











Removal of the large inverted repeat from the plastid genome reveals gene dosage effects and leads to increased genome copy number

Received: 22 November 2023

Accepted: 24 April 2024

Published online: 27 May 2024

 Check for updates

Carolin Krämer , Christian R. Boehm , Jinghan Liu ,
Michael Kien Yin Ting , Alexander P. Hertle, Joachim Forner ,
Stephanie Ruf , Mark A. Schöttler , Reimo Zoschke  & Ralph Bock  

The chloroplast genomes of most plants and algae contain a large inverted repeat (IR) region that separates two single-copy regions and harbours the ribosomal RNA operon. We have addressed the functional importance of the IR region by removing an entire copy of the 25.3-kb IR from the tobacco plastid genome. Using plastid transformation and subsequent selectable marker gene elimination, we precisely excised the IR, thus generating plants with a substantially reduced plastid genome size. We show that the lack of the IR results in a mildly reduced plastid ribosome number, suggesting a gene dosage benefit from the duplicated presence of the ribosomal RNA operon. Moreover, the IR deletion plants contain an increased number of plastid genomes, suggesting that genome copy number is regulated by measuring total plastid DNA content rather than by counting genomes. Together, our findings (1) demonstrate that the IR can enhance the translation capacity of the plastid, (2) reveal the relationship between genome size and genome copy number, and (3) provide a simplified plastid genome structure that will facilitate future synthetic biology applications.

Plastids (chloroplasts) contain their own genetic information, which is stored in the plastid genome (plastid DNA or ptDNA). The ptDNA maps as a circular genome that, in most embryophytes (land plants), has a tripartite structure. A large single-copy region (LSC) is separated from a small single-copy region (SSC) by a pair of large inverted repeats^{1,2} (IR_A and IR_B). The IRs are identical in sequence and, in a typical seed plant plastid genome, account for approximately a third of the genome size. For example, in the model plant tobacco (*Nicotiana tabacum*), the total size of the ptDNA is 156 kb, and the two IRs together are 50.6 kb (ref. 3 and Extended Data Fig. 1). In seed plants, exceptions to the conserved tripartite genome organization are found in a few taxa that secondarily lost one of the IRs for unknown reasons^{4,5}. The size of the IR region is variable, and during evolution, both expansions and contractions have occurred. Very large IRs are found, for example, in some species of the family Geraniaceae^{6,7}, where the IR size has tripled and become larger than 70 kb. Although the size of the IR region shows substantial

variation between species, and the IR borders with the two single-copy regions of the genome can be variable even between closely related species⁸, a nearly universally conserved feature is the presence of the ribosomal RNA (rRNA) operon within the IR^{1,2}. The rRNA operon typically encodes the full set of rRNA species contained in the chloroplast ribosome: the 16S rRNA of the small ribosomal subunit and the 23S, 5S and 4.5S rRNAs of the large ribosomal subunit.

The two copies of the IR constantly undergo homologous recombination. Due to their inverted orientation, recombination does not generate subgenomic circles but instead changes the relative orientation of the LSC and SSC (that is, it inverts one single-copy region relative to the other). This phenomenon has been termed flip-flop recombination^{9,10}. Its immediate consequence is that the ptDNA occurs in two conformations that are interconvertible and, due to the high efficiency of homologous recombination in plastids, typically co-exist in a 1:1 ratio⁹.

The function of the large IR region in plastid genomes has remained enigmatic. The presence of the rRNA operon within the IR raises the possibility that, by increasing rRNA gene dosage, the IR supports ribosome biogenesis¹¹ given that rRNAs account for more than 90% of the cellular RNA content in both prokaryotes and eukaryotes. Interestingly, genes in the IR show substantially slower nucleotide substitution rates than genes in the single-copy regions of the plastid genome^{12–16}. This is probably due to copy correction by gene conversion^{12,17} and could be the reason for the presence of highly conserved genes in the IR (rRNA and transfer RNA genes) that need to be protected from mutation. Studies in the few exceptional taxa that lack the IR have revealed elevated mutation rates and increased frequency of genome rearrangements over evolutionary timescales^{4,18}, possibly suggesting that the IR may be important for long-term plastid genome stability. However, it is important to note that there are also rearranged plastid genomes that contain an IR (and non-rearranged genomes that lack an IR), suggesting that other factors such as small direct repeats could be more relevant causes of structural instability^{5,11,19,20}.

The proposed hypothetical functions of the IR are not necessarily mutually exclusive, and it seems possible that low mutation rates, increased genome stability and elevated gene dosage of IR-resident genes together explain the presence of the IR in plastid genomes. In this work, we have addressed possible functions of the IR by removing one of its two copies from the plastid genome of tobacco (*N. tabacum*). We show that homoplasmic plant lines with a uniform population of IR-deleted plastid genome copies (Δ IR plants) can be obtained. Under standard growth conditions, the Δ IR plants are indistinguishable from wild-type plants. However, they show a developmental reduction in plastid ribosome abundance, indicating that doubled rRNA gene dosage may be one of the benefits linked to IR presence. The substantial reduction in plastid genome size resulting from IR deletion also prompted us to investigate the effect of IR removal on the control of ptDNA content and genome replication. We report that the reduced plastid genome size in the Δ IR plants is compensated by an increase in ptDNA copy number, suggesting that chloroplast genome replication is regulated on the basis of total ptDNA content rather than the number of DNA molecules per organelle.

Results

Development of a transplastomic strategy for IR removal

The complete deletion of the IR is complicated by the location of two of the borders between the IRs and the single-copy regions in essential genes: the large open reading frame *ycfI* (ref. 21) at one of the two IR–SSC borders and a large operon of ribosomal protein genes³ at one of the two IR–LSC borders (Fig. 1a and Extended Data Fig. 1). *ycfI* encodes a component of the protein import machinery in the inner membrane of the chloroplast envelope²², and its disruption results in the loss of cell viability²¹. In addition to the border regions with the LSC and the SSC, the two different genome conformations resulting from flip-flop recombination⁹ need to be considered for the development of an IR deletion strategy (Fig. 1a).

IR removal can theoretically be achieved with four different deletion constructs that target either IR_A or IR_B in either of the two genome conformations resulting from flip-flop recombination (Fig. 1a). Three of the four possible deletion strategies are expected to result in lethality, because they lead to the disruption of *ycfI*, the ribosomal protein operon or both (Fig. 1a). The fourth strategy leaves all genes intact by combining them in the single IR copy that is retained (IR_B in Fig. 1a).

On the basis of these considerations, we constructed a chloroplast transformation vector to remove IR_A (Fig. 1a,b). The vector contains the appropriate flanking regions taken from the LSC (*psbA* region) and the SSC (*ndhF* region) to enable precise IR excision by homologous recombination. Crossover events in the two flanks replace the IR with the selectable marker gene for plastid transformation, a chimeric *aadA* gene conferring spectinomycin resistance²³ (Fig. 1b). The *aadA* cassette was equipped with *loxP* sites immediately upstream and downstream to enable post-transformation marker removal by site-specific recombination mediated by the Cre recombinase²⁴ (Fig. 2b,c). As a control for neutrality of the marker insertion site in the chloroplast genome, an additional vector was constructed (*aadA* control) that leaves both IRs intact and solely integrates the *aadA* cassette (with its flanking *loxP* sites) into the border between the IR and the SSC (Fig. 1).

Generation of homoplasmic transplastomic IR-lacking plants

The constructed transformation vectors were loaded onto microparticles and biolistically bombarded into tobacco cells. Selection for spectinomycin resistance conferred by the *aadA* marker gene cassette yielded numerous antibiotic-resistant lines that were subjected to additional rounds of regeneration in the presence of spectinomycin to eliminate residual wild-type copies of the highly polyploid chloroplast genome²⁵. Homoplasmy of the transformed plastid genomes was assessed by Southern blotting (Fig. 2a). Two independent homoplasmic Δ IR-*aadA* transplastomic lines and an *aadA* control line (Fig. 1b) were subsequently subjected to selectable marker gene elimination by site-specific recombination. To this end, crosses with a plant line that expresses a plastid-targeted Cre recombinase (linked to a gentamycin resistance gene²⁴) were conducted, and the progeny were assayed for loss of the *aadA* marker gene. Homoplasmy of the marker-free plastid genomes was evidenced by both DNA gel blot analysis (Fig. 2a) and inheritance tests (Fig. 2b) that revealed homogeneous loss of spectinomycin resistance in the offspring. Finally, to obtain plant lines that are completely free of transgenes, the gentamycin resistance marker was removed from the nuclear genome by crosses followed by seed assays and PCR tests to identify lines that are no longer resistant to gentamycin (Fig. 2b). A set of PCR assays with primer pairs that distinguish between IR_A-containing and IR_A-free genomes confirmed the absence of IR_A in the Δ IR plants (Extended Data Fig. 3).

To provide ultimate proof of the successful deletion of the IR from the plastid genome, pulsed-field gel electrophoresis (PFGE) of purified chloroplasts was conducted to directly visualize the reduced plastid genome size. The removal of the IR results in a substantially smaller plastid genome of 131 kb (instead of 156 kb in the wild type);

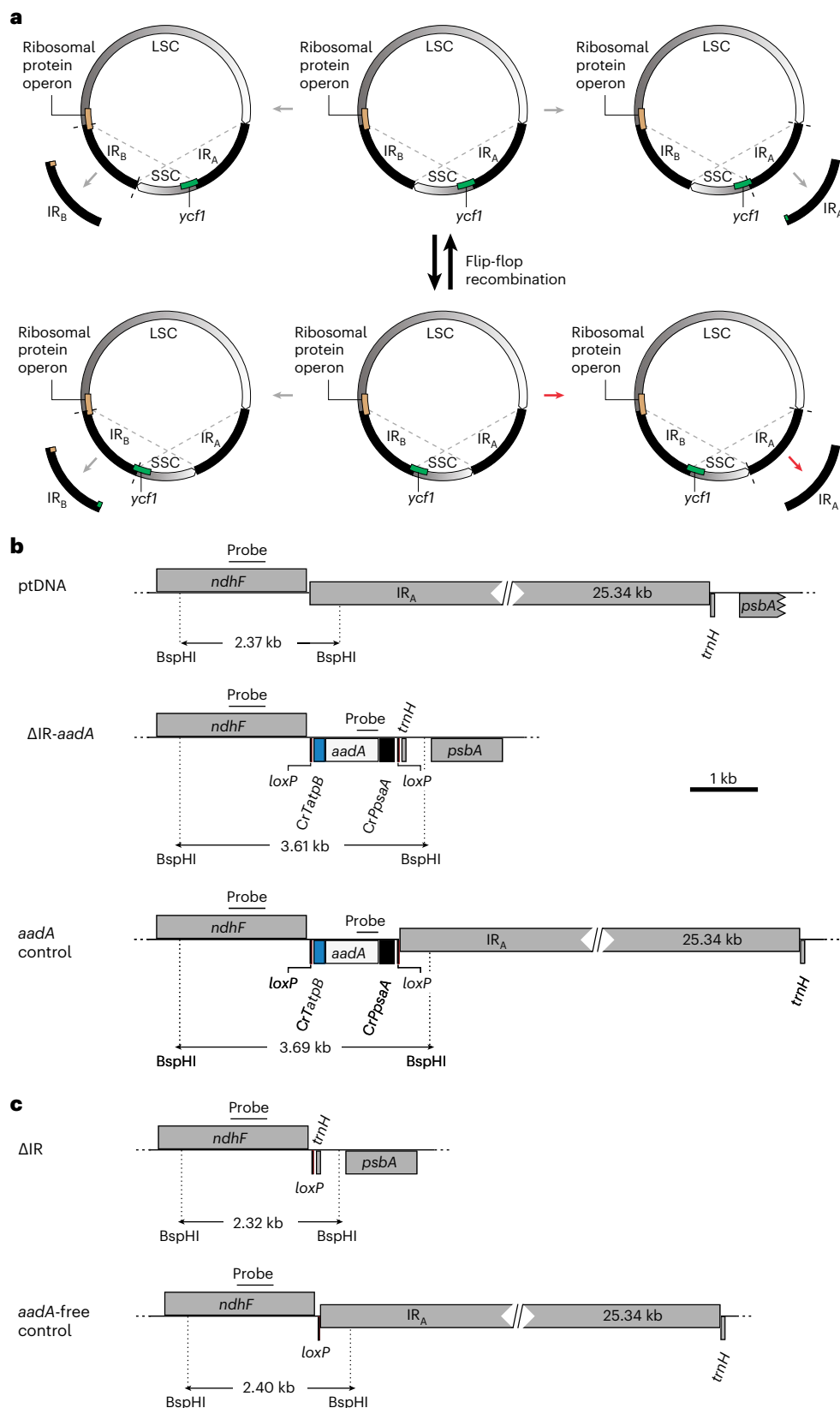
Fig. 1 | Experimental strategies for IR removal and subsequent selectable marker gene elimination from the plastid genome. a, Isoforms of the plastid genome and analysis of the four possible IR deletion scenarios. The ptDNA constantly undergoes flip-flop recombination (shown by dashed lines) resulting in two genome isoforms (depicted in the middle) that differ in the relative orientations of the LSC and SSC. The four theoretical possibilities for IR removal are shown (IR_B deletions are on the left; IR_A deletions are on the right). Genes spanning the borders between the IRs and the single-copy regions of the genome are indicated to illustrate the expected effects of each deletion scenario. Note that only the genome configuration marked with red arrows retains the integrity of the ribosomal protein operon (brown box) at the LSC–IR border and the essential gene *ycfI* (green box) at the SSC–IR border. See the text for the details. **b**, Physical maps of transformation vectors for the deletion of IR_A from the tobacco plastid genome and construction of a control line for neutrality of *aadA* marker

insertion. The maps show the targeting region in the tobacco plastid genome (ptDNA) and the engineered regions in the IR deletion lines (Δ IR-*aadA*) and the control lines (*aadA* control). In Δ IR-*aadA* plants, IR_A is eliminated and replaced with the *aadA* cassette. The *aadA* control plants harbour the *aadA* cassette in the same site but have two intact IR copies (Methods). Genes above the line are transcribed from left to right; genes below the line are transcribed in the opposite direction. For simplicity, the 25.34-kb IR is indicated by a horizontally broken box (and the genes in the IR are not shown). The *aadA* cassette in both constructs is flanked by *loxP* sites. The expected sizes of DNA fragments in restriction fragment length polymorphism analyses with the restriction enzyme BspHI are given. Hybridization probes (derived from *ndhF* and *aadA*) are indicated by black bars. **c**, Physical maps of the plastid genomes after selectable marker removal mediated by the Cre recombinase. *aadA* excision results in transgene-free plastid genomes that retain only a 34-bp *loxP* scar at the excision site.

Extended Data Figs. 1 and 2)—a size difference that should be readily resolvable by PFGE. Indeed, the chloroplast genomes of the IR deletion lines showed the expected shift in electrophoretic mobility compared with the wild-type ptDNA (Fig. 2c). The size reduction was seen both in the monomeric genomes and in the genome dimers that arise from homologous recombination between two monomers (Fig. 2c).

Phenotype and physiological characterization of Δ IR lines

Having isolated homoplasmic marker-free plants whose chloroplast genomes are identical to that of the wild type except for the absence of one of the 25.3-kb IRs (Extended Data Figs. 1 and 2), we next wanted to determine the phenotypic consequences of the IR removal. To this end, the Δ IR lines were compared to wild-type plants under a variety of



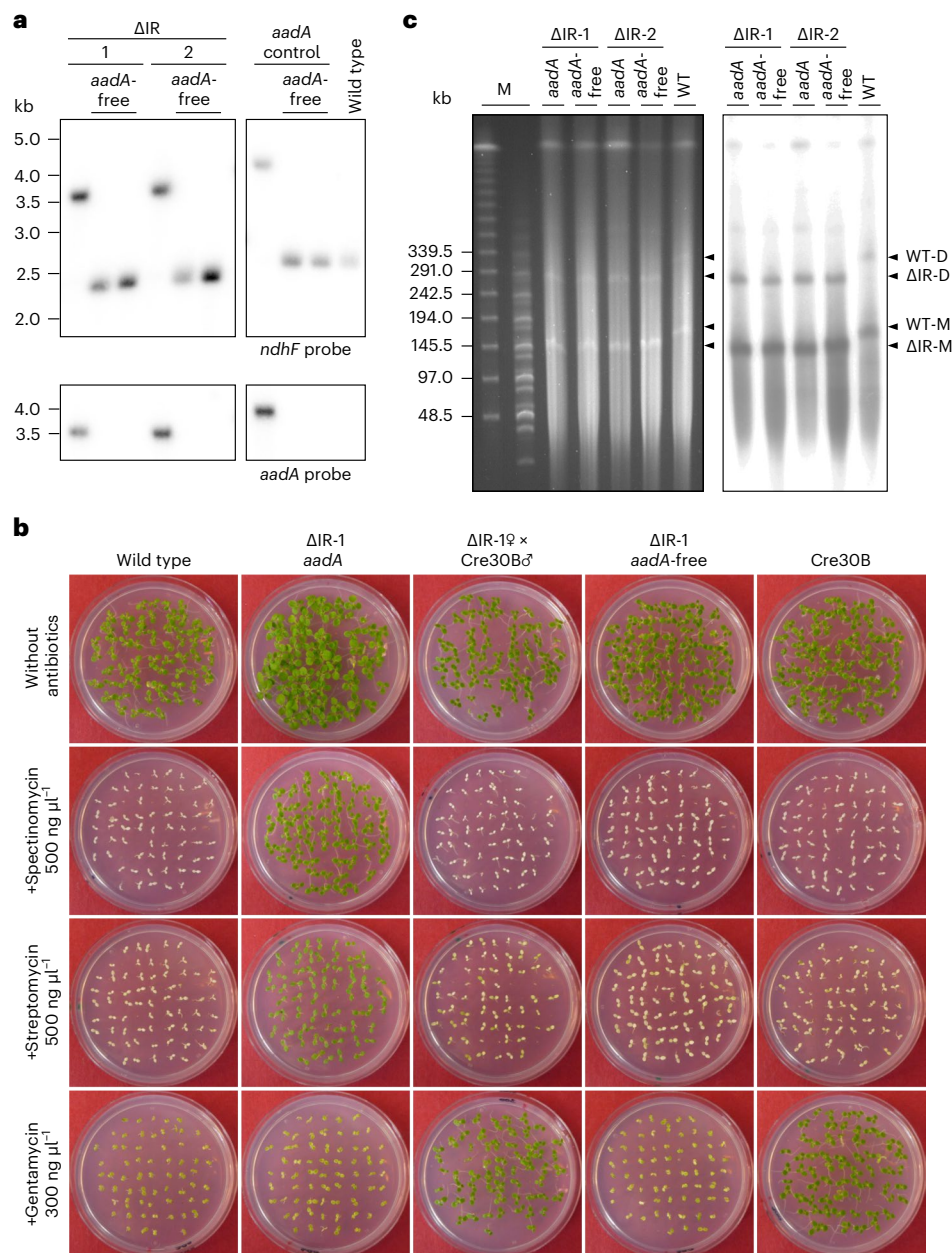


Fig. 2 | Homoplasmy of transplastomic lines, selectable marker gene elimination and size reduction of the plastid genome resulting from IR deletion. **a**, Restriction fragment length polymorphism analysis of *aadA*-containing and *aadA*-free ΔIR and control lines. Total DNA was digested with BspHI and hybridized to radiolabelled *ndhF* or *aadA* probes (Fig. 1b,c). Homoplasmy is evidenced by the absence of the 2.4-kb *ndhF* fragment diagnostic of the wild-type genome in the *aadA*-containing lines and exclusive detection of the expected fragment sizes (ΔIR lines, 3.6 kb; *aadA* control line, 3.7 kb). Successful marker gene elimination is evidenced by a reduction in fragment size that corresponds to the size of the *aadA* cassette (1.3 kb) in the hybridization to the *ndhF* probe (top) and by hybridization to an *aadA*-specific probe (bottom). All samples were run in the same gel (irrelevant lanes were digitally removed). **b**, Seed assay for all steps in the marker elimination process. Seeds were sown on antibiotic-free medium or on media supplemented with spectinomycin, streptomycin or gentamycin. *aadA*-containing transplastomic seedlings are resistant to spectinomycin and

streptomycin. The nuclear-transgenic Cre30B line²⁴ contains a gentamycin resistance gene. The F₃ generation is sensitive to spectinomycin and streptomycin but resistant to gentamycin, indicating successful elimination of the *aadA* marker. The F₃ generation is sensitive to all antibiotics and indistinguishable from the wild-type control (ΔIR-1 *aadA*-free). The seedlings were photographed 18 days after sowing. **c**, PFGE showing the size reduction of the plastid genome resulting from IR deletion in the ΔIR mutants. Two *aadA*-containing IR deletion lines and the corresponding *aadA*-free lines were analysed. The ethidium-bromide-stained pulsed-field gel and the Southern blot analysis of the blotted gel by hybridization to a radiolabelled *psaB* probe are shown. The expected sizes for the monomeric form of the plastid genome are 156 kb for the wild type (WT), 132 kb for the *aadA*-containing IR deletion mutants (ΔIR *aadA*) and 130.7 kb for the *aadA*-free IR deletion mutants (ΔIR *aadA*-free). The monomeric (-M) and dimeric (-D) forms of the plastid genomes are detectable in all lines. M, DNA size markers. The experiments in **a** and **c** were performed once.

different growth conditions (Fig. 3 and Methods). No distinctive growth phenotype was seen under standard greenhouse conditions, and the ΔIR plants were phenotypically indistinguishable from the wild type (Fig. 3a). Exposure to various abiotic stress conditions also did not cause

a visible mutant phenotype in the IR deletion plants. Most importantly, chilling stress, an environmental condition that is challenging to both photosynthesis^{26,27} and plastid gene expression activity^{28–30}, did not affect the growth or development of the ΔIR lines more severely than

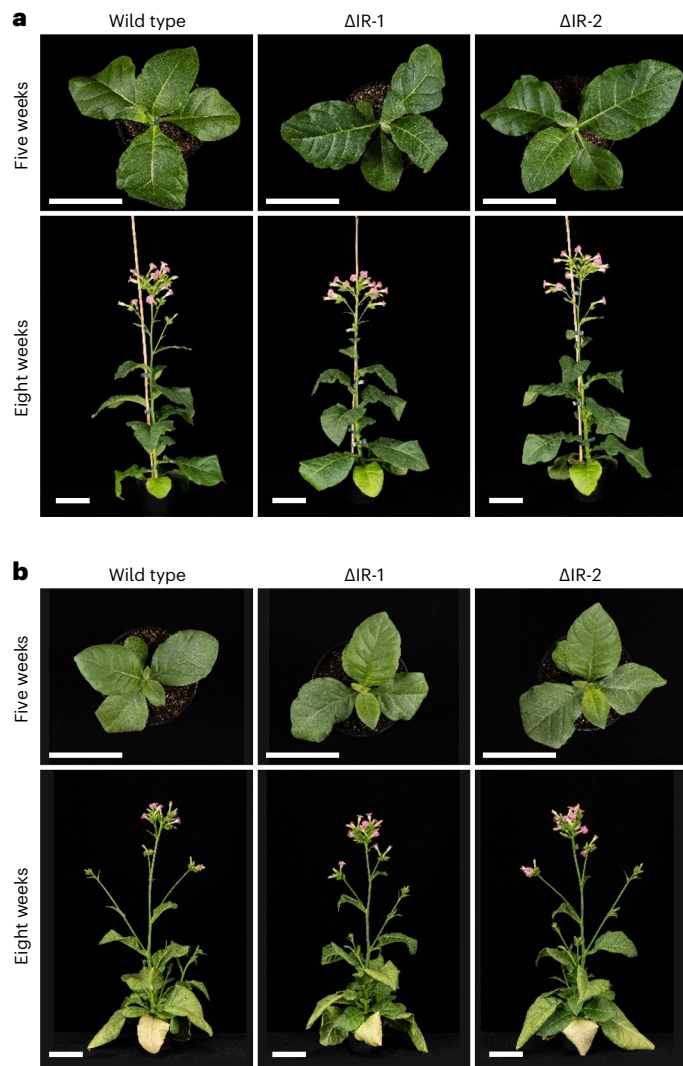


Fig. 3 | IR deletion plants show a wild-type-like phenotype under different growth conditions. **a**, Plants grown under standard greenhouse conditions (average light intensity, $\sim 300 \mu\text{mol photons per m}^2 \text{ per s}$). **b**, Plants grown under high-light conditions ($1,000 \mu\text{mol photons per m}^2 \text{ per s}$). Seedlings were transferred to high-light conditions at the age of four weeks after sowing. In both panels, the photos show a representative wild-type plant and two independently generated marker-free IR deletion lines ($\Delta\text{IR-1}$ and $\Delta\text{IR-2}$). The same set of plants is shown at the ages of five and eight weeks. Scale bars, 10 cm.

those of the wild-type control (Fig. 5a). Finally, to trigger maximum growth rates (which are expected to require optimum performance of photosynthesis and plastid gene expression), plants were exposed to high-light conditions. This also did not lead to a visibly different phenotype (Fig. 3b), suggesting that the absence of the IR does not entail a serious disadvantage under a variety of growth conditions.

As photosynthesis is the main function of chloroplasts, and most of the plastid-genome-encoded genes are directly or indirectly involved in photosynthesis, we next characterized various photosynthesis-related parameters in high-light-acclimated ΔIR plants. To capture possible developmental differences, all measurements were performed with both young leaves (which developed after the high-light shift) and fully developed leaves (which had already developed prior to the shift; Extended Data Fig. 4). While most parameters were very similar in ΔIR plants and the wild type, a subtle difference was noticed in the light response curves of non-photochemical quenching in young leaves. The ΔIR plants showed a slightly higher capacity of linear electron transport (ETR_{II}), and, accordingly, the induction of photoprotective

non-photochemical quenching was shifted to higher light intensities (Extended Data Fig. 4c). Since this effect was seen only in young leaves, it may be related to the biogenesis of the photosynthetic apparatus (which is highly active in developing leaves but largely completed in fully expanded leaves^{31,32}). As the biogenesis of the photosynthetic apparatus during leaf development absorbs most of the gene expression capacity of the chloroplast, subtle reductions in plastid gene expression activity often translate into measurable physiological phenotypes only in young leaves^{28,29}. On the basis of these observations and considerations, and in view of the presence of the entire rRNA operon in the IR region of the plastid genome, we next turned our attention to the analysis of gene expression in the ΔIR plants.

IR deletion leads to reduced plastid ribosome numbers

The majority of the genes in the IR of the tobacco plastid genome are involved in translation. In addition to all rRNA genes (16S, 23S, 5S and 4.5S rRNAs), there are four ribosomal protein genes and seven transfer RNA genes contained in the IR (Extended Data Fig. 1). Only two IR-resident genes, the *ndhB* gene (encoding a subunit of the plastid NAD(P)H dehydrogenase complex³³) and the *ycf2* reading frame (presumably involved in protein import³⁴), are not involved in chloroplast translation.

To determine whether the deletion of one of the IR copies affects ribosome biogenesis, we quantified plastid ribosome abundance during leaf development. As rRNA abundance is a reliable proxy of ribosome abundance (in that plastid rRNAs are stable only when incorporated into ribosomal subunits^{29,35}), the ratio of chloroplast rRNAs to cytosolic rRNAs can serve as a sensitive indicator of alterations in plastid ribosome numbers^{36,37}. We therefore analysed a developmental series of leaves of ΔIR and wild-type plants (Fig. 4a) and determined the ratio of (1) the 16S rRNA in the small subunit of the chloroplast ribosome to the 18S rRNA in the small subunit of the cytosolic ribosome and (2) a fragment of the 23S rRNA in the large subunit of the chloroplast ribosome to the 18S rRNA of the cytosolic ribosome (Fig. 4b–f). The 23S rRNA undergoes post-transcriptional fragmentation (by a process known as hidden break processing^{37,38}), resulting in the presence of three fragmentation products in the mature chloroplast ribosome (Fig. 4b).

Both quantifications revealed a reduction of chloroplast ribosome abundance in the ΔIR lines that was particularly pronounced in developing leaves (Fig. 4c–f). While the fully expanded leaves (leaf number 3 (L3)) had very similar numbers of plastid ribosomes, developing leaves (L5, L7 and L9) showed a reduced abundance of plastid ribosomes. The reduction was most pronounced (and statistically significant) in L7 (Fig. 4e,f), but a similar trend was also seen in the slightly older (but also not yet fully expanded) L5 and in the youngest leaf analysed (L9). These results suggest that the reduced gene dosage of IR-encoded ribosomal components limits ribosome biogenesis during leaf development and results in lower numbers of chloroplast ribosomes in our ΔIR plants.

As explained above, chilling stress represents a particularly challenging environmental condition that affects both photosynthesis²⁶ and plastid gene expression^{28,29}. We therefore also analysed plastid ribosome abundance in developing leaves of young plants growing under chilling stress (Fig. 5). These analyses confirmed the reduced plastid ribosome abundance in the ΔIR plants and, in the most strongly affected leaf (L2), revealed an up to 50% lower number of chloroplast ribosomes in the young plants (Fig. 5c).

Genome-wide plastid transcription and translation analysis

To examine whether the reduced number of plastid ribosomes in the ΔIR plants can be directly related to gene dosage effects (that is, reduced expression levels of IR-resident genes), we conducted comparative transcript and ribosome profiling experiments with two ΔIR lines and the wild type (Fig. 6). Consistent with the absence of a visibly different phenotype, no strong differences in transcript abundances

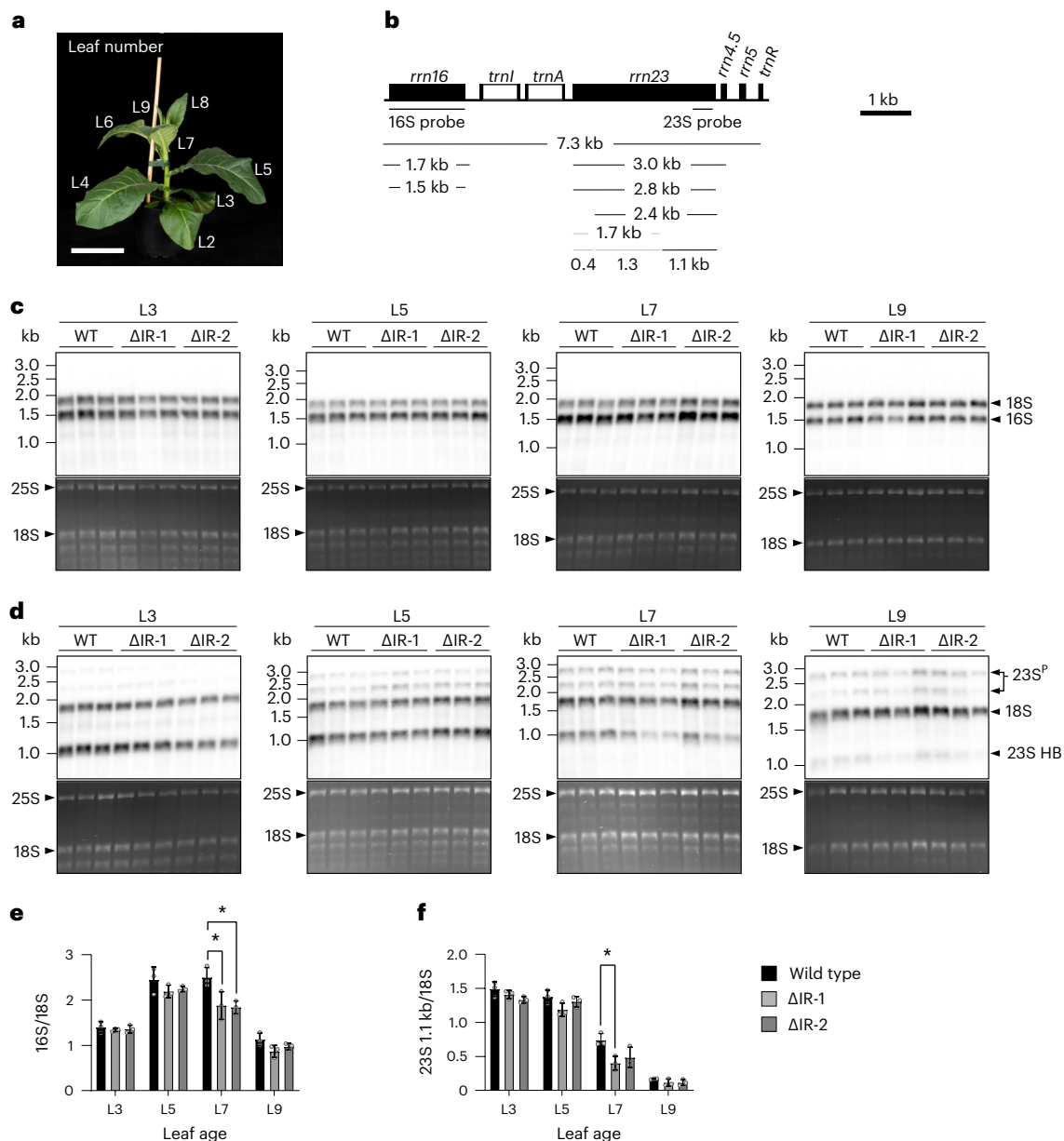


Fig. 4 | Analysis of plastid ribosome abundance during leaf development. **a**, Developmental stages of tobacco leaves analysed. The leaves were consecutively numbered from the bottom of the plant (excluding the cotyledons), and L3, L5, L7 and L9 were harvested for RNA isolation. A Δ IR-2 plant, grown for six weeks under standard greenhouse conditions, is shown here as a representative example. Scale bar, 10 cm. **b**, Physical map of the plastid rRNA operon, and its transcription and RNA processing products^{38,64}. The sizes of major RNA species are given, and the two hybridization probes are indicated by horizontal bars. The 23S rRNA probe used here is derived from the 1.1-kb hidden break (HB) fragment³⁸, to simplify the band pattern and facilitate relative quantification. Black bars indicate expected band sizes, grey bars represent 23S rRNA HB fragments and processing products that are not detected by the 23S rRNA probe used here. **c,d**, RNA gel blot analyses to quantify plastid rRNAs as a proxy of ribosome abundance. Total cellular RNA was extracted from every other leaf of three biological replicates (individual plants) per plant line (the wild type and

two marker-free Δ IR mutant lines, Δ IR-1 and Δ IR-2), separated by denaturing gel electrophoresis, and simultaneously hybridized to a radiolabelled probe for the cytosolic 18S rRNA (serving as a reference for relative quantification) and a radiolabelled probe for either the plastid 16S rRNA (**c**) or 23S rRNA (**d**). The ethidium-bromide-stained gels prior to blotting are shown below each blot as a control for RNA integrity and equal loading. The positions of the cytosolic 25S and 18S rRNAs are marked. The 23S rRNA probe detects, in addition to the 1.1-kb HB product, two precursors ($23S^P$) of 2.8 and 2.4 kb (cf. **b**). **e**, Quantification of plastid 16S rRNA abundance relative to the cytosolic 18S rRNA (Δ IR-1 L7 $P = 0.047$; Δ IR-2 L7 $P = 0.0123$). **f**, Quantification of plastid 23S rRNA abundance (1.1-kb HB fragment) relative to the cytosolic 18S rRNA (Δ IR-1 L7 $P = 0.0135$). The calculations are based on measured band intensities of the hybridized blots. The error bars indicate the standard deviation of the results obtained from the three independent biological replicates. The asterisks denote statistically significant differences ($P < 0.05$, two-sided Student's t -test).

were detected between the IR deletion and wild-type plants (Extended Data Fig. 5a,b). At the translational level, only three chloroplast genes (*rps14*, *rps8* and *rpl16*) displayed significant differences in ribosome footprint numbers (Extended Data Fig. 5c,d). In general, the measured changes in gene expression were relatively moderate (less than

twofold), consistent with the mild alterations in ribosome abundance (Supplementary Data 1). Although most changes in the expression of individual genes were not statistically significant (Extended Data Fig. 5), a plastid-genome-wide view of protein-coding genes revealed a pronounced general trend towards an overall decrease in transcript

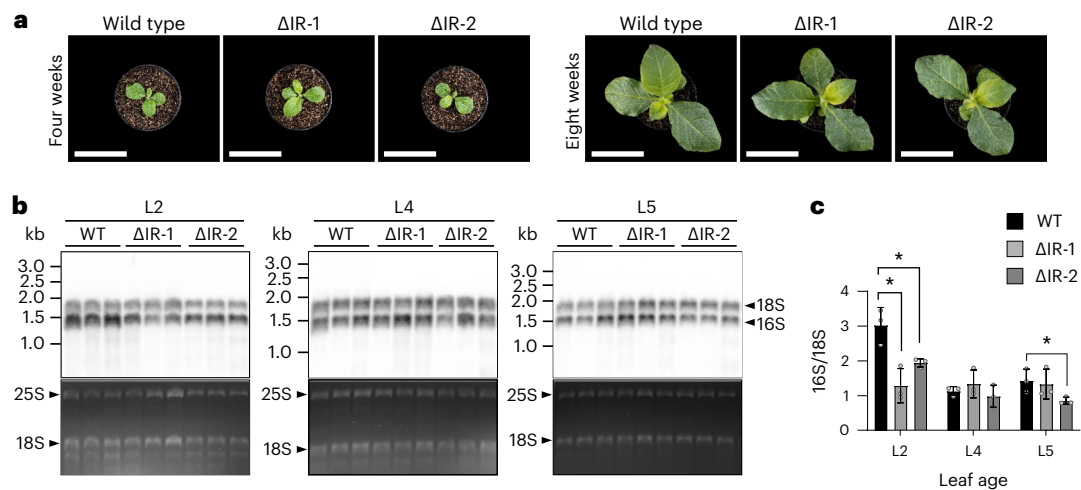


Fig. 5 | Analysis of plastid ribosome abundance during leaf development under chilling stress. **a**, Phenotypes of plants grown under chilling stress conditions (12 °C, 250 μ mol photons per m^2 per s). Plants were transferred to the cold conditions at the age of four weeks. The photographs show the plants at the time of transfer to the cold (four weeks) and at the time of leaf harvest (eight weeks). The experiment was performed once. Scale bars, 10 cm. **b**, RNA blot analyses to quantify rRNAs as a proxy for chloroplast ribosome abundance under cold conditions. L2, L4 and L5 (numbered from the bottom; Fig. 4a) were analysed in three biological replicates. RNA gel blots were simultaneously hybridized to a radiolabelled probe for the cytosolic 18S rRNA (serving as a reference for relative

quantification) and a radiolabelled probe for the plastid 16S rRNA (for the details, see Fig. 4). The ethidium-bromide-stained gels prior to blotting are shown below each blot as a control for RNA integrity and equal loading. The bands representing the cytosolic 25S and 18S rRNAs are indicated. **c**, Quantification of plastid 16S rRNA abundance relative to the cytosolic 18S rRNA (Δ IR-1 L2 $P = 0.0132$; Δ IR-2 L2 $P = 0.0228$; Δ IR-2 L5 $P = 0.0450$). The calculations are based on measured band intensities of the hybridized blots. The error bars indicate the standard deviation of the results obtained from the three independent biological replicates. The asterisks denote statistically significant differences ($P < 0.05$, two-sided Student's t -test).

accumulation and translation of IR-resident genes (Fig. 6a,b). There was also a clear tendency for non-IR genes to display an overall increase in translation in the Δ IR mutants (Fig. 6b), which may be part of a regulatory response to the IR deletion (see below).

Taken together, our analyses of chloroplast ribosome abundance and our comparative studies of mRNA abundances and translational output provide strong evidence for a function of the IR in gene expression. By doubling the gene dosage of key components of the plastid translational apparatus, the IR enhances plastid ribosome biogenesis and contributes to the efficiency of plastid gene expression during leaf development, when the maximum gene expression capacity is needed.

Control of ptDNA replication by genome size

The generation of plants with a substantially reduced plastid genome size provides unique genetic material to investigate the correlation between genome size and genome copy number. In this regard, two fundamentally different scenarios are theoretically possible. If no correlation between genome size and copy number exists, the copy number of the plastid genome in the IR deletion plants should be identical to that in the wild type. This would imply that plastid genome replication is regulated by the cell 'measuring' the number of plastid genome molecules, and if a sufficient number has been reached, DNA replication stops. In the alternative scenario, the cell would measure not plastid genome copy numbers but ptDNA content (and would adjust the DNA amount per organelle to a fixed value). In this scenario, the copy number of the plastid genome should increase in the Δ IR plants proportionally to the reduction in genome size. The size reduction of the plastid genome in the IR deletion lines is approximately 16%—a value that is substantial enough to be able to determine whether or not the cell responds to this reduction by increasing plastid genome copy numbers.

To distinguish between the two possible modes of replication control, genome copy numbers were measured using quantitative real-time PCR (qPCR) (Fig. 7 and Methods). To this end, sets of amplicons were derived from (1) three genes residing in the IRs, (2) three genes residing in the LSC, (3) three genes residing in the SSC, (4) two nuclear genes that

were used for normalization and (5) two genes in the mitochondrial genome (as additional controls). Genome copy number quantification was performed in seedlings (Fig. 7a), green leaves (Fig. 7b) and roots as a representative non-green tissue (Fig. 7c).

In all three tissue types, genome copy determination in the wild type showed the expected behaviour in that IR-resident plastid genes were present in approximately twice as many copies as LSC and SSC genes. The data also confirmed previous studies revealing that mitochondrial genome copy numbers are far lower than plastid genome copy numbers per cell^{39,40} and that non-photosynthetic plastids contain fewer plastid genome copies than leaf chloroplasts⁴¹. Estimates in *Arabidopsis* have indicated that, in roots, the copy number of the plastid genome is only one fifth of that in green leaves⁴¹. The difference determined by our qPCR measurements in tobacco is in the same range (Fig. 7).

When genome copy numbers in the Δ IR plants were determined, a striking pattern was seen. All six amplicons derived from the two single-copy regions of the plastid genome consistently showed a substantial increase in copy number in all three tissue types investigated (Fig. 7). By contrast, all IR-derived amplicons showed a copy-number reduction, which, however, was significantly lower than the expected reduction by 50%. Interestingly, both the increase in copy numbers of the LSC and SSC marker genes and the lower-than-expected reduction in the copy number of the IR marker genes correlated well with the genome size reduction resulting from IR removal, in that the -16% reduction in genome size led to a similar increase in plastid genome copy numbers in the Δ IR plants (Supplementary Data 2). To exclude the possibility that chloroplast size and/or number per cell are altered in the Δ IR plants, chloroplasts in cells of the wild type and the IR deletion lines were observed by confocal laser-scanning microscopy and counted (Extended Data Fig. 6). No difference in organelle size or number could be detected, strongly suggesting that the observed regulation occurs at the level of genome copy number.

Together, these data strongly suggest that ptDNA replication is not controlled on the basis of genome copy numbers but rather regulated by adjusting the DNA content per plastid to a fixed value.

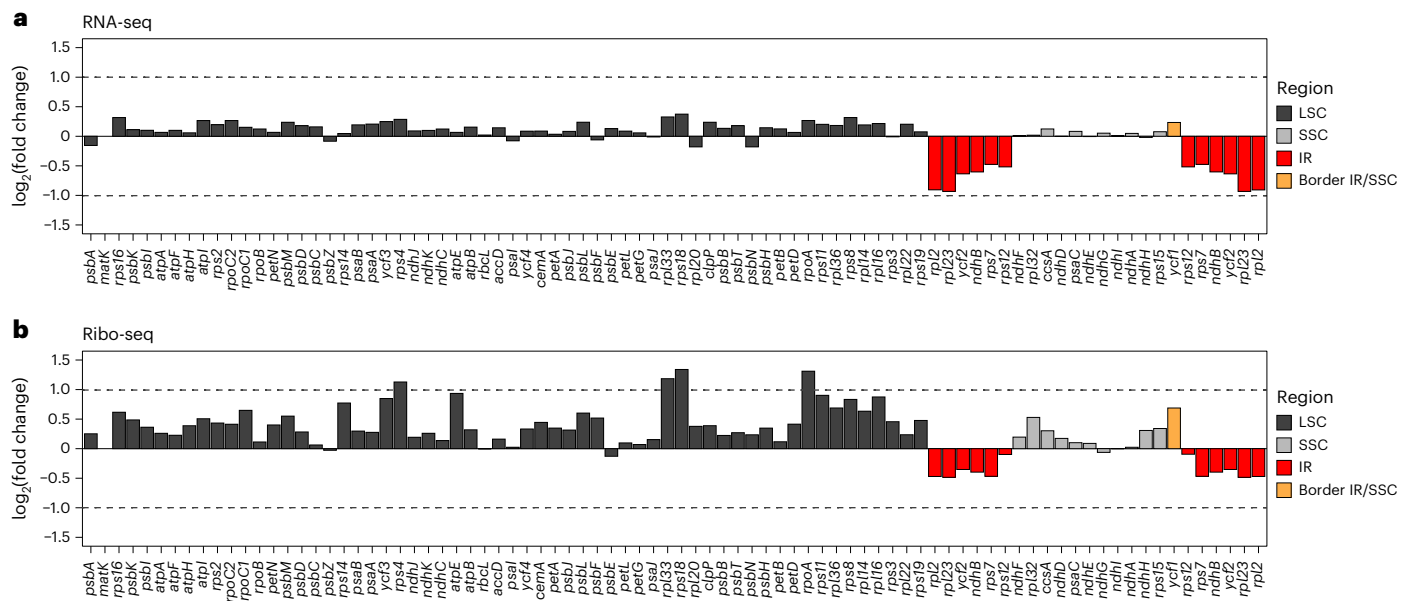


Fig. 6 | Plastome-wide overview of transcriptional and translational changes in protein-coding genes for Δ IR plants compared with the wild type. a, b, L5, L6 and L7 from greenhouse-grown plants (Fig. 4a) were harvested and pooled at equal mass for extraction of total RNA and isolation of ribosome footprints. The data summarize results from three biological replicates (that is, three individual plants) for each of the two independent Δ IR lines. Log₂ fold changes for transcript abundance are shown in a, and log₂ fold changes for ribosome footprint abundance are shown in b. Note that the changes measured were

significant (false discovery rate, <0.05) at the level of translation only for the genes *rps14*, *rps8* and *rpl16* (Extended Data Fig. 5). Reads from the wild type and the Δ IR mutants were mapped to the same reference genome; therefore, IR reads in the Δ IR mutants are distributed over two IR copies. The data were normalized to all obtained reads for nuclear, chloroplast and mitochondrial genes. For *matK*, no reads were obtained in RNA-seq and ribo-seq experiments. Log₂ fold changes were obtained from analyses using the R package edgeR.

Discussion

In this study, we have developed an experimental approach to address the functional relevance of one of the most enigmatic features of plastid genomes: the presence of a large IR region. The IR doubles the gene dosage of a substantial number of chloroplast genes, including the entire rRNA operon, and greatly increases the size of an otherwise extremely streamlined and gene-dense genome^{1–3} (Extended Data Fig. 1). By removing one copy of the IR from the tobacco plastid genome, we reduced genome size by as much as 25.3 kb (Extended Data Figs. 1 and 2) and eliminated flip-flop recombination (Fig. 1a), a salient feature of plastid genomes^{9,10}. We have shown that IR deletion and the concomitant genome size reduction are well tolerated by the plant, thus boding well for future efforts to design minimal-size synthetic plastid genomes^{42,43}.

Our molecular analyses of the IR deletion plants revealed that plastid gene expression was mildly affected. The effects observed were most pronounced in rapidly growing leaf tissues (Fig. 4) and can largely be attributed to slightly less efficient ribosome biogenesis in the absence of the second copy of the rRNA operon and a set of ribosomal protein genes (Figs. 4–6). Together, our data reported here suggest that, by providing increased gene dosage for key components of the translational apparatus, an important function of the IR lies in the maximization of the capacity of the plastid gene expression machinery. This is needed, for example, in the expansion phase of leaf development, when rapid cell division occurs, necessitating high levels of chloroplast biogenesis and high synthesis rates of the many plastid-encoded components of the photosynthetic apparatus.

While our study has uncovered the functional relevance of the IR to plastid gene expression, the long-term functional consequences of IR removal remain to be determined. It will be interesting to use the plant material generated in this work to grow the IR deletion lines along with the wild type over many generations to be able to assess the long-term stability of the plastid genome. Two aspects seem

particularly interesting in this regard: (1) the nucleotide substitution rates in genes that are part of the IR in the wild type and (2) the structural stability of the plastid genome. Since IR-resident genes show substantially lower nucleotide substitution rates than genes residing in the single-copy regions of the genome^{13,16}, it would be worthwhile to resequence the plastid genomes in a larger population of IR deletion plants at regular intervals to see whether the expected increase in the nucleotide substitution rate of the remaining copy of the IR can be observed under greenhouse and/or field conditions. Phylogenetic data have also suggested that the loss of the IR can be associated with an increased frequency of genome rearrangements^{4,18}. By large-scale genome resequencing over many generations, it may also be possible to observe this effect in real time and thus provide direct evidence for a role of the IR in long-term plastid genome stability.

In addition to revealing the function of the IR in plastid gene expression, our work has provided new insights into the control of plastid genome copy number and the regulation of DNA replication in plastids. Our quantitative analyses of genome copy numbers revealed that in both chloroplasts and non-photosynthetic plastids, IR removal results in substantially increased numbers of plastid genomes (Fig. 7). Remarkably, this increase in copy number is proportional to the decrease in genome size resulting from IR removal. These data demonstrate that chloroplast genome copy number is regulated by measuring the total ptDNA content rather than by counting genome molecules. Our findings also suggest that ptDNA replication is regulated by adjusting the DNA amounts per plastid to a fixed value (which depends on the tissue type; Fig. 7). It will be interesting to identify the factors that are responsible for this regulation. Plastid DNA replication is controlled by nuclear genes¹⁰, but which gene or genes determine the genome copy number and how this occurs is currently completely unknown. In addition to negative regulatory mechanisms similar to those that control the copy number of bacterial plasmids (at the level of initiation of DNA replication⁴⁴), it seems possible that plastid genome

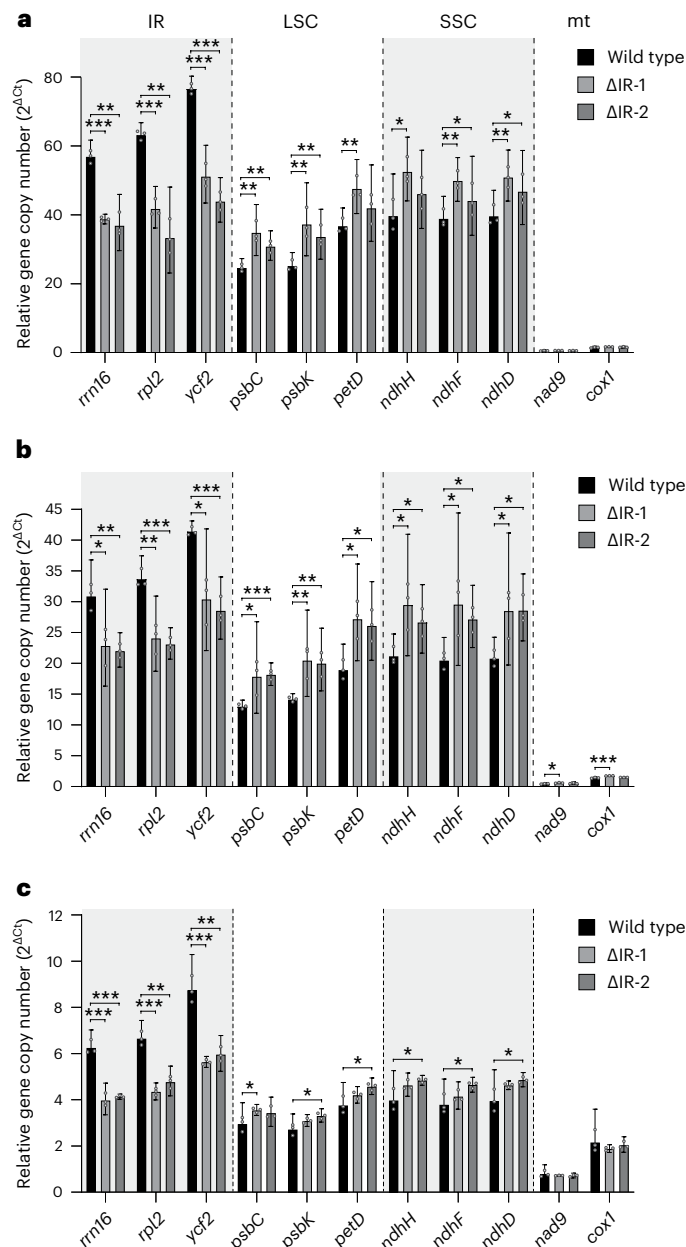


Fig. 7 | Relative quantification of plastid genome copy numbers in wild-type plants and Δ IR plants by qPCR. **a–c.** The analyses were performed with extracted total DNA from pools of five 25-day-old seedlings (**a**); pooled L5, L6 and L7 from greenhouse-grown plants (Fig. 4a) (**b**); and roots from pools of six or seven plants grown in hydroponic culture under aseptic conditions (**c**). The relative gene copy numbers are shown as $2^{\Delta Ct}$, where ΔCt values are Ct (cycle threshold) values normalized to the nuclear genes coding for 18S rRNA and ribosomal protein L25. The error bars indicate the 95% confidence intervals ($n = 3$ biologically independent samples). The locations of the genes analysed are indicated above the plot in **a**. In addition to gene sets in the IR, the LSC and the SSC, two mitochondrial (mt) genes were analysed (which have much lower copy numbers⁴⁰). The determined relative ptDNA copy numbers were analysed using unpaired two-tailed t -tests performed on ΔCt values ($*P < 0.05$; $**P < 0.01$; $***P < 0.001$; for the corresponding P values, see Supplementary Data 2).

copy number is passively determined by the abundance (and processivity) of the replication enzymes.

Finally, the demonstration that the IR can be eliminated without serious deleterious consequences has implications for future applications in plastid synthetic biology and synthetic genomics^{42,43}. Removal of the IR not only makes the plastid genome substantially smaller but

also greatly simplifies plastid genome structure. The assembly of synthetic genomes is much more straightforward if sequence duplications can be avoided. As our IR deletion plants are marker-free, they provide an excellent starting point for further size reductions of the genome using transplastomic approaches, including intron removal, deletion of non-essential genes, and refactoring of operons and gene clusters (by what is commonly referred to as top-down synthetic genomics⁴²).

In summary, our work reported here (1) demonstrates that the chloroplast genome of seed plants can be substantially reduced in size by excision of the large IR, without entailing noticeable phenotypic consequences; (2) reveals that the IR, via a gene dosage effect, supports high expression rates of the plastid genome; and (3) uncovers how the copy number of plastid genomes per cell is regulated.

Methods

Plant material and growth conditions

Nicotiana tabacum cultivar Petit Havana plants were grown under sterile conditions on agar-solidified Murashige and Skoog (MS) medium containing 30 g l^{-1} sucrose⁴⁵. The same medium was used for root induction and initial growth of transplastomic lines. For analysis of plant phenotypes and seed production, plants were grown in soil under standard greenhouse conditions ($\sim 300 \mu\text{mol photons per m}^2 \text{ per s}$ light intensity under a 16 h light/8 h dark regime with an average day temperature of 25°C and an average night temperature of 20°C). For inheritance assays, surface-sterilized seeds were germinated on agar-solidified MS medium supplemented with the appropriate antibiotics. For analysis of root tissue, plants were grown in hydroponic culture under aseptic conditions in liquid MS medium. For growth tests under different environmental conditions, seeds were germinated in soil, and seedlings were grown for four weeks under standard conditions prior to transfer to the stress condition. For growth under high-light conditions, plants were exposed to $1,000 \mu\text{mol photons per m}^2 \text{ per s}$ light intensity (in a 16 h light/8 h dark diurnal cycle). Chilling stress was applied by shifting plants to a day temperature of 12°C and a night temperature of 8°C (at $250 \mu\text{mol photons per m}^2 \text{ per s}$ light intensity in a 16 h light/8 h dark diurnal cycle).

Construction of transformation vectors for IR removal

The construct designed for the deletion of IR_A (ΔIR ; Fig. 1) contains an *aadA* resistance cassette flanked by 1-kb fragments derived from the regions upstream and downstream of IR_A in the targeted chloroplast genome isoform (Fig. 1a). The *aadA* cassette is driven by the *Chlamydomonas reinhardtii* plastid *psaA* promoter (*CrPpsaA*) and the *atpB* terminator (*CrTatpB*)⁴⁶. To generate appropriate transplastomic control lines, the same *aadA* cassette was inserted into the junction of IR_A and the SSC (Fig. 1). In this construct, the marker gene is flanked by 1-kb fragments of the SSC and IR_A . In both plastid transformation vectors, the *aadA* cassette is flanked by *loxP* sites to facilitate post-transformation marker gene excision (Fig. 1). All plastid targeting sequences were PCR amplified and purified by agarose gel electrophoresis followed by extraction from excised gel slices (Nucleospin Gel and PCR Clean-up kit; Macherey-Nagel). The vectors are based on pBlueScript II SK (+), and the PCR fragments were inserted between the *Acc65I* and *SacI* restriction sites in the polylinker by Gibson assembly reactions (New England Biolabs).

Maps of the plastid genomes of the wild type and the IR deletion lines were generated with OGDRAW v.1.3.1^{47–49}.

Plastid transformation and selection of transplastomic lines

Gold particles were coated with $20 \mu\text{g}$ of vector DNA and delivered into young, aseptically grown *N. tabacum* cv. Petit Havana leaves via particle bombardment²³. After biolistic bombardment, the leaves were cut into pieces and placed on regeneration medium containing MS elements with modified vitamins (Duchefa M0245), sucrose (30 g l^{-1}), 6-benzylaminopurine (1 mg l^{-1}), 1-naphthaleneacetic acid (0.1 mg l^{-1})

and agar (5.4 g l⁻¹) at pH 5.8. To facilitate selection of transplastomic lines, the medium was supplemented with 500 µg ml⁻¹ spectinomycin (Duchefa). Primary spectinomycin-resistant shoots were subjected to additional regeneration rounds to select for homoplasmy. To this end, leaf pieces from antibiotic-resistant lines were transferred to fresh regeneration medium containing 500 µg ml⁻¹ spectinomycin. Leaf pieces were additionally cultured on regeneration medium containing 500 µg ml⁻¹ streptomycin (Duchefa) to distinguish transplastomic clones from spontaneously arising spectinomycin-resistant mutants⁵⁰. Homoplasmic shoots were transferred to MS medium supplemented with sucrose (3%) and 500 µg ml⁻¹ spectinomycin to induce rooting. Rooted shoots were transferred to the greenhouse to produce seeds by selfing.

Post-transformation excision of the selectable marker gene

The removal of the selectable marker gene via Cre-mediated *loxP* site-specific recombination was conducted as previously described^{24,51}. Briefly, marker gene excision was achieved by pollination with a homozygous *Cre*-expressing line that contains a gentamycin resistance marker²⁴. The progeny resulting from this cross were selfed, and the next generation was tested for the absence of the *aadA* marker gene and the recombinase gene (linked to the gentamycin resistance gene). To confirm marker gene removal from the plastid genome and outcrossing of the nuclear recombinase cassette, seeds were germinated on MS medium supplemented with spectinomycin (500 µg ml⁻¹), streptomycin (500 µg ml⁻¹) or gentamycin (300 µg ml⁻¹; Fig. 2).

PFGE

PFGE with isolated chloroplasts was conducted to visualize genome size differences between the plastid genomes of the wild type and the ΔIR mutants. All steps of chloroplast isolation were performed at 4 °C or on ice. About 25 g of young leaves were homogenized in 300 ml of grinding buffer (330 mM sorbitol, 50 mM HEPES, 2 mM EDTA, 1 mM MgCl₂, 1 mM sodium pyrophosphate tetrabasic decahydrate, 1 mM 2-mercaptoethanol), and the homogenate was filtered through gauze and Miracloth (Merck). The filtrate was centrifuged at 1,700 g for 2 min, and the pellet was resuspended in 10 ml of grinding buffer and loaded onto Percoll density gradients (40–80% (v/v) Percoll in grinding buffer, 10% (w/v) PEG6000, 45 mM (+)-sodium L-ascorbate). After centrifugation at 5,600 g for 10 min, the intact chloroplasts were collected and washed in wash buffer (350 mM D-sorbitol, 50 mM Tris-HCl pH 7.6). Subsequently, the chloroplasts were collected by centrifugation, resuspended in 300 µl of wash buffer and mixed with 300 µl of 2% (w/v) low-melting-point agarose (InCert agarose, Lonza). The solution was transferred to plug moulds (Bio-Rad), and the solidified plugs were placed in lysis buffer (2% (w/v) Na-laroyl-sarcosine, 0.45 mM EDTA, 10 µg ml⁻¹ proteinase K) and incubated at 50 °C for one day, with regular lysis buffer exchange. The plugs were then washed in TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA) and loaded onto a 1% (w/v) agarose gel (Biozym LE, prepared in 0.5× TBE (40 mM Tris-HCl, 50 mM boric acid, 1.25 mM EDTA)), and the wells were sealed with 1% InCert agarose. The DNA was separated in 0.5× TBE using the CHEF Mapper XA System (Bio-Rad). The PFGE conditions were 6 V cm⁻¹, 14 °C, an angle of 120°, 2.9 to 56 s switch time, 22 h. Finally, the DNA was blotted and analysed by hybridization as described below.

DNA and RNA gel blot analyses

Total DNA was isolated using a cetyltrimethylammoniumbromide-based DNA extraction method⁵². For Southern blot analyses, samples of 2 µg of DNA were digested with the restriction enzyme BspHI and separated by electrophoresis in a 0.8% (w/v) agarose gel. In preparation for blotting, the gel was incubated in depurination solution (0.25 M HCl) for 15 min, in 0.5 M NaOH for 30 min, in denaturation solution (0.5 M NaOH, 1.5 M NaCl) for 30 min and finally in neutralization solution (1 M Tris, 3 M NaCl, pH 6.5 with HCl) for 15 min.

For total plant RNA extraction, the starch content of the sampled leaf material was reduced by adding 500 µl of SDS buffer (100 mM Tris-HCl pH 8.0, 150 mM NaCl, 50 mM EDTA, 1.5% (w/v) SDS, 2% (v/v) 2-mercaptoethanol) to ~100 mg of ground tissue, followed by incubation on ice for 15 min (ref. 53). After centrifugation, the supernatant was mixed with 500 µl of TRIzol reagent (Thermo Fisher Scientific), and RNA extraction was completed following the manufacturer's instructions. Samples of 500 mg of denatured RNA were separated in a 1.3% (w/v) denaturing agarose gel (containing 16% (v/v) formaldehyde (Sigma), prepared in 1× MOPS buffer (100 mM 3-(*N*-morpholino) propanesulfonic acid, 300 mM NaOAc, 1 mM EDTA, pH 7)).

Electrophoretically separated DNA and RNA samples were transferred to Hybond-XL membranes (GE Healthcare) by overnight capillary blotting in 10× SSC buffer for DNA (1.5 M NaCl, 150 mM trisodium citrate, pH 7.0) or 5× SSC buffer for RNA (0.75 M NaCl, 75 mM trisodium citrate, pH 7). Nucleic acids were covalently bound to the membrane by UV crosslinking (0.12 J cm⁻²; UV-crosslinker BLX-254, Vilber Lourmat).

Probes for DNA and RNA blot analyses were generated by PCR amplification and purified from excised gel slices using the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel). The primers for probe amplification are listed in Supplementary Table 1. The probes were radiolabelled with α [³²P]dCTP by random priming (GE Healthcare). DNA and RNA blots were hybridized in Church buffer (1% BSA, 1 mM EDTA, 7% SDS, 500 mM Na₂HPO₄, pH 7.2) at 65 °C overnight. The membranes were exposed to a storage phosphor screen, and radioactive signals were detected in a Typhoon scanner (GE Healthcare).

Relative ribosome abundances were inferred from hybridization signals of plastid and cytosolic rRNA species on the basis of quantification of signal intensities with Image Lab version 6.1.0 (Bio-Rad).

Measurements of photosynthetic electron transport

Parameters of photosynthetic electron transport were determined with the modular version of the Dual-PAM-100 instrument (Heinz Walz GmbH) at 22 °C. After 30 min of dark adaptation, the maximum quantum efficiency of PSII in the dark-adapted state (F_v/F_m) was determined. Then, light response curves of the chlorophyll *a* fluorescence parameters linear electron transport, non-photochemical quenching and *qL* (a measure of the redox state of the PSII acceptor side⁵⁴) were measured⁵⁵. In addition, the donor-side limitation of PSI (Y(ND)) was determined. The measuring times at each actinic light intensity were 150 s under light-limited and 60 s under light-saturated conditions.

Ribosome profiling and RNA-seq

L5, L6 and L7 (numbered from the bottom of the plant, excluding the cotyledons) were harvested individually from greenhouse-grown, six-week-old tobacco plants and immediately frozen in liquid nitrogen. The frozen leaf material was ground, and equal amounts (w/w) of leaf powder of the three leaves harvested (L5, L6 and L7) were pooled. For each plant line, three biological replicates (that is, three individual plants) were analysed.

Ribosome profiling and rRNA removal were performed as previously described⁵⁶, using 360 mg of pooled leaf tissue. For RNA-seq, an aliquot of total RNA was used as input material for the Zymo-Seq RiboFree total RNA library kit (Zymo), following the manufacturer's instructions. Data analysis was performed as previously described⁵⁶. In short, genomes and gene annotations were derived from various sources. The tobacco nuclear genome (Edwards assembly) was derived from the Sol Genomics Network (<https://solgenomics.net/>). The chloroplast and mitochondrial genomes were derived from NCBI (NC_001879.2, and NC_006581.1, respectively). Cutadapt (ref. 57) was used to remove adapters and UMIs. Ribo-seq contaminants (rRNA and transfer RNA) were filtered out using STAR aligner⁵⁸ with the following parameters: outFilterMismatchNoverLmax, 0.1; outReadsUnmapped, Fastx. After filtering, the remaining unmapped reads were simultaneously aligned to the three genomes (nuclear, chloroplast and mitochondrial) using STAR aligner

with the following parameters: `outFilterMismatchNoverLmax, 0.1; alignIntronMax, 8000; outMultimapOrder, Random`. The alignment files were formatted using SAMtools⁵⁹. Mapped reads were counted over all nuclear, chloroplast and mitochondrial genes using FeatureCounts with the following parameters: `M, fraction`⁶⁰. Differential expression analysis was performed using edgeR, using the default parameters (that is, filtering out genes with less than ten counts⁶¹). Data visualization was performed in R v.3.5.3 (ref. 62) using Tidyverse and ggplot2.

Relative quantification of plastid genome copy numbers

Total plant DNA samples were used for all qPCR measurements. Assays were performed in technical triplicates for each of the three biological replicates. qPCR measurements were conducted in a LightCycler 480 Real-Time PCR system (Roche), using the associated software for data analysis. The reactions were carried out in a volume of 5 µl containing 2.5 µl of LightCycler 480 SYBR Green I Master (Roche), 500 nM of each primer and -15 ng of DNA. The PCR conditions were 95 °C for 10 min followed by 45 cycles of 95 °C for 10 s, 60 °C for 10 s and 72 °C for 10 s. For quantification of relative plastid genome copy numbers, a set of three primer pairs was used for each region of the plastid genome (IR, SSC and LSC). In addition, two primer pairs for mitochondrial genes were included. The results were normalized against the *N. tabacum* nuclear *rrn18* (AJ236016.1) gene and the nuclear gene for ribosomal protein L25 (L18908.1). All primers are listed in Supplementary Table 1. Prior to all experiments, the PCR efficiencies were determined for each gene from the slope of DNA serial dilution curves ($E = (10^{-(1/\text{slope})} - 1) \times 100\%$). Primer pairs with an efficiency of 90–100% were used for the analyses.

Microscopy

Microscopic analysis of leaf sections was performed as previously described⁶³. Briefly, a young leaf was cut into small pieces and submerged in fixation buffer (3.5% glutaraldehyde in 0.1 M phosphate buffer, pH 7.2). After vacuum infiltration, the samples were incubated at 4 °C for 1 h. The fixation buffer was then replaced with 0.1 M Na₂E-DTA, and the samples were kept at 4 °C overnight. The final incubation was done at 60 °C for 2 h and with vigorous shaking. After placing the sample between the microscope slide and coverslip, soft pressure was applied to release and separate individual cells. Image stacks along the z axis were captured using a Leica SP8 confocal microscope and the Leica Application Suite X software.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data supporting the findings of this study are available within the paper and its Supplementary Information. Sequences from *N. tabacum* are available through Sol Genomics Network (<https://solgenomics.net/>, Edwards assembly 2017) and NCBI (NC_001879.2 and NC_006581.1). The NGS sequencing data are available under accession number PRJNA1035449 (<http://www.ncbi.nlm.nih.gov/bioproject/1035449>). Source data are provided with this paper.

References

- Sugiura, M. The chloroplast genome. *Plant Mol. Biol.* **19**, 149–168 (1992).
- Bock, R. Structure, function, and inheritance of plastid genomes. *Top. Curr. Genet.* **19**, 29–63 (2007).
- Yukawa, M., Tsudzuki, T. & Sugiura, M. The 2005 version of the chloroplast DNA sequence from tobacco (*Nicotiana tabacum*). *Plant Mol. Biol. Rep.* **23**, 359–365 (2005).
- Magee, A. M. et al. Localized hypermutation and associated gene losses in legume chloroplast genomes. *Genome Res.* **20**, 1700–1710 (2010).
- Blazier, J. C. et al. Variable presence of the inverted repeat and plastome stability in *Erodium*. *Ann. Bot.* **117**, 1209–1220 (2016).
- Palmer, J. D., Nugent, J. M. & Hebron, L. A. Unusual structure of geranium chloroplast DNA: a triple-sized inverted repeat, extensive gene duplications, multiple inversions, and two repeat families. *Proc. Natl Acad. Sci. USA* **84**, 769–773 (1987).
- Chumley, T. W. et al. The complete chloroplast genome sequence of *Pelargonium × hortorum*: organization and evolution of the largest and most highly rearranged chloroplast genome of land plants. *Mol. Biol. Evol.* **23**, 2175–2190 (2006).
- Goulding, S. E., Olmstead, R. G., Morden, C. W. & Wolfe, K. H. Ebb and flow of the chloroplast inverted repeat. *Mol. Gen. Genet.* **252**, 195–206 (1996).
- Stein, D. B., Palmer, J. D. & Thompson, W. F. Structural evolution and flip-flop recombination of chloroplast DNA in the fern genus *Osmunda*. *Curr. Genet.* **10**, 835–841 (1986).
- Day, A. & Madesis, P. DNA replication, recombination, and repair in plastids. *Top. Curr. Genet.* **19**, 65–119 (2007).
- Mower, J. P. & Vickrey, T. L. Structural diversity among plastid genomes of land plants. *Adv. Bot. Res.* **85**, 263–292 (2018).
- Birky, C. W. Jr. & Walsh, J. B. Biased gene conversion, copy number, and apparent mutation rate differences within chloroplast and bacterial genomes. *Genetics* **130**, 677–683 (1992).
- Maier, R. M., Neckermann, K., Igloi, G. L. & Kössel, H. Complete sequence of the maize chloroplast genome: gene content, hotspots of divergence and fine tuning of genetic information by transcript editing. *J. Mol. Biol.* **251**, 614–628 (1995).
- Yukawa, M., Tsudzuki, T. & Sugiura, M. The chloroplast genome of *Nicotiana sylvestris* and *Nicotiana tomentosiformis*: complete sequencing confirms that the *Nicotiana sylvestris* progenitor is the maternal genome donor of *Nicotiana tabacum*. *Mol. Genet. Genom.* **275**, 367–373 (2006).
- Yamane, K., Yano, K. & Kawahara, T. Pattern and rate of indel evolution inferred from whole chloroplast intergenic regions in sugarcane, maize and rice. *DNA Res.* **13**, 197–204 (2006).
- Zhu, A., Guo, W., Gupta, S., Fan, W. & Mower, J. P. Evolutionary dynamics of the plastid inverted repeat: the effects of expansion, contraction, and loss on substitution rates. *N. Phytol.* **209**, 1747–1756 (2016).
- Khakhlova, O. & Bock, R. Elimination of deleterious mutations in plastid genomes by gene conversion. *Plant J.* **46**, 85–94 (2006).
- Palmer, J. D. & Thompson, W. F. Chloroplast DNA rearrangements are more frequent when a large inverted repeat sequence is lost. *Cell* **29**, 537–550 (1982).
- Tsuji, S. et al. The chloroplast genome from a lycophyte (michophyllophyte), *Selaginella uncinata*, has a unique inversion, transpositions and many gene losses. *J. Plant Res.* **120**, 281–290 (2007).
- Weng, M.-L., Blazier, J. C., Govindu, M. & Jansen, R. K. Reconstruction of the ancestral plastid genome in Geraniaceae reveals a correlation between genome rearrangements, repeats, and nucleotide substitution rates. *Mol. Biol. Evol.* **31**, 645–659 (2013).
- Drescher, A., Ruf, S., Calsa Jr, T., Carrer, H. & Bock, R. The two largest chloroplast genome-encoded open reading frames of higher plants are essential genes. *Plant J.* **22**, 97–104 (2000).
- Kikuchi, S. et al. Uncovering the protein translocon at the chloroplast inner envelope membrane. *Science* **339**, 571–574 (2013).
- Svab, Z. & Maliga, P. High-frequency plastid transformation in tobacco by selection for a chimeric *aadA* gene. *Proc. Natl Acad. Sci. USA* **90**, 913–917 (1993).
- Corneille, S., Lutz, K., Svab, Z. & Maliga, P. Efficient elimination of selectable marker genes from the plastid genome by the CRE-lox-site-specific recombination system. *Plant J.* **27**, 171–178 (2001).

25. Greiner, S. et al. Chloroplast nucleoids are highly dynamic in ploidy, number, and structure during angiosperm leaf development. *Plant J.* **102**, 730–746 (2020).
26. Kudoh, H. & Sonoike, K. Irreversible damage to photosystem I by chilling in the light: cause of the degradation of chlorophyll after returning to normal growth temperature. *Planta* **215**, 541–548 (2002).
27. Petersen, K., Schöttler, M. A., Karcher, D., Thiele, W. & Bock, R. Elimination of a group II intron from a plastid gene causes a mutant phenotype. *Nucleic Acids Res.* **39**, 5181–5192 (2011).
28. Rogalski, M., Schöttler, M. A., Thiele, W., Schulze, W. X. & Bock, R. Rpl33, a nonessential plastid-encoded ribosomal protein in tobacco, is required under cold stress conditions. *Plant Cell* **20**, 2221–2237 (2008).
29. Fleischmann, T. T. et al. Nonessential plastid-encoded ribosomal proteins in tobacco: a developmental role for plastid translation and implications for reductive genome evolution. *Plant Cell* **23**, 3137–3155 (2011).
30. Gao, Y. et al. Chloroplast translational regulation uncovers nonessential photosynthesis genes as key players in plant cold acclimation. *Plant Cell* **34**, 2056–2079 (2022).
31. Hojka, M. et al. Inducible repression of nuclear-encoded subunits of the cytochrome b6f complex in tobacco reveals an extraordinarily long lifetime of the complex. *Plant Physiol.* **165**, 1632–1646 (2014).
32. Schöttler, M. A., Toth, S. Z., Boulouis, A. & Kahlau, S. Photosynthetic complex stoichiometry dynamics in higher plants: biogenesis, function, and turnover of ATP synthase and the cytochrome b6f complex. *J. Exp. Bot.* **66**, 2373–2400 (2015).
33. Strand, D. D., D'Andrea, L. & Bock, R. The plastid NAD(P) H dehydrogenase-like complex: structure, function and evolutionary dynamics. *Biochem. J.* **476**, 2743–2756 (2019).
34. Kikuchi, S. et al. A Ycf2–FtsHi heteromeric AAA-ATPase complex is required for chloroplast protein import. *Plant Cell* **30**, 2677–2703 (2018).
35. Tiller, N. & Bock, R. The translational apparatus of plastids and its role in plant development. *Mol. Plant* **7**, 1105–1120 (2014).
36. Walter, M. et al. Knockout of the plastid RNase E leads to defective RNA processing and chloroplast ribosome deficiency. *Plant J.* **64**, 851–863 (2010).
37. Tiller, N. et al. The plastid-specific ribosomal proteins of *Arabidopsis thaliana* can be divided into non-essential proteins and genuine ribosomal proteins. *Plant J.* **69**, 302–316 (2012).
38. Kössel, H. et al. in *Molecular Form and Function of the Plant Genome* (eds Vloten-Doting, L. V. et al.) 183–198 (Plenum, 1985).
39. Zoschke, R., Liere, K. & Börner, T. From seedling to mature plant: *Arabidopsis* plastidial genome copy number, RNA accumulation and transcription are differentially regulated during leaf development. *Plant J.* **50**, 710–722 (2007).
40. Preuten, T. et al. Fewer genes than organelles: extremely low and variable gene copy numbers in mitochondria of somatic plant cells. *Plant J.* **64**, 948–959 (2010).
41. Isono, K., Niwa, Y., Satoh, K. & Kobayashi, H. Evidence for transcriptional regulation of plastid photosynthesis genes in *Arabidopsis thaliana* roots. *Plant Physiol.* **114**, 623–630 (1997).
42. Scharff, L. B. & Bock, R. Synthetic biology in plastids. *Plant J.* **78**, 783–798 (2014).
43. Boehm, C. R. & Bock, R. Recent advances and current challenges in synthetic biology of the plastid genetic system and metabolism. *Plant Physiol.* **179**, 794–802 (2019).
44. del Solar, G. & Espinosa, M. Plasmid copy number control: an ever-growing story. *Mol. Microbiol.* **37**, 492–500 (2000).
45. Murashige, T. & Skoog, F. A revised medium for rapid growth and bio assays with tobacco tissue culture. *Physiol. Plant.* **15**, 473–497 (1962).
46. Hoelscher, M. P. et al. Expression strategies for the efficient synthesis of antimicrobial peptides in plastids. *Nat. Commun.* **13**, 5856 (2022).
47. Lohse, M., Drechsel, O. & Bock, R. OrganellarGenomeDRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. *Curr. Genet.* **52**, 267–274 (2007).
48. Lohse, M., Drechsel, O., Kahlau, S. & Bock, R. OrganellarGenomeDRAW—a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Res.* **41**, W575–W581 (2013).
49. Greiner, S., Lehwark, P. & Bock, R. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Res.* **47**, W59–W64 (2019).
50. Bock, R. Transgenic plastids in basic research and plant biotechnology. *J. Mol. Biol.* **312**, 425–438 (2001).
51. Oey, M., Lohse, M., Scharff, L. B., Kreikemeyer, B. & Bock, R. Plastid production of protein antibiotics against pneumonia via a new strategy for high-level expression of antimicrobial proteins. *Proc. Natl Acad. Sci. USA* **106**, 6579–6584 (2009).
52. Doyle, J. J. & Doyle, J. L. Isolation of plant DNA from fresh tissue. *Focus* **12**, 13–15 (1990).
53. Dutta, S., Muthusamy, V., Chhabra, R., Zunjare, R. U. & Hossain, F. Two-step method for isolation of high-quality RNA from stored seeds of maize rich in starch. *3 Biotech* **10**, 433 (2020).
54. Kramer, D. M., Johnson, G., Kiirats, O. & Edwards, G. E. New fluorescence parameters for the determination of QA redox state and excitation energy fluxes. *Photosynth. Res.* **79**, 209–218 (2004).
55. Schöttler, M. A., Flügel, C., Thiele, W., Stegemann, S. & Bock, R. The plastome-encoded Psaj subunit is required for efficient photosystem I excitation, but not for plastocyanin oxidation in tobacco. *Biochem. J.* **403**, 251–260 (2007).
56. Ting, M. K. Y. et al. Guidelines for performing ribosome profiling in plants including structural analysis of rRNA fragments. Preprint at *bioRxiv* <https://doi.org/10.1101/2023.11.16.567332> (2023).
57. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10–12 (2011).
58. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
59. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
60. Liao, Y., Smyth, G. K. & Shi, W. FeatureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
61. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2009).
62. R Core Team. R: A Language and Environment for Statistical Computing. <http://www.R-project.org/> (R Foundation for Statistical Computing, 2019).
63. Lu, Y. et al. Horizontal transfer of a synthetic metabolic pathway between plant species. *Curr. Biol.* **27**, 3034–3041 (2017).
64. Strittmatter, G. & Kössel, H. Cotranscription and processing of 23S, 4.5S and 5S rRNA in chloroplasts from *Zea mays*. *Nucleic Acids Res.* **12**, 7633–7647 (1984).

Acknowledgements

We thank K. Jäger, C. Voigt and W. Thiele for excellent technical assistance; L. Yaneva-Roder for help with PFGE; V. Loiacono for help with the RNA gel blots; Q. Luo, B. Ren and O. A. Sandoval Ibanez (all MPI-MP) for help with the DNA extraction and PCR assays; and the MPI-MP GreenTeam for plant cultivation. This research was financed by the Max Planck Society, a grant from the European Research Council under the European Union's Horizon 2020 research and

innovation programme (ERC-ADG-2014; grant agreement no. 669982) to R.B., and a grant from the Deutsche Forschungsgemeinschaft (ZO 302/5-1) to R.Z.

Author contributions

C.K., J.F., S.R. and R.B. designed the research. C.K., C.R.B., J.L., A.P.H. and M.A.S. performed the experiments. All authors (C.K., C.R.B., J.L., M.K.Y.T., A.P.H., J.F., S.R., M.A.S., R.Z. and R.B.) participated in data evaluation. R.B. conceived of the study and wrote the manuscript, with input from C.K. and M.A.S. All authors commented on the manuscript draft.

Funding

Open access funding provided by Max Planck Society.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41477-024-01709-9>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41477-024-01709-9>.

Correspondence and requests for materials should be addressed to Ralph Bock.

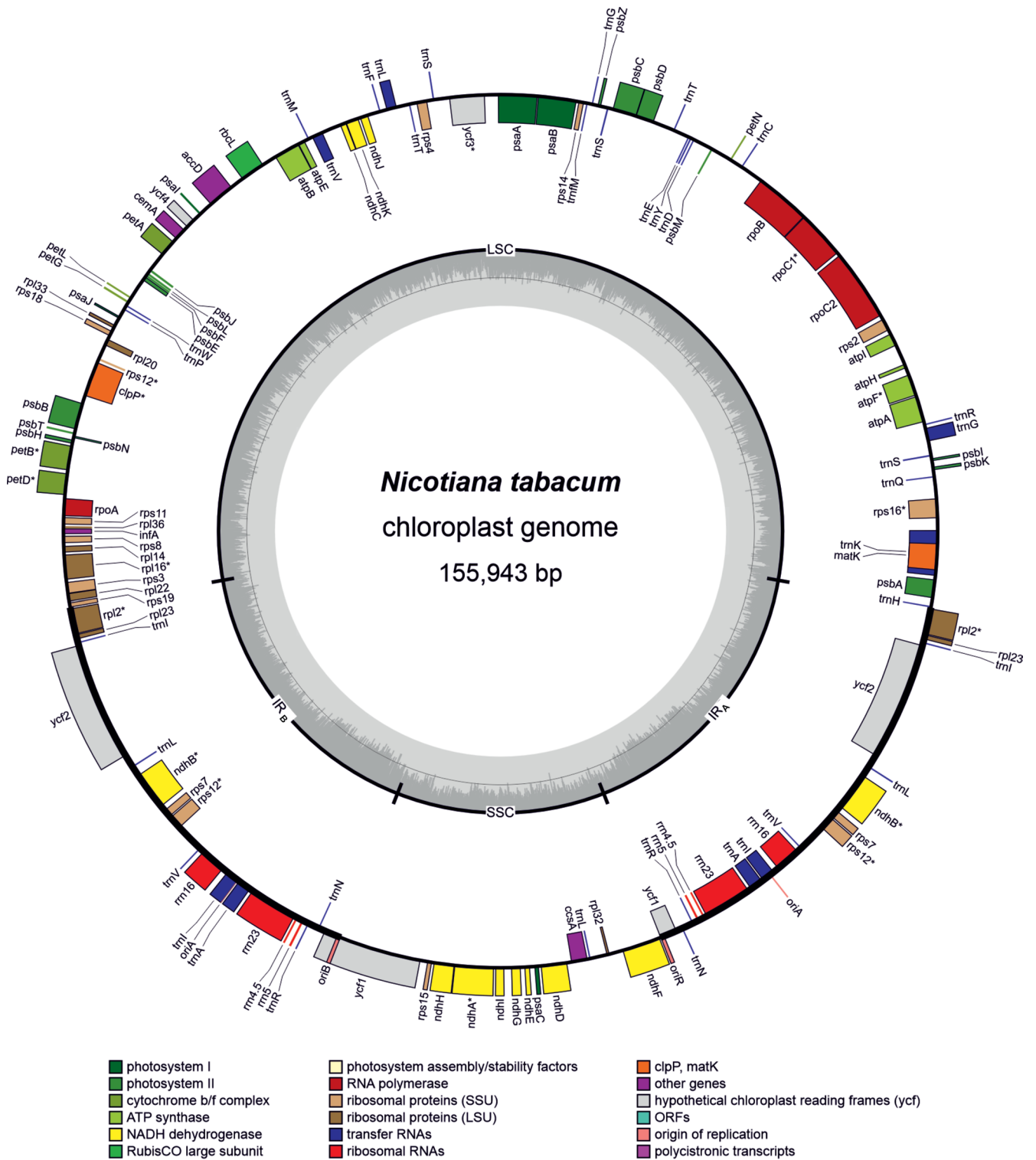
Peer review information *Nature Plants* thanks Shin-ichi Arimura and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

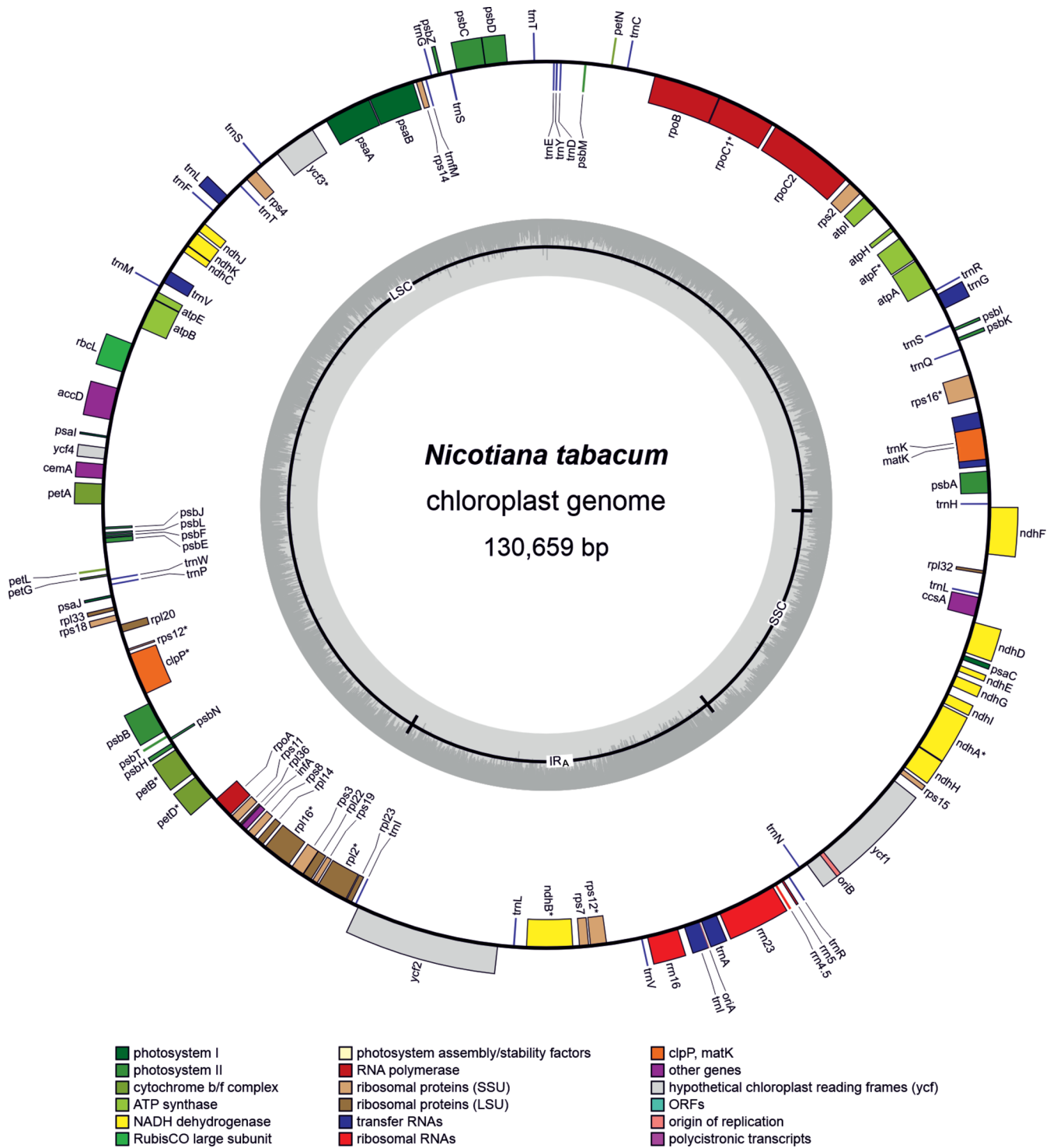
Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024

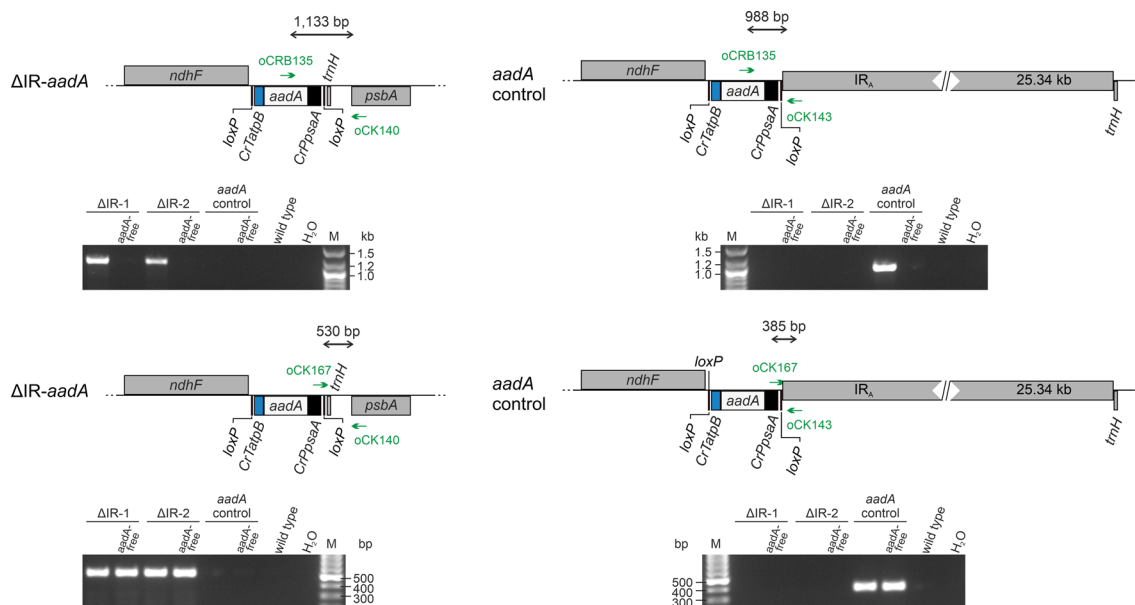


Extended Data Fig. 1 | Physical map of the wild-type plastid genome of tobacco (*Nicotiana tabacum*). Genes in the inner circle are transcribed clockwise, genes in the outer circle are transcribed counterclockwise. Intron-

containing genes are denoted by asterisks. Genes are color-coded according to functional classes (47). The grey graph in the center indicates the GC content (48). The map was drawn with OGDRAW version 1.3.1 (49).

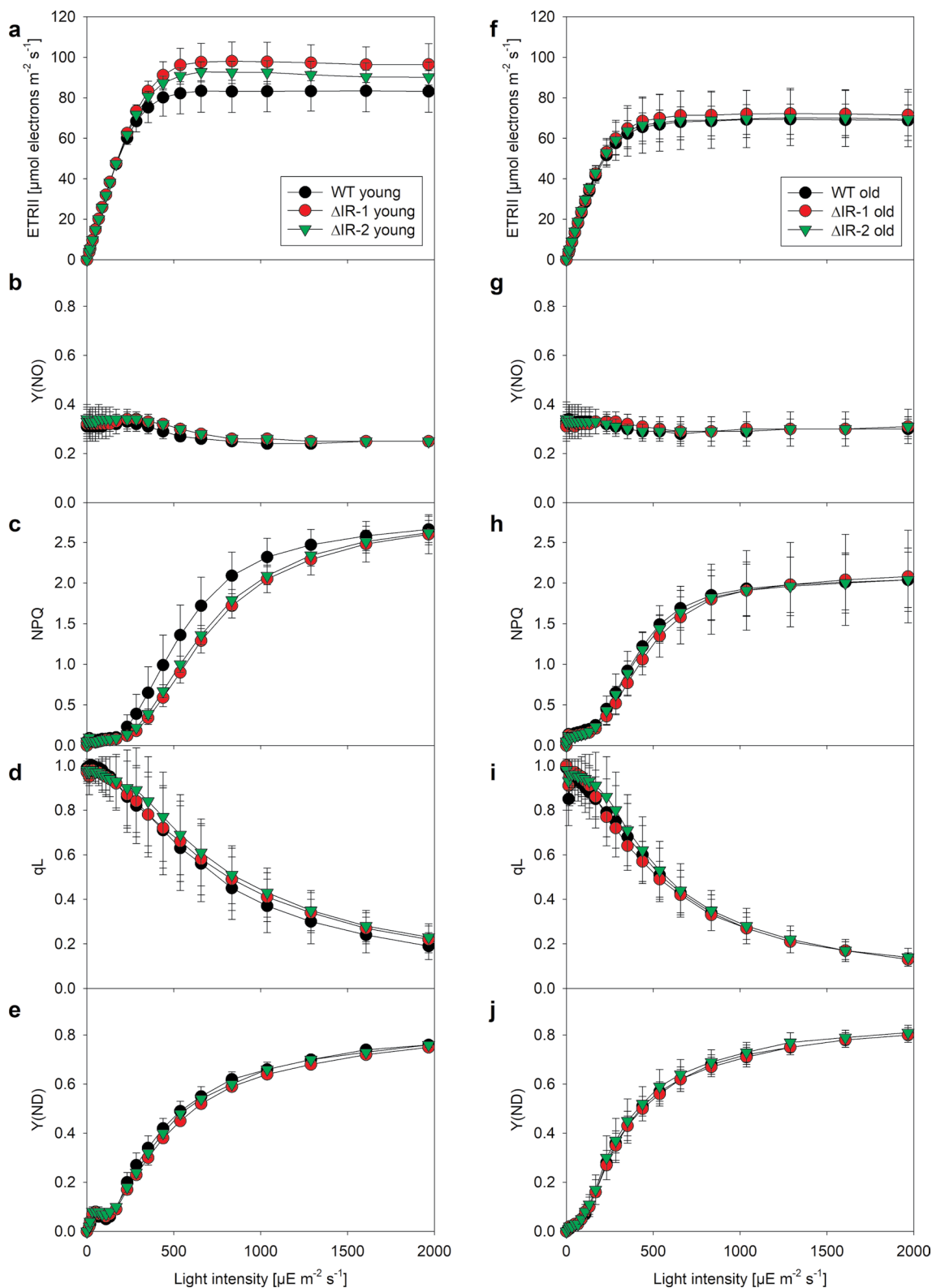


Extended Data Fig. 2 | Physical map of the tobacco plastid genome in the IR deletion plants. Note the substantial reduction in genome size compared to the wild-type genome (cf. Extended Data Fig. 1).



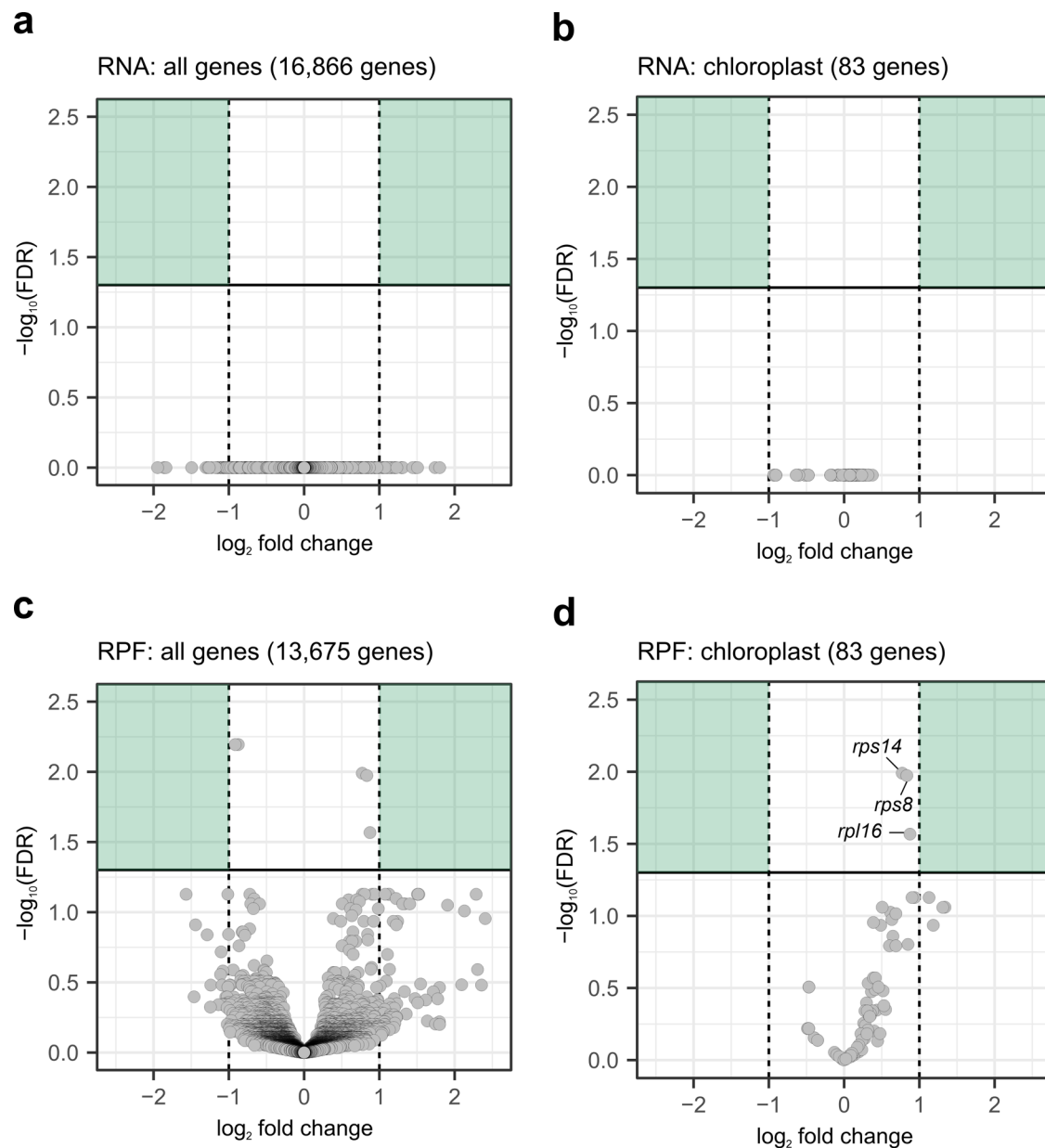
Extended Data Fig. 3 | PCR analysis to test for the absence of IR_A from the ΔIR plants. Primer pairs flanked the IRA excision site, or the border between SSC and IRA as a control. The binding sites of the PCR primers, along with the expected amplicon sizes, are indicated. The upper panel shows the PCR results for a primer binding within the *aadA* gene, while the bottom panel displays the results for a primer binding within the *loxP* site. The *aadA* control line served as a

positive control for amplification of the SSC / IR_A border-spanning amplicon. The wild type served as negative control. Note that, due to the abundant presence of promiscuous plastid DNA in the nuclear genome, one of the two PCR primers needs to be derived from either the *aadA* sequence or the *loxP* sequence (that do not occur in the nuclear genome). M, DNA size marker; H_2O , buffer control (with water instead of DNA added).



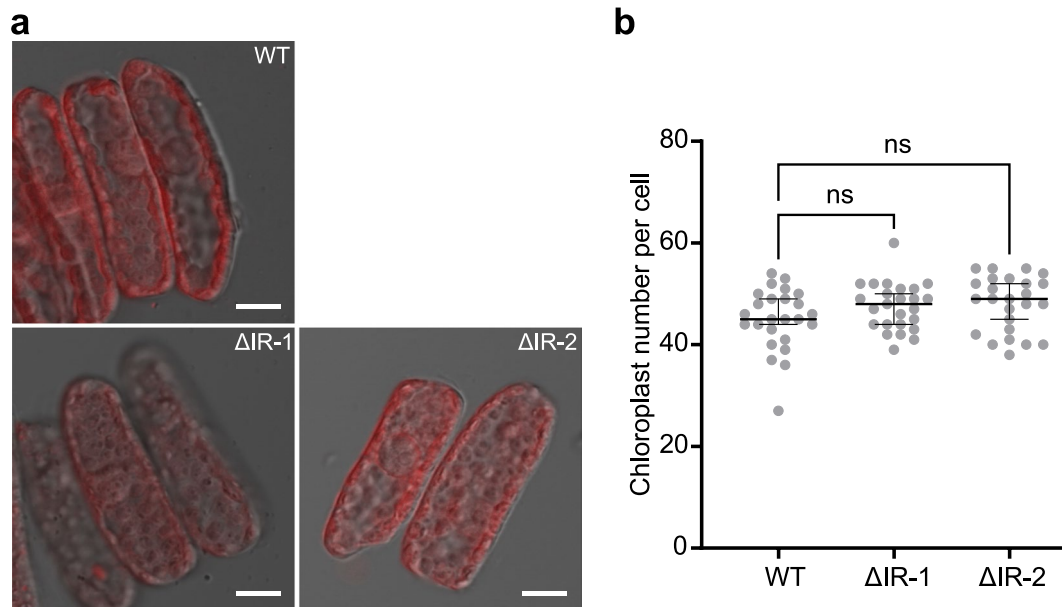
Extended Data Fig. 4 | Photosynthetic parameters of high-light acclimated wild-type and Δ IR plants. (a–e) Photosynthetic parameters analyzed in the youngest leaves (of ≥ 17.5 cm length) that fully developed under high-light conditions. (f–j) Photosynthetic parameters analyzed in older leaves. The fourth leaf below the young leaf measured in (a–e) was analyzed. Unlike the young leaf, this leaf had already developed prior to the high-light shift. Shown are the chlorophyll-a fluorescence parameters linear electron transport (ETR_{II};

a, f), Y(NO), the non-regulated dissipation of excitation energy in PSII (b, g), the non-photochemical quenching (NPQ; c, h), and the redox state of the PSII acceptor side (qL; d, i). Y(ND), the donor-side limitation of PSI was determined from difference transmittance measurements of P700 (e, j). Data are presented as mean values with error bars indicating the standard deviation (for wild-type plants and Δ IR-2 plants $n = 7$, for Δ IR-1 plants $n = 6$; individual data points for the three plant lines are provided as Source Files).



Extended Data Fig. 5 | Analysis of transcript accumulation and translational output in wild-type and Δ IR plants. Leaves 5, 6 and 7 from greenhouse-grown plants (cf. Figure 4a) were harvested and pooled at equal mass for extraction of total RNA and isolation of ribosome footprints. The data shown summarize results of three biological replicates (that is, three individual plants) for each of the two independent Δ IR lines. **(a)** Volcano plot of the FDR values versus fold-changes in transcript abundance for all genes analyzed (in the nuclear, plastid and mitochondrial genomes; Supplementary Dataset 1). **(b)** Volcano plot derived from the results in (a), but only showing plastid-encoded protein-coding genes. **(c)** Volcano plot showing the FDR values versus fold-changes in ribosome

footprint abundance (RPF: ribosome-protected mRNA fragments) for all genes. **(d)** Volcano plot derived from results in (c), but only showing the plastid-encoded protein-coding genes. Significantly changed plastid-encoded genes with an FDR < 0.05 are indicated with their respective gene names. Green colored rectangles mark areas with an FDR < 0.05 and a fold change below 0.5 or above 2.0. Note that none of the detected genes shows a more than twofold up- or down-regulation with statistical significance. Data were normalized to all obtained reads for nuclear, chloroplast and mitochondrial genes. Differential gene expression was analyzed with the R package edgeR.



Extended Data Fig. 6 | Plastid genome size does not affect chloroplast size and number per cell. (a) Confocal microscopic images showing the chlorophyll fluorescence of chloroplasts in mesophyll cells of the wild type (WT) and two ΔIR mutant lines. z-stacks of these images were used for quantification of chloroplast numbers per cell. Scale bars: 10 μm . (b) Chloroplast numbers per cell. The

black bar indicates the median, shown with the corresponding 95% confidence interval. Results for each individual cell analyzed are shown as grey dots ($n = 25$). Chloroplast numbers are not significantly different between the three plant lines analyzed (ns; unpaired two-tailed Student's *t*-test, $\Delta IR-1 P = 0.111$, $\Delta IR-2 P = 0.072$).

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a | Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection | The qPCR data was obtained using LightCycler 480 and the corresponding software in version 1.5.0 SP4.

Data analysis | Signal intensities of RNA blots were quantified using Image Lab version 6.1.0 (Bio-Rad).
Data was analyzed with Microsoft Excel 2019 MSO v16.0.10377.20023.
Data analysis of RNA-seq and ribosome profiling data was performed with cutadapt 2.10 with Python 3.7.6; STAR v2.7.7a; samtools v0.1.19; featureCounts v2.0.1; edgeR v3.38.4; ggplot2 v3.4.2; tidyverse v2.0.0 with R v3.5.3.
Chloroplast genomes were visualized using OGDRAW v1.3.1.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The data supporting the findings of this study are available within the paper and its supplementary information files.

Sequences from *Nicotiana tabacum* are available through Sol Genomics Network (<https://solgenomics.net/>, Edwards assembly 2017) and NCBI (NC_001879.2 and NC_006581.1)

The NGS sequencing data are available under accession number PRJNA1035449 (<http://www.ncbi.nlm.nih.gov/bioproject/1035449>).

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	<input type="text" value="n/a"/>
Reporting on race, ethnicity, or other socially relevant groupings	<input type="text" value="n/a"/>
Population characteristics	<input type="text" value="n/a"/>
Recruitment	<input type="text" value="n/a"/>
Ethics oversight	<input type="text" value="n/a"/>

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample size calculations were performed. The analysis of two independently obtained transplastomic lines is standard in the field, as all transplastomic lines are genetically identical due to homologous recombination. All analyses included three independent biological replicates to allow statistical analysis.
Data exclusions	No data were excluded from the analysis.
Replication	The transformation to obtain IR deletion was performed once, successfully generating several independent IR deletion lines. Of these, two lines were used for in-depth analysis. For each condition, three independent biological replicates were analyzed and statistically assessed. In qPCR experiments, each biological replicate was measured in technical triplicate. All attempts of replication are shown in the manuscript, and were successfully performed.
Randomization	For growth under different environmental conditions, plants from different lines were equally distributed in the growth chamber to ensure equal growth conditions. Under all conditions, all genotypes were equally represented, i.e., the same number of plants per transplastomic line and wild type were grown simultaneously under the same growth conditions.
Blinding	Data collection and analysis was not blinded as no subjective parameters were measured that would have benefited from blinding.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

Methods

- n/a Involved in the study
- Antibodies
- Eukaryotic cell lines
- Palaeontology and archaeology
- Animals and other organisms
- Clinical data
- Dual use research of concern
- Plants

- n/a Involved in the study
- ChIP-seq
- Flow cytometry
- MRI-based neuroimaging

Plants

Seed stocks

Nicotiana tabacum cv. Petit Havana, kept as lab stock for 20 years.

Novel plant genotypes

Transplastomic IR deletion lines were obtained by biolistic transformation via homologous recombination and spectinomycin selection. For post-transformation elimination of the selectable marker gene a previously described homozygous Cre-expressing line was used for crossings (Corneille et al. 2001).
Two independent lines were analyzed in this study, both from the T4 generation.

Authentication

Homoplasmy of transplastomic lines was confirmed by RFLP analysis. The excision sites of the IR and the selectable marker were verified by sequencing of PCR products.