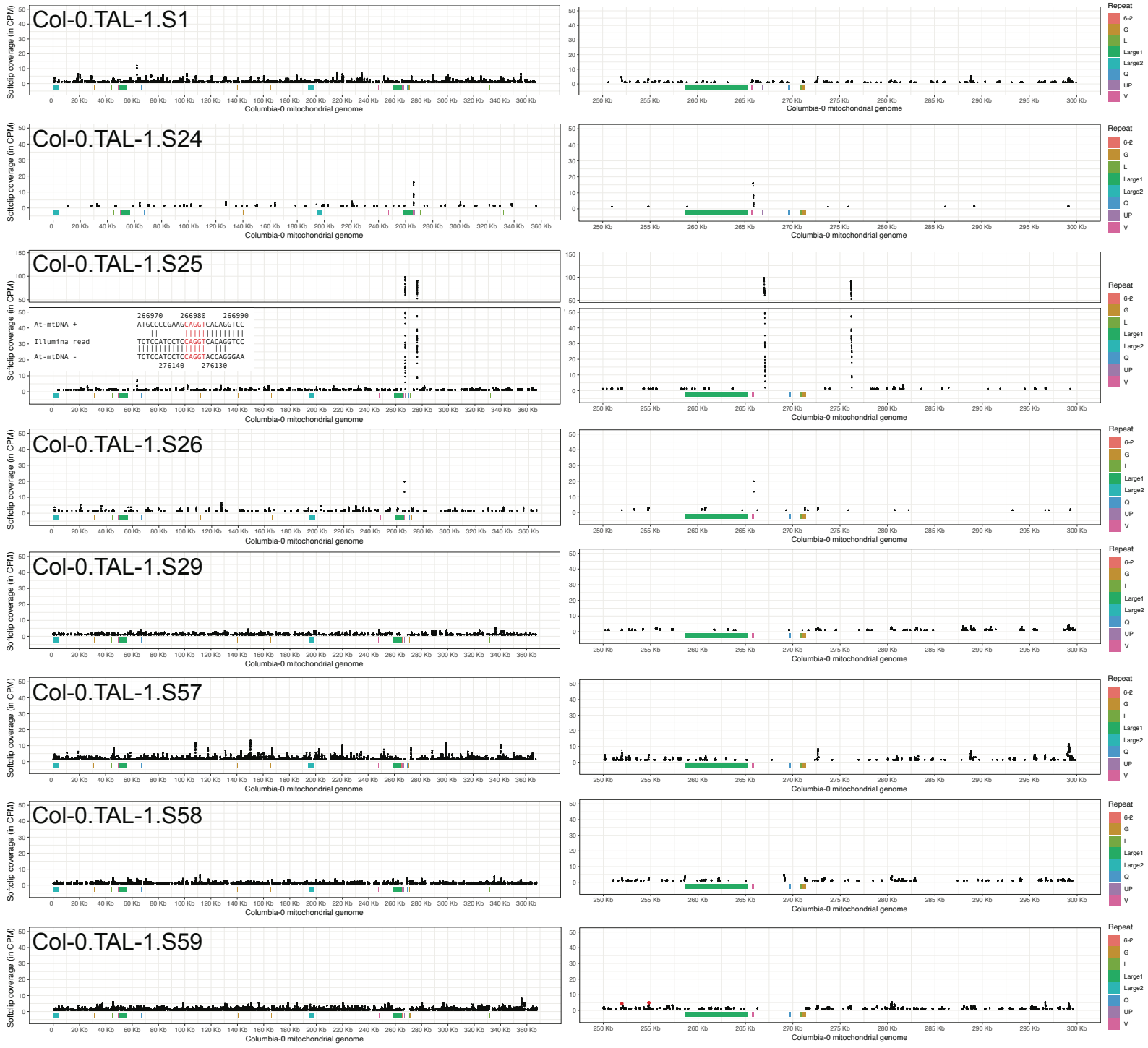


Appendix.4. Repeat windows for calculating reads confirming the corresponding recombination event

Repeat	Type	Copy	Size (bp)	window.start (bp)	window.end (bp)	R1_site_start (bp)	R1_site_end (bp)	R2_site_start (bp)	R2_site_end (bp)
A	direct	1	555	18882	21037	18882	19682	20237	21037
AA	direct	1	222	246684	248506	246684	247484	247706	248506
B	inverted	1	532	40664	42796	40664	41464	41996	42796
C	inverted	1	456	35562	37618	35562	36362	36818	37618
CC	direct	1	147	78423	80170	78423	79223	79370	80170
D	inverted	1	451	5318	7369	5318	6118	6569	7369
DD	inverted	1	127	112864	114591	112864	113664	113791	114591
EE	direct	1	126	64747	66473	64747	65547	65673	66473
F	direct	1	349	205295	207244	205295	206095	206444	207244
G	inverted	1	334	30138	32072	30138	30938	31272	32072
GG	inverted	1	100	165032	166732	165032	165832	165932	166732
H	NA	1	341	30925	32866	30925	31725	32066	32866
HH	inverted	1	101	23051	24752	23051	23851	23952	24752
I	inverted	1	280	29642	31522	29642	30442	30722	31522
II	inverted	1	92	8121	9813	8121	8921	9013	9813
J.S	direct	1	250	134581	135993	134581	134637	135193	135993
JJ	direct	1	114	170886	172600	170886	171686	171800	172600
K	inverted	1	250	22453	24303	22453	23253	23503	24303
KK	NA	1	112	127490	129202	127490	128290	128402	129202
L	direct	1	248	270488	271823	270488	270775	271023	271823
M	inverted	1	257	233372	235229	233372	234172	234429	235229
MM	direct	1	81	133700	134637	133700	134500	134581	134637
MT	NA	1	86	147800	149486	147800	148600	148686	149486
N	inverted	1	286	48099	49985	48099	48899	49185	49985
NN	direct	1	69	95928	97597	95928	96728	96797	97597
O	inverted	1	310	331860	333755	331860	332660	332955	333755
OO	NA	1	90	87645	89335	87645	88445	88535	89335
P	direct	1	464	138110	140174	138110	138910	139374	140174
Q	direct	1	204	66058	67862	66058	66858	67062	67862
QQ	direct	1	238	200001	201839	200001	200801	201039	201839
R	direct	1	205	181929	183734	181929	182729	182934	183734
SS	direct	1	76	60112	61788	60112	60912	60988	61788
T	inverted	1	184	67860	69644	67860	68660	68844	69644
TT	inverted	1	54	237591	239245	237591	238391	238445	239245
U	direct	1	186	165148	166934	165148	165948	166134	166934
UU	inverted	1	62	222669	224331	222669	223469	223531	224331
V	NA	1	175	48443	50218	48443	49243	49418	50218
W	direct	1	183	79377	81160	79377	80177	80360	81160
X	direct	1	203	287515	289318	287515	288315	288518	289318
Z	direct	1	141	10594	12335	10594	11394	11535	12335
A	direct	2	555	345408	347563	345408	346208	346763	347563
AA	direct	2	222	264888	266696	264888	265688	265896	266696
B	inverted	2	532	320635	322767	320635	321435	321967	322767
C	inverted	2	456	143153	145209	143153	143953	144409	145209
CC	direct	2	147	232846	234589	232846	233646	233789	234589
D	inverted	2	451	83289	85340	83289	84089	84540	85340
DD	inverted	2	127	275692	277419	275692	276492	276619	277419
EE	direct	2	126	72811	74537	72811	73611	73737	74537
F	direct	2	349	245966	247915	245966	246766	247115	247915
G	inverted	2	334	270488	272195	270488	271061	271395	272195
GG	inverted	2	100	206425	208132	206425	207225	207332	208132
H	NA	2	341	269064	270486	269064	269864	270199	270486
HH	inverted	2	101	221808	223505	221808	222608	222705	223505
I	inverted	2	280	254042	255922	254042	254842	255122	255922
II	inverted	2	92	172144	173836	172144	172944	173036	173836
J.S	direct	2	250	257579	258943	257579	257620	258143	258943
JJ	direct	2	114	274190	275904	274190	274990	275104	275904
K	inverted	2	250	330781	332631	330781	331581	331831	332631
KK	NA	2	112	285905	287608	285905	286705	286808	287608
L	direct	2	248	331077	332925	331077	331877	332125	332925
M	inverted	2	257	269773	271630	269773	270573	270830	271630
MM	direct	2	81	256711	257620	256711	257511	257579	257620
MT	NA	2	86	306931	308610	306931	307731	307810	308610
N	inverted	2	286	109503	111396	109503	110303	110596	111396
NN	direct	2	69	242528	244197	242528	243328	243397	244197
O	inverted	2	310	358182	360092	358182	358982	359292	360092
OO	NA	2	90	283484	285169	283484	284284	284369	285169
P	direct	2	464	257329	259358	257329	258129	258558	259358
Q	direct	2	204	268770	270575	268770	269570	269775	270575
QQ	direct	2	238	346547	348375	346547	347347	347575	348375
R	direct	2	205	206529	208334	206529	207329	207534	208334
SS	direct	2	76	233245	234924	233245	234045	234124	234924
T	inverted	2	184	360069	361849	360069	360869	361049	361849
TT	inverted	2	54	287819	289468	287819	288619	288668	289468
U	direct	2	186	307132	308918	307132	307932	308118	308918
UU	inverted	2	62	284328	285990	284328	285128	285190	285990
V	NA	2	175	246694	248473	246694	247494	247673	248473
W	direct	2	183	277658	279440	277658	278458	278640	279440
X	direct	2	203	306169	307964	306169	306969	307164	307964
Y	direct	2	174	5051	6661	5051	5769	5943	6661
Z	direct	2	141	87137	88878	87137	87937	88078	88878
G	NA	3	334	110840	112524	110840	111640	111724	112524
H	NA	3	341	331381	333322	331381	332181	332522	333322
KK	NA	3	112	357922	359622	357922	358722	358822	359622
L	NA	3	248	43816	45556	43816	44616	44756	45556
OO	NA	3	90	357724	359413	357724	358524	358613	359413
V	NA	3	175	264898	266667	264898	265698	265867	266667
Y	direct	3	174	6661	8271	6661	7379	7553	8271
G	NA	4	334	139530	141085	139530	140306	140363	141085
G	NA	5	334	164805	166450	164805	165605	165650	166450

Appendix.5. Coverage plots of NGS analysis using soft-clips to find MHMR events in WT and mutant plants produced in this project

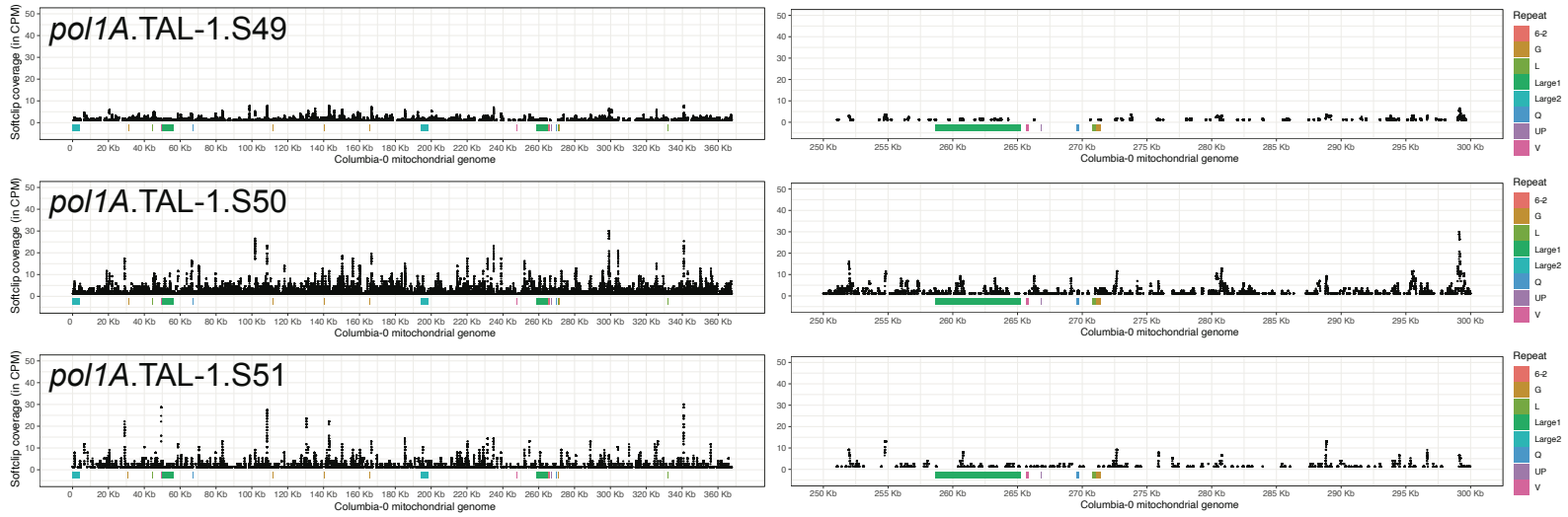
Col-0 TAL-1



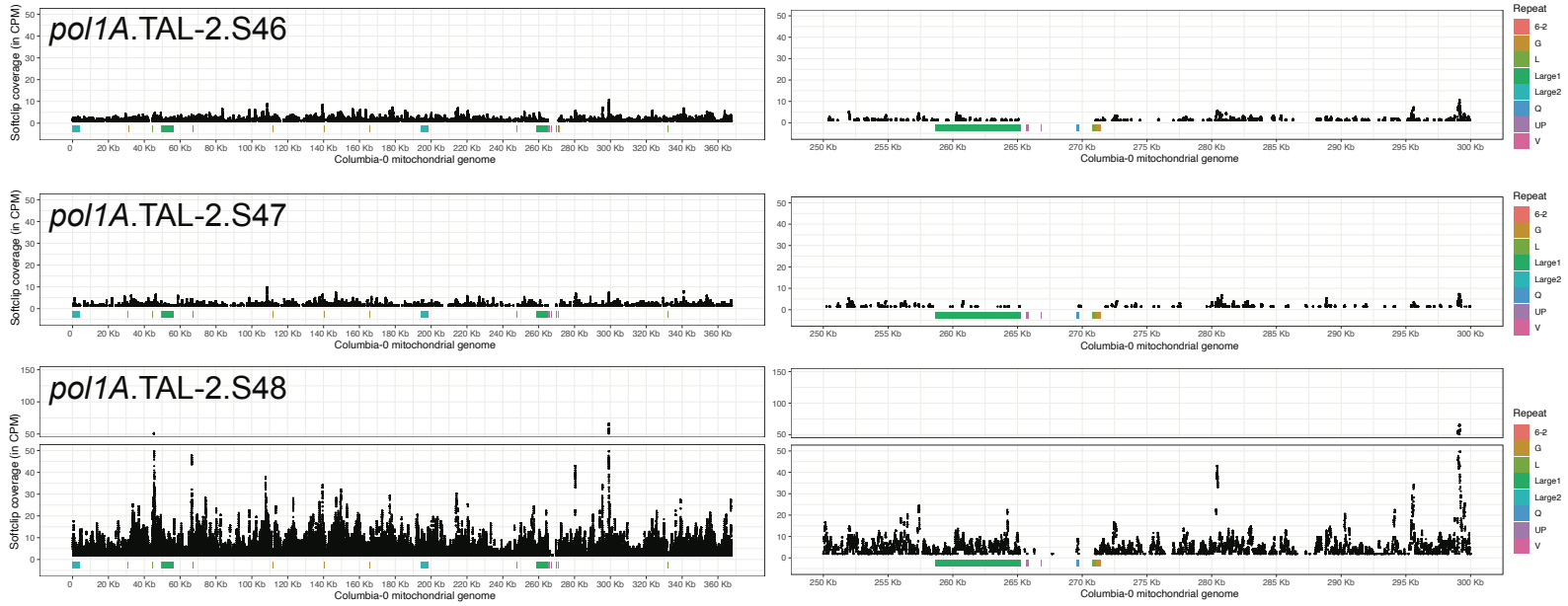
Col-0 TAL-2


Col-0.TAL-2.S1
Col-0.TAL-2.S2
Col-0.TAL-2.S3
Col-0.TAL-2.S27
Col-0.TAL-2.S28
Col-0.TAL-2.S55
Col-0.TAL-2.S56
Softclip coverage (in CPM)
Columbia-0 mitochondrial genome
Repeat
6-2
G
L
Large1
Large2
Q
UP
V

pol1A TAL-1



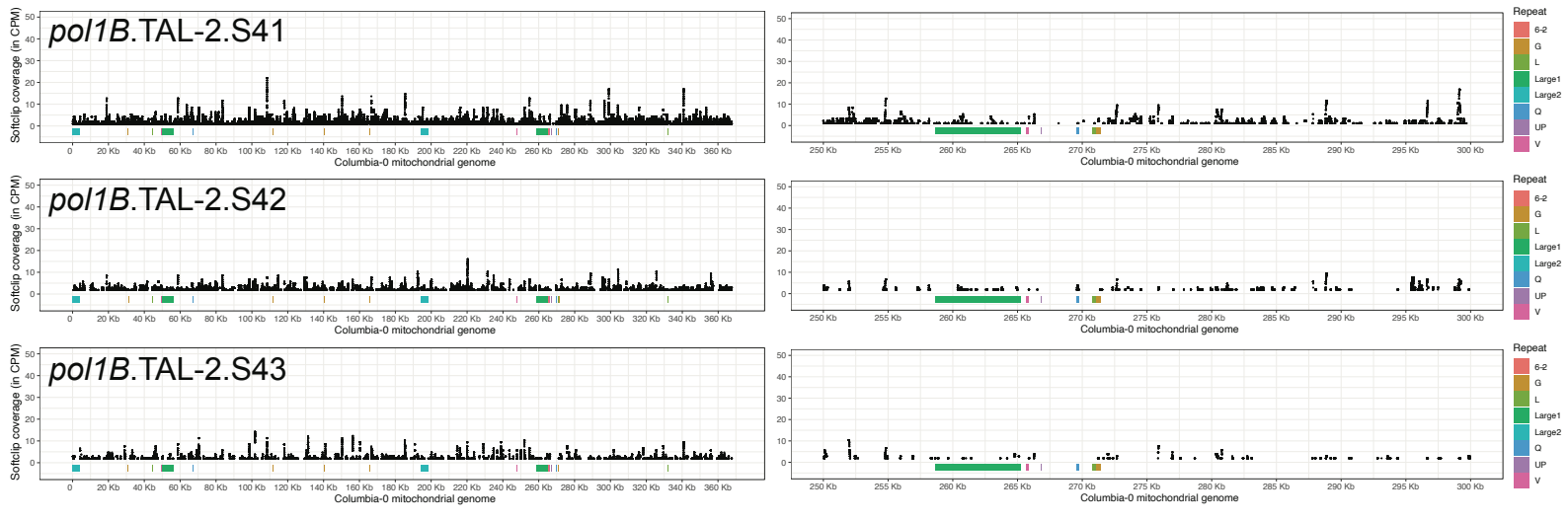
pol1A TAL-2



pol1B TAL-1



pol1B TAL-2



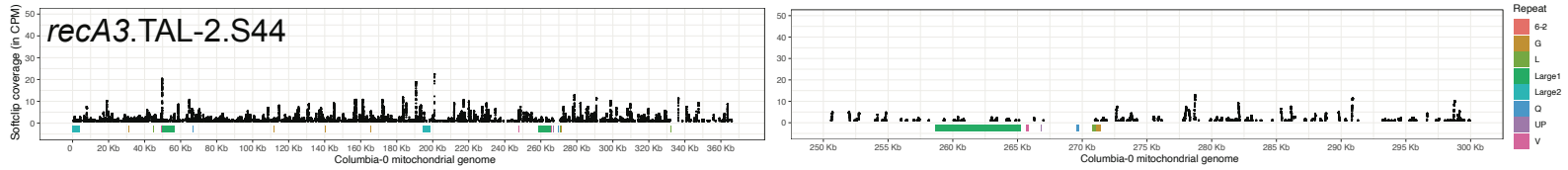
odb1 TAL-1



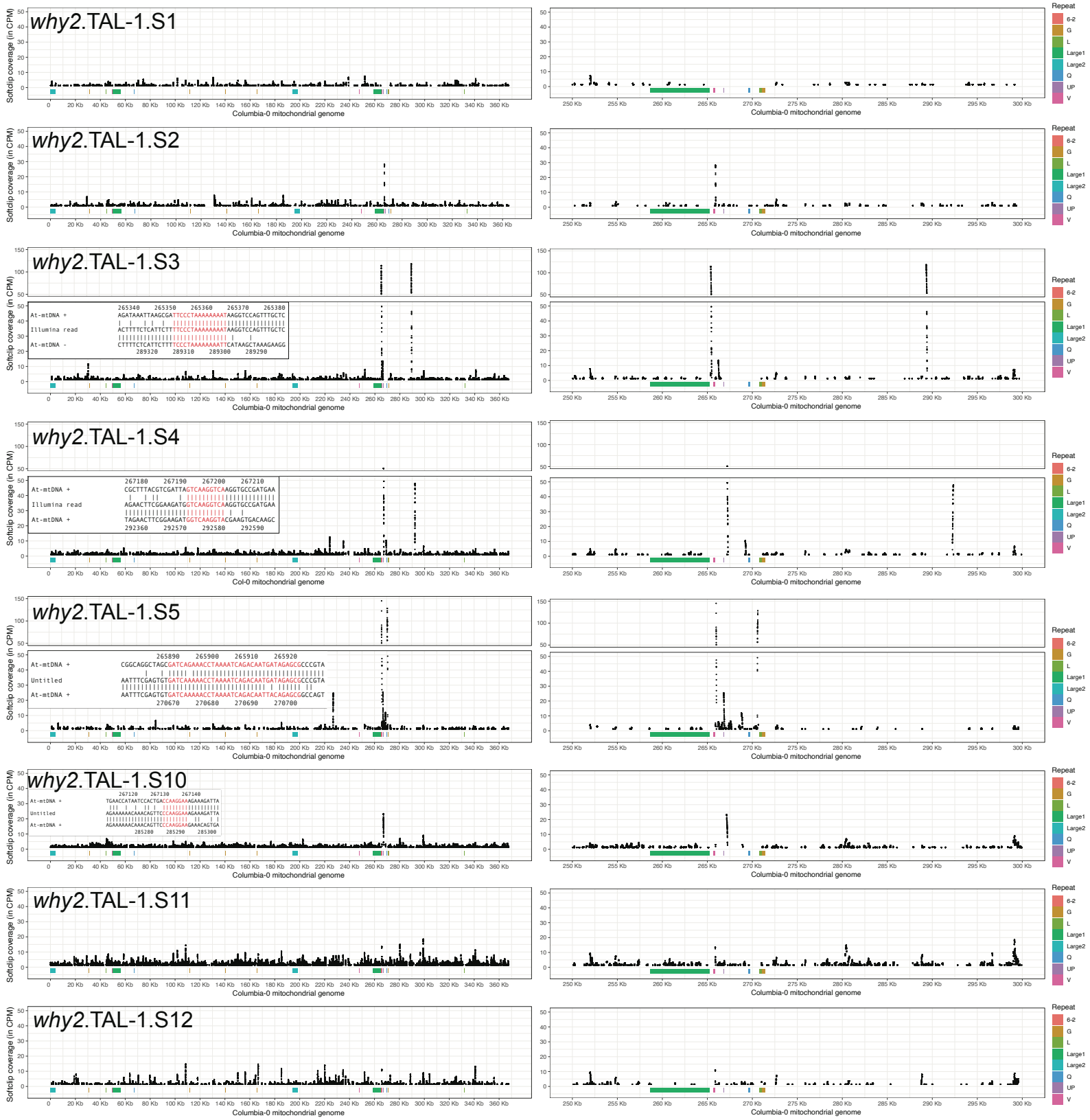
odb1 TAL-2



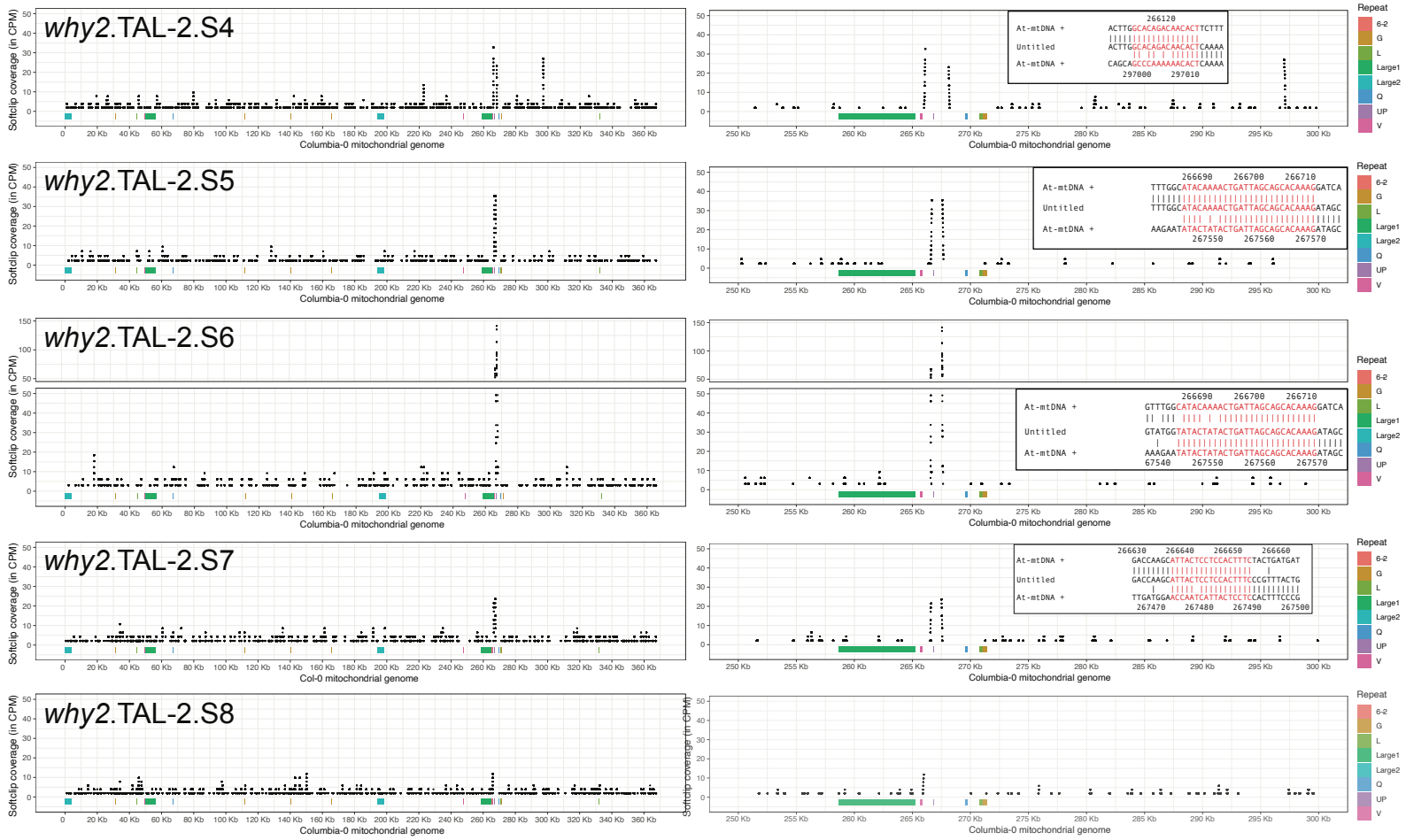
recA3 TAL-2



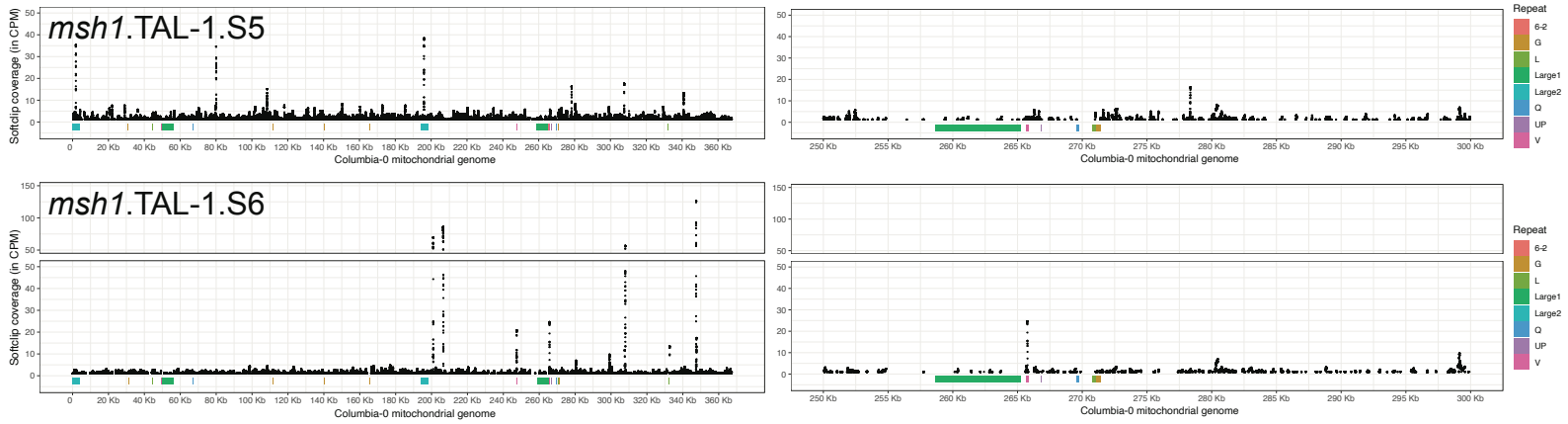
why2 TAL-1


why2.TAL-1.S1
why2.TAL-1.S2
why2.TAL-1.S3
why2.TAL-1.S4
why2.TAL-1.S5
why2.TAL-1.S10
why2.TAL-1.S11
why2.TAL-1.S12
Softclip coverage (in CPM)
Columbia-0 mitochondrial genome
Col-0 mitochondrial genome
Repeat
6-2
G
L
Large1
Large2
Q
UP
V
At-mtDNA +
Illumina read
At-mtDNA -
Untitled

why2 TAL-2


why2.TAL-2.S4
why2.TAL-2.S5
why2.TAL-2.S6
why2.TAL-2.S7
why2.TAL-2.S8
Softclip coverage (in CPM)
Columbia-0 mitochondrial genome
Col-0 mitochondrial genome
Repeat
6-2
G
L
Large1
Large2
Q
UP
V

msh1 TAL-1



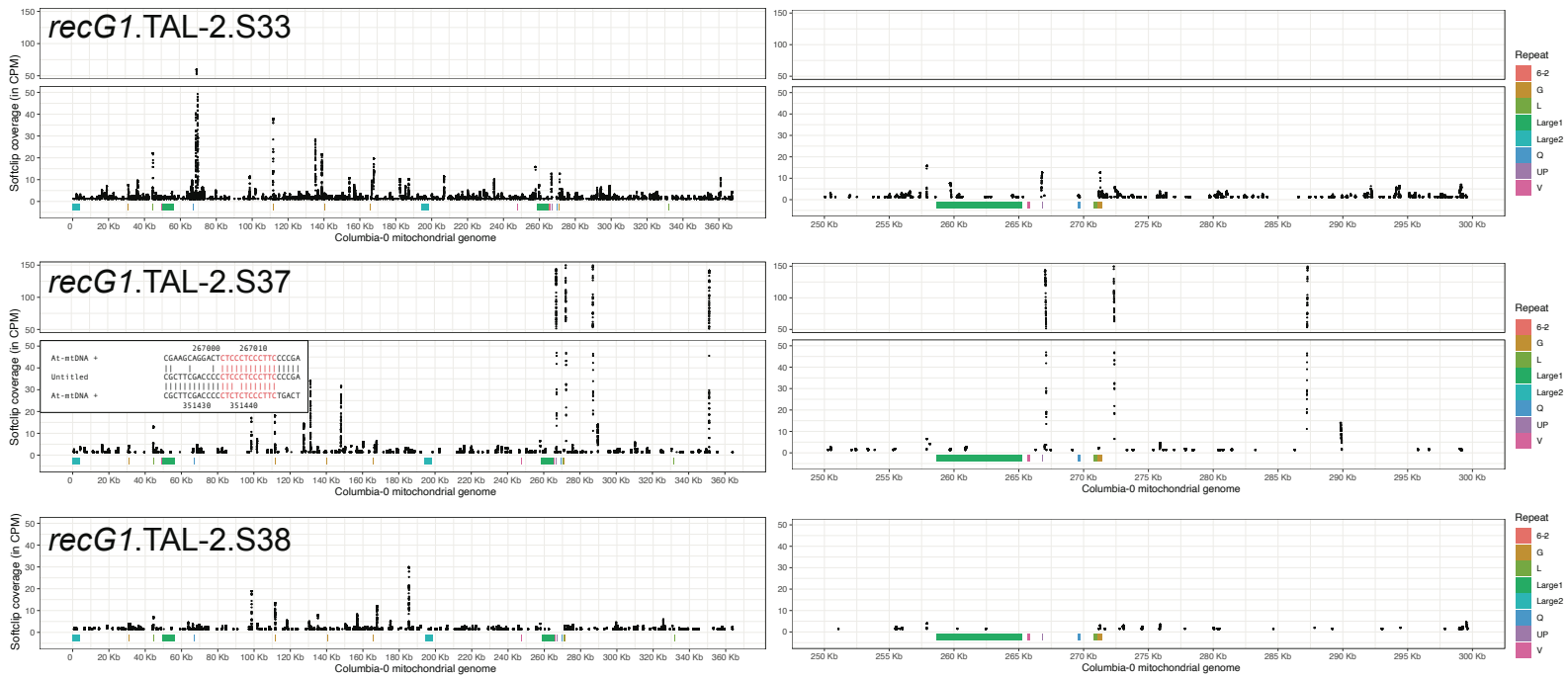
msh1 TAL-2



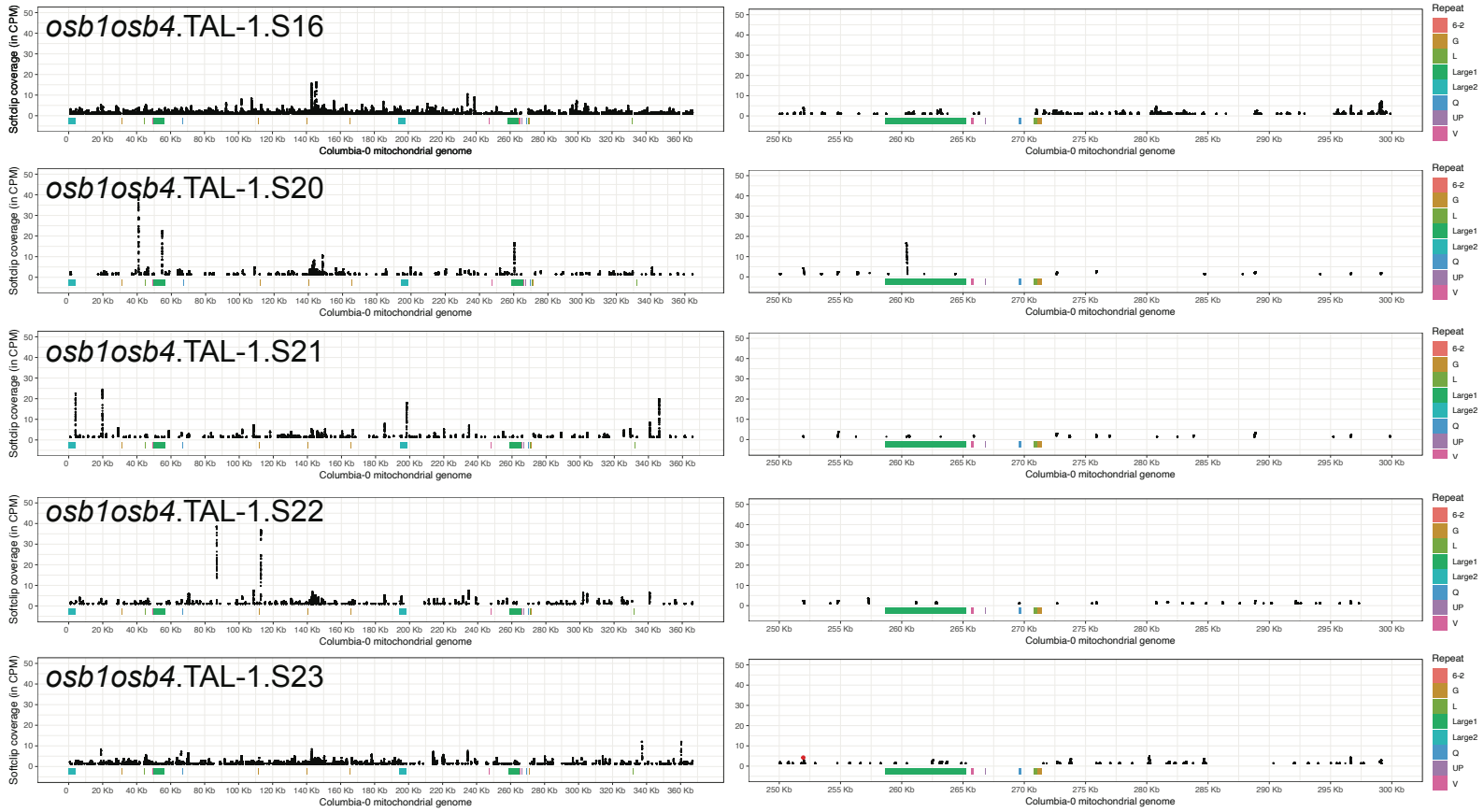
recG1 TAL-1



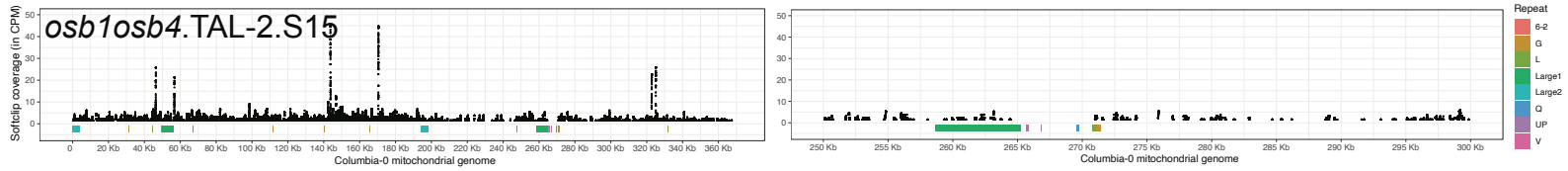
recG1 TAL-2



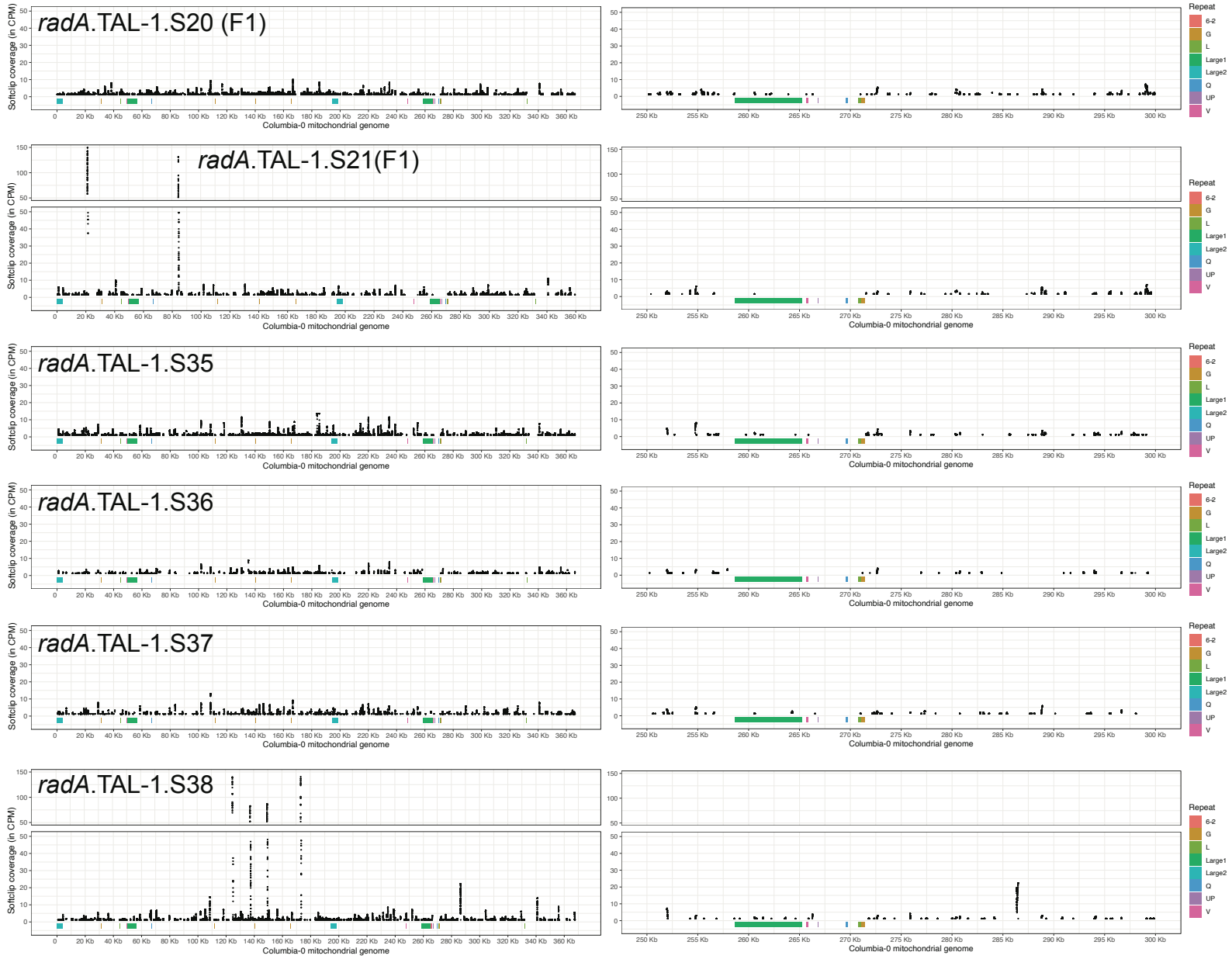
osb1osb4 TAL-1



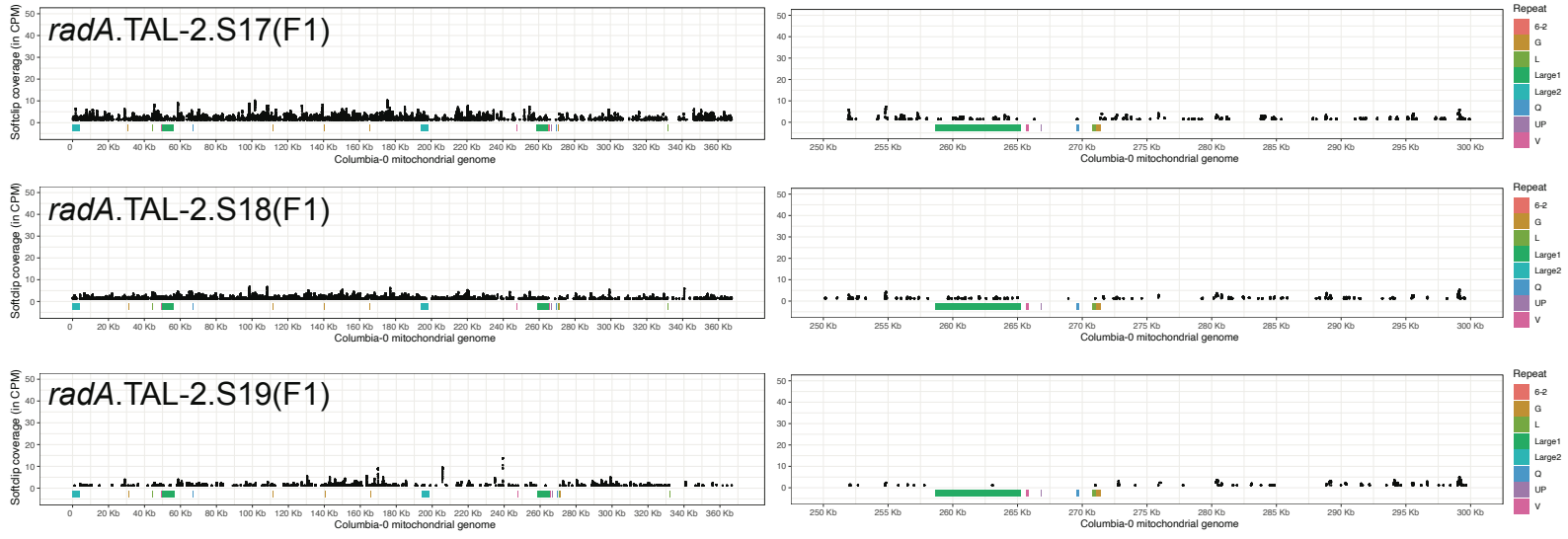
osb1osb4 TAL-2



radA TAL-1

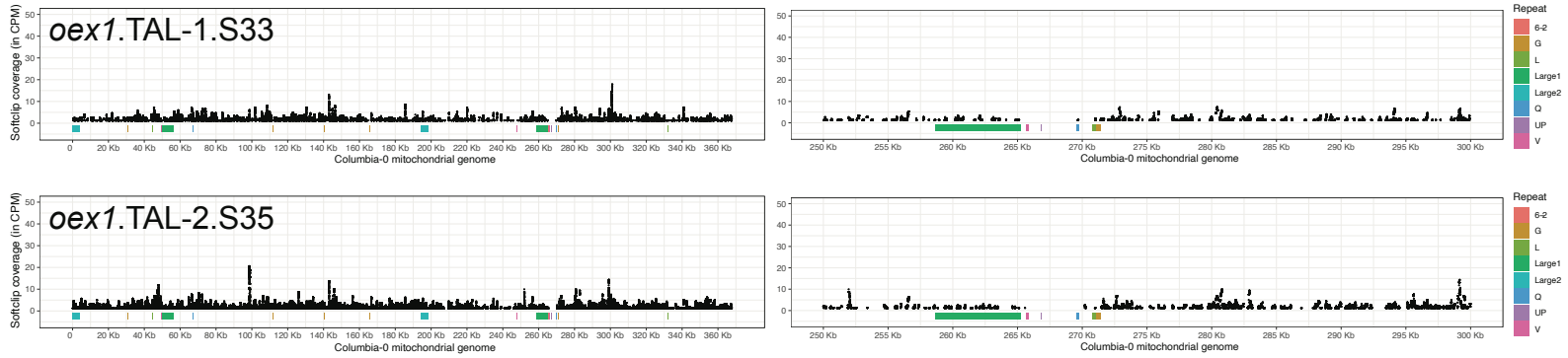


radA TAL-2

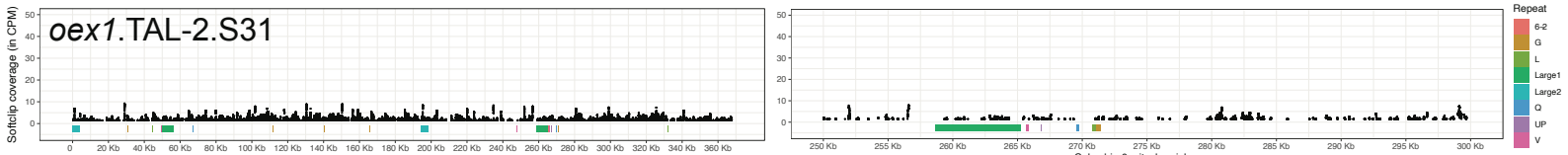


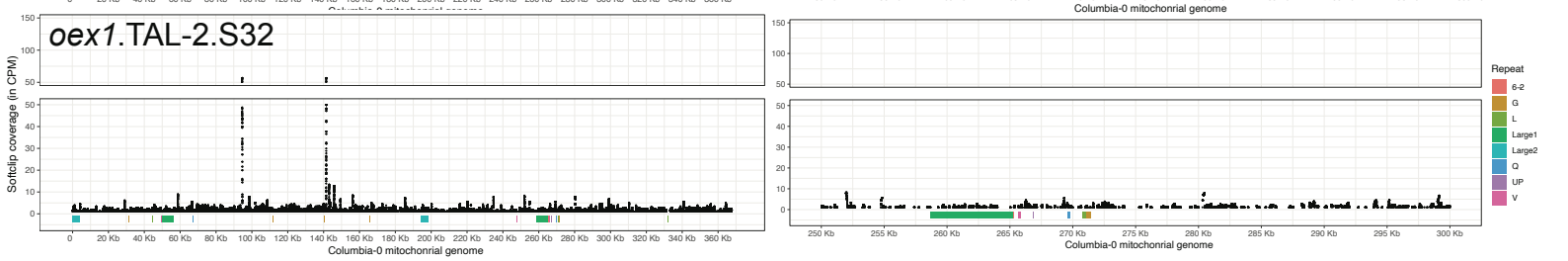
oex1$^{-/+}$ T1 x *oex1*$^{-/+}$

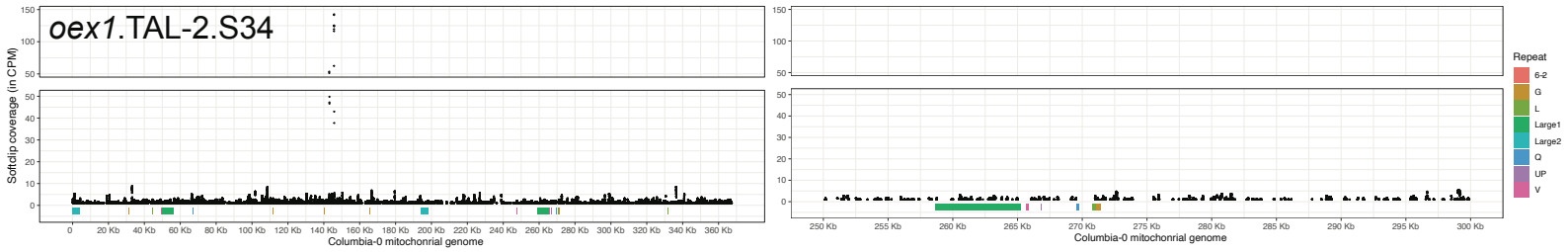
F1



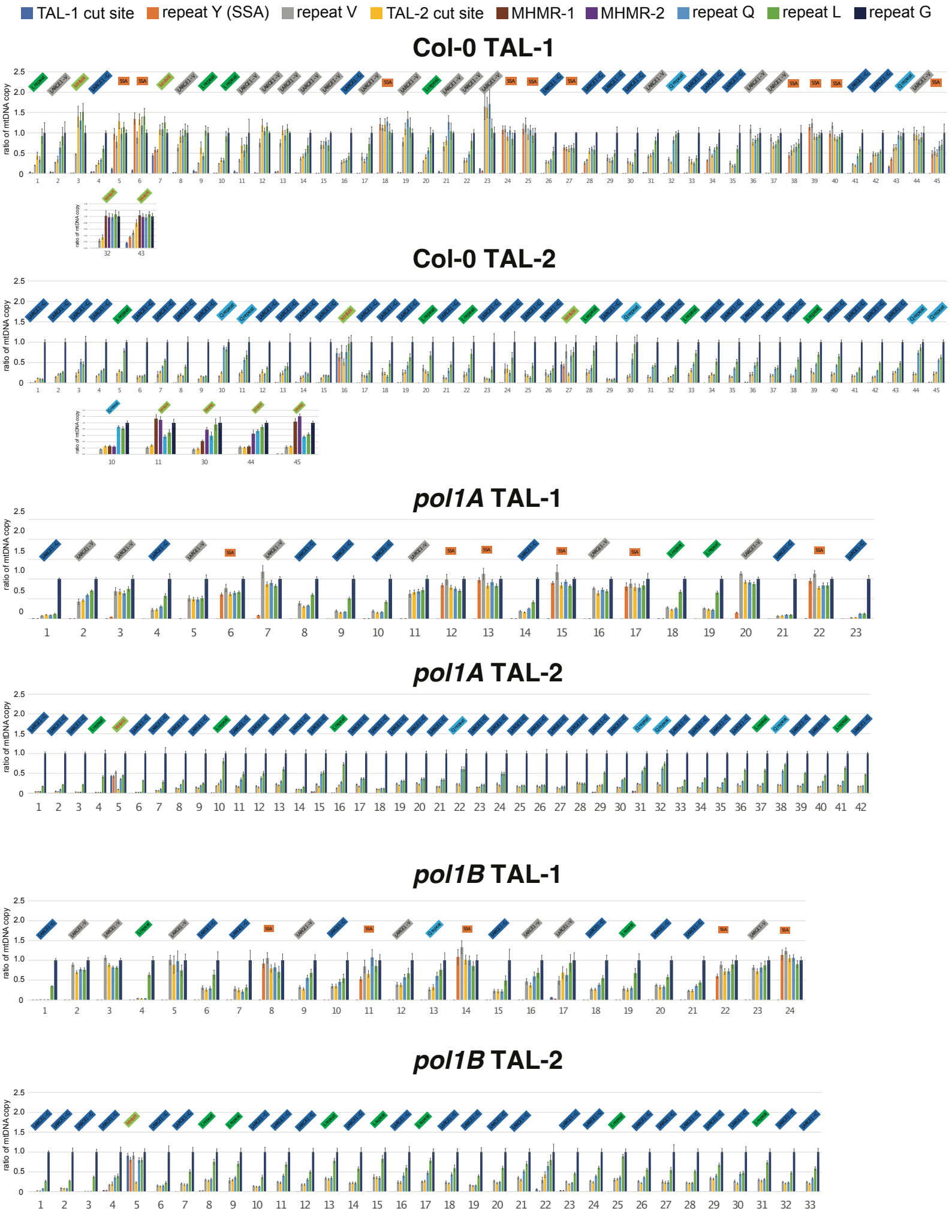
Heteroplasmic *oex1* TAL-2

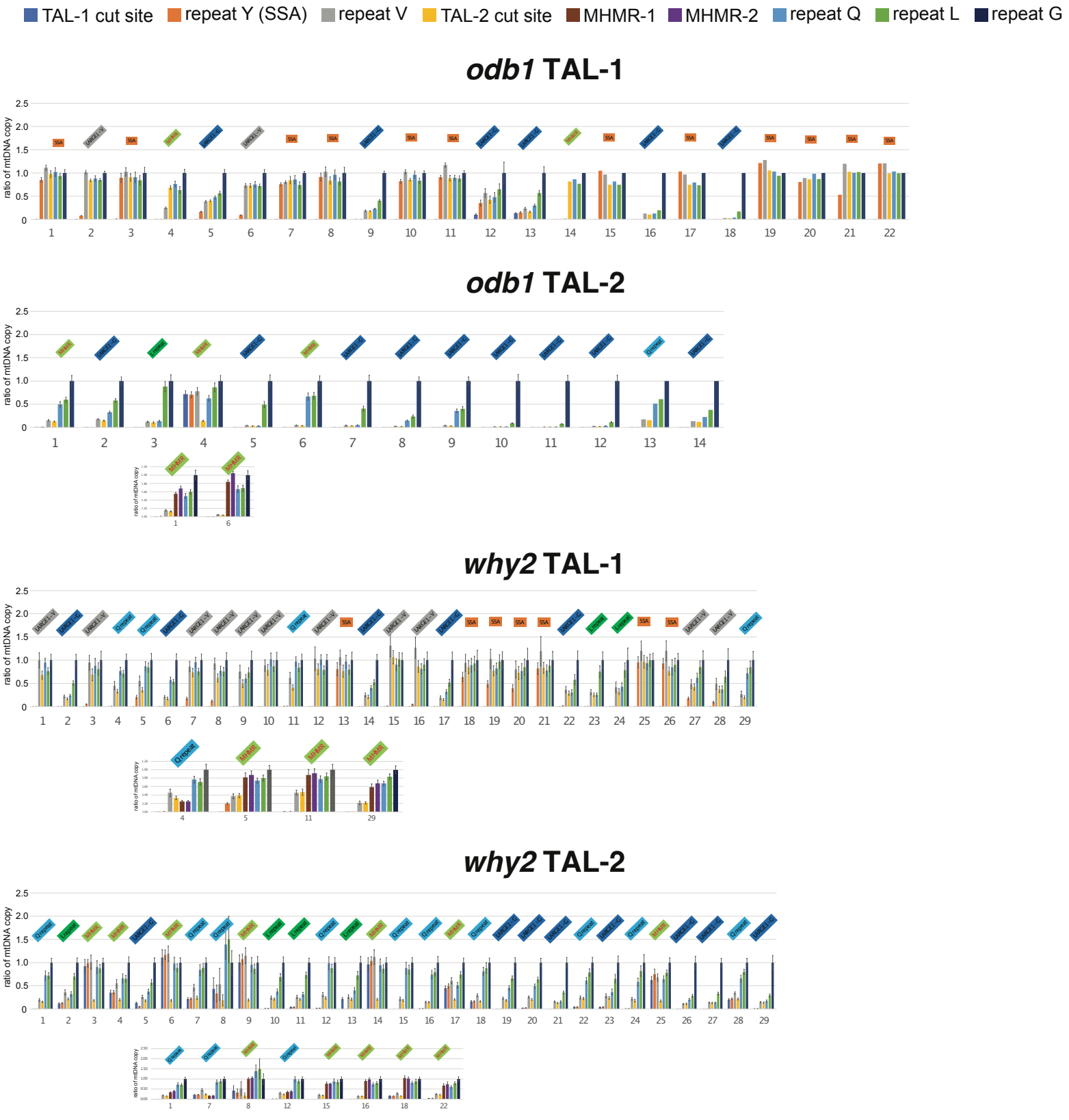






Appendix.6. qPCR Analysis of repair pathway active to repair DSB in mitoTALEN mutants




TAL-1 cut site
repeat Y (SSA)
repeat V
TAL-2 cut site
MHMR-1
MHMR-2
repeat Q
repeat L
repeat G
odb1 TAL-1
odb1 TAL-2
why2 TAL-1
why2 TAL-2
ratio of mtDNA copy

Résumé en Français

Introduction

Les mitochondries, que l'on trouve dans presque toutes les cellules eucaryotes, sont l'un des rares organites à posséder leur propre ADN. On pense que les génomes mitochondriaux dérivent d'un ancêtre symbiotique, les α-protéobactéries (Sagan, 1967). Cependant, les mitochondries des plantes sont très différentes de celles des animaux. En comparaison avec les mitochondries humaines, petites et compactes, les génomes mitochondriaux des plantes sont grands et riches en ADN non codant. Le plus petit génome mitochondrial végétal séquencé, celui de *Brassica napus*, d'une longueur de 220 kb, est plus de 10 fois plus grand que celui des humains (Handa, 2003). Le plus grand génome végétal séquencé à ce jour, celui de *Larix sibirica*, est estimé à 11,7 Mbp de longueur ; cependant, malgré sa taille, il ne contient que 77 gènes (Putintseva et al., 2020). De plus, les mitochondries des plantes ont un taux de mutation très faible tout en présentant des variations structurelles fréquentes (Drouin et al., 2008). Ces réarrangements structurels sont dus à l'abondance de séquences répétées impliquées dans des événements de recombinaison réciproque. Les événements de recombinaison produisent une collection de sous-génomes qui peuvent être rassemblés dans un génome unique et uniforme contenant toutes les séquences codantes de la mitochondrie, souvent appelé cercle maître (Sloan, 2013).

La réparation de l'ADN mitochondrial (ADNmt) suscite un grand intérêt car, chez la plupart des plantes supérieures, les séquences d'ADNmt sont hautement conservées et évoluent très lentement. Plusieurs voies de réparation de l'ADN ont été identifiées dans les mitochondries des plantes, telles que la réparation des mésappariements et la réparation par excision de bases. Cependant, on pense que la recombinaison homologue, en particulier, est responsable de ces faibles taux de mutation et de l'évolution structurelle rapide du ADNmt des plantes (Davila et al., 2011). Ainsi, le principal objectif de mon doctorat était de mieux comprendre les voies de réparation impliquées dans la réparation des cassures double-brin (DSB) et d'acquérir une meilleure compréhension de la fonction de facteurs clés impliquées dans la réparation de l'ADNmt par recombinaison homologue, en utilisant des outils de génie génomique.

Jusqu'à présent, il n'existait pas de méthodes fiables pour transformer les mitochondries des plantes. Auparavant, la seule transformation mitochondriale réussie avait été réalisée par une approche biolistique chez la levure (Fox et al., 1988; Johnston et al., 1988; Zhou et al., 2010) et les algues vertes (Bonnefoy et al., 2007; Randolph-Anderson et al., 1993; Remacle et al., 2006). Par conséquent, les études antérieures sur la réparation du



Figure 1. Analyse par qPCR des sites de coupure mitoTALEN dans les semis T1 mutants TAL-1 et TAL-2 (n=45). Le nombre relatif de copies de l'ADNmt à proximité des sites de coupure TAL-1 et TAL-2 est normalisé par rapport au nombre de copies d'une séquence d'ADNmt distante (gène cox2).



Figure 2. Analyse par qPCR des sites de coupure mitoTALEN dans les graines matures T1 mutantes TAL-1 et TAL-2 (n=150). Le nombre relatif de copies de l'ADNmt à proximité des sites de coupure TAL-1 et TAL-2 est normalisé par rapport au nombre de copies d'une séquence d'ADNmt distante.

ADNmt ont été effectuées en utilisant des agents chimiques comme la ciprofloxacine (CIP) pour introduire des DSB. Cependant, la CIP introduit de multiples DSB à des positions aléatoires et cible à la fois les chloroplastes et les mitochondries. L'étude exclusive des DSB en utilisant la CIP dans les mitochondries n'est pas possible en raison des difficultés à identifier les sites cibles et des effets secondaires dus aux mutations délétères aléatoires dans les gènes.

Les récents progrès réalisés par l'équipe du Prof. Arimura (Université de Tokyo) sur l'outil d'édition génétique TALEN leur ont permis de créer une nouvelle technologie, les mitoTALENs, capable de produire une cassure double brin (DSB) spécifique dans le génome mitochondrial, générant des mutants quasi homochondriomiques ou homoplasmiques (Arimura et al., 2020). Quelques années plus tard, une nouvelle approche de mutagenèse par "Gene drive mutagenesis" (GDM) a été développée en utilisant le mitoTALEN, pour produire des plantes mutantes mitochondriales génétiquement stables et homoplasmiques (Forner et al., 2022). Une telle avancée dans les outils d'édition génétique et le séquençage de nouvelle génération (NGS) m'a fourni les outils et ouvert la possibilité d'analyser la maintenance et la réparation du génome mitochondrial après l'introduction de DSB spécifiques.

L'équipe Arimura a aimablement fourni deux constructions mitoTALEN qui ciblent des régions non-essentielles de l'ADN mitochondrial et introduisent des DSB (cassures double brin). TAL-1 et TAL-2 correspondent aux constructions annotées respectivement comme "ATP6-2 et ATP6-2UP" dans (Arimura et al., 2020). Pour étudier la réparation des DSB, j'ai introduit les constructions TALEN dans des plantes d'*Arabidopsis thaliana* sauvages, accession Col-0, et dans des mutants de la voie de recombinaison homologue mitochondriale (*odb1, why2, osb1-1 osb4-1, radA-3, msh1, pol1A, pol1B, recA3, oex1-1, recG1-2*), afin de mieux comprendre la fonction des protéines respectives.

Résultats du Projet

Tout d'abord, la réparation des cassures double brin (DSB) a été analysée dans le fond génétique sauvage Col-0. Une étude préliminaire des plantules T1 par qPCR a montré une perte presque complète des sites cibles des mitoTALENs (Fig. 1). Pour identifier quand une DSB est réparée, l'ADN des graines T1 a été extrait et utilisé pour la même analyse qPCR (Fig. 2). La séquence d'ADN au site cible était perdue dans les graines mûres comme dans les plantules, ce qui suggère que les mitoTALENs sont déjà très actifs tôt dans le développement des graines. De plus, les sites de coupure de TAL-1 et TAL-2 étaient affectés dans les plantes transformées avec l'une ou l'autre des constructions, ce qui suggère que la réparation des DSB a entraîné la perte de grandes régions de l'ADNmt.

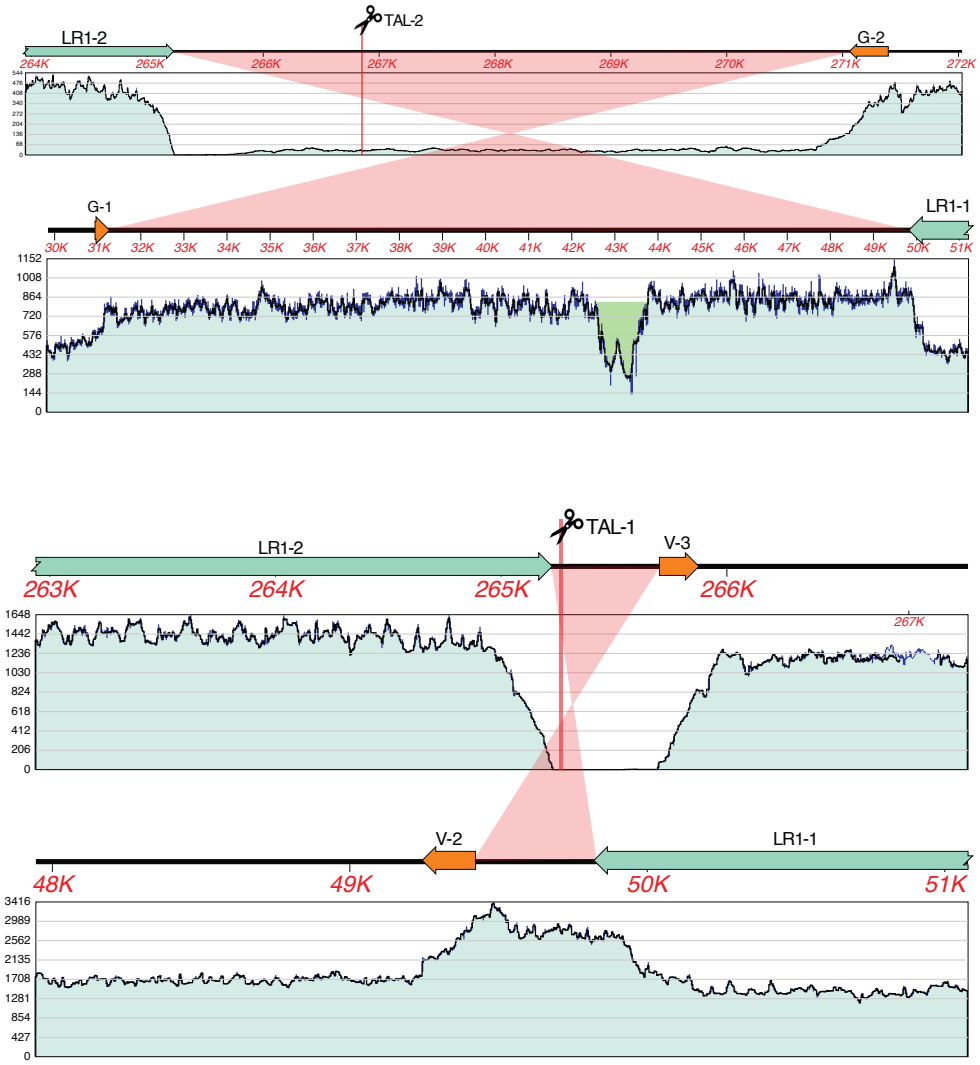


Figure 3. Diagrammes de couverture des du séquençage de Col-0 TAL-2 (panneau supérieur) et du séquençage de Col-0 TAL-1 (panneau inférieur), mises en correspondance avec le génome mitochondrial de Col-0. Produit par l'aligneur Bowtie2 implémenté dans MacVector Assembler. Les encadrés rouges représentent la région comprenant le site de coupure TALEN remplacée par la région en dessous, situées loin l'une de l'autre dans l'ADNmt (voir les coordonnées du génome). Il s'agit des régions situées entre la grande répétition LR1 et la petite répétition G (panneau supérieur, pour TAL-2) et entre LR1 et la petite répétition V (panneau inférieur, pour TAL-1). La zone en vert représente des séquence supprimées, correspondant à des insertions de séquences chloroplastiques dans l'ADNmt.

L'approche principale de l'étude consistait à séquencer l'ADNmt des lignées mitoTALEN et, selon les empreintes génétiques résultantes à côté du site de coupure, déduire comment l'ADNmt avait été réparé. Pour ce faire, j'ai réalisé un séquençage d'ADN Illumina sur des échantillons mitoTALEN partiellement enrichis en mitochondries. Plusieurs raisons m'ont conduit à choisir de séquencer des plantes individuelles T1. Tout d'abord, les plantes T1 sont la première génération après l'activité des mitoTALENs, et différentes plantes T1 pourraient avoir utilisé différentes voies de réparation des DSB. Ensuite, il était nécessaire d'étudier des plantes individuelles, car les mitotypes mineurs auraient été difficiles à détecter dans un pool de plantes. Le regroupement de plantes aurait dilué les résultats alternatifs mineurs de la réparation des DSB. Enfin, le séquençage de la progéniture T2 ségrégant aurait introduit un biais, en raison de l'héritage maternel de l'ADNmt.

Cartographie des séquences Illumina des lignées TALEN dans le fond WT du génome mitochondrial a confirmé que la recombinaison homologue (HR) est la principale voie de réparation de l'ADN dans les mitochondries d'Arabidopsis. Dans la plupart des plantes séquencées, la perte du site de coupure était accompagnée d'une augmentation respective de la copie d'ADNmt de régions éloignées encadrées par les mêmes ensembles de séquences répétées (Fig. 3). Ces régions étaient également encadrées par les mêmes répétitions (séquences répétées G ou V à la bordure gauche et LR1-1 à la bordure droite). Étant donné l'amplitude équivalente des changements de couverture ainsi que l'implication des mêmes paires de répétitions, cela suggérait que la région avec une couverture accrue était dupliquée, remplaçant la région contenant le site de coupure qui avait été perdue. Pour prouver cette hypothèse, la région encadrée par les répétitions a été amplifiée par PCR longue distance, et le fragment résultant a été purifié sur gel et séquencé (Fig. 4). Le séquençage Sanger a confirmé que le fragment amplifié dans TAL-1 avait la région dupliquée entre la répétition LR1 et la séquence répétée V. Nous avons proposé le mécanisme de recombinaison homologue impliquant deux ensembles de séquences répétées dans la réplication induite par rupture (BIR) (Fig. 5). Le modèle suggère qu'un des sites d'homologie est utilisé lors de la capture de brin, tandis que la seconde région d'homologie verrouille et stabilise l'extension par capture de la deuxième extrémité, entraînant la duplication de l'ADN matrice au site de coupure. En conséquence, ce modèle de réparation des DSB explique comment de nouvelles grandes répétitions peuvent être créées, devenant prédominantes dans les génomes mitochondriaux des plantes, et contribuant grandement au contenu élevé en ADN non codant et à la grande taille des génomes d'ADNmt des plantes.

Outre la recombinaison homologue, plusieurs voies de réparation alternatives ont été identifiées dans les plantes WT : l'appariement de simple-brins ("Single-Strand Annealing, SSA) et la réparation médiée par microhomologie (MHMR). Dans certaines plantes WT

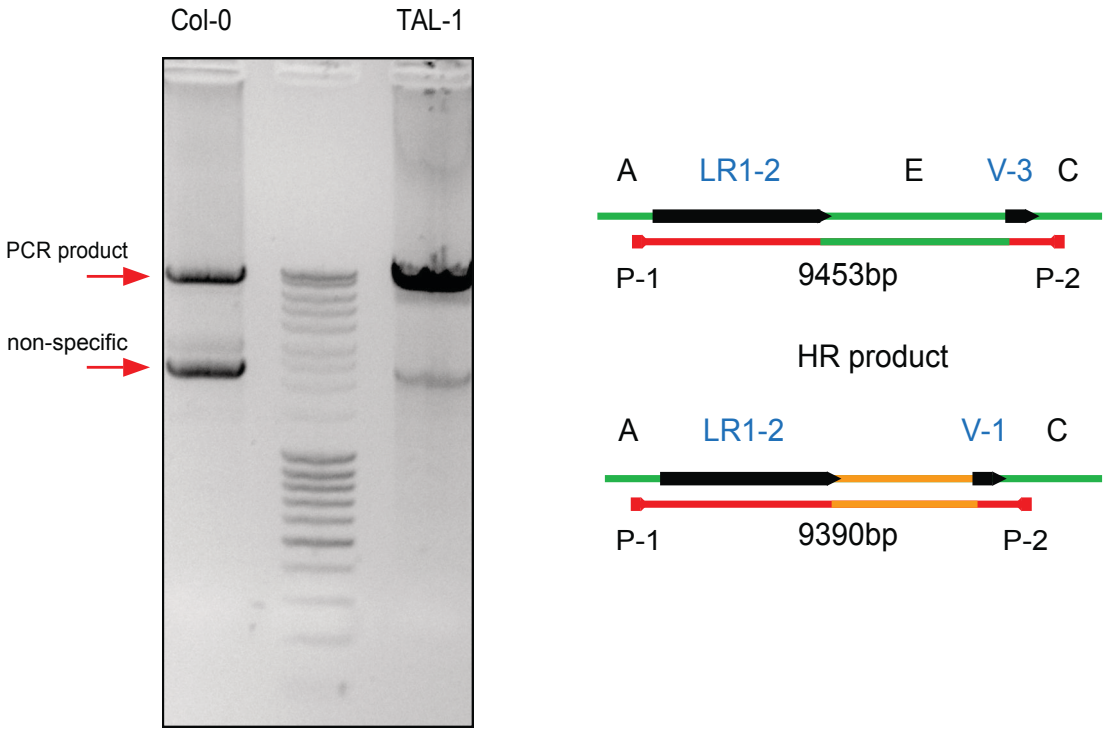


Figure 4. Deux événements de recombinaison homologue impliquant respectivement la grande séquence répétée 1 (LR1) et la séquence répétée V entraînent le remplacement de la région E par une séquence localisée dans un locus éloigné du génome, qui devient dupliquée. Analyse sur gel des produits PCR et schémas (à droite) de l'amplification PCR des régions en jaune et en rouge, utilisant les mêmes amorces P-1 et P-2, mais produisant des produits différents dans l'échantillon TAL-1 par rapport au WT (Col-0). Le modèle a été confirmé par la séquence Sanger des fragments de PCR.

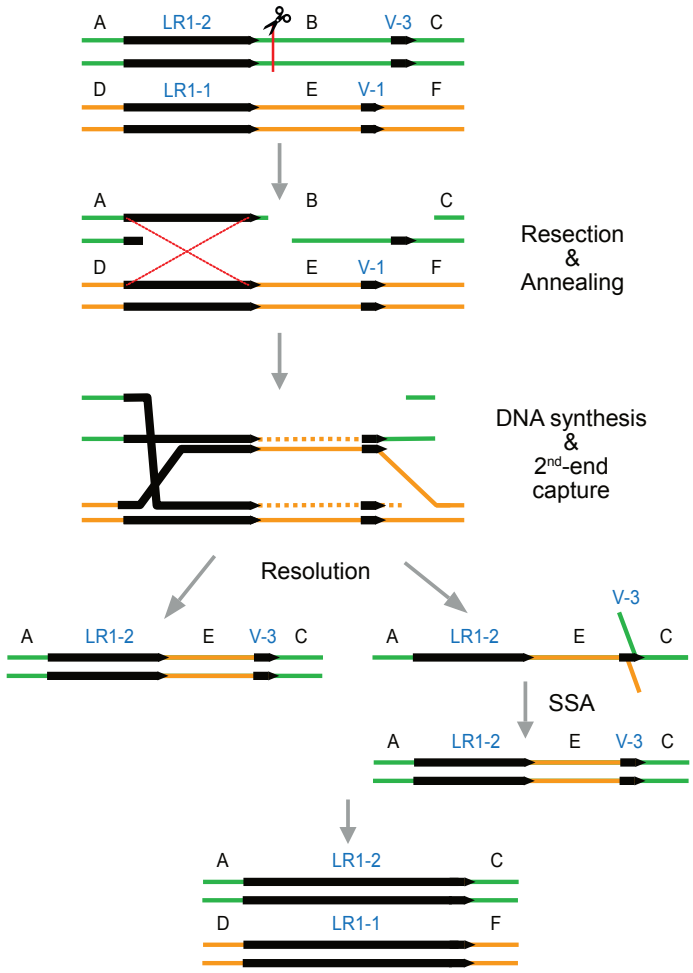


Figure 5. Mécanisme proposé de réparation de la DSB induite par la mitoTALEN par HR, impliquant deux paires de séquences répétées, dans ce cas, la grande répétition LR1 et la petite séquence répétée V. Les régions de l'ADNmt représentées en vert et en orange sont très éloignées l'une de l'autre dans le génome, mais elles partagent les mêmes répétitions, en noir.

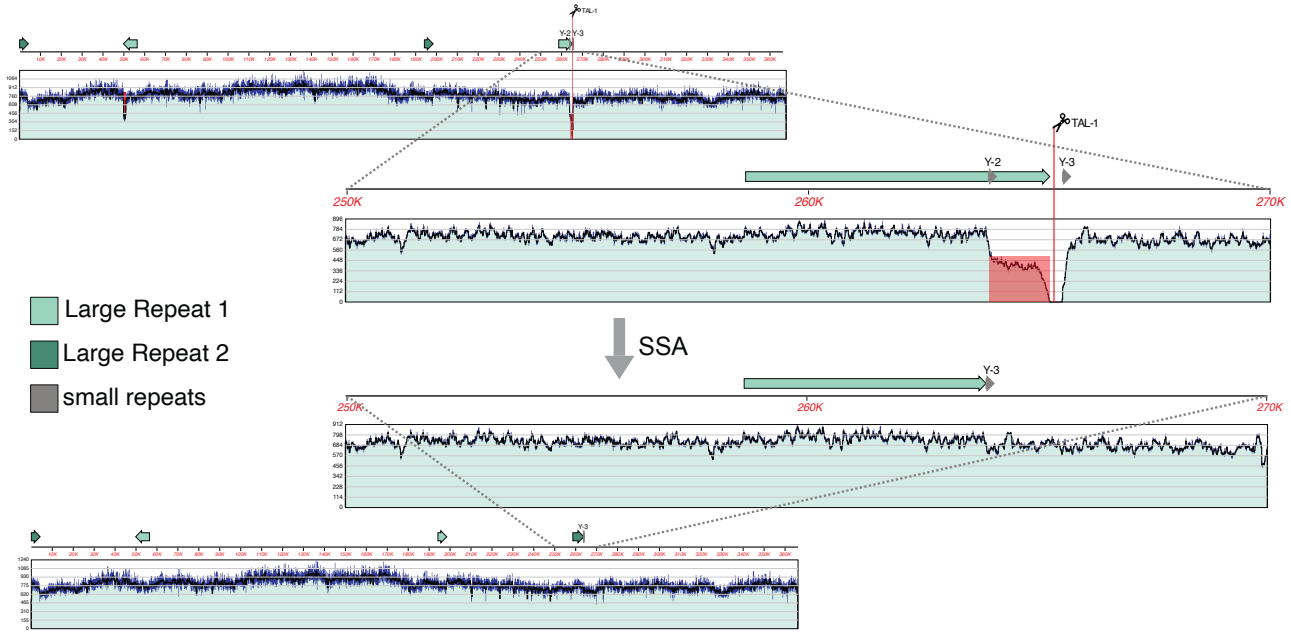


Figure 6. Cartographie des séquences des lignées mitoTALEN au nouveau mitotype théorique, qui résulte de la réparation de DSB par "Single Strand Annealing" (SSA). Représentation graphique de la couverture du séquençage de Col-0 TAL-1, mises en correspondance avec l'ADNmt de Col-0 WT (panneau supérieur) et avec le nouveau mitotype proposé, respectivement (résultant de la suppression de la région entre les répétitions Y). La région en rouge entre les répétitions Y est perdue par réparation SSA. La réduction de la couverture (en rouge) est due à la multi-cartographie des séquences, résultant d'une distribution égale entre les sites de cartographie dupliqués, bien que l'un des sites soit perdu par SSA. Produit par l'aligneur Bowtie2 implémenté dans le package MacVector Assembler.

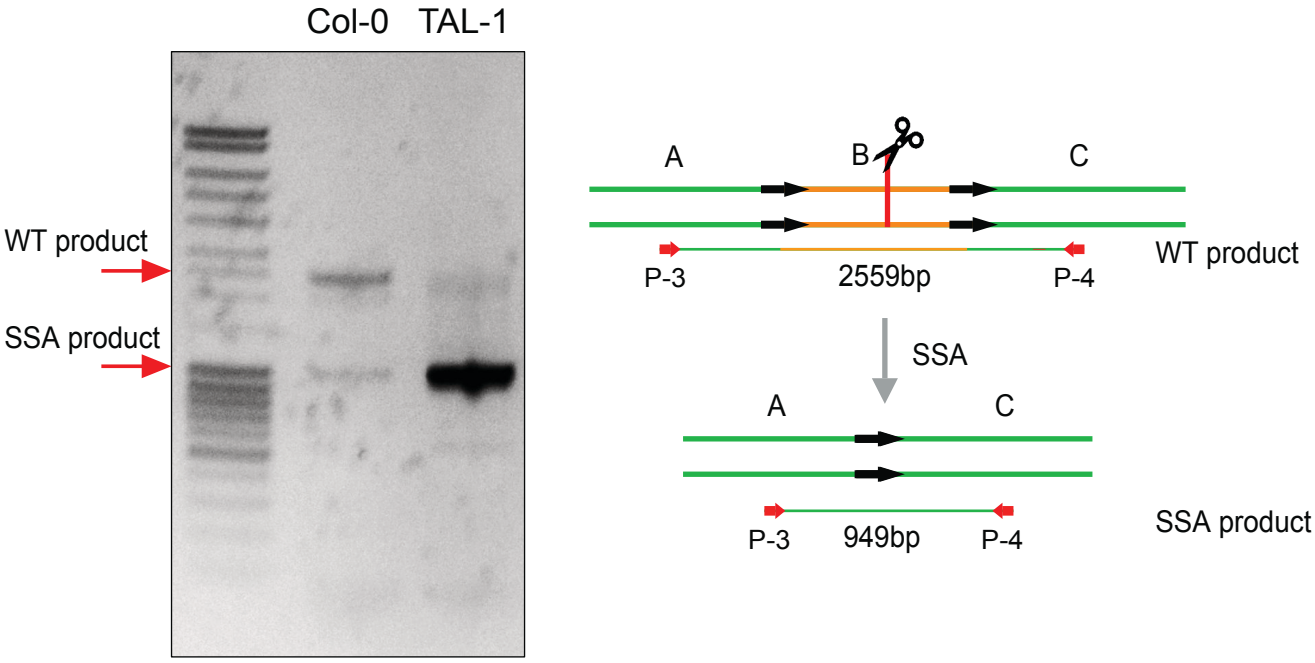


Figure 7. La SSA impliquant la séquence répétée Y entraîne la suppression de la région B, contenant la cassure double brin induite par TAL-1. Analyse par électrophorèse sur gel d'agarose et schémas de l'amplification par PCR des régions correspondantes, en utilisant les amorces P-3 et P-4, provenant de la lignée TAL-1 et de la lignée sauvage Col-0. Confirmé par la séquence Sanger.

transformées avec la construction TAL-1, les cassures double brin (DSB) ont été réparées par une paire de répétitions directes, Y-2 et Y-3, situées à seulement 1,4 kb l'une de l'autre dans l'ADNmt. Il y avait une diminution de la couverture flanquée par les répétitions Y, tandis que le reste de l'ADNmt était en équilibre stœchiométrique (Fig. 6). J'ai émis l'hypothèse que la région entre les répétitions avait été perdue via SSA. Pour confirmer cette hypothèse, j'ai amplifié la région du site de coupure par PCR et, comme prévu, le fragment PCR résultant était réduit en taille chez le mutant mitoTALEN comparé à la plante WT (Fig. 7). Le séquençage des fragments PCR a confirmé que chez le mutant mitoTALEN, la région comprise entre les deux copies de la répétition Y est perdue.

Dans certains transformants mitoTALEN séquencés, il y avait une diminution de la couverture de séquence autour du DSB, sans lien avec aucune séquence répétée annotée (Fig. 8). Une analyse plus approfondie des séquences a révélé des micro-homologies qui étaient utilisées, dans la réparation médiée par microhomologie (Fig. 9). Pour une analyse complète des événements MHMR, une chaîne de traitement pour l'analyse des "soft-clips" dans les données NGS a été développée. Certaines lectures s'alignent incomplètement au génome de référence, et une extrémité non-alignée d'une lecture est appelée "soft-clip". Une des raisons pourrait être que les séquences peuvent se mapper de manière discontinue à deux régions différentes du génome matrice, en raison de la réorganisation du génome à ces loci. La cartographie de ces segments soft-clip à l'ADNmt résulte en leur regroupement à travers le génome. Les pics résultants indiquent des sites de recombinaison ectopique impliquant soit des micro-homologies, soit de très petites répétitions (plus petites que la taille de la séquence, qui dans notre cas était de 150 pb) (Fig. 10).

L'activité des mitoTALENs dans certaines plantes WT conduit également à un développement anormal : les feuilles sont frisées et panachées, parfois réduites en taille, et, selon la gravité du phénotype, les plantes affectées ne parviennent pas à développer des tiges florales et des organes floraux. Les phénotypes pléiotropes des plantes T1 transformées avec les constructions mitoTALEN confirment que différents mitotypes majeurs sont créés dans des plantes individuelles. On a proposé que le phénotype soit dû à l'un des mitotypes produits, dans lequel la DSB est réparée via la recombinaison impliquant la séquence répétée G, ce qui conduit à la perte de la partie C-terminale du gène *atp9* et crée potentiellement une fusion entre le 5' de *atp9* et l'orf315 (Fig. 11). Des résultats de qPCR ont confirmé une diminution de l'extrémité 3’ de atp9 chez les plantes TAL-2 phénotypiques (Fig. 12). Pour analyser plus en détail si une diminution de l'expression du gène *atp9* est corrélée avec les phénotypes de développement défectueux, nous avons effectué des analyses par RT-qPCR et Northern blot (Fig. 13 et 14). Ces analyses ont révélé une diminution des niveaux du transcript d'*atp9* chez les plantes TAL-1 et TAL-2 avec des développements défectueux,

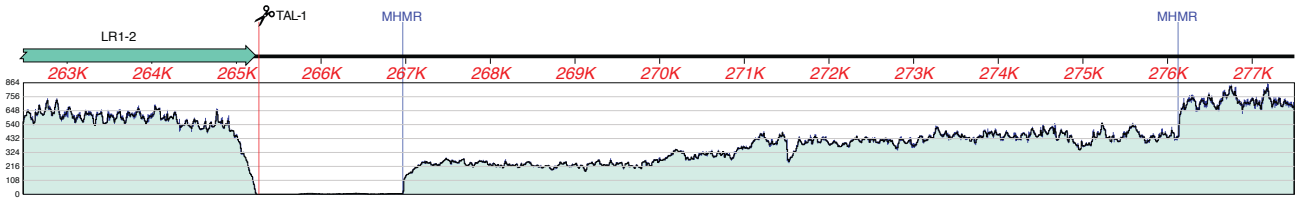


Figure 8. Représentation graphique de la couverture du séquençage de TAL-1, à partir d'un transformant T1 présentant des preuves de réparation par recombinaison impliquant des microhomologique (MHMR). La ligne rouge et les ciseaux indiquent le site cible de TALI-1. Les lignes bleues indiquent les régions de microhomologie qui subissent la MHMR. Produit par l'aligneur Bowtie2 de MacVector.

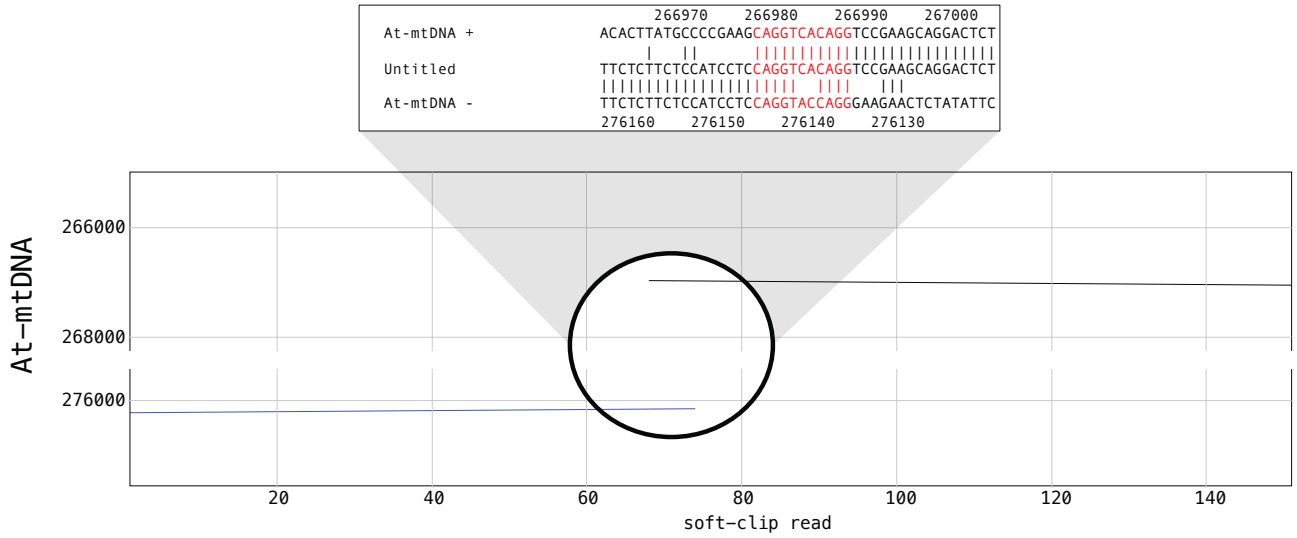


Figure 9. Alignement d'une séquence Soft-clip sur deux loci de l'ADNmt différents, qui partagent une microhomologie (11 pb avec deux mésappariements). L'appariement des bases de la microhomologie reliant les deux sites de l'ADNmt est indiqué en rouge. L'alignement du dot-plot a été généré avec la fonction MacVector Pustel DNA Matrix.

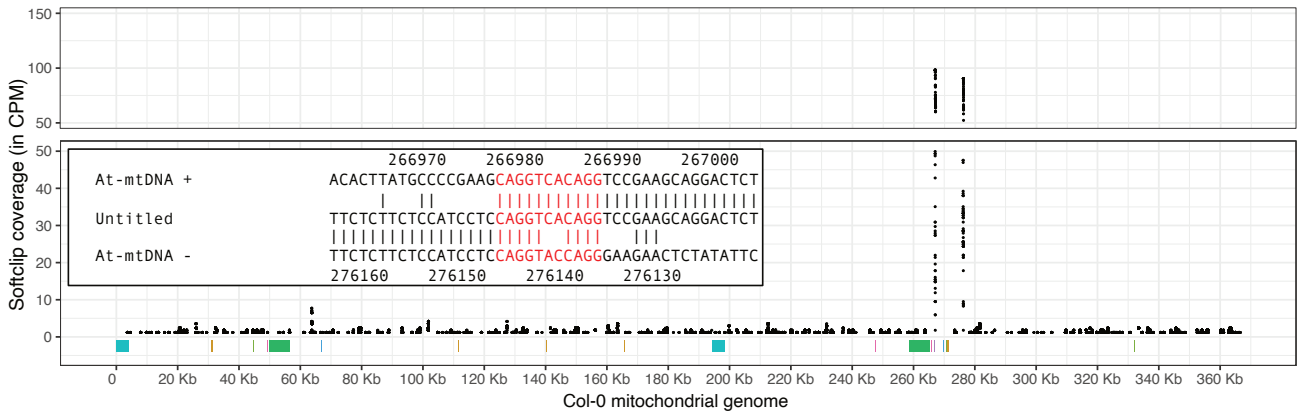


Figure 10. Représentation graphique de la couverture des séquences illumina de Col-0 TAL-1 avec des soft-clips > 20 pb, cartographiées sur l'ADNmt de Col-0 WT. Les deux pics de couverture élevée correspondent aux sites de MHMR. En rouge, l'appariement des bases reliant les deux sites de l'ADNmt.

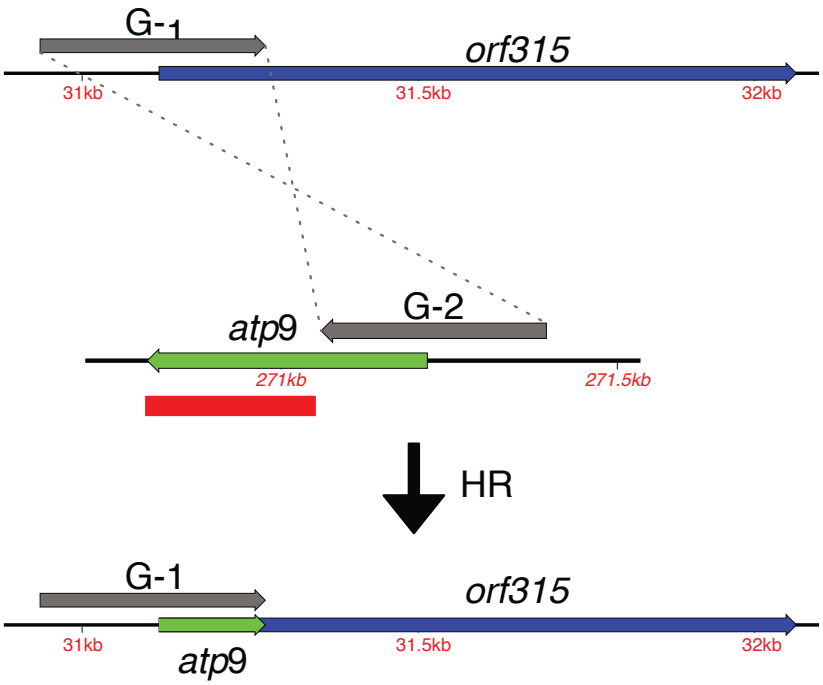


Figure 11. Représentation schématique de la recombinaison impliquant la répétition G. L'extrémité 3' d'atp9 est perdue par recombinaison ; cependant, comme elle fait partie de la répétition G, une version tronquée est fusionnée en phase à l'orf315, ce qui donne un gène chimérique qui peut être transcrit par le promoteur situé en amont de l'atp9. La région en rouge de l'atp9 est perdue.

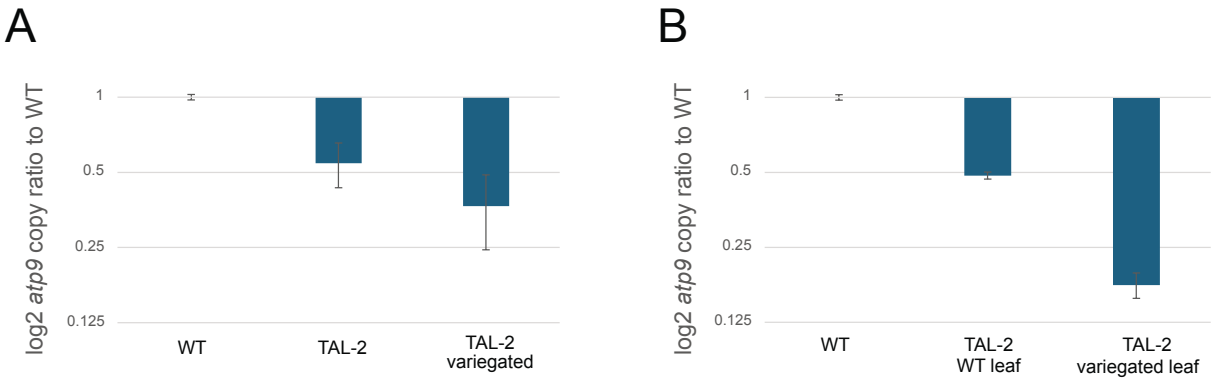


Figure 12. Quantification par qPCR de l'extrémité 3' d'atp9. (A) Comparaison dans les transformants T1 TAL-2 avec (feuilles déformées et panachées) et sans phénotype défectueux. (B) Comparaison entre des feuilles présentant le phénotype et des feuilles de type sauvage, provenant de la même plante chimérique TAL-2.

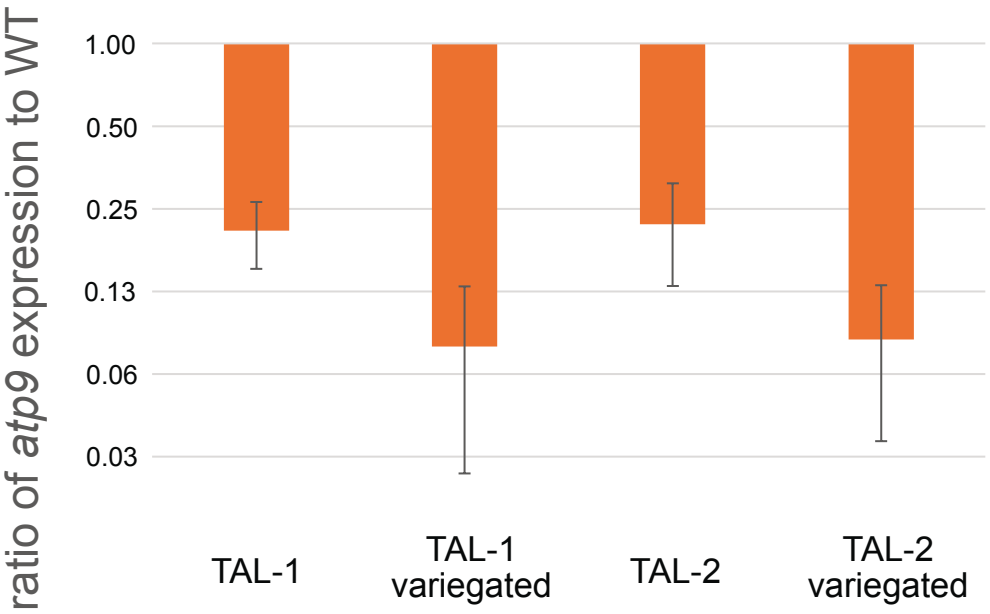


Figure 13. Analyse RT-qPCR de l'expression d'atp9 dans les transformants TAL-1 et TAL-2 présentant ou pas le phénotype délétère.

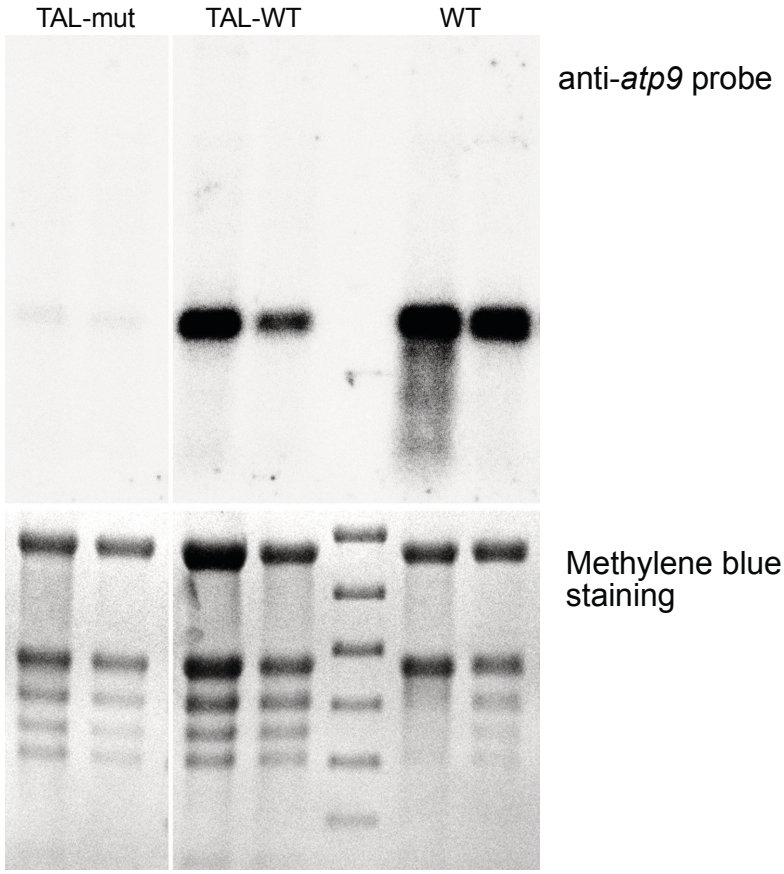


Figure 14. Analyse par Northern blot du transcrit d'atp9 dans transformants TAL-2 avec (TAL-mut) et sans (TAL-WT) défaut de développement. La membrane a été colorée au bleu de méthylène après transfert, comme contrôle de charge.

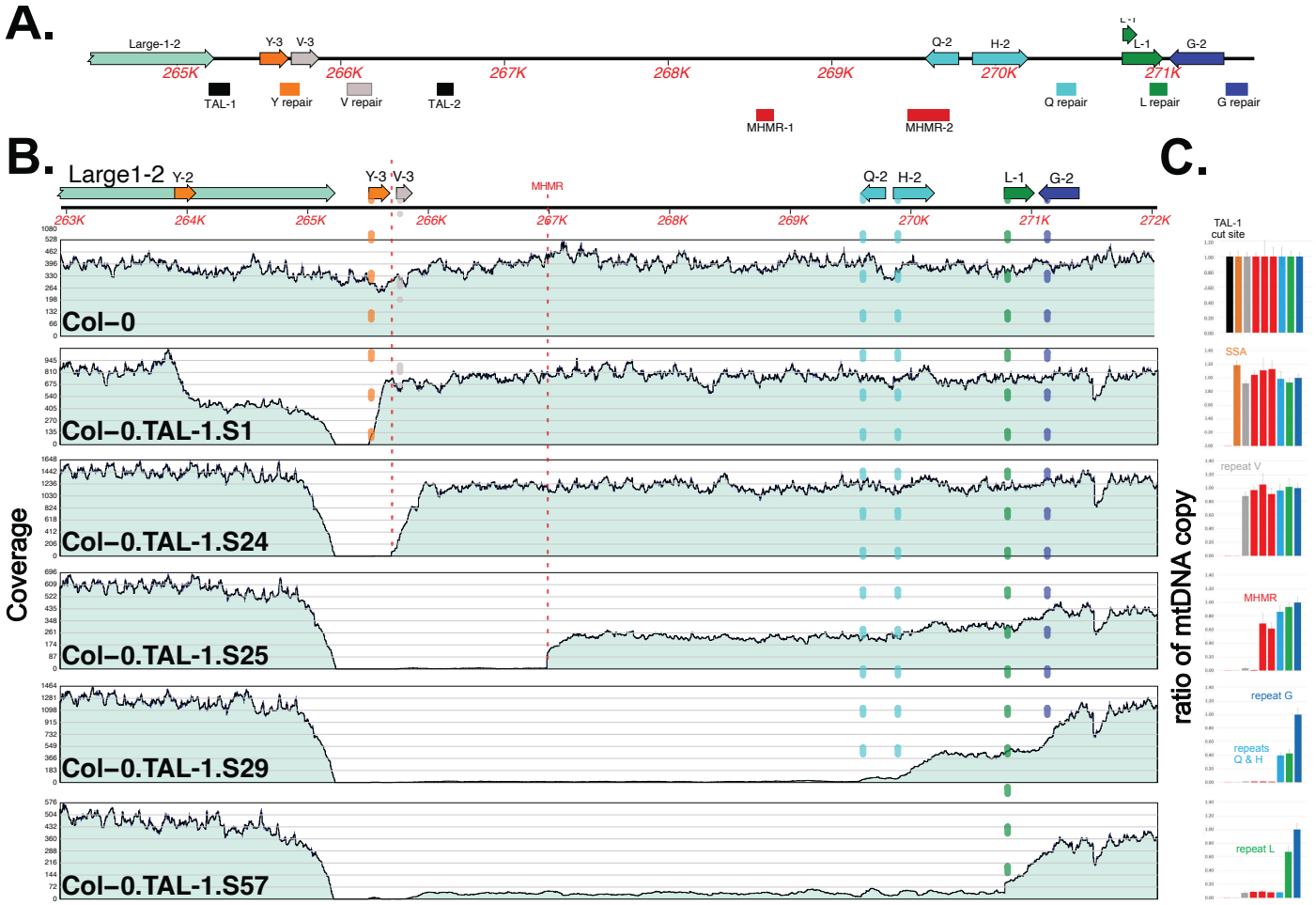


Figure 15. Validation de l'analyse qPCR sur des échantillons préalablement analysés par séquence Illumina. (A) Sondes sélectionnées pour l'analyse qPCR des voies de réparation (SSA et MHMR) et des répétitions impliquées dans la réparation par HR de la DSB. (B) Diagrammes de couverture, avec des lignes pointillées codées en couleur indiquant les répétitions ou les microhomologies liées à une baisse correspondante de la couverture. (C) Analyse par qPCR du nombre relatif de copies des séquences de l'ADNmt indiquées en (A).

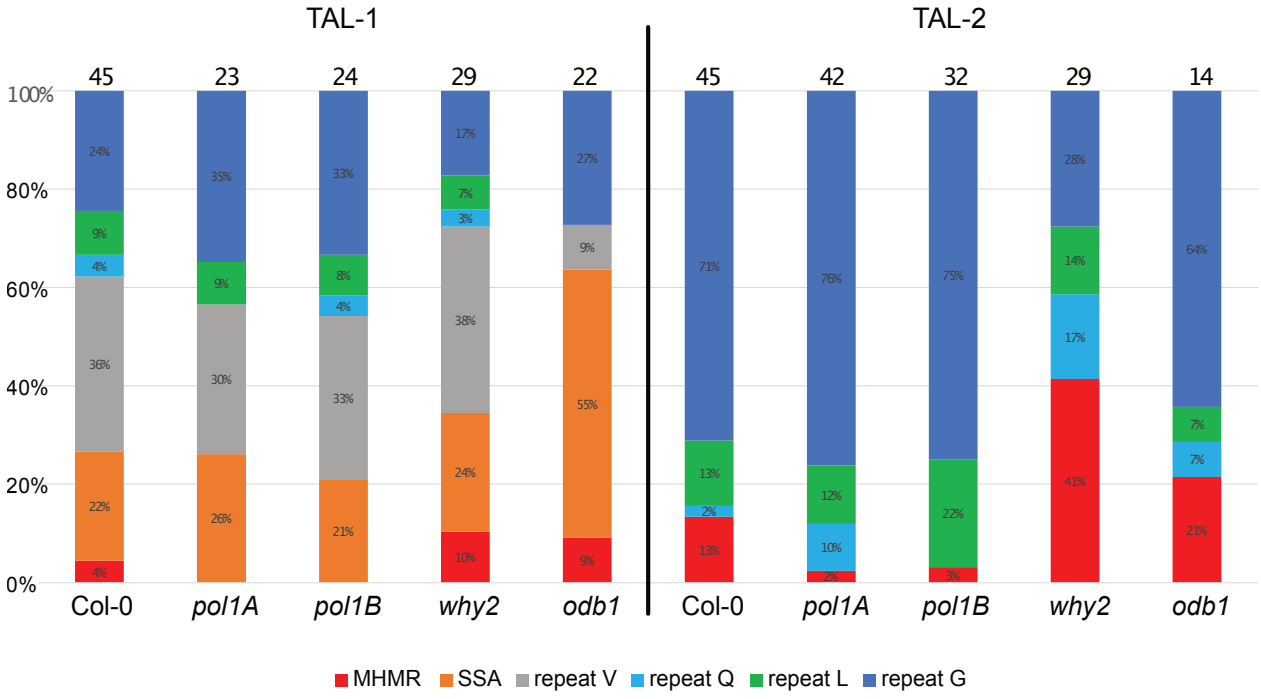


Figure 16. Fréquences des différents mécanismes de réparation impliqués dans la réparation des DBS introduits par les MitoTALEN, dans les lignées mutantes transformées avec les constructions mitoTALEN (MHMR, SSA, et HR impliquant les répétitions V, Q, L, ou G). Évalué par analyse qPCR, Le nombre de plantes T1 indépendantes analysées est indiqué.

comparé aux transformants sans phénotype et aux plantes WT Col-0. Cependant, aucune transcription de fusion entre *atp9* et orf315 n'a été détectée par Northern blot. Ces résultats confirment collectivement que les défauts de développement dans les plantes mitoTALEN sont dus à la perte du gène *atp9* et à une diminution de son expression.

De plus, l'observation de plantes avec des phénotypes chimériques a suggéré l'hétéroplasmie de l'ADNmt, avec la possibilité d'avoir deux mitotypes majeurs différents dans des tissus différents de la même plante. Les feuilles panachées ont montré une diminution plus importante du nombre de copies du 3' de *atp9* par rapport aux feuilles de type sauvage de la même plante (Fig. 12). En outre, l'analyse de la progéniture des plantes T1 hétéroplasmiques a révélé que les révertants vers l'ADNmt de type sauvage peuvent se ségréger si le mitotype WT a été retenu dans la plante maternelle. Cela suggère qu'il pourrait y avoir une préférence et un contrôle inconnu de l'équilibre entre les mitotypes coexistant.

Pour étudier comment une DSB est réparée chez les mutants dépourvus de facteurs spécifiques impliqués dans la voie de recombinaison homologue (HR) mitochondriale, les lignées mutantes correspondantes ont été transformées avec les constructions TAL1 et TAL2, et l'ADNmt des plantes T1 a été séquencé, comme cela a été fait pour les plantes WT. L'analyse NGS a révélé que lorsque la réparation par HR est limitée chez les plantes mutantes, différentes voies de réparation sont capables de reconstituer un génome stable via SSA (observé chez le mutant *odb1*) ou par la recombinaison médiée par microhomologie (observée chez le mutant *why2*). Pour confirmer les résultats du séquençage NGS, une approche par qPCR a été développée pour analyser de manière économique comment les DSB sont réparée dans un nombre important de réplicats de plantes mutantes (Fig. 15). L'analyse a révélé que la perte des ADN polymérases POL1A ou POL1B ne conduit à aucun changement significatif dans la façon dont les DSB sont réparées (Fig. 16). En revanche, la perte de la protéine RAD52-like ODB1, a entraîné une augmentation de plus de 2 fois la réparation par SSA chez les mutants TAL-1. Nous proposons un modèle où, en présence d'ODB1, les régions ssDNA homologues sont masquées par la liaison de protéines de liaison à l'ADN simple brin (ssDNA) à haute affinité et abondantes, supprimant ainsi la SSA. Cependant, en l'absence d'ODB1, l'appariement des bases d'homologie devient accessible (Fig. 17). De manière surprenante, cela diffère de ce qui a été trouvé chez la levure et les animaux pour RAD52/BRCA2, où les mutations *rad52* conduisent à une SSA défectueuse (Stark et al., 2004 ; Symington, 2002). La perte de WHY2 a entraîné une augmentation de plus de 3 fois de la réparation par MHMR. Nous avons proposé un mécanisme de réparation semblable à celui d'ODB1, selon lequel les sites d'homologie sont masqués par la protéine WHY2 (Fig. 18). Cependant, les différences fonctionnelles entre ces deux protéines de liaison à l'ADN simple brin des mitochondries des plantes restent inconnues.

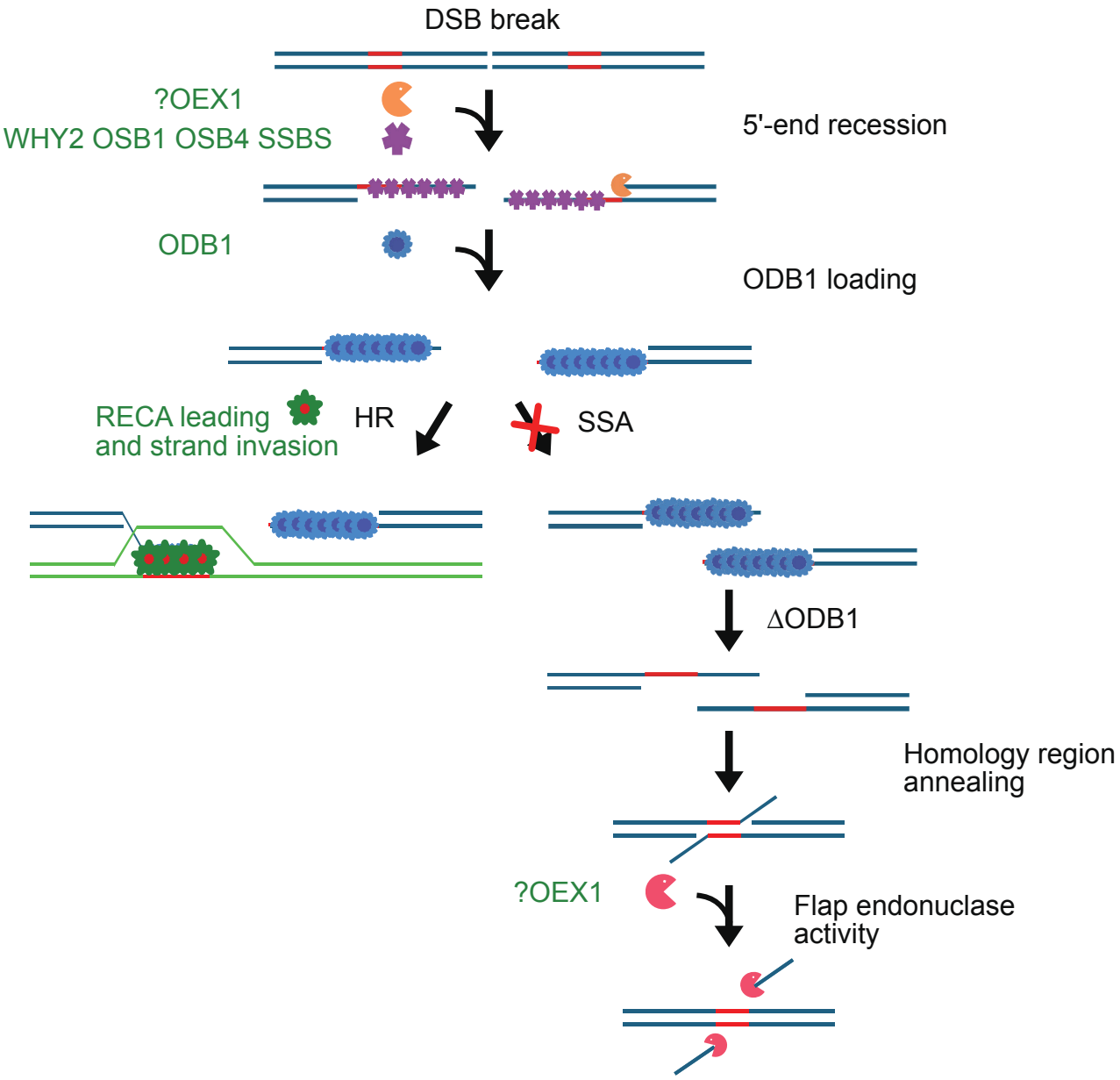


Figure 17. Rôle proposé pour ODB1 dans le mécanisme de réparation par SSA, dans les mitochondries d'Arabidopsis.

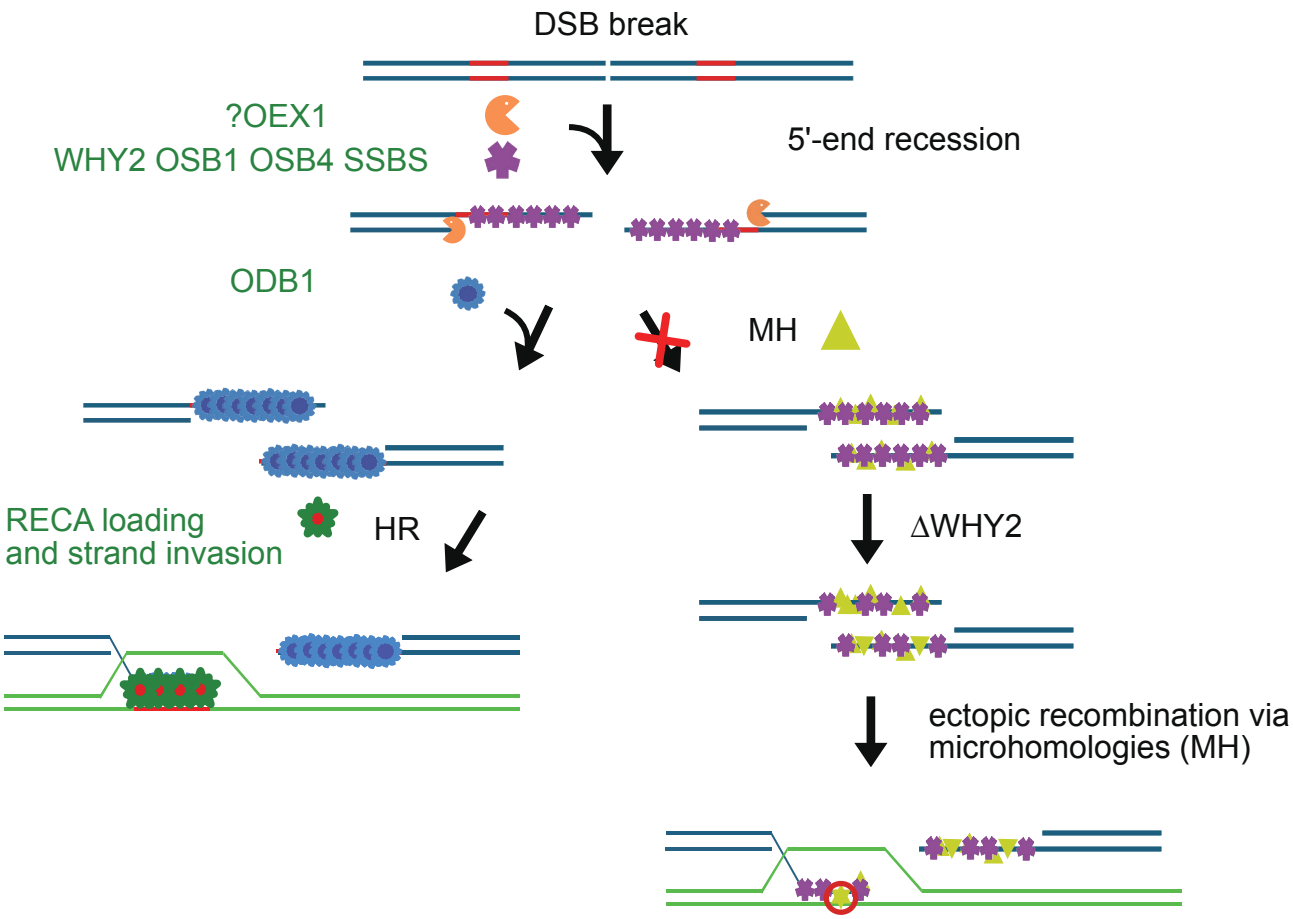


Figure 18. Rôle proposé pour WHY2 dans le mécanisme de réparation par MHMR, dans les mitochondries d'Arabidopsis.

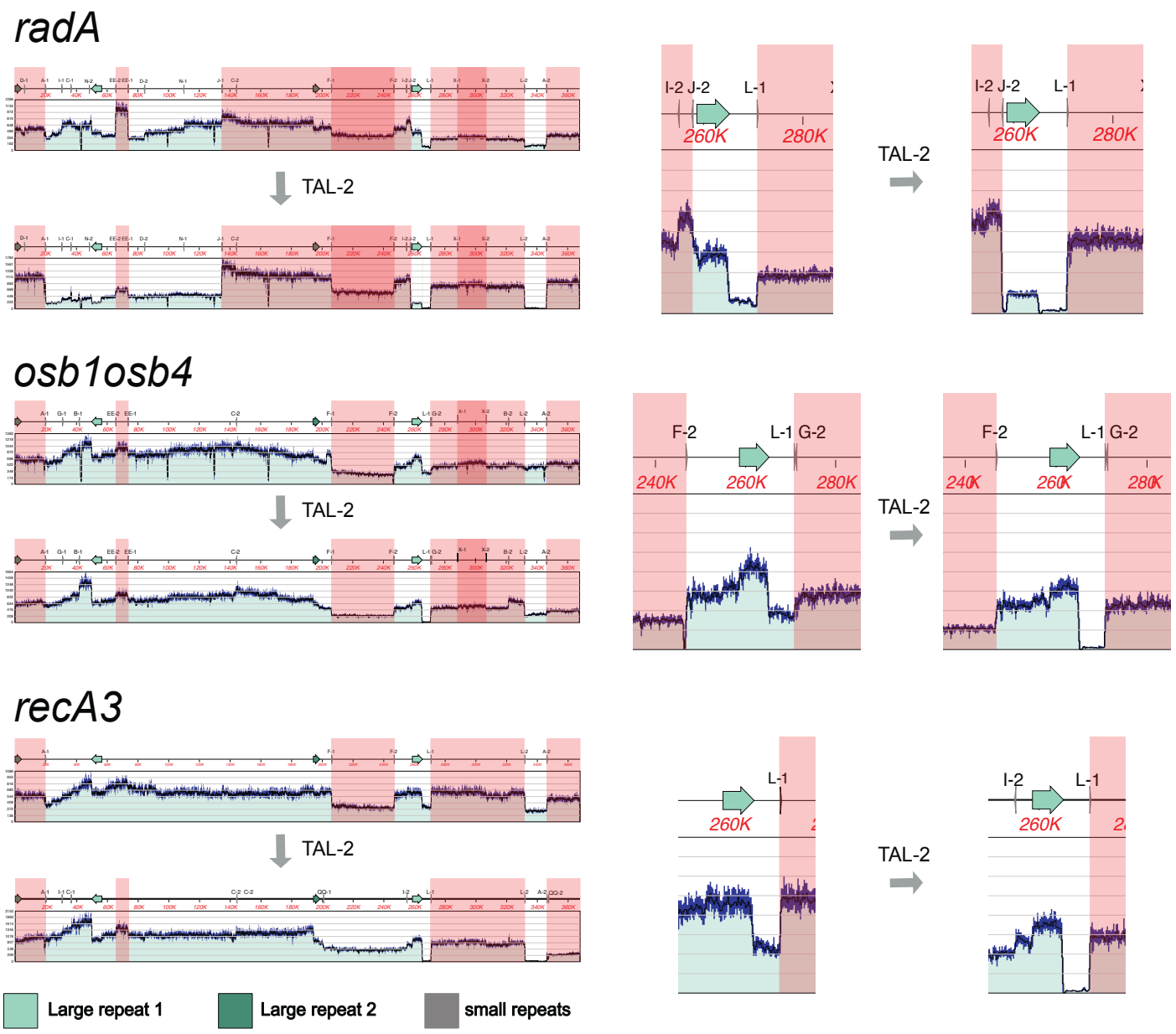


Figure 19. Couverture du séquençage Illumina des mutants d'Arabidopsis radA, osb1osb4 et recA3, avant et après l'introduction de la construction TAL-2. La cartographie des séquences sur l'ADNmt a été faite à l'aide de l'aligneur Bowtie2, implémenté dans l'application MacVector Assembler. Les boîtes rouges indiquent les régions flanquées par séquence répétées en orientation directe, supposées former des sous-génomes se répliquant indépendamment.

De plus, chez les mutants *radA*, *recA* et *osb1osb4*, aucune preuve de réparation dépendante de la recombinaison homologue (HR) n'a été observée (Fig. 19). Cela suggère que ces lignées mutantes pourraient avoir perdu la capacité de réparer les DSB par HR. Cependant, ces mutants sont viables, bien qu'ils présentent des phénotypes de développement prononcés et une réorganisation globale de l'ADNmt. Nous proposons que dans ces plantes, un processus alternatif permet la survie lorsque l'intégrité de l'ADNmt est compromise. Ce processus impliquerait la ségrégation d'un génome mitochondrial viable, reconstruit à partir de sous-génomes mineurs préexistants (Fig. 20). Ces sous-génomes semblent être créés via la HR impliquant des petites séquences répétées en orientation directe, conduisant à la formation d'épisomes qui peuvent répliquer et ségréger indépendamment.

Enfin, chez les mutants *recG1* et *msh1*, des preuves de réparation ont été trouvées ; cependant, les profils de couverture résultants indiquaient une perte de contrôle soit dans la copie des mitotypes ségrégant de manière autonome, soit dans la surveillance de la HR à travers l'ensemble du génome mitochondrial (Fig. 21).

Conclusion

Au cours de mes études doctorales, j'ai présenté la manière dont des mitoTALENs pouvaient être utilisées pour étudier la réparation de l'ADNmt. Je confirme que la HR est la voie de réparation majeure pour la réparation des DSB dans les mitochondries d'Arabidopsis. Je présente également un modèle de comment la HR peut produire de nouvelles séquences répétées plus grandes et augmenter la taille totale du génome mitochondrial. En outre, je montre à quelle vitesse l'ADNmt peut évoluer en une seule génération à partir d'un parent hétéroplasmique, ce qui indique également qu'il existe un contrôle de la sélection du mitotype majeur de l'ADNmt, qui maintient l'équilibre entre les différents mitotypes qui existent dans la cellule. La perte de la voie HR chez les mutants *radA* et *osb1osb4* a conduit à la reconstitution du génome à partir de sous-génome mineurs préexistantes, ce qui a également conduit à la déstabilisation de l'équilibre génomique. Enfin, la perte des protéines de liaison à l'ADN simple brin WHY2 et ODB1 entraîne une augmentation de la MHMR et de la SSA dans les mutants respectifs. Cela suggère que l'abondance de ces protéines stabilise l'ADN simple brin et cache les régions d'homologie. Enfin, des lignées mitoTALEN développées ont été montrées pouvoir conduire à des mutants partiels d'*atp9*, qui pourraient être utilisés dans l'étude de la perte de fonction du complexe V de la chaine respiratoire et des voies de signalisation associées.

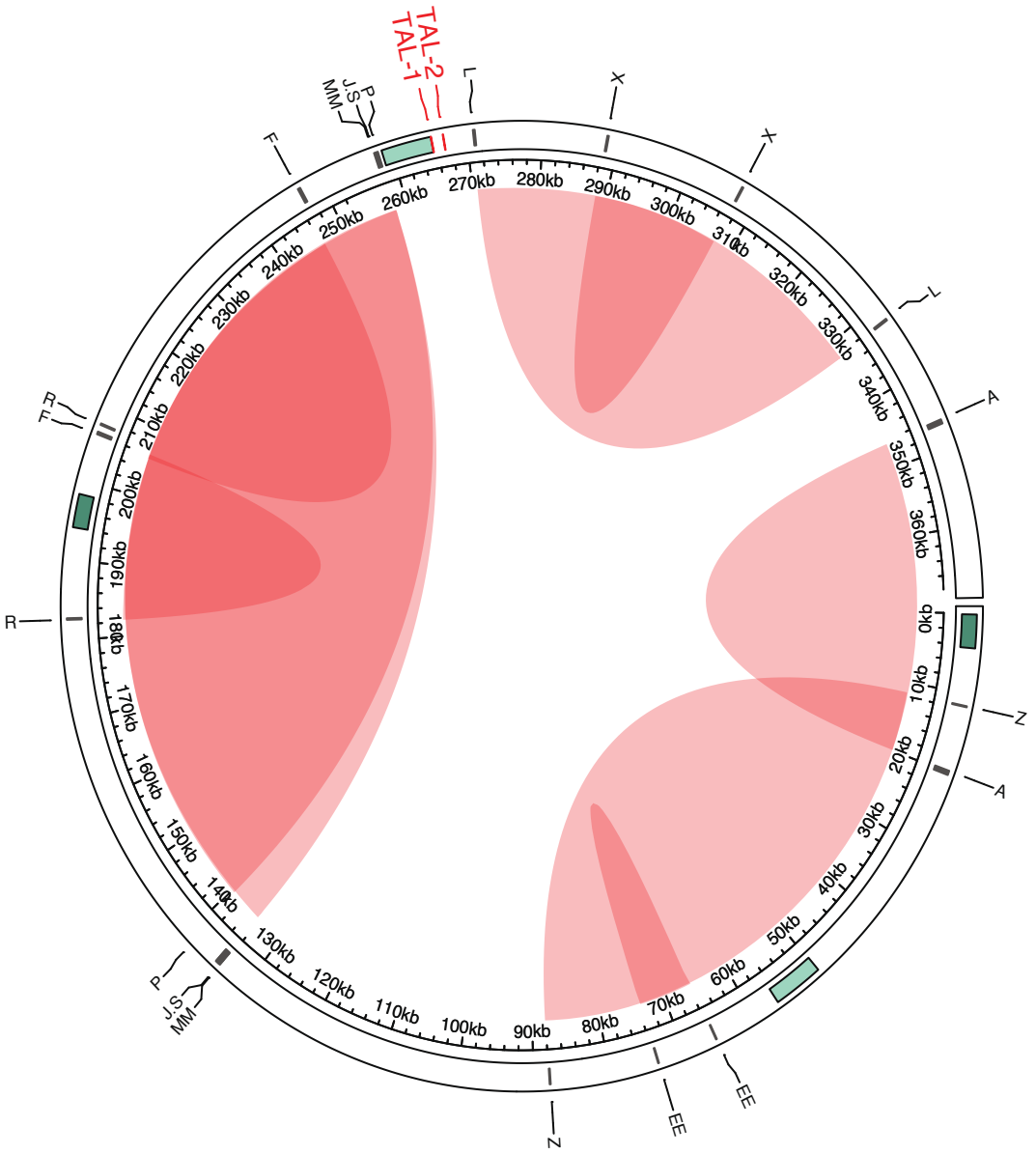


Figure 20 Représentation circulaire du génome mitochondrial d'Arabidopsis Col-0, indiquant en rouge les sous-génomes qui pourraient être formés par recombinaison homologue impliquant des répétitions directes.

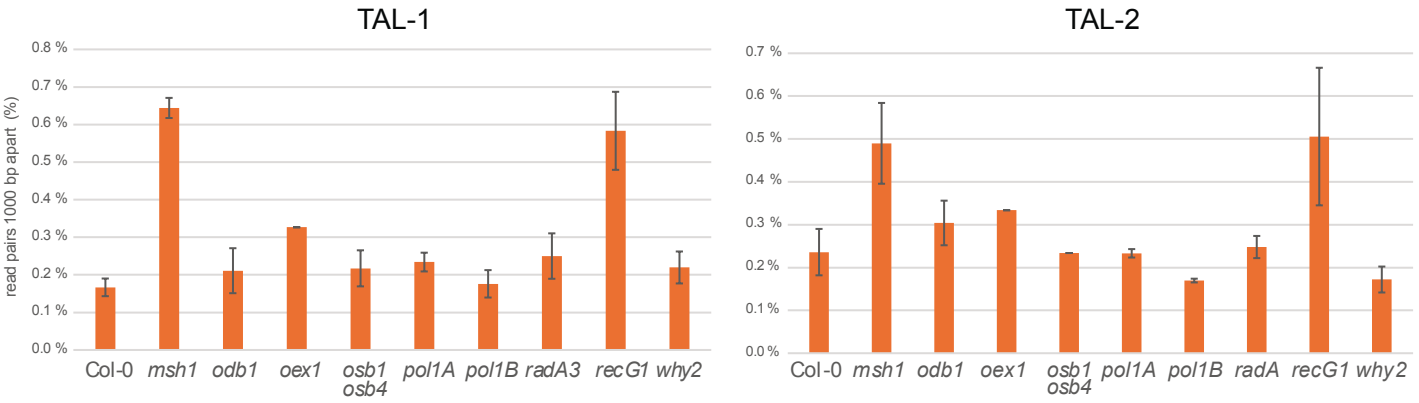


Figure 21. Proportion des paires de lecture Illumina distantes de plus de 1000 pb, indiquant des variantes structurelles du génome mitochondrial, dans les mutants transformés avec les constructions TAL-1 ou TAL-2. Calculée par rapport au nombre total de séquences cartographiées sur le génome mitochondrial WT Col-0.

Rokas KUBILINSKAS

MitoTALENs to explore mitochondrial DNA repair and segregation

Résumé

Pendant longtemps, il n'était pas possible de manipuler le génome mitochondrial des plantes (mtDNA), jusqu'aux récentes avancées en édition des génomes utilisant des "Transcription Activator-Like Effector Nucleases" (TALEN). Dans ce travail, j'ai utilisé des TALENs spécifiquement ciblés sur les mitochondries (mitoTALENs) pour étudier la réparation et la ségrégation du mtDNA des plantes. Les constructions de mitoTALEN ont été introduites dans 10 lignées mutantes différentes d'*Arabidopsis thaliana*, déficientes en divers facteurs impliqués dans la réparation mitochondriale des plantes par recombinaison homologue. Les lignées résultantes ont été analysées par séquençage Illumina et par des approches de qPCR. Chez les plantes de type sauvage, les cassures double brin (DSB) de l'ADN mitochondrial induites par les mitoTALENs ont été réparées par recombinaison homologue, entraînant le remplacement de la région contenant la DSB par une séquence distale, non-affectée, du mtDNA, flanquée par les mêmes séquences répétées. Chez les mutants déficients en facteurs de réparation, la réparation pourrait se déplacer vers des voies alternatives, telles que le "Single-Strand Annealing" (SSA) et "Microhomology-mediated recombination" (MHMR). De plus, chez certains mutants, les données n'ont révélé aucune trace de réparation des DSB, mais ont plutôt suggéré que les plantes déficientes en facteurs de réparation essentiels pourraient survivre en reconstituant un génome mitochondrial alternatif viable, à partir de sous-génomes préexistants se répliquant de manière autonome.

Mots-clés : mitochondries végétales, réparation de l'ADN mitochondrial, ségrégation de l'ADN mitochondrial, génie génétique, NGS, mitoTALEN, recombinaison homologue, SSA, MHMR

Résumé en anglais

For long, the plant mitochondrial genome (mtDNA) was not amenable to manipulation, until recent advancements in genome engineering using Transcription Activator-Like Effector Nucleases (TALEN). In this work I used TALENs specifically targeted to mitochondria (mitoTALENs) to study plant mtDNA repair and segregation. MitoTALEN constructs were transformed into the background of 10 different *Arabidopsis thaliana* mutant lines, deficient in various factors involved in plant mitochondrial repair by homologous recombination. The resulting lines were analysed by Illumina sequencing and qPCR approaches. In wild type plants, the mtDNA double-strand-break (DSB) induced by MitoTALENs was repaired by homologous recombination, resulting in the replacement of the region containing the DSB by a distal unaffected sequence of the mtDNA, flanked by the same set of repeated sequences. In mutants deficient in repair factors, repair could shift to alternative pathways, such as Single-Strand Annealing (SSA) and Microhomology-mediated recombination (MHMR). Furthermore, in some mutants, the data revealed no evidence of DSB repair, but rather suggested that plants deficient in key repair factors could survive by reconstituting an alternative viable mitochondrial genome, from pre-existing autonomously replicating sub-genomes.

Keywords: plant mitochondria, mtDNA repair, mtDNA segregation, Genome engineering, NGS, mitoTALEN, Homologous recombination, Single-Strand Annealing, Microhomology mediated recombination