# R-Loop control and mitochondria genome stability requires the 5'-3' exonuclease/flap-endonuclease OEX1

Déborah Schatz-Daas[1], Frédérique Weber-Lotfi[1*], Anaïs Le Blevenec[1], Fabio G. Moratti[2], Pierre Mercier[1], Rana Khalid Iqbal[1], Ralph Bock[2] André Dietrich[1], José M. Gualberto[1*]

[1]Institut de Biologie Moléculaire des Plantes, CNRS, Université de Strasbourg, Strasbourg, France
[2]Max Planck Institute of Molecular Plant Physiology, Potsdam-Golm, Germany
*Corresponding authors: lotfif@unistra.fr, jose.gualberto@cnrs.fr

Short title: Multiple Functions of OEX1 in mitochondrial genome maintenance and stability

**Abstract**

The maintenance of the plant organelle genomes involves factors that are mostly inherited from their bacterial symbiotic ancestors. In bacteria, a major player in genome maintenance is a 5-3 exonuclease/flap-endonuclease required for multiple replication and repair functions. This activity is given by the 5-3 exonuclease domain of DNA Polymerase I. Plant organellar DNA polymerases POL1A and POL1B are derived from DNA Pol I, but lack this domain. In Arabidopsis, proteins similar to this missing domain were identified, named OEX1 and OEX2 (Organellar Exonucleases 1 and 2), targeted to mitochondria and chloroplasts, respectively. An *oex1* mutant shows severe developmental and fertility defects, which aggravate in subsequent generations. These defects correlate with mitochondrial genome (mtDNA) instability, resulting in differential segregation of subgenomes generated by recombination, suggesting that OEX1 processes replication and/or recombination intermediates whose accumulation compromises genome stability. OEX1 is expressed in two isoforms via alternative splicing, which can differentially interact with POL1A and POL1B and variably affect mtDNA repair.

Recombinant OEX1 digests double-strand DNA in a 5-3 direction and has flap-endonuclease activity, a key function in replication and repair processes. In addition, OEX1 has a high affinity for RNA:DNA hybrid substrates, rapidly degrading RNA in Okazaki-like structures and in R-loops. Consistent with a role in suppressing R-loops, *oex1* mutant plants accumulate R-loops in highly transcribed regions of the mtDNA. Overall, our results support that OEX1 fulfills multiple important roles in the processes required to maintain mtDNA stability.

## Introduction

Mitochondria and chloroplasts have their own genomes (mtDNA and cpDNA), remnants of their past autonomous life. The current mitogenomes contain only a limited number of genes (37 in humans and 50-60 in plants), yet these genes are crucial for cell survival, as they code for essential subunits of OXPHOS complexes or are required for their expression (Knoop, 2004; Chevigny et al., 2020). Although plant and animal mtDNA code for a similar number of genes, plant mitogenomes are considerably larger, with sizes ranging from 200 kb up to 11 Mb (Sloan et al., 2012). They have been observed in diverse structural and organizational forms, including linear, circular, and branched configurations (Oldenburg and Bendich, 1996; Backert et al., 1997; Manchekar et al., 2006). These specificities in size and structure result from rearrangements through homologous recombination (HR) involving repeated sequences, which are abundant in plant mitogenomes (Kubo and Newton, 2008). These repeated sequences, ranging in size from several kilobases to less than 100 bp, engage in HR with differential frequency according to their size and genomic localization, leading to the highly dynamic structure of the higher plant mitogenomes (Small et al., 1989; Hanson and Bentolila, 2004). Recombination is also the main driving force for the rapid evolution of the plant mtDNA organization, which can significantly differ between closely related species and even among different accessions of the same species (Mower et al., 2007; Fertet et al., 2021). Recombination processes also facilitate the incorporation of external sequences through horizontal transfer (Rice et al., 2013; Garcia et al., 2021).

Despite its organizational plasticity, plant mtDNA coding sequences remain remarkably conserved across most species, indicating the presence of highly effective maintenance and repair mechanisms. Both Base Excision Repair (BER) and HR-dependent repair processes have been revealed in plant mitochondria (Boesch et al., 2009; Ferrando et al., 2018; Trasvina-Arenas et al., 2018). In particular, Break-Induced replication (BIR) and illegitimate recombination through Microhomology-mediated BIR (MMBIR) have been proposed to be predominant processes in driving both replication and repair of plant mitogenomes (Cappadocia et al., 2010; Christensen, 2013; Gandini et al., 2023). At the generational scale, defects in factors that influence HR efficiency or specificity can significantly affect plant fitness, resilience to stress, and fertility (Gualberto and Newton, 2017).

Due to its symbiotic origin, mtDNA replication and stability involve a patchwork of factors derived from its prokaryotic ancestor, from its host, or that might be of bacteriophage origin (Shutt and Gray, 2006). These factors vary between species, with plants, fungi, and animals each possessing distinct sets of proteins required for the maintenance of their respective mitogenomes. In plants, this complexity is further compounded by the sharing of genes for the maintenance of the plastid genome. As an example, in animals, the mtDNA is replicated by a single DNA polymerase, DNA Pol Gamma, while higher plants such as Arabidopsis have two organellar DNA polymerases (POL1A and POL1B), derived from prokaryotic DNA Pol I. They are required for the replication of both the mtDNA and the cpDNA and have evolved into multifunctional enzymes that differentially participate in replication and repair processes (Parent et al., 2011; Peralta-Castro et al., 2020). Unlike animal DNA Pol Gamma, which is a heterodimeric enzyme constituted by a catalytic core and an accessory factor (Kaguni, 2004), plant POL1A and POL1B have no apparent need for accessory factors.

Plant mitochondrial HR pathways recruit factors that many are orthologs of bacterial proteins, with isoforms targeted to mitochondria, chloroplasts, or dual targeted to both organelles. These include RECA-like recombinases (Miller-Messmer et al., 2012), SSB single-strand DNA-binding proteins (Edmondson et al., 2005; Qian et al., 2022), a MutS-like (MSH1) involved in the rejection of homoeologous recombination intermediates (Wu et al., 2020; Penafiel-Ayala et al., 2023) , and the RECG and RADA branch-migration helicases (Odahara et al., 2015; Wallet et al., 2015; Chevigny et al., 2022). Some are essential for plant development, such as RECA2, while mutation of several others (ex. MSH1, RECG, RECA3, RADA…) results in increased promiscuous recombination involving small repeats, with major impacts on plant fitness (Gualberto and Newton, 2017).

However, key nucleases that should play critical roles in replication and repair have yet to be identified in plant mitochondria. Such activities should be critical during replication, for the processing of Okazaki-like primers prior to ligation of the neo-synthesized DNA segments, or otherwise the residual RNA segments would lead to severe genome instability. Their removal could be either by strand displacement by DNA polymerase, which by generating a flap structure would recruit a Flap-endonuclease for its processing (Zheng and Shen, 2011; Ma et al., 2022), or by RNase H, which degrades the RNA segments hybridized to DNA (Randall et al., 2019). Both processes can coexist, as in animal mitochondria, where Flap endonuclease FEN1, RNase H1, endonuclease DNA2, and ssDNA-specific nuclease MGME1 play roles in potentially redundant

pathways for primer removal (Uhler and Falkenberg, 2015). RNAse H can, in addition, be involved in the control of R-loops, three-stranded structures generated during transcription in which an RNA segment is hybridized to one of the DNA strands (Roy and Lieber, 2009; Ginno et al., 2012). If not regulated, R-loops and unprocessed Okazaki primers can lead to replication fork stalling, genome rearrangements, and increased recombination (Gan et al., 2011; Aguilera and Garcia-Muse, 2012; Wimberly et al., 2013). To date, no flap-endonuclease or functional analogs of MGME1 or DNA2 have been identified in plant mitochondria. The Arabidopsis RNase H RNH1B has been described as targeted to mitochondria, but its mutant is not affected in development or mtDNA stability, as would be expected if this activity was crucial for mtDNA maintenance (Cheng et al., 2021). By comparison, knockout of human mitochondrial RNase H1 results in mtDNA depletion and embryonic lethality (Misic et al., 2022).

Additionally, a FEN or 5'-3' exonuclease should be involved in plant mitochondrial HR-dependent repair pathways. HR repair of double-strand breaks is initiated by strand recession by a 5'-3' exonuclease, which remains to be identified in plant mitochondria, while a FEN1-like flap endonuclease is a key factor for Single Strand Annealing (SSA) and the processing of recombination intermediates. Furthermore, a FEN or 5'-3' exonuclease may also play a role in long-patch BER (Ranalli et al., 2002).

Here, we describe OEX1, an Arabidopsis mitochondrial nuclease conserved in the entire plant lineage and phylogenetically related to the 5'-3' Exo activity of prokaryotic DNA Pol I. OEX1 is essential for mtDNA stability, and its loss leads to severe plant growth defects. *In vitro*, it is capable of degrading dsDNA distributively from 5' to 3' and has efficient FEN activity, capable of processing substrates generated by the displacement activity of DNA polymerase. From activity tests and DRIP-RT-qPCR assays, we show that OEX1 can degrade RNA in RNA:DNA hybrids and that *oex1* mutants have increased accumulation of R-loops within highly transcribed mtDNA regions. OEX1 exists as two isoforms resulting from alternative splicing. These have similar substrate specificities *in vitro*, but differ in processivity and interaction with the two organellar DNA polymerases. In agreement with the differential roles of the two OEX1 isoforms, they differentially affect mtDNA segregation and repair. Our results show that OEX1 is a jack-of-all-trades exonuclease/flap endonuclease with multiple roles in mtDNA maintenance.

## Results

*Two Arabidopsis genes code for potential organellar exonucleases.*

The plant organellar DNA polymerases POL1A and POL1B1B share a relationship with bacterial DNA polymerase I (DNA Pol I) (Kimura et al., 2002; Mori et al., 2005; Ono et al., 2007). However, unlike bacterial polymerases belonging to the Pol I family, they lack an N-terminal 5'-3' exonuclease domain. In bacteria, this domain plays roles in DNA repair and the maturation of Okazaki fragments during replication (Lovett, 2011). The importance of this activity for maintaining the integrity of bacterial genomes is demonstrated by the conditional lethality observed in corresponding mutants, while mutations affecting the DNA polymerase or 3'-5' exonuclease domains of DNA Pol I are viable (Lehman and Uyemura, 1976). Synthetic lethality was also observed in Pol I knockout (KO) mutants and mutants affecting the *YgdG* gene, which encodes the sequence-related nuclease Exo IX (Fukushima et al., 2007). A screen of the Arabidopsis genome for potential counterparts of the 5'-3' exonuclease domain of bacterial DNA Pol I revealed two candidate genes, At3g52050 and At1g34380, which we named *OEX1* and *OEX2*, respectively, for Organellar Exonuclease 1 and 2 (Figure 1A). These genes encode proteins with a distinctive PIN-like 5' nuclease domain and an H3TH (helix-3-turn-helix) DNA-binding domain, a structure also found in the 5'-3' exonuclease domain of type A DNA polymerases, eukaryotic flap endonuclease FEN1 and RNase H of bacteriophage T4 (Matelska et al., 2017). Phylogenetic analysis revealed that OEX1 is present in all groups of the green lineage, including land plants and several green algae. In contrast, OEX2 is exclusively found in land plants (Figure 1B). The similarity between OEX1 and OEX2 is lower than the similarity between OEX1 and the orthologous bacterial sequences (Figure 1C), suggesting that OEX2 was acquired independently from OEX1, probably later in evolution through horizontal gene transfer.

*OEX1 is targeted to mitochondria and OEX2 to chloroplasts*

Both OEX1 and OEX2 feature N-terminal extensions compared to their bacterial homologs. Predictions suggest that these extensions are mitochondrial and chloroplast targeting sequences, respectively (http://suba.plantenergy.uwa.edu.au). To confirm this, we expressed fusion proteins OEX1-GFP, OEX1-mCherry, and OEX2-GFP. In Arabidopsis transgenic plants, constitutive expression of OEX1-GFP or of OEX1-mCherry showed mitochondrial targeting, confirmed by co-localization with the mitochondrial marker mt-GFP (Figures 2A and 2B). DAPI staining additionally revealed co-localization of OEX1-mCherry with mitochondrial nucleoids. This

observation was facilitated by examining the leaves of plants grown under conditions that induce elongated mitochondria, as described (Jaipargas et al., 2015). Constitutive expression of OEX2-GFP in Arabidopsis was unsuccessful, but transient expression in leaf epidermal cells showed strict chloroplast localization, mainly in speckles characteristic of nucleoids (Figure 2B). The nucleoid localization of Arabidopsis OEX2 is in agreement with the enrichment of the maize OEX2 ortholog in the nucleoid proteome of maize plastids (Majeran et al., 2012). These results suggest that OEX1 and OEX2 are isoforms that likely perform similar roles in the two cytoplasmic genomic compartments of plants.

*OEX1 is required for normal plant development and fertility.*

We further focused on characterizing mitochondrial OEX1. Differential OEX1 gene expression in various tissues and during development was investigated by expressing the ß-glucuronidase (GUS) gene under the control of the *OEX1* promoter. The observations showed expression in cotyledons and very young developing leaves, as well as in the central vascular cylinder of roots, in young pistils, in stamens and pods (Figure S1). No expression was detected in gametophytic tissues such as pollen and ovules. These results are consistent with published RNAseq data, which showed that OEX1 is predominantly expressed in young leaves, petals, and pistils (Klepikova et al., 2016).

The consequences of *OEX*1 loss were analyzed on a T-DNA insertion mutant. From publicly available collections, we identified a single Arabidopsis mutant line that we could validate, *oex1-1* (GABI_911E05), with a T-DNA insertion that was mapped to the ninth exon of *OEX1*. This mutation disrupts the protein-coding sequence at the level of the PIN-like nuclease active site. RT-PCR showed that the last six exons are no longer transcribed, resulting in the loss of 152 codons out of 425, including the entire H3TH DNA-binding domain, indicating that *oex1-1* is a true knock-out line. Homozygous *oex1-1* plants are viable, but severely impaired in development (Figure 3A). Compared to WT plants, they display stunted growth, distorted leaves and flowers, and have very low fertility. Although *oex1-1* plants can bolt and produce flowers, the flowers are abnormal, with anthers that produce small amounts of pollen and pistils often curved with few pollen grains attached to the papilla. The small siliques mainly contain aborted embryos, and the few mature seeds obtained are mostly unviable (Figure 3A). The observed phenotypes are fully penetrant, affecting all homozygous *oex1-1* plants, although with some variation in the severity of the phenotypes among individual plants (Figure 3A). In following generations the

phenotypes worsened, and third-generation homozygous mutants were extremely affected in growth and completely unable to flower. Reciprocal backcross attempts with WT plants were unsuccessful, suggesting that both male and female *oex1-1* gametes are affected. Similar phenotypes have been described for mutants of other genes involved in the maintenance of Arabidopsis mtDNA, particularly *radA* plants deficient in the RADA branch migration helicase, which exhibit identical developmental phenotypes with full penetrance as *oex1* mutants (Chevigny et al., 2022).

To explore possible changes in mitochondria morphology, *oex1-1* rosette leaf sections were examined by transmission electron microscopy (TEM). The TEM images showed that in *oex1-1* plants mitochondria were smaller and more electron-dense, when compared to mitochondria in WT plants of equivalent size, while the morphology of chloroplasts remained unaffected (Figure S2).

The mutant phenotype was fully complemented by expressing the *OEX1* cDNA under the native *OEX1* promoter. Complemented plants (*oex1 OEX1*) were phenotypically normal and indistinguishable from WT plants (Figure 3B). Because OEX1-GFP transient expression suggested a possible dual localization in chloroplasts, hemicomplementation experiments were also conducted, where *oex1-1* plants were transformed with constructs designed to specifically target OEX1 to either mitochondria or chloroplasts (see Methods). As shown in Figure 3B, full complementation of the mutant phenotype was achieved when OEX1 was specifically targeted to mitochondria. However, no complementation was observed when OEX1 was only targeted to chloroplasts. Together, all results support that the *oex1-1* mutation primarily affects mitochondria.

*OEX1 plays a critical role in maintaining mitochondrial genome stability*

Mutants deficient in factors involved in mtDNA maintenance and repair can accumulate cross-over products from recombination mediated by small (100-500 bp) repeated sequences, leading to genome instability that correlates with varying severity of developmental defects, similar to those observed in *oex1* (Gualberto and Newton, 2017). Such molecular phenotype was therefore tested in *oex1* plants. The mtDNA sequence from T2 *oex1* plants (first homozygote mutant generation) revealed significant changes in the relative stoichiometry of mtDNA sequences (Figure 4A). In contrast, the plastid genome (cpDNA) sequence showed no evidence of instability (Figure 4B). Several mtDNA regions with increased copy number in *oex1* are flanked by directly

oriented repeats (e.g., repeats A (556 bp), L (249 bp), and EE (127 bp)). This suggests the looping-out of subgenomes through recombination, a process seen in other mtDNA recombination mutants (Wallet et al., 2015; Chevigny et al., 2022). To confirm this, the accumulation of cross-over products resulting from recombination involving repeats L, F, and EE was tested. These repeats were selected because they are consistently mobilized in other mutants and are small enough for qPCR analysis. As expected, a significant increase in cross-over products was observed compared to WT levels (Figure 4C). An asymmetrical accumulation of primarily one type of cross-over was seen, consistent with previous findings in mutants with mtDNA recombination defects. This asymmetry is likely due to the involvement of the error-prone break-induced replication (BIR) repair pathway (Christensen, 2013; Gualberto and Newton, 2017; Kockler et al., 2021).

To assess the range of mtDNA changes in individual *oex1* plant mutants, qPCR was used to measure the relative copy number of different mtDNA regions, as previously described (Schatz-Daas et al., 2022). Individual *oex1* T3 with severe growth and developmental impairments were analyzed (Figure 4D). All plants showed high mtDNA instability, with significant changes in the stoichiometry of different mtDNA regions, though the extent of these changes varied between plants (Figure 4E). Most had changes in the regions flanked by repeats A, L, F, and EE (plants #1 to #4 in Figure 4E). However, some plants (plants #5 and #6) showed a large increase in copy number of sequences not associated with annotated repeats. Further investigation revealed that these regions were subgenomes formed by illegitimate recombination., with one involving an imperfect repeat of 85 bp and the other a microhomology of 14 bp (Figure 4E). These results indicate that in the absence of OEX1, illegitimate recombination activities are unleashed, which can recruit small repeats and microhomologies. This leads to the formation of subgenomes that replicate autonomously through unknown mechanisms. Similar consequences of recombination dysfunction were described in the *recG1* and *radA* mutants (Wallet et al., 2015; Chevigny et al., 2022). Thus, the severe growth phenotype of *oex1* plants correlates to defects in mtDNA maintenance, which impact mtDNA stability through recombination.

*Two OEX1 isoforms generated by alternative splicing*

Cloning and sequence of *OEX1* cDNAs from Arabidopsis leaves revealed two mRNA variants resulting from alternative splicing of the 6th exon. This splicing was observed in 6 out of 11 leaf cDNA clones and is supported by publicly available EST sequences (https://www.arabidopsis.org).

The two mRNA variants encode proteins that differ by 18 amino acids, which we named OEX1a and OEX1b, respectively (Figure 5A).

The 6th exon-encoded extension in OEX1a is conserved across all flowering plants, though sequence conservation is limited. Corresponding transcript variants have been identified, among others, in soybean (XM_006574959.4), grapevine (XM_010652395.3), maize (XM_008660501.4), rice (XM_015776708.3) and Amborella (XM_020676298.1). However, this extension is absent in OEX2 and in the 5'-3' exonuclease domain of bacterial DNA Pol I (Figure S3A). Structural modeling suggests that the extension is located near the active site and H3TH DNA-binding domains (Figure S3B), potentially influencing substrate specificity and/or OEX1 activity. The relative abundance of the two transcripts in different tissues was assessed by RT-qPCR, using variant-specific primers (Figure S3C). The results indicated that both mRNA forms are present at similar levels in most tissues. However, the ratio of OEX1b to OEX1a was higher in flower buds, implying that OEX1 may be developmentally regulated through alternative splicing (Figure S3C).

The impact of the two OEX1 isoforms on mtDNA maintenance was further examined by complementing the *oex-1* mutant with either OEX1a or OEX1b cDNAs, under the control of the OEX1 promoter. Both isoforms successfully complemented the development defects of *oex1-1*, and the resulting plants (*cOEX1a* and *cOEX1b*) were indistinguishable from WT (Figure 5B), suggesting functional redundancy. However, at the molecular level, while stoichiometric mtDNA replication was fully restored to wild-type levels in *cOEX1a* plants, *cOEX1b* plants showed a slight imbalance in the relative abundance of the two mtDNA subgenomes flanked by large repeat 2 (LR2, of 4.2 kb in size) (Figure 5C). The complemented *cOEX1a* and *cOEX1b* plants were further challenged for repair capacity by growth under genotoxic stress induced by ciprofloxacin (CIP), as described (Schatz-Daas et al., 2022). At the diagnostic concentration of 0.5 µM CIP, most WT plants were still able to develop normal first leaves, whereas cOEX1b plants, and to a lesser extent cOEX1a plants, either failed to develop first leaves or those leaves were mostly wrinkled and distorted. Thus, defects in replication, repair, and/or segregation of mtDNA subgenomes were not fully restored in *cOEX1b* plants, suggesting that OEX1a and OEX1b are two partially redundant isoforms with potentially distinct activities or substrate affinities, leading to different effects on mtDNA maintenance.

*OEX1a and OEX1b may differentially interact with the organellar DNA polymerases*

Because of the relation of OEX1 to the 5'-3' exo domain of type A bacterial DNA polymerases, a yeast two-hybrid (Y2H) assay was used to test whether OEX1 could form complexes with the organellar DNA polymerases POL1A or POL1B. No interaction was observed when OEX1a or OEX1b were fused to the GAL4 activating domain (AD) and POL1A or POL1B to the binding domain (BD). However, in the reverse configuration, both BD-OEX1a and BD-OEX1b interacted with AD-POL1A, while only BD-OEX1b interacted with AD-POL1B (Figure S4B). These results further suggest different *in vivo* roles for OEX1a and OEX1b. Because POL1A and POL1B were described as having differential roles in repair (Parent et al., 2011), the different effects of OEX1a and OEX1b in repair align with their differential interactions with DNA polymerases.

Yet, proteomic data (https://complexomemap.de/; (Senkler et al., 2017)) showed that OEX1 migrates as a single band in Blue-Native gels, much smaller than 100 kDa (Figure S4C). This size matches that of monomeric OEX1 (about 43 kDa) rather than a complex with POL1A or POL1B, both of which are larger than 100 kDa. A quantitative proteomic analysis further revealed that OEX1 is more abundant than the DNA polymerases, with about 40 molecules of OEX1 per mitochondrion compared to 5 and 2 molecules of POL1A and POL1B, respectively (Fuchs et al., 2019). This difference in abundance supports the idea that OEX1 primarily functions as a single-subunit holoenzyme and does not form stable complexes with the DNA polymerases. Nevertheless, the Y2H assay suggests that OEX1 may transiently interact with POL1A or POL1B, with the two OEX1 isoforms exhibiting different affinities for these polymerases.

*The loss of OEX1 has no apparent detrimental effect on mitochondrial transcripts*

To assess whether mtDNA instability in *oex1* affects mitochondrial transcript levels, they were analyzed by RT-qPCR, as described elsewhere (Baudry, 2022). Results revealed changes in the abundance of some transcripts, but mostly increases, with no significant reduction in any gene transcript (Figure 6). For genes with introns, exon borders were examined, but no splicing defects were found. This suggests that the severe growth defects in *oex1* plants are unlikely to be caused by problems in the expression of a specific mitochondrial gene. Interestingly, several gene transcripts, including *rps4*, *rps7*, *ccmC*, *matR*, *ccmFN1*, and *ccmFN2,* were more than two-fold higher in *oex1*. In certain cases, this increase could be attributed to increased gene copy number, such as for *ccmC*, *ccmFN1*, and *ccmFN2,* which are clustered in the subgenome defined by repeat L, or for *atp1* that is coded by the episome EE. In plants complemented with either *OEX1a* or *OEX1b*, transcript levels were restored to WT levels (Figure 6).

*OEX1 is a 5’-3’ exonuclease/Flap endonuclease*

Recombinant OEX1a and OEX1b proteins were expressed in *E. coli* and purified as described in the Methods section (Figure S5A). A catalytic mutant protein (D250A) was also expressed and purified (Figure S5B). The activities of OEX1a and OEX1b were tested on various DNA structures, selected for their relevance to the potential roles of OEX1 in mtDNA maintenance.

Exonuclease activity was first tested on double-strand DNA (dsDNA), and the optimal conditions for OEX1 activity on 5’-labelled dsDNA were determined, regarding temperature, pH and divalent ion requirements (Figure S6). OEX1a rapidly degraded a 25-mer dsDNA substrate, showing strict 5’-3’ exonuclease activity, with no evidence of endonuclease activity, as no intermediate products were seen using 5’-labelled substrates (Figure 7A). To confirm this activity was from OEX1a and not from bacterial contaminants, the D250A catalytic mutant was tested under identiacla conditions and, as expected, showed no activity. OEX1b showed the same specificity as OEX1a, but with less processivity, and a significant portion of the substrate remained intact after 30 minutes, (Figure 7B). The degradation pattern on a 3’-labelled dsDNA substrate (5x[GTCA]GTCCC) suggested sequence-dependant distributive activity, with pausing at G-C base pairs. This was confirmed by testing substrates with 1, 2 or 3 consecutive Gs (Figure S7A).

Another function tested was flap-endonuclease activity, which had not been previously identified in plant organelles, but necessary for processing recombination intermediates. Furthermore, during long patch base excision repair (BER), a flap endonuclease may be required to process the flap formed by strand displacement during DNA synthesis. A flap-endonuclease may also be needed for the maturation of Okazaki fragments, which are displaced by DNA polymerase during replication. A flap-like substrate was prepared and labeled either at the 5’ or 3’ end of the flap strand. Precise processing of the flap should produce an 18- or 25-nucleotide labeled product, respectively. Both OEX1a and OEX1b, but not the catalytic mutant, could precisely process the flap (Figure 7C and 7D). As for the 5'-3' exonuclease activity, OEX1a showed higher processivity on flaps than OEX1b.

Following flap cleavage OEX1a released additional degradation products, suggesting further degradation of the single strand DNA (ssDNA) product. This prompted testing OEX1's activity on ssDNA. Both OEX1a and OEX1b, but not the D250A catalytic mutant, showed some 5’-3’ and 3’-5’ exonuclease activities on ssDNA substrates (Figure S8). However, these activities were much

weaker than on dsDNA, and after 30 minutes most of the ssDNA substrate remained intact. This low *in vitro* activity on ssDNA may not be biologically relevant.

We tested whether OEX1 could play a role in base excision repair (BER), using substrates with single-strand gaps mimicking BER intermediates. When tested on similar substrates with gaps of 1, 3 or 10 nucleotides, OEX1 showed minimal activity. However, it was still able to recognize and extend the gaps, regardless of their size (Figure S7B). The processivity on gaps was much lower than at the 5' end of dsDNA, suggesting that OEX1 alone is inefficient in extending single-strand gaps on DNA. To see if gap extension requires strand displacement by a DNA polymerase, the same experiment was repeated in the presence of the Klenow enzyme, added after OEX1a (Figure 7E). Under these conditions the gaps were processed much faster, indicating that it is the flap-nuclease activity of OEX1 that processes the displacement products generated by the DNA polymerase. No significant activity was observed with Klenow and the D250A catalytic mutant, confirming that the substrate processing was not due to a contaminant or the Klenow's 3'-5' exonuclease activity. Therefore, *in vitro* results on DNA substrates supported OEX1's 5'-3' exonuclease activity on dsDNA, its flap-endonuclease function, and its role in gap-extension in cooperation with the displacement activity of DNA polymerase (Figure 7F).

*OEX1 has high activity on RNA:DNA hybrids and is involved in the processing of R-loops*

Given the potential role of OEX1 in the processing of Okazaki fragments, its activity on RNA:DNA hybrids was tested. When tested on an 3'-labeled RNA paired to a complementary DNA primer (RNA:DNA substrate), OEX1a rapidly degraded the RNA (Figure 8A), with apparently much higher processivity than on dsDNA (as seen in the 3'-labeled product in Figure 7A). On a substrate mimicking an Okazaki fragment, consisting of a dsDNA region separated from an RNA:DNA duplex by a small gap, OEX1a also rapidly degraded the RNA (Figure 8B). This contrasts with its weak activity on a similar substrate made only of DNA duplexes, where OEX1a required strand displacement by a DNA polymerase (Figure 7E). When tested on single-stranded RNA no significant nuclease activity was detected (Figure 8C). These results strongly suggest that OEX1 prefers RNA:DNA hybrids as substrates. These can form during replication, such as Okazaki fragments, or as by-products of transcription, leading to R-loops, which are known to destabilize genome replication (Crossley et al., 2019). To explore the potential role of OEX1 in resolving R-loops, a substrate mimicking an R-loop was used. In this substrate the 5' dsDNA ends were biotinylated, to protect them from OEX1 activity. Remarkably, OEX1a promptly degraded the

RNA component of this R-loop substrate (Figure 8D). To validate the hypothesis that one of the *in vivo* functions of OEX1 is the processing of R-loops, the accumulation of R-loops in *oex1-1* plants was tested by DNA:RNA immunoprecipitation (DRIP). Nucleic acids were extracted from both WT and *oex1-1* plants, cut into smaller pieces by restriction enzymes, and RNA:DNA hybrids were immunoprecipitated using the S9.6 antibody, which is specific for RNA:DNA hybrids. The precipitated DNA was analyzed by qPCR, targeting the highly transcribed mtDNA regions coding for the ribosomal RNAs. Results were normalized against the quantification of non-transcribed regions. To ensure the results validity, a negative control was performed by treating the samples with RNAse H prior to immunoprecipitation. In our hands, and as reported by others (Sanz and Chedin, 2019), the signal-to-background ratio in DRIP experiments is typically low. Nevertheless, a clear and significant increase was observed in *oex1-1* samples, as compared to WT. This accumulation of R-loops is potentially linked to the mtDNA instability observed in *oex1-1* plants.

**Discussion**

*Origin of OEX1 and OEX2 genes and relationship to possible bacterial counterparts*

In eubacteria, a 5'-3' exonuclease activity is essential in all species, provided either by the N-terminal domain of Pol I or by ExoIX (YpcP in *B. subtilis*) orthologs (Fukushima et al., 2007). In species such as *E. coli,* both enzymes co-exist and either is sufficient for cell survival, although a double mutation leads to synthetic lethality (Fukushima et al., 2007). Thus, in bacteria that lack an ExoIX/YpcP ortholog the 5'-3' exonuclease activity of Pol I is indispensable. Structurally, plant OEX1/2 are similar to the 5'-3' exonuclease domain of Pol I, much like ExoIX/YpcP. Some bacterial genera, corresponding to class 2 in Fukushima's classification, only have ExoIX/YpcP exonucleases and lack a repair polymerase. They are all symbiotic or infectious bacteria, much like the probable symbiotic ancestor of mitochondria (Fukushima et al., 2007). It is thus tempting to speculate that plant organelles fall on class 2 systems and that OEX1/2 derived from bacterial ExoIX/YpcP. However, phylogeny suggests that OEX1/2 are more closely related to the N-term of Pol I than to ExoIX/YpcP, implying that plant OEX1/2 evolved from a Pol I that lost its 3'-5' exo and polymerase domains. The phylogenetic analysis also indicates that OEX2 is not a paralog of OEX1, as their similarity is lower than with bacterial orthologs, suggesting a later acquisition.

Our results strongly supported that OEX1 is exclusively localized in the mitochondria, specifically in nucleoids. This nucleoid localization is consistent with mtDNA maintenance functions of OEX1. In agreement with the GFP results, the full complementation of the developmental and molecular phenotypes of the *oex1* mutant through hemicomplementation with OEX1 solely targeted to mitochondria indicates that OEX1 is active only in mitochondria. Regarding OEX2, the strict chloroplast localization of OEX2-GFP in nucloid-like speckles fits with the identification of the maize ortholog in the plastid nucleoid proteome (Majeran et al., 2012). Taken together, these results suggest that OEX1 and OEX2 are isoforms fulfilling equivalent functions in mitochondria and plastids. Still, they might have specialized functions, depending on their interplay with other factors such as DNA polymerases POL1A and POL1B and other nucleases such as RNase H1. Notably, chloroplastic OEX2 lacks the extension added to OEX1a by alternative splicing.

*Major potential role of OEX1 in the mtDNA Replication*

To date, details of how mtDNA is replicated has been elucidated in only a few model systems. In animals, replication of the small and compact mtDNA is highly derived compared to other systems. A mitochondrial RNase H1 creates RNA primers by cleaving longer transcripts in R-

loops, which are stabilized by mitochondrial ssDNA-binding protein mtSSB (Falkenberg et al., 2024). After replication, removal of the primers requires RNase H1 and nuclease MGME1, which cuts ssDNA flaps. Mitochondrial DNA Polymerase γ (POL γ) can displace the upstream DNA end into a short flap structure, which is removed by MGME1 in combination with the 3'-5' exonuclease activity of POL γ, leaving a nick that is sealed by LIG3 (Uhler et al., 2016) (Macao et al., 2015). Because of the central role of RNase H1 in animal mitochondria, its loss results in mtDNA depletion and embryonic lethality (Misic et al., 2022). Other nucleases dually targeted to both the nucleus and mitochondria, such as FEN1 and DNA2, may also be involved.

In plants however, there is no ortholog of MGME1, and mtDNA replication and maintenance relies on many prokaryotic-like factors. There is still no definite model for how the complex plant mitogenomes replicates, but at present OEX1 is the primary candidate enzyme for processing replication RNA primers bound to DNA during replication. These RNA fragments, produced by a primase, enable DNA polymerases to bind and start DNA synthesis. This in both unidirectional replication, such as for the D-loop responsible for initiation of mtDNA replication in animals (Falkenberg et al., 2024), and in bidirectional DNA replication, where Okazaki fragments are required for lagging strand synthesis. In bacteria the elimination of the RNA part of such primers is one of the tasks of the 5’-3’ exo domain of DNA Pol I, which is actually a structure-specific endonuclease that cleaves at the junction between a 5’ single strand and a duplex. The DNA polymerase I activity at a nick displaces the DNA ahead of the enzyme, forming a 5’ flap that is subsequently cleaved by the 5’ endo/exonuclease. ExoIX/YpcP orthologs have also been described as having similar structure-specific FEN activity (Anstey-Gilbert et al., 2013; Randall et al., 2019). The FEN and gap-extension activity of OEX1, promoted by DNA polymerase, fits with an ability to cleave the small flaps formed by strand displacement when DNA polymerase reaches a downstream 5'-end. Such activities support the idea that OEX1 plays a predominant role in replication. Consistent with this role, the recombinant protein recognized and efficiently processed the RNA strand of an RNA:DNA hybrid *in vitro*, as well as structures mimicking Okazaki fragment in a replication fork. Through these experiments OEX1 was more processive on RNA than on DNA and was only processive on double strand DNA containing a gap when in the presence of DNA polymerase activity. These results suggest that the elimination of RNA sequences hybridized to DNA is a main role of OEX1. The maturation of plant mitochondrial Okazaki fragments by OEX1 could therefore be either through exonucleolytic degradation or FEN activity on the flaps formed by strand displacement by DNA polymerase. Both systems may

coexist. Interestingly, bacterial YpcP was also shown to be most active on RNA:DNA duplexes, working on Okazaki fragments and flap substrates (Randall et al., 2019). Organellar RNase H1 enzymes may also contribute to the removal of replication primers. However, the loss of mitochondrial RNase H1B alone does not have adverse effects on plant development, in stark contrast to the vital importance of mitochondrial RNase H1 in animals (Cheng et al., 2021; Misic et al., 2022).

*Other potential roles of OEX1 in the mtDNA maintenance*

HR is the main mechanism in the repair and maintenance of plant mitogenomes, requiring several nucleases for function. A 5'-3' exonuclease is typically needed for strand resection at double-strand breaks. In bacteria this function is handled by the RecBCD complex or by RecJ in the RecFOR pathway, both of which are highly processive (Kowalczykowski, 2015). OEX1 can recognize blunt-ended dsDNA and degrade the 5' to 3' strand, but the *in vitro* tests showed that OEX1 has rather a distributive and not processive activity on these substrates, making its involvement in strand resection unlikely. Nonetheless, it is possible that OEX1, *in vivo,* works together with other factors such as helicases and ssDNA-binding proteins to enhance its processivity, similar to the bacterial RecBCD and RecFOR complexes. FEN activity is also required in HR pathways, particularly in single-strand annealing (SSA) or in the resolution of D-loops,. Therefore, OEX1 is a primary candidate for resolving such recombination intermediates.

The FEN activity of OEX1 is also likely required for BER and ribonucleotide excision repair (RER). BER is a principal mechanism for repairing damaged mtDNA in both plant and animal mitochondria (Szczesny et al., 2008; Boesch et al., 2011; Ferrando et al., 2018). Okazaki fragment processing and long-patch BER share similar mechanisms, as both involve DNA polymerase displacing the existing strand to form a flap, which is then cleaved by a FEN. Our results show that OEX1 can recognize and cleave such flaps, as required in long-patch BER. In long-patch BER, two models can be proposed: in the first, a repair-associated DNA polymerase such as POL1B fills the gap left by glycolsylases and AP-endonuclease, displacing the existing strand to form a flap that is then cleaved by the FEN activity of OEX1. In the second model, POL1B and OEX1 work together: POL1B fills the gap while OEX1 degrades the strand downstream of the lesion. Our enzymatic tests suggest that OEX1 can recognize and extend gaps, but inefficiently, making the first model more likely. Like for BER, the FEN activity of OEX1 could also be involved in RER, a repair process responsible for the removal of ribonucleotides mis-inserted instead of

deoxyribonucleotides, a most frequently occurring error during DNA synthesis. If left unrepaired such DNA-embedded ribonucleotides ultimately lead to genome instability. RER requires the action of an RNAse H to cleave the strand containing the incorporation, and of the strand-displacement activity of a DNA polymerase to create a flap that can be processed by the FEN activity of OEX1.

Finally, OEX1 is apparently implicated in the removal of R-loops, which can form in transcribed regions and are known to cause genomic instability (Gan et al., 2011; Wimberly et al., 2013). The strong *in vitro* activity of OEX1 degrading the RNA strand in RNA:DNA hybrids suggest such a role, confirmed by the DRIP-qPCR results, showing an accumulation of R-loops in the mitogenome of *oex1-1,* particularly in the highly transcribed regions coding for the rRNAs.

*The Absence of OEX1 is Critical for plant development because of mtDNA instability*

The study of the *oex1* mutant phenotypes confirmed the importance of OEX1 in plant growth, development, and fertility. By the third generation of homozygote mutants, plants could no longer produce offspring. This quasi-sterility can be attributed to severely deformed pistils and stigmas, along with reduced pollen viability, as evidenced by failed attempts to pollinate wild-type pistils with *oex1-1* pollen. Although the overall phenotype was consistent, its severity varied, correlating with mitochondrial genome instability that resulted in altered stoichiometry of mtDNA sequences. While these changes were generally similar among plants, significant differences correlated with the severity of the phenotype and likely depending on which mtDNA regions were affected. Typically, the changes in sequences stoichiometry involved sub-genomes created by recombination across small repeats larger than 100 bp. However, as seen in some *oex1* plants, much smaller and imperfect repeats as well short microhomologies can also drive recombination. Such molecules may result from the accumulation of unprocessed DNA ends in *oex1*, triggering pathological replication re-initiation via Break-induced replication (BIR) and/or Microhomology Mediated BIR (MMBIR). As previously suggested, processes that generate mtDNA breaks potentially promote error-prone repair pathways, explaining the increased heteroplasmy in recombination surveillance mutants (Marechal and Brisson, 2010; Gualberto and Newton, 2017). Similar shifts in sequences stoichiometry of sub-genomes generated by recombination have been described in mutants of other mtDNA recombination and repair factors, such as RADA, REG1, RECA2, RECA3, MSH1, and OSB1. However, only *oex1* and *radA* showed 100% penetrance of developmental defects, while most first-generation *recA3*, *msh1*,

and *osb1* mutants have WT-like phenotypes. The severe effects of *oex1* and *radA* are possibly due to their critical and non-redundant functions in mtDNA maintenance.

The changes in relative copy numbers can potentially impact the mitochondrial transcriptome. However, no mitochondrial genes have shown significantly reduced expression, and the mtDNA regions with lower copy numbers in *oex1* plants do not contain any identified genes. One possibility is that mtDNA rearrangements activate ORFs that could be toxic to mitochondria, as proposed for *radA* mutants (Chevigny et al., 2022). Conversely, several gene transcripts were found to be more abundant, possibly due to an increased copy number of the corresponding coding sequences. Thus, the harmful effects of the *oex1* mutation on plant development likely derive from problems in mtDNA replication and the stoichiometric segregation of recombination-generated subgenomes. These seem to affect mitochondrial biogenesis, though not necessarily through clear problems in gene expression. Promoter:GUS results showed that *OEX1* is primarily expressed in young, rapidly dividing tissues, and its absence leads to defects in mitochondrial biogenesis and proliferation, as revealed by TEM images. Similar effects on mitochondrial size have also been observed in mutants deficient in mitochondrial ssDNA-binding protein OSB1 (Zaegel et al., 2006). It is possible that mtDNA replication and segregation are linked to mitochondrial division, as in mammals, where nucleoids involved in active mtDNA replication localize to contact points with the endoplasmic reticulum, where mitochondrial division occurs (Lewis et al., 2016). However, this has not been demonstrated in plants. An additional component contributing to the developmental defects in mutants with impaired mtDNA maintenance might be retrograde responses triggering the expression of nuclear genes that suppress cell proliferation, as observed in animals, chloroplasts and in the plant *radA* mutant (Tigano et al., 2021; Chevigny et al., 2022; Fu et al., 2023).

*Differential roles for the two alternative-splicing OEX1 isoforms*

Both our experimental data and public transcriptomic data confirm that Arabidopsis has two OEX1 isoforms coded by alternatively spliced transcripts. These isoforms seem conserved among flowering plants, as similar variants have been found in other species including monocots, indicating biological relevance. Since the extension found in mitochondrial OEX1a is not found in chloroplast OEX2 or any of the bacterial orthologs, this raises the question of why mtDNA maintenance requires the two isoforms. Our analysis suggested developmental regulation, with a higher OEX1b/OEX1a ratio in developing flowers. Both proteins are functional, as shown by

their ability to complement the *oex1*_developmental defects. However, they affect mtDNA segregation differently, and plants complemented with OEX1b maintained a slight imbalance in mtDNA subgenomes, which was fully corrected by OEX1a. This suggested distinct roles or substrate preferences for the two isoforms. Structural modeling also supports this assumption, by placing the OEX1a extension near the active site and DNA-binding domain. Indeed, *in vitro* activity tests showed that although both isoforms process the same substrates, they differed in processivity, with OEX1a having higher specific activity on all tested substrates, including 5'-3' exonuclease on DNA ends and Flap-endonuclease (FEN). Nevertheless, we cannot rule out that OEX1b is specifically required to process certain specific recombination intermediates. The existence of two isoforms may also relate to their different interactions with other factors involved in mtDNA maintenance. Proteomic data (Senkler et al., 2017; Fuchs et al., 2019) suggest that OEX1 acts as a holoenzyme without forming stable protein complexes, though its localization in nucleoids suggests it is part of larger nucleo-protein structures where mtDNA replication, repair, and transcription occur. The requirement of DNA polymerase activity for efficient processing of DNA gaps, such as long-patch BER intermediates, further supports the idea that OEX1a and OEX1b act in coordination with other factors, but probably only establishing transient interactions. The two isoforms apparently interact differently with the organellar DNA polymerases POL1A and POL1B. Y2H results indicate that both can interact with POL1A, while only OEX1b interacts with POL1B. POL1A and POL1B have distinct roles in mtDNA maintenance, being redundant for genome replication but differing in processivity, fidelity, and involvement in repair (Parent et al., 2011; Baruch-Torres and Brieba, 2017; Ayala-Garcia et al., 2018; Garcia-Medel et al., 2019; Schatz-Daas et al., 2022). Our findings also suggest that the two OEX1 isoforms have differing roles in mtDNA repair. At present, the picture that emerges from our present work and other studies, is that mtDNA maintenance relies on the fine-tuned interplay between multiple factors, several of which (RecA recombinases, SSB and OSB ssDNA-binding proteins, DNA polymerases, and OEX nucleases) are available in different isoforms with specialized roles shaped by tissue-specific expression, activity profiles, and distinct functions in replication and repair pathways. These isoforms can arise from distinct gene paralogs, dual targeting of the same protein to both organelles, or, in the case of OEX1, from alternative splicing.

## Materials and Methods

*Plant material and constructions*

Sequence data from this article can be found in the Arabidopsis Genome Initiative or GenBank/EMBL databases under the following accession numbers: OEX1, At3g52050; OEX2, At1g34380. *Arabidopsis thaliana* T-DNA insertion line *oex1-1* (GK-911E05) derives from accession Col-0 and was obtained from the Nottingham Arabidopsis Stock Centre (NASC). Plants were grown on soil or *in vitro* under a 16 h light/8 h dark photoperiod (65–85 µmol photons m−2 s−1) at 21 °C/18 °C. For *in vitro* cultures, surface-sterilized seeds were sown on agar plates containing MS255 media (Duchefa) supplemented with 1% (w/v) sucrose and stratified during 2 days at 4°C in the dark. For DNA and RNA extraction, plant tissue frozen in liquid nitrogen were ground with a TissueLyser II (Quiagen). DNA was extracted using the cetyltrimethylammonium bromide method and RNA was extracted using TRI Reagent (Molecular Research Centre, Inc.) following the manufacturer's protocol. RNA was further extracted with phenol-chloroform before ethanol precipitation. For mutant complementation, the WT *OEX1* gene and promoter sequence were cloned into the binary vector pGWB613 and used to transform heterozygous *oex1-1* plants. For complementation with either OEX1a or OEX1b sequences, the gene sequence including the promoter, 5'UTR and the first two exons and introns was fused to the corresponding *OEX1a* or *OEX1b* cDNA sequences. For the hemicomplementation constructs, the organellar targeting sequence (OTS) of OEX1 (86 codons). was replaced with either the OTS (49 codons) of mitochondrial Alternative Oxidase 1 from soybean (AOX1; NM_001249237) or the OTS (80 codons) of the chloroplast small subunit of ribulose bisphosphate carboxylase (RBCS; At1g67090). The constructs were assembled using MultiSite Gateway cloning in vector pB7m34GW (https://gatewayvectors.vib.be/), giving AOX1:OEX1 and RBCS:OEX1 constructs. These were used for *Agrobacterium tumefaciens* transformation of heterozygous *oex1-1* plants by floral dip.

*Phylogenetic Analysis*

For the building of phylogenetic trees, bacterial and plant orthologous sequences were aligned with T-Coffe implemented in the Macvector package with default parameters (Myers-Miller dynamic algorithm, gap open penalty of -50 and gap extension penalty of -50) and a consensus maximum likelihood tree was built with IQ-TREE (http://iqtree.cibiv.univie.ac.at), from 1000 bootstrap iterations. Graphical representation and edition of the tree was with TreeDyn (v198.3).

The Arabidopsis OEX1 sequence was modelled on the structure of the 5'-3' exonuclease domain of *Thermus aquaticus* DNA Polymerase I (Kim et al., 1995), using Modeler (Sali and Blundell, 1993); http://salilab.org/modeller/about_modeller.html), and the modeled structures represented with PyMol (https://pymol.org/2/)..

*Subcellular localization and Promoter-GUS fusion*

For *in vivo* intracellular localization, the *OEX1* and *OEX2* gene sequences were cloned into the pUBC-paGFP-Dest vector, upstream and in frame with the GFP coding sequence, under the control of the UBQ10 promoter (Grefen et al., 2010). Leaf pavement cells of independent transformants were observed at the confocal microscope (Zeiss LSM700). For colocalization of OEX1 with mitochondria, the *OEX1* cDNA was amplified and cloned into the pUBC-mCherry-Dest binary vector (Grefen et al., 2010), to obtain the pUbi::OEX1-mCherry construct. The resulting plasmid was used for Agrobacterium-mediated transformation of plants expressing the mitochondrial marker pCaMV35S::mTP-GFP, corresponding to the transit peptide of ATP citrate lyase (At2g20420) fused to GFP, under the control of the 35S promoter. The resulting seeds were selected on MS medium supplemented with phosphinotricin at a final concentration of 10 mg/L. To visualize mitochondrial nucleoids, cells with elongated mitochondria were investigated. To this end, Arabidopsis seedlings expressing both pCaMV35S::mTP-GFP and pUbi::OEX1-mCherry transgenes were grown in the dark for 8 days on 0.5x MS medium without sugar, as described (Jaipargas et al., 2015). DNA was stained with DAPI solution (1 µg/mL) in phosphate-buffered saline for 30 min at 25 °C in the dark, and leaf pavement cells were imaged at a Leica TCS SP8 confocal laser-scanning microscope. Excitation of DAPI was at 405 nm, GFP and chlorophyll at 488 nm and mCherry at 561 nm. Fluorescence emission was collected at 435–475 nm for DAPI, 500–545 nm for GFP, 595-645 nm for mCherry and beyond 650 nm for chlorophyll.

For promoter-GUS histochemical analysis, the 5' upstream region of OEX1 (595 nucleotides comprising 161 nucleotides of *OEX1* 5'-UTR and a large part of upstream gene At3g52040) was cloned upstream of the *GUS* gene in the binary vector pMDC162 (Curtis and Grossniklaus, 2003), and the construct was used to generate stable Arabidopsis transformants. Tissues from six independent lines were stained with 5-bromo-4-chloro-3-indolyl-ß-D-glucuronic acid (X-Gluc; Biosynth) as described (Zaegel et al., 2006) and observed by stereomicroscopy and transmission microscopy.

*Transmission Electron Microscopy*

Leaf tissue samples were fixed overnight in 3% glutaraldehyde, treated 2 h with 10% (w/v) picric acid, 2 h with 2% uranyl acetate and stained with 0.1% (v/v) osmium tetroxide in 150 mM phosphate buffer, pH 7.2. After dehydration through an ethanol series, samples were infiltrated with EPON812 medium-grade resin (Polysciences). Ultrathin 70 µm sections were collected on grids coated with formvar (Electron Microscopy Sciences) and visualized with a Hitachi H-600 electron microscope operating at 75 kV.

*Yeast two-hybrid assays*

The cDNA sequences of Arabidopsis OEX1a, OEX1b, POL1A and POL1B were cloned into pGADgwT7-AD and pGBKgwT7-BD vectors. The yeast strain AH109 was co-transformed with different combinations of AD and BD vectors and transformants were selected on synthetic defined (SD)/-Leu/-Trp (SD-LW) medium for 4 days at 30 °C. Colonies were grown overnight in SD-LW liquid media, and after adjustment to $OD_{600nm}$= 0.5, serial dilutions were plated on SD-LW and on SD/-Leu/-Trp/-His (SD-LWH) plates. The interaction between the F-box protein EBF2 and transcription factor EIN3 was used as positive control].

*qPCR / RT-qPCR Analysis*

For RT-qPCR experiments, 5 µg of RNA were depleted from contaminating DNA by treatment with RQ1 RNase-free DNase (Promega) and reverse-transcribed with Superscript IV Reverse Transcriptase (Thermo Fisher Scientific), according to the manufacturer's protocol using random hexamers. The qPCR assays were performed with the LightCycler480 (Roche) on a reaction mix of 6 µL containing 1X LightCycler 480 SYBRGreen I MasterMix (Roche) and 0.5 µM of each primer. Three technical replicates were performed for each experiment and Cp values were determined from the second derivative maximum. Quantification of mtDNA and cpDNA copy numbers used a set of primer pairs located along the organellar genomes, as described previously (Wallet et al., 2015; Chevigny et al., 2022). Results were normalized against *UBQ10* (At4G05320) and *ACT1* (At2G37620) nuclear genes. The accumulation of ectopic recombination in mtDNA was quantified using primers flanking each repeat, as described (Miller-Messmer et al., 2012), and results normalized against the *COX2* (AtMG00160) and 18S rRNA (AtMG01390) mitochondrial genes. RT-qPCR experiments were normalized using the *GAPDH* (At1G13440) and *ACT2* (At3G18780) transcripts as standards.

*Sequencing*

Total leaf DNA of WT and *oex1-1* plants was quantified with a QuBit Fluorometer (Life Technologies) and libraries prepared with the Nextera Flex Library kit according to the manufacturer instructions (Illumina), using 100 ng DNA. Final libraries were quantified, quality checked on a Bioanalyzer 2100 (Agilent) and sequenced on an Illumina Miseq system (2 × 150 paired-end reads). Public RNAseq data was analyzed using the JBrowse platform https://jbrowse.arabidopsis.org/.

*Expression and purification of recombinant proteins*

The coding sequence of the At*OEX1* gene minus the first 85 codons was amplified from cDNA and both alternative cDNA sequences, *OEX1a and OEX1b,* were cloned in the pET28a expression vector (Novagen), to express recombinant proteins containing a N-terminal hexa-histidine tag. The recombinant proteins were soluble and were purified from the *E. coli* BL21(DE3) extracts by nickel ion affinity chromatography, followed by gel filtration (Figure S5). Briefly, protein expression was induced with 0.5 mM IPTG and after 2 h at 37 °C the bacterial cells were sedimented, resuspended in 150 mM Tris-HCl pH 8.0, 5% glycerol, 150 mM NaCl (Buffer A) supplemented with 1 mM PMSF, and lysed with a LM20 Digital Microfluidizer® Processor (Microfluidics™) under 1200 PSI. The crude lysate was clarified by centrifugation at 17 000 g for 20 min, adjusted to 30 mM imidazole. Proteins were purified by affinity on cOmplete™ His-Tag Purification Resin (Roche®). Columns were washed with Buffer A containing 30 mM imidazole, followed by washing with 50 mM imidazole. The recombinant proteins were eluted with a 90-310 mM imidazole step gradient, followed by gel filtration on a Superdex 200 10/300 GL column equilibrated with Buffer A (Figure S5C). Protein aliquots were flash frozen in liquid nitrogen and stored at -80 °C. A catalytic mutant protein was equally expressed and purified, which contained a substitution of an essential aspartic acid with alanine in the active site (protein D250A, Figure S5B). Following purification, the recombinant proteins were free of detectable contaminants and purified as a single monodispersed peak by gel filtration, of a size consistent with a monomeric protein (Figure S5C). The optimal activity conditions of OEX1 on 5'-labelled dsDNA were tested (Figure S6). The protein requires $Mg^{2+}$ ions, that could be poorly substituted by $Mn^{2+}$, but not by $Zn^{2+}$ or $Ca^{2+}$. It showed optimal activity at about 37 °C but was inhibited at higher temperature. Finally, it worked in a broad pH range, between pH 7.4 and pH 9.0. Based

on these results we established the assay conditions for further experiments as 2 mM $Mg^{2+}$, pH 8.0 at 37 °C.

*Nuclease assays*

Structure-specific nuclease assays were performed using oligonucleotides synthesized by Integrated DNA Technologies (IDT). DNA oligonucleotides were either radiolabelled in 5' using T4 polynucleotide kinase (Fermentas) and [γ-$^{32}$P]ATP, or in 3', by template-directed addition of one to three cytosines by the Klenow fragment of *Escherichia coli* DNA polymerase I and [α-$^{32}$P]dCTP. RNA oligonucleotides were labelled in 3' using poly(U) polymerase (NEB) and [α-$^{32}$P]UTP. When necessary, labelled products where annealed with cold oligonucleotides by incubation in 10 mM Tris-HCl pH 7.5, 1 mM EDTA, 100 mM NaCl (Buffer B) during 5 min at 95 °C, followed by slow cooling to room temperature. Oligonucleotide sequences used to form the different structures are listed in supplementary Table 1*. Substrates were purified by electrophoresis on non-denaturing 15% polyacrylamide gels and eluted in Buffer B, overnight at 4 °C under agitation. Standard nuclease reactions were performed with 50 fmol of labelled structure in reaction buffer (50 mM Tris-HCl pH 8.0, 1 mM DTT, 0.1 mg/mL BSA and 4% glycerol) supplemented with 2 mM bivalent ion and 1.5 µM enzyme. The reaction was incubated at 37 °C and stopped by addition of an equal volume of stop-mix (95% formamide, 20 mM EDTA, 0,05% bromophenol blue, 0,05% xylene cyanol). Reaction products were analysed by denaturing PAGE (17% 19:1 acrylamide/bisacrylamide, 7 M urea, 1 X TBE) and revealed using a Amersham Typhoon biomolecular imager (GE Healthcare Life Sciences).

*DNA:RNA Immunoprecipitation*

Total DNA was extracted with Plant II Midi kit (Macherey-Nagel®) and fragmented with EcoRI, BamHI, HindIII and XbaI to cut the mtDNA into pieces between 200 to 5,000 bp. About 5 µg were either treated or not with Ribonuclease H (RNase H 5 U/µL) (Thermo Fisher Scientific®), and incubated with 2 µL S9.6 antibody for 15 h at 4 °C under rotation (10 rpm), followed by incubation with 25 µL of Protein G Sepharose beads (Sigma-Aldrich®) for 2 h at 4 °C (Yang et al., 2017). Immunoprecipitated fragments were detached from antibodies by proteinase K treatment (Thermo Fisher Scientific®) and purified by phenol/chloroform extraction followed by ethanol precipitation. The isolated DNA was analysed by qPCR using primers specific to highly transcribed regions and results were normalized against non-transcribed regions of the mtDNA.

## Figure Legends

**Figure 1. OEX1 and OEX2 are plant organellar orthologs of bacterial exonucleases.**

**A)** Two Arabidopsis genes encode putative organellar proteins similar to the 5'-3' exonuclease domain of bacterial DNA polymerase I, a domain absent in plant organellar DNA polymerases POL1A and POL1B (OTS: organellar targeting sequence). **B)** Phenogram showing that OEX1 orthologs are present across the green lineage, including land plants and green algae, while OEX2 orthologs are restricted to land plants. **C)** Similarity between OEX1 and OEX2 as compared to the respective similarities to bacterial proteins, suggesting that the two plant genes did not diverge from a common ancestor and may have distinct origins.

**Figure 2. OEX1 and OEX2 are respectively targeted to mitochondria and chloroplasts.**

**A)** Mitochondrial targeting of OEX1:GFP constitutively expressed in Arabidopsis. **B)** OEX1 is localized in mitochondrial nucleoids, as revealed by the colocalization of OEX1-mCherry with DAPI-stained nucleoids (mT-GFP: mitochondria marker protein). **C)** Chloroplast targeting of OEX2:GFP in *Nicotiana benthamiana* leaf guard cells following biolistic transfection. The protein is mainly localized in nucleoid-like speckles

**Figure 3. Loss of OEX1 has severe effects on plant growth and development.**

**A)** Representative images of 4-week-old wild-type (WT) and *oex1-1* mutant plants, showing growth retardation, distorted leaves with chlorotic sectors, flower deformities, and reduced fertility. The severity of the phenotypes can vary significantly among individual plants within the first homozygote mutant generation (T2 plants), as highlighted in the box plots showing the high variation in rosette size in *oex1-1*. Flowers show few pollen grains bound to the papillae of the stigma. Siliques are small and mainly contain aborted embryos. **B)** Developmental phenotypes are fully complemented in *oex1-1* plants expressing OEX1 under its native promoter (*oex1-1 OEX1* plant) and in mutants expressing OEX1 targeted only to mitochondria, by fusion to the targeting sequence of AOX1 (*oex1-1 AOX1:ΔN-OEX1* plant). However, chloroplast-specific OEX1 expression by fusion to the targeting sequence of RBCS does not rescue the phenotype (*oex1 RBCS:ΔN-OEX1* plant).

**Figure 4. Structural alterations of the mtDNA in *oex1* mutants**

**A**) Sequence coverage of the mtDNA from *oex1-1* vs. WT, showing regions with changed stoichiometry flanked by direct repeats. The coordinates are those of the Col-0 mtDNA sequence. The position of the mtDNA large repeats LR1 and LR2 are shown. Regions with changed stoichiometry are shadowed and annotated according to the repeat name (repeats A, F, L, X and EE, of 556 bp, 350bp, 249 bp, 204 bp and 127 bp, respectively). **B**) As in A, but for the plastidial genome, showing that there are no effects from OEX1 loss on cpDNA stability. LSC: large single copy region, SSC: small single copy region, IR: inverted repeat (only one copy is included). **C)** qPCR confirmation of the accumulation in *oex1* plants of recombination products across repeats L, F and EE. Parental sequences 1/1, 2/2 and corresponding crossover products 1/2 and 2/1 were quantified. Error bars correspond to *SD* values from three biological replicates. **D** and **E)** Analysis of the relative copy numbers of the different mtDNA regions in individual severely affected T4 generation *oex1* plants. Sequences spaced about 5 kb apart on the mtDNA were quantified by qPCR. Regions with changed stoichiometry flanked by direct repeats are shadowed as in A. In plants #5 and #6, regions that are highly increased in copy number correspond to subgenomes resulting from recombination involving a microhomology of 14 bp and an imperfect repeat of 85bp, respectively. The parental and recombined sequences are shown.

**Figure 5. Alternative splicing results in the expression of two OEX1 isoforms.**

**A**) Representation of the two proteins resulting from alternative splicing of exon 6. **B**) Complementation of the *oex1-1* mutant with OEX1a or OEX1b coding sequences (*cOEX1a* and *cOEX1b* plants, respectively), under control of the endogenous *OEX1* promoter. Both could complement the deleterious developmental phenotypes of *oex1*. **C)** Analysis of the mtDNA of *cOEX1a* and *cOEX1b* plants by qPCR, as described in Figure 4E, showing full complementation by *cOEX1a,* but retention of a slightly remodeled mtDNA in *oex1 cOEX1b*, with differential copy number of the subgenomes defined by large repeat LR2. Results represent the average and SD of three biological replicates. **D** and **E) C)** Phenotypes of plants grown under genotoxic stress induced by 0.5 µM ciprofloxacin (CIP). Plants complemented with *cOEX1b* were more sensitive, as assessed by the percentage of plants still able to develop normal first true leaves, at a CIP concentration that had a mild effect on WT. Asterisks indicate statistical significance according to the $\chi^2$ test (*: $p<0.05$, **: $p<0.01$).

**Figure 6. Accumulation of mitochondrial transcripts in *oex1***

**A)** RT-qPCR analysis of steady-state levels of mitochondrial proteins transcripts and rRNAs in *oex1-1* and complemented lines *cOEX1a* and *cOEX1b.* Results were normalized to a set of nuclear housekeeping genes. The exons borders were quantified for genes containing introns, to asses for possible splicing defects. Results are in log2 scale and are the average and *SD* from three biological replicates.

**Figure 7. OEX1 exonuclease and Flap endonuclease activities**

**A**) Analysis of OEX1a exonuclease activity and of catalytic mutant D250A on a dsDNA substrate, represented above, labeled either on 5' or 3' (red stars). Substrates and degradation products were analyzed on denaturing polyacrylamide gels. Size of size markers are indicated. **B**) As in A, for isoform OEX1b. **C**) As in A, for a substrate miming flap structures. **D**) As in C, for OEX1b. **E**) As in A, for a substrate with a 10 bp gap, in the absence and presence of a DNA polymerase (Klenow). **F**) Summary of OEX1 activities on the various DNA substrates.

**Figure 8. OEX1 exonuclease activity on RNA:DNA hybrids and *in vivo* contribution to the clearance of R-loops in the mtDNA.**

**A**) Analysis of OEX1a exonuclease activity and of catalytic mutant D250A on a RNA:DNA hybrid labeled on the 3' of the RNA (red star). Substrates and degradation products were analyzed on denaturing polyacrylamide gels. **B**) As in A, on a substrate representing an Okazaki RNA fragment. **C**) As in A, showing that OEX1 has no activity on a single-strand RNA substrate. **D**) As in A, showing OEX1 high activity on the RNA moiety of an R-loop. **E**) DRIP-qPCR analysis of R-loops in *oex1* plants as compared to WT. Results are the average of three biological replicates and showed an increase of R-loops in *oex1*, in the mtDNA region corresponding to the highly transcribed rRNAs 18S and 26S. Results were normalized against non-transcribed regions of the mtDNA. Asterisks represent statistical significance by Student's-t-test (*: $p<0.05$, **: $p<0.01$, ***: $p<0.001$).

**Supplemental Figure S1. Widespread expression of *oex1*, according to promoter-GUS fusion.**

Promoter::GUS fusion reporter plants of OEX1 (pOEX1::GUS) were stained for GUS gene expression at different developmental stages. A) three-day-old seedling, B) root, C) 10-day-old seedling, D) inflorescences, E) flower, and F) developing silique.

**Supplemental Figure S2. Transmission electron microscope images showing smaller, more electron-dense mitochondria in *oex1* compared to WT.**

**A)** Cells from leaves of same size showed morphologically normal chloroplasts (cp) in *oex1,* while mitochondria (mt) were smaller and more electron-dense as compared to mitochondria in WT cells. **B)** Distribution of the organelle sections (in $\mu m^2$) extracted from TEM images. Asterisks indicate statistical significance ($p<0.001$) according to Student's t-test.

**Supplemental Figure S3. Alternative splicing results in the expression of two OEX1 isoforms.**

**A)** Alternatively spliced exon 6 of OEX1 encodes an 18 aa insertion that is absent from the sequences of OEX2 and the 5'-3' exonuclease of bacterial DNA Pol I. An insertion of comparable size is found in the OEX1 sequences of land plants. **B)** Comparison between the known structure of the 5'-3' exonuclease domain of DNA Pol I from *Thermophilus aquaticus* and the modeled structure of OEX1. The 18 aa extension is in green and indicated by an arrow, with the H3TH domain in purple. The two essential aspartic acids in the PIN active domain are highlighted in red. **C)** Relative quantification of alternative spliced *OEX1a* and *OEX1b* transcripts in different tissues.

**Supplemental Figure S4. Interaction between OEX1 and organellar DNA polymerases.**

**A)** Interaction between OEX1a and OEX1b with the organellar DNA polymerase POL1A, tested by yeast two-hybrid assay. BD and AD are the binding and activating domains of GAL4, respectively. Serial dilutions of cell suspensions were spotted on SD-LWH plates. The interaction between EBF2 and EIN3 served as positive control. Both BD-OEX1a and BD-OEX1b could interact with AD-POL1A. **B)** As in A, but with organellar DNA polymerases POL1B. BD-OEX1b could interact with AD-POL1B, but not BD-OEX1a. Controls with empty vectors confirmed no significant autoactivation by either BD-OEX1a/1b or by AD-POL1A/1B. **C)** Data from complexome database (complexomemap.de/at_mito_leaves) indicate that OEX1 migrates as a band smaller than 100 kDa on Blue-Native gels. For comparison, the results obtained for the TWINKLE helicase, subunit B2 of gyrase and ssDNA-binding proteins SSB1 and SSB2 are also shown.

**Supplemental Figure S5. Expression and purification of recombinant OEX1 proteins.**

**A)** SDS-PAGE analysis of the expression and solubility of OEX1a and OEX1b, indicated by arrowheads. **B)** expression of a catalytic mutant where an essential aspartic acid of the catalytic

PIN domain was replaced by an alanine. **C)** Purification of recombinant protein OEX1a by gel filtration, following metal-affinity purification on a Ni-column. The apparent molecular weight is consistent with a monomeric protein.

**Supplemental Figure S6. Optimization of OEX1 activity conditions.**

**A)** Activity of OEX1a on a 5'-labeled dsDNA substrate, tested at varying concentrations of divalent ions. Substrates and degradation products were analyzed on denaturing polyacrylamide gels, after 3 minutes of incubation. Maximum activity was observed with $Mg^{2+}$. **B)** As in A, according to incubation temperature. The optimal temperature observed was 37 °C. **C)** As in A, but according to pH.

**Supplemental Figure S7. Sequence-dependent processivity of OEX1 and activity on DNA gaps.**

**A)** Activity of OEX1a on dsDNA substrates differing by 1, 2 or 3 consecutive Gs, showing pausing pausing at Gs. **B)** Minimal activity of OEX1 alone on DNA gap of 1, 3 or 10 nucleotides, but no detectable activity of the catalytic mutant (D250A).

**Supplemental Figure S8. Activity of OEX1 on ssDNA**

Little activity of OEX1a and OEX1b on ssDNA substrates.

## References


**Aguilera, A., and Garcia-Muse, T.** (2012). R loops: from transcription byproducts to threats to genome stability. Mol. Cell **46,** 115-124.

**Anstey-Gilbert, C.S., Hemsworth, G.R., Flemming, C.S., Hodskinson, M.R., Zhang, J., Sedelnikova, S.E., Stillman, T.J., Sayers, J.R., and Artymiuk, P.J.** (2013). The structure of Escherichia coli ExoIX--implications for DNA binding and catalysis in flap endonucleases. Nucleic Acids Res. **41,** 8357-8367.

**Ayala-Garcia, V.M., Baruch-Torres, N., Garcia-Medel, P.L., and Brieba, L.G.** (2018). Plant organellar DNA polymerases paralogs exhibit dissimilar nucleotide incorporation fidelity. FEBS J **285,** 4005-4018.

**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.

**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.

**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.

**Boesch, P., Weber-Lotfi, F., Ibrahim, N., Tarasenko, V., Cosset, A., Paulus, F., Lightowlers, R.N., and Dietrich, A.** (2011). DNA repair in organelles: Pathways, organization, regulation, relevance in disease and aging. Biochim. Biophys. Acta **1813,** 186-200.

**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.

**Cheng, L.L., Wang, W.J., Yao, Y., and Sun, Q.W.** (2021). Mitochondrial RNase H1 activity regulates R-loop homeostasis to maintain genome integrity and enable early embryogenesis in Arabidopsis. PLoS biology **19,** e3001357.

**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,** 328.

**Chevigny, N., Weber-Lotfi, F., Le Blevenec, A., Nadiras, C., Fertet, A., Bichara, M., Erhardt, M., Dietrich, A., Raynaud, C., and Gualberto, J.M.** (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.

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

**Crossley, M.P., Bocek, M., and Cimprich, K.A.** (2019). R-Loops as Cellular Regulators and Genomic Threats. Mol. Cell **73,** 398-411.

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

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

**Falkenberg, M., Larsson, N.G., and Gustafsson, C.M.** (2024). Replication and Transcription of Human Mitochondrial DNA. Annu. Rev. Biochem. **93,** 47-77.

**Ferrando, B., Furlanetto, A., Gredilla, R., Havelund, J.F., Hebelstrup, K.H., Moller, I.M., and Stevnsner, T.** (2018). DNA repair in plant mitochondria - a complete base excision repair pathway in potato tuber mitochondria. Physiol. Plant. **166,** 494-512.

**Fertet, A., Graindorge, S., Koechler, S., de Boer, G.J., Guilloteau-Fonteny, E., and Gualberto, J.M.** (2021). Sequence of the Mitochondrial Genome of Lactuca virosa Suggests an Unexpected Role in Lactuca sativa's Evolution. Front. Plant Sci. **12,** 697136.

**Fu, Y., Sacco, O., DeBitetto, E., Kanshin, E., Ueberheide, B., and Sfeir, A.** (2023). Mitochondrial DNA breaks activate an integrated stress response to reestablish homeostasis. Mol. Cell **83,** 3740-3753 e3749.

**Fuchs, P., Rugen, N., Carrie, C., Elsasser, M., Finkemeier, I., Giese, J., Hildebrandt, T.M., Kuhn, K., Maurino, V.G., Ruberti, C., Schallenberg-Rudinger, M., Steinbeck, J., Braun, H.P., Eubel, H., Meyer, E.H., Muller-Schussele, S.J., and Schwarzlander, M.** (2019). Single organelle function and organization as estimated from Arabidopsis mitochondrial proteomics. Plant J. **101,** 420-441.

**Fukushima, S., Itaya, M., Kato, H., Ogasawara, N., and Yoshikawa, H.** (2007). Reassessment of the in vivo functions of DNA polymerase I and RNase H in bacterial cell growth. J. Bacteriol. **189,** 8575-8583.

**Gan, W., Guan, Z., Liu, J., Gui, T., Shen, K., Manley, J.L., and Li, X.** (2011). R-loop-mediated genomic instability is caused by impairment of replication fork progression. Genes Dev. **25,** 2041-2056.

**Gandini, C.L., Garcia, L.E., Abbona, C.C., Ceriotti, L.F., Kushnir, S., Geelen, D., and Sanchez-Puerta, M.V.** (2023). Break-induced replication is the primary recombination pathway in plant somatic hybrid mitochondria: a model for mitochondrial horizontal gene transfer. J. Exp. Bot. **74,** 3503-3517.

**Garcia, L.E., Edera, A.A., Palmer, J.D., Sato, H., and Sanchez-Puerta, M.V.** (2021). Horizontal gene transfers dominate the functional mitochondrial gene space of a holoparasitic plant. New Phytol. **229,** 1701-1714.

**Garcia-Medel, P.L., Baruch-Torres, N., Peralta-Castro, A., Trasvina-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.

**Ginno, P.A., Lott, P.L., Christensen, H.C., Korf, I., and Chedin, F.** (2012). R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. Mol. Cell **45,** 814-825.

**Grefen, C., Donald, N., Hashimoto, K., Kudla, J., Schumacher, K., and Blatt, M.R.** (2010). A ubiquitin-10 promoter-based vector set for fluorescent protein tagging facilitates temporal stability and native protein distribution in transient and stable expression studies. Plant J. **64,** 355-365.

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

**Hanson, M.R., and Bentolila, S.** (2004). Interactions of mitochondrial and nuclear genes that affect male gametophyte development. Plant Cell **16 Suppl,** S154-169.

**Jaipargas, E.A., Barton, K.A., Mathur, N., and Mathur, J.** (2015). Mitochondrial pleomorphy in plant cells is driven by contiguous ER dynamics. Front. Plant Sci. **6,** 783.

**Kaguni, L.S.** (2004). DNA polymerase gamma, the mitochondrial replicase. Annu. Rev. Biochem. **73,** 293-320.

**Kim, Y., Eom, S.H., Wang, J., Lee, D.S., Suh, S.W., and Steitz, T.A.** (1995). Crystal structure of Thermus aquaticus DNA polymerase. Nature **376,** 612-616.

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

**Klepikova, A.V., Kasianov, A.S., Gerasimov, E.S., Logacheva, M.D., and Penin, A.A.** (2016). A high resolution map of the Arabidopsis thaliana developmental transcriptome based on RNA-seq profiling. Plant J. **88,** 1058-1070.

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

**Kockler, Z.W., Osia, B., Lee, R., Musmaker, K., and Malkova, A.** (2021). Repair of DNA Breaks by Break-Induced Replication. Annu. Rev. Biochem. **90,** 165-191.

**Kowalczykowski, S.C.** (2015). An Overview of the Molecular Mechanisms of Recombinational DNA Repair. Cold Spring Harb. Perspect. Biol. **7**, a016410.

**Kubo, T., and Newton, K.J.** (2008). Angiosperm mitochondrial genomes and mutations. Mitochondrion **8,** 5-14.

**Lehman, I.R., and Uyemura, D.G.** (1976). DNA polymerase I: essential replication enzyme. Science **193,** 963-969.

**Lewis, S.C., Uchiyama, L.F., and Nunnari, J.** (2016). ER-mitochondria contacts couple mtDNA synthesis with mitochondrial division in human cells. Science **353,** aaf5549.

**Lovett, S.T.** (2011). The DNA Exonucleases of Escherichia coli. EcoSal Plus **4**.

**Ma, L., Sun, H., Abeywardana, T., Zheng, L., and Shen, B.** (2022). Structure-specific nucleases: role in Okazaki fragment maturation. Trends Genet. **38,** 793-796.

**Macao, B., Uhler, J.P., Siibak, T., Zhu, X., Shi, Y., Sheng, W., Olsson, M., Stewart, J.B., Gustafsson, C.M., and Falkenberg, M.** (2015). The exonuclease activity of DNA polymerase gamma is required for ligation during mitochondrial DNA replication. Nat Commun **6,** 7303.

**Majeran, W., Friso, G., Asakura, Y., Qu, X., Huang, M., Ponnala, L., Watkins, K.P., Barkan, A., and van Wijk, K.J.** (2012). Nucleoid-enriched proteomes in developing plastids and chloroplasts from maize leaves: a new conceptual framework for nucleoid functions. Plant Physiol. **158,** 156-189.

**Manchekar, M., Scissum-Gunn, K., Song, D., Khazi, F., McLean, S.L., and Nielsen, B.L.** (2006). DNA recombination activity in soybean mitochondria. J. Mol. Biol. **356,** 288-299.

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

**Matelska, D., Steczkiewicz, K., and Ginalski, K.** (2017). Comprehensive classification of the PIN domain-like superfamily. Nucleic Acids Res. **45,** 6995-7020.

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

**Misic, J., Milenkovic, D., Al-Behadili, A., Xie, X., Jiang, M., Jiang, S., Filograna, R., Koolmeister, C., Siira, S.J., Jenninger, L., Filipovska, A., Clausen, A.R., Caporali, L., Valentino, M.L., La Morgia, C., Carelli, V., Nicholls, T.J., Wredenberg, A., Falkenberg, M., and Larsson, N.G.** (2022). Mammalian RNase H1 directs RNA primer formation for mtDNA replication initiation and is also necessary for mtDNA replication completion. Nucleic Acids Res. **50,** 8749-8766.

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

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

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

**Oldenburg, D.J., and Bendich, A.J.** (1996). Size and Structure of Replicating Mitochondrial DNA in Cultured Tobacco Cells. Plant Cell **8,** 447-461.

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

**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.

**Penafiel-Ayala, A., Peralta-Castro, A., Mora-Garduno, J., Garcia-Medel, P., Zambrano-Pereira, A.G., Diaz-Quezada, C., Abraham-Juarez, M.J., Benitez-Cardoza, C.G., Sloan, D.B., and Brieba, L.G.** (2023). Plant organellar MSH1 is a displacement loop specific endonuclease. Plant Cell Physiol. **65,** 560-575

**Peralta-Castro, A., García-Medel, P.L., Baruch-Torres, N., Trasviña-Arenas, C.H., Juarez-Quintero, V., Morales-Vazquez, C.M., and Brieba, L.G.** (2020). Plant Organellar DNA Polymerases Evolved Multifunctionality through the Acquisition of Novel Amino Acid Insertions. Genes **11**.

**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.

**Ranalli, T.A., Tom, S., and Bambara, R.A.** (2002). AP endonuclease 1 coordinates flap endonuclease 1 and DNA ligase I activity in long patch base excision repair. J. Biol. Chem. **277,** 41715-41724.

**Randall, J.R., Nye, T.M., Wozniak, K.J., and Simmons, L.A.** (2019). RNase HIII Is Important for Okazaki Fragment Processing in Bacillus subtilis. J. Bacteriol. **201**, e00686-18.

**Rice, D.W., Alverson, A.J., Richardson, A.O., Young, G.J., Sanchez-Puerta, M.V., Munzinger, J., Barry, K., Boore, J.L., Zhang, Y., dePamphilis, C.W., Knox, E.B., and Palmer, J.D.** (2013). Horizontal transfer of entire genomes via mitochondrial fusion in the angiosperm Amborella. Science **342,** 1468-1473.

**Roy, D., and Lieber, M.R.** (2009). G clustering is important for the initiation of transcription-induced R-loops in vitro, whereas high G density without clustering is sufficient thereafter. Mol. Cell Biol. **29,** 3124-3133.

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

**Sanz, L.A., and Chedin, F.** (2019). High-resolution, strand-specific R-loop mapping via S9.6-based DNA-RNA immunoprecipitation and high-throughput sequencing. Nat. Protoc. **14,** 1734-1755.

**Schatz-Daas, D., Fertet, A., Lotfi, F., and Gualberto, J.M.** (2022). Assessment of Mitochondrial DNA Copy Number, Stability, and Repair in Arabidopsis. Methods Mol. Biol. **2363,** 301-319.

**Senkler, J., Senkler, M., Eubel, H., Hildebrandt, T., Lengwenus, C., Schertl, P., Schwarzlander, M., Wagner, S., Wittig, I., and Braun, H.P.** (2017). The mitochondrial complexome of Arabidopsis thaliana. Plant J. **89,** 1079-1092.

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

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

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

**Szczesny, B., Tann, A.W., Longley, M.J., Copeland, W.C., and Mitra, S.** (2008). Long patch base excision repair in mammalian mitochondrial genomes. J. Biol. Chem. **283,** 26349-26356.

**Tigano, M., Vargas, D.C., Tremblay-Belzile, S., Fu, Y., and Sfeir, A.** (2021). Nuclear sensing of breaks in mitochondrial DNA enhances immune surveillance. Nature **591,** 477-481.

**Trasvina-Arenas, C.H., Baruch-Torres, N., Cordoba-Andrade, F.J., Ayala-Garcia, V.M., Garcia-Medel, P.L., Diaz-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.

**Uhler, J.P., and Falkenberg, M.** (2015). Primer removal during mammalian mitochondrial DNA replication. DNA Repair (Amst.) **34,** 28-38.

**Uhler, J.P., Thorn, C., Nicholls, T.J., Matic, S., Milenkovic, D., Gustafsson, C.M., and Falkenberg, M.** (2016). MGME1 processes flaps into ligatable nicks in concert with DNA polymerase gamma during mtDNA replication. Nucleic Acids Res. **44,** 5861-5871.

**Wallet, C., Le Ret, M., 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 **27,** 2907–2925.

**Wimberly, H., Shee, C., Thornton, P.C., Sivaramakrishnan, P., Rosenberg, S.M., and Hastings, P.J.** (2013). R-loops and nicks initiate DNA breakage and genome instability in non-growing Escherichia coli. Nat Commun **4,** 2115.

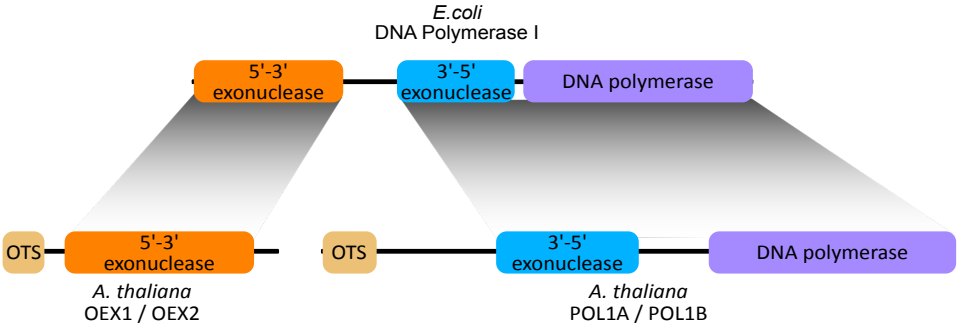
**Wu, Z., Waneka, G., Broz, A.K., King, C.R., and Sloan, D.B.** (2020). MSH1 is required for maintenance of the low mutation rates in plant mitochondrial and plastid genomes. Proc. Natl. Acad. Sci. USA. **117,** 16448-16455.

**Zaegel, V., Guermann, B., Le Ret, M., 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.

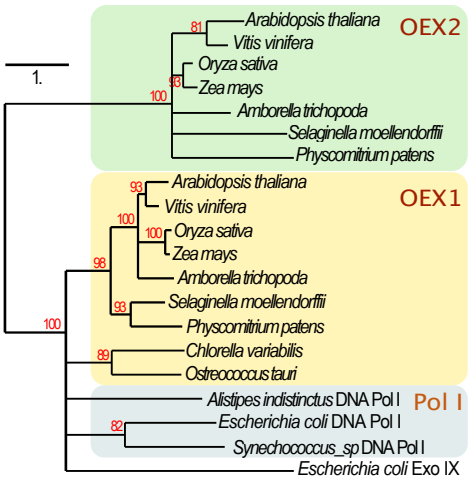
**Zheng, L., and Shen, B.** (2011). Okazaki fragment maturation: nucleases take centre stage. J Mol Cell Biol **3,** 23-30.

**Figure 1.**

A




B




C

| | *A. thaliana* OEX1 | *A. thaliana* OEX2 | *A. indistinctus* Pol I | *E. coli* Pol I | *Synechococcus sp* Pol I | *E. coli* Exo IX |
|---|---|---|---|---|---|---|
| *A. thaliana* OEX1 | 100 | 16 | 26 | 22 | 24 | 18 |
| *A. thaliana* OEX2 | 29 | 100 | 17 | 16 | 13 | 14 |
| *A. indistinctus* Pol I | 39 | 28 | 100 | 28 | 31 | 19 |
| *E. coli* Pol I | 32 | 29 | 44 | 100 | 33 | 24 |
| *Synechococcus sp* Pol I | 33 | 25 | 43 | 46 | 100 | 22 |
| *E. coli* Exo IX | 29 | 26 | 33 | 38 | 36 | 100 |

Identity %

Similarity %

**Figure 2.**

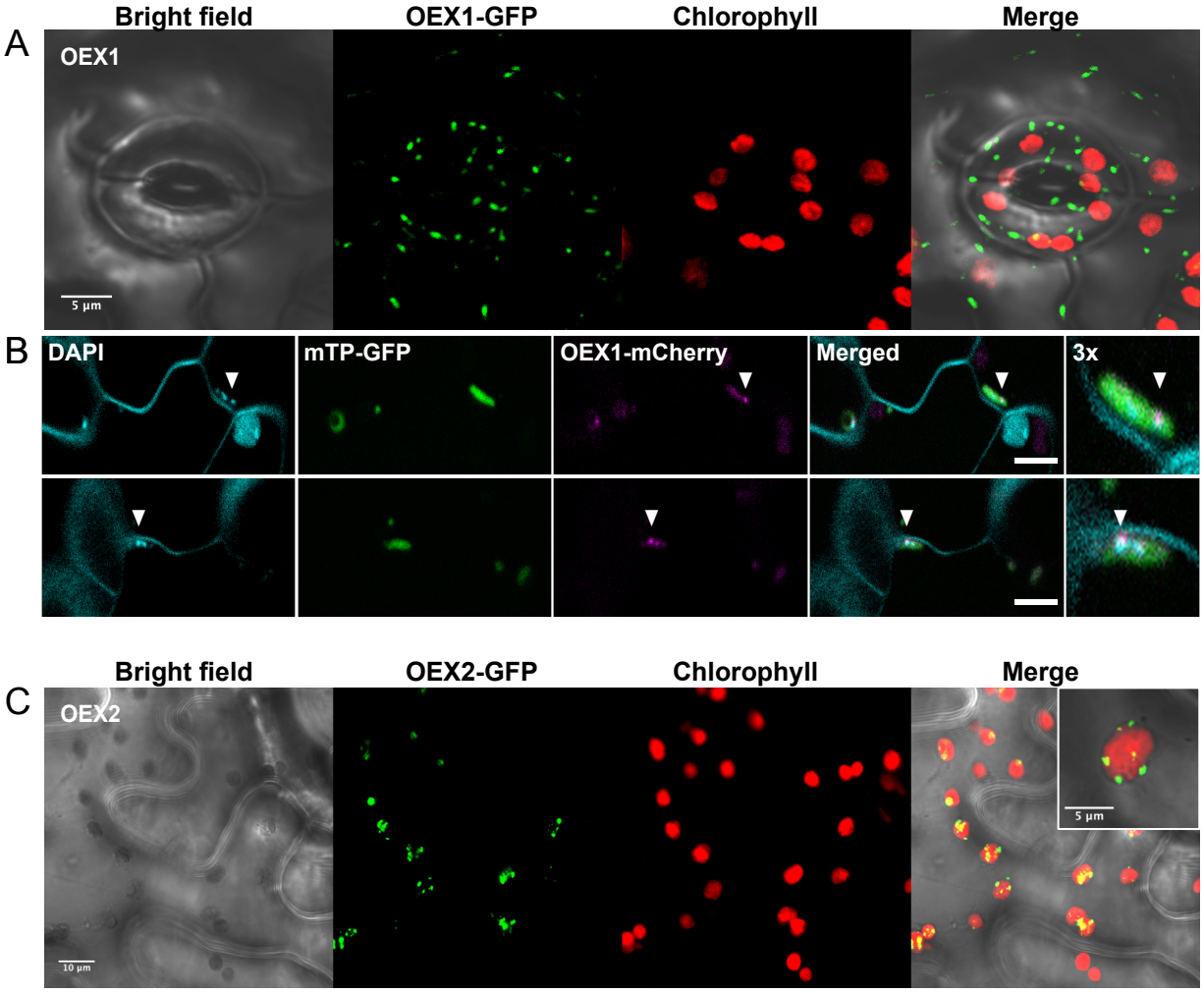

A
Bright field
OEX1-GFP
Chlorophyll
Merge
OEX1
5 µm
B
DAPI
mTP-GFP
OEX1-mCherry
Merged
3x
C
Bright field
OEX2-GFP
Chlorophyll
Merge
OEX2
10 µm
5 µm

**Figure 3.**

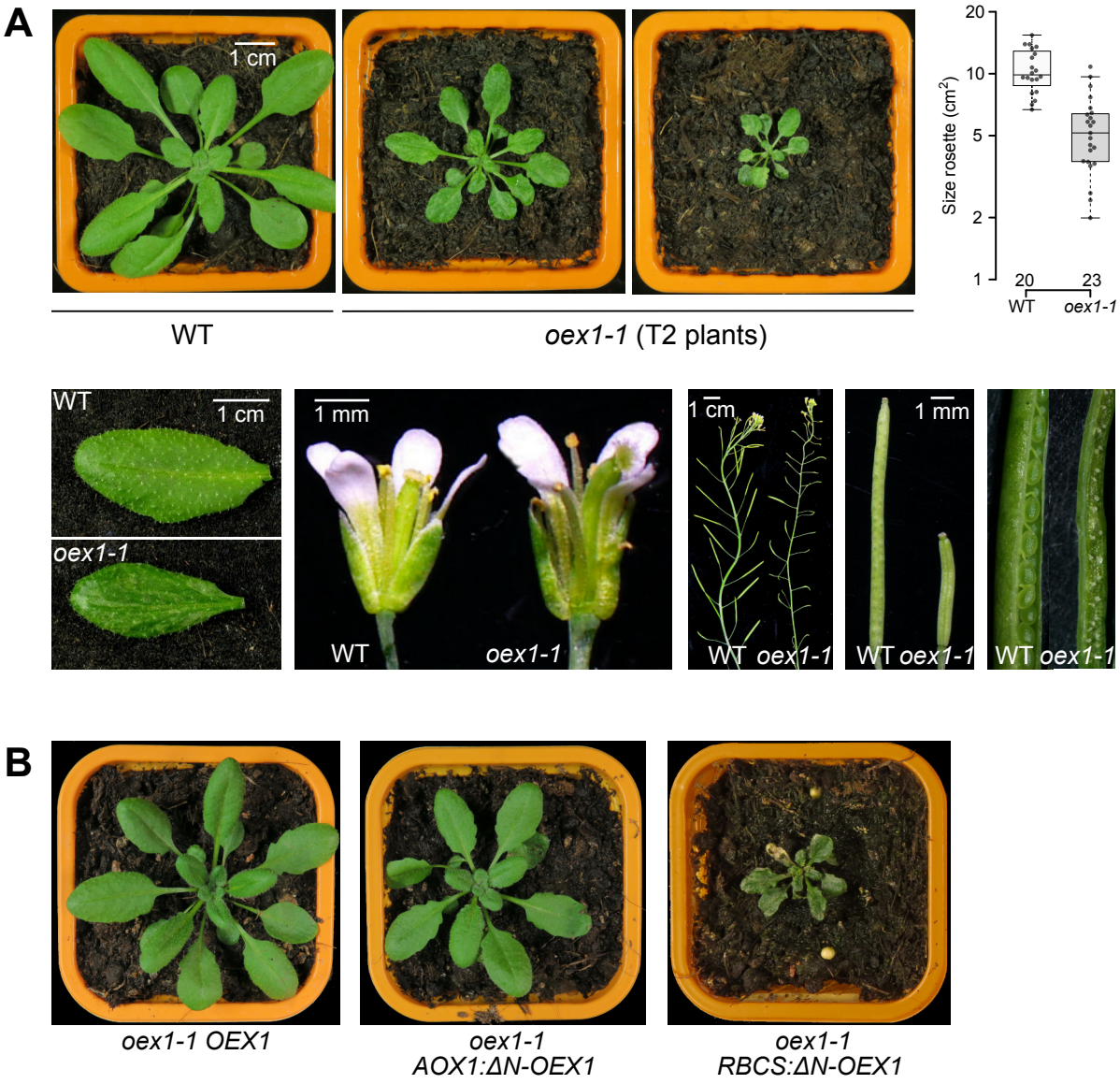

A
1 cm
WT
oex1-1 (T2 plants)
Size rosette (cm²)
20
10
5
2
1
20
23
WT
oex1-1
WT
1 cm
oex1-1
1 mm
WT
oex1-1
1 cm
WT oex1-1
1 mm
WT oex1-1
WT oex1-1
B
oex1-1 OEX1
oex1-1
AOX1:ΔN-OEX1
oex1-1
RBCS:ΔN-OEX1

**Figure 4.**

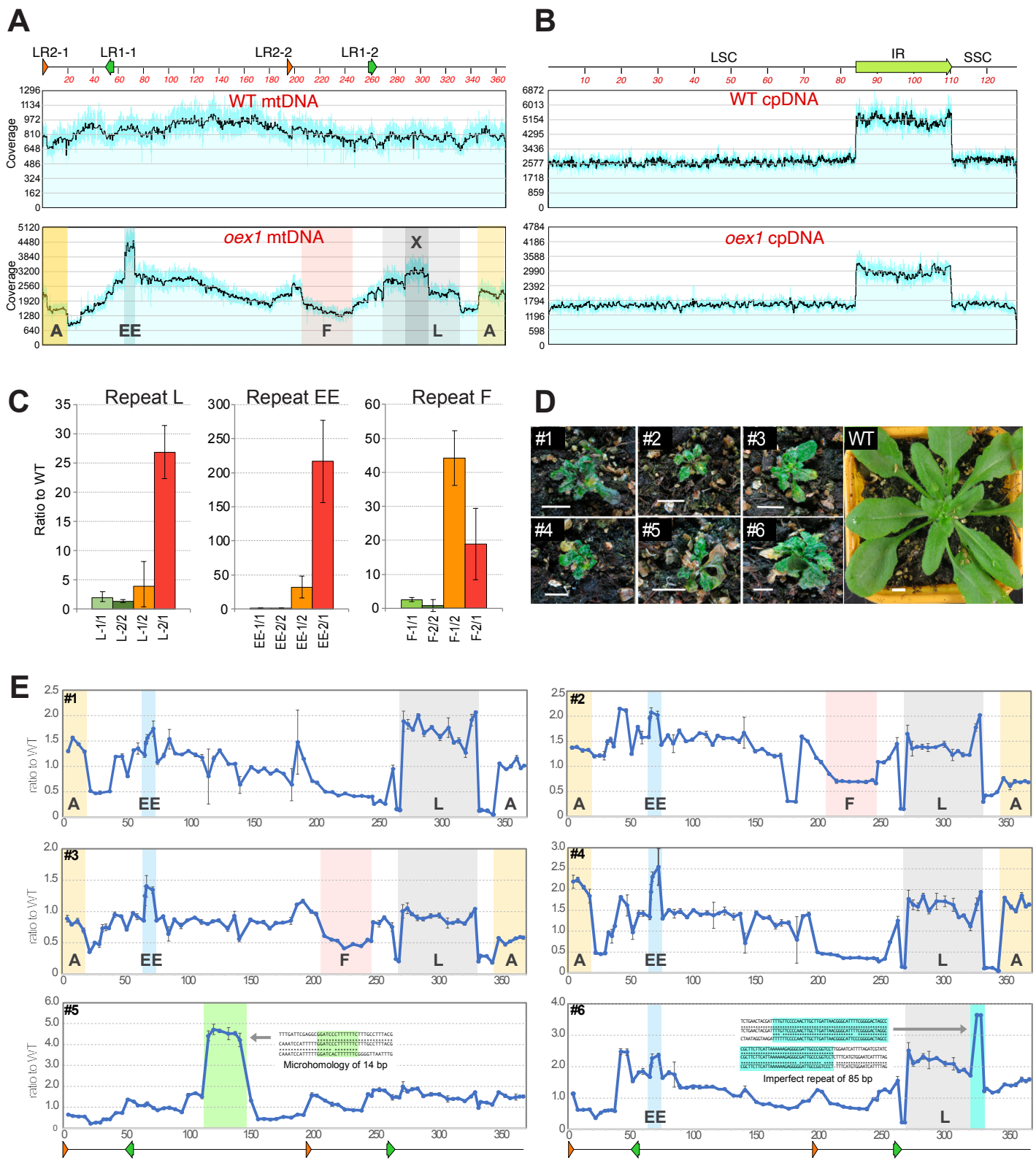

A
LR2-1
LR1-1
LR2-2
LR1-2
WT mtDNA
Coverage
oex1 mtDNA
X
A
EE
F
L
A
B
LSC
IR
SSC
WT cpDNA
oex1 cpDNA
C
Repeat L
Repeat EE
Repeat F
Ratio to WT
L-1/1
L-2/2
L-1/2
L-2/1
EE-1/1
EE-2/2
EE-1/2
EE-2/1
F-1/1
F-2/2
F-1/2
F-2/1
D
#1
#2
#3
#4
#5
#6
WT
E
ratio to WT
Microhomology of 14 bp
Imperfect repeat of 85 bp

**Figure 5.**

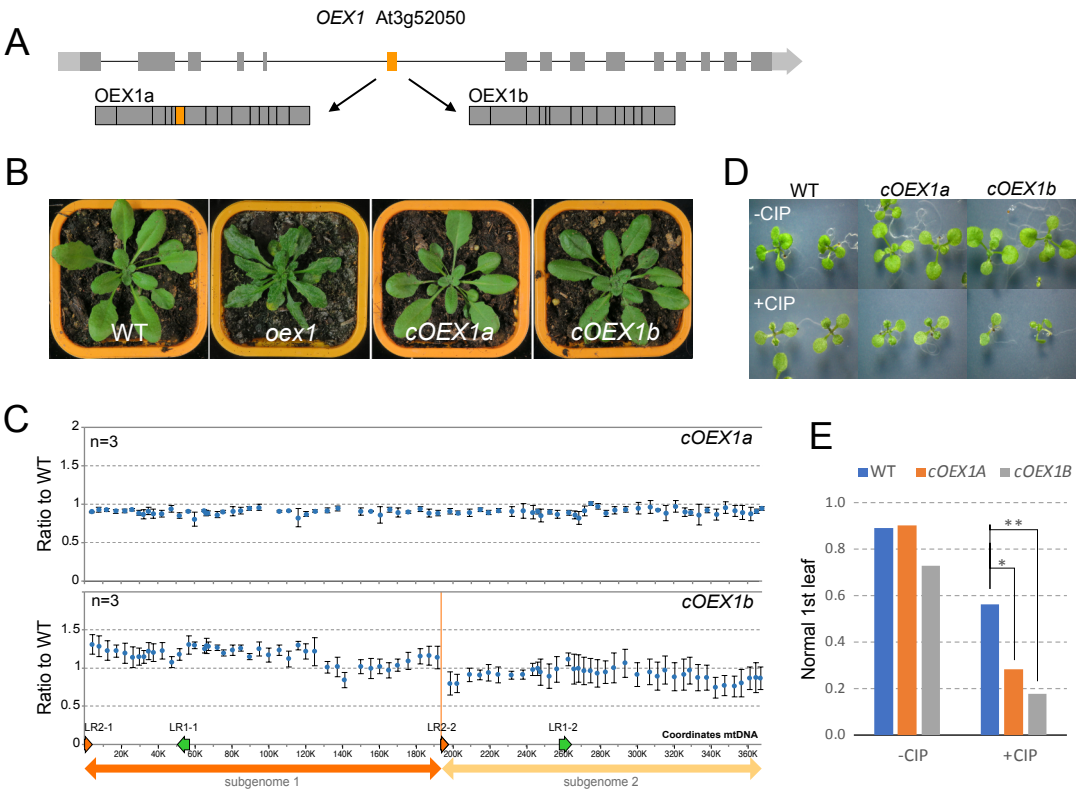

A
OEX1 At3g52050
OEX1a
OEX1b
B
WT
oex1
cOEX1a
cOEX1b
D
WT
cOEX1a
cOEX1b
-CIP
+CIP
C
n=3
cOEX1a
Ratio to WT
n=3
cOEX1b
Ratio to WT
LR2-1
LR1-1
LR2-2
LR1-2
Coordinates mtDNA
subgenome 1
subgenome 2
E
WT
cOEX1A
cOEX1B
Normal 1st leaf
-CIP
+CIP

**Figure 6.**

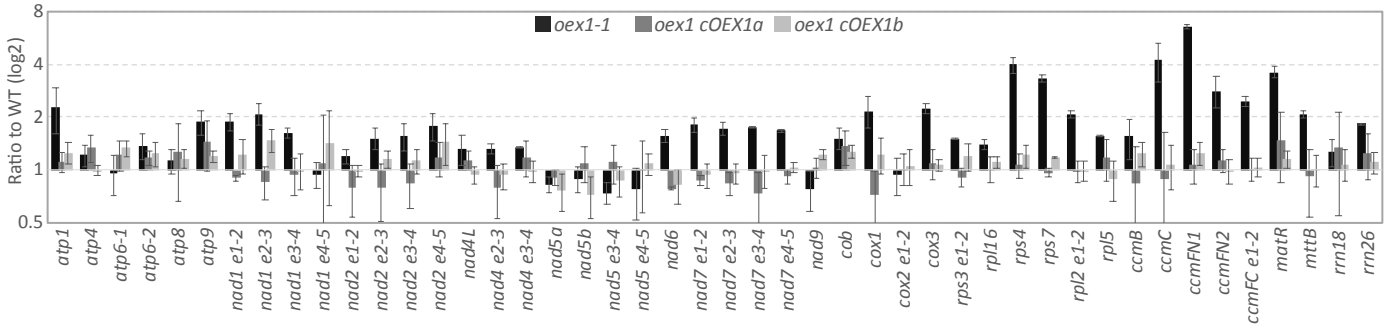

oex1-1
oex1 cOEX1a
oex1 cOEX1b
Ratio to WT (log2)
8
4
2
1
0.5
atp1
atp4
atp6-1
atp6-2
atp8
atp9
nad1 e1-2
nad1 e2-3
nad1 e3-4
nad1 e4-5
nad2 e1-2
nad2 e2-3
nad2 e3-4
nad2 e4-5
nad4L
nad4 e2-3
nad4 e3-4
nad5a
nad5b
nad5 e3-4
nad5 e4-5
nad6
nad7 e1-2
nad7 e2-3
nad7 e3-4
nad7 e4-5
nad9
cob
cox1
cox2 e1-2
cox3
rps3 e1-2
rpl16
rps4
rps7
rpl2 e1-2
rpl5
ccmB
ccmC
ccmFN1
ccmFN2
ccmFC e1-2
matR
mttB
rrn18
rrn26

**Figure 7.**

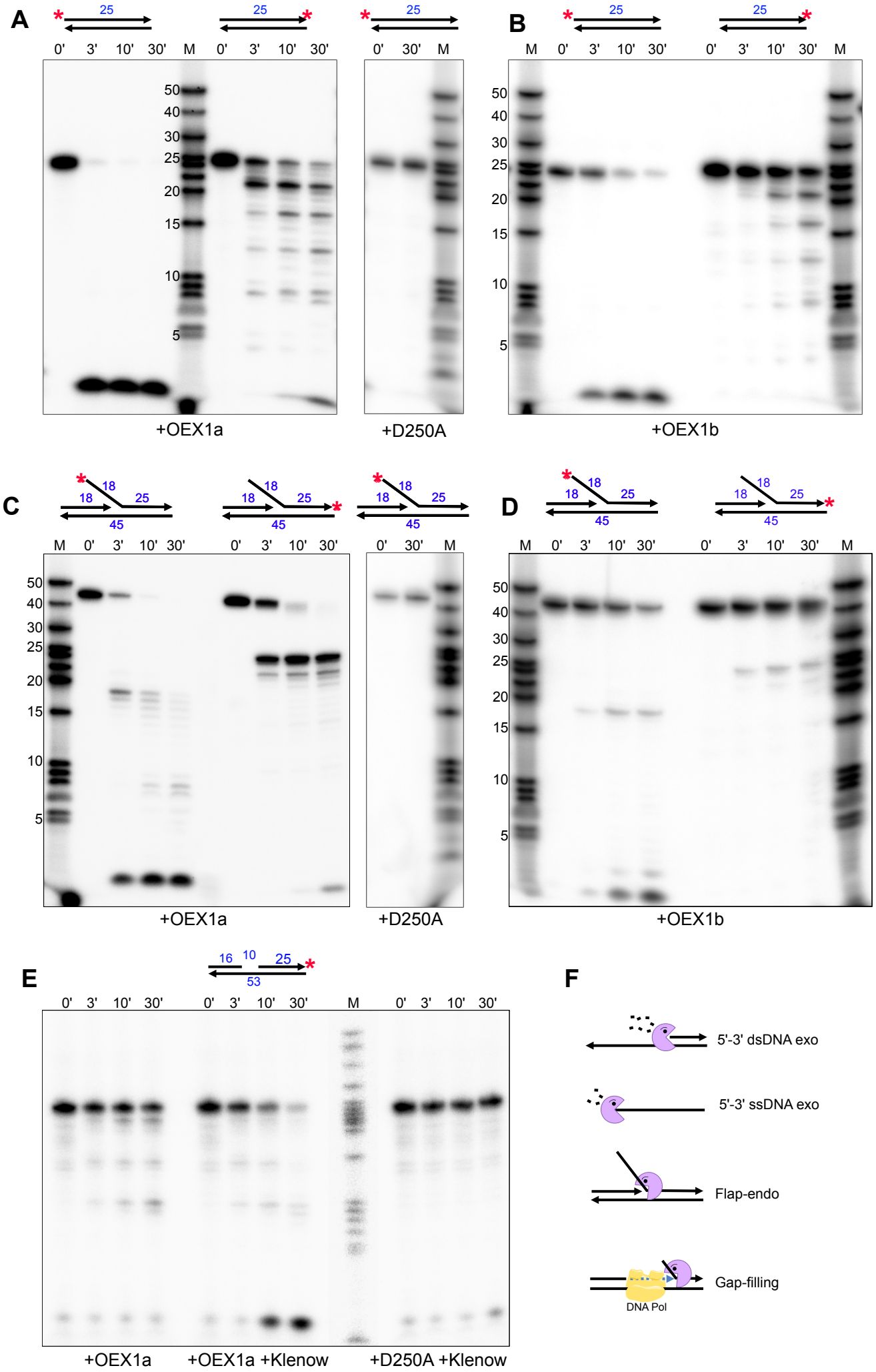
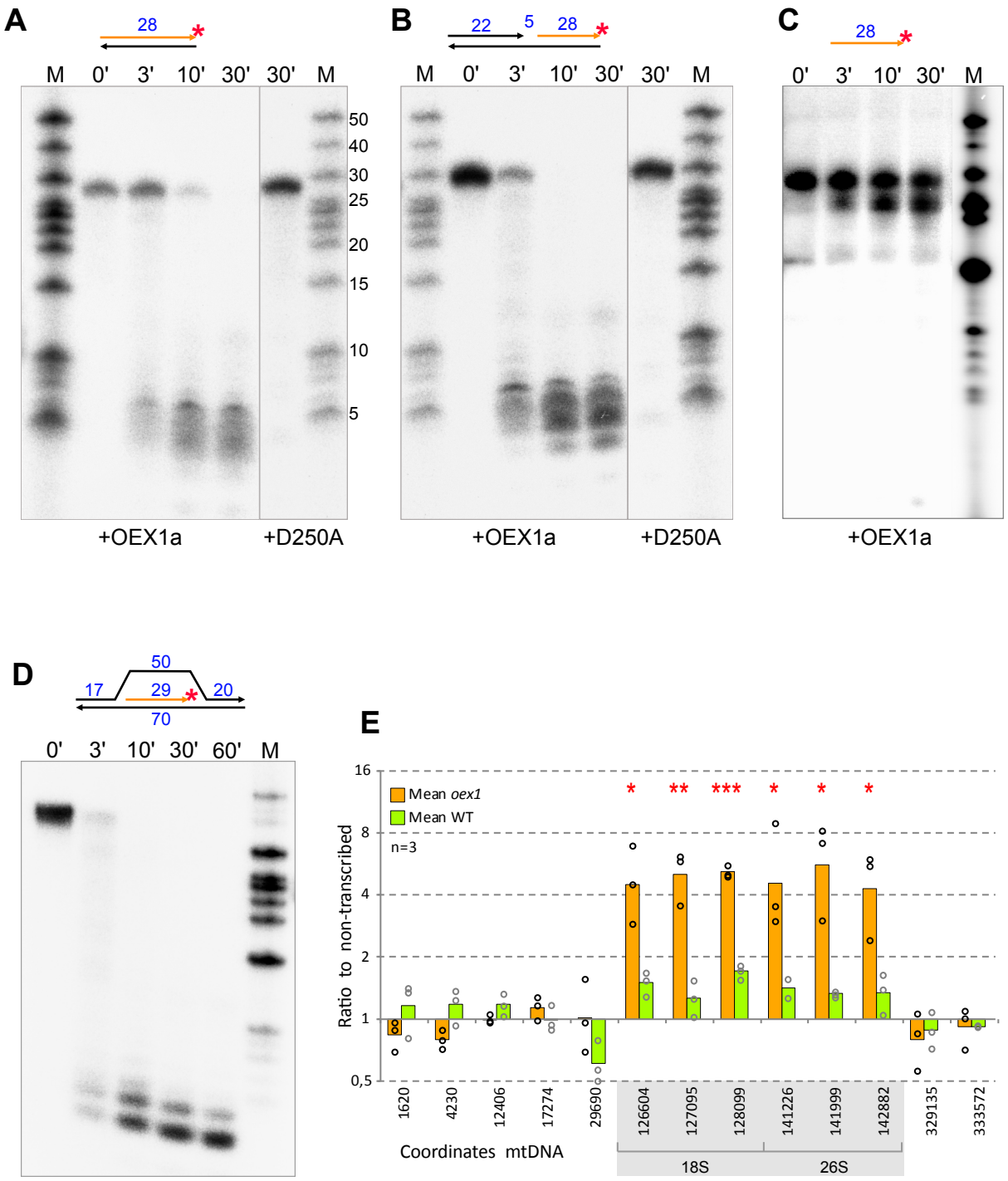

A
25
0' 3' 10' 30' M 0' 3' 10' 30'
0' 30' M
50
40
30
25
20
15
10
5
+OEX1a
+D250A
B
M 0' 3' 10' 30' 0' 3' 10' 30' M
+OEX1b
C
18
25
45
M 0' 3' 10' 30' 0' 3' 10' 30'
0' 30' M
+OEX1a
+D250A
D
M 0' 3' 10' 30' 0' 3' 10' 30' M
+OEX1b
E
16 10 25
53
0' 3' 10' 30' 0' 3' 10' 30' M 0' 3' 10' 30'
+OEX1a
+OEX1a +Klenow
+D250A +Klenow
F
5'-3' dsDNA exo
5'-3' ssDNA exo
Flap-endo
Gap-filling
DNA Pol

**Figure 8.**


A
28
M 0' 3' 10' 30' 30' M
50
40
30
25
20
15
10
5
+OEX1a
+D250A
B
22 5 28
M 0' 3' 10' 30' 30' M
+OEX1a
+D250A
C
28
0' 3' 10' 30' M
+OEX1a
D
50
17 29 20
70
0' 3' 10' 30' 60' M
E
Mean oex1
Mean WT
n=3
Ratio to non-transcribed
16
8
4
2
1
0,5
1620
4230
12406
17274
29690
126604
127095
128099
141226
141999
142882
329135
333572
Coordinates mtDNA
18S
26S