Targeted introduction of heritable point mutations into the plant mitochondrial genome

Joachim Forner[1], **Dennis Kleinschmidt**[1], **Etienne H. Meyer**[1,3], **Axel Fischer**[1], **Robert Morbitzer**[2], **Thomas Lahaye**[2], **Mark A. Schöttler**[1], **Ralph Bock**[1,*]

[1]Max-Planck-Institut für Molekulare Pflanzenphysiologie, Am Mühlenberg 1, D-14476 Potsdam-Golm, Germany

[2]ZMBP, Allgemeine Genetik, Universität Tübingen, Auf der Morgenstelle 32, D-72076 Tübingen, Germany

Abstract

The development of technologies for the genetic manipulation of mitochondrial genomes remains a major challenge. Here we report a method for the targeted introduction of mutations into plant mitochondrial DNA (mtDNA) that we refer to as TALEN-Gene-Drive Mutagenesis (TALEN-GDM). The method combines TALEN-induced site-specific cleavage of the mtDNA with selection for mutations that confer resistance to the TALEN cut. Applying TALEN-GDM to the tobacco mitochondrial *nad9* gene, we isolated a large set of mutants carrying single amino acid substitutions in the Nad9 protein. The mutants could be purified to homochondriomy and stably inherited their edited mtDNA in the expected maternal fashion. TALEN-GDM induces both transitions and transversions, and can access most nucleotide positions within the TALEN binding site. Our work provides an efficient method for targeted mitochondrial genome editing that, for the first time, produced genetically stable, homochondriomic and fertile plants with specific point mutations in their mtDNA.

Introduction

Basic and applied research on plant mitochondria continues to suffer from the lack of tools to engineer plant mitochondrial genomes. Similar to transgenic research with chloroplasts [1,2], the availability of technologies for mitochondrial genome engineering would greatly facilitate functional genomics research (by reverse genetics, e.g., ref. [3]),

[*]For correspondence: Ralph Bock, Max-Planck-Institut für Molekulare Pflanzenphysiologie, Am Mühlenberg 1, D-14476 Potsdam-Golm, Germany. Tel.: +49(0)331-567-8700, Fax: +49(0)331-567-8701. rbock@mpimp-golm.mpg.de .
[3]present address: Institut für Pflanzenphysiologie, Martin-Luther-Universität Halle-Wittenberg, Weinbergweg 10, D-06120 Halle (Saale), Germany

allow the *in vivo* analysis of all steps in mitochondrial gene expression (e.g., refs. [4–6]) and also enable the exploitation of mitochondria in biotechnology and synthetic biology [7,8]. Unfortunately, despite enormous efforts in the academic and industrial sectors, a technology for plant mitochondrial transformation is still not within reach. In addition to the lack of a mitochondria-specific selectable marker gene [9,10], there are unknowns involved in the design of suitable expression cassettes (especially with regard to the largely enigmatic *cis*-elements for translation and translational regulation [11,12]), concerns about the low copy number of the mitochondrial genome [13], and uncertainties about the efficiency of homologous recombination in mitochondria (although it is known that a homologous recombination activity exists [14–16]). In the absence of technologies for direct genetic transformation of the mitochondrial genome, indirect methods for introducing changes into mitochondrial DNA (mtDNA) of seed plants deserve serious consideration.

Genome editing tools have proven to provide powerful and nearly universally applicable methods for targeted mutagenesis [17–19]. Initially, these systems relied on site-directed DNA cleavage and, therefore, were restricted to the generation of insertions and/or deletions in the target sequence. Later, combination of the DNA recognition properties of CRISPR/Cas systems with nucleoside deaminases has produced new genome editing reagents, so-called base editors, that introduce specific point mutations without cutting the target DNA [20–23]. Since all CRISPR/Cas systems rely on RNA-mediated recognition of the target site in the genome by complementary base pairing, they cannot easily be adapted for mitochondrial genome editing. Although import of RNA molecules, especially tRNAs, into mitochondria occurs in a number of organisms [24,25], the import systems appear to have rather strict requirements for RNA structure, sequence and/or size [25,26], and their utilization for the import of CRISPR RNAs has not been accomplished yet. In principle, CRISPR/Cas-based editing tools could be used in those few unicellular organisms (yeast, *Chlamydomonas reinhardtii*) where mitochondrial transformation is available, but they would not be very useful there, given that mitochondrial transformation occurs by homologous recombination [27,28], which represents an even more powerful, versatile and precise genome editing and engineering tool than CRISPR/Cas. Instead, the great attraction of mitochondrial genome editing is that it potentially can make non-transformable species amenable to site-specific sequence alterations in mitochondrial genes.

Given that CRISPR gRNAs cannot readily be targeted to mitochondria, protein-only genome editing methods offer greater promise for mitochondrial genome engineering. Transcription activator-like effector nucleases (TALENs) targeted to mitochondria have been shown to be capable of inducing deletions in the mitochondrial genome [29]. Very recently, a base editing method originally developed for animal mitochondria [30] has been tested in plant tissue culture and shown to be capable of introducing C-to-T mutations in chloroplasts and mitochondria [31,32]. However, genetically stable plants with a homogeneous population of mitochondrial genomes carrying the point mutation (referred to as homochondriomic plants) have not been obtained. Also, fertile base-edited plants that would transmit the genetic alteration in their mtDNA to the progeny have not been reported.

Here we describe the development of a novel method for the targeted introduction of stable point mutations into plant mitochondrial genomes. Combining site-specific genome cleavage

by TALENs with a gene drive-like selection strategy for mutations conferring resistance to TALEN cleavage, we have generated a large set of fertile homochondriomic mutant lines that harbor specific base substitutions in the mitochondrial NADH dehydrogenase subunit 9 (*nad9*) gene of tobacco, *Nicotiana tabacum*.

Results

TALEN-induced targeted cleavage of the mitochondrial genome

Transcription activator-like effectors (TALEs) recognize their target sequence by tandem arrays of 34 amino acid repeats that collectively form a superhelical structure. Each repeat pairs with a single base in the target sequence [33]. Base specificity is conferred by amino acids residues 12 and 13 of the repeat which have been designated as the repeat variable diresidue (RVD). When TALE repeat arrays are fused to the catalytic domain of the restriction endonuclease FokI that induces DNA cleavage upon dimerization, TALE nucleases (TALENs) designed to bind in tail-to-tail orientation on opposite strands of a given target sequence can induce DNA double-strand breaks at specific sites [34,35].

In order to test if TALENs targeted to mitochondria are capable of cleaving the tobacco mitochondrial genome, we selected the mitochondrial *nad9* gene as target for genome editing. *nad9* was chosen because it encodes a subunit of mitochondrial complex I, the NADH-ubiquinone oxidoreductase complex in the inner mitochondrial membrane, which is not essential for photoautotrophic growth [36,37]. This feature should allow the unbiased recovery of all mutations induced in the gene.

Three anti-*nad9* mitochondrial TALEN pairs were designed and the corresponding gene sequences were inserted into vectors for transformation of the plant nuclear genome (see Methods; Fig. 1a). The 19 bp long target sequences of the TALENs should be more than sufficient to confer target site-specific cleavage in the comparably small mitochondrial genome, without causing off-target effects. A mitochondrial presequence encoding a transit peptide for protein import into the mitochondrial compartment was added, and distinct epitope tags (HA and FLAG, respectively) were inserted to facilitate detection of the two TALEN proteins (HA for the 'left' and FLAG for the 'right' TALEN) in transgenic plants. The three constructs (pJF1004, pJF1005 and pJF1006; Fig. 1a) were introduced into tobacco (*Nicotiana tabacum*) cells by *Agrobacterium*-mediated transformation, and approximately 30 transgenic lines per construct were selected (Extended Data Figs. 1 and 2). All T_0 plants were phenotypically inconspicuous, and indistinguishable from wild-type plants when grown under standard greenhouse conditions.

Initial genotyping of the lines by DNA sequencing of the *nad9* locus revealed no indication of the presence of mutations (InDels or SNPs) at or in vicinity of the predicted TALEN cut sites. To test whether cleavage of the mitochondrial genome occurred, Southern blot analyses with *nad9*-specific probes were conducted. Interestingly, several lines (all from the same construct, pJF1006) displayed additional hybridization signals whose sizes were consistent with the predicted sizes of TALEN cleavage products (Fig. 1b; Extended Data Fig. 2a). The same hybridizing fragments were consistently observed in all T_1 progeny of these lines (Extended Data Fig. 2b).

To ultimately confirm that the additional hybridizing fragments originate from TALEN cleavage of the mitochondrial genome, we attempted to map the ends of the fragments. To this end, extracted total plant DNA was treated with mung bean exonuclease to remove single-stranded overhangs from double-strand breaks, followed by ligation to double-stranded DNA adapters (Extended Data Fig. 3a; see Methods for details). Subsequent PCR amplification of ligation products and DNA sequencing indeed revealed cleavage of the *nad9* locus at the TALEN binding sites (Extended Data Fig. 3b,c).

The hybridization intensities of the two TALEN cleavage products relative to the signal intensity of the uncut genome were remarkably similar in all transgenic lines that displayed detectable TALEN activity (Fig. 1b; Extended Data Fig. 2), possibly suggesting that the observed level of cleavage is close to the maximum tolerable by the plant, and/or represents the achievable balance of double-strand break and repair. Interestingly, genotyping of the lines revealed that three out of five Nt-JF1006 lines that showed TALEN-induced cleavage of the mitochondrial genome did not harbor both TALEN genes. Instead, they contained only the gene for one of the two TALEN arms (Fig. 1c; Extended Data Fig. 4), which however, did not appreciably affect TALEN cleavage efficiency. This observation is in line with previous research that has demonstrated that (i) homodimeric TALENs function efficiently in yeast [38], and (ii) a DNA-bound FokI monomer can recruit a second monomer from solution to facilitate cleavage of double-stranded DNA [39]. Since one TALEN arm was obviously sufficient to induce genome cleavage in the mtDNA, most subsequent experiments were conducted with a line (Nt-JF1006-30; Fig. 1b,c) that harbors only one TALEN gene (encoding the 'right' arm; Fig. 1). All mutants presented here are derived from that line.

TALEN cleavage of nuclear DNA typically results in short insertions and/or deletions, due to imperfect double-strand break repair. Our failure to detect such mutations at the *nad9* TALEN cleavage site is consistent with the previously proposed absence of a religation pathway of DNA double-strand break repair from both plastids [40] and plant mitochondria [29]. Instead, double-strand breaks in plant organellar genomes appear to be repaired by homologous recombination or, in rare instances, by microhomology-mediated illegitimate recombination [29,40].

A gene drive-like strategy for induction of point mutations

Given the likely absence of (error-prone) religation of TALEN-induced cuts in the mtDNA and the remarkably stable low-level genome cleavage frequency in all transgenic lines expressing active TALENs (Fig. 1b; Extended Data Fig. 2), we reasoned that any point mutation in the TALEN binding site within *nad9* should confer a strong selective advantage in that it would make the mutated mitochondrial genome less susceptible to the TALEN cut. This is because repair events that restore the wild-type sequence remain TALEN targets and will be cleaved again. By contrast, repair products that contain mutations at the TALEN target site should become fixed since they are no longer TALEN substrates. Thus, continued cleavage activity of TALENs on the wild-type sequence is expected to select for mutant alleles that weaken TALEN binding to the target sequence.

To set up such a gene drive-like system, repeated regeneration rounds in tissue culture were performed to increase the chances of mutant recovery by single-cell passage of mtDNA

(Fig. 2a). To this end, three T_1 seedlings each from five TALEN-active lines (Fig. 1b) were passed through altogether 10 consecutive cycles of regeneration and selection. In each round, regenerating shoots were assayed by sequencing their *nad9* locus, and subsequently used for the next regeneration round. After the sixth round, a single mutant was recovered from TALEN-expressing line Nt-JF1006-30 (Fig. 1b). The line harbored an A-to-G point mutation at position 271 of the *nad9* coding region (A271G; Fig. 2b), leading to an amino acid substitution from serine to glycine (S91G). The corresponding mitochondrial mutant line was named S91G-1 (as the first line recovered with this particular point mutation).

Assuming that TALEN activity was not causally responsible for the appearance of the point mutation but rather selected for it through the above-described gene drive-like mechanism, we reasoned that application of mutagens should greatly promote the recovery of TALEN-resistant mitochondrial mutants (Fig. 2). We, therefore, exposed seeds and leaf explants to two mutagens that have been shown to specifically affect organellar genomes: N-nitroso-N-ethylurea (NEU [41]) and ethidium bromide (EtBr [42]). Seed imbibition in 5 mM NEU or 0.2% EtBr did not result in increased recovery of *nad9* mutants. In each experiment, 50 T_1 seedlings were sequenced, but no mutations in *nad9* were found. Next, we conducted three successive rounds of plant regeneration with leaf explants that either had been dipped into NEU solution prior to their exposure to regeneration medium or were regenerated in the presence of EtBr (Fig. 2a). NEU dipping did not result in a substantial increase in *nad9* mutant recovery. Only a single mutant was identified from 76 sequenced regenerated shoots (Table 1; Supplementary Table 1). By contrast, plant regeneration in the presence of EtBr strongly increased the appearance of *nad9* mutations within the TALEN target sequence, and six new mutants could be identified by sequencing of 94 regenerants (Table 1; Supplementary Table 1). A comparable number of mutants was identified from cultures kept in constant light and cultures exposed to a light-dark cycle (Table 1), suggesting that the light regime does not strongly affect the selection efficiency of mitochondrial mutants.

Homochondriomy of *nad9* mutants

The suspected gene drive-like effect should also strongly accelerate the attainment of the homochondriomic state. Homochondriomy refers to the presence of a homogeneous population of mutated mitochondrial genomes and the absence of residual wild-type copies of the (polyploid) mtDNA [13]. Indeed, nearly all of our mitochondrial mutants quickly reached a state in which the mutated genome represented the prevailing genome type. Typically, already upon initial detection of the mutation by DNA sequencing of the *nad9* locus, the mutated nucleotide accounted for >50% of the peak intensity in the sequencing chromatogram (Fig. 3a). However, a weak signal for the wild-type nucleotide persisted even after multiple additional cycles of plant regeneration, and even after multiple rounds of crosses and sexual propagation.

Previous work on plastid transformation has shown that weak wild-type like signals (detected by DNA sequencing and/or Southern blot hybridization) are not due to persistent heteroplasmy, but are rather explained by the presence of promiscuous plastid DNA sequences in the nuclear genome that originate from rampant gene transfer from the plastid to the nuclear genome [43–46]. This explanation (and the homoplasmy of the transplastomic

plants) were ultimately confirmed by chloroplast isolation and analysis of purified plastid DNA [47,48]. As mtDNA sequences are known to also frequently escape to the nuclear genome [49–51], we reasoned that the small peak for the wild-type nucleotide in our *nad9* mutants (Fig. 3a) may not come from persistent heterochondriomy, but rather from promiscuous mitochondrial *nad9* sequences that reside in the tobacco nuclear genome. To confirm this hypothesis, we isolated mitochondria from mutant plants and sequenced the *nad9* locus from purified mtDNA. This analysis revealed the complete absence of the wild-type peak (Fig. 3a), indicating presence of a homogeneous population of mutated mitochondrial genomes (homochondriomy).

The *nad9* mutant plants flowered readily and produced abundant amounts of seeds. To ultimately confirm homochondriomy, mutant plants were selfed and crossed to wild-type plants. The *nad9* mutations were uniformly present in the progeny of selfed mutants and all crosses using a mitochondrial mutant as maternal parent, in line with the maternal inheritance of the mitochondrial genome [52]. Deep sequencing of *nad9* RT-PCR products (that allowed exclusion of promiscuous *nad9* sequences in the nucleus based on the absence of RNA editing) further confirmed the homochondriomic state of the mutations (Supplementary Data 1).

Taken together, these data suggest that, in our selection system (Fig. 2), mitochondrial mutations attain the homochondriomic state very quickly, presumably due to the strong selective advantage conferred by mutations in the TALEN binding sites (Fig. 3c) and the resulting gene drive-like mechanism. We, therefore, refer to our TALEN-based mutagenesis method as TALEN-GDM (for TALEN-Gene-Drive Mutagenesis).

Characterization of *nad9* mutants

In the experiments described above (Table 1) and additional sets of experiments that are summarized in Supplementary Table 1, altogether 20 independent *nad9* mutants were isolated (Fig. 3b; Supplementary Table 1). Some point mutations were recovered independently up to three times (see Supplementary Table 1). A wide spectrum of mutations (including both transitions and transversions) was obtained, indicating that the method has no pronounced bias in the mutation types that can be recovered. As expected, the mutations clustered in the TALEN binding site, although two point mutations were located a few nucleotides outside of the binding motif. It appears likely that at least one of these mutations (C265A = Q89K) also weakens TALEN binding to the target sequence, as evidenced by quantitation of the TALEN cleavage activity in the mutant *nad9* sequences (Fig. 3c; Extended Data Fig. 5), whereas the other one (G254T = R85L) may be the only mutation that was not obtained by a (pronounced) gene drive effect. This R85L mutant was also the only line that yielded many regenerating shoots (9 out of 22) that displayed only the wild-type *nad9* allele when the initial heterochondriomic mutant plant was subjected to a new regeneration round - something never observed for any of the other mutants.

Interestingly, also a double mutant of two adjacent nucleotides (C279A and G280A; Fig. 3b) was obtained. The two mutations affect neighboring codons and are both non-synonymous substitutions, thus resulting in two adjacent amino acid exchanges in Nad9 (D93E and E94K; Fig. 3b). The two nucleotide substitutions may have arisen independently (in that

each conferred a selective advantage by reducing the sensitivity of the genome to TALEN cleavage), or alternatively, were introduced together upon imperfect repair of a TALEN cut by homologous recombination.

All *nad9* mutants were crossed using the wild type as pollen donor to eliminate the transgenes from the nuclear genome. The plants were also characterized at the cDNA level (by sequencing the entire *nad9* coding region and parts of the 5' and 3' untranslated regions) to assay for the presence of off-target mutations or possible defects in *nad9* mRNA editing that, in theory, could result from the introduced point mutations. No off-target mutations in *nad9* or editing defects were found in any of the mutant lines, suggesting high specificity of the TALEN-GDM method for the targeted sequence in the mitochondrial genome. To confirm the specificity of TALEN-GDM at the genome-wide level, three mutant lines were additionally subjected to next-generation sequencing of the whole mitochondrial genome (Supplementary Data 2; Extended Data Fig. 6). No mutations other than the point mutations in *nad9* were detected in any of the three sequenced genomes. 12 single-nucleotide polymorphisms compared to the reference genome (NC_006581.1) from cultivar Bright Yellow 4 were found in all three mutant lines, but all of these are also present in the wild type of our cultivar Petit Havana, and thus represent intraspecific variation rather than off-target mutations. Furthermore, no sequences corresponding to the wild-type *nad9* allele were detected in the mutants (with nearly 200x coverage), further confirming homochondriomy.

Next, we analyzed the phenotype of the *nad9* mutants. To this end, the entire set of mutants was grown under standard greenhouse conditions, and also under high-light conditions (see Methods) to promote maximum growth rates. No obvious phenotypes were detected, and all mutants were indistinguishable in growth and development from the wild-type controls (Extended Data Figs. 7 and 8). These observations are in line with the surface-exposed location of the targeted Nad9 domain in complex I and its unlikely involvement in complex biogenesis [37].

Selected lines were subjected to detailed physiological and biochemical analyses (Figs. 4 and 5; Extended Data Fig. 9). Measurements of gas exchange rates of plants grown at 350 $\mu E\ m^{-2}\ s^{-1}$ light intensity in a controlled environment chamber revealed no detectable defects in mitochondrial respiration (Fig. 5a). Because defects in complex I function are known to also affect photosynthesis [53], photosynthetic parameters were determined in wild-type and mutant plants. These measurements provided no evidence of impaired electron transfer and/or photosynthetic carbon fixation (Fig. 5a,b).

To specifically examine complex I accumulation and activity, we also analyzed mitochondrial protein complexes in three selected mutants (S91N, D93N and E94K) by native gel electrophoresis and immunoblotting (Fig. 4; Extended Data Fig. 9). While two of the mutants (S91N and D93N) showed no pronounced alterations compared to the wild type, the E94K mutant plants displayed reduced complex I accumulation (Fig. 4a; Extended Data Fig. 9). The functional importance of E94 is in line with the high evolutionary conservation of this position in prokaryotes and eukaryotes (Extended Data Fig. 10). Immunodetection of Nad9 with specific antibodies revealed lower protein accumulation levels in E94K

plants (approximately 60% of wild-type levels; Extended Data Fig. 9), suggesting reduced stability of the mutant Nad9 protein. Interestingly, the Nad9 protein of the E94K mutant also displayed altered electrophoretic mobility (Fig. 4b), thus providing direct biochemical evidence of the significance of the mutation at the protein level.

Discussion

Due to their independence of RNA molecules, transcription activator-like effectors (TALEs) currently provide the method of choice for mitochondrial genome editing [54]. In previous studies, TALENs have been shown to induce large deletions in the mitochondrial genome of plants, presumably due to double-strand break repair by microhomology-mediated illegitimate recombination [29,55]. The difficulties with predicting the breakpoints of such deletions and the risk of the deletion spanning multiple genes and/or including essential mitochondrial genes, limit the usefulness of the method in reverse genetics and mitochondrial genome engineering for biotechnology. Base editing technologies [20–22,56] potentially provide more precise and versatile methods for the targeted introduction of genetic changes into mitochondrial genomes. A recent study demonstrated the possibility to introduce C-to-T mutations into human mtDNA with a bacterial cytidine deaminase fused to TALEs [30]. Preliminary evidence suggests that the method can also be applied to plants, although homochondriomic mitochondrial mutants that are fertile and transmit the mutated mitochondrial genomes into the next generation have not been reported [31]. Moreover, the TALE-cytidine deaminase fusion has several serious limitations, including its restriction to a single mutation type (C-to-T transitions), and a pronounced sequence context preference [30] that makes most positions in the mitochondrial genome inaccessible to editing.

In the course of this work, we have developed a novel method that addresses these limitations. TALEN-GDM relies on site-specific cleavage of the mitochondrial genome and the assumption that the resulting double-strand breaks cannot be fixed by religation, but rather are repaired by homologous recombination (using an uncut genome as repair template [40]) or, in rare cases by microhomology-mediated illegitimate recombination [29]. Since recombination repair involves DNA polymerization, repair errors by nucleotide misincorporation can occur that lower the sensitivity of the target sequence to TALEN-induced cleavage and, in this way, can set off a gene drive-like selection effect. We have shown that application of the DNA-intercalating agent EtBr accelerates the appearance of such mutations, thus enhancing the efficiency of TALEN-GDM (Table 1; Supplementary Table 1).

Applying TALEN-GDM, we successfully isolated a large set of homochondriomic mutants in the mitochondrial *nad9* gene of tobacco (Fig. 3b; Supplementary Table 1). The mutants stably transmitted their genetic changes to the progeny in the expected non-Mendelian (uniparentally maternal) fashion. The frequency of mitochondrial mutant recovery was high: With the EtBr protocol, more than 5% of the sequenced regenerants carried a mutation in the target sequence (Table 1; Fig. 3b). This frequency makes TALEN-GDM a workable method for the efficient introduction of mutations into the plant mitochondrial genome.

In this work, we identified mitochondrial mutants by sequencing of individual regenerated shoots (Fig. 2). However, in view of the fast enrichment of the mutated mitochondrial genome (Fig. 3a), it would be straightforward to set up high-throughput strategies based on sequencing of pooled regenerants, which would greatly increase the power of the method and facilitate the large-scale identification of nearly unlimited numbers of mutations in the target sequence.

Importantly, our TALEN-GDM method does not suffer from the limitations that base editors have with respect to mutation types and sequence context requirements. We have recovered a wide spectrum of different transitions and transversions, and mutations in many different positions of the TALEN binding site (Fig. 3b). These findings suggest that TALEN-GDM provides a relatively unbiased method of site-directed mutagenesis. This feature should make TALEN-GDM a useful tool for the analysis of mitochondrial gene functions by reverse genetics, in that it allows the recovery of viable homochondriomic mutants even for essential mitochondrial genes. Unbiased mutagenesis in a target sequence covering 6-7 codons (Fig. 3b) greatly increases the chances of obtaining useful hypomorphic mutants (in which the mutated gene product possesses a reduced level of activity), similar to our E94K mutant that reduces Nad9 protein accumulation to ~60% of the wild-type level.

The TALEN-GDM method developed in the course of this work will be applicable to all species in which nuclear transformation is available. Thus, TALEN-GDM may also become a valuable tool in applied research aiming at the engineering of agronomically valuable traits into plant mitochondrial genomes. Most importantly, TALEN-GDM may facilitate identification of novel mitochondrial mutations that confer cytoplasmic male sterility (CMS), a trait that is of paramount importance in plant breeding [57–59]. Finally, the TALEN-GDM method will be applicable to all organisms that are recalcitrant to direct transformation of their organellar genomes, including for example, the mitochondria of animals and humans, and the plastids of cereals.

Methods

Construction of TALEN vectors

TALEN sequences were designed according to established principles. Briefly, the TALEN effector binding elements (EBEs) are 18 bp long, preceded by a thymine, and separated by a 12 bp spacer sequence. Suitable target sequences in *nad9* were identified by eye, preferring cytosines and avoiding guanines, because the C-binding repeats have high affinity and specificity, whereas the G-binding repeats have relatively low affinity and specificity. Stretches of identical bases were also avoided. Selection of suitable TALEN EBEs was additionally based on the presence of restriction endonuclease recognition sites (e.g., BsrGI) within the spacer sequence (i.e., the sequence between the two TALEN binding sites).

For the construction of TALEN vectors, TALE-encoding modules lacking the repeats were assembled with BsaI site-flanked modules encoding the short CaMV 35S promoter (from vector pICH51277; ref. [60]), a mitochondrial presequence (AtIVD; ref. [61]) with a triple HA-tag (forward TALEN; 'left' TALEN) or a triple FLAG-tag (reverse TALEN; 'right' TALEN), the truncated TALE N-terminus and C-terminus, the FokI endonuclease [62] and the

OCS terminator from *Agrobacterium tumefaciens* (from plasmid pICH41432; ref. [60]) into vectors pICH47732 and pICH47742 (ref. [60]), yielding vectors pICH47732-TALEN Rep and pICH47742-TALEN Rep, respectively. The repeat domains of the TALENs were created using a previously described method [63] and cloned via BpiI sites into pICH47732-TALEN Rep and pICH47742-TALEN Rep. TALEN modules with repeats were assembled together with pICH47751-Kanamycin, conferring *in planta* resistance to kanamycin, and pICH47766 into pICH50505 (ref. [60]) using the BpiI restriction sites. The plasmid sequences of the final plant transformation vectors pJF1004, pJF1005 and pJF1006 were deposited in GenBank under accession numbers MZ417370, MZ417371 and MZ417372. In all constructs, the RVDs NI (for A), HD (for C), NN (for G), and NG (for T) were used.

Plant material and growth conditions

Nicotiana tabacum cultivar Petit Havana was used for all experiments. Plants were raised from seeds under aseptic conditions on synthetic medium (MSsuc3 medium). The medium consists of Murashige and Skoog (MS) salts [64] supplemented with modified MS vitamins (Duchefa M0245) and 3% sucrose. The pH was adjusted to 5.8 and the medium solidified with 0.68% agar (Duchefa M1002). Calli and shoots were induced on medium NtSIM1, which is MSsuc3 medium supplemented with 0.1 g L^{-1} 1-naphthaleneacetic acid and 1 mg L^{-1} 6-benzylaminopurine. Rooting of regenerated shoots was induced in MSsuc3 medium.

If not stated otherwise, the photoperiod was 16 h of daylight (light intensity 25-50 μE m^{-2} s^{-1}) at 25°C followed by 8 h of darkness at 20°C. Seeds were surface-sterilized in a desiccator with chlorine gas (released by adding 3% v/v of 37% HCl to a 13% solution of NaOCl) for 5 h or, alternatively, by incubation in 70% ethanol (+0.3% Tween20) followed by incubation in 6% NaOCl (+0.3% Tween20) for 2 min each. In the greenhouse, plants were grown in soil under standard conditions (16 h day length; temperature regime: 21°C during the day, 18°C at night; average light intensity: 300 μE m^{-2} s^{-1}). For some growth analyses and physiological measurements, plants were grown in a controlled environment chamber under long-day conditions (16 h light) at 22°C and 75% relative humidity during the day, and 18°C and 60% relative humidity during the night. For the high-light growth experiments, the light intensity was set to 1000 μE m^{-2} s^{-1}. For the gas exchange and photosynthetic analyses, plants were grown at 350 μE m^{-2} s^{-1} light intensity.

Plant transformation and characterization of transgenic lines

Agrobacterium tumefaciens strain GV2260 was used for plant transformation experiments. Transgenic lines were selected on an MS-based regeneration medium supplemented with 50 μg/mL kanamycin.

To identify plants that functionally express the TALEN transgenes (i.e., display cleavage of the mitochondrial *nad9* locus), total DNA was isolated by a cetyltrimethylammoniumbromide (CTAB)-based method [65] from T_1 plants and subjected to RFLP analysis by Southern blotting. To this end, DNA samples were digested with restriction enzymes, separated in 1% or 2% agarose gels by electrophoresis and blotted onto Hybond-XL membranes (GE Healthcare). Primers oJF271 and oJF272 (Supplementary Table 2) were used to produce an *nad9*-specific hybridization probe by PCR amplification.

The probe was purified with the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel) and labelled with α[^{32}P]dCTP by random priming (GE Healthcare) according to the instructions of the manufacturer. Hybridizations were performed at 65°C in Church buffer (1% BSA, 1 mM EDTA, 7% SDS, 500 mM Na$_2$HPO$_4$, pH 7.2) overnight. Signals were detected with a Typhoon Scanner (GE Healthcare). The *nad9* cleavage efficiency was quantified using the Fiji (ImageJ) image processing package by calculating the amount of signal in the two TALEN-cut bands (712 bp and 436 bp) as percentage of the total *nad9* signal (uncut 1148 bp band plus the two TALEN-cut bands).

The *nad9* sequences in the T$_0$ plants were screened for mutations by PCR amplification with primers oJF271 and oJF272, followed by bulk Sanger sequencing of the amplification products (LGC Genomics). Template DNA was extracted with Extraction Solution Molecular Biology (Sigma-Aldrich). To address the possibility that plants might be mosaic and/or heterochondriomic, DNA for PCR analyses was always extracted from a pool of material harvested from three consecutive leaves (to likely cover offspring of all stem cells in all layers).

Presence of the genes encoding each TALEN arm in the TALEN-active lines was analyzed by PCR with primers oJF1147 and oJF1323 (Supplementary Table 2). β-TUBULIN as control locus was amplified with primers oJF1028 and oJF1029. The DNA size marker shown in all PCR gel images is the GeneRuler 100 bp Plus DNA Ladder (Thermo Scientific). RT-PCR with cDNA synthesized from total RNA with primers oJF1026 (β-TUBULIN) and oJF1300 (TALENs) was performed to assess expression of the HA- and FLAG-TALEN arms using primers oJF1027 and oJF748 (for β-TUBULIN as internal control), oJF1298 and oJF748 (HA-tag specific), oJF1299 and oJF748 (FLAG-tag specific), and oJF1147 and oJF748 (for amplification of both HA-tag and FLAG-tag), respectively.

To verify that the double-strand break in *nad9* detected in our Southern blot analyses is indeed caused by TALEN cuts, total DNA was blunted using mung bean exonuclease (NEB) and ligated to a short synthetic double-stranded DNA fragment created by annealing oligonucleotides oJF650 and oJF651 (Supplementary Table 2). Ligation products with *nad9* were amplified using an adapter primer (oJF060) combined with the following gene-specific primers: oJF271 (upstream 1), oJF317 (upstream 2), oJF272 (downstream 1) and oJF499 (downstream 2; Supplementary Table 2; Extended Data Fig. 3).

Identification of mitochondrial mutants

To isolate mitochondrial mutants, verified TALEN-expressing lines were raised from seeds under aseptic conditions, leaves were cut into pieces and placed onto shoot induction medium (NtSIM1; see above). Shoots were transferred to Magenta boxes with rooting medium (MSsuc3; see above) and genotyped. If no mutation in *nad9* was found, a new regeneration round was initiated with leaf pieces from these plants. This process was repeated up to ten times. If a mutation was found, an additional regeneration round was conducted to reduce actual or apparent heterochondriomy. After rooting, confirmed homochondriomic plants were transferred to the greenhouse for seed production by selfing or fertilization with wild-type pollen (to remove possible tissue culture-induced

somaclonal variation from the nuclear genome). Descendants of this backcross were used for phenotyping.

For experiments with mutagens, the mutagenic chemicals were either added to the culture medium or used for seed imbibition or treatment of leaf pieces. Ethidium bromide (EtBr) was applied as aqueous solution at concentrations of 0.2% for seed imbibition, and 0.0002% to 0.002% in the culture medium. N-nitroso-N-ethylurea (NEU) was dissolved in a mixture of 70% ethanol and 0.1% acetic acid and used at a final concentration of 5 mM.

To investigate the apparent heterochondriomy caused by the presence of promiscuous mtDNA in the nuclear genome, genotyping was performed with DNA extracted from purified mitochondria (see below). In addition, total RNA was isolated from mutant plants with the NucleoSpin RNA Plant kit (Macherey-Nagel) and reverse transcribed with SuperScript III Reverse Transcriptase (Invitrogen) using the *nad9*-specific primer oJF1113 (Supplementary Table 2). Full-length *nad9* cDNA was then amplified with primers oJF311 and oJF748 (Supplementary Table 2), RT-PCR products were purified and bulk sequenced as described above. In addition to the homo- versus heterochondriomic status of the mutation of interest, the cDNA sequences were also analyzed for possible aberrations in mRNA editing patterns and the possible presence of additional (off-target) mutations elsewhere in the coding sequence of *nad9*. For amplicon deep sequencing, total RNA isolated from kanamycin-resistant seedlings was reverse transcribed with primer oJF1302 and used as a template for PCR amplification with primers oJF1304 and oJF1301 (Supplementary Table 2).

Isolation of mitochondria

Mitochondria were purified from 40 g of fresh leaf material. The leaves were homogenized in extraction buffer (300 mM sucrose, 15 mM potassium pyrophosphate, 2 mM EDTA, 10 mM KH_2PO_4, 1% (w/v) PVP-40, 1% (w/v) BSA, 20 mM sodium ascorbate, 5 mM cysteine, pH 7.5) in a Waring blender. The homogenate was filtered through two layers of Miracloth and then centrifuged at 2000 g. The supernatant was collected, centrifuged at 20,000 g, and the pellet was resuspended in wash buffer (300 mM sucrose, 1 mM EGTA, 10 mM MOPS-KOH, pH 7.2). Following centrifugation at 2000 g, the supernatant was collected, centrifuged at 20,000 g, and the pellet obtained was resuspended in a small volume of wash buffer and then loaded onto a Percoll step gradient consisting of 1 volume of 50% (v/v) Percoll, 5 volumes of 25% (v/v) Percoll and 1 volume of 18% (v/v) Percoll (all solutions prepared in gradient buffer: 300 mM sucrose, 10 mM MOPS-KOH, pH 7.2). The gradient was centrifuged at 40,000 g for 45 min, the mitochondria were collected from the interface between the 50% and 25% solutions, and subsequently washed several times with wash buffer. Protein concentration was estimated using the Bradford method (ROTI-Quant, Carl Roth GmbH, Karlsruhe, Germany).

Protein extraction and immunoblot analysis

Samples of 20 μg extracted mitochondrial protein were solubilized in loading buffer (80 mM Tris-HCl, pH 6.8, 25 mM EDTA, 0.1 M DTT, 1% (v/v) glycerol, 2% (w/v) SDS, 0.05% (w/v) bromophenol blue) and separated by electrophoresis in 12% SDS-PAA gels.

Proteins were then transferred to PVDF membranes (Immobilon-P, Merck Millipore) in a tank blotter (transfer buffer: 192 mM glycine, 25 mM Tris, pH 8.3). Equal loading and successful transfer were confirmed by Coomassie staining of the membrane. The blotted proteins were then immunodecorated with specific antibodies, and the signals detected by chemiluminescence with secondary horseradish peroxidase-conjugated antibodies (ECL Prime, GE Healthcare). Images were acquired using the Fusion-FX system (Vilber Lourmat GmbH, Eberhardzell, Germany). Polyclonal antibodies against Nad9 (ref. [66]) were used at a dilution of 1:20,000, polyclonal antibodies against Cox1 (ref. [67]) at a dilution of 1:10,000. Anti-rabbit-HRP conjugate from Sigma (A0545) was used as secondary antibody used at a dilution of 1:10,000.

BN-PAGE and NADH oxidase staining

For analysis of mitochondrial membrane protein complexes via blue-native polyacrylamide gel electrophoresis (BN-PAGE), mitochondria equivalent to 200 μg protein were solubilized with 5% (w/v) digitonin (solubilization buffer: 150 mM potassium acetate, 10% (v/v) glycerol, 30 mM HEPES, pH 7.4) and loaded on a 4.5-16% gradient BN-PAA gel. Following electrophoretic separation of protein complexes, the gel was stained with colloidal Coomassie. For NADH oxidase staining, the gel was washed twice in water, and then incubated in staining solution (0.14 mM NADH, 1 mg/mL nitroblue tetrazolium, 100 mM Tris-HCl, pH 7.4). When satisfactory staining intensity had been reached, the reaction was stopped with fixing solution (50% (v/v) methanol, 10% (v/v) acetic acid). Stained gels were scanned using a flatbed scanner. Protein complexes from replicate gels were transferred to PVDF membranes (Immobilon-P) by tank blotting (transfer buffer: 192 mM glycine, 25 mM Tris, pH 8.3). Equal loading and successful protein transfer were confirmed by Coomassie staining of the membranes. Proteins were labelled with specific antibodies and detected by chemiluminescence as described above. Polyclonal antibodies against the CA2 subunit of complex I (ref. [68]) were used at a dilution of 1:10,000.

Measurements of gas exchange and photosynthetic parameters

Gas exchange measurements were performed with a GFS-3000 gas exchange system equipped with the LED array unit 3056-FL as an actinic light source (Heinz Walz GmbH, Effeltrich, Germany). Light response curves of CO_2 assimilation were recorded at 22°C cuvette temperature and 17,500 ppm humidity. Measurements were performed at two CO_2 concentrations: 400 ppm, and 2000 ppm (a saturating CO_2 concentration) to fully repress photorespiration and determine the capacity of photosynthesis. The youngest fully expanded leaves of plants grown at 350 μE m^{-2} s^{-1} light intensity were used for the measurements. At this developmental state, tobacco leaves have their highest capacity of both respiration and assimilation. After leaf respiration had been determined in darkness, the actinic light intensity was increased stepwise to 100, 350, 750, and finally 1500 μE m^{-2} s^{-1}. At each light intensity, gas exchange was recorded until the steady state of leaf assimilation had been reached. After the final saturating illumination step at 1500 μE m^{-2} s^{-1}, the actinic light was switched off, and respiration was recorded until a constant respiration rate in darkness had been reached.

Parameters of photosynthetic electron transport were determined with the MODULAR version of the Dual-PAM-100 instrument (Heinz Walz GmbH) at 22°C. Light response curves of chlorophyll-a fluorescence and PSI parameters were measured after 20 min of dark adaptation, to first determine the maximum quantum efficiency of PSII in the dark-adapted state (F_V/F_M). Afterwards, light response curves of linear electron transport, the chlorophyll-a fluorescence parameters non-photochemical quenching (qN; ref. [69]) and qL, a measure of the redox state of the PSII acceptor side [70], and of the acceptor-side (Y(NA)) and donor-side limitation of PSI (Y(ND)) were determined [71]. The measuring times at each actinic light intensity were 150 s under light-limited conditions, and 60 s under light-saturated conditions.

Linear electron transport was corrected for leaf absorptance, which was calculated as 100% incident light minus light transmitted through the leaf (%) minus light reflected on the leaf surface (%). Transmittance and reflectance spectra between 400 and 700 nm wavelength were recorded using an integrating sphere attached to a photometer (V650, Jasco GmbH, Groß-Umstadt, Germany). The spectral bandwidth was set to 1 nm, and the scanning speed was 200 nm min^{-1}. Measurements of chlorophyll content and the chlorophyll a/b ratio were done with a Jasco V-630 photometer (Jasco GmbH, Groß-Umstadt, Germany) in 80% (v/v) acetone [72].

DNA sequence analyses and artwork

The Lasergene suite (DNASTAR) and the SnapGene Viewer (https://www.snapgene.com/snapgene-viewer/) were used for DNA sequence analyses. Figure 2a and Extended Data Figure 1 were created with pre-drawn icons from BioRender (https://biorender.com).

Next-generation sequencing (NGS)

NGS was performed at the Sequencing Core Facility of the Max Planck Institute for Molecular Genetics (Berlin). For NGS analysis of mitochondrial genomes (mtDNA-seq), DNA was extracted from an aliquot of the mitochondria isolated for protein analysis. After initial quality control with the Bioanalyzer (Agilent), sequencing libraries were prepared from 30 ng of DNA per sample following the KAPA DNA HyperPrep Kit (Roche) library preparation protocol for double-indexed Illumina libraries. Briefly, DNA was sheared using a Covaris S2 system (duty cycle 5%, intensity 5, 40 s run time). After end repair and A-tailing, Illumina sequencing-compatible adapters carrying unique dual indices were ligated (NEXTFLEX® Unique Dual Index Barcodes). Following bead-based clean-up steps, the libraries were amplified using 7 cycles of PCR. Library quality and size were checked with qBit, the Bioanalyzer and qPCR. Sequencing was carried out on an Illumina MiSeq Nano flow cell in PE150bp mode, yielding between 240,000 and 260,000 sequenced fragments per sample. The read coverage across the mitochondrial genome is displayed in Extended Data Fig. 10.

For NGS amplicon sequencing (amplicon-seq), RT-PCR products were purified using the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel). After initial quality control with the Bioanalyzer, sequencing libraries were prepared from 50 ng of DNA per sample according to the KAPA DNA HyperPrep Kit (Roche) library preparation protocol for

double-indexed Illumina libraries. After end repair and A-tailing, Illumina sequencing-compatible adapters carrying unique dual indices were ligated (NEXTFLEX® Unique Dual Index Barcodes). Following bead-based clean-up steps, the libraries were amplified using 5 cycles of PCR, and library quality and size were checked with qBit, the Bioanalyzer and qPCR. Sequencing was carried out on an Illumina MiSeq Micro flow cell in PE150bp mode, yielding between 140,000 and 200,000 sequenced fragments per sample.

Bioinformatic analyses of NGS data

An initial quality check of the sequence data of all samples from both NGS datasets (mtDNA-seq and amplicon-seq) was performed with FastQC v0.11.9 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Three additional pre-mapping steps were conducted for the amplicon-seq data. First, read pairs were joined using PEAR (ref. [73]) v0.9.11 [-v 65] to avoid doubled coverage counting and increase the quality of the overlapping part of the pairs, in which the introduced mutations are located. Second, Flexbar (ref. [74]) v3.5 was used to remove the additional amplicon-specific sequence (CGTCACGAACGTACTTTGGA; introduced during the cDNA synthesis step as 5'extension of primer oJF1302) with the adapted options -at ANY -m 213. Third, reads representing the reference DNA allele [C] instead of the known RNA-edited sequence [T] within the sequences GTTT[C/T]GAT, GGC[C/T]GGTG and TCT[C/T]CGGT were considered to originate from contaminating DNA and removed.

Fastq files of both datasets were then mapped against the NCBI NC_006581.1 reference with BWA v0.7.17 in MEM mode [75]. Sample-wise alignment files were merged dataset-wise with SAMtools (ref. [76]) v1.12. Single nucleotide polymorphisms were afterwards called with Freebayes (ref. [77]) v1.3.2 with slightly adapted options (-p 1 -n 2) and annotated with SnpEff (ref. [78]) v4.3k. The database was built from GenBank file of NC_006581.1 using the -genbank build option in SnpEff. To confirm homochondriomy, alignment files of the amplicon-seq dataset were further processed with SAMtools mpileup (-d 200000 -Q 40), and nucleotide-wise coverage was extracted with a user-defined Perl script for all positions at which SNPs were introduced and 4 additional randomly chosen positions as non-SNP control sites.

Extended Data
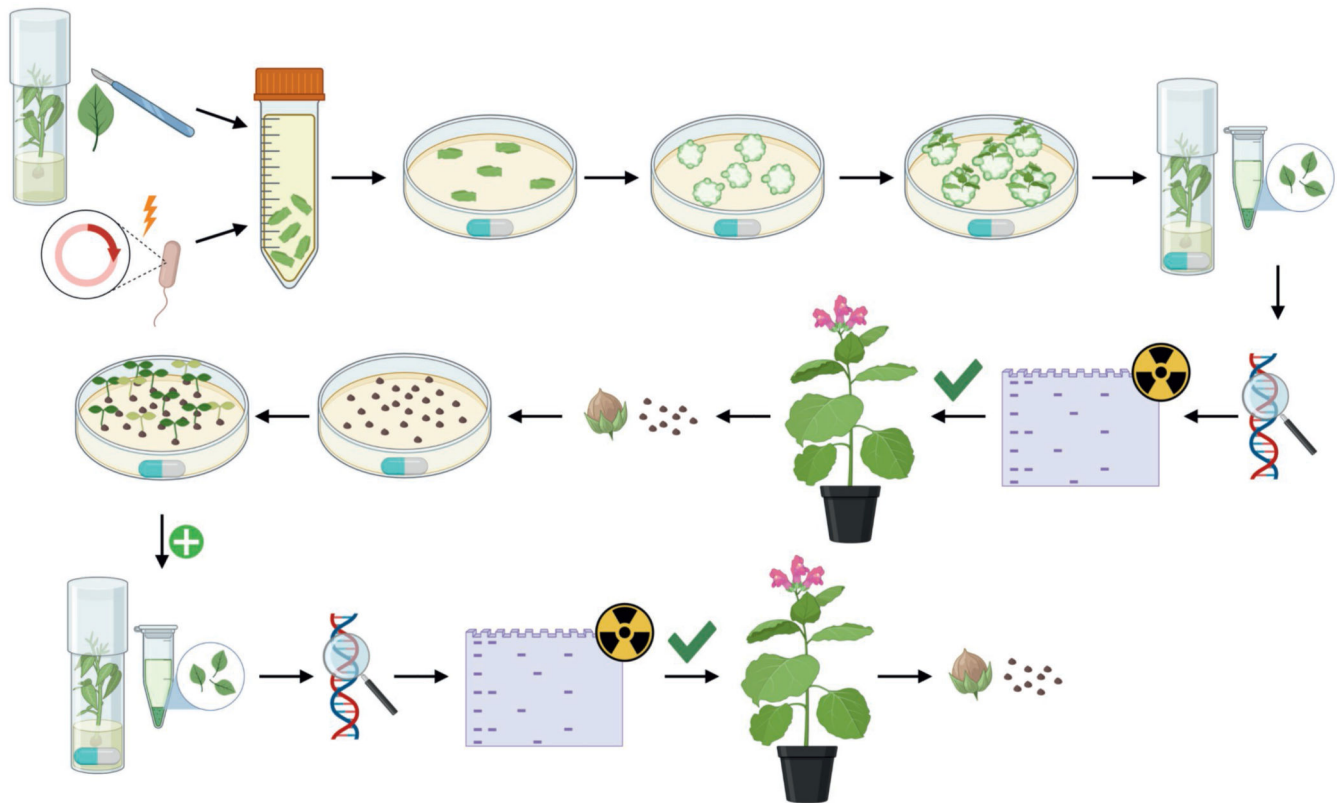
Extended Data Fig. 1. Isolation of TALEN-expressing lines.
The workflow illustrating the procedures involved in the generation of transgenic TALEN-expressing lines and the detection of TALEN cleavage activity in the mitochondrial genome is schematically shown. Tobacco leaves are transformed with *Agrobacterium tumefaciens* cells harboring TALEN-encoding plasmids. Putative transgenic T_0 shoots are regenerated on medium containing kanamycin, rooted and transferred to the greenhouse. DNA is isolated from these plants and assayed by Southern blotting for TALEN activity (i.e., cleavage of *nad9*). Individual T_1 seedlings derived from these plants (and selected for kanamycin resistance) are screened again for TALEN activity and then used for T_2 seed production. The blue-white pill indicates presence of kanamycin in the medium.
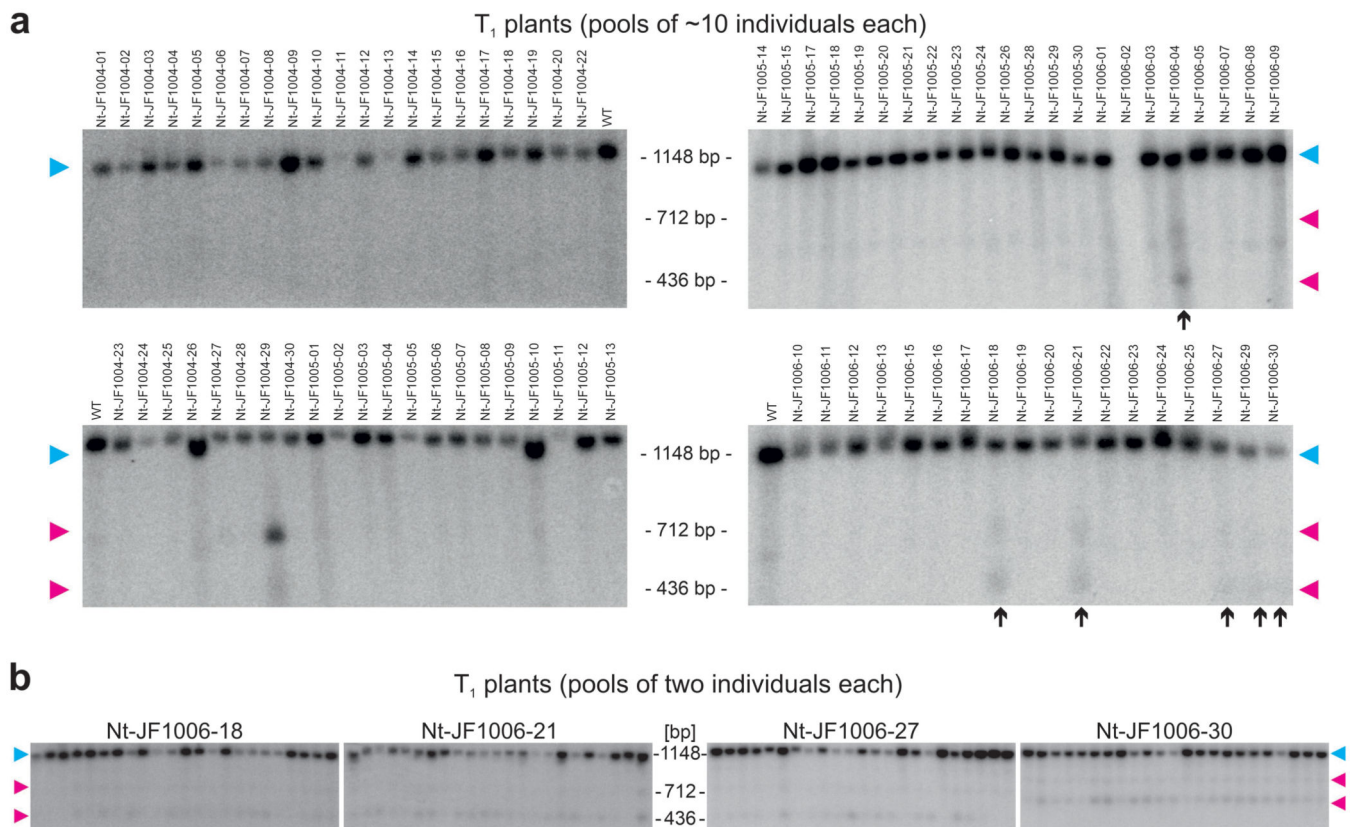
Extended Data Fig. 2. Southern blot analyses to identify lines with stable functional TALEN expression. For details, see Fig. 1b.
(**a**) The blots in the upper panel show the full screen of all transgenic plants obtained with the three TALEN constructs. Pools of approximately ten kanamycin-resistant T_1 seedlings were used for DNA extraction. Candidate lines showing a hybridization pattern consistent with TALEN cleavage activity are marked by black arrows. (**b**) Analysis of additional T_1 descendants of four candidate T_0 lines. Each lane represents a pool of two individual plants. The 1148 bp fragment (*nad9* not cut by TALENs) is marked by a cyan arrowhead, the two fragments resulting from the TALEN cut (712 bp and 436 bp) are indicated by magenta arrowheads. All Southern blots shown here were performed once.
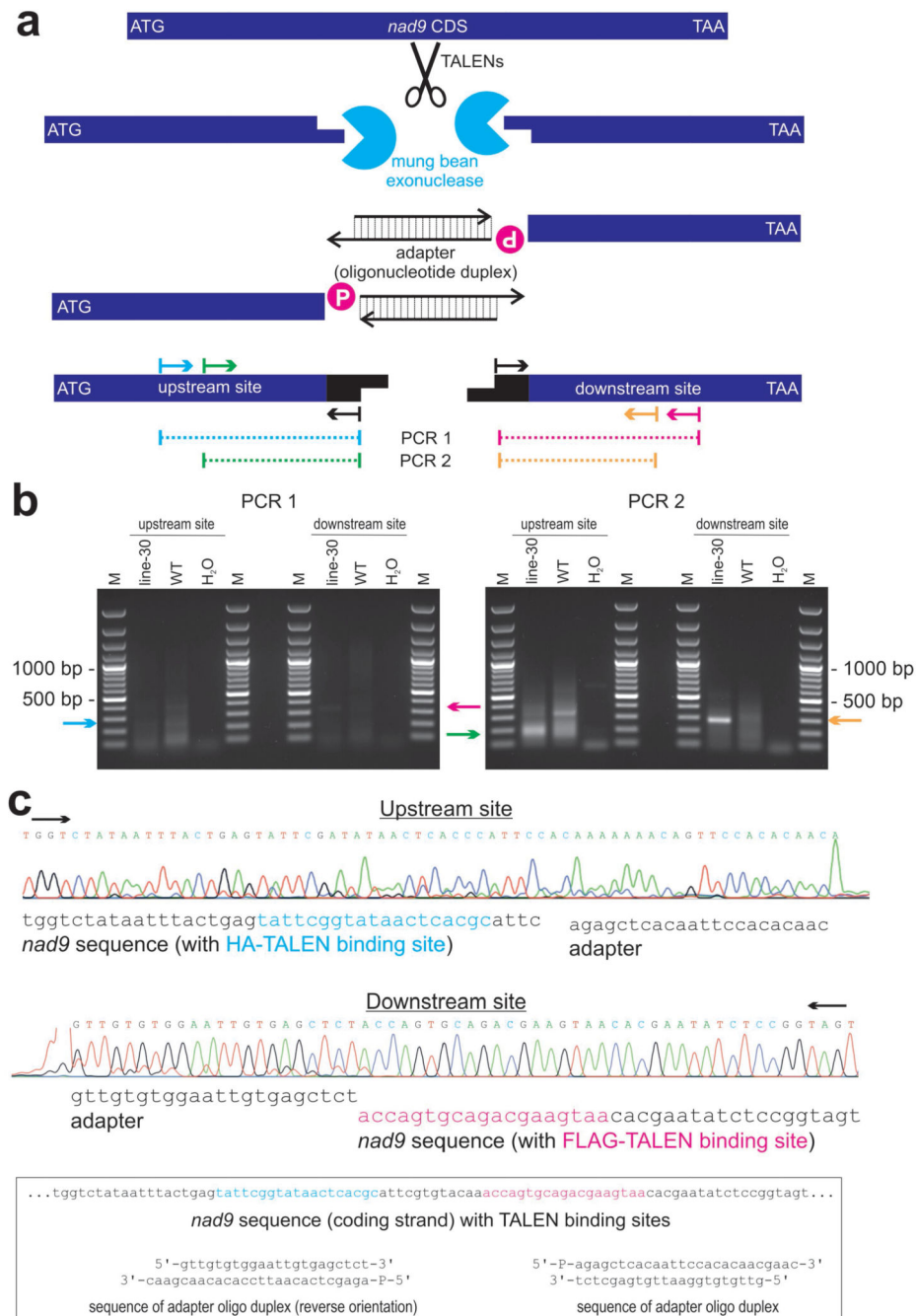
Extended Data Fig. 3. Detection of TALEN-induced DNA double-strand breaks.
(**a**) End-mapping strategy to verify that the TALENs cut at the predicted sites. Arrows indicate oligonucleotides or primer-binding sites. mtDNA cut by the TALENs *in vivo* is blunted by mung bean exonuclease *in vitro*, followed by adapter ligation and PCR amplification. Note that PCR 2 is a half-nested PCR using the products of PCR 1 as template. P, 5' monophosphate group of oligonucleotides. (**b**) PCR results obtained from execution of the strategy depicted in panel (a). Arrows indicate the expected sizes of products derived from TALEN-cut mtDNA. Cyan arrow, expected size (270 bp) of the

amplification product of upstream site PCR 1; magenta arrow, expected product size (350 bp) for downstream site PCR 1; green arrow, expected amplicon size (192 bp) for upstream site PCR 2; orange arrow, expected amplicon size (288 bp) for downstream site PCR 2; line-30, line Nt-JF1006-30; WT, wild-type control; H₂O, water control; M, DNA size marker. One of two similar experiments with similar results is shown. (**c**) Sequencing chromatograms for the products of PCR 2. The chromatograms show the DNA sequence obtained from sequencing the PCR products marked by the green (upstream site) and orange (downstream site) arrows, respectively, in panel (b). For comparison, the sequences of the adapter oligonucleotide duplex and the *nad9* TALEN target site are also given (boxed). Black arrows indicate the direction of the sequencing reactions.
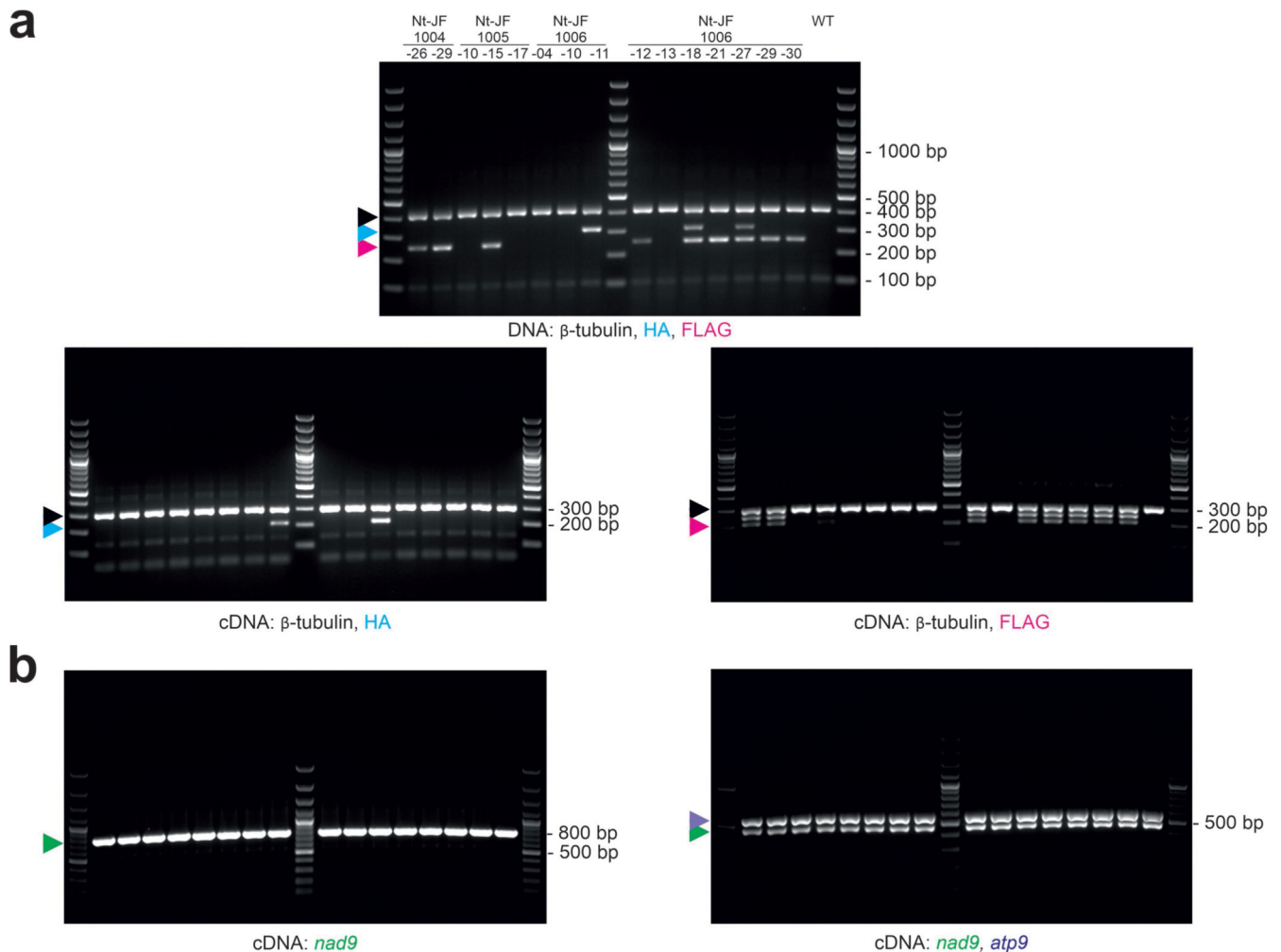
Extended Data Fig. 4. Tests for presence of HA- and FLAG-TALEN arms, and analysis of *nad9* expression in a subset of Nt-JF1004, Nt-JF1005 and Nt-JF1006 lines.
(**a**) Extension of the analyses shown in Fig. 1c to a larger set of plants. See legend to Fig. 1c for details. (**b**) Analyses of the same plants as shown in (a) by amplification of full-length *nad9* cDNA with primers oJF311 and oJF748 (yielding a product of 769 bp, indicated by the green arrowhead; left panel) for subsequent analysis by Sanger sequencing, and partial *nad9* cDNA with primers oJF496 and oJF748 (465 bp, indicated by the green arrowhead;

right panel) and full-length *atp9* cDNA with primers oJF1145 and oJF748 (568 bp, purple arrowhead) for semiquantitative analysis of expression strength. Primers for first strand cDNA synthesis were oJF1113 (*nad9*) and oJF1144 (*atp9*). Sequencing analyses showed all full-length *nad9* RT-PCR products to be fully edited and free of mutations. The loading scheme is identical for all five gels. For primer sequences, see Extended Data Table 2. The experiments shown in (a) and (b) were performed once.
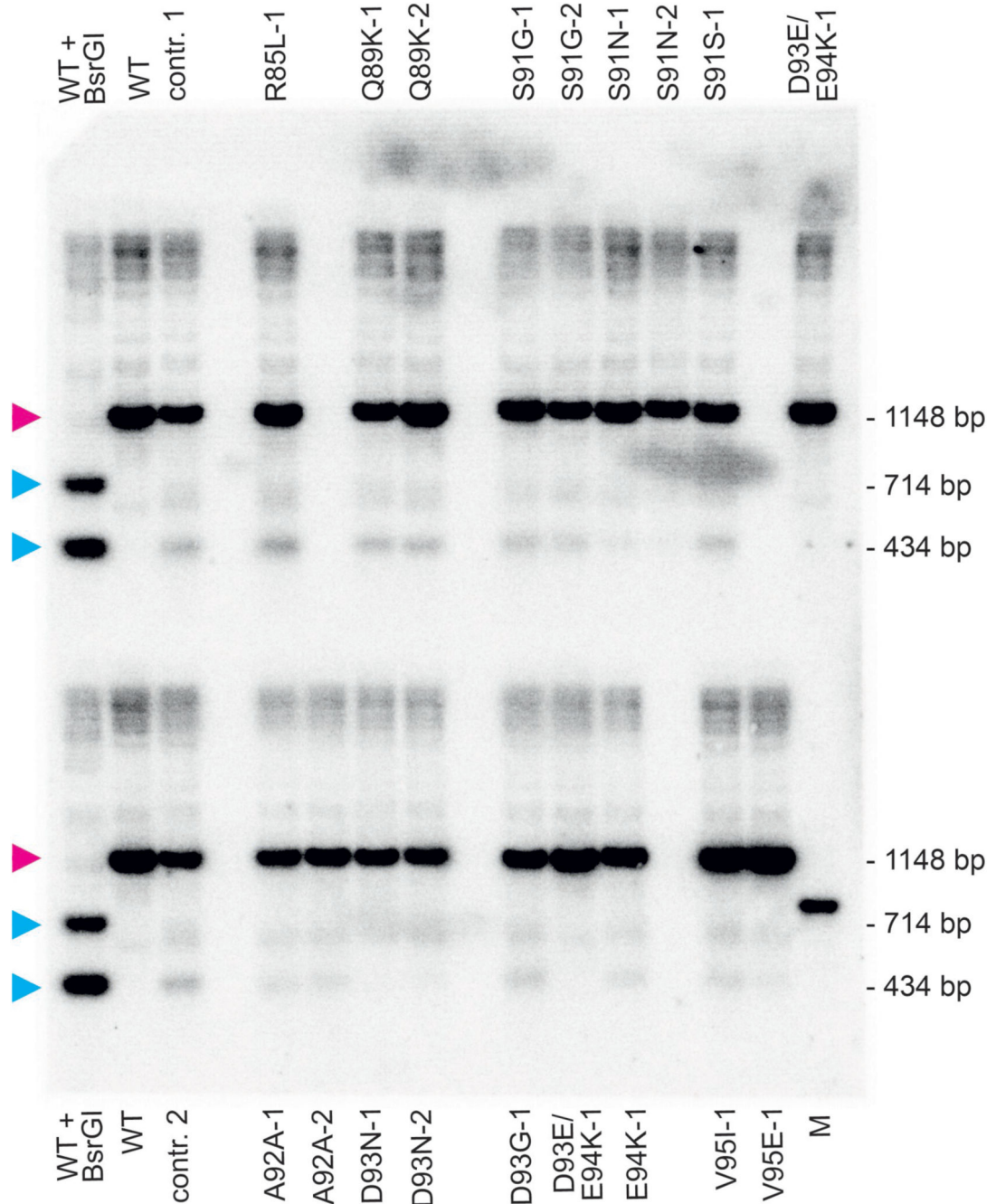
Extended Data Fig. 5. Southern blot from Fig. 3c shown with enhanced contrast to better visualize weak hybridization signals.

M, DNA size marker (with the 800 bp marker fragment cross-hybridizing to the probe).
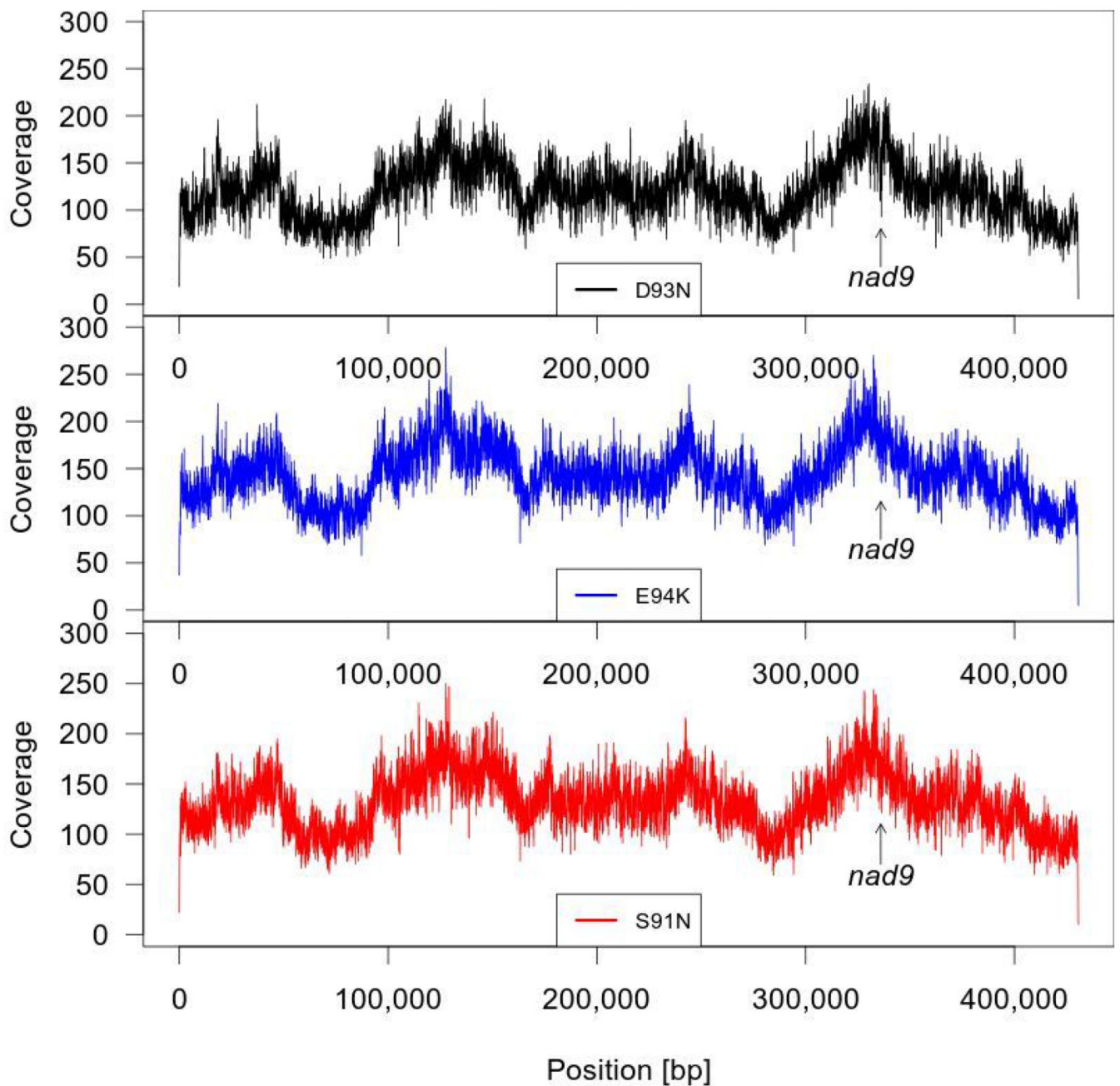
Extended Data Fig. 6. Read coverage across the whole mitochondrial genome in nextgeneration sequencing experiments.

The mitochondrial genomes of mutant lines D93N-1, E94K-1 and S91N-1 were sequenced. Nucleotide positions [bp] on the x-axis are according to the tobacco mitochondrial reference genome sequence (NC_006581). The y-axis shows the number of reads per position. Upper panel: D93N-1 (black), middle panel: E94K-1 (blue), lower panel: S91N-1 (red). The arrow indicates the position of *nad9*.
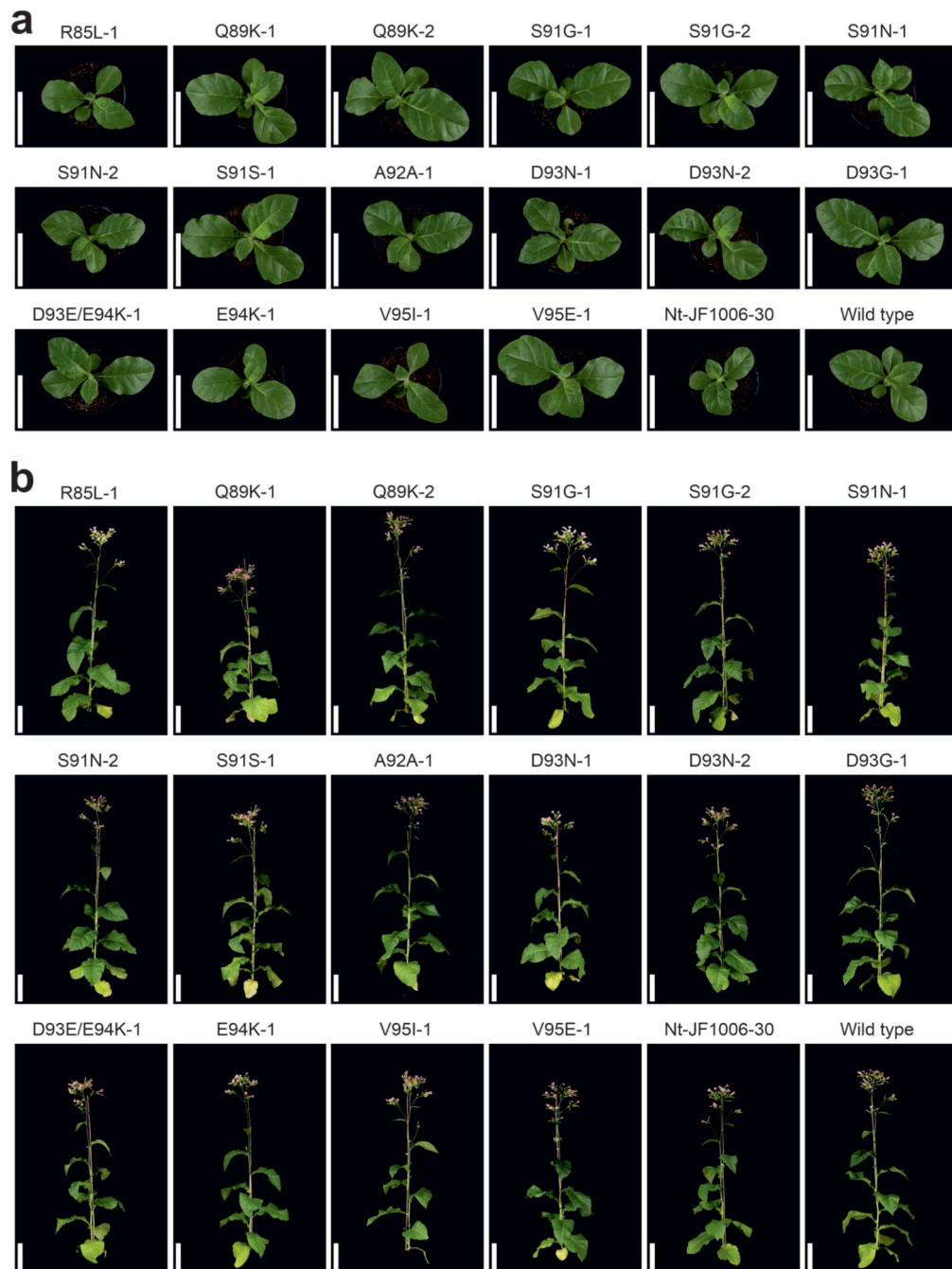
Extended Data Fig. 7. Phenotypes of the full set of *nad9* mutant plants.
(**a**) Photographs taken 33 days after sowing. Plants were raised from seeds, pregrown for 30 days in a nursery chamber, and then transferred to the greenhouse. Line names indicate the amino acid substitutions in Nad9. Scale bars: 10 cm. Nt-JF1006-30, TALEN control line with wild-type *nad9*. (**b**) The same plants as in (a) photographed 61 days after sowing. Scale bars: 20 cm. The images of E94K-1 and the wild type are as in Figure 4c,d.
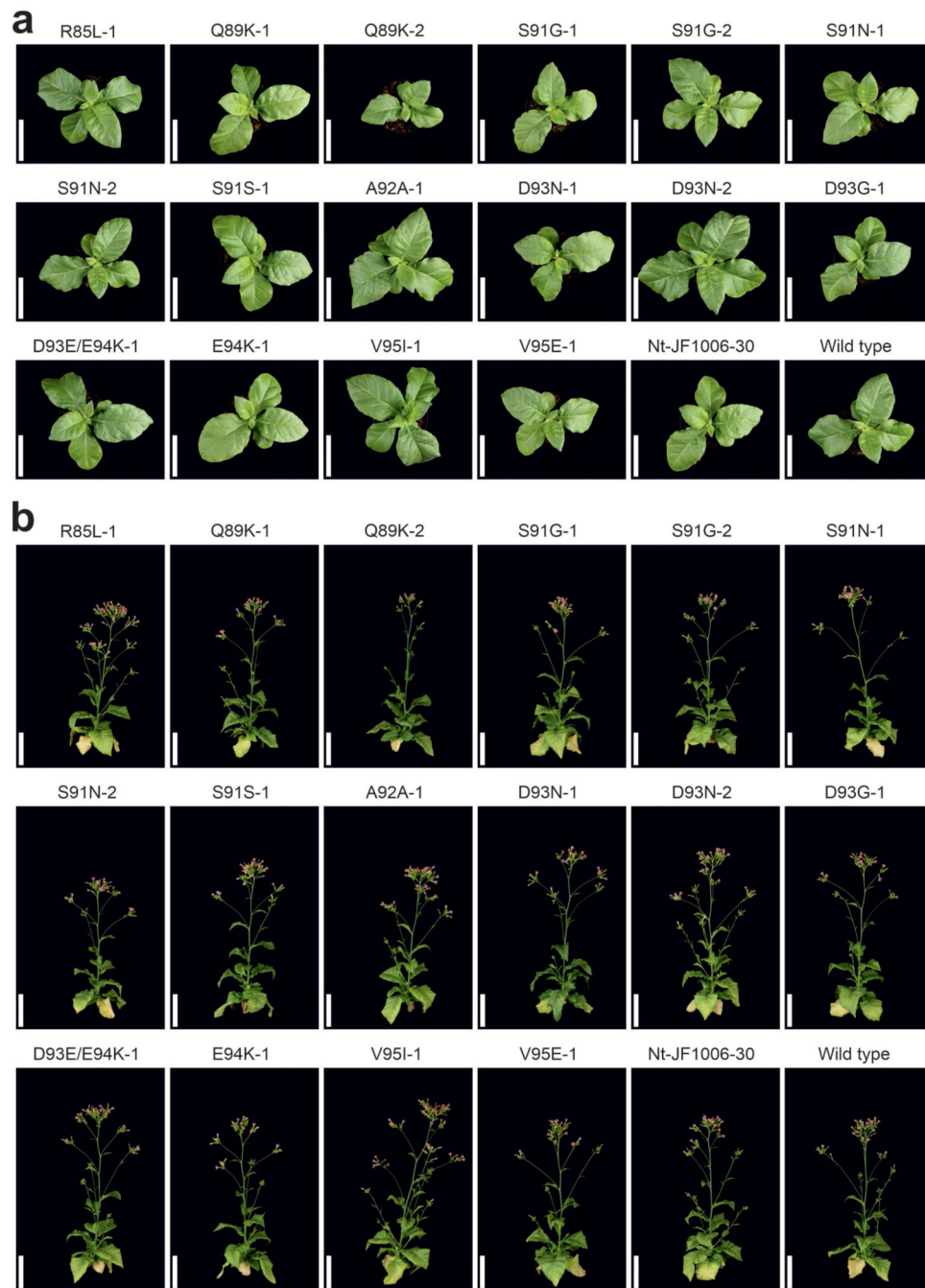
Extended Data Fig. 8. Phenotypes of the full set of *nad9* mutant lines under high-light conditions (cf. Extended Data Fig. 7).
(**a**) Photographs taken 40 days after sowing. Plants were raised from seeds, pre-grown for 28 days in a nursery chamber, and then transferred to a growth chamber with a light intensity of $1000 \ \mu E \ m^{-2} \ s^{-1}$. Scale bars: 10 cm. Nt-JF1006-30, TALEN control line with wild-type *nad9*. (**b**) The same plants as in (a) photographed 63 days after sowing. Scale bars: 20 cm.
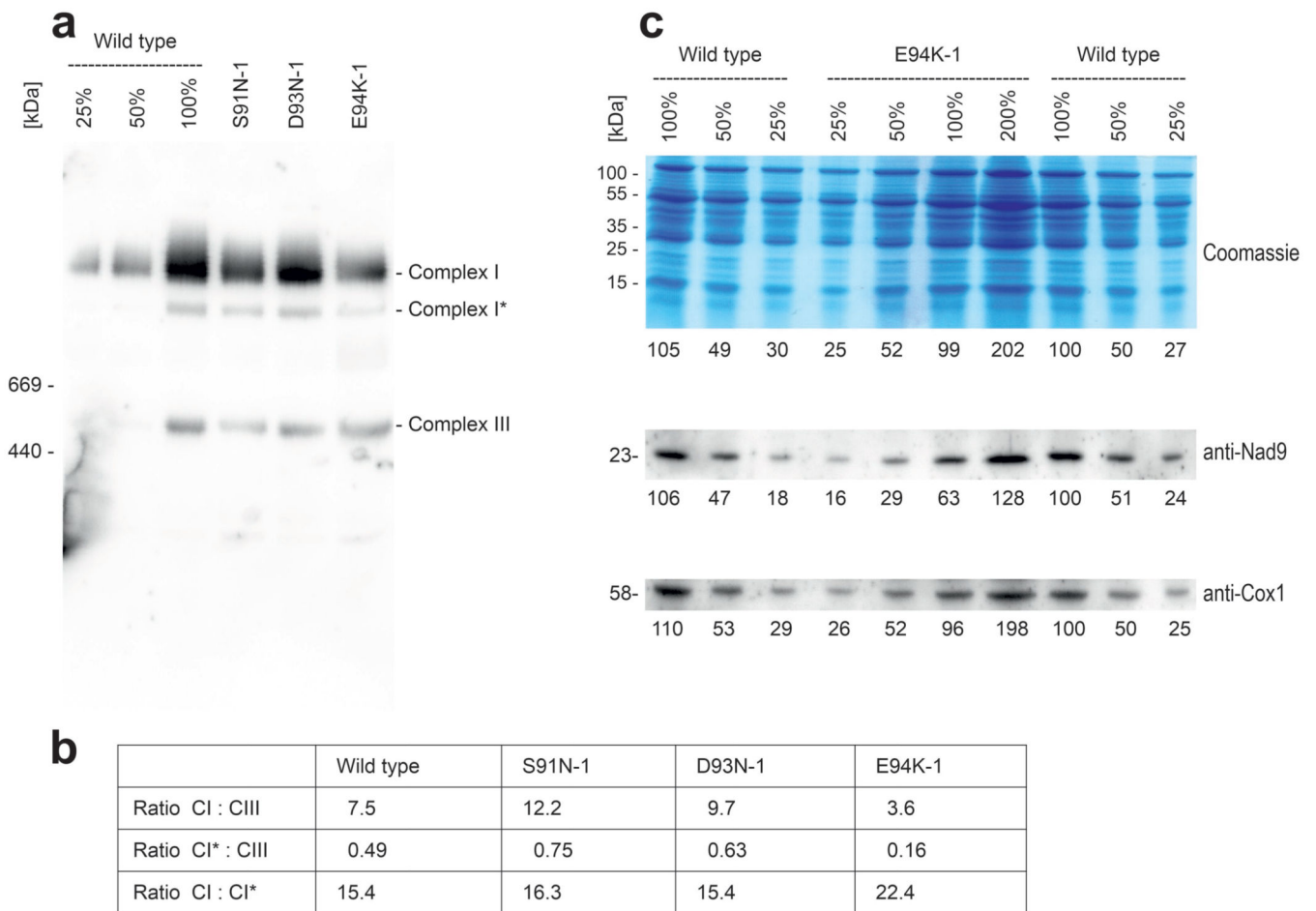
Extended Data Fig. 9. Characterization of complex I in a set of selected mitochondrial *nad9* mutants.
(**a**) Analysis of complex I accumulation by BN-PAGE and immunoblotting. A replicate gel of the gel shown in Fig. 4a was blotted and hybridized to an anti-CA2 antibody recognizing one of the carbonic anhydrase subunits of complex I. One of two technical replicates with similar results is shown. (**b**) Quantification of the ECL signal intensities in the immunoblot shown in panel (a). (**c**) Quantification of Nad9 protein amounts in the *nad9* E94K-1 mutant after SDS-PAGE analysis (in a 12% polyacrylamide gel) and immunoblotting. The percentages above the Coomassie-stained gel indicate the relative protein amounts loaded in each lane. The numbers below each panel represent the relative staining intensities (upper panel) and the intensities of the hybridization signals (middle and bottom panels). Quantification was done with the Fiji image processing package. The 100% wildtype sample in lane 8 served as reference. For information about the total number of SDS-PAGE and anti-Nad9 western blot experiments, see Fig. 4. The anti-Cox1 western blot was performed once. Quantifications were done based on the images shown here.

```
Yarrowia lipolytica    MLSRFARIGSMGIRPVAAARATFVTSARAAQAAPSWENIKDIRL------DPKV------   48
Nicotiana tabacum      ------------------------------------------------------------    0
Homo sapiens           -----------------------MAAAAVARLWWRGILGASALTRGTGRPSVLLLPVR   35
Escherichia coli       -----------------------MTDLTAQEPAWQTRDHLD----------------   18
Thermus thermophilus   ----------------------------------------MRL----------------    3


Yarrowia lipolytica    --HVDEVYEPIVNP-ADRYLQHVSDLHQYAKYIMAALPK-YIQGFSVWK-DELTLHVAPS  103
Nicotiana tabacum      ------------------M--D--NQFIFKYSWETLPKKWVKKMERSE-HGNRSDTNTD   36
Homo sapiens           RESAGADTRPTVRP-RNDVA--HKQLSAFGEYVAEILPK-YVQQVQVSCFNELEVCIHPD   91
Escherichia coli       --------DPVIGELRNRFGP----DAFTVQATRTGVPVVWIKR--------------E   51
Thermus thermophilus   --------ERVLEEA----RA----KGYPI--EDNGLGNLWVVLP------------RE   32
                                                                         :    ::                .


Yarrowia lipolytica    AVIPVTTFLRDNTSTQYKSIIDITAVDYPSR-------ENRFEVVYNFLSVRH-----NS  151
Nicotiana tabacum      YLFQLLCFLKLHTYTRVQVSIDICGVDYPSR-------KQRFEVVYNLLSIRY-----NS   84
Homo sapiens           GVIPVLTFLRDHTNAQFKSLVDLTAVDVPTR-------QNRFEIVYNLLSLRF-----NS  139
Escherichia coli       QLLEVGDFLKK-LPKPYVMLFDLHGMDERLRTHREGLPAADFSVFYHLISIDR-----NR  105
Thermus thermophilus   RFKEEMA--HY-KAMGFNFLADIVGLDYLTY---PDPRPERFAVVYELVSLPGWKDGDGS   86
                              .         :          *:  .:*        * :.*.::*:        .


Yarrowia lipolytica    RIRLKTYATEV-TPVPSITCLYEGANWFEREAYDMYGVFFEGHPDLRRIMTDYGFEGHPL  210
Nicotiana tabacum      RIRVQTSADEV-TRISPVVSLFPSAGRWEREVVWDMFGVSSINHPDLRRISTDYGFEGHPL  143
Homo sapiens           RIRVKTYTDEL-TPIESAVSVFKAANWYEREIWDMFGVFFANHPDLRRILTDYGFEGHPF  198
Escherichia coli       DIMLKVALAENDLHVPTFTKLFPNANWYERETWDLFGITFDGHPNLRRIMMPQTWKGHPL  165
Thermus thermophilus   RFFVRVYVPEEDPRLPTVTDLWGSANFLEREVYDLFGIVFEGHPDLRKILTPEDLEGHPL  146
                         : ::.    *       :      .  ::   *.   *** :*::*:     .**:**:*         :***:


Yarrowia lipolytica    RKDFPLTGYTEVRWDEEKRRVVYEPLELTQAFRNFS-----AGSTAWEPVGPGRDDRPDS  265
Nicotiana tabacum      RKDLPLSGYVEVRYDDPEKRVVSEPIEMTQEFRYFD-----FAS-PWEQRSDG-------  190
Homo sapiens           RKDFPLSGYVELRYDDEVKRVVAEPVELAQEFRKFD-----LNS-PWEAFPVYRQ-PPES  251
Escherichia coli       RKDYPARATEFSPF-----------ELTKAKQDLEMEALTFKPEEWGMKRGTEN--EDF  211
Thermus thermophilus   RKDYPLGETPTLFR-----------E-----------GRYIIPAEFRAALTGKD--PGL  181
                       *** *                        *                           :


Yarrowia lipolytica    FKLPTP--KPEEK-------------------------------EGDKK-----------  281
Nicotiana tabacum      ------------------------------------------------------------  190
Homo sapiens           LKLEAGDKKPDAK-----------------------------------------------  264
Escherichia coli       MFLNLGPNHPSAHGAFRIVLQLDGEEIVDCVPDIGYHHRGAEKMGERQSWHSYIPYTDRI  271
Thermus thermophilus   TFYKGG--------------------------------------SRKGYRSLWADLKKARE--  204
```

Extended Data Fig. 10. Amino acid sequence alignment of the Nad9 protein from tobacco with homologues from different species of prokaryotes and eukaryotes.
The highly conserved glutamate residue 94 is highlighted in red. Asterisks (*) indicate perfect conservation, colons (:) denote residues belonging to a group of amino acids exhibiting strong similarity, and dots (.) mark residues belonging to a group of amino acids exhibiting weak similarity. For the tobacco Nad9 protein, the (unedited) genomic sequence was used.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Data availability

The data supporting the findings of this study are available within the paper and its supplementary information files. The NGS datasets were deposited under BioProject ID PRJNA787054 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA787054). Sequences of plasmids pJF1004, pJF1005 and pJF1006 are accessible via GenBank accession numbers MZ417370, MZ417371 and MZ417372. GenBank accession number NC_006581.1 was used as tobacco mitochondrial reference genome.

Code availability

The Perl code (extract_BaseCovInformation.pl) employed is available under the github repository https://github.com/AxelMacFoly/ngs_perl_scripts/.

References

1. Maliga P. Plastid transformation in higher plants. Annu Rev Plant Biol. 2004; 55: 289–313. [PubMed: 15377222]

2. Bock R. Engineering plastid genomes: Methods, tools, and applications in basic research and biotechnology. Annu Rev Plant Biol. 2015; 66: 211–241. [PubMed: 25494465]

3. Ruf S, Kössel H, Bock R. Targeted inactivation of a tobacco intron-containing open reading frame reveals a novel chloroplast-encoded photosystem I-related gene. J Cell Biol. 1997; 139: 95–102. [PubMed: 9314531]

4. Allison LA, Maliga P. Light-responsive and transcription-enhancing elements regulate the plastid psbD core promoter. EMBO J. 1995; 14: 3721–3730. [PubMed: 7641691]

5. Bock R, Hermann M, Kössel H. In vivo dissection of cis-acting determinants for plastid RNA editing. EMBO J. 1996; 15: 5052–5059. [PubMed: 8890178]

6. Staub JM, Maliga P. Translation of the psbA mRNA is regulated by light via the 5'-untranslated region in tobacco plastids. Plant J. 1994; 6: 547–553. [PubMed: 7987413]

7. Lu Y, Stegemann S, Agrawal S, Karcher D, Ruf S, Bock R. Horizontal transfer of a synthetic metabolic pathway between plant species. Curr Biol. 2017; 27: 3034–3041. [PubMed: 28943084]

8. Boehm CR, Bock R. Recent advances and current challenges in synthetic biology of the plastid genetic system and metabolism. Plant Physiol. 2019; 179: 794–802. [PubMed: 30181342]

9. Altpeter F, Baisakh N, Beachy R, Bock R, Capell T, Christou P, Daniell H, Datta K, Datta S, Dix PJ, Fauquet C, et al. Particle bombardment and the genetic enhancement of crops: myths and realities. Mol Breeding. 2005; 15: 305–327.

10. Li W, Ruf S, Bock R. Chloramphenicol acetyltransferase as selectable marker for plastid transformation. Plant Mol Biol. 2011; 76: 443–451. [PubMed: 20721602]

11. Pring DR, Mullen JA, Kempken F. Conserved sequence blocks 5' to start codons of plant mitochondrial genes. Plant Mol Biol. 1992; 19: 313–317. [PubMed: 1320424]

12. Scharff LB, Childs L, Walther D, Bock R. Local absence of secondary structure permits translation of mRNAs that lack ribosome-binding sites. PLoS Genet. 2011; 7 e1002155 [PubMed: 21731509]

13. Preuten T, Cincu E, Fuchs J, Zoschke R, Liere K, Börner T. Fewer genes than organelles: extremely low and variable gene copy numbers in mitochondria of somatic plant cells. Plant J. 2010; 64: 948–959. [PubMed: 21143676]

14. Manchekar M, Scissum-Gunn K, Song D, Khazi F, McLean SL, Nielsen BL. DNA recombination activity in soybean mitochondria. J Mol Biol. 2006; 356: 288–299. [PubMed: 16376379]

15. Shedge V, Arrieta-Montiel M, Christensen AC, Mackenzie SA. Plant mitochondrial recombination surveillance requires unusual RecA and MutS homologs. Plant Cell. 2007; 19: 1251–1264. [PubMed: 17468263]

16. Miller-Messmer M, Kühn K, Bichara M, Le Ret M, Imbault P, Gualberto JM. RecA-dependent DNA repair results in increased heteroplasmy of the Arabidopsis mitochondrial genome. Plant Physiol. 2012; 159: 211–226. [PubMed: 22415515]

17. Bogdanove AJ, Voytas DF. TAL effectors: customizable proteins for DNA targeting. Science. 2011; 333: 1843–1846. [PubMed: 21960622]

18. Kanchiswamy CN, Maffei M, Malnoy M, Velasco R, Kim J-S. Fine-tuning next-generation genome editing tools. Trends Biotechnol. 2016; 34: 562–574. [PubMed: 27167723]

19. Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotechnol. 2014; 32: 347–355. [PubMed: 24584096]

20. Nishida K, Arazoe T, Yachie N, Banno S, Kakimoto M, Tabata M, Mochizuki M, Miyabe A, Araki M, Hara KY, Shimatani Z, et al. Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. Science. 2016; 353 1248

21. Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature. 2016; 533: 420–424. [PubMed: 27096365]

22. Gaudelli NM, Komor AC, Rees HA, Packer MS, Badran AH, Bryson DI, Liu DR. Programmable base editing of A·T to G·C in genomic DNA without DNA cleavage. Nature. 2017; 551: 464–471. [PubMed: 29160308]

23. Tan J, Zhang F, Karcher D, Bock R. Engineering of high-precision base editors for site-specific single nucleotide replacement. Nat Commun. 2019; 10: 439. [PubMed: 30683865]

24. Schneider A. Import of RNA into mitochondria. Trends Cell Biol. 1994; 4: 282–286. [PubMed: 14731591]

25. Duchêne A-M, Pujol C, Maréchal-Drouard L. Import of tRNAs and aminoacyl-tRNA synthetases into mitochondria. Curr Genet. 2009; 55: 1–18. [PubMed: 19083240]

26. Val R, Wyszko E, Valentin C, Szymanski M, Cosset A, Alioua M, Dreher TW, Barciszewski J, Dietrich A. Organelle trafficking of chimeric ribozymes and genetic manipulation of mitochondria. Nucleic Acids Res. 2011; 39: 9262–9274. [PubMed: 21768127]

27. Johnston SA, Anziano PQ, Shark K, Sanford JC, Butow RA. Mitochondrial transformation in yeast by bombardment with microprojectiles. Science. 1988; 240: 1538–1541. [PubMed: 2836954]

28. Randolph-Anderson BL, Boynton JE, Gillham NW, Harris EH, Johnson AM, Dorthu M-P, Matagne RF. Further characterization of the respiratory deficient dum-1 mutation of Chlamydomonas reinhardtii and its use as a recipient for mitochondrial transformation. Mol Gen Genet. 1993; 236: 235–244. [PubMed: 8437570]

29. Kazama T, Okuno M, Watari Y, Yanase S, Koizuka C, Tsuruta Y, Sugaya H, Toyoda A, Itoh T, Tsutsumi N, Toriyama K, et al. Curing cytoplasmic male sterility via TALEN-mediated mitochondrial genome editing. Nat Plants. 2019; 5: 722–730. [PubMed: 31285556]

30. Mok BY, de Moraes MH, Zeng J, Bosch DE, Kotrys AV, Raguram A, Hsu F, Radey MC, Peterson SB, Mootha VK, Mougous JD, et al. A bacterial cytidine deaminase toxin enables CRISPR-free mitochondrial base editing. Nature. 2020; 583: 631–637. [PubMed: 32641830]

31. Kang B-C, Bae S-J, Lee S, Lee JS, Kim A, Lee H, Baek G, Seo H, Kim J, Kim J-S. Chloroplast and mitochondrial DNA editing in plants. Nat Plants. 2021; 7: 899–905. [PubMed: 34211132]

32. Nakazato I, Okuno M, Yamamoto H, Tamura Y, Itoh T, Shikanai T, Takanashi H, Tsutsumi N, Arimura S-i. Targeted base editing in the plastid genome of Arabidopsis thaliana. Nat Plants. 2021; 7: 906–913. [PubMed: 34211131]

33. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A, Bonas U. Breaking the code of DNA binding specificity of TAL-type III effectors. Science. 2009; 326: 1509–1512. [PubMed: 19933107]

34. Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, Baller JA, Somia NV, Bogdanove AJ, Voytas DF. Efficient design and assembly of custom TALEN and other TAL

effector-based constructs for DNA targeting. Nucleic Acids Res. 2011; 39: e82. [PubMed: 21493687]

35. Li T, Huang S, Zhao X, Wright DA, Carpenter S, Spalding MH, Weeks DP, Yang B. Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes. Nucleic Acids Res. 2011; 39: 6315–6325. [PubMed: 21459844]

36. Kühn K, Obata T, Feher K, Bock R, Fernie AR, Meyer EH. Complete mitochondrial complex I deficiency induces an up-regulation of respiratory fluxes that is abolished by traces of functional complex I. Plant Physiol. 2015; 168: 1537–1549. [PubMed: 26134164]

37. Ligas J, Pineau E, Bock R, Huynen MA, Meyer EH. The assembly pathway of complex I in Arabidopsis thaliana. Plant J. 2019; 97: 447–459. [PubMed: 30347487]

38. Aouida M, Piatek MJ, Bangarusamy DK, Mahfouz MM. Activities and specificities of homodimeric TALENs in Saccharomyces cerevisiae. Curr Genet. 2014; 60: 61–74. [PubMed: 24081604]

39. Sanders KL, Catto LE, Bellamy SRW, Halford SE. Targeting individual subunits of the FokI restriction endonuclease to specific DNA strands. Nucleic Acids Res. 2009; 37: 2105–2115. [PubMed: 19223323]

40. Kohl S, Bock R. Transposition of a bacterial insertion sequence in chloroplasts. Plant J. 2009; 58: 423–436. [PubMed: 19144000]

41. Hagemann R. Milestones in plastid genetics of higher plants. Prog Bot. 2002; 63: 1–51.

42. Gillham NW, Boynton JE, Harris EH. Specific elimination of mitochondrial DNA from Chlamydomonas by intercalating dyes. Curr Genet. 1987; 12: 41–47. [PubMed: 3449222]

43. Timmis JN, Scott NS. Spinach nuclear and chloroplast DNAs have homologous sequences. Nature. 1983; 305: 65–67.

44. Huang CY, Ayliffe MA, Timmis JN. Direct measurement of the transfer rate of chloroplast DNA into the nucleus. Nature. 2003; 422: 72–76. [PubMed: 12594458]

45. Stegemann S, Hartmann S, Ruf S, Bock R. High-frequency gene transfer from the chloroplast genome to the nucleus. Proc Natl Acad Sci USA. 2003; 100: 8828–8833. [PubMed: 12817081]

46. Stegemann S, Bock R. Experimental reconstruction of functional gene transfer from the tobacco plastid genome to the nucleus. Plant Cell. 2006; 18: 2869–2878. [PubMed: 17085684]

47. Hager M, Biehler K, Illerhaus J, Ruf S, Bock R. Targeted inactivation of the smallest plastid genome-encoded open reading frame reveals a novel and essential subunit of the cytochrome b6f complex. EMBO J. 1999; 18: 5834–5842. [PubMed: 10545095]

48. Ruf S, Biehler K, Bock R. A small chloroplast-encoded protein as a novel architectural component of the light-harvesting antenna. J Cell Biol. 2000; 149: 369–377. [PubMed: 10769029]

49. Sun C-W, Callis J. Recent stable insertion of mitochondrial DNA into an Arabidopsis polyubiquitin gene by nonhomologous recombination. Plant Cell. 1993; 5: 97–107. [PubMed: 8382544]

50. Timmis JN, Ayliffe MA, Huang CY, Martin W. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nature Rev Genet. 2004; 5: 123–135. [PubMed: 14735123]

51. Bock R. Witnessing genome evolution: experimental reconstruction of endosymbiotic and horizontal gene transfer. Annu Rev Genet. 2017; 51: 1–22. [PubMed: 28846455]

52. Greiner S, Sobanski J, Bock R. Why are most organelle genomes transmitted maternally? BioEssays. 2015; 37: 80–94. [PubMed: 25302405]

53. Dutilleul C, Driscoll S, Cornic G, De Paepe R, Foyer CH, Noctor G. Functional mitochondrial complex I is required by tobacco leaves for optimal photosynthetic performance in photorespiratory conditions and during transients. Plant Physiol. 2003; 131: 264–275. [PubMed: 12529534]

54. Bhardwaj A, Nain V. TALENs - an indispensable tool in the era of CRISPR: a mini review. J Genet Eng Biotechnol. 2021; 19: 125. [PubMed: 34420096]

55. Arimura, S-i; Ayabe, H; Sugaya, H; Okuno, M; Tamura, Y; Tsuruta, Y; Watari, Y; Yanase, S; Yamauchi, T; Itoh, T; Toyoda, A; , et al. Targeted gene disruption of ATP synthases 6-1 and 6-2 in the mitochondrial genome of Arabidopsis thaliana by mito TALENs. Plant J. 2020; 104: 1459–1471. [PubMed: 33098708]

56. Tan J, Zhang F, Karcher D, Bock R. Expanding the genome-targeting scope and the site selectivity of high-precision base editors. Nat Commun. 2020; 11

57. Chase CD. Cytoplasmic male sterility: a window to the world of plant mitochondrial-nuclear interactions. Trends Genet. 2006; 23: 81–90. [PubMed: 17188396]

58. Caruso CM, Case AL, Bailey MF. The evolutionary ecology of cytonuclear interactions in angiosperms. Trends Plant Sci. 2012; 17: 638–643. [PubMed: 22784826]

59. Greiner S, Bock R. Tuning a ménage à trois: Co-evolution and co-adaptation of nuclear and organellar genomes in plants. BioEssays. 2013; 35: 354–365. [PubMed: 23361615]

60. Weber E, Engler C, Gruetzner R, Werner S, Marillonnet S. A modular cloning system for standardized assembly of multigene constructs. PLoS One. 2011; 6 e16765 [PubMed: 21364738]

61. Däschner K, Couée I, Binder S. The mitochondrial isovaleryl-coenzyme a dehydrogenase of Arabidopsis oxidizes intermediates of leucine and valine catabolism. Plant Physiol. 2001; 126: 601–612. [PubMed: 11402190]

62. Mussolino C, Morbitzer R, Lütge F, Dannemann N, Lahaye T, Cathomen T. A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. Nucleic Acids Res. 2011; 39: 9283–9393. [PubMed: 21813459]

63. Morbitzer R, Elsaesser J, Hausner J, Lahaye T. Assembly of custom TALEtype DNA binding domains by modular cloning. Nucleic Acids Res. 2011; 39: 5790–5799. [PubMed: 21421566]

64. Murashige T, Skoog F. A revised medium for rapid growth and bio assays with tobacco tissue culture. Physiol Plant. 1962; 15: 473–497.

65. Doyle JJ, Doyle JL. Isolation of plant DNA from fresh tissue. Focus. 1990; 12: 13–15.

66. Lamattina L, Gonzalez D, Gualberto J, Grienenberger J-M. Higher plant mitochondria encode an homologue of the nuclear-encoded 30-kDa subunit of bovine mitochondrial complex I. Eur J Biochem. 1993; 217: 831–838. [PubMed: 8223639]

67. Meyer EH, Lehmann C, Boivin S, Brings L, De Cauwer I, Bock R, Kühn K, Touzet P. CMS-G from Beta vulgaris ssp. maritima is maintained in natural populations despite containing an atypical cytochrome c oxidase. Biochem J. 2018; 475: 759–773. [PubMed: 29358189]

68. Perales M, Eubel H, Heinemeyer J, Colaneri A, Zabaleta E, Braun H-P. Disruption of a nuclear gene encoding a mitochondrial gamma carbonic anhydrase reduces complex I and supercomplex I + III2 levels and alters mitochondrial physiology in Arabidopsis. J Mol Biol. 2005; 350: 263–277. [PubMed: 15935378]

69. Krause GH, Weis E. Chlorophyll fluorescence and photosynthesis: the basics. Annu Rev Plant Physiol Plant Mol Biol. 1991; 42: 313–349.

70. Kramer DM, Johnson G, Kiirats O, Edwards GE. New fluorescence parameters for the determination of Q_A redox state and excitation energy fluxes. Photosynth Res. 2004; 79: 209–218. [PubMed: 16228395]

71. Klughammer C, Schreiber U. Analysis of photosystem I donor and acceptor sides with a new type of online-deconvoluting kinetic LED array spectrophotometer. Plant Cell Physiol. 2016; 57: 1454–1467. [PubMed: 27053032]

72. Porra RJ, Thompson WA, Kriedemann PE. Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b extracted with four different solvents: Verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. Biochim Biophys Acta. 1989; 975: 384–394.

73. Zhang J, Kobert K, Flouri T, Stamatakis A. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. Bioinformatics. 2014; 30: 614–620. [PubMed: 24142950]

74. Roehr JT, Dieterich C, Reinert K. Flexbar 3.0 - SIMD and multicore parallelization. Bioinformatics. 2017; 33: 2941–2942. [PubMed: 28541403]

75. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv, arXiv. 2013. 1303.3997.v2

76. 1000 Genome Project Data Processing Subgroup. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009; 25: 2078–2079. [PubMed: 19505943]

77. Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing. arXiv, arXiv. 2012. 1207.3907v2

78. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly. 2012; 6: 80–92. [PubMed: 22728672]
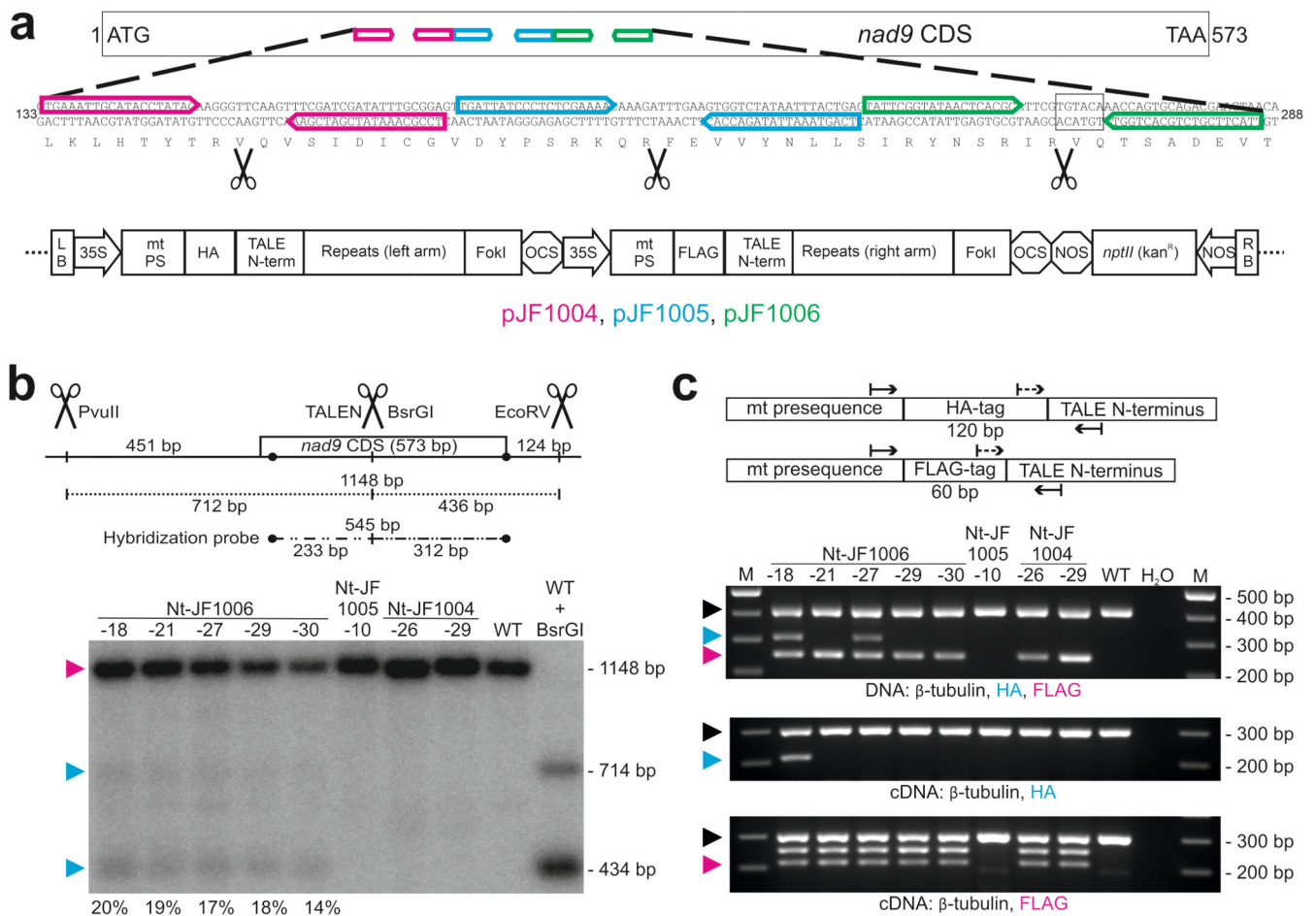
Fig. 1. Generation of transgenic tobacco lines expressing TALENs targeted to the mitochondrial *nad9* locus.

(**a**) Binding sites of the three designed TALEN pairs in the *nad9* CDS (coding sequence), and schematic view of the binary vectors used for plant transformation. Nucleotide positions refer to the start codon (<u>A</u>TG=1). Scissors indicate the predicted TALEN cut sites, the BsrGI restriction site is boxed. LB/RB, T-DNA left/right borders; 35S, CaMV 35S promoter; mt PS, mitochondrial presequence (from the *Arabidopsis* IVD protein); HA, 3xHA-tag; TALE N-term, TALE N-terminus; Repeats, TALE DNA-binding units; FokI, endonuclease domain; OCS, octopine synthase terminator (*Agrobacterium*); FLAG, 3xFLAG-tag; NOS (octagon), nopaline synthase terminator (*Agrobacterium*); *nptII* (kan^R), neomycin phosphotransferase II encoding kanamycin resistance; NOS (arrow), nopaline synthase promoter (*Agrobacterium*); arrows, TALEN binding sites. (**b**) Detection of TALEN activity by Southern blot analysis. Predicted sizes of hybridizing restriction fragments and the location of the probe are depicted. Scissors mark the cut sites of the restriction enzymes PvuII, EcoRV and BsrGI, and of the pJF1006-encoded TALEN pair. BsrGI cuts 2 bp away from the predicted TALEN cut site. In addition to PvuII/EcoRV, BsrGI was included in a digestion reaction with wild-type DNA (WT + BsrGI; right) to simulate the TALEN cut. The numbers below the gel give the relative intensities of the TALEN cleavage products). Arrowheads mark the positions of the three expected signals (magenta: uncut,

cyan: TALEN-cut). (**c**) PCR and RT-PCR assays to test for the presence of TALEN arms in the lines shown in panel (b). The binding sites of the PCR primers are shown in the upper panel (dashed lines: RT-PCR primers), the lower panels display the PCR results. A sequence from the β-TUBULIN gene was amplified as internal control. Arrowheads indicate the expected sizes for β-TUBULIN (black, 412 bp/303 bp for DNA/cDNA), the HA arm (cyan, 311 bp/210 bp) and the FLAG arm (magenta, 251 bp/212 bp), respectively. H_2O, water control; M, DNA size marker. Only the FLAG-TALEN arm is present in Nt-JF1006-30. Pools of kanamycin-resistant T_1 seedlings obtained by selfing were used for nucleic acid extractions. See also Extended Data Fig. 4. The experiments in (b) and (c) were performed once.
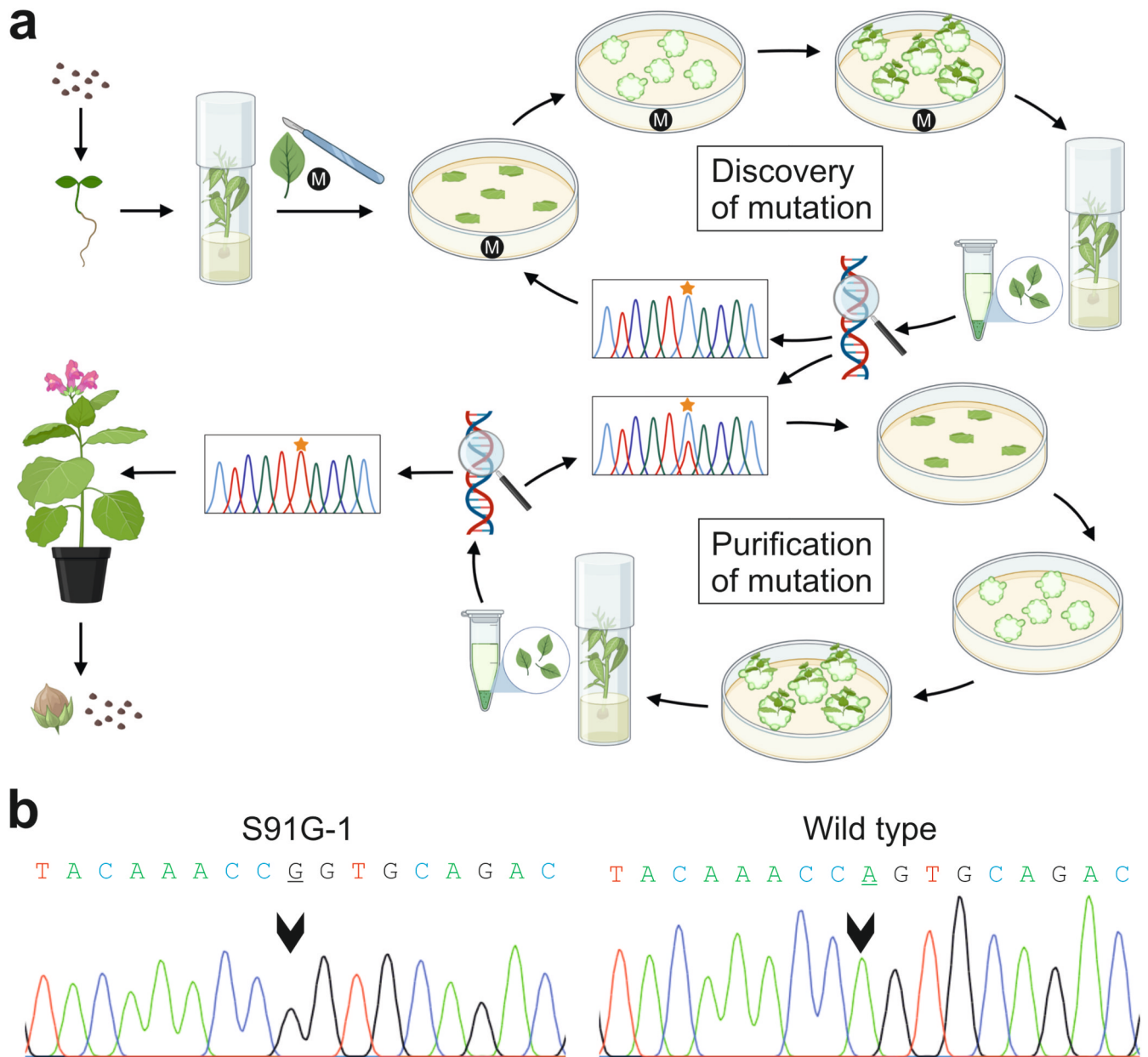
Fig. 2. Workflow for the isolation of tobacco lines with TALEN-induced point mutations in the mitochondrial *nad9* gene.

(**a**) Schematic overview of the experimental procedures. Transgenic plants with confirmed TALEN activity (cf. Extended Data Fig. 1) were raised from seeds, leaves were harvested, cut into pieces and placed onto shoot induction medium. Leaves from regenerated shoots were genotyped for mutations in *nad9* ('Discovery of mutation'). If no mutation was detected, a new regeneration round was initiated. When a point mutation was found, an additional regeneration round was conducted to promote genome segregation and facilitate isolation of homochondriomic lines ('Purification of mutation'). Homochondriomic mutant shoots were rooted and transferred to the greenhouse for seed production. In some experiments, the mutagens (M) ethidium bromide or N-nitroso-N-ethylurea were added

to the shoot induction medium or applied during seed imbibition. See text, Methods and Supplementary Methods for details. (**b**) Exemplary sequencing chromatograms showing the successful isolation of a homochondriomic *nad9* mutant carrying a Ser-to-Gly exchange in amino acid position 91 of the Nad9 protein (corresponding to an A-to-G substitution in nucleotide position 271 of the *nad9* reading frame). The mutated position is indicated by arrowheads in the sequencing chromatograms of the S91G-1 mutant line and the wild type. The sequencing primer was oJF271 (Supplementary Table 2).
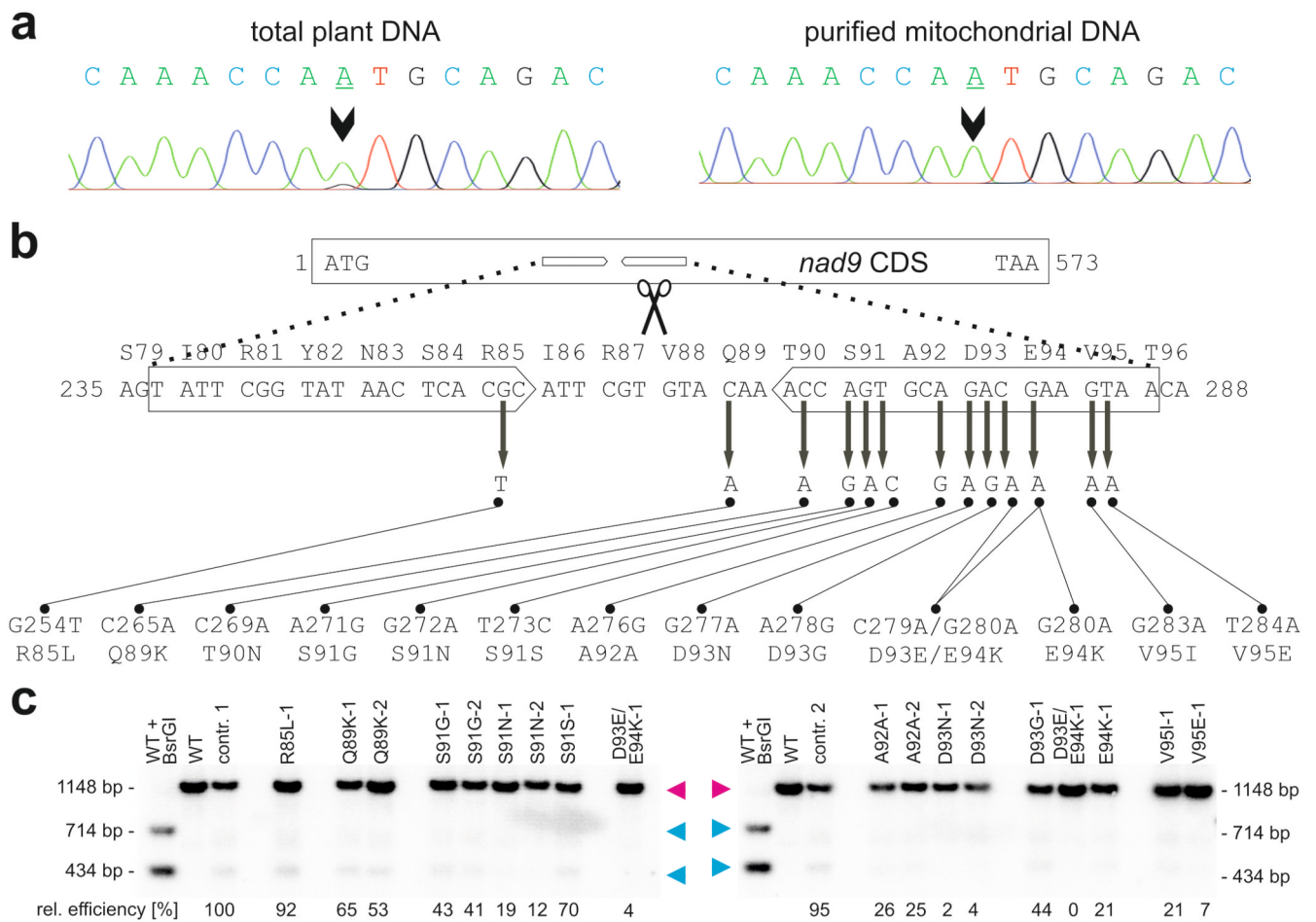
Fig. 3. Apparent heterochondriomy, location of the TALEN-induced point mutations in *nad9*, and effects of the point mutations on TALEN cleavage activity.
(**a**) Genotyping of an S91N-1 mutant plant carrying the G272A mutation. Sequencing of PCR products amplified from total plant DNA (with primers oJF271 and oJF272; Supplementary Table 2) suggested persistent heteroplasmy (i.e., presence of residual copies of wild-type *nad9* alleles), as evidenced by an A+G double peak (arrowhead, left sequence). However, when the same PCR product was generated from purified mtDNA as template, a clean single A peak is seen (right sequence). Sequencing was done with primer oJF271. (**b**) Overview of all TALEN-induced mutations obtained in *nad9*. Nucleotide positions in the *nad9* coding sequence (<u>A</u>TG = 1) and amino acid positions in the Nad9 protein are given. Block arrows denote the pJF1006 TALEN binding sites. The 'left' (HA-)TALEN (not present in line Nt-JF1006-30) extends from nucleotide position 237 to position 255, the 'right' (FLAG-)TALEN from position 268 to 286. Scissors point to the predicted TALEN cut site. (**c**) Analysis of TALEN cleavage activity by Southern blotting using pools of kanamycin-resistant seedlings (from back-crosses of the original mutants with the wild type). The same restriction enzymes and hybridization probe as in Fig. 1b were used, but electrophoretic separation was done in a 2% agarose gel. To determine relative cutting efficiencies, the percentage of cut *nad9* was calculated for each lane and then divided by the respective value for control line 1 (contr. 1). The smaller of the two *nad9* cleavage products

was used for quantification, because it is covered by a larger portion of the hybridization probe. Contr. 1 and contr. 2, descendants of Nt-JF1006 plants having gone through the same mutagenesis procedure as the point mutants, but retaining a wild-type *nad9* sequence. Line D93E/E94K-1 is represented twice (descendants of two different vegetative clones of the original mutant plant). The full Southern blot with enhanced contrast settings is presented in Extended Data Fig. 5. In contrast to all other mutants, line R85L-1 may not be the result of a gene drive effect. A technical replicate of the Southern blot yielded similar results.
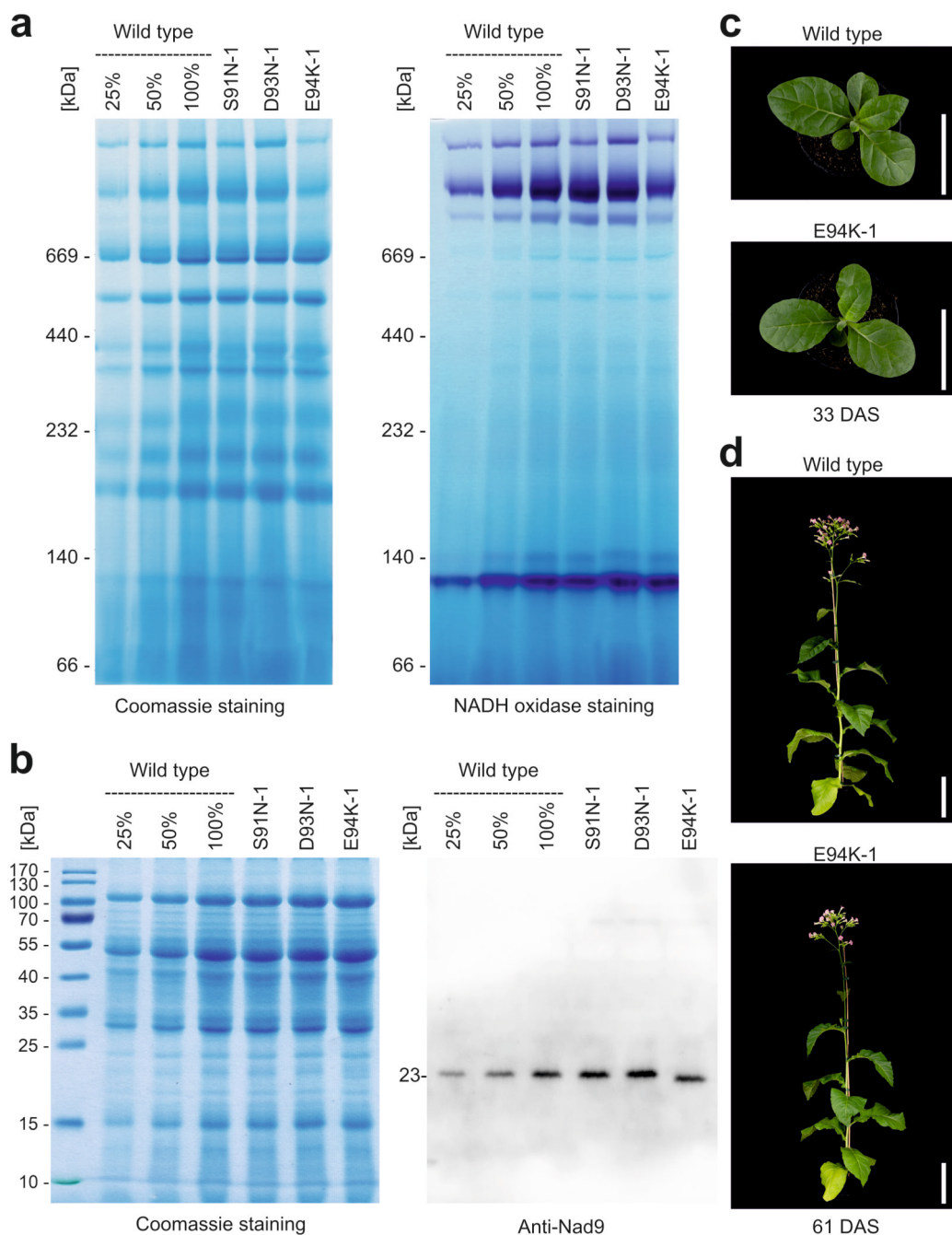
Fig. 4. Biochemical characterization and phenotype of mitochondrial *nad9* mutants.
(**a**) Analysis of mitochondrial protein complexes and NADH oxidase (complex I) staining
in selected *nad9* mutants by blue-native polyacrylamide gel electrophoresis (BN-PAGE).
Protein sizes are given in kDa. A dilution series of the wild-type sample (25%, 50%
and 100%) was loaded to allow for semiquantitative assessment of protein accumulation.
A technical replicate of the BN-PAGE (stained with Coomassie) yielded similar results;
the NADH oxidase (complex I) staining was performed once. (**b**) Accumulation and
electrophoretic mobility of Nad9 proteins in *nad9* mutants assessed by SDS-PAGE. Note

faster migration of the E94K variant of Nad9. The Coomassie-stained gel is shown to confirm equal loading. The SDS-PAGE (including Coomassie staining) was done seven times and the anti-Nad9 western blot five times (technical replicates) with similar results (see also Extended Data Fig. 9). (**c**) Plant phenotypes. Plants were grown on soil and cultivated in long-day conditions in the greenhouse. A wild-type plant is shown in comparison to an E94K-1 mutant plant (harboring a G280A mutation in *nad9*). Photographs were taken 33 days after sowing (DAS). Scale bars: 10 cm. (**d**) Plants photographed 61 DAS. Scale bars: 20 cm.
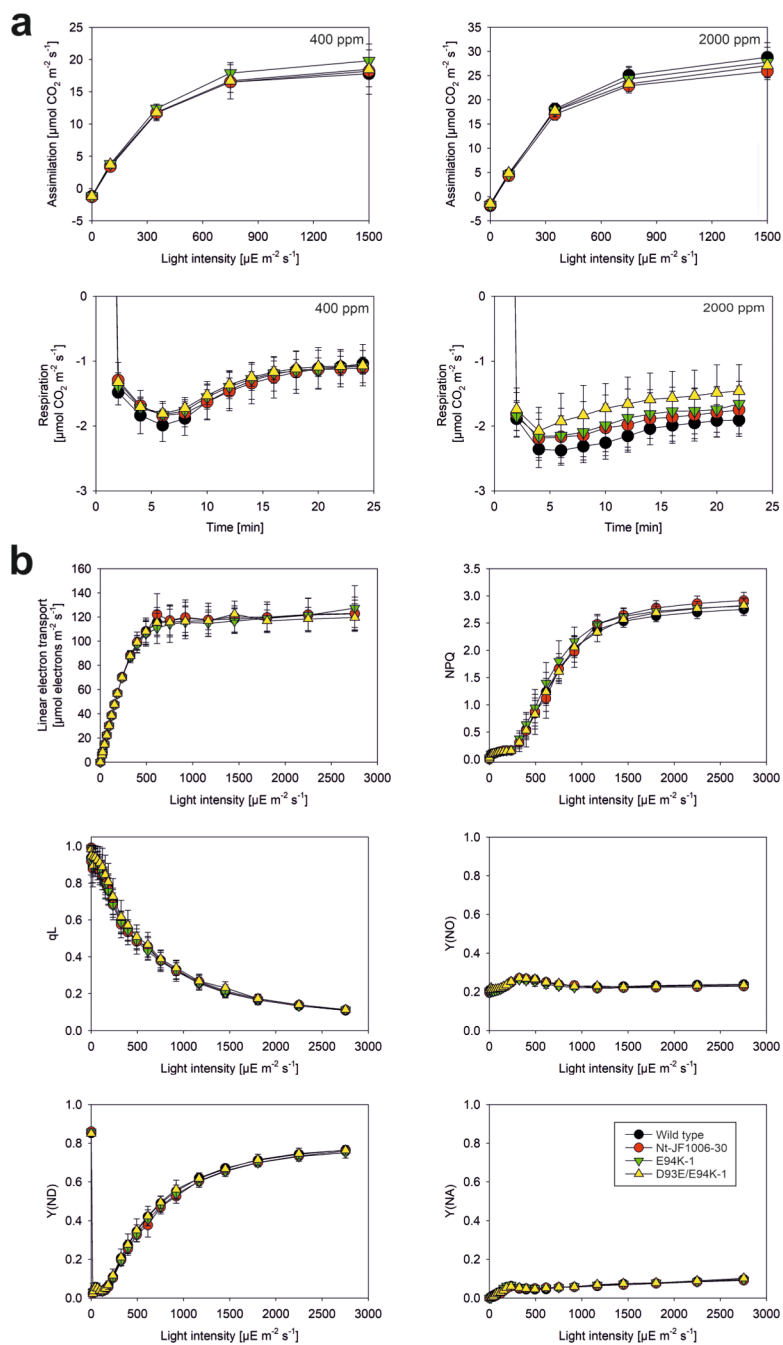
Fig. 5. Determination of gas exchange rates and photosynthetic parameters of two mitochondrial mutants grown at 350 μE m^{-2} s^{-1} light intensity in a controlled environment chamber. The youngest fully expanded leaves were analyzed.
(**a**) Light response curves of assimilation of plants measured at 400 ppm and 2000 ppm CO_2 concentration, and stimulated respiration during the first minutes after the end of illumination with saturating actinic light. ppm, CO_2 concentration in parts per million. (**b**) Photosynthetic parameters, only measured at 400 ppm CO_2 concentration. Nt-JF1006-30 served as TALEN-control harboring the unaltered *nad9* allele. ppm, CO_2 concentration in parts per million; NPQ, non-photochemical quenching (69); qL, measure for the redox state

of the photosystem II acceptor side (70); Y(NO), measure for the non-regulated thermal dissipation of excitation energy, an indicator of PSII photoinhibition (if increased); Y(ND), measure of the donor side limitation of photosystem I by linear electron transport; Y(NA); measure of the acceptor side limitation of photosystem I by the Calvin-Benson cycle and other downstream metabolic reactions [71] . For the wild type and the TALEN-control line Nt-JF1006-30, always n=6 biological replicates (independent plants) were measured. For the single mutant E94K-1 and the double mutant D93E/E94K-1, n=5 independent plants (biological replicates) were measured, except for the assimilation measurements at 2000 ppm CO_2 for the single mutant E94K-1, where n=4 biological replicates were analyzed. Average values and standard deviations are shown. Where error bars are not visible, the standard deviation was smaller than the size of the symbol.

Table 1

Effects of light conditions and mutagen application on recovery of mitochondrial mutants.

	16 h light / 8 h dark			24 h light / 0 h dark		
Treatment	none	EtBr	NEU	none	EtBr	NEU
st						
1st reg. round	0/10	2^a/10	0/10	0/10	3^b/10	0/10
2nd reg. round	0/10	0/10	0/10	0/10	0/10	0/10
3rd reg. round	0/22	0/16	1^c/19	0/24	1^d/38	0/20
Total	0/42	2/36	1/36	0/44	4/58	0/40

The numbers give the number of mutants obtained per number of regenerated shoots assayed. EtBr, ethidium bromide [0.001%]; NEU, N-nitroso-N-ethylurea [5 mM]; reg. round, regeneration round.

[a] Lines S91G-2 and D93E/E94K-1

[b] Initially, two mutations were obtained. One gave rise to line Q89K-1, the other (T90N-1) was lost due to random mitochondrial genome segregation. However, one regenerated shoot later displayed a new mutation (line D93G-1) that was recovered.

[c] Line V95I-1

[d] Line R85L-1