Biparental inheritance of chloroplasts is controlled by lipid biosynthesis

Johanna Sobanski[1], Patrick Giavalisco[1], Axel Fischer[1], Dirk Walther[1], Mark Aurel Schöttler[1], Tommaso Pellizzer[1], Hieronim Golczyk[2], Toshihiro Obata[3], Ralph Bock[1], Barbara B. Sears[3], and Stephan Greiner[1]

[1]Max Planck Institute of Molecular Plant Physiology, Am Mühlenberg 1, 14476 Potsdam-Golm, Germany

[2]Department of Molecular Biology, Institute of Biotechnology, John Paul II Catholic University of Lublin, Konstantynów 1I, 20-708, Poland

[3]Center for Plant Science Innovation and Department of Biochemistry, University of Nebraska-Lincoln, 1901 Vine Street, Lincoln, Nebraska 68588, USA

[4]Department of Plant Biology, Michigan State University, East Lansing, MI 48824-1312, USA

Abstract

In most eukaryotes, organellar genomes are transmitted preferentially by the mother, but the reasons underlying this fundamental biological principle are far from being understood. It is believed that biparental inheritance promotes competition between the cytoplasmic organelles and allows the spread of selfish cytoplasmic elements. Those can be, for example, fast replicating chloroplasts (plastids) that are incompatible with the hybrid nuclear genome and therefore maladaptive.

Here we show, that plastid competition is a metabolic phenotype determined by extremely rapidly evolving regions in the plastid genome of the evening primrose *Oenothera*. Polymorphisms in repeats of the regulatory region of *accD* (the only plastid-encoded subunit of the acetyl-CoA carboxylase, which catalyzes the first step of lipid biosynthesis), are responsible for the differences in competitive behavior of different plastid genotypes. These cytoplasmic drive loci change lipid synthesis and consequently lipid profiles of the plastid envelope membrane. This most likely determines plastid division and/or turn-over rates and hence competitiveness. Since plastid competition can result in uniparental inheritance (through elimination of the "weak" plastid) or biparental inheritance (when two similarly "strong" plastids are transmitted), this work uncovers for the first time a genetic determinant of organelle inheritance.

Main Text

When organelles are transmitted to the progeny by both parents, evolutionary theory predicts that the organelle genomes will go into competition. As they do not undergo sexual recombination, one organelle genome will "try" to overgrow the other one [1-5]. In a population, the ensuing arms race can lead to the spread of so-called selfish cytoplasmic elements that potentially harm the host cell. The mechanisms and molecular factors underlying this phenomenon are elusive, yet there is solid evidence available for a widespread presence of competing organelles in nature [1]. Nonetheless, only a few cases have been worked out genetically in some detail in model organisms [6-9]. One of them is the evening primrose (genus *Oenothera*). In this genus, biparental plastid inheritance is the rule and the system is a prime example of naturally occurring selfish chloroplasts. Based on extensive crossing studies, five genetically distinguishable chloroplast genome (plastome) types were shown to exist. Those were designated by Roman numerals (I to V) and grouped into three classes according to their inheritance strength or assertiveness rates in crosses (strong: plastomes I and III, intermediate: plastome II, weak: plastomes IV and V). The plastome types were initially identified based on their (in)compatibility with certain nuclear genomes. Strong plastomes provide hitchhiking opportunities for incompatible and maladaptive chloroplasts, especially since interspecific hybridization is frequent in the genus. It was further shown that the loci, which determine the differences in competitive ability

of the chloroplast, are encoded by the chloroplast genome itself [1,2,7,10,11].

To pinpoint the underlying genetic determinants, we developed a novel association mapping approach that correlates local sequence divergence to a phenotype. We analyzed 14 complete chloroplast genomes from *Oenothera* wild type lines, whose inheritance strength had been previously classified in exhaustive crossing analyses [6,7,12]. This enabled us to correlate the experimentally determined inheritance strengths to sequence divergence in a given alignment window (Materials and Methods; Supplementary Text for details). Both applied correlation metrics (Pearson's and Spearman's) associated the genetic determinants of inheritance strength with four discrete sites: the regulatory region of the fatty acid biosynthesis gene *accD* (promoter, 5'-UTR and protein N-terminus), the *origin of replication B* (*oriB*), *ycf1* and *ycf2* (including its promoter/5'-UTR; Fig. 1, Data S1, and Supplementary Text). The *ycf1* and *ycf2* genes are two ORFs of unknown function, although *ycf1* (or *tic214*) has tentatively been identified as part of the chloroplast protein import machinery [13]. The detected polymorphisms represent larger insertions/deletions (indels), which are in frame in all coding sequences (Figs. S1 and S2, Data S2, Supplementary Text). Several other polymorphisms in intergenic spacers of photosynthesis and/or chloroplast translation genes were correlated with assertiveness rates (see Fig. 1, Data S1, and Supplementary Text for details). However, these are unlikely to affect chloroplast inheritance, since no such impact has been observed for chlorophyll-deficient mutants of *Oenothera* [7,14] (Supplementary Text). By contrast, the large ORFs of unknown function, the origins of replication and *accD,* which is involved in catalyzing the rate limiting step of lipid biosynthesis [15], are serious candidates.

Due to the lack of measurable sexual recombination in seed plant chloroplast genomes [1], our correlation mapping method associates polymorphisms that are fixed in the slow plastome type IV with weak inheritance strength, regardless of their functional relevance (Supplementary Text). Hence, to test whether the loci predicted in wild type plastomes affect inheritance strength, we conducted a genetic screen for weak chloroplast mutants derived from a strong chloroplast genome by employing a *plastome mutator* (*pm*) allele (Materials and Methods). The *pm*-based mutagenesis approach yields indels in repetitive regions [16], similar to those identified by our association mapping (see Supplementary Text for details). This led to the isolation of 24 plastome I variants with altered inheritance strength (Fig. 2, Supplementary Text). Since we selected for green and photosynthetically competent chloroplast in the mutagenesis (Materials and Methods for details), none of those mutants differed from the wild type in their photosynthetic parameters, chloroplast size or chloroplast volume per cell. Also, the plants also did not display any growth phenotype (Figs. S3-S5, Supplementary Text). Analysis of a subset of 18 variants, spanning the range of observed variation in inheritance strength, confirmed the relevance of *accD* in chloroplast inheritance. Moreover, sites in the *ycf2* gene also seem

to be associated with inheritance strength (Figs. S1, S2, S6-S10; Supplementary Text). Similar to the natural variation observed in the wild types, mutations detected in the coding regions of these genes were in-frame indels (Figs. S1 and S2, Data S2). This analysis did not support an involvement *oriB* or *ycf1* in the inheritance phenotype (Supplementary Text).

Interestingly, in the mutants, the strong sequence variation of the *oriB* region is not associated with inheritance strength. Furthermore, the second replication origin (*oriA*) was found to be nearly identical within all sequenced wild type or mutant plastomes. This finding argues against plastid DNA (ptDNA) replication *per se* being responsible for differences in chloroplast competitiveness. This conclusion is in line with previous analyses of the *Oenothera* replication origins, which had suggested that their variability does not correlate with the competitive strength of the plastids [17,18] (also see Supplementary Text). We further confirmed this by determining the relative ptDNA amounts of chloroplasts with different inheritance strengths in a constant nuclear background. No significant variation of ptDNA amounts was observed over a developmental time course in these lines, thus excluding ptDNA stability and/or turnover as a potential mechanism (Fig. S11 and Supplementary Text). Moreover, no significant differences in nucleoid numbers per chloroplast or in nucleoid morphology was observed, as judged from DAPI staining (Figs. S12 and S13, Supplementary Text).

Next, we conducted a more detailed analysis of *accD* and *ycf2*. In a constant nuclear background, the weak wild type plastome IV appeared to be an *accD* overexpressor when compared to the strong wild type plastome I, as judged from northern blot analyses. However, this overexpression, could not be detected in the plastome I variants, which differ in their competitive ability. Similar results were obtained for *ycf2*. Here, a band of about 7 kb, reflecting the predicted size of the full-length transcript, is observable (Fig. S14 and Supplementary Text). Interestingly, lower bands, probably reflecting transcript processing and/or degradation intermediates, differ between the strong and the weak wild types I and IV. They can be further correlated with competitive ability in the plastome I variants.

Since these analyses did not allow conclusions about the functionality of AccD or Ycf2 in our lines, we decided to determine the acetyl-CoA carboxylase (ACCase) activity in isolated chloroplasts. It appeared that mutations/polymorphisms in the reading frame of *accD* influence ACCase enzymatic activity. Surprisingly, mutations/polymorphisms in *ycf2* also have an influence on ACCase activity, as revealed by lines that are not affected by mutations in *accD* (Fig. 3A). The molecular nature of this functional connection between Ycf2 and ACCase activity is currently unclear. A simple correlation of ACCase activity with the competitive ability of plastids is not present, but alterations in the earliest step of fatty acid biosynthesis can conceivably result in various changes in lipid metabolism (see also Supplementary Text).

To examine the alterations in lipid biosynthesis, we determined the lipid composition of seedlings harboring strong and weak wild type chloroplasts as well as our variants differing in inheritance strength (Table S1; Supplementary Text). Then, we employed a LASSO regression model to predict competitive ability of a given chloroplast. Indeed, out of 102 lipids analyzed, 20 predictive ones for inheritance strengths were identified (Fig. 3C, Table S2, and Supplementary Text). Strikingly, the signal is independent of greening (i.e. an intact thylakoid membrane system and the photosynthetic capacity), which is in line with the genetic data (Supplementary Text). This result hints at the chloroplast envelope determining assertiveness rates, a view that is also supported by the fact that the 20 predictive lipids include the lipids abundant in the chloroplast envelope such as DGDG, PG, PC and PE (although statistical significance could not be established due to low numbers; Table S2 and S3; Supplementary Text). Especially in the lipid class PC, which is the dominant lipid class in the chloroplast outer envelope [19], 4 out of 13 detected lipids were found to be predictive, making it very likely that the lipid composition of the envelope membrane affects chloroplast competition. A possible explanation could be that strong and weak chloroplasts differ in division rates, for example, due their effectiveness to recruit chloroplast division rings, which are anchored by membrane proteins [20]. Alternatively, chloroplast stability might depend on envelope membrane composition.

The present work can explain plastid competition following biparental inheritance from a mechanistic perspective and, for the first time, points to genetic loci that appear to be responsible for these differences. Moreover, since chloroplast competition can result in uniparental inheritance through the elimination of weak chloroplasts [7,14,21], the mechanistic explanation can be extended to uniparental transmission. Arguably, the most surprising finding from our work is the discovery that chloroplast competition in evening primroses is essentially a metabolic phenotype and not connected to ptDNA replication or copy number [22]. The underlying molecular loci are rapidly evolving genes. Due to their highly repetitive sequence composition, they are very sensitive to replication slippage, the major mechanism of spontaneous chloroplast mutations [23,24] (also see Supplementary Text). This result is somewhat reminiscent of recent findings in *Drosophila* where sequence variation in the non-coding regulatory region of the mitochondrial genome, containing the origins of replication, was associated with different competitive abilities. Similar to *Oenothera*, these sequences are highly repetitive, hypervariable and contribute to cytoplasmic drive. In *Drosophila*, they are among the most divergent ones in Metazoa pointing to their positive selection [9,25] (also see Supplementary Text). Our analyses show that, due to their high mutation rates, these loci can evolve and become fixed in a population very quickly. In *Oenothera*, the divergence time between the strong plastome I and the weak plastome IV is estimated to be approximately 830,000 years. Moreover, the strong plastomes I and III do not form a phylogenetic clade. The strong plastome I clusters with the intermediate plastome

II, with III as outgroup. The divergence time of II and III, however, is only about 420,000 years [24]. Hence, evolution and fixation of selfish cytoplasmic drive loci happened twice independently within a very short time frame.

Acknowledgments

We thank Werner Dietrich and Wilfried Stubbe who compiled the comprehensive *Oenothea* collection that allowed us to use the original material of the Renner school. John B. Ohlrogge is acknowledged for advice with the ACCase activity assay and Michael Tillich for help with the MassArray® system. We further thank the MPI-MP GreenTeam for their support, the Ornamental Plant Germplasm Centre for providing the line chicaginensis de Vries and Liliya Yaneva-Roder for technical assistance. This research was supported by the Max Planck Society to S.G., P.G., D.W., M.A.S., and R.B. DAPI fluorescent patterns were analysed with microscopy equipment funded by the Polish National Science Centre (NCN) to H.G. (2015/19/B/N22/01692).

Author Contributions

J.S. performed the main experimental work. P.G., A.F., D.W., M.A.S., H.G., T.O., T.P., B.B.S, and S.G. provided supportive data. A.F. and S.G. developed the correlation mapping approach. All authors analysed and discussed the data. B.B.S. and S.G. designed the study. S.G. and J.S. wrote the manuscript. P.G., A.F., D.W., M.A.S., H.G., R.B. and B.B.S. participated in writing.

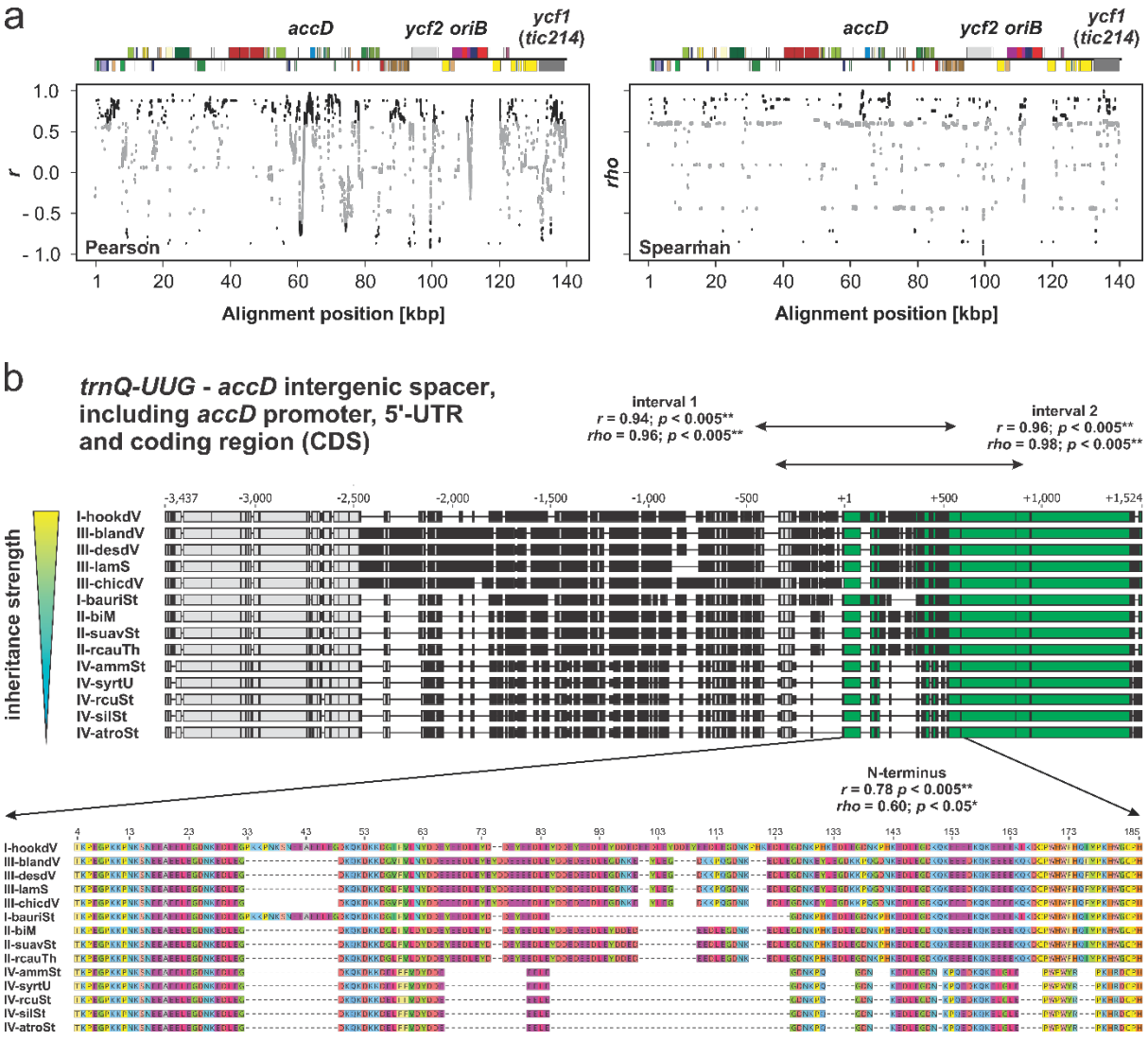


Fig. 1. Correlation mapping to identify chloroplast loci for inheritance strength. (**A**) Spearman and Pearson correlation to inheritance strength plotted against alignment windows of the wild type plastomes. Relevant genes or loci with significant correlation are indicated in the linear plastome maps above. Significant correlations ($p < 0.05$) are shown in black. Correlations to *k*-means classes are shown. For details see Materials and Methods, Supplementary Text and Main Text. (**B**) Correlation to inheritance strength at the *accD* region in the wild type plastomes. Individual sequences are sorted according to their competitive ability. Disagreement to the consensus in a given alignment windows is indicated in black. Upper panel: Alignment of the *trnQ-UUG* - *accD* intergenic spacer (- 3,437 to -1) and the *accD* gene, including promoter, 5'-UTR and coding region (CDS). The *accD* CDS, starting from +1, is highlighted in green. Regions marked by "interval 1" and "interval 2" display nearly absolute correlation to inheritance strength (see Supplementary Text and Data S1 for details). Note that these sequence intervals span the promotor region, the 5'-UTR and the 5'-end of *accD*. All three are considered to play a regulatory role [15,26-29]. Lower panel: Amino acid sequence of the AccD N-terminus and correlation to inheritance strength. Most variation in the sequence is conferred by repeats encoding glutamic acid-rich domains (also see Supplementary Text).

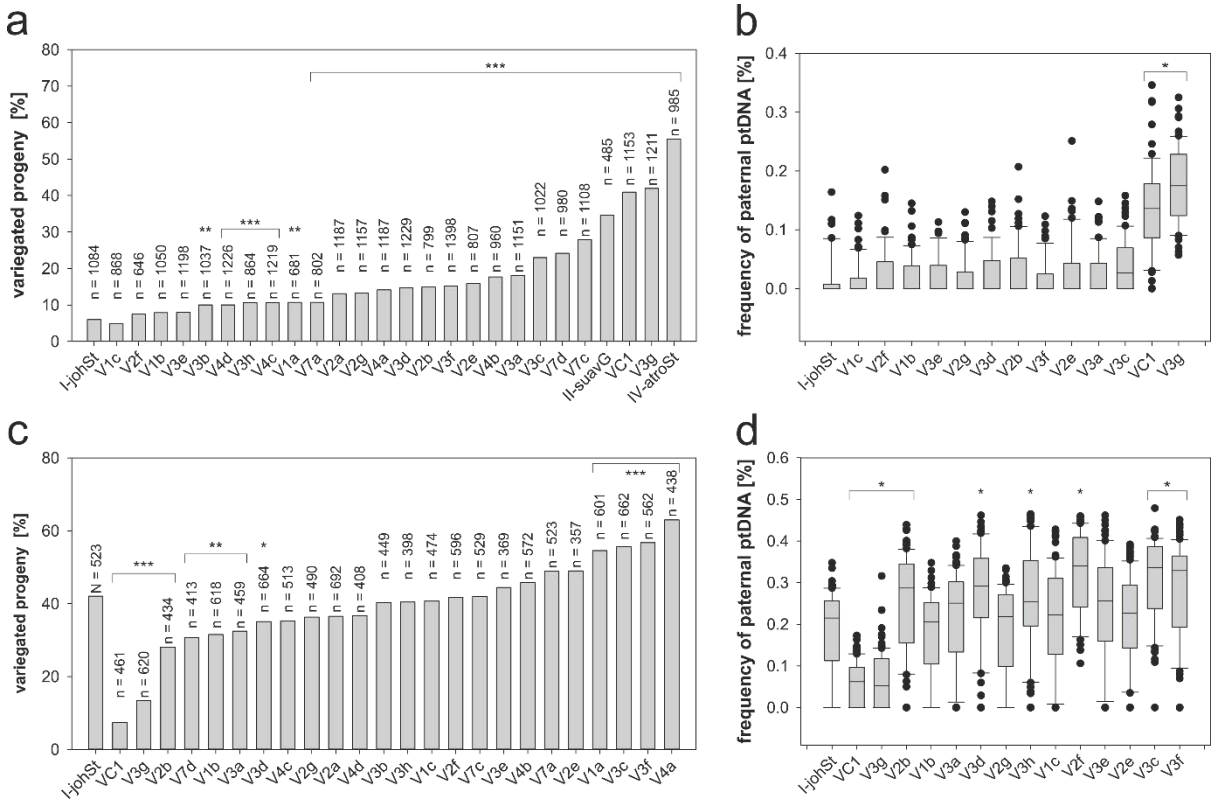


Fig. 2. Transmission efficiencies of plastome I variants and the wild type plastomes I-johSt, II-suavG and IV-atroSt as determined by crosses with I-chi/I-hookdV and IV-delta/IV-atroSt. (**A**,**C**) Classical approach based on counting of variegated seedlings. (**B**,**D**) MassARRAY® assay quantifying maternal and paternal ptDNA. Wild types and green variants were crossed (**A**) as the seed parent to I-chi as male parent and (**C**) crossed as the pollen parent to IV-delta as female parent. Depicted is the average percentage of variegated progeny (= percentage biparental inheritance) obtained from three seasons (2013, 2014 and 2015). Significance of differences between variant and I-johSt was calculated by Fisher`s exact test (*** $p < 0.0001$, ** $p < 0.001$, * $p < 0.01$). (**B**) Crosses of I-johSt and variants with I-hookdV as male and, (**D**) with IV-atroSt as female parent. Box-plots represent the transmission frequencies of the paternal plastomes measured by MassARRAY®. To account for significant differences to I-johSt Kruskal-Wallis, one-way ANOVA on ranks was performed (* $p < 0.05$).

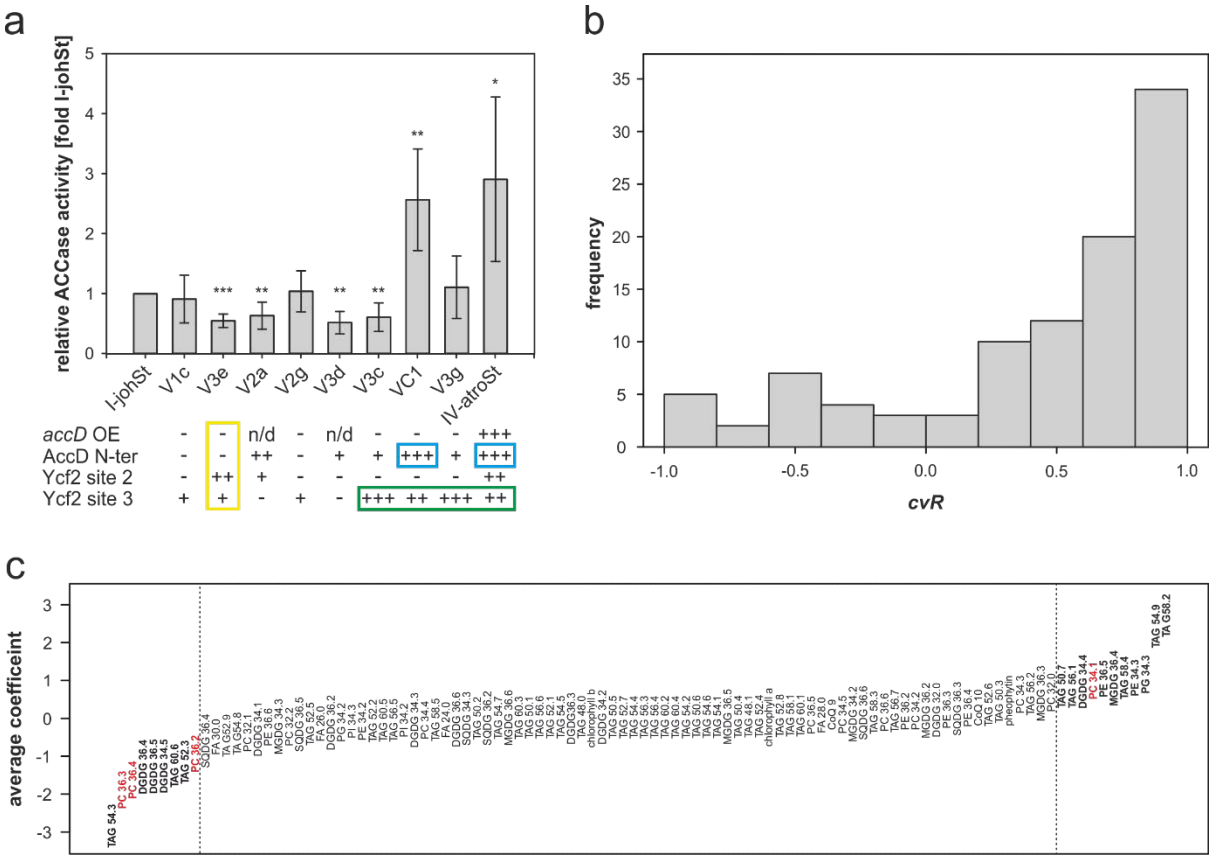


Fig. 3. ACCase activity and prediction of inheritance strengths based on lipid level. (**A**) ACCase activity of chloroplasts with difference inheritance strengths, sorted according to competitive abilities (percentage biparental inheritance I-chi; cf. Fig. 2); *accD* OE = *accD* overexpressor; AccD N-ter = AccD N-terminus; Ycf2 site 2 and Ycf2 site 3; see Supplementary Text. Compared to I-johSt: - = not affected, + = mildly affected, ++ = intermediately affected, and +++ = strongly affected, n/d = not determined (cf. Figs. S1,S2, S8 - S10 and S14). Note the influence of mutations in *ycf2* on AccD activity in a non-mutated *accD* background (yellow box), the strong correlation of mutations in the AccD N-terminus with ACCase activity (green box), and the correlation of mutations in site 3 of Ycf2 with inheritance strengths (blue box). Significance of difference compared to I-johSt was calculated using paired two-tailed *t-test* (*** $p < 0.0005$, ** $p < 0.005$, * $p < 0.05$). (**B**) Histogram of Pearson correlation coefficient *cvR* between actual and predicted inheritance strengths obtained from 100 cross-validation runs. Note shift towards 1.0, pointing to the predictive power of lipid data in the LASSO regression model. (**C**) Average linear LASSO model coefficients of the 102 lipids/molecules available for analysis (cf. Table S2). The 20 identified predictive lipids are marked in boldface font. PCs, the dominant lipids of the chloroplast outer envelope, are indicated in red. In general, large absolute values show predictive power either negatively or positively correlating with inheritance strength. Predictive lipids are designated when their absolute average weight was larger than one standard deviation of all weigh values obtained for the 102 lipids.

Materials and Methods

Plant material

Throughout this work, the terms *Oenothera* or evening primrose refer to subsection *Oenothera* (genus *Oenothera* section *Oenothera*) [30]. Plant material used here is derived from the *Oenothera* germplasm collection harboured at the Max Planck Institute of Molecular Plant Physiology, Potsdam-Golm, Germany [31]. Part of this collection is the so-called Renner Assortment, a collection of lines thoroughly characterized by the genetic school of Otto Renner [12,32]. Therefore, the original material of Franz Schötz in which he determined the distinct classes of chloroplast replication speeds [6,33] was available for correlation mapping. For all other genetic or physiological work presented here, the nuclear genetic background of *O. elata* ssp. *hookeri* strain johansen Standard [34] was used. The employed chloroplast (genomes) are either native or were introgressed into that race by Wilfried Stubbe or Stephan Greiner. The wild type chloroplast genomes (I-johSt, I-hookdV, II-suavG, and IV-atroSt) are compatible with and hence green when combined with the johansen Standard nucleus. The chloroplast genome III-lamS confers a reversible bleaching, so-called *virescent*, phenotype in this genetic background [30,35] (Fig. S15). The white chloroplast mutants I-chi and IV-delta (Fig. S15). are part of the huge collection of spontaneous plastome mutants compiled by Wilfried Stubbe and co-workers [23,36,37]. Both mutants harbour a similar single locus mutation in the *psaA* gene (encoding a core subunit of photosystem I) and derive from the strong and weak wild type plastomes I-hookdV and IV-atroSt, respectively [23]. The *plastome mutator* line is a descendant of the original isolate E-15-7 of Melvin D. Epp. The nuclear *pm* allele was identified after an ethyl methanesulfonate mutagenesis [38] in johansen Standard. From descendants of this line the variant chloroplast genomes (V1a, V1b, V2a, etc., VC1, III-V1 and III-V2) with altered inheritance strengths were generated from the strong chloroplast genomes I-johSt and III-lamS, respectively [14] (see below). A summary of all strains and origins of the chloroplast genomes is given in Tables S4-S9.

Germination, plant cultivation and cross pollination

Fresh seeds were germinated on wet filter paper at 27°C at 100-150 µE m^{-2} s^{-1}. With this method essentially 100% germination was achieved after 1-3 day. If desired, seedlings were then cultivated to the appropriate developmental stage in a glasshouse. Crosses with flowering plants were performed as published earlier [23,31].

Plastome mutator mutagenesis

Chloroplast genome mutagenesis was conducted as previously described [16]. In brief, johansen Standard plants newly restored to homozygosity for the nuclear *plastome mutator* allele (*pm*/*pm*) were employed to mutagenize the chloroplast genome (I-johSt). When homozygous the *plastome mutator* causes a 200-1000x higher rate of chloroplast mutants compared to the spontaneous frequency. The underlying mutations mostly represent indels originated from replication slippage events [16,38-41].

Homozygous *pm* plants were identified when new mutant chlorotic sectors were observed on them. On those plants, flowers on green shoots were backcrossed to the wild type *PM* allele as pollen donor. In the resulting *pm*/*PM* populations the chloroplast mutations were stabilized. This led (after repeated backcrosses with the *PM* allele and selection with appropriate markers against the paternal chloroplast) to homoplasmic green variants of the strong chloroplast genome I-johSt that differed by certain indels or combination of indels. The material was designated V1a, V1b, V2a, etc., where "V" stands for variant, the Arabic number for the number of the backcrossed plant in the experiment, and the small Latin letter for the shoot of a given plant. An additional line, named VC1, derived from a similar *plastome mutator* mutagenesis of I-johSt, but the mutagenesis was conducted over several generations. Due to this fact VC1, which is also a green variant, carries a much higher mutational load than do variants V1a, V1b, V2a, etc. (Table S8). The two variant chloroplast genomes III-V1 and III-V2 (Table S9) have a derivation similar to VC1. They are derived from the strong wild type chloroplast genome III-lamS, which displays a reversible bleaching (*virescent* phenotype) in the johansen Standard nuclear genetic background. To mutagenize this chloroplast genome, it was introgressed into the *pm*/*pm* background of johansen Standard by Wilfried Stubbe and selfed for a number of generations. When stabilized with the *PM* allele, it still displayed a *virescent* phenotype that is comparable to the original wild type plastome III-lamS (Fig. S15).

Determination of plastid inheritance strength

In the evening primroses biparental transmission of plastids shows maternal dominance, i.e. F1 plants are either homoplasmic for the maternal chloroplast or heteroplasmic for the paternal and maternal chloroplasts. If in such crosses one of the chloroplasts is marked by a mutation, resulting in a white phenotype, the proportion of variegated (green/white; i.e. heteroplasmic) seedlings can be used to determine chloroplast inheritance strength (as percentage of biparental inheritance). Moreover, if in such crosses one of the crossing partners is kept constant, the inheritance strength of all tested chloroplasts in respect to the constant one is determined [6,7,14,33]. For example, in the I-chi crosses (where the strong white plastid is donated by the father; see below), when more variegated seedlings

are found in the F1, it indicates that more paternal (white) chloroplasts were able to out-compete the dominating maternal green chloroplasts. Hence, in this crossing direction small biparental percentace values indicate strong (assertive) plastomes and high biparental values indicate weak variants. The situation is reversed in the reciprocal cross where the white chloroplast is donated by the mother, as is the case in the IV-delta crosses. Here, the weak white chloroplast is maternal and strong green variants contributed by the pollen give high fractions of variegated plants in the F1, whereas low percentages of biparental progeny result when weak green variants are carried by the pollen donor.

Crossing studies

All crossing studies between chloroplast genomes were performed in the constant nuclear background of the highly homozygous johansen Standard strain (see above). Germination efficiency in all populations was 100% (see above). Transmission efficiencies of the green plastome I variants (V1a, V1b, V2a, etc.) were determined using the white chloroplast I-chi (strong inheritance strength) and IV-delta (weak inheritance strength) as crossing partners, respectively. This allows the determination of the inheritance strength of a given green chloroplast relative to a white one based on counting the number of green, variegated (green/white), or white seedlings in the F1 [6,14,33] (also see above). In the I-chi crosses, the green plastome I variants, as well as the wild type chloroplast genomes I-johSt (strong inheritance strength; native in the genetic background of johansen Standard and the original wild type chloroplast genome used for mutagenesis), II-suavG (intermediate inheritance strength) and IV-atroSt (weak inheritance strength) were crossed as female parent to I-chi in three following seasons (2013, 2014, and 2015). In the IV-delta crosses green variants and I-johSt were crossed as male parent to IV-delta, again in three independent seasons, 2013, 2014 and 2015. From each cross of each season randomized populations of 100-300 plants were grown twice independently, followed by rating of green, variegated and white seedlings/plantlets 14-21 days after sowing (DAS; I-chi crosses) or 7-14 DAS (IV-delta crosses). Based on this counting, percentage of variegated progeny was calculated for each individual cross. To determine statistically significant differences between the transmission efficiencies of the plastome I variants and I-johSt, counting results from all three seasons were summed for a particular cross and a Fisher's exact test was employed.

A very similar experiment was performed to determine the inheritance strength of the two green variants III-V1 and III-V2, which derive from the strong chloroplast genomes III-lamS. Here in two independent seasons (2015 and 2016) the wild type I-johSt was used as pollen donor to induce variegation between the maternal plastome III (giving rise to a *virescent* phenotype) and the green plastome I (native in the background of johansen Standard; see above).

To determine transmission efficiencies independent of white chloroplast mutants or other bleached material, the plastome I variants (including their wild type I-johSt) were crossed to the green wild type plastomes IV-atroSt (weak inheritance strength) as female parent and to I-hookdV (strong inheritance strength) as male one in two independent seasons (2013 and 2014). F1 progeny was harvested 6 DAS by pooling 60-80 randomized seedlings and the ratios of the plastome types in the pool were analysed via MassARRAY® (Agena Bioscience, Hamburg, Germany) as described below.

MassARRAY®: multiplexed genotyping analysis using iPlex Gold

SNP genotyping to distinguish plastome I-johSt and I-hookdV/I-chi or I-johSt and IV-atroSt/IV-delta and subsequent quantification of their plastome ratios in appropriate F1s was carried out with the MassARRAY® system (Agena Bioscience, Hamburg, Germany). The system was used to analyse chloroplast transmission efficiencies in different crosses. For this, total DNA was prepared from 60-80 randomized pooled plantlets 6 DAS. Then, 10 SNPs distinguishing the plastomes I-johSt and I-hookdV/I-chi (I/I assay) and 15 SNPs between I-johSt and IV-atroSt/IV-delta (I/IV assay) were selected. Two appropriate primers flanking the SNP and one unextended primer (UEP; binding an adjacent sequence to the SNP) were designed using MassARRAY® Assay Design v4.0 (Agena Bioscience, Hamburg, Germany). Primer sequences, SNPs and their position in I-johSt are listed in Table S10. Plastome regions were amplified in a 5 µl PCR reaction containing PCR buffer (2 mM $MgCl_2$, 500 µM dNTP mix, 1 U HotStartTaq; Agena Bioscience, Hamburg, Germany), 10 ng DNA and 10 (I/I assay) or 15 (I/IV assay) PCR primer pairs, respectively, at concentrations ranging from 0.5-2.0 µM. The reaction mix was incubated for 2 min at 95°C in 96 well plates, followed by 45 cycles of 30 sec at 95°C, 30 sec at 56°C and 60 sec at 72°C, and a final elongation for 5 min at 72°C. Excess nucleotides were removed by adding 0.5 U Shrimp alkaline phosphatase (SAP enzyme) and SAP buffer (Agena Bioscience, Hamburg, Germany), followed by an incubation for 40 min at 37°C and 5 min at 85°C. For the primer extension reaction the iPLEX reaction mixture (containing Buffer Plus, Thermo Sequenase and termination mix 96; Agena Bioscience, Hamburg, Germany) and, depending on the primer, the extension primers at a concentration of 7-28 µM were added. Sequence-specific hybridization and sequence-dependent termination were carried out for 30 sec at 94°C, followed by 40 cycles of 5 sec at 94°C plus five internal cycles of 5 sec at 52°C and 5 sec at 80°C, and finally 3 min at 72°C. After desalting with CLEAN resin (Agena Bioscience, Hamburg, Germany) the samples were spotted on 96-pad silicon chips preloaded with proprietary matrix (SpectroCHIP; Agena Bioscience, Hamburg, Germany) by using the Nanodispenser RS1000 (Agena Bioscience, Hamburg, Germany). Subsequently, data were acquired with MALDI-TOF mass spectrometer MassARRAY® Analyzer 4 (Agena Bioscience, Hamburg, Germany) and analyzed with the supplied software. To identify significant differences in the

frequencies of paternal ptDNA Kruskal-Wallis one-way analysis of variance (ANOVA) on ranks was performed.

k-means clustering to classify inheritance strength

For the wild type chloroplasts, inheritance strength was classified using the paternal transmission frequencies (percentage of variegated plants in F1, see above) of the chloroplasts "biennis white" and "blandina white" according to Schötz [6]. Both crossing series included the same 25 wild type chloroplasts, 14 of which had fully sequenced genomes and were employed for correlation mapping (Table S5 and below). For original data, see Schötz (1968) [6], summaries in Cleland (1972, p. 180) [12], or Table S11. Based on the two transmission frequencies the tested wild type plastomes were clustered using the *k*-means algorithm with Euclidean distance as distance dimension. The optimal number of centers was calculated with the pamk function of the fpc package, as implameted in R v.3.2.1 [42]. Strikingly, essentially the same three classes (strong, intermediate, and weak) were obtained that had been previously determined by Schötz [7,12] (Fig. S16 and Supplementary Text).

For the variants, we used the transmission frequencies from I-chi and IV-delta crosses obtained from this work (Table S8 and Supplementary Text). Since the data-driven determination of the optimal number of clusters (k = 2; see above) does not reflect the biological situation, upon repeated *k*-means runs, we chose the number of centres with the best trade-off between lowest swapping rate of the samples between the clusters and the biological interpretability. This approach resulted in four classes (see Supplementary Text and Fig. S16 for details).

DNA isolation

Total DNA for Illumina sequencing was isolated described before [23]. For PCR analyses, quantitative real-time PCR and MassARRAY® experiments total DNA was isolated with the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) [43].

DNA sequencing

Illumina sequencing were performed at the Max Planck-Genome-Centre in Cologne (Germany), as described previously [23,44]. Exclusively "PCR-free" paired-end libraries (375 bp insert size) generated from total DNA isolations were employed. 100 bp Illumina paired-end reads were generated to determine the sequences of the green plastome I variants, whereas 150 bp Illumina paired-end reads were used to assemble the wild type plastomes. Sanger sequence were obtained from Eurofins MWG Operon (Ebersberg, Germany).

Plastome assembly, sequence annotation and repeat analysis

Complete chloroplast genomes were assembled with SeqMan NGen v12.1.0 or v13.0.0 (DNASTAR, Madison, WI, USA). Wild type ones were done *de novo*. The sequence of the green plastome I variants derive from reference-guided assemblies. For this, their original wild type chloroplast genome I-johSt of *O. elata* ssp. *hookeri* strain johansen Standard (AJ271079.4) was used as reference [23]. The notoriously repetitive regions in the *Oenothera* plastome, where individual repetitive elements can span more than an Illumina read length and/or the insert size of the employed paired-end libraries, namely upstream and/or within *accD*, *ycf1* (*tic214*), *ycf2*, and the *rrn16* - *trnI-GAU* spacer (*oriB*) were determined and/or confirmed by Sanger sequencing in all chloroplast genomes discussed in this work. Finalized sequences were annotated by GeSeq v0.9 [45] and the wild type chloroplasts genomes were prepared for NCBI submission using GenBank 2 Sequin v1.3 [46]. Repeat structures were analysed based on an inspection by eye and/or employing the EMBOSS suite [47] as previously described [24].

Correlation mapping

For correlation mapping in the wild types, 14 complete plastomes (GenBank accession numbers EU262890.2, EU262891.2, KT881170.1, KT881171.1, KT881172.1, KT881176.1, KU521375.1, KX687910.1, KX687913.1, KX687914.1, KX687915.1, KX687916.1, KX687917.1, KX687918.1) with known inheritance strength [6,7,12] (Table S5) were employed. The eight chloroplast genomes assigned to accession numbers starting with KU or KX were newly determined in the course of this work. Mapping of genetic determinants in the green variants was done in 18 fully sequenced mutagenized plastomes (V1a, V1b, V1c, V2a, V2b, V2g, V3a, V3b, V3c, V3d, V3e, V3f, V3g, V3h, V4b, V4c, V7a, and VC1), as well as their wild type reference (I-johSt; GenBank accession number AJ271079.4).

In both sequence sets, divergence at a given alignment window was correlated to the experimentally determined inheritance strengths of a chloroplast genome. For the wild types, inheritance strength was measured using the paternal transmission frequencies (percentage of variegated plants in the F1) of the chloroplasts “biennis white” or “blandina white” according to Schötz [6], or *k*-means classes combining the two datasets by clustering (Table S11, Supplementary Text and see above). For the variants, we used the transmission frequencies from the I-chi and IV-delta crosses determined in this work, as well as the obtained *k*-means classes thereof (Table S8, Supplementary Text and above).

For correlation of these transmission frequencies to loci on the chloroplast genome, the redundant inverted repeat A (IR_A) was removed from all sequences. Then, plastomes were aligned with ClustalW [48] and the alignments were curated manually (Data S2). Subsequently, using a script

in R v3.2.1 [42] (Data S3), nucleotide changes (SNPs, insertion and deletions) relative to a chosen reference sequence plastome [I-hookdV (KT881170.1) for the wild type set and I-johSt (AJ271079.4) for the variant set] were counted window-wise by two approaches: (i) segmenting the reference sequence in overlapping windows using a sliding window approach with a window size of 1 kb and a step size of 10 bp, yielding a matrix of 13,912 x 13 (wild type set) and 13,668 x 18 (variants), respectively, or (ii) defining regions of interest with correspondingly chosen window sizes. Then, Pearson's and Spearman's correlation coefficients were calculated between (i) the total count of nucleotide changes for every plastome in the aligned sequence window compared to the reference (total sequence divergence) and (ii) the determined inheritance strength of the plastomes (Fig. S17, Data S1). For the sliding window approach, *p*-values were adjusted for multiple testing using Benjamini-Hochberg correction. To reduce the number of *p*-value adjustments, adjacent alignment windows with identical count vectors were collapsed into one. For visualization (Fig. 1A, Fig. S6 and 7), correlation coefficients obtained for every alignment window were plotted as a function of the alignment position. Correlation values above the *p*-value threshold (> 0.05) are greyed out. Linear chloroplast genome maps were derived from the annotation of the consensi of both sequence sets (Data S2) and drawn by OrganellarGenomeDRAW v1.2 [49] in a linear mode using a user-defined configuration XML file. Alignments of selected plastome regions were visualized in Geneious v10.2.3 [50] and subsequently, as the output of OrganellarGenomeDRAW, edited in CorelDraw X8 (Corel Corporation, Ottawa, ON, Canada).

Chloroplast isolation

For isolation of chloroplasts, *Oenothera* leaf tissue was harvested 7-8 weeks after sowing and processed as described previously [23]. However, minor modifications were applied to allow a rapid isolation of chloroplasts from six plant lines in parallel: 35 g of leaf material was homogenized in 500 ml BoutHomX buffer. The pellet from the first centrifugation step was re-suspended in 100 ml ChloroWash and, after one filtration, the volume was adjusted to 150 ml before the second filtration. After the second centrifugation step re-suspended chloroplasts were loaded on two Percoll step gradients (each: 7 ml 85% Percoll, 14 ml 45% Percoll). Subsequent to gradient centrifugation the recovered chloroplasts were washed with 30 ml ChloroWash, followed by three additional washing steps with smaller volumes and a final re-suspension of the chloroplasts in 300-500 µl ChloroWash for the following ACCase activity measurements.

ACCase activity assay

ACCase activity was measured in isolated chloroplasts [51,52]. Suspensions of isolated chloroplasts were diluted to gain 400 µg chlorophyll/ml, determined as described above. To validate equilibration to chlorophyll, protein concentration using a Bradford assay (Quick StartTM Bradford 1x Dye Reagent; Bio-Rad, Hercules, CA, USA; with BSA solutions of known concentrations as standards) and chloroplast counts per ml suspension were determined for the same samples. For chloroplast counting, the suspension gained above was further diluted 1:10, with 15 µl subsequently loaded on a CELLOMETER™ Disposable Cell Counting Chamber (Electron Microscopy Sciences, Hatfield, PA, USA) and analysed under a Zeiss Axioskop 2 (Zeiss, Oberkochen, Germany). For each sample six "B squares" were counted and chloroplast concentration was calculated as chloroplasts/ml = 10 x average count per "B square" / 4 x 10-6. All three equilibration methods gave comparable results.

ACCase activity was measured as the acetyl-CoA-dependent fixation of $H^{14}CO_3^-$ into acid-stable products. For each plant line (I-johSt, V1c, V3e, V2a, V2g, V3d, V3c, VC1, V3g and IV-atroSt) three independent chloroplast isolations (= three biological replicates) were analysed in triplicates, including individual negative controls (minus acetyl-CoA) for each measurement. 10 µl of chloroplast suspensions were incubated with 40 µl of reagent solution having a final concentration of 100 mM Tricine-KOH pH 8.2, 100 mM potassium chloride, 2 mM magnesium chloride, 1 mM ATP, 0.1 mM Triton X-100, 10 mM sodium bicarbonate, 0.5 mM acetyl-CoA and 40 mM radioactively labelled sodium bicarbonate ($NaH^{14}CO_3$, ca. 4000 dpm/nmol; Amresham, Little Chalfont, UK) at room temperature for 20 min. For the negative control, acetyl-CoA in the reaction mixture is replaced with water. Reactions were stopped by adding 50 µl 2 M hydrochloric acid. The sample was transferred to a scintillation vial and acid labile radioactivity (i.e. remaining $H^{14}CO_3^-$) was evaporated by heating for 20 min at 85°C. After addition of 3 ml scintillation cocktail (Rotiszint® eco plus, Carl Roth, Karlsruhe, Germany), the acid stable radioactivity from incorporation of $H^{14}CO_3^-$ (^{14}C dpm) were detected by liquid scintillation counter (LS6500, Beckman Coulter, Brea, CA). ACCase activity is represented as the ^{14}C incorporation rate into acid stable fraction (dpm min^{-1}) calculated by dividing the total fixed radioactivity by 20 min. The rates in three replicated reactions were averaged and corresponding values from negative control samples were subtracted and normalized by the number of chloroplast to gain ACCase activity in individual samples. The average rates were calculated for each line. To combine all measurements, relative ACCase activities were calculated for each experiment as [fold I-johSt], and significant differences between each line and the wild type were identified using two-tailed paired *t*-test, followed by *p*-value adjustment using the Benjamini-Hochberg procedure.

Lipid extraction, mass spectrometry sample preparation and measurements

Metabolites were extracted according to published protocols [53] from 50 mg *Oenothera* seedlings harvested 6 DAS. In brief, frozen tissue was homogenized by a ball mixer mill and transferred to cooled 2.0 ml round bottom microcentrifuge tubes. Subsequently, each sample was re-suspended in 1.0 ml of a -20°C methanol:methyl-*tert* butyl-ether [1:3 (v/v)] mixture, containing 0.5 µg of 1,2-diheptadecanoyl-*sn*-glycero-3-phosphocholine (Avanti Polar Lipids, Alabaster, AL, USA) as an internal standard. Samples were then immediately vortexed before incubation for 10 min at 4°C on an orbital shaker. This step was followed by ultra-sonication in an ice-cooled bath-type sonicator for an additional 10 min. To separate the organic from the aqueous phase, 650 µl of a H_2O:methanol mix [3:1 (v/v)] was added to the homogenate, which was shortly vortexed before being centrifuged for 5 min at 14,000 *g*. Finally, 500 µl of the upper methyl-*tert* butyl-ether phase, containing the hydrophobic (lipid) compounds, was placed in a fresh 1.5 ml microcentifuge tube. This aliquot was either stored at -20°C for up to several weeks or immediately concentrated to complete dryness in a speed vacuum concentrator at room temperature. Prior to analysis the dried pellets were re-suspended in 400 µL acetonitrile:isopropanol [7:3 (v:v)], ultra-sonicated and centrifuged for 5 min at 14,000 *g*. The cleared supernatant was transferred to fresh glass vials and 2 µl of each sample injected onto a C_8 reverse phase column (100 mm x 2.1 mm x 1.7 µm particles) using a Acquity UPLC system (Waters, Manchester, UK). In addition to the individual samples, we prepared pooled samples, in which 10 µl of each sample from the whole sample collection was mixed. These pooled samples were measured after every 20^{th} sample, to provide information on system performance including sensitivity, retention time consistency, sample reproducibility and compound stability.

The mobile phase for our chromatographic separation consisted of Buffer A (1% 1 M NH_4-acetate and 0.1% acetic acid in UPLC MS grade water), while Buffer B contained 1% 1 M NH_4-acetate and 0.1% acetic acid in acetonitrile/isopropanol [7:3 (v:v)] (BioSolve, Valkenswaard, Netherlands). The flow rate of the UPLC system was set to 400 µl/min with a Buffer A/Buffer B gradient of 1 min isocratic flow at 45% Buffer A (55% Buffer B), 3 min linear gradient from 45% to 25% Buffer A (55% to 75% Buffer B), 8 min linear gradient from 25% to 11% Buffer A (75% to 89% Buffer B), and 3 min linear gradient from 11% to 1% Buffer A (89% to 99% Buffer B). After cleaning the column for 4.5 min at 1% Buffer A/99% Buffer (B) the solution was set back to 45% Buffer A/55% Buffer (B) and the column was re-equilibrated for 4.5 min, resulting in a final run time of 24 min per sample.

Mass spectra were acquired with an Orbitrap-type mass spectrometer (Exactive; Thermo-Fisher, Bremen, Germany) and recorded in the full scan mode, covering a mass range from 100-1,500 m/z. The resolution was set to 60,000 with 2 scans per second, restricting the maximum loading time to 100 ms. Samples were injected using the heated electrospray ionization source (HESI) at a capillary

voltage of 3.5 kV in positive and negative ionization mode. A sheath gas flow value of 40 was used, with an auxiliary gas flow value at 20 and a capillary temperature of 200°C, while drying gas temperature in the heated electro spray source was 350°C. The skimmer voltage was set to 20 V with tube lens value at 140 V. The spectra were recorded from 0 to 20 min of the UPLC gradients.

Data processing and normalization of lipid data

Data analysis of the raw files (.raw) was performed using QI for metabolomics v2.3 (Nonlinear Dynamics, Newcastle upon Tyne, UK) according to the vendor description. Data were normalized to the internal standard (1,2-diheptadecanoyl-*sn*-glycero-3-phosphocholine) and the exact fresh weight of each sample. Lipid annotation was performed manually as described[54]. Statistical data analysis was performed using Excel 2013 (Microsoft, Redmond, WA, USA), R v3.2.1 [42] and SIMCA-P v13.0 (Umetrics, Umea, Sweden).

Predictability of inheritance strength based on lipid-level data as explanatory variables

Lipidomics data from *Oenothera* seedlings of the strain johansen Standard, harboring chloroplast genomes with different assertiveness rates, were analyzed jointly to test for predictability of inheritance strength based on lipid levels. For this, 33 probes representing 16 genotypes whose chloroplast genomes ranged from inheritance strength class 1 to 5 (see Supplementary Text) were measured in five replicates in three independent experimental series (Table S1; Supplementary Text). In this dataset, a total of 184 different lipids/molecules could be annotated (Data S4; see above). Then, the data from each series were log-transformed and median-centered based on genotypes with inheritance strengths = 1, i.e. for every lipid/molecule, its median level across all inheritance strengths = 1 genotypes was determined and subtracted from all genotypes tested in the respective experimental series. Inheritance strength 1 was then selected to serve as a common reference across all three experimental series. Subsequently, the three experimental series were combined into a single set. Only those lipids/molecules were considered further, for which level-data were available across all three datasets, leaving 102 lipids/molecules for analysis (Data S4)

LASSO regression model: Inheritance strength was predicted based on the median-centered lipid level data using LASSO, a regularized linear regression approach [55] as implemented in the "glmnet" R-software package (R v3.2.1) [42]. Glmnet was invoked with parameter α set to 1 to perform LASSO regression (Data S4). The penalty parameter λ was determined from the built-in cross-validation applied to training set data (i.e. all but two randomly selected genotypes) and set to the obtained one-standard-error estimate deviation from the optimal (minimal error) value and assuming Gaussian response type. All other parameters were taken as their default values.

Predictive lipids: As a regularized regression method, LASSO aims to use few predictor variables, which allows better identification of truly predictive lipids. Summarized from all 100 performed cross-validation runs, lipids/molecules were ranked by their mean linear model coefficients assigned to them in the LASSO regression with their absolute value indicating influence strength, and their sign indicating positive or negative correlation of their abundance to inheritance strength.

Test for enrichment of predictive lipids/molecules in lipid classes: Across all 100 cross-validation runs, the importance of each of the 102 molecules was assessed based on their average absolute weight factor (avgW) by which they entered the 100 LASSO models. Molecules with avaW of greater than one standard deviation obtained across all 102 molecules were considered important. Then, all lipid/molecule were assigned to their respective class (MGDG, DGDG, SQDG, PG, PC, PI, PE, FA, PE, TAG, CoQ, chlorophyll, and pheophytin) and every class was tested for enrichment in the lipid/molecule set considered to be important. This was done by employing Fisher's exact test yielding both *p*-values and odds-ratios. The *p*-values express enrichment the odds-ratio express the relative enrichment or depletion of a particular class among the set of important lipids.

Determination of photosynthetic parameters

Gas exchange measurements were performed with a GFS-3000 open gas exchange system equipped with the LED array unit 3055-FL as actinic light source for simultaneous chlorophyll a fluorescence measurements (Heinz Walz GmbH, Effeltrich, Germany). Light response curves of CO_2 assimilation were measured at 22°C cuvette temperature with 17,500 ppm humidity and a saturating CO_2 concentration of 2,000 ppm, to fully repress photorespiration. Plants were dark-adapted for a minimum of 30 min. Then, the maximum quantum efficiency of photosystem II in the dark-adapted state (F_V/F_M) and leaf respiration were determined. Afterwards, the actinic light intensity was first set to the growth light intensity of 200 µE m^{-2} s^{-1}, followed by measurements at 500, 1,000, and finally 1,500 µE m^{-2} s^{-1}. At each light intensity, gas exchange was recorded until the steady state of transpiration and leaf assimilation was reached. Maximum leaf assimilation was corrected for the respiration measured in darkness. After the end of the gas exchange measurements, the chlorophyll content and chlorophyll a/b ratio of the measured leaf section were determined in 80% (v/v) acetone according to [56]. Leaf absorptance was calculated from leaf transmittance and reflectance spectra as 100% minus transmittance (%) minus reflectance (%). Spectra were measured between 400 and 700 nm wavelength using an integrating sphere attached to a photometer (V650, Jasco Inc., Groß-Umstadt, Germany). The spectral bandwidth was set to 1 nm, and the scanning speed was 200 nm min^{-1}.

Fluorescence microscopy and differential interference contrast to analyse ptDNA nucleoids, chloroplast volume and number per cell

We investigated leaf material from the central laminal region of the first true leaf 25 DAS. For this, four pieces of 5 mm^2 were excised from five individual plants per line and fixed. DAPI (4',6-diamidino-2-phenylindole) stains of ptDNA nucleoids and fluorescence microscopy was conducted as previously described [57,58] with minor modifications: In brief, excised leaf fragments were fixed with 3% glutaraldehyde in 50 mM phosphate buffer (pH 7.2), washed in 1x PBS buffer (phosphate-buffered saline, 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.2) and macerated in 1% (w/v) cellulase and 1% (w/v) pectinase solution (both Sigma-Aldrich, St. Louis, MO, USA) in the 1x PBS for 30 min at 37°C. Subsequently, explants were washed in the 1x PBS and stored at 4°C until use. For microscopy, small tissue sectors were gently squeezed in a drop of the 1x PBS between a microscope slide and a cover glass. Then preparations were frozen in liquid nitrogen and, after removing the cover glasses, air dried and mounted in a drop of DAPI solution [5 µg/ml DAPI (Sigma-Aldrich, St. Louis, MO, USA) in 1x PBS buffer in 70% glycerol ("for fluorescence microscopy"; Merck, Darmstadt, Germany)]. DAPI as fluorochrome is considered to be sensitive enough to detect DNA of a single plastid genome copy [59]. The preparations were sealed with Fixogum rubber cement (Marabu, Tamm, Germany) and examined with a Nikon Eclipse Ni-U upright epifluorescence microscope equipped with a cooled monochrome camera (Nikon, Chiyoda, Japan) under a 100x UV objective. For each investigated cell, five to seven picture frames were digitally captured, each at a different focal plane. The frames were stacked and combined using standard macro commands of the Combine ZP software v1.0 developed by Alan Hadley (http://combinezp.software.informer.com). Nucleoids were counted manually in Adobe Photoshop CS3 (Adobe Systems San Jose, CA, USA). For this, chloroplasts in a cell were delimited and nucleoids identified (Fig. S13). Only chloroplasts with non-overlapping nucleoids were subjected to analysis. Overall we investigated 22 cells per plant line and 6-29 chloroplasts per cell. In total 26,120 nucleoids were counted (Table S12). Significance of difference between all lines was tested by one-way ANOVA. In addition, to test differences between the strong plastome I-johSt and weak plastome V3g, VC1, and IV-atroSt, respectively, two-tailed homoscedastic *t*-test was calculated followed by *p*-value adjustment according to Benjamini-Hochberg.

For differential interference contrast (DIC) microscopy, explants excised as described above were transferred to 10% formalin in phosphate buffer (Tissue-Prep Buffered 10% Formalin; Electron Microscopy Sciences, Hatfield, PA, USA). Samples were then evaporated for at least 1 h and incubated at 4°C overnight. After washing with sterile water, leaf fragments were incubated under rotation in 0.1 M EDTA for 2 h at room temperature, followed by incubation at 4°C overnight. Directly before analysis, samples were incubated for 3 h at 60°C while shaking (500 rpm). Leaf pieces were mounted

in water on a slide and cells were released by softly tapping on the top of the cover slide. Analysis was performed on a motorized epifluorescence microscope Olympus BX61 under a 40x objective (Olympus, Shinjuku, Japan) using DIC optics. This allows capturing of a focal plane in which all chloroplasts within a cell are visible. Subsequently, counting of chloroplasts per cell and measurement of chloroplast length and width was carried out with the Olympus cellSensTM Dimension software v1.7 (Olympus, Shinjuku, Japan). 4-5 individuals per plant line, and at least two leaf pieces per individual plant and a minimum of 5 spongy mesophyll cells per piece were investigated. To determine chloroplast numbers per cell, 55 cells per plant line were analysed and overall 11,645 chloroplasts counted. For measuring chloroplast length and width 20 cells per plant line were chosen and three plastids per cell measured resulting in 480 individual measurements. Chloroplast volume index was calculated according length x width^2 [60]. For statistical analysis of each experiment, for comparison of all lines, one-way ANOVA was performed. In addition, to test differences between the weak plastome I variants or IV-atroSt and the strong wild type I-johSt using a two-tailed homoscedastic *t*-test followed by adjustment of *p*-values according to Benjamini-Hochberg was done. (Tables S13 and S14).

Quantitative real-time PCR to determine plastid DNA amounts

Plastid DNA content of wild type and variants with different inheritance strength were analysed in a developmental series including 5, 21, and 32 DAS. For each line, three DNA preparations from independent pools of 20 individuals (5 DAS), or two DNA preparations from independent pools of three plantlets (21 DAS) or two DNA preparations from independent pools of three second leaves (32 DAS) were isolated, and subsequently analysed two times in triplicate by qPCR: Reactions containing LightCycler® 480 SYBR Green I Master (Roche Diagnostics GmbH, Mannheim, Germany), 0.75 ng total DNA and 0.5 µM primers were incubated for 5 min at 95°C, followed by 40 cycles of 10 sec at 95°C, 10 sec at 58°C and 15 sec at 72°C in a LightCycler® 480 II instrument (Roche Diagnostics GmbH, Mannheim, Germany). Primer sequences, accession numbers of target genes/loci, size of amplification products, and primer efficiencies (that have been determined based on dilution series of total johansen Standard DNA harbouring either plastome I-johSt or IV-atroSt) are provided in Table S15. ptDNA copy numbers were quantified for the plastid genes *rbcL*, *psbB* and *ndhI*, which are topographically well separated on the plastid genome, and normalized to three nuclear loci M02, M19, and *pgiC* [43,61]. The nuclear loci are only present once in the nuclear genome of johansen Standard, as judged from coverage analysis of Illumina libraries. Additionally, three markers for mitochondrial DNA (mtM03, mtM04, mtM06) [62] were included in the calculation. Data were analysed with the LightCycler® 480 software v1.5.0 SP4 (Roche Diagnostics GmbH, Mannheim, Germany) employing the

"Advanced Relative Quantification" method that incorporates primer efficiencies. Target/Reference values calculated by the software were used to determine the proportion of ptDNA per total DNA (including ptDNA, mtDNA and nuclear genome) by employing the approximate size of the nuclear (C1 about 1 GB) [44], mitochondrial (about 400 kb) [62,63] and chloroplast (about 160 kb) [24] genomes, respectively. For illustration, arbitrary units were calculated by setting the I-johSt 5 DAS values of the plastid targets (*rbcL*, *psbB*, *ndhI*) to 1. Hence, the relative amount of ptDNA for all other genotypes and developmental stages is expressed as "fold I-johSt 5 DAS": *((Target^{pt}/Ref)*plastome size [bp]/((Target/Ref)*plastome size [bp] + C1 nuclear genome size [bp] + (Target^{mt}/Ref)* chondriome size [bp]))/(value 5 DAS).* Significance of the differences between I-johSt and IV-atroSt/plastome I variants for each developmental stage was calculated with a one-sample *t*-test followed by multiple testing *p*-values correction according to Benjamini-Hochberg.

Detection of RNA via RNA gel blot analyses

Total RNA was isolated as described previously [23] and 2.5 µg separated in 1% (w/v) agarose gels under denaturing conditions and transferred to a nylon membrane (Hybond™-N; GE Healthcare, Chicago, IL, USA) by capillary transfer with 10x SSC (1x SSC: 0.015 M sodium citrate, 0.15 M NaCl). RNA was covalently linked to the membrane using a UV cross linker (Vilber Lourmat, Ebertharfzell, Germany). Transfer success and equal loading was visualized by methylene blue staining (0.03% methylene blue, 0.3 M Na-acetate). Membranes were incubated in Church buffer [0.5 M NaH_2PO_4, 7% SDS (w/v), 1 mM EDTA; pH 7.2] for 1 h. Hybridization was performed after addition of radiolabeled *accD* or *ycf2* probes overnight at 65°C. Double-stranded DNA probes were obtained by PCR amplification with gene specific primers AbaccDfor (5´-AGTATGGGATCCGTAGTCGG-3´) and accD_cons_rev (5´-ATTCAGCCGTTTGTGAACCCTC-3´) for *accD*, and the primers ycf2Vno6 (5´-TAATGATCGAGTGACATTGC-3´) and Ycf2_VP30rev (5´-CTCTTCGTCTTCCTCTTCAAGC-3´) for *ycf2*. Purification of PCR fragments was done with the NucleoSpin® Gel and PCR Clean-up Kit (Machery-Nagel GmbH & Co. KG, Düren, Germany) and random-prime ^{32}P-labelling with the Megaprime DNA Labelling System (GE Healthcare, Chicago, IL, USA). Membranes were washed first with 1x SSC/0.2% SDS and two times with 0.5x SSC/0.2% SDS for 20 min each prior to exposition overnight. Radioactive signals were visualized with Typhoon TRIO+ ImageQuant (GE Healthcare, Chicago, IL, USA).

Supplementary Text

1. Generation of green variants with altered inheritance strength

For functional validation of loci predicted by correlation mapping in the wild types (see Main Text and below), mutagenesis of the strong chloroplast genome I-johSt was conducted using the *Oenothera plastome mutator* (see Materials and Methods for details). Inheritance strengths of the obtained green variants were determined in crosses to the white chloroplast mutants I-chi or IV-delta as pollen or seed parent, respectively (Materials and Methods for details, Fig. 2, Table S8, and Main Text).

The progeny of three seasons were analysed for the two crossing series. From these experiments it appeared that the lines VC1 and V3g (together with the weak wild type IV-atroSt) have very low assertiveness rates in the F1. As already judged by eye, they form a distinct class from all other variants in both crossing directions (Fig. 2; Table S8). For the reciprocal cross, no significant differences to the strong wild type I-johSt were found for the variants V1c, V2f, and V3e. Interestingly, these variants had the same transmission efficiency as the wild type, although they underwent a mutagenesis approach and carry a mutational load. This makes them a particularly valuable material to identify *plastome mutator*-induced mutations that do not affect chloroplast inheritance. All other variant plastids showed a significantly decreased competitive ability from at least one parent when compared to the wild type chloroplast genome I-johSt. In general, the plastome I variants cannot be grouped easily into the distinct classes of Schötz (strong, intermediate and weak). Besides VC1 and V3g forming a separate weak group (see above), their transmission frequencies steplessly align between the strong and intermediate wild types (cf. I-johSt and II-suavG of the I-chi cross; Fig. 2A). Hence, only two classes, a weak and a strong/intermediate one are present in the variants (analysed below in more detail).

The results of the classical experimental set up in which bleached chloroplast mutants are used, could be confirmed in a MassARRAY® approach employed for the crosses of the plastome I variants with the strong wild type plastome I-hookdV as male parent or the weak wild type IV-atroSt as the female parent (see Materials and Methods for details). Due to the detection threshold of the method (5-10%) when I-hookdV was transmitted through the pollen, most variants showed the same or slightly decreased transmission efficiency as their wild type I-johSt, with the progeny having increased amounts of paternal plastid DNA (ptDNA). However, only for VC1 and V3g is the difference of the ratio of paternal and maternal cpDNA in the pool large enough to result in the detection of a significantly lower assertiveness rate (Fig. 2B). The lines behave similarly in the other crossing direction, where most variants seem to be of wild type competitive ability. Again VC1 and V3g can clearly be confirmed as weak lines, while V3c and V3f (which appear as weak to intermediate when contributed by the female), show a higher transmission efficiency than I-johSt. This is the same reciprocal difference that is observed also in the classical experimental set up (Fig. 2). Altogether,

especially due to the detection limit, the classical approach using bleached chloroplast mutants gives more reliable results and allows a much finer discrimination of transmission efficiencies. Moreover, there is no qualitative difference in the assertiveness rates between the green wild types I-hook/IV-atroSt and their corresponding bleached mutants I-chi/IV-delta. This is in agreement with the classical literature [7] and investigated in more detail below.

2. The green plastome I variants do not display impaired growth, altered chloroplast morphology, or a photosynthetic phenotype

To rule out the possibility that the observed differences in chloroplast inheritance strength result from secondary effects in the green variants we performed several controls: First, we monitored growth behaviour of plants with the green chloroplasts of different inheritance strength in the common nuclear background of johansen Standard. Second, to access the physiological status of the material, we measured photosynthesis parameters. Third and last, we performed detailed microscopy to investigate chloroplast size, number per cell and morphology.

2.1 No growth, germination or macroscopic phenotypes are present in the plastome I variants

To ensure that the green variants are not impaired in development, cultures of johansen Standard plants harbouring various variant chloroplasts were compared side-by-side to plants with their strong wild type chloroplast genome I-johSt and the weak one IV-atroSt. It appeared that seeds from all plant lines germinated at 100% within 3 days after sowing (DAS). After transfer to soil, no differences in growth were observed during whole plant development under standard greenhouse conditions. Also no macroscopic phenotype such as altered leaf coloration was observed (Fig. S3).

2.2 Photosynthetic parameters are unaltered in the plastome I variants

To gain insights into the physiological status of our materials, we determined several photosynthetic parameters and plotted them against competitive ability (Fig. S4). From these analyses it became clear that differences in photosynthesis capability, if present at all, cannot be interpreted as a function of inheritance strength: We could not detect significant differences between plants nor dependencies of inheritance strengths on chlorophyll content per leaf area or for chlorophyll a/b ratio. The latter reflects the ratio of the photosynthetic reaction centres (exclusively binding chlorophyll a) to the antenna proteins (which bind both chlorophyll a and b). Also F_V/F_M, the maximum quantum efficiency of photosystem II (PSII) in the dark-adapted state, did not show any changes with inheritance strengths. All measured values were above 0.8 indicating that PSII was intact and that its antenna proteins were efficiently coupled to the reaction centre. There was a minor tendency towards a

decrease of leaf respiration in darkness with higher assertiveness rates. However, neither for leaf assimilation rates measured at the growth light intensity of 200 µE m^{-2} s^{-1}, nor for assimilation capacity measured under light-saturated conditions, were changes dependent on competitive ability observed. Similarly, for other photosynthetic parameters tested, including leaf absorptance, the chlorophyll a fluorescence parameters qN (non-photochemical quenching, a measure for the thermal dissipation of excess excitation energy in the antenna bed of PSII) and qL (a measure for the redox state of the PSII acceptor side), no clear differences dependent on competitive ability were found.

2.3 Chloroplast sizes, numbers or volumes per cell are unchanged in the plastome I variants

To test if differences in competitive ability are a side effect of a putative chloroplast division phenotype, our strong wild type I-johSt was compared to three lines with weak transmission efficiency (VC1, V3g and IV-atroSt). For this, we performed light microscopy using DIC optics at a developmental stage where the first 3-4 true leaves have developed (25 DAS). The chloroplast numbers in spongy mesophyll cells varied between 40-70 plastids per cell. This variance was found in all lines. None of the analysed weak lines showed a significant differences compared to I-johSt in chloroplast morphology, average chloroplast number per cell or chloroplast volume index (length x $width^2$) [60]. For chloroplast number per cell, one-way analysis of variance (ANOVA) yielded $p = 0.93$ among all four lines (Fig. S5). Adjusted *p*-values obtained by *t*-test and multiple testing correction in the comparison of each single line with I-johSt again did not uncover significant differences (Table S13). Very similar results were obtained for the chloroplast volume, for which one-way ANOVA gave a value of 0.51 in the comparison of all four lines. Comparing I-johSt with the weak plastomes also did not uncover significant differences, as judged from multiple *t*-testing (Table S14)

3. Correlation mapping

As described above, the green variants do not display any phenotype other than an altered inheritance of the chloroplast in crosses. Together with the wild type chloroplasts of different inheritance strengths, this makes them a valuable material to pinpoint molecular loci for chloroplast transmission encoded on the plastome. In contrast to algae or fungi, however, organelle genomes of higher plants or animals are not amendable to linkage mapping [1]. Consequently, in these materials, identification of functional relevant loci can only be based on correlation of a polymorphism within a given sequence interval to a phenotype in a mapping panel. To our best knowledge, this has been done only manually so far [64,65], which somewhat limits these analyses to a manageable number of organelle sequences, as well as to simple phenotypes, such as the presence or absence of sterility [66]. We therefore developed a novel mapping approach that fills this methodological gap. Conceivably, this approach

could be applied to map loci conferring cytoplasmic male sterility [66], mitochondrial diseases [67], cytonuclear incompatibility or to analyse adaptive cytoplasms [21,68,69].

The method is based on Spearman's rank and/or Pearson's correlation (Materials and Methods), with the latter capturing linear dependencies more directly. Since (i) presence or absence of linear dependencies in our data structure is a matter of speculation, and (ii) as a rank-based correlation metric, Spearman correlation yields more statistically robust results that are less influenced by outliers, we have used both approaches. For this, we calculated sequence divergence (total count of nucleotide changes, i.e. SNP, insertions and deletions) in respect to a reference sequence for every sequence in an alignment at a given alignment window. The value thus obtained is then correlated with a phenotype. In our case, this is a class of inheritance strength or a percentage value expressing transmission efficiency of a given chloroplast genome (see above and Materials and Methods for details). For example, if the reference sequence represents a strong chloroplast genome and, relative to it, certain weak plastomes contain polymorphisms in the same alignment window, this window is identified as highly relevant for inheritance strength (Fig. 1A; Fig. S17). Subsequently, individual polymorphisms or regions within this window are analysed separately (Fig. 1B). Since full organelle genomes are analysed, more than one relevant site can be identified. However, as for any other association mapping approach, the presence of two or more genetically independent loci that confer the same phenotype can complicate the conclusion. Perfect correlation coefficients of 1 or -1 might not be achievable at a single site. On the other hand, in a non-recombining system, such as an organelle genome, even absolute correlation at a single site may be due to genetic hitchhiking via linkage disequilibrium, and not necessarily due to functional relevance. Hence, experimental verification of predicted loci by independent methods is necessary.

3.1 Division of wild type and mutant plastomes into classes of inheritance strength

The datasets that measure inheritance strength of wild type chloroplasts or of the green variants were either deduced from the literature or produced in this work. They represent percentage values of heteroplasmic seedlings in an F1 generation that reflect inheritance strength of a given chloroplast genome (Tables S8 and S11; see above). The numbers can be directly applied to Spearman's/Pearson's correlation. If datasets of more than one crossing series are to be combined, clustering of the crossing data into classes is necessary.

For the wild type plastomes, we used the original data of Franz Schötz, where two sets of crosses "biennis white" and "blandina white" are available [6,33] (Table S11; Materials and Methods); inheritance strength of 25 wild type chloroplasts was determined using these previously described tester lines. Clustering of the two datasets with the *k*-means algorithm using the optimal number of

centres (*k* = 3) confirmed the original classifications suggested by Schötz, with the exception of the I-bauriSt and II-corSt plastomes (Fig. S16A; for details see Materials and Methods). These plastomes were borderline genotypes in Schötz's classification system, and according to our data, they might be reassigned. Besides these minor discrepancies, clustering conclusively supports the presence of the three distinct classes of inheritances strengths (strong, medium and weak) in the wild type plastomes of *Oenothera*, as previously described.

The clustering of the green variants is less clear. When data from the I-chi and IV-delta crossing experiments are combined, the pamk function identified *k* = 2 as the optimal number of clusters, clearly separating the weak from the stronger materials (Fig. S16B). However, finer clustering of the stronger variants leads to ambiguous class membership. This is likely due to the higher variation in the IV-delta crosses compared to the I-chi crosses (Fig. 2 and above). This seemingly weakens the combination of the two datasets. Finally, we chose four clusters to classify the variants: (i) At this number of clusters, individual samples swap at the lowest rate between the clusters, when *k*-means (a clustering method that includes a random element) was applied repeatedly. (ii) This number of classes best reflects the transmission abilities of the variants, since many variants are somewhat in between the strong and the intermediate genotypes (see positions of the strong and intermediate wild types I-johSt and II-suavG in the I-chi crosses; Fig. 2A and above). This justifies the definition of a fourth class and material of that kind does not exist in the wild types (cf. Fig. S16A vs. 16A). Hence, based on the wild type classes, the variants are classified as follows: class 1 = strong, class 2 = strong to intermediate, class 3 = intermediate, and class 4 = weak.

3.2 Correlation mapping in the wild type plastomes

Pearson's correlation generally identified more windows than Spearman's, but both predict essentially the same regions relevant for inheritance strengths. Interestingly, there was no notable difference between the methods if either *k*-means classes or the "biennis/blandina white" crossing data were used for correlation (Fig. 1A, Fig. S6, and Data S1). Largely based on theoretical considerations (presence of three clearly ranked classes in the wild types and stronger experimental base if the "biennis white" and "blandina white" crossing experiments are combined; see above), we discuss here Spearman's rank correlation to *k*-means classes in more detail. According to the latter, sequence windows in the *ycf1* and *ycf2* genes (between alignment positions 99011-100000 and 134641-135640) show nearly absolute correlation to inheritance strengths (*rho* = -0.99, $p < 0.0005$; Data S1). In both genes, the correlation oscillates from *rho* = 0.86 to -0.99 ($p < 0.0005$), and the positive and negative correlation should be interpreted as equally important. Another nearly absolute correlation (*rho* = 0.98, $p < 0.0005$) was measured in alignment windows containing the promoter, 5'-UTR and 5'-end of

accD (positions 63501-64760). Window further upstream containing the same features (positons 63391-64490) also correlates with *rho* = 0.96, $p < 0.0005$ (Fig. 1B). However, highly significant correlations of 0.96 were also found in intergenic regions of photosynthesis genes and/or tRNA genes, for example between *ycf3* and *psaA* (encoding a photosystem I assembly factor and core subunit, respectively) [70]. In addition, significant correlations were measured from the spacers of the photosystem II and cytochrome b_6f subunit genes *psbE* and *petL*, and in a sequence interval contacting *trnR-UCU* and *trnG-UCC*. In contrast, no significant correlation was observed for *oriA*. For *oriB*, three sequence windows (partially) containing the *oriB* correlate with 0.90, 0.88 (both $p < 0.0005$) and 0.81 ($p< 0.005$). If Pearson's correlation to *k*-means classes is applied to the wild type data, the described pattern can be reproduced but more windows with significant correlation are identified (Fig. S6 and see above). The highest observed Pearson correlation in the wild type dataset is $r = 0.96$ ($p < 0.0005$) in a sequence window again containing the promoter, 5'-UTR and the 5'-end of *accD* (Fig. 1B; Data S1).

3.3 Correlation mapping in the green variants

When correlation mapping results are compared between the wildtype and the green variants, the most striking difference in the variants is the loss of significance after *p*-value adjustment for Spearman's but not for Pearson's correlation. Here, windows with significant correlations were obtained (cf. Fig. S6 vs. 7; Data S1). This is probably because the rank-based Spearman correlation being less influenced by the VC1 and V3g data points. These two single genotypes, however, form the weak and, therefore, most predictive class, whereas the other variants do not differ noticeably from their wild type progenitor (cf. Fig. 2 and Fig. S16B; also discussed above). This leads to a relatively weak correlations to inheritance strengths which appears to be an under-estimation and a consequence of the multiple testing correction (> 13,000 tests). The weak correlations also contradict the observations, which clearly indicate that the plastomes of the green variants must contain mutated loci for inheritance strength. A similar argument applies to correlation of the *k*-means classes in the variants. As discussed above, definition of these classes is less clear than in the wild type, which weakens their predictive power. We therefore think that the Pearson correlation of the I-chi crosses (which yield a better resolution than the reciprocal IV-delta crosses; Fig. 2 and above) represents the best approach to identify the relevant loci that alter inheritance strength in this material. Notably, this approach yields the most significant correlations, but all approaches (including Spearman) identify the same regions in the plastome with the highest correlation values (Fig. S7; Data S1).

In the variants, Pearson's correlation of the I-chi crosses predicts a sequence window in the 5'-end of *accD* as significantly correlated to inheritance strengths ($r = 0.78$ and $p < 0.005$). The

strongest correlation for this dataset is observed for the 5'-UTR of *ycf2* (r = 0.91, p < 0.0005). In addition, a highly repetitive region in the coding region of the same gene also shows good correlation values (r = 0.71; p < 0.05; Fig. S7; Data S1). Two insertions/deletions (indels) in *ycf1* are also found significant (r = 0.61; p < 0.05). They represent a single insertion and a deletion in the weak variant VC1, located relatively close to each other (see below). The functional relevance of the two mutations for inheritance strength can be questioned, however. Since the second weakest variant of the dataset displays a wild type *ycf1* sequence (Fig. S10A), the above described mutations are likely a result of the higher mutation load present in VC1 (see Materials and Methods for details). For the same reason, a contribution to the phenotype by the *oriB* can be excluded (r = 0.41, p = 0.25) in the variants. Taken together, our results narrow down the regions identified in the wild types to the two genes *accD* and *ycf2*. All other mutated loci in the variants seem to be of minor importance.

3.4 Correlation analysis at selected loci

When correlation mapping is applied to selected loci within the above identified alignment windows, the general observation is that correlation values drop to some extent (cf. Data S1). This is probably best explained by looking at the highly correlating sequence intervals spanning the promoter, 5'-UTR and 5'-end of *accD* in the wild types (Fig. 1B). When analysed as functional units (promoter/5'-UTR region and protein N-terminus; Fig. S8), correlation of the individual segments (promoter/5'-UTR region: r = 0.80 or *rho* = 0.74; p < 0.005 for both; full N-terminus: r = 0.78 or *rho* = 0.60; p < 0.005 or p < 0.05), is much lower than for the original sequence intervals (r = 0.94 or *rho* = 0.96 and r = 0.96 or *rho* = 0.98 with p < 0.005 for all), which led to the identification of these regions. As discussed below, experimental evidence is available that promoter/5'-UTR and N-terminus interact to affect the inheritance phenotype, while the individual regions display weaker correlations.

In spite of these complications, to get an impression how well certain coding or promoter/5'-UTR regions, as well as segments of *oriB* correlate with inheritance strengths, we calculated correlation values for *accD*, *ycf1*, *ycf2*, and *oriB* for polymorphisms that are present in both wild type and the variants (Figs. S8-S10; Data S1). We also included two prominent sites in the *ycf2* gene present in the wild type, for which we found no mutation in the variants. Please note that at three sites (AccD N-terminus, the AccD site 2 and the *ycf2* promoter/5'-UTR region) in addition to Person's correlation the Spearman's correlation analysis yields significant correlations in the variants. This is in contrast to the whole plastome approach described above, where p-values corrections due to multiple testing were applied.

The best correlating region in both sequence sets (wild type and variants) is site 2 of the AccD N-terminus (Fig. S8B). Its prediction is extremely robust in that significant Pearson's and Spearman's

correlation were obtained for all crossing series and *k*-means classes (Fig. S8; Data S1). Less clear is the contribution of *ycf2*. In the wild types, the *ycf2* site 1 and site 2, but not site 3 can be associated with inheritance strength, but in the green variants site 3, exerts the most influence on the competitive ability of chloroplasts (Fig. S9B and below).

In summary, the refined analyses at selected loci clearly confirm the contribution of *accD* on inheritance strengths and might have even identified the most important region. It also shows that *ycf2* may contribute to the phenotype. Without chloroplast transformation in the evening primrose, a technology currently not available, the influence of the individual sites remains speculative.

4. Repeat structure, sequence evolution and divergence of *accD*, *ycf1*, *ycf2,* and the *oriB*

The four genes or loci partially span rapidly evolving regions of the *Oenothera* plastome that are characterized by large repetitive regions. Those can be of up to 1 kb in size as it is the case for site 3 in the *ycf2* gene. They are comprised mostly of tandem or direct repeats (and less pronounced of palindrome or inverted repeats) as described earlier [24]. Due to their repetitive nature these regions are very prone to replication slippage [23,71] and sequence divergence at these regions substantially contributes to the overall sequence variation of the *Oenothera* chloroplast DNA (Greiner et al. 2008. Fig. 3 therein) [24]. The presence of repeats also makes them a preferred target of the *plastome mutator* allele [16,39].

Sequence evolution is extremely fast at these repeats. In case of the repetitive regions of *accD* and *ycf1* phenotypically neutral spontaneous mutations were isolated repeatedly at very similar sites [23]. Moreover, the *oriB* (which is essentially located in the *rrn16* - *trnI-GAU* spacer) is used as hypervariable marker allele that allows discrimination among a huge variety of *Oenothera* strains [43]. The repeat structure of the *oriB* region was analysed earlier and is comprised of 7 direct repeat classes that can be divided into various subtypes [16,17] (and below). In the *accD* gene mostly tandem or direct repeats span the promoter/5'-UTR and N-terminal region (Fig. 1); all three are considered to contribute to the regulation of the gene [15,26-29]. In fact, sequence variation induced by theses repeats is so high that upstream of the *accD* start codon a window of about 1.4 kb cannot be aligned between the weak plastome IV and the stronger plastomes I-III (Data S2). This is to some extent also observed for site 2 in the N-terminus of the wild type AccD. In plastome IV major portions of this site is missing and about half of the remaining sequence is polymorphic (Fig. S1A). In the *ycf2* coding sequence, the most prominent repeats are present in site 2 and 3. At the first site the number of PEKRKEKK tandem repeats can be correlated with inheritance strengths in the wild types, but not in the variants. The situation is reversed for site 3, in which tandem repeats exist as two subtypes 5'-

GAGGAAGtAGAAGGGACAGAA-3' and 5'-GAGGAAGgAGAAGGGACAGAA-3' associated with a GAT linker, and correlate with inheritance strengths in the variants, but not in the wild types (Fig. S2).

5. Variation at the chloroplast origins of DNA replication is not causative for chloroplast competition

As elaborated above, our correlation mapping already points to a connection of lipid biosynthesis and chloroplast competition. However, one might still argue that *a priori* differences in the origins of replication are the simplest mechanistic explanation for organelle competition. At least some evidence supporting this claim is available for yeast and *Drosophila* [9,72,73]. The location and repetitive nature of the *oriB* in evening primroses (see above) are reminiscent of the non-coding displacement loop (D-loop) of metazoan mitochondrial DNA (mtDNA). In many animal taxa the D-loop is the most variable sequence of mtDNA and is in the proximity of tRNA or rRNA genes [9,74,75].

5.1. Sequence variation in the oriB cannot explain differences in inheritance strength

Previous work in the evening primrose did not support an involvement of the origins of replication in chloroplast competition. First, the number of D-loop initiation sites (i.e. *oris*) does not differ between weak and strong plastomes and their locations in the chloroplast genome is identical [18]. Second, in a previous association mapping study that investigated the hypervariable repeat region of *oriB*, a short repeat series was identified as the sole determinant that could explain the difference between the strong and intermediate plastomes I, III and II on one side, and the weak plastome IV on the other side [17]. The sequence (5'-ACGACACGACGATTAGATTAGCTCATTGGTAGGACGACGATTAGCTCATTGGT AGGACGACG-3') is 62 bp in size and is capable of forming of weak hairpins. Our study, analysing a greater number of plastome sequences, confirms the absence of this sequence in the weak plastome IV. However, in none of the green plastome I variants with alerted inheritance strength is the sequence partially or fully deleted (Data S2). We therefore do not think that a genetic determinant within the *oriB* of *Oenothera* is able to explain the observed huge differences in competitive behaviour.

To substantiate this view, also on the level of DNA, we investigated the dynamics by which ptDNA increases during plant development in more detail. In addition, we analysed chloroplast nucleoid structure and number per cell.

5.2 Changes in plastid DNA amounts during development do not correlate with inheritance strength

To investigate if potential differences in ptDNA increase during development and/or changed ratios of plastid/nuclear DNA are able to explain chloroplast competition, we performed quantitative real-time PCR. In general, plastid DNA amounts are not static during ontogenies [58,76]. They increase as leaves grow, starting from 0.4% in meristematic tissue to more than 20% in mature leaves [57]. If differences were observed in DNA abundance during development in different *Oenothera* lines

harbouring chloroplast with different inheritance strengths, it might hint towards DNA replication or replication speed as an underlying mechanism for plastid competition. To monitor this process we analysed total DNA of the johansen Standard strain equipped with the strong and the weak wild type chloroplast I-johSt and IV-atroSt, respectively. In addition, we included selected lines of our plastome I variants: V1c, V2a, V2g, and V3e (all strong), V3c (strong to intermediate), V3d (intermediate) and VC1, V3g (all weak). Plants tissue of different developmental stages was analysed, seedlings 5 DAS (when seeds have just germinated and cotyledons have developed), plantlets 21 DAS (after development of the first two true leaves), and the second true leaf of young rosettes 32 DAS. ptDNA amounts were calculated relative to I-johSt (5 DAS) per one haploid genome (for details see Materials and Methods).

Depending on the plastome target region, at 5 DAS a small increase of ptDNA amounts was observed in the lines V1c, V3e, V3c, VC1 and V3g. These differences, however, are not significant. At 21 DAS the weaker variants V3c, V3g and VC1 showed an increase in relative ptDNA amounts compared to wildtype I-johSt, but only for the target *ndhI* which was again not significant. In general, from 5 to 21 DAS only a minor or no increase of plastid DNA amount was observed for each particular line and each plastid target, while in the young rosette at 32 DAS the ptDNA amount doubled (Fig. S11). These results echo previous work in *Arabidopsis* and sugar beet [57,76]. Since the same results were obtained for plant lines carrying strong and weak plastids, no developmental differences in ptDNA copy numbers correlates with differential transmission efficiencies.

In summary, all minor increases in ptDNA amount are not significant nor correlate with transmission efficiencies nor with the DNA variations described previously (Fig. S10B). Moreover, no differences can be detected in IV-atroSt compared to wildtype I-johSt, although plastome IV is the weakest of all genotypes tested. Therefore, the ptDNA amounts in vegetative tissues do not indicate different replication speeds, suggesting that replication *per se* is not the underlying mechanism for different transmission efficiencies.

5.3 Nucleoid number and structure is identical in lines with different inheritance strength

Under the premise that ptDNA amounts are constant, there is still the possibility that strong and faster replicating plastomes have altered numbers of nucleoids, which could impact their ability to divide. To exclude this possibility we quantified nucleoids in the central laminal region of the first true leaf 25 DAS. After staining with DAPI, nucleoids were clearly visible as small dots with their fluorescence sharply delimiting them from the dark cellular background, even when forming tight associations like clumps or threads (Figs. S12 and S13). One strong (I-johSt) and three weak lines (V3g, VC1, and IV-atroSt) were investigated. The mean number of nucleoids per chloroplast ranges between 17.7 and

18.1 with no significance differences between the lines. One-way ANOVA gave $p = 0.53$; multiple *t*-tests comparing I-johSt with each of the weaker plastomes did not point to significant differences as well (Fig. S14 and Table S12). Moreover, we did not observe any difference in nucleoid morphology between the lines.

6. Expression and transcript maturation of *accD* and *ycf2*

Since the polymorphisms in *oriB* cannot explain differences in competitive ability, we investigated the accumulation of *accD* and *ycf2* transcripts. For this, we used leaves of plants 26 DAS from the lines I-johSt, V1c, V3e, V2g (all strong), V3c (strong to intermediate), and VC1, V3g and IV-atroSt (all weak; Fig. S14). A probe specific for the conserved C-terminal part of *accD* detected the mature transcript at about 3 kb. No differences in transcript accumulation were observed between I-johSt and the plastome I variants. However, for IV-atroSt two additional bands running below the mature transcript were present. Moreover, the mature transcript clearly over-accumulated in this weak, but phylogenetically more distant plastome. This transcript over-accumulation appears to be a result of the high sequence variation observed in the *accD* promotor/5'-UTR that strongly correlates with inheritance strength (see Fig. 1, Fig. S8 and above). A similar analysis was conducted for *ycf2*, where again a probe specific for the C-terminal part of the gene detected the mature transcript at the expected size of about 9 kb. The very small differences in size between the lines perfectly mirrors the occurrence of in-frame deletions in the lines IV-atroSt, V3c, V3g, and VC1 (Data S2). In IV-atroSt as well as in the plastome I variants, no difference in accumulation of the mature transcript compared to I-johSt was found. However, transcript stability/processing seems to vary between the strong plastome I-johSt and the weak IV-atroSt. Interestingly, whereas the strong variants V1c, V3e and, to same, extent V2g showed exactly the same transcript pattern as the wildtype, the weak variants showed a pattern more similar to IV-atroSt. This indicates a correlation between transmission efficiency of mutations in site 3 of *ycf2* (cf. Fig. S2), that might result from altered mRNA degradation and/or processing.

7. ACCase activity in lines harbouring chloroplasts of different inheritance strength

As the above described analysis indicates an involvement of *accD* and/or *ycf2* in the inheritance phenotype, we decided to determine ACCase activity in our lines. From these measurements, it appeared that the strong variants (V1c, V3e, V2a, and V2g) display a similar or even lower ACCase activity than their wild type I-johSt. The same holds true for the strong to intermediate or intermediate genotypes (V3c and V3d). In the weak materials, however, a 2-3 fold increase of ACCase activity is observed for VC1 and IV-atroSt, although V3g shows wild type enzyme activity (Fig. 3A). Although

there is no simple linear correlation between inheritance strengths and ACCase activity, the strong increase in VC1 and IV-atroSt is hard to ignore. In fact, both inheritance strength and ACCase activity seem to depend on the particular mutation pattern: (i) Mutations in *ycf2* seem to influence ACCase activity, as judged from the variant V3e that is wild type for the *accD* segments but is mutated in *ycf2* (Fig. 3A, yellow box). (ii) Larger mutations in the AccD N-terminus have higher ACCase activity, whereas the presence of a more diminutive AccD N-terminus correlates with lower activity (Fig. 3A, cf. blue boxes vs. the remaining pattern). (iii) There must be an influence of *ycf2* on inheritance strength (cf. Fig. 3A, green boxes associated with the weaker materials). Hence, if ACCase and/or Ycf2 result in changes lipid levels, one would expect that lipid composition is predictive of inheritance strengths.

8. Predictability of inheritance strength based on lipid-levels

To test for predictability of inheritance strength from lipid level data, we analyzed 16 chloroplast genotypes of different inheritance strength in a LASSO regression model (Table S1; Materials and Methods). Since chloroplast inheritance strength is independent of photosynthetic competence (see below), we included pale lines. The aim was to enrich the lipid signal responsible for inheritance strength, i.e. to deplete for the structural lipids of the thylakoid membrane [77]. Namely, we used our bleached *psaA* mutants I-chi and IV-delta impaired in photosystem I assembly, as well as the pale green *virescent* genotypes III-lamS, III-V1 and III-V2 (Tables S1 and S7; Fig. S15). For such Methods perturbed thylakoid membrane formation was shown previously [78-81]. Moreover, as elaborated in the following chapters, we could confirm the independence of a pale phenotype to inheritance strength with these plastomes.

8.1. Chloroplast inheritance strength is independent from bleaching

Intuitively one might expect that bleached chloroplast mutants would be less successful in crosses than their corresponding green wild types. However, previous analyses in evening primroses showed that differences in chloroplast inheritance strength are largely independent of the chloroplast mutant used for the analyses [14,82]. At least for *Oenothera*, it is therefore generally accepted that mutations in a chloroplast genome that result in bleaching essentially do not affect chloroplast assertiveness rates [7] (also see Fig. 1B,D and above). Due to technical limitations, however, this hypothesis was never tested directly. Since closure of this gap is of general relevance for this work, and to provide further evidence that chloroplast inheritance strength is largely independent of the photosynthetic status of the chloroplast, we directly compared the wild type chloroplast I-hookdV and its bleached derivative I-chi, as well as IV-atroSt and the corresponding mutant IV-delta. For this, we investigated appropriate F1 populations crossed to the chloroplast genomes I-johSt, VC1, and V3g with the

MassARRAY® system (Fig. S18; see Materials and Methods for details on the material). As expected, transmission efficiencies were found of the same range for nearly all six pairs of crosses under investigation. Only in one cross with VC1 as a mother, the bleached mutant I-chi actually behaved even stronger than its corresponding green wild type.

Taken together, we could confirm that chloroplast assertiveness rates are independent of photosynthetic capability. Moreover, these results make it very unlikely that the differences in inheritance strengths observed for the mutated plastome I variants (all sharing a green phenotype), are due to a secondary effect.

8.2 The very weak variants III-V1 and III-V2

While the plastome I variants and their wild type I-johSt are native in and compatible with the nuclear background of the johansen Standard race, III-lamS and its *plastome mutator* variants III-V1 and III-V2 are foreign and incompatible in this background, meaning that tissues carrying them do not develop a normal green colour (Materials and Methods, Fig. S15, Table S9). The wild type III-lamS plastome appears to be strong, as judged from crosses to I-johSt as pollen donor, its derivative variants III-V1 and III-V2 are weak (cf. Fig. 2A vs. Fig. S19; cf. Table S8,S9 and S11) [14]. Although the fraction of plants showing biparental inheritance in the crosses of III-V1 and III-V2 to I-johSt as pollen donor are somewhat low (37.3% and 38.0%, respectively) for a combination of a weak and a strong plastome (Fig. 2, Table S8) [7], the striking difference of this cross to all other crosses described is that some seedlings inherit only paternal chloroplasts (Fig. S19; Table S9). As mentioned previously, biparental inheritance in the evening primrose shows maternal dominance, in which progeny are either homoplasmic for the maternal chloroplast or heteroplasmic for the maternal and the paternal chloroplasts, but they are never homoplasmic for the paternal one. The appearance of homoplasmic offspring having the paternal chloroplast in the III-V1/III-V2 crosses to I-johSt is the only reported case in the evening primrose where an exception to maternal dominance occurs. This justifies the definition of a new inheritance class for these plastomes.

8.3 Classes of inheritance strength employed in the LASSO regression model

To predict chloroplast inheritance strength from lipid-level data, the genotypes of the plants needed to be ranked according to their inheritance strengths (Materials and Methods; Table S1). For the green variants (V1c, V2a, V2g, V3e, V3c, V3d, VC1, and V3g) and the wildtypes I-johSt and IV-atroSt were used the existing *k*-means classes 1 - 4 already employed in our association mapping approach (see above). The remaining plastome (I-chi, IV-delta, I-hookdV, III-lamS, III-V1 and III-V2) were rendered consistent with this framework based on the classification of Schötz and our own data. This adds the

plastome I-chi, its wild type I-hookdV and III-lamS to the strong class 1. The mutant IV-delta was placed into the weak class 4. As a result of the exceptions to maternal dominance, when III-V1 and III-V2 were seed parents, these plastome are placed in a new class 5 (very weak). Taken together, the 16 genotypes are classified into 5 classes of descending inheritance strengths (strong = 1, strong to intermediate = 2, intermediate = 3, weak = 4, and very weak = 5). Material in class 1 and class 4/5 are over-represented, since (as for the inclusion of the bleached material; see above) we expect to enhance the signal for predictive lipids (Table S1)

8.4 Predictability of inheritance strength based on lipid-level data as explanatory variables

The rationale of the predictive approach is as follows: To test for predictability of inheritance strength based on lipid-level data, a linear model (LASSO) was trained and its performance tested in a cross-validation setting on two randomly selected genotypes with differing inheritance strength (see Materials and Methods). If the proposed regression model has predictive power, the actual inheritance strength-values associated with the two test genotypes should be positively correlated with their predicted ones. Note that for each genotype repeated measurements were available. Thus, regression was performed over more than two points and correlation coefficients could assume absolute values differing from 1. Testing was done in a cross-validation setting, i.e. the two test genotypes were not included in the model training. This procedure was repeated 100 times, with each run corresponding to two new randomly selected genotypes of differing inheritance strength and all others used for model training.

If, indeed, inheritance strengths can be predicted based on lipid levels, on average, a positive correlation (Pearson correlation coefficient, *r*) between actual and predicted inheritance strength-values of the two test set genotypes should be obtained. To test for this outcome, 100 Pearson correlation coefficients are classified as positive (success) or negative (failure). Then, they were compared to the null hypothesis of no predictive value, corresponding to a 50% chance of obtaining a positive correlation and significant deviations from this expected probability tested by performing a binomial test.

8.5 The lipid classes DGDG, PG, PC, and PE are enriched for predictive lipids

From the 100 cross-validation runs, using the combined dataset from three independent experimental series (Table S1), a median Pearson correlation coefficient (*cvR*) between actual and predicted inheritance strength values of cvR_{median} = 0.7 was obtained (Fig. 3A) with 82 being positive, i.e. successful predictions. This corresponds to $p_{binomial} = 2.17 \times 10^{-9}$ vs. the null hypothesis of 0.5 (no predictive value). Thus, lipid levels proved indeed predictive relative to inheritance strength.

Individual lipids were ranked with regard to their predictive value based on the coefficients by which they entered the regression model (Fig. 3C; Table S2). Averaged over all 100 cross-validation runs, 20 lipids/molecules were identified as predictive as judged by their average absolute weight. They were considered predictive if their absolute average weight was greater than one standard deviation (SD = 0.7) of the average weights of all 102 lipids/molecules. Among those, lipid/molecule classes DGDG, PG, PC, and PE were found enriched (odds ratio > 1), albeit statistical significance could not be established (Table S3)

8.6. A model for the predictability of inheritance strength based on lipid-levels

Taking into account the data of Fig. 3, we propose the following model to explain how certain changes in lipid abundance influence inheritance strengths: Increased activity of acetyl-CoA carboxylase in the chloroplast (Fig. 3A) increases fatty acid concentrations and subsequently fatty acid export to the endoplasmic reticulum (ER). The combination of increased fatty acid synthesis and export leads to an upregulation of the eukaryotic lipid biosynthesis pathway in the ER [83]. This is seen in increased amounts or shifted proportions of diverse lipid classes, including phospholipids or storage lipids. Since PC is the dominant lipid class of in the chloroplast outer envelope [19] those changes affect the structural and physiological properties of the envelope. It in turn impacts chloroplast division and/or stability processes, thus ultimately determining inheritance strengths (see Main Text). Conceivably, the Ycf2 protein, which is located in the envelope [84], participates directly in the transport of fatty acids or its function may be responsive to changes in the lipid composition of the envelope, thus influencing other transport processes and/or growth and division of the chloroplast. The observed shift in the proportions of storage lipids or other changes in the lipidome (cf. DGDGs or TAGs in Fig. 3C) might occur in response to altered fatty acid pools, although the storage lipids are likely not relevant for the variation in chloroplast inheritance strengths.

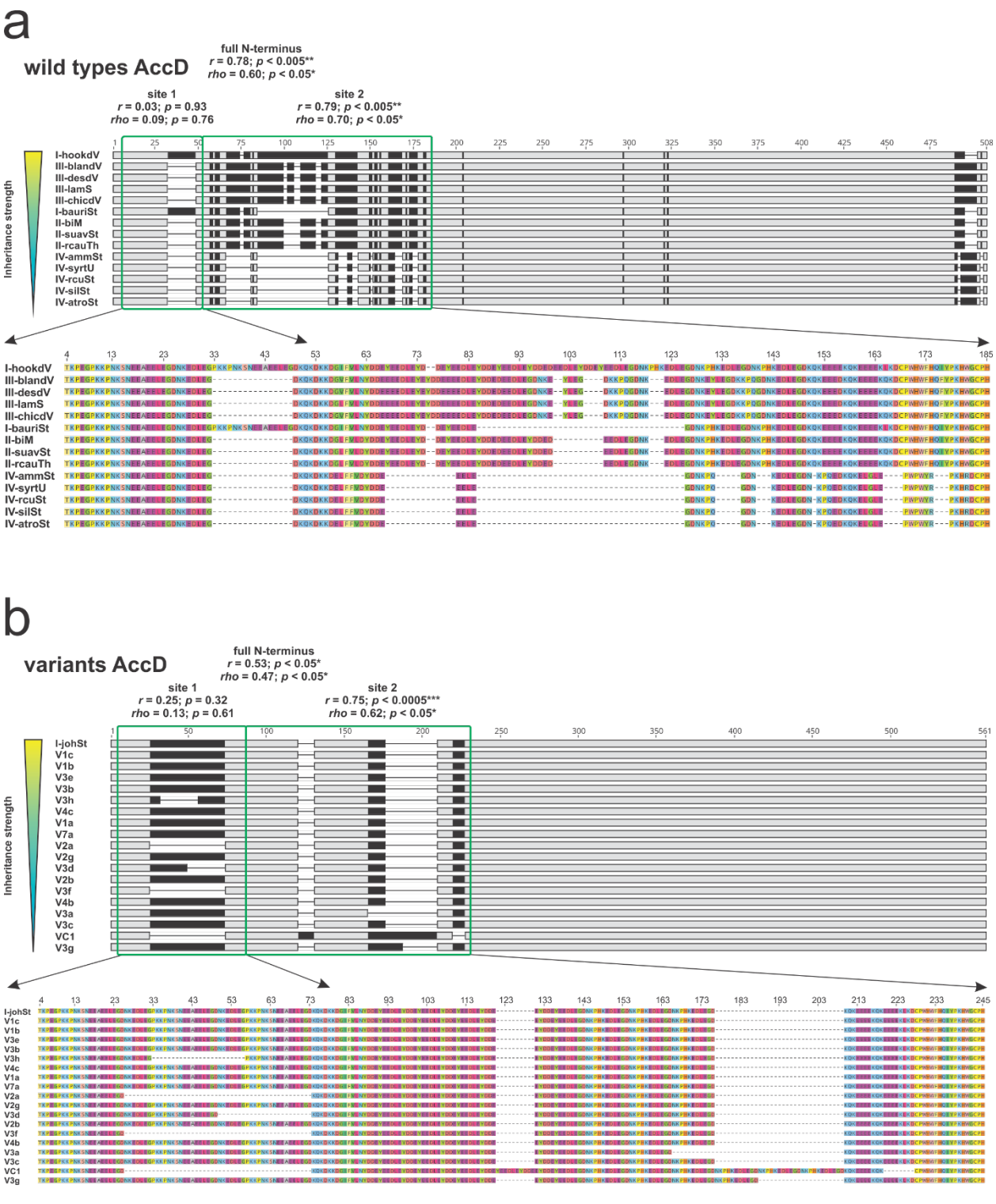
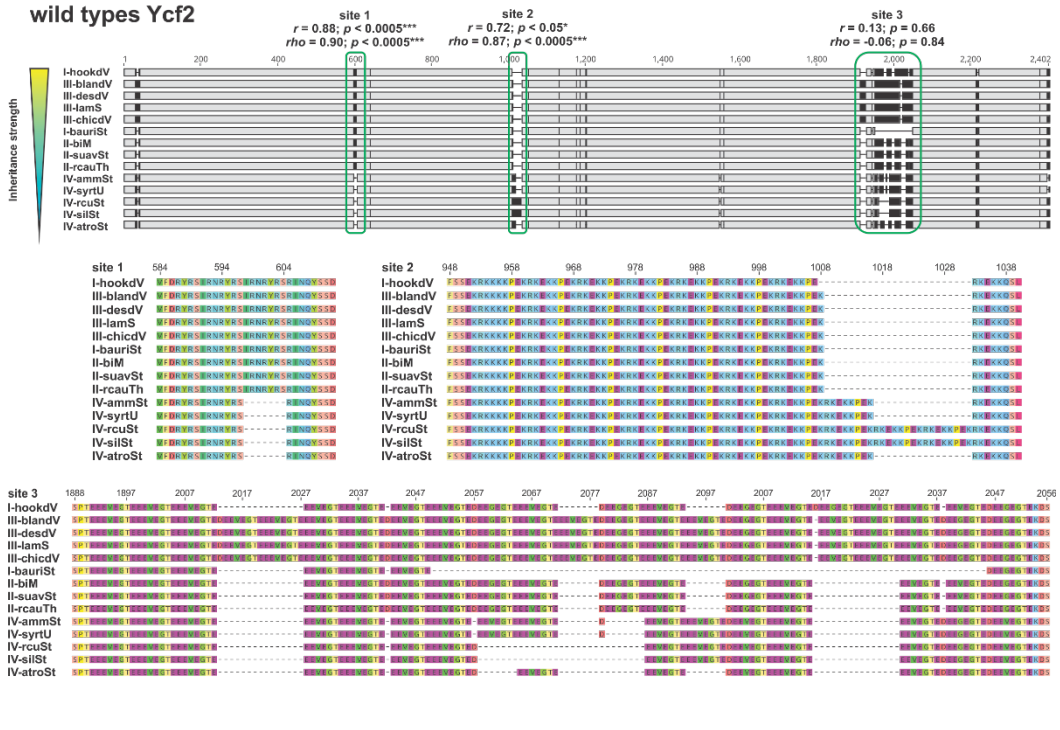


Fig. S1. Amino acid sequence of the AccD N-terminus and correlation to inheritance strength. Individual sequences are sorted according to their competitive ability. Disagreement to the consensus in a given alignment window is indicated in black, alignments of identical sequences in grey. (A) Wild types. (B) Variants. Sequence variation in both sequence sets is conferred by large tandem or direct repeats.

a



b

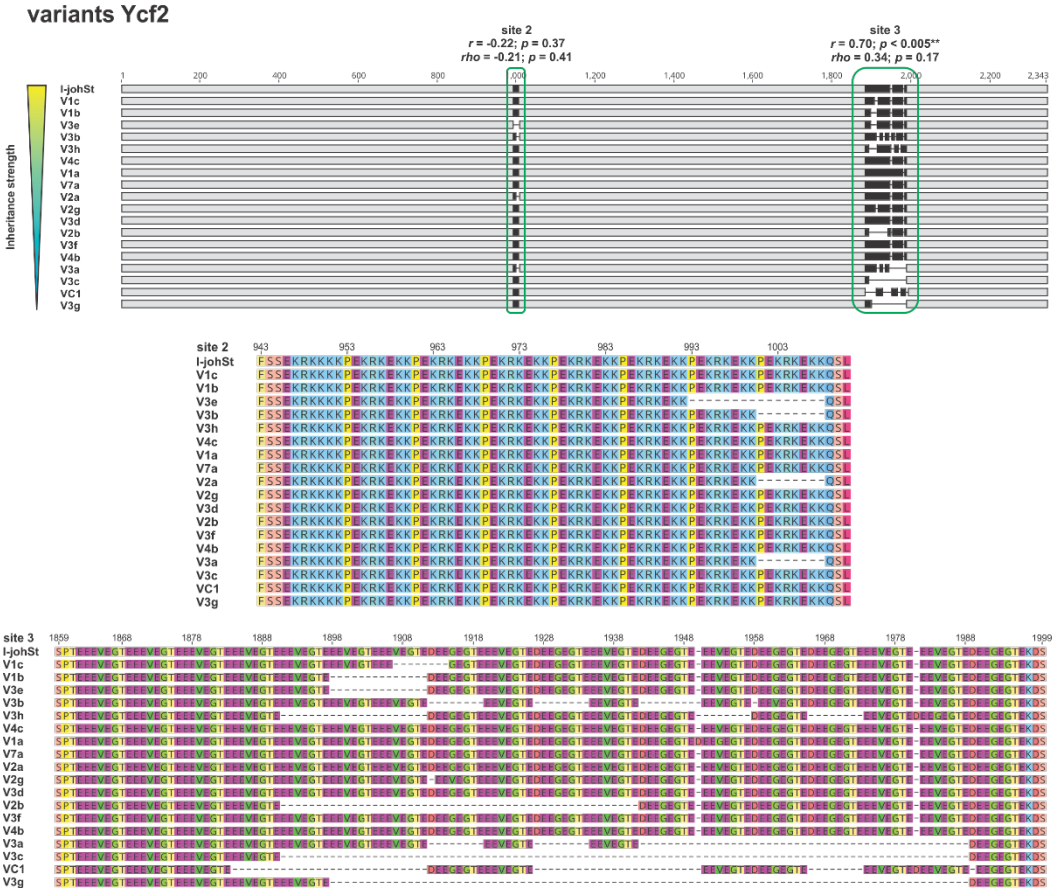


Fig. S2. Amino acid sequence of the Ycf2 protein and correlation to inheritance strength. Individual sequences are sorted according to their competitive ability. Disagreement to the consensus in a given alignment window is indicated in black, alignments of identical sequences in grey. (A) Wild types. (B) Variants. Sequence variation in both sequence sets is conferred by large tandem or direct repeats. Note that sites 1 and 2, but not site 3 correlate with inheritance strengths in the wild types, whereas multiple deletions in site 3 are associated with the weaker inheritance phenotype of the variants.

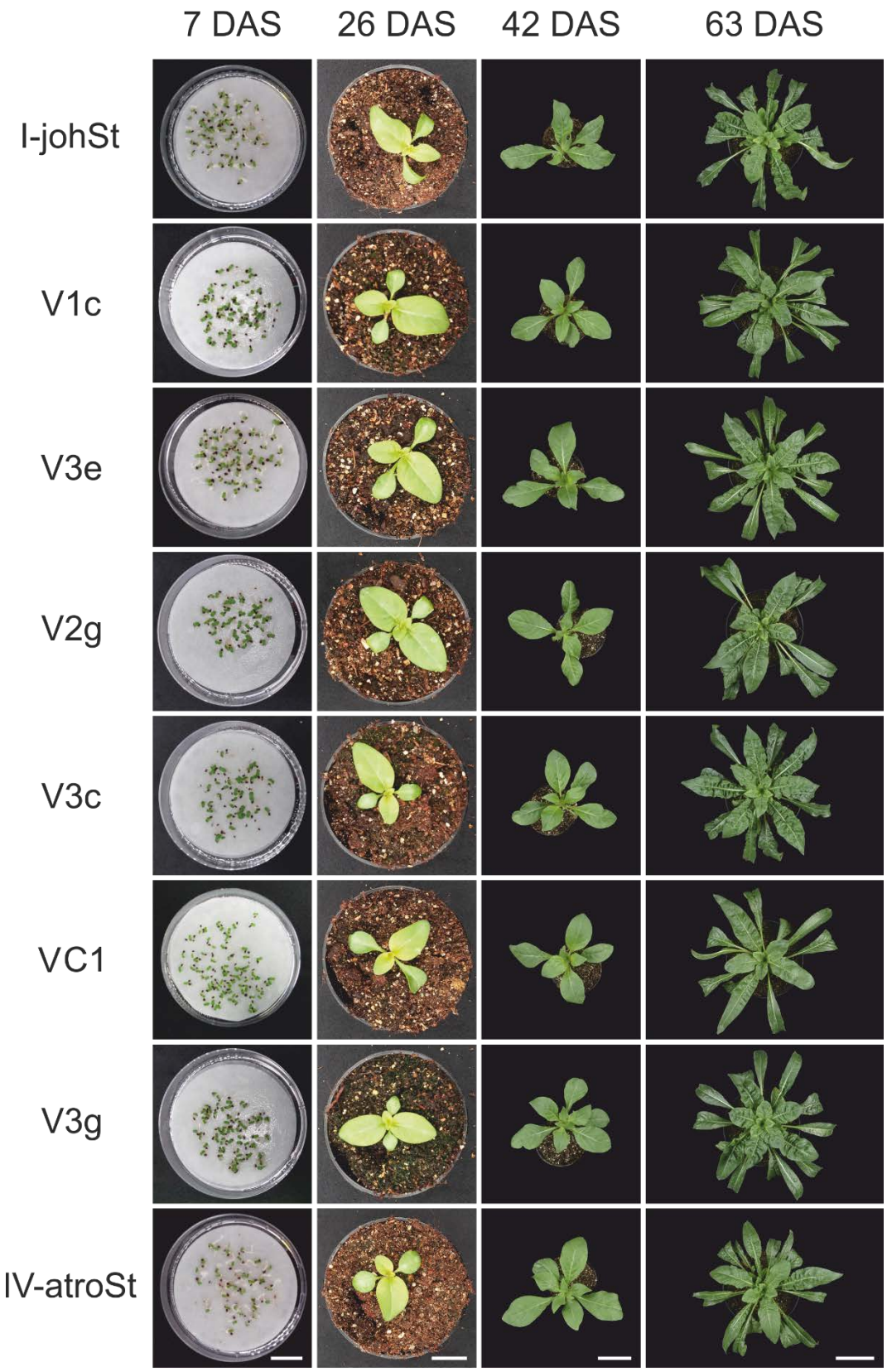


Fig. S3. Developmental series of johansen Standard plants with wild type or variant chloroplast genomes of different inheritance strength. Columns from left to right. First column: seedlings 4 days after geminating or 7 days after sowing (DAS); scale bar = 2 cm. Second column: plantlets 26 DAS; scale bar = 2 cm. Third column: end of early rosette stage, 42 DAS; scale bar = 4 cm. Forth column: mature rosettes, 63 DAS; scale bar = 10 cm.

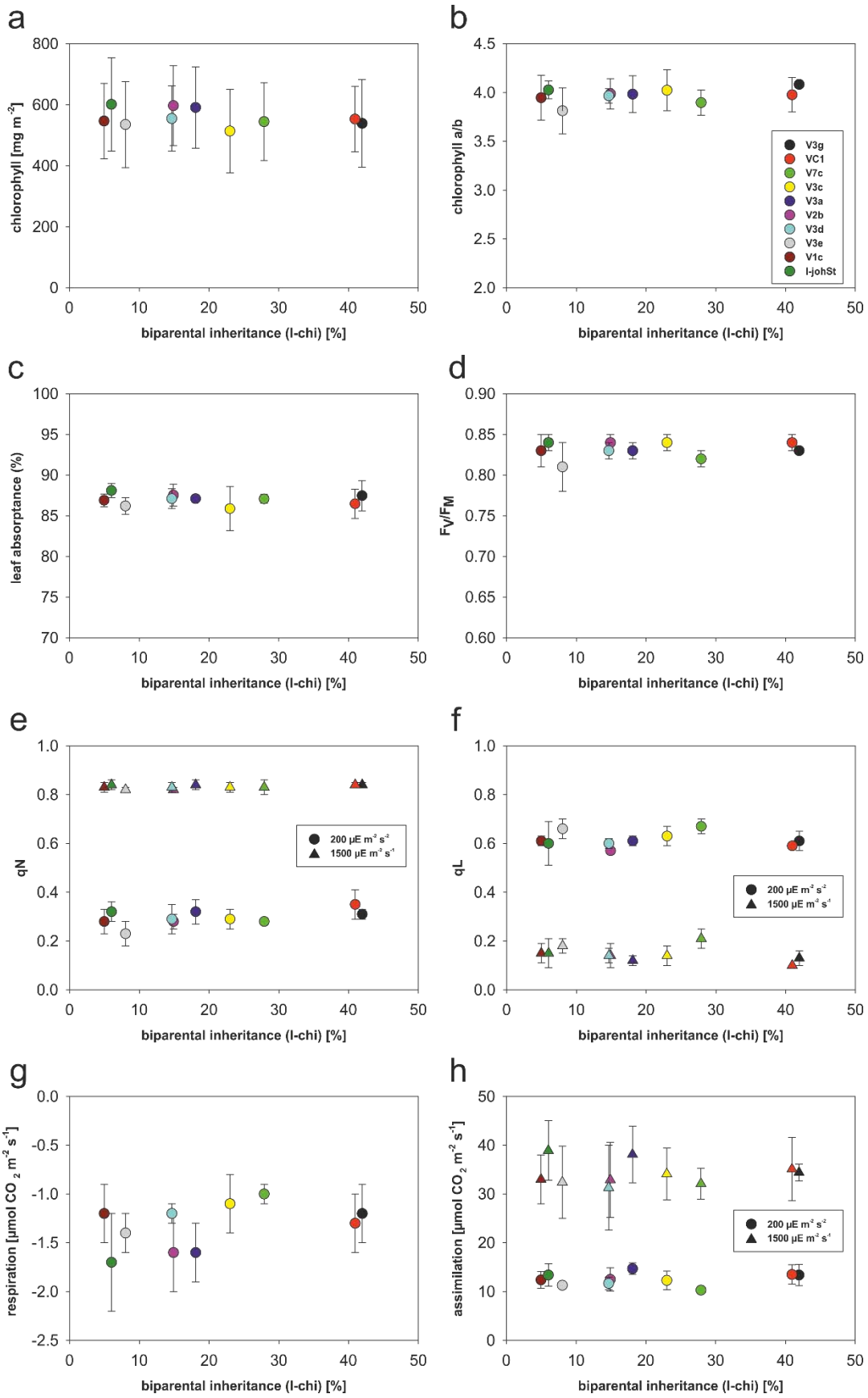


Fig. S4. Photosynthetic parameters of fully expanded leaves from 7-8 week-old I-johSt plants and several plastome I variants. (A) Chlorophyll content. (B) Chlorophyll a/b ratio. (C) Leaf absorptance, (D) Fv/Fm. **e**, qN under low light intensity at 200 µE m^{-2} s^{-1} and saturating light intensity at 1,500 µE m^{-2} s^{-1}. **f,** qL under low light intensity 200 µE m^{-2} s^{-1} and saturating light intensity at 1,500 µE m^{-2} s^{-1}. **g,** Respiration measured after 30 min of dark adaption. **e,** CO_2 assimilation under low light intensity at 200 µE m^{-2} s^{-1} and saturation light intensity at 1500 µE m^{-2} s^{-1}. Values are plotted against inheritance strength (biparental inheritance I-chi [%], for details see Fig. 2 and Supplementary Text).

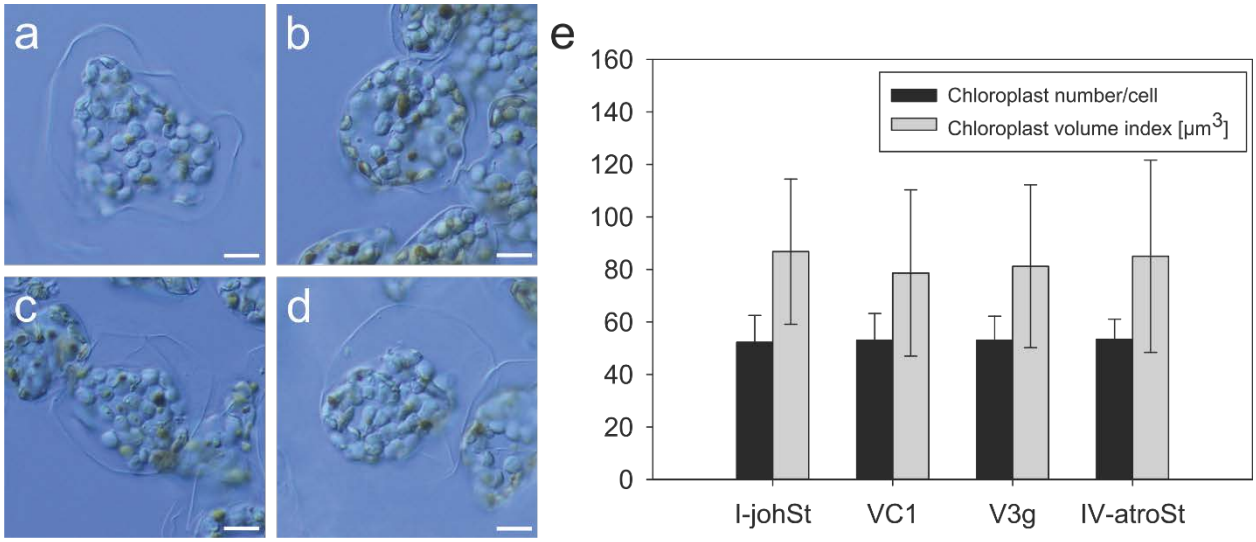


Fig. S5. Chloroplast size and number per cell of johansen Standard lines with wild type or variant chloroplasts of different inheritance strength. Representative cells of (A) I-johSt (strong wild type). (B) VC1 and (C) V3g (weak variants derived from I-johSt), and (D) IV-atroSt (weak wild type). e, Comparison between the lines. Note lack of statistically significant differences (Tables S13 and S14; Supplementary Text for details). Scale bar = 10 µm

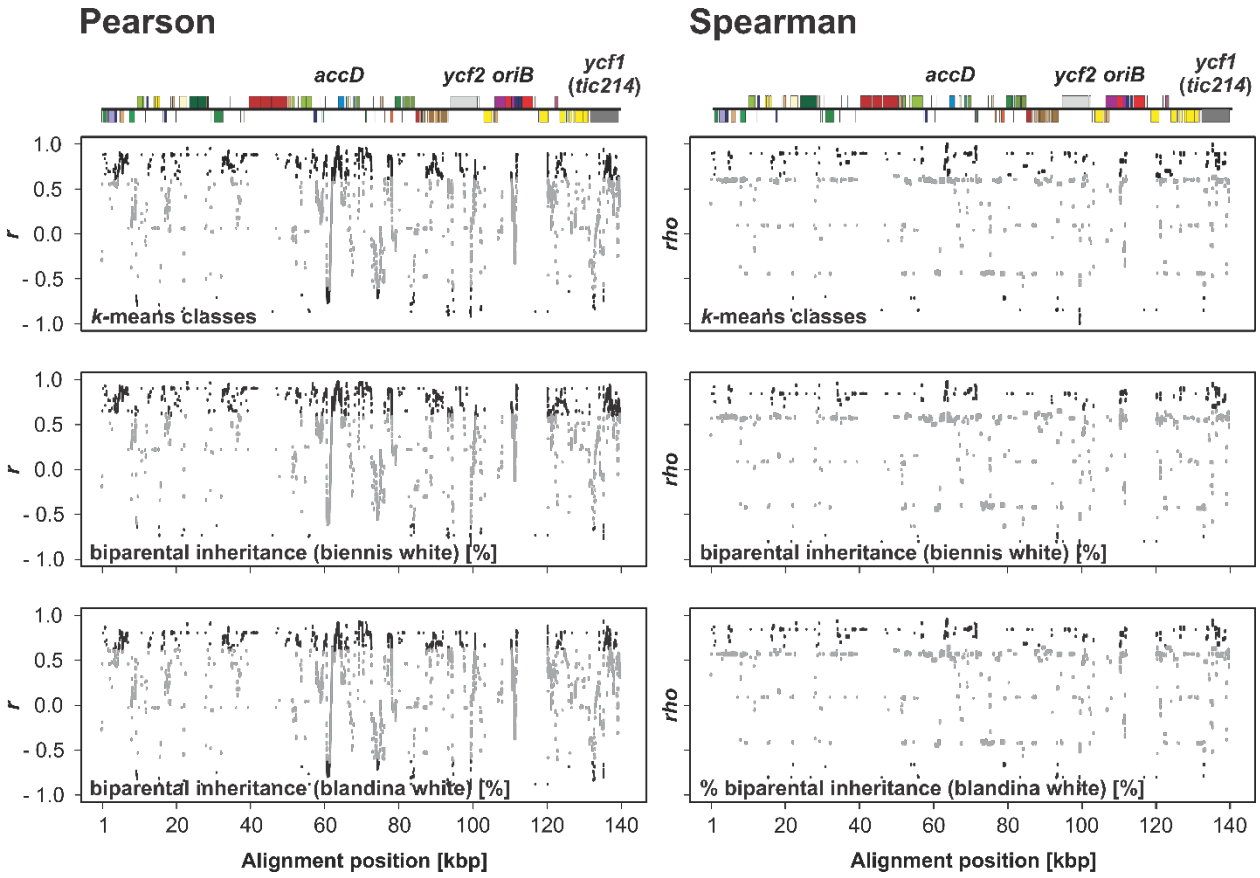


Fig. S6. Correlation mapping in wild type chloroplasts. Pearson's and Spearman's correlation to the *k*-means classes and the "biennis/blandina white" crossing data. Relevant genes or loci with significant correlation are noted on the linear plastome maps above. Non-significant correlations after *p*-value adjustment ($p > 0.05$) are displayed in gray. For details see Supplementary Text.

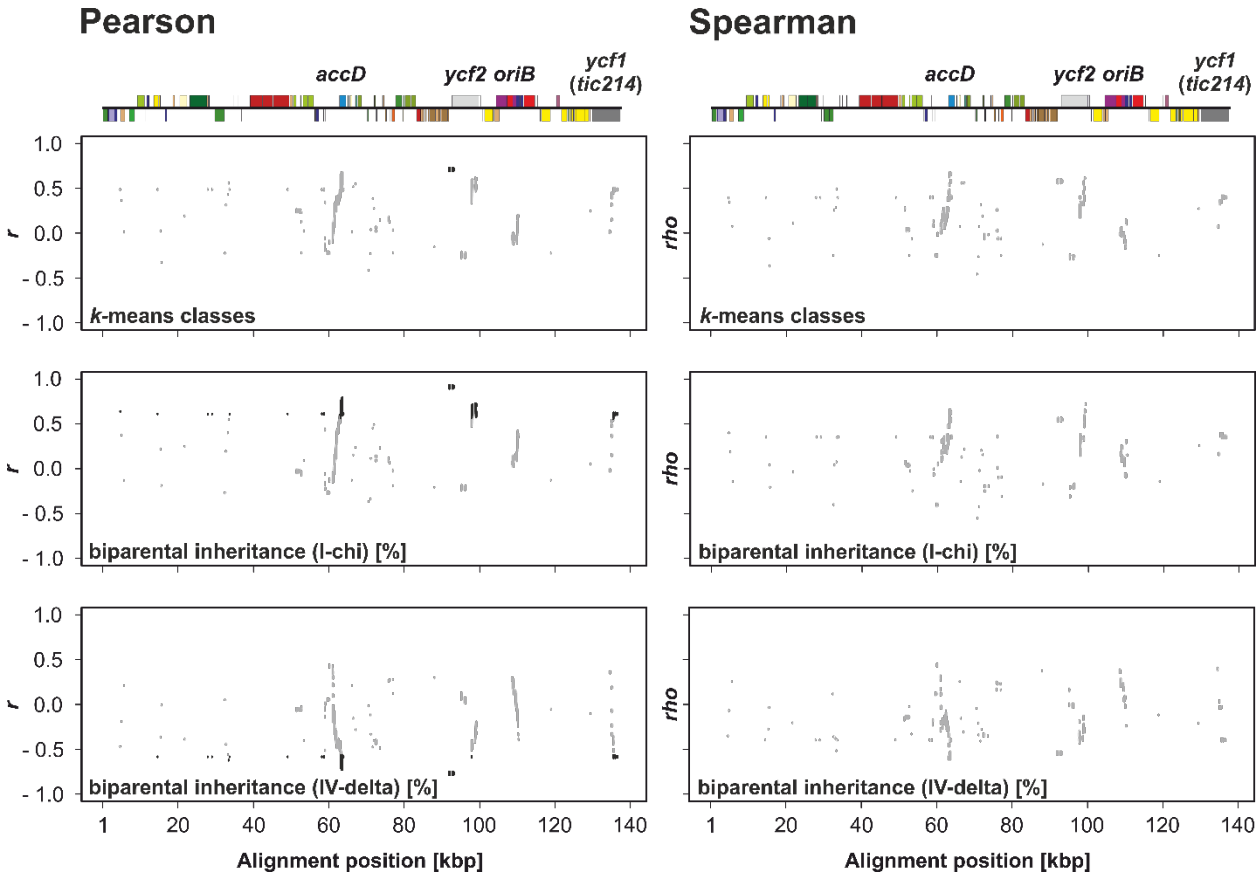
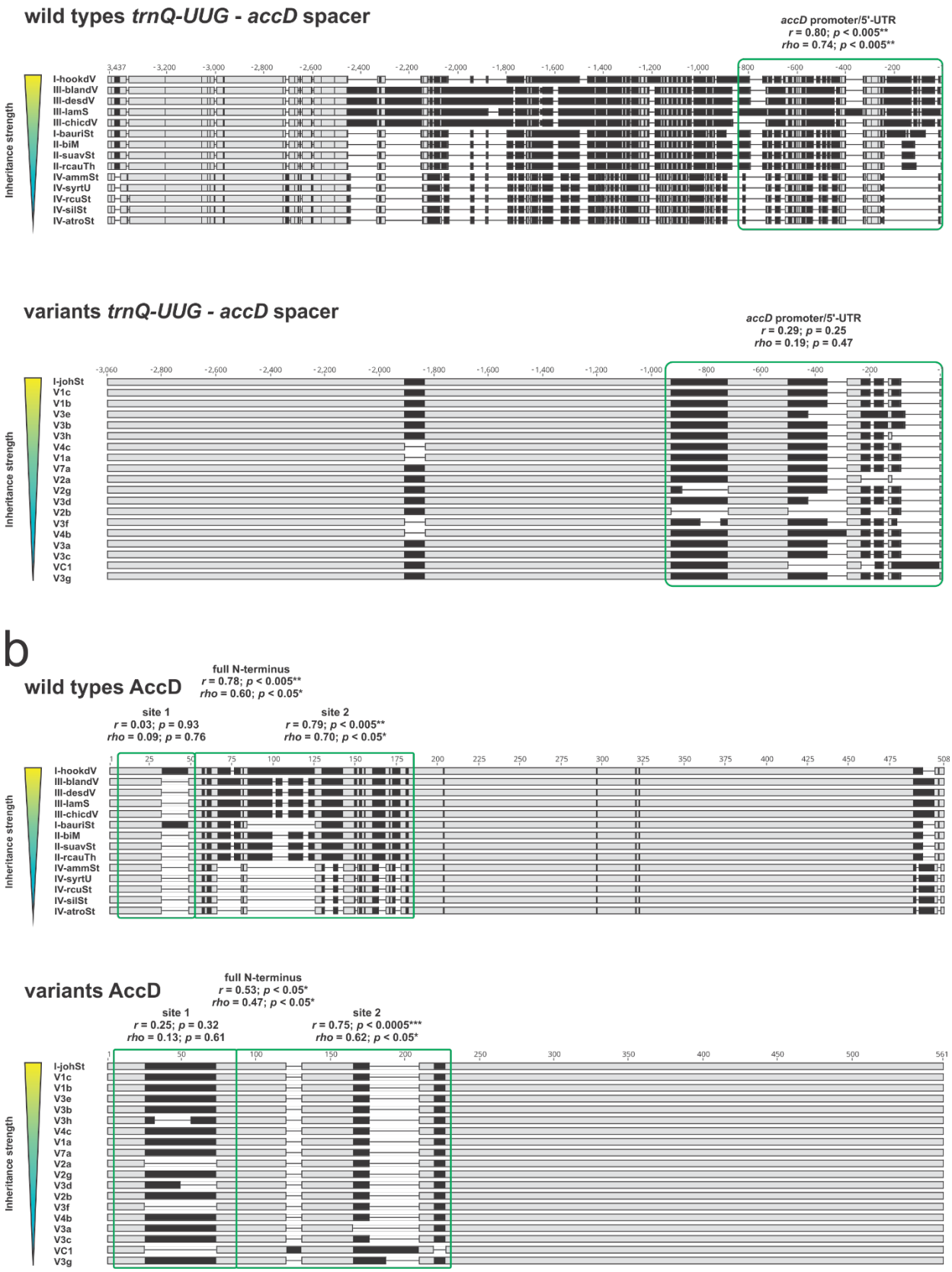


Fig. S7. Correlation mapping in the plastome I variants. Pearson's and Spearman's correlation to the *k*-means classes and the I-chi/IV-delta crossing data. Relevant genes or loci with significant correlation are designated in the linear plastome maps above. Non-significant correlations after *p*-value adjustment ($p > 0.05$) are displayed in grey. Note that regions positively correlated in the I-chi cross are negatively correlated in the reciprocal IV-delta cross. For details see Supplementary Text.

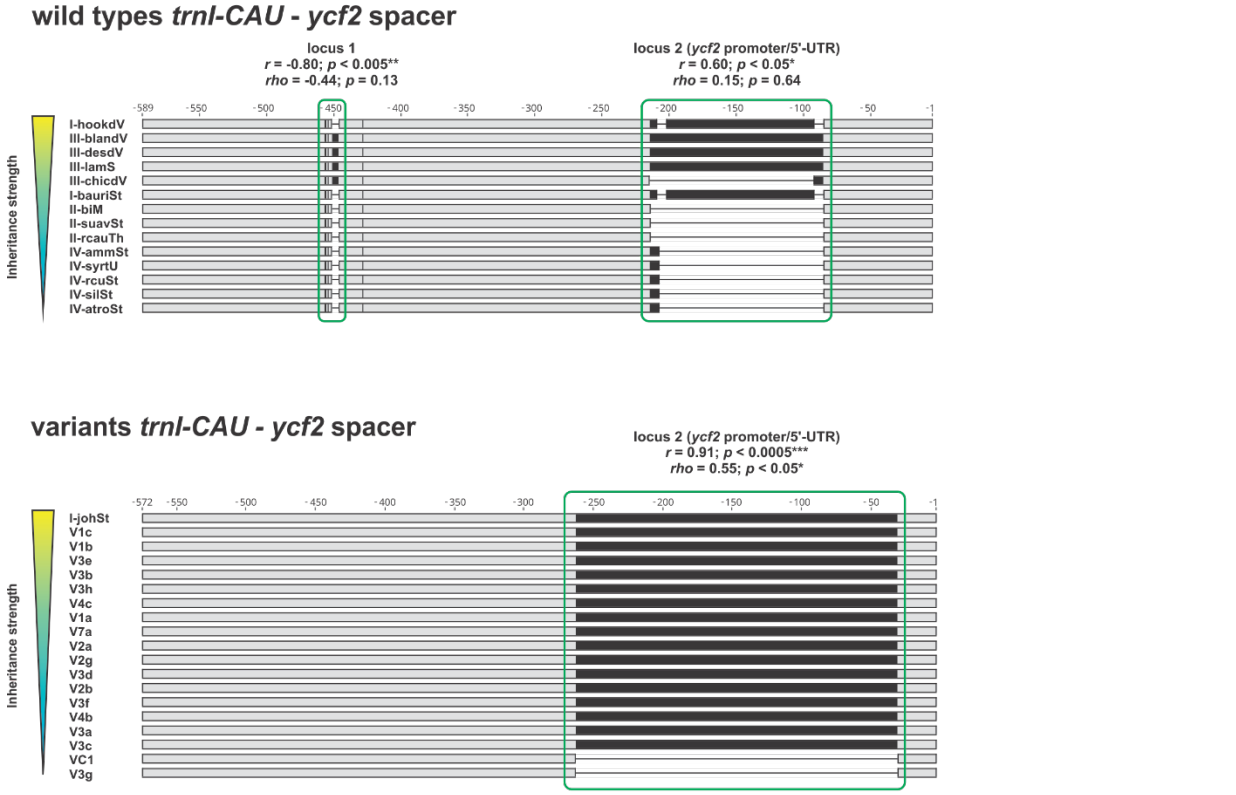
a



b

Fig. S8. Pearson's/Spearman's correlation to *k*-means classes (wild type) and I-chi (variants) at selected sites of the *accD* gene. Individual sequences are sorted according to their inheritance strength. Disagreement to the consensus in a given alignment window is indicated in black, alignments of identical sequences in grey. (A) *trnQ-UUG* - *accD* spacer (*accD* promoter/5'-UTR). (B) *accD* coding region. For better presentability, the protein alignment is shown. For details see Main Text and Supplementary Text.

a



b

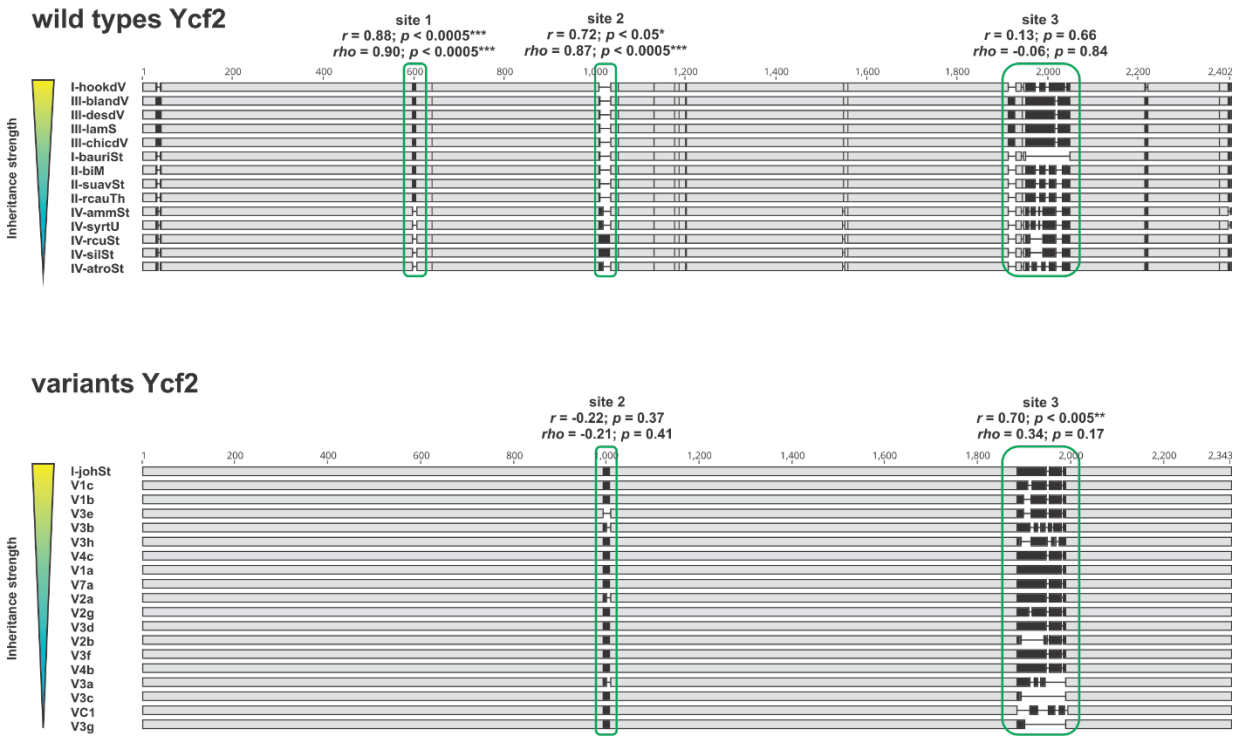


Fig. S9. Pearson's/Spearman's correlation to *k*-means classes (wild type) and I-chi (variants) at selected sites of the *ycf2* gene. Individual sequences are sorted according to their inheritance strength. Disagreement to the consensus in a given alignment window is indicated in black, alignments of identical sequences in grey. (A) *trnI-CAU* - *ycf2* spacer (*ycf2* promoter/5'-UTR). (B) *ycf2* coding region. Note that sites 1 and 2, but not site 3 correlate with inheritance strength in the wild types, whereas site 3 is associated with the inheritance phenotype in the variants. For better presentability, the protein alignment is shown. For details see Main Text and Supplementary Text.

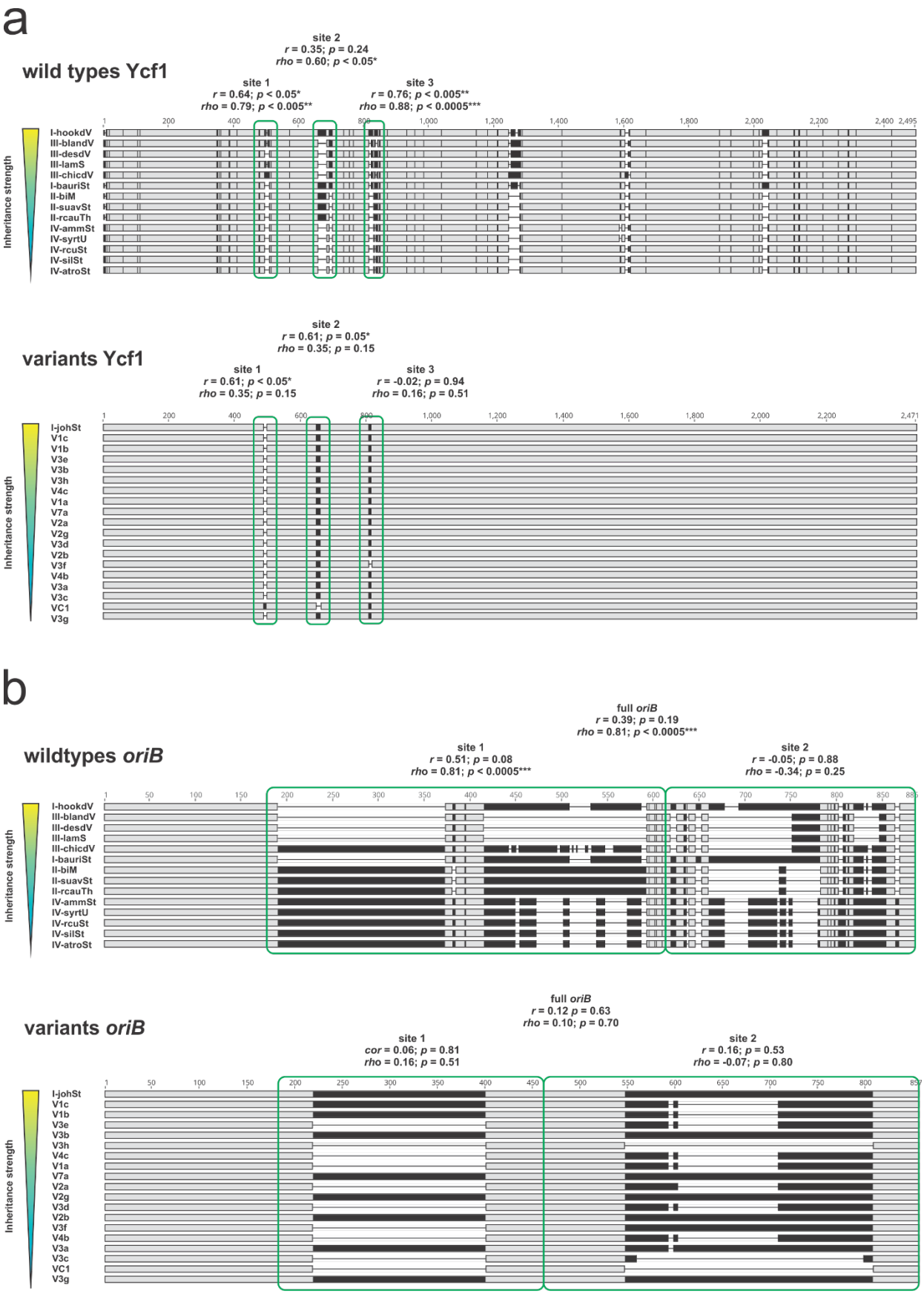


Fig. S10. Pearson's/Spearman's correlation to *k*-means classes (wild type) and I-chi (variants) at selected sites of the *ycf1* coding region and *oriB*. Individual sequences are sorted according to their inheritance strength. Disagreement to the consensus in a given alignment window is indicated in black, alignments of identical sequences in grey. (A) *ycf1* coding region. For better presentability, the protein alignment is shown. (B) *oriB*. For details see Main Text and Supplementary Text.

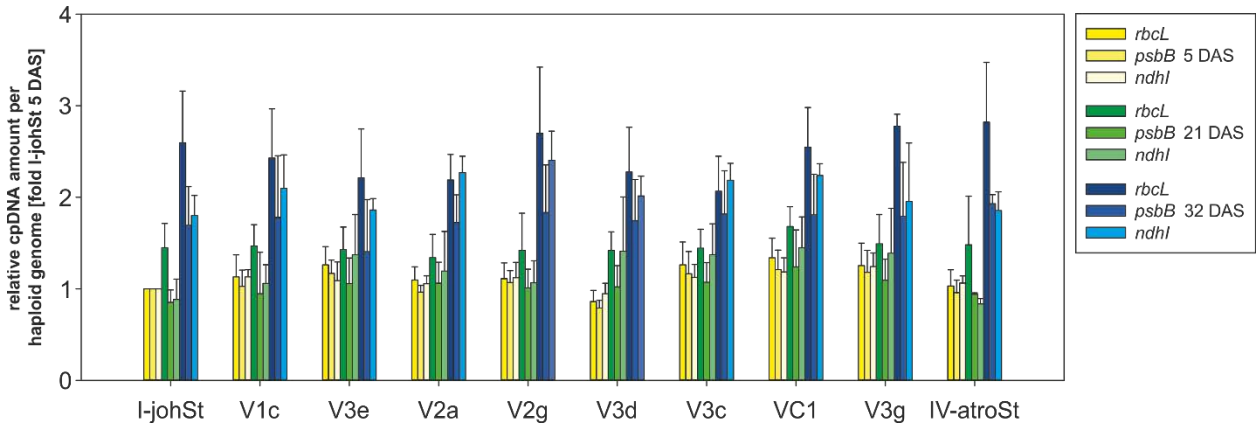


Fig. S11. Relative ptDNA content of the *rbcL*, *psbB* and *ndhI* loci at developmental stages 5 DAS (cotyledons), 21 DAS (cotyledons and first and second true leave) and 32 DAS (second true leaf) as judged from quantitative real-time PCR. Data for plastid markers were normalized to the mean amounts of three nuclear markers (M02, M19, and *pgiC*) and expressed as relative values [fold I-johSt 5 DAS]. For details see Materials and Methods. No statistical significant differences between the lines at a given developmental stage are observed (also see Supplementary Text).

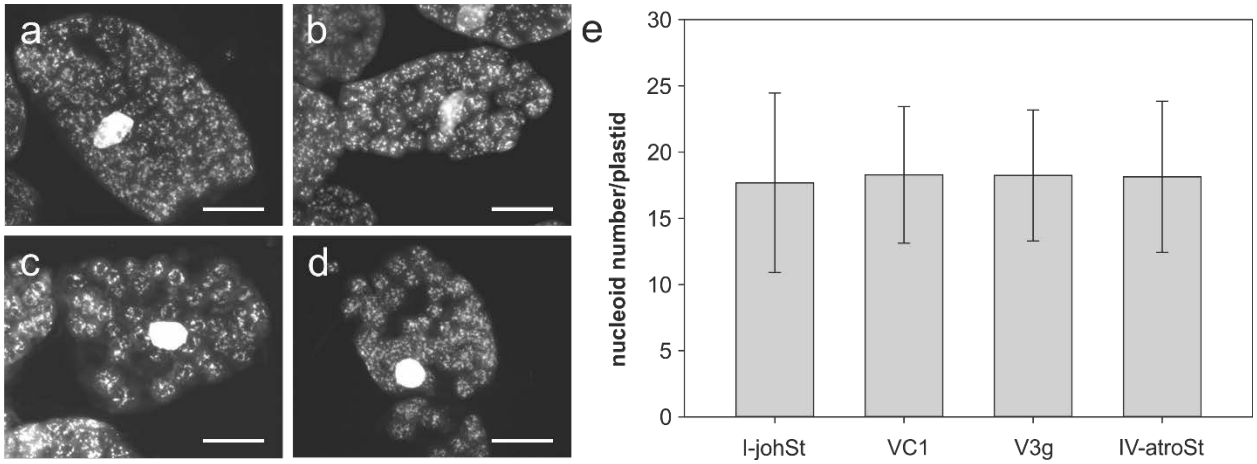


Fig. S12. ptDNA nucleoids visualized by DAPI and summary of nucleoid counts in chloroplast of different inheritance strengths. Representative cells of (A) I-johSt (strong wild type). (B) VC1 and (C) V3g (weak variants derived from I-johSt), and (D) IV-atroSt (weak wild type). **e**, Comparison of nucleoid number/plastid in the lines. Note lack of statistically significant differences (cf. Fig. S13, Table S12, Materials and Methods, and Supplementary Text). Scale bar = 10 µm

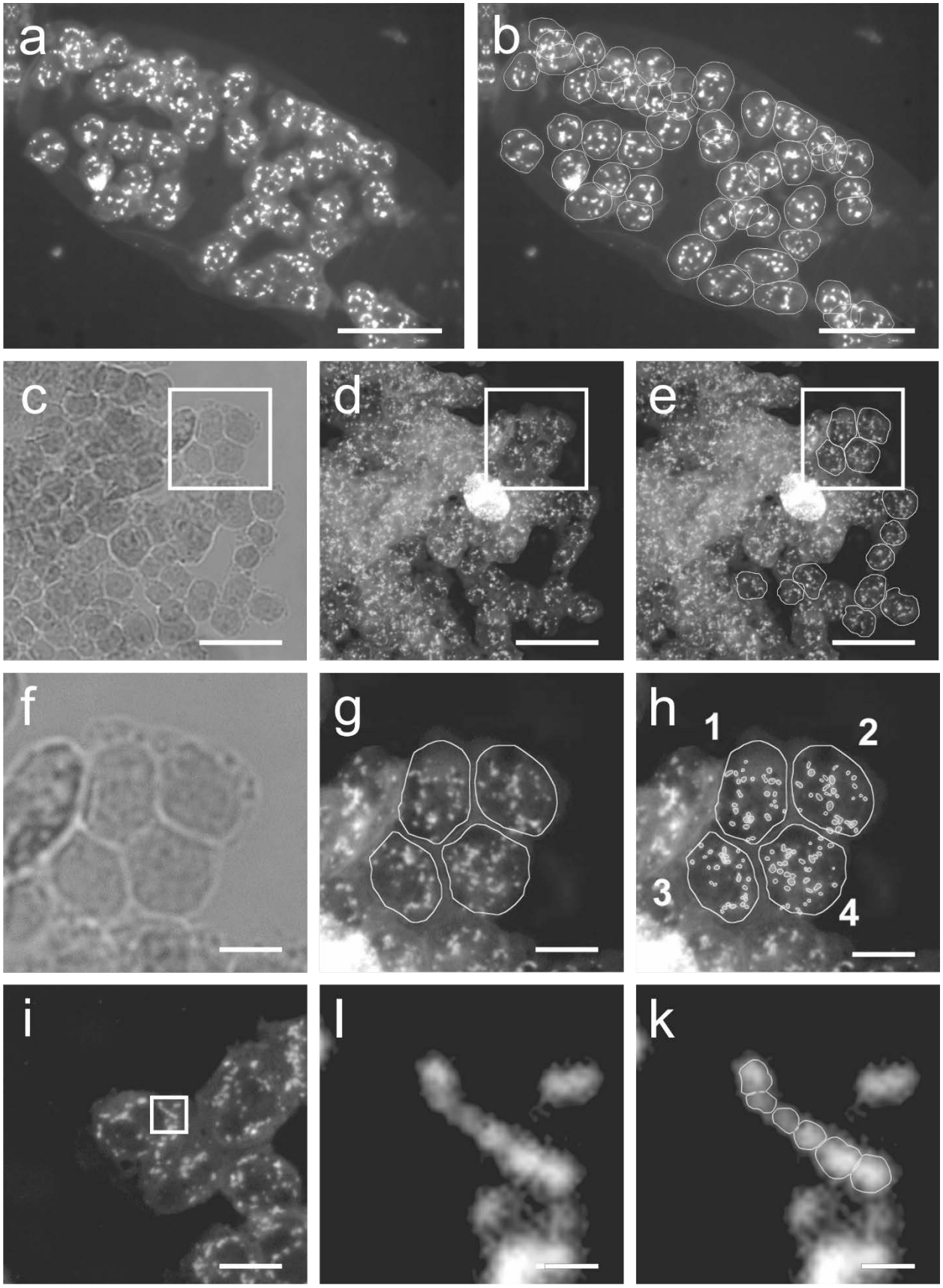


Fig. S13. Representative examples of nucleoid counting in DAPI-stained chloroplasts. (A) Original image of VC1 cells. (B) Manual delimitation of single chloroplasts in **a**. Only chloroplasts with non-overlapping nucleoids were analysed. **c-k**, Nucleoid counting in IV-atroSt chloroplasts. **c-e**, Regions from which cell fragments in **f-h** derive. **f-h** Chloroplast delimitation and counting. **f,** Bright field. **g,** Delimitation of chloroplast. **h,** Counting. Chloroplast 1: 26 nucleoids, chloroplast 2: 28 nucleoids, chloroplast 3: 27 nucleoids, and chloroplast 4: 39 nucleoids. **i-k,** Detangling of nucleoids aggregated in clumps and/or threads. **i**, Region from which the nucleoid thread in **j** and **k** derives from. **j,** Nucleoid thread. **k,** Delimitation and counting of single nucleoids in the thread. Scale bar = 10 µm in **a-e**, 2.5 µm in **f-h**, and 0.25 µm in **i-k**.

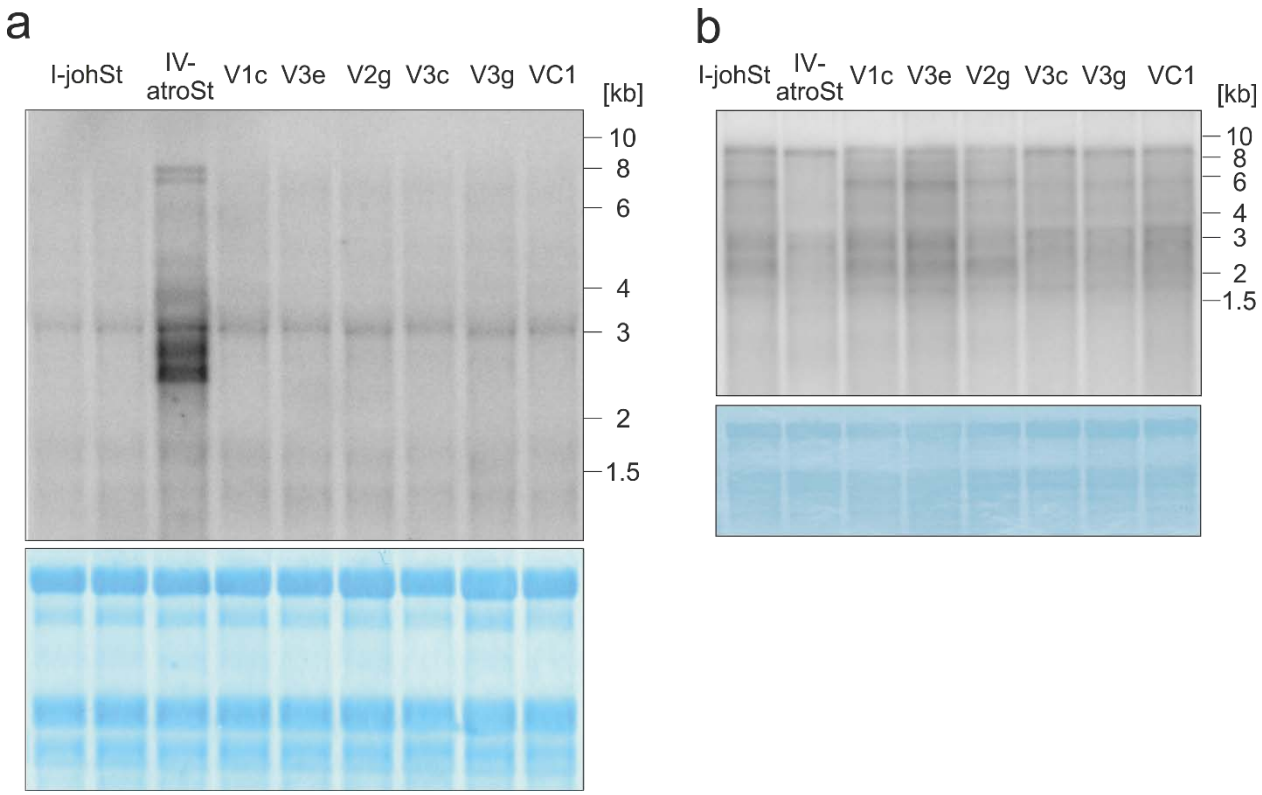


Fig. S14. Transcript accumulation of the *accD* and *ycf2* mRNA in lines of different inheritance strengths. (A) *accD.* Note the strong over-expression in the weak wild type IV-atroSt, whereas the plastome I variants do not differ from their wild type I-johSt. (B) *ycf2*; mature transcript (9 kb) and degradation/processing products. The latter differ between the strong and the weak wild types (I-johSt vs. IV-atroSt). The patterns are similar in the stronger (V1C, V3e and V2g) and weaker (V3c, V3g, and VC1) materials, correlating with the presence of mutations in site 3 of *ycf2* (Fig. S2).

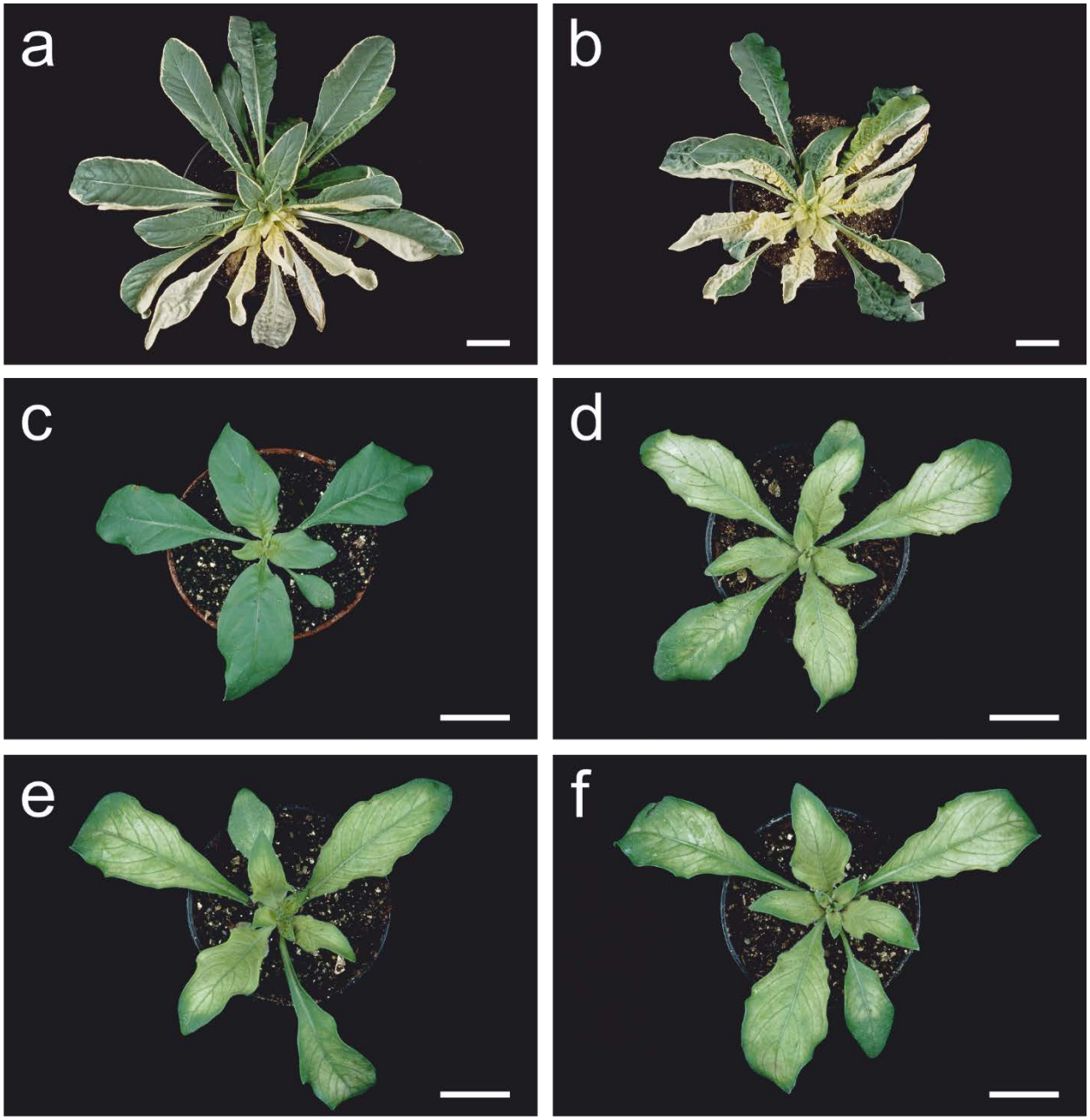


Fig. S15. Plants with chloroplast genotypes resulting in bleaching used in this work. a,b, Variegated lines combining wild type (green) and mutated (white) chloroplasts in one individual. **c-f,** Homoplastic material with only one chloroplast type. (A) Plastome mutant I-chi (white) and wild type IV-atroSt (green). (B) Plastome mutant IV-delta (white) and wild type I-johSt (green). (C) Chloroplast genome I-johSt in its native nuclear background **d-f,** *Virescent* phenotype of III-lamS and derived variants in the background of johansen Standard. (D) Wild type genotype III-lamS, **e,** III-V1, and **f,** III-V2; two independent very weak variants derived from the strong plastome III-lamS. **a,b,** Mature rosette. **c-f,** Early rosette stage. Scale bar = 5 cm in **a,b,** and 2 cm in **c-f**.

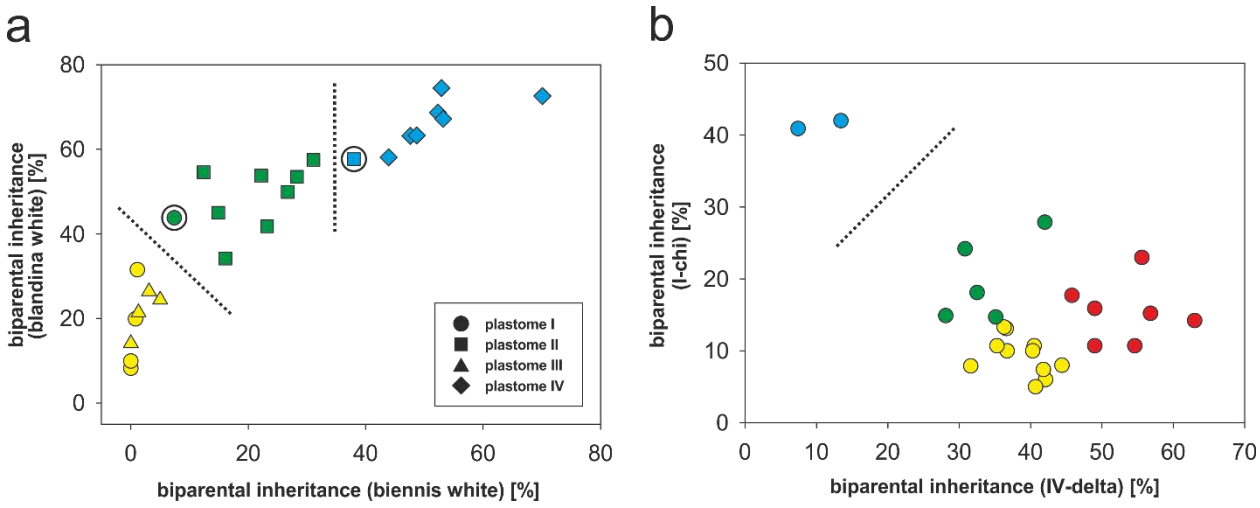


Fig. S16. Inheritance strength of the wild type and variant plastomes based on crossing data with white tester chloroplasts. The two crossing series, "biennis/blandina white" (wild types) and I-chi/IV-delta (variants), are plotted against each other. (A) Wild type plastomes of Table S11: *k*-means clustering with three centres identified by the pamk function (dotted lines) confirms the classes of Schötz (strong, intermediate and weak), indicated as yellow, green, and blue and comprised of plastome I and III, plastome II, and plastome IV, respectively. The two exceptions I-bauriSt and II-corSt are circled. (B) Plastome I variants of Table S8: *k*-means clustering with two centres (again obtained from the pamk function) clearly separates the dataset (dotted line), whereas finer clustering into four groups results in a new class (red) not present in the wild types (strong = yellow, strong to intermediate = red, intermediate = green, weak = blue). For details see Supplementary Text.

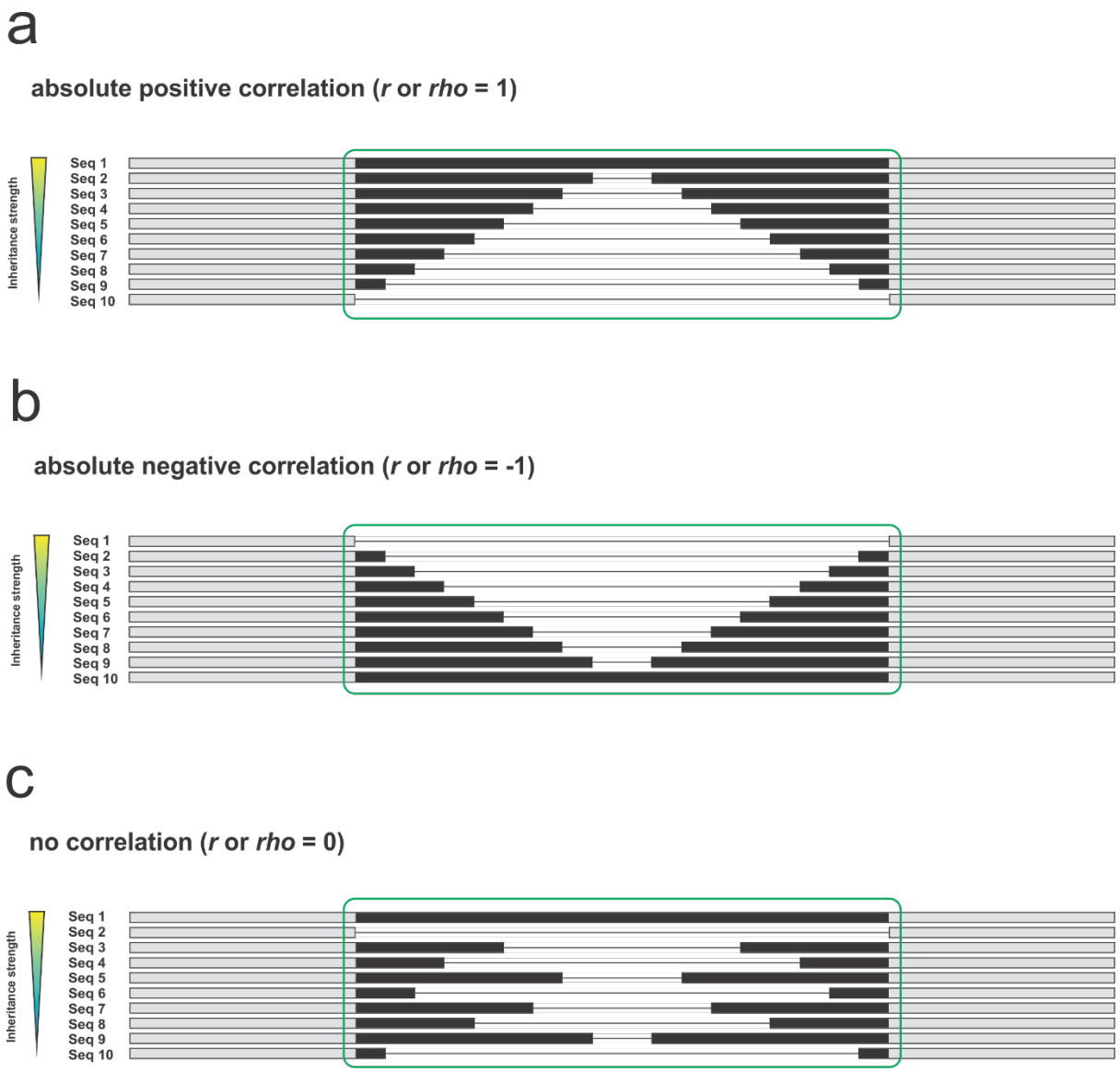


Fig. S17. Principal of the correlation mapping approach. Individual sequences are sorted according to their inheritance strength. Disagreement to the consensus in a given alignment windows is indicated in black, alignments of identical sequences in grey. Sequence alignment window with absolute positive (**A**) absolute negative (**B**) and no correlation (**C**) to inheritance strengths. For details see Materials and Methods and Supplementary Text.

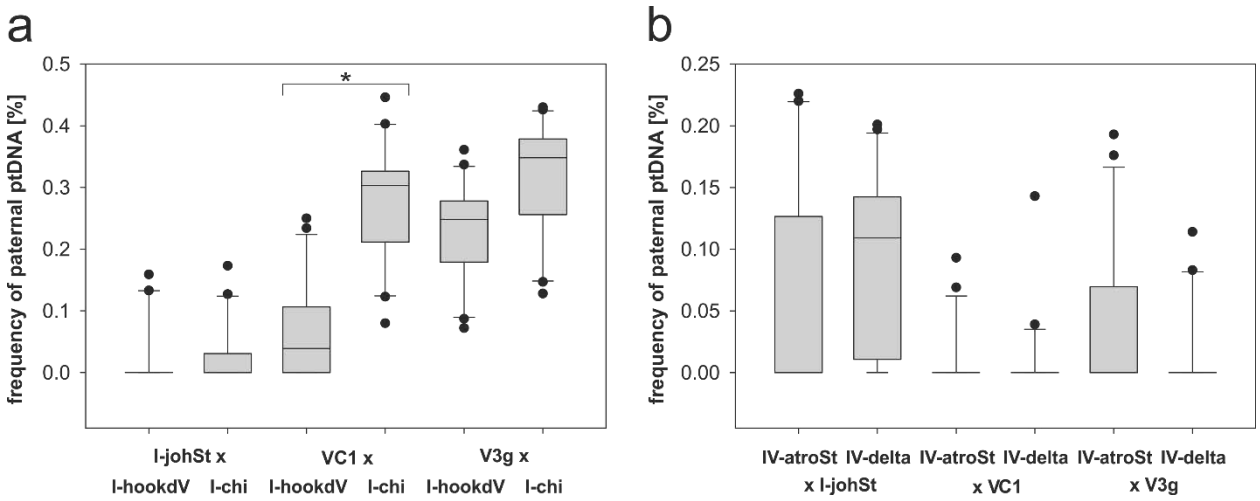


Fig. S18. Transmission efficiencies of wild type and derived bleached mutant chloroplasts in crosses to I-johSt, VC1 and V3g. (A) Wild type/mutant pair (I-hookdV/I-chi) as pollen parent. (B) Wild type/mutant pair (IV-atroSt/IV-delta) as seed parent. Box-plots represent frequencies of the paternal cpDNA [%] measured by the MassARRAY® system. Significance of difference (VC1 x I-hookdV vs. VC1 x I-chi) was tested with Kruskal-Wallis one-way ANOVA on ranks (* $p < 0.05$).

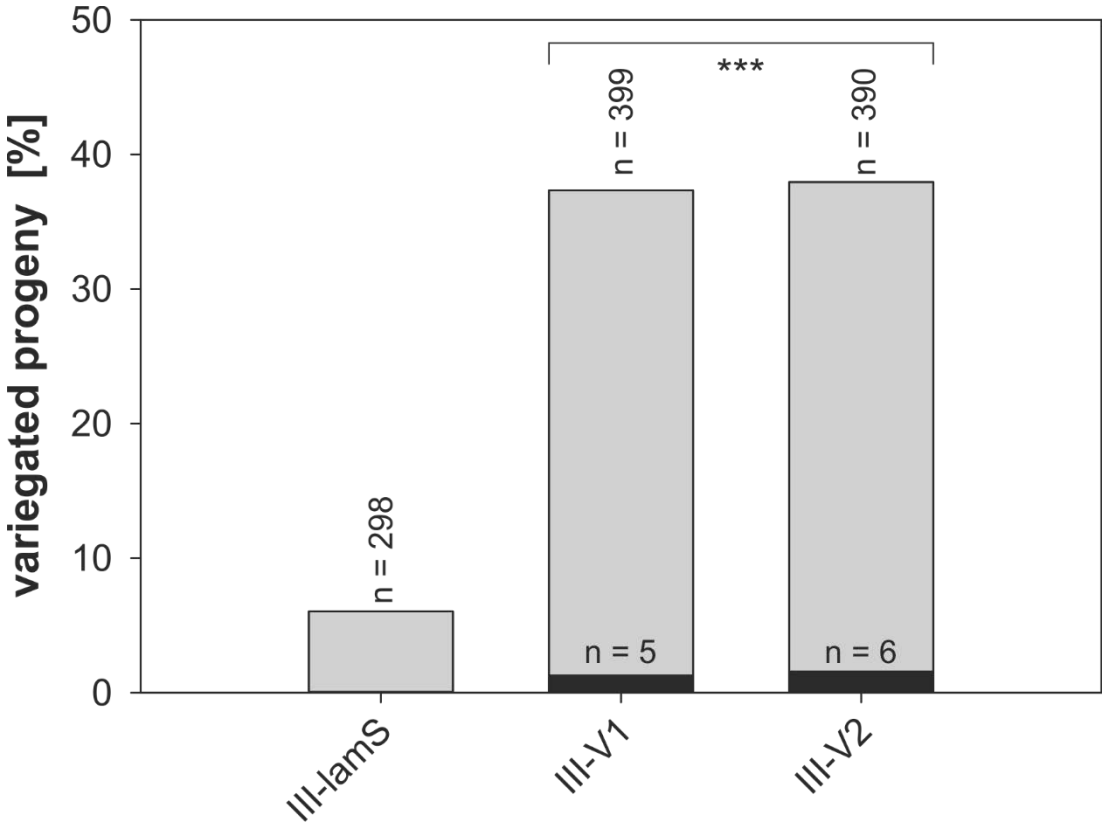


Fig. S19. Transmission efficiencies of wild type plastome III-lamS and its variants III-V1 and III-V2 in crosses to I-johSt as pollen donor. Depicted in grey is the average percentage of variegated progeny (= biparental inheritance [%]) obtained from two seasons (2015 and 2016). The black bars indicate seedlings with only the paternal chloroplast. Significance of differences between variegated progeny of III-lamS and the variants was calculated by Fisher`s exact test (*** $p < 0.0001$).

Table S1. Genotypes used for the individual experimental series of the lipid-level measurements described in this work[1)]

Variant/mutant/ plastome[2)]	Experiment 1	Experiment 2	Experiment 3	# total samples measured	Inheritance class	Phenotype
I-johSt	x	x	x	15	1	green
I-hookdV	-	x	x	10	1	green
III-lamS	-	x	x	10	1	*virescent*
I-chi	-	x	x	10	1	white to yellowish
V1c	x	-	x	10	1	green
V2a	-	x	x	10	1	green
V2g	x	-	x	10	1	green
V3e	x	-	x	10	1	green
V3c	x	-	x	10	2	green
V3d	-	x	x	10	3	green
VC1	x	x	x	15	4	green
V3g	x	-	x	10	4	green
IV-atroSt	x	x	x	15	4	green
IV-delta	-	x	x	10	4	white to yellowish
III-V1	-	x	-	5	5	*virescent*
III-V2	-	x	-	5	5	*virescent*

x = included in the experiment **-** = not included in the experiment

[1)] For details on the chloroplast genomes, inheritance class and phenotype of the nuclear background of johansen Standard see Fig. 2, Figs. S15,S18 and S19, Tables S4-S9, and Supplementary Text.

Table S2. Average weight and corresponding *p*-value of the 102 lipids/molecules analyzed in the LASSO regression model to predict inheritance strengths. The 20 predictive lipid (absolute average weight greater than one standard deviation of all lipids/molecules) are marked in italics

Lipid/ molecule	Average weight	Adjusted *p*-value[1)]	Lipid/ molecule	Average weight	Adjusted *p*-value[1)]	Lipid/ molecule	Average weight	Adjusted *p*-value[1)]
TAG 54.3	*-2.89*	*1.27×10^{-30}*	SQDG 34.3	-0.04	1.27×10^{-1}	FA 28.0	0.06	1.50×10^{-4}
PC 36.3	*-1.93*	*6.96×10^{-46}*	TAG 50.2	-0.03	1	CoQ 9	0.07	9.11×10^{-6}
PC 36.4	*-1.66*	*3.05×10^{-27}*	SQDG 36.2	-0.03	6.92×10^{-2}	PC 34.5	0.07	2.90×10^{-1}
DGDG 36.4	*-1.32*	*8.84×10^{-31}*	TAG 54.7	-0.02	1	MGDG 34.2	0.08	9.61×10^{-4}
DGDG 36.5	*-1.30*	*1.01×10^{-33}*	MGDG 36.6	-0.02	1	SQDG 36.6	0.11	6.52×10^{-3}
DGDG 34.5	*-1.29*	*1.16×10^{-28}*	TAG 60.3	-0.02	1	TAG 58.3	0.12	1.24×10^{-4}
TAG 60.6	*-1.29*	*2.82×10^{-38}*	TAG 50.1	-0.01	1	PC 36.6	0.13	3.54×10^{-5}
TAG 52.3	*-1.19*	*5.40×10^{-31}*	TAG 56.6	-0.01	1	TAG 56.7	0.16	6.84×10^{-2}
PC 36.2	*-0.97*	*1.18×10^{-18}*	TAG 52.1	0.00	1	PE 36.2	0.16	1
SQDG 36.4	-0.63	2.77×10^{-29}	TAG 54.5	0.00	1	PC 34.2	0.16	6.19×10^{-8}
FA 30.0	-0.52	5.76×10^{-33}	DGDG 36.3	0.00	1	MGDG 36.2	0.17	5.08×10^{-6}
TAG 52.9	-0.49	2.07×10^{-22}	TAG 48.0	0.00	1	DGDG 32.0	0.19	4.82×10^{-8}
TAG 54.8	-0.43	2.80×10^{-13}	chlorophyll b	0.00	1	PE 36.3	0.20	1.05×10^{-7}
PC 32.1	-0.40	8.73×10^{-13}	DGDG 34.2	0.00	1	SQDG 36.3	0.22	1.60×10^{-7}
DGDG 34.1	-0.31	3.82×10^{-15}	TAG 50.5	0.00	1	PE 36.4	0.24	2.88×10^{-7}
PE 36.6	-0.26	1.64×10^{-8}	TAG 52.7	0.00	1	CoQ10	0.26	1.85×10^{-7}
MGDG 34.3	-0.25	7.70×10^{-6}	TAG 54.4	0.00	1	TAG 52.6	0.32	1.15×10^{-11}
PC 32.2	-0.25	1.53×10^{-5}	TAG 56.3	0.00	1	TAG 50.3	0.40	7.34×10^{-14}
SQDG 36.5	-0.20	7.40×10^{-9}	TAG 56.4	0.00	1	pheophytin	0.49	8.42×10^{-39}
TAG 52.5	-0.19	1.12×10^{-10}	TAG 60.2	0.00	1	PC 34.3	0.60	4.03×10^{-19}
FA 26.0	-0.16	1.15×10^{-11}	TAG 60.4	0.00	1	TAG 56.2	0.61	2.89×10^{-15}
DGDG 36.2	-0.14	5.00×10^{-2}	TAG 54.2	0.00	1	MGDG 36.3	0.63	2.21×10^{-46}
PG 34.2	-0.13	6.70×10^{-3}	TAG 50.6	0.00	1	PC 32.0	0.65	3.71×10^{-21}
PI 34.3	-0.11	1.53×10^{-8}	TAG 54.6	0.00	1	*TAG 50.7*	*0.80*	*2.43×10^{-20}*
PE 34.2	-0.11	1.10×10^{-1}	TAG 54.1	0.01	1	*TAG 56.1*	*0.80*	*1.14×10^{-34}*
TAG 52.2	-0.10	5.54×10^{-5}	MGDG 36.5	0.01	1.72×10^{-1}	*DGDG 34.4*	*0.94*	*9.52×10^{-29}*
TAG 60.5	-0.09	1.68×10^{-2}	TAG 50.4	0.01	1	*PC 34.1*	*0.96*	*3.43×10^{-29}*
TAG 56.5	-0.09	6.92×10^{-2}	TAG 48.1	0.02	1	*PE 36.5*	*1.01*	*1.26×10^{-23}*
PI 34.2	-0.09	3.75×10^{-8}	TAG 52.4	0.02	7.39×10^{-1}	*MGDG 36.4*	*1.02*	*2.77×10^{-34}*
DGDG 34.3	-0.08	2.90×10^{-2}	chlorophyll a	0.03	1	*TAG 58.4*	*1.03*	*7.23×10^{-32}*
PC 34.4	-0.08	1.16×10^{-2}	TAG 52.8	0.04	1	*PE 34.3*	*1.03*	*3.58×10^{-23}*
TAG 58.5	-0.08	6.38×10^{-7}	TAG 58.1	0.05	7.72×10^{-6}	*PG 34.3*	*1.12*	*2.13×10^{-32}*
FA 24.0	-0.06	9.17×10^{-8}	TAG 60.1	0.05	2.16×10^{-4}	*TAG 54.9*	*2.39*	*1.30×10^{-55}*
DGDG 36.6	-0.06	5.02×10^{-1}	PC 36.5	0.06	3.40×10^{-13}	*TAG 58.2*	*2.68*	*1.75×10^{-37}*

[1)] *p*-values represent significance of deviation of average coefficients (weights) from zero as obtained from the 100 cross-validation runs. Raw *p*-values have been adjusted for multiple testing (102 lipids) using Benjamini and Hochberg procedure.

Table S3. Enrichment/depletion of lipid/molecule classes in the set of 20 predictive lipids based on Fisher's exact test *p*-value and odds ratio

Lipid class	Odds ratio	*p*-value[1)]	# predictive lipids/molecules	# total lipids/ molecules in classes
PG	4.18	0.35	1	2
DGDG	2.64	0.22	4	11
PC	2.01	0.28	4	13
PE	1.70	0.62	2	7
TAG	0.81	0.80	8	45
MGDG	0.67	1.00	1	7
FA	0.00	0.58	0	4
SQDG	0.00	059	0	6
PI	0.00	1.00	0	2
chlorophyll	0.00	1.00	0	2
CoQ	0.00	1.00	0	2
pheophytin	0.00	1.00	0	1

[1)] *p*-values have not been corrected for multiple testing as none of the raw *p*-values proved below $p < 0.05$. After correction, *p*-values shall be set to 1.

Table S4. *Oenothera* strains with native wild type chloroplasts used in this work[1)]

Species	Strain	Locality	Collection/ isolation date	Collector/ isolated by	Reference
O. elata ssp. *hookeri*	johansen Standard	USA, CA, Sutter Co., roadside between Nicholas and Yuba City	1927	C. B. Wolf	[34]
O. elata ssp. *hookeri*	hookeri de Vries	USA, CA, Alameda Co., near Berkeley	1904	H. de Vries	[85]
O. villosa ssp. *villosa*	bauri Standard	Poland, Kujawsko-Pomorskie, near Toruń	before 1942	R. Hölscher	[86]
O. glazioviana	blandina de Vries[2)]	chromosome translocation mutant of the original lamarckiana de Vries[3)]	isolated in 1908	H. de Vries	[12,87]
O. glazioviana	deserens de Vries[4)]	chromosome translocation mutant of the original lamarckiana de Vries[3)]	isolated in 1913	H. de Vries	[12,88,89]
O. glazioviana	*r*/*r*-lamarckiana Sweden	Sweden, Skåne Län, garden in Almaröd	1906	N. Heribert-Nilsson	[90]
O. biennis	chicaginensis de Vries	USA, IL, Cook Co., Chicago, near Jackson Park	1904	H. de Vries	[85]
O. biennis	biennis Muenchen	Germany, Bavaria, Munich, Nymphenburg Garden	1914	O. Renner	[91]
O. biennis	rubricaulis Thorn	Poland, Kujawsko-Pomorskie, Vistula River near Toruń	before 1941	R. Hölscher	[92]
O. biennis	suaveolens Standard	France, Seine-et-Marne, forest of Fontainebleau	1912	L. Blaringhem	[93]
O. biennis	suaveolens Grado	Italy, Friuli-Venezia Giulia, dune near Grado at the Adriatic sea	before 1950	H. Zeidler	[94]
O. oakesiana	ammophila Standard	Germany, Schleswig-Holstein, Helgoland	1922	E. Hoeppener	[95]
O. oakesiana	*r*/*r*-syrticola Ulm	Germany, Baden-Württemberg, Danube River near Ulm	1917	O. Renner	[96]
O. parviflora	rubricuspis Standard	Germany, Hessen, railway embankment between Neu-Isenburg and Luisa near Frankfurt on the Main	1942/1943	O. Burck, F. Laibach, and E. Fischer	[97]
O. parviflora	silesiaca Standard	Poland, Dolnośląskie, bank of Bóbr River near Nowogrodziec	1937	O. Renner	[92]
O. parviflora	atrovirens Standard	USA, NY, Erie Co., Sandy Hill near Lake George[5)]	1902/1903	D. T. MacDouglas	[85,98]

[1)] For details on *Oenothera* taxonomy and strain designation see [12,30,32].

[2)] *Syn*: *O. blandina* de Vries, *O. lamarckiana* de Vries mut. *blandina*, *O. lamarckiana* de Vries mut. *veluntina.*

[3)] The original lamarckiana de Vries was first collected in 1886 at Graveland near Hilversum (The Netherlands, Noord-Holland) by H. de Vries [85,99].

[4)] *Syn*: *O. deserens* de Vries, *O. lamarckiana* de Vries mut. *deserens.*

[5)] According to Renner [98] this line was originally "received from Amsterdam" by N. v. Gescher in 1907. The material is quite likely identical to that collected by D. T. MacDouglas in 1902/1903 as described in [85]. Also see [100].

Table S5. Summary on the wild type *Oenothera* chloroplast genomes used in this work[1)]

Species	Strain[1)]	Abbreviation	Basic plastome[2)]	[%] biparental inheritance (biennis white)[3)]	[%] biparental inheritance (blandina white)[3)]	Inheritance class (Schötz)[4)]	Inheritance class (this work)[5)]	GenBank/EMBL accession number
O. elata ssp. *hookeri*	johansen Standard	I-johSt	I	n/a	n/a	strong	1	AJ271079.4
O. elata ssp. *hookeri*	hookeri de Vries	I-hookdV	I	0.0	8.2	strong	1	KT881170.1
O. villosa ssp. *villosa*	bauri Standard	I-bauriSt	I	7.4	43.8	strong	3	KX687910.1
O. glazioviana	blandina de Vries	III-blandV	III	0.0	14.1	strong	1	KT881171.1
O. glazioviana	deserens de Vries	III-desdV	III	1.3	21.4	strong	1	KT881172.1
O. glazioviana	*r*/*r*-lamarckiana Sweden	III-lamS	III	3.1	26.4	strong	1	EU262890.2
O. biennis	chicaginensis de Vries	III-chicdV	III	5.0	24.4	strong	1	KX687913.1
O. biennis	biennis Muenchen	II-biM	II	22.2	53.8	intermediate	3	KU521375.1
O. biennis	rubricaulis Thorn	II-rcauTh	II	23.2	41.8	intermediate	3	KX687914.1
O. biennis	suaveolens Standard	II-suavSt[6)]	II	26.7	49.9	intermediate	3	KX687915.1
O. biennis	suaveolens Grado	II-suavG[6)]	II	n/a	n/a	intermediate	3	EU262889.2
O. oakesiana	ammophila Standard	IV-ammSt	IV	43.9	58.1	weak	4	KT881176.1
O. oakesiana	*r*/*r*-syrticola Ulm	IV-syrtU	IV	48.7	63.3	weak	4	KX687918.1
O. parviflora	rubricuspis Standard	IV-rcuSt	IV	52.3	68.7	weak	4	KX687916.1
O. parviflora	silesiaca Standard	IV-silSt	IV	53.2	67.2	weak	4	KX687917.1
O. parviflora	atrovirens Standard	IV-atroSt	IV	70.1	72.6	weak	4	EU262891.2

[2)]In *Oenothera,* five genetically distinguishable plastome types (I-V) can be recognized based on their compatibility with three nuclear genomes (A, (B) C) in either homozygous (AA, BB, CC) or stable heterozygous (AB, AC, BC) states. The basic plastome genotype is accompanied by a given inheritance strength (strong, intermediate and weak). Basic plastome and nuclear genome type are an important factor of species definition in *Oenothera*. For details see, e.g. [6,7,11,12,14,30,35].

[3)] Percentage of variegated seedlings, heteroplasmic due to the paternal transmission of the bleached chloroplast mutants “biennis white” or “blandina white” and the maternal transmission of the green wild chloroplast of the seed parent. Data accoding to Schötz [6]. For details see therein, reviews in [7,12] and Supplementary Text.

[4)] Class of inhertiance strength as determined by F. Schötz. For reviews see [7,10,12].

[5)] For details on the definition of these classes see Supplementary Text.

[6)] The chloroplast genomes of the two suaveolens stains are identical.

Table S6. Wild type chloroplast genomes introgressed into the nuclear background of johansen Standard used in this work[1)]

Plastome[2)]	Basic plastome[3)]	Inheritance class (Schötz)[4)]	Inheritance class (this work)[5)]	Chloroplast donor strain[6)]	Phenotype in the nuclear background of johansen Standard	Produced by	Reference
I-johSt	I	strong	1	johansen Standard	green	wild type	[14,34]
I-hookdV	I	strong	1	hookeri de Vries	green	S. Greiner	[23]; this work
II-suavG	II	intermediate	3	suaveolens Grado	green	W. Stubbe	[14,35]
III-lamS	III	strong	1	*r*/*r*-lamarckian Sweden	*viresencet*	W. Stubbe	[43]
IV-atroSt	IV	weak	4	atrovirens Standard	green	W. Stubbe	[14,35]

[1)] Wild type plastids with different transmission efficiencies were introgressed into the constant nuclear background of the johansen Standard strain (*O. elata* ssp. *hookeri*). See Table S4 for details on this line.

[2)] See Table S5 for details.

[3)] In *Oenothera,* five genetically distinguishable plastome types (I-V) can be recognized based on their compatibility with three nuclear genomes (A, (B) C) in either homozygous (AA, BB, CC) or stable heterozygous (AB, AC, BC) states. The basic plastome genotype is accompanied by a given inheritance strength (strong, intermediate and weak). Basic plastome and nuclear genome type are an important factor of species definition in *Oenothera*. For details, see e.g. [6,7,11,12,14,30,35].

[4)] Classes of inhertiance strength as determined by F. Schötz. For reviews see [7,10,12].

[5)] For details on the definition of these classes, see Supplementary Text.

[6)] For details on the donor strains, see Table S4.

Table S7. Spontaneous mutants in the chloroplast gene *psaA* in the nuclear background of johansen Standard used in this work[1)]

Mutant/ plastome	Wild type background plastome[2)]	Basic plastome[3)]	Inheritance class (Schötz)[4)]	Inheritance class (this work)[5)]	Phenotype	Mutation	Reference
I-chi	I-hookdV	I	strong	1	white to yellowish	5-bp duplication (CCGCT), +1420 to +1424 *psaA*, truncated PsaA	[23]
IV-delta	IV-atroSt	IV	weak	4	white to yellowish	5-bp duplication (TTAAC), +667 to +671 *psaA*, truncated PsaA	[23]

[1)] Plastome mutants with different transmission efficiencies were introgressed into the constant nuclear background of the johansen Standard strain (*O. elata* ssp. *hookeri*) by W. Stubbe. See Table S4 for details on this line.

[2)] See Table S5 for details on the wildtype plastomes.

[3)] In *Oenothera,* five genetically distinguishable plastome types (I-V) can be recognized based on their compatibility with three nuclear genomes (A, (B) C) in either homozygous (AA, BB, CC) or stable heterozygous (AB, AC, BC) states. The basic plastome genotype is accompanied by a given inheritance strength (strong, intermediate and weak). Basic plastome and nuclear genome type are an important factor of species definition in *Oenothera*. For details see, e.g. [6,7,11,12,14,30,35].

[4)] Classes of inhertiance strength as determined by F. Schötz. For reviews see [7,10,12].

[5)] For details on the definition of these classes see Supplementary Text.

Table S8. Inheritance strength of the wild type chloroplast genomes I-johSt, II-suavG, IV-atroSt[1)] and the green variants of plastomes I-johSt, determined in crosses to the chloroplast mutants I-chi and IV-delta[2)] as pollen donor and seed parents, respectively[3)]

Variant/ plastome	Generations of *pm* mutagenesis	[%] biparental inheritance (I-chi)[4)]	[%] biparental inheritance (IV-delta)[5)]	Inheritance class (Schötz)[6,7)]	Inheritance class (this work)[7)]	Sequence determined
I-johSt	wild type	6.0	42.1	strong	1	AJ271079.4
II-suavG	wild type	34.6	n/a	intermediate	3	EU262889.2
IV-atroSt	wild type	55.4	n/a	weak	4	EU262891.2
V1a	one	10.7	54.6	strong to intermediate	2	yes
V1b	one	7.9	31.6	strong	1	yes
V1c	one	5.0	40.7	strong	1	yes
V2a	one	13.1	36.6	strong	1	yes
V2b	one	14.9	28.1	intermediate	3	yes
V2e	one	15.9	49.0	strong to intermediate	2	no
V2f	one	7.4	41.8	strong	1	no
V2g	one	13.3	36.3	strong	1	yes
V3a	one	18.1	32.5	intermediate	3	yes
V3b	one	10.0	40.3	strong	1	yes
V3c	one	23.0	55.6	strong to intermediate	2	yes
V3d	one	14.7	35.1	intermediate	3	yes
V3e	one	8.0	44.4	strong	1	yes
V3f	one	15.2	56.8	strong to intermediate	2	yes
V3g	one	42.0	13.4	weak	4	yes
V3h	one	10.7	40.5	strong	1	yes
V4a	one	14.2	63.0	strong to intermediate	2	no
V4b	one	17.7	45.8	strong to intermediate	2	yes
V4c	one	10.7	35.3	strong	1	yes
V4d	one	10.0	36.7	strong	1	no
V7a	one	10.7	49.0	strong to intermediate	2	yes
V7c	one	27.9	42.0	intermediate	3	no
V7d	one	24.2	30.8	intermediate	3	no
VC1	several	40.9	7.4	weak	4	yes

[1)] See Table S6 for details on the wild type plastomes.

[2)] See Table S7 for details on the chloroplast mutants.

[3)] All crosses were performed in the nuclear genetic background of *O. elata* ssp. *hookeri* strain johansen Standard. For details on the crosses see Materials and Methods, on the line see Table S4.

[4)] Percentage of variegated seedlings, heteroplasmic due to the paternal transmission of the bleached chloroplast mutant I-chi and the maternal transmission of the green wild type or variant chloroplast of the seed parent. For details see Supplementary Text and Fig. 2

[5)] Percentage of variegated seedlings, heteroplasmic due to the paternal transmission of the green wild type or variant chloroplast and the maternal transmission of the bleached chloroplast mutant IV-delta of the seed parent. For details see Supplementary Text and Fig. 2.

[6)] Classes of inhertiance strength as determined by F. Schötz. For reviews, see [7,10,12].

[7)] For details on the definition of the classes in the variants see Supplementary Text.

Table S9. Wild type chloroplast genome III-lamS[1)] and its green variants created by the *plastome mutator* in crosses to I-johSt[1)] as pollen donor[2)]

Variant/ plastome	Generations of *pm* mutagenesis	Cross to I-johSt as pollen donor		Inheritance class (Schötz)[5,6)]	Inheritance class (this work)[6)]
		[%] biparental inheritance[3)]	[%] paternal inheritnace[4)]		
III-lamS	wild type	6.0	0.0	strong	1
III-V1	several	37.3	1.3	very weak	5
III-V2	several	38.0	1.6	very weak	5

[1)] See Table S6 for details on the wild type plastomes.

[2)] All crosses were performed in the constant nuclear genetic background of *O. elata* ssp. *hookeri* strain johansen Standard. For details on the crosses see Materials and Methods, on the line see Table S4.

[3)] Percentage of variegated seedlings, heteroplasmic due to the paternal transmission of the green wild type chloroplast I-johSt and the maternal transmission of the *viresencet* chloroplast of the seed parent. For details see Supplementary Text and Fig. S19.

[4)] Percentage of homoplasmic green seedlings resulting from paternal inheritance of the green wild type chloroplast I-johSt that fully out-competed the *virescent* chloroplast of the seed parent. For details see Supplementary Text and Fig. S19.

[5)] Classes of inhertiance strength as determined by F. Schötz. For reviews, see [7,10,12].

[6)] For details on the definition of the classes in the variants see Supplementary Text.

Table S10. Oligonucleotides employed in the MassARRAY and SNPs distinguishing I-johSt (AJ271079.4) and I-hookdV/I-chi (KT881170.1; I/I assay) or I-johSt and IV-atroSt/IV-delta (EU262891.2; I/IV assay)

Primer name	Assay	Sequence (5‘ - 3‘)	SNP (position in I-johSt)	Function
jh_atpI_W1-1	I/I	ACGTTGGATGGCCCTTGCTAACCAAGTCAT		PCR
jh_atpI_W1-2	I/I	ACGTTGGATGCGTCCCTGTATACAATCTAC	A/C (51,812)	PCR
jh_atpI_W1UEP	I/I	CAATCTACATGTAGGATACTG		Extension
jh_matK_W1-1	I/I	ACGTTGGATGATCATTCCGGGTCGGCTTAC		PCR
jh_matK_W1-2	I/I	ACGTTGGATGTGCCTCTGATTGGATCGTTG	T/C (2,235)	PCR
jh_matK_W1UEP	I/I	GCGAAATTTTGTAATGCATTAGG		Extension
jh_rpl32_W1-1	I/I	ACGTTGGATGCGTATTCGTAAAACTATTTGG		PCR
jh_rpl32_W1-2	I/I	ACGTTGGATGCCAGTAGAAAGAGACTTCGC	T/G (119,269)	PCR
jh_rpl32_W1UEP	I/I	CGCTAACGAAAAAGCTTTCAAGG		Extension
jh_rps16_W1-1	I/I	ACGTTGGATGAGTTGACAATTTCGGTACTG		PCR
jh_rps16_W1-2	I/I	ACGTTGGATGAGTATGTCAAGTCAACGTCC	G/A (6,919)	PCR
jh_rps16_W1UEP	I/I	CATCCATATTCTAATCTACCCGT		Extension
jh_rps19_W1-1	I/I	ACGTTGGATGATAGGCAAATGCTCTTTTCC		PCR
jh_rps19_W1-2	I/I	ACGTTGGATGTAAGAACTTGGTCTAGGGCG	G/C (89,483)	PCR
jh_rps19_W1UEP	I/I	CGATTATACTTACAATGATCGG		Extension
jh_trnK_W1-1	I/I	ACGTTGGATGGCGATTGTATCTACACATAG		PCR
jh_trnK_W1-2	I/I	ACGTTGGATGAGACGCACTTAAAAGCCGAG	C/A (4,124)	PCR
jh_trnK_W1UEP	I/I	TTGAGTTAGCAACCCCCCC		Extension
jh_trnQ_W1-1	I/I	ACGTTGGATGCATCATCGAAGTCATCATGC		PCR
jh_trnQ_W1-2	I/I	ACGTTGGATGTCGCGAACTTTATACTCCAC	A/G (61,685)	PCR
jh_trnQ_W1UEP	I/I	TGAAGGCATCAGACCAT		Extension
jh_trnS_W1-1	I/I	ACGTTGGATGCGGATCGTTTGATTCATTTC		PCR
jh_trnS_W1-2	I/I	ACGTTGGATGATTTAGGTGGTTCCCCACCC	T/G (20,001)	PCR
jh_trnS_W1UEP	I/I	CCCACCCTTTTTCTTTC		Extension
jh_ycf1-2_W1-1	I/I	ACGTTGGATGTCCGAGCTGAGCAAGTTGTG		PCR
jh_ycf1-2_W1-2	I/I	ACGTTGGATGTATGCTTCCTGCATAGCTCG	C/T (135,506)	PCR
jh_ycf1-2_W1UEP	I/I	CGCAGCTTCGATTTCCAT		Extension
jh_ycf1-3_W1-1	I/I	ACGTTGGATGCTATCGAACGAAACCCAAGC		PCR
jh_ycf1-3_W1-2	I/I	ACGTTGGATGCTCCTTTGCGCTTCGAATTT	G/T (135,768)	PCR
jh_ycf1-3_W1UEP	I/I	CTTTCTTCAAGTTTCTTCTCT		Extension
ja_atpH_W1-1	I/IV	ACGTTGGATGTTCTTAGCAACGGAAACCCC		PCR
ja_atpH_W1-2	I/IV	ACGTTGGATGTAGAATCTAGGGCGGGTTTC	A/G (52,971)	PCR
ja_atpH_W1UEP	I/IV	CTTGTACCAATCCCTAAAAAAA		Extension
ja_ndhD-1_W1-1	I/IV	ACGTTGGATGAGCGATACGGGACTTAATGG		PCR
ja_ndhD-1_W1-2	I/IV	ACGTTGGATGCCGGCCAAGAAAAAAAGAGC	G/T (121,802)	PCR
ja_ndhD-1_W1UEP	I/IV	AGAGCCGCCCCAATAAA		Extension
ja_ndhD-2_W1-1	I/IV	ACGTTGGATGAAACCAATACCCAGAACTCC		PCR
ja_ndhD-2_W1-2	I/IV	ACGTTGGATGCGTCTCTATTCAGCTACAAAG	A/C (122,300)	PCR
ja_ndhD-2_W1UEP	I/IV	AAAGTTCATTTTGTACACGGCGGG		Extension

Table S10. (continued)

Primer name	Assay	Sequence (5‘ - 3‘)	SNP (position in I-johSt)	Function
ja_ndhH_W1-1	I/IV	ACGTTGGATGTATTCAGCAGGCTCTGGAAG		PCR
ja_ndhH_W1-2	I/IV	ACGTTGGATGTTCAAAACCGTCCCATTCTG	C/A (127,723)	PCR
ja_ndhH_W1UEP	I/IV	CCATTCTGGATCCTTTTCT		Extension
ja_ndhJ_W1-1	I/IV	ACGTTGGATGACGTTCCCAATGTGCTTATG		PCR
ja_ndhJ_W1-2	I/IV	ACGTTGGATGGCTTGATCCACACCATACTC	T/C (15,205)	PCR
ja_ndhJ_W1UEP	I/IV	ACACTAGCTAACAGTCC		Extension
ja_petD-1_W1-1	I/IV	ACGTTGGATGCACACCTACATGAATGAATCC		PCR
ja_petD-1_W1-2	I/IV	ACGTTGGATGAAATTCCTTGCAATGGGTAG	G/A (82,065)	PCR
ja_petD-1_W1UEP	I/IV	AATGGGTAGTTGCAACTGC		Extension
ja_psaA_W1-1	I/IV	ACGTTGGATGGCCCTGCTAAATGGTGATTC		PCR
ja_psaA_W1-2	I/IV	ACGTTGGATGGCTTTTTGCTGGTTGGTTCC	G/T (23,840)	PCR
ja_psaA_W1UEP	I/IV	CTGGTTGGTTCCATTATCACAAAGC		Extension
ja_psaB_W1-1	I/IV	ACGTTGGATGAGATGTATTACCGCATCCCG		PCR
ja_psaB_W1-2	I/IV	ACGTTGGATGGGCATAAAGATTCCACTGGC	A/C (26,244)	PCR
ja_psaB_W1UEP	I/IV	TGGCCTGTAAAAAGGGG		Extension
ja_rpl116_W1-1	I/IV	ACGTTGGATGAACCCCGAATATTGGGTAGC		PCR
ja_rpl116_W1-2	I/IV	ACGTTGGATGTATTTTCTGCGACTCCACCC	T/G (86,534)	PCR
ja_rpl116_W1UEP	I/IV	ACTCCACCCATTTCATAAAG		Extension
ja_rpoB_W1-1	I/IV	ACGTTGGATGAGCGCGAACTGCTAGTTTTC		PCR
ja_rpoB_W1-2	I/IV	ACGTTGGATGCTTCAGACGTGTCAATTGGG	T/C (40,519)	PCR
ja_rpoB_W1UEP	I/IV	GTCAATTGGGCAAATACGTCCATAGT		Extension
ja_rps4_W1-1	I/IV	ACGTTGGATGTTTATGTCGCGTTACCGAGG		PCR
ja_rps4_W1-2	I/IV	ACGTTGGATGGGCTCTAGGCCTTTTACTGG	C/A (18,825)	PCR
ja_rps4_W1UEP	I/IV	GCCTTTTACTGGTTAGTCC		Extension
ja_trnW_W1-1	I/IV	ACGTTGGATGACTGAACTAAGAGCGCTTTC		PCR
ja_trnW_W1-2	I/IV	ACGTTGGATGAGCATACAAGAGGTATTGGG	A/G (72,243)	PCR
ja_trnW_W1UEP	I/IV	TTGGGATTACAAAACAAAAGA		Extension
ja_ycf1_W1-1	I/IV	ACGTTGGATGGCAATTCTCTACGCCGTTTG		PCR
ja_ycf1_W1-2	I/IV	ACGTTGGATGTATGCCTAAGACCAATGCGG	G/A (129,698)	PCR
ja_ycf1_W1UEP	I/IV	AGTGATTCTGATTTGTTTGTTCC		Extension
ja_ycf2_W1-1	I/IV	ACGTTGGATGCGAAAGAACGGAACTTAGCC	C/A (94,219; 161,260)	PCR
ja_ycf2_W1-2	I/IV	ACGTTGGATGATTCGAGTGATCGGTCTGAG		PCR
ja_ycf2_W1UEP	I/IV	CGGTCTGAGGTTAGCGAC		Extension
jah_atpF_W1-1	I/IV	ACGTTGGATGGGGCTTTTTCCAGCTGTTCG		PCR
jah_atpF_W1-2	I/IV	ACGTTGGATGATCGAAAACAGAGGATCTTG	T/C (54,206)	PCR
jah_atpF_W1UEP	I/IV	GAGGATCTTGAATACTATTCG		Extension

Table S11. Summary of F. Schötz's original transmission frequencies in hybrid populations of *Oenothera*, from which the different classes of inheritance strength were determined [6][1)]

Species	Strain	Abbreviation	Basic plastome[2)]	[%] biparental inheritance (biennis white)[3)]	[%] biparental inheritance (blandina white)[3)]	Inheritance class (Schötz)[4)]	Inheritance class (this work)[5)]
O. elata ssp. *hookeri*	hookeri de Vries	I-hookdV	I	0.0	8.2	strong	1
O. elata ssp. *hookeri*	franciscana de Vries	I-frandV	I	0.8	19.9	strong	1
O. villosa ssp. *villosa*	cockerelli de Vries	I-cockdV	I	0.0	9.9	strong	1
O. villosa ssp. *villosa*	mollis Standard	I-molSt	I	1.1	31.5	strong	1
O. villosa ssp. *villosa*	bauri Standard	I-bauriSt	I	7.4	43.8	strong	3
O. glazioviana	blandina de Vries	III-blandV	III	0.0	14.1	strong	1
O. glazioviana	deserens de Vries	III-desdV	III	1.3	21.4	strong	1
O. glazioviana	*r*/*r*-lamarckiana Sweden	III-lamS	III	3.1	26.4	strong	1
O. glazioviana	coronifera Standard	II-corSt	II	38.0	57.7	intermediate	4
O. biennis x *O. glazioviana*	conferta Stanard	II-conSt	II	16.1	34.2	intermediate	3
O. biennis x *O. villosa* ssp. *villosa*	hoelscheri Standard	II-hoeSt	II	31.1	57.5	intermediate	3
O. biennis	chicaginensis de Vries	III-chicdV	III	5.0	24.4	strong	1
O. biennis	nuda Standard	II-nudaSt	II	12.4	54.6	intermediate	3
O. biennis	purpurata Standard	II-pupSt	II	14.9	45.0	intermediate	3
O. biennis	biennis Muenchen	II-biM	II	22.2	53.8	intermediate	3
O. biennis	rubricaulis Thorn	II-rcauTh	II	23.2	41.8	intermediate	3
O. biennis	suaveolens Standard	II-suavSt	II	26.7	49.9	intermediate	3
O. biennis	grandiflora de Vries	II-gradV	II	28.3	53.5	intermediate	3
O. oakesiana	ammophila Standard	IV-ammSt	IV	43.9	58.1	weak	4
O. oakesiana	*r*/*r*-syrticola Ulm	IV-syrtU	IV	48.7	63.3	weak	4
O. oakesiana	germanica Standard	IV-gerSt	IV	52.9	74.5	weak	4
O. parviflora	parviflora Waldenburg	IV-parW	IV	47.6	63.2	weak	4
O. parviflora	rubricuspis Standard	IV-rcuSt	IV	52.3	68.7	weak	4
O. parviflora	silesiaca Standard	IV-silSt	IV	53.2	67.2	weak	4
O. parviflora	atrovirens Standard	IV-atroSt	IV	70.1	72.6	weak	4

[1)] For reviews see [7,10,12].

[2)] In *Oenothera,* five genetically distinguishable plastome types (I-V) can be recognized based on their compatibility with three nuclear genomes (A, (B) C) in either homozygous (AA, BB, CC) or stable heterozygous (AB, AC, BC) states. The basic plastome genotype is accompanied by a given inheritance strength (strong, intermediate and weak). Basic plastome and nuclear genome type are an important factor of species definition in *Oenothera*. For details, see e.g. [6,7,11,12,14,30,35].

[3)] Percentage of variegated seedlings, heteroplasmic due to the paternal transmission of the bleached chloroplast mutants “biennis white” or “blandina white” and the maternal transmission of the green wild chloroplast of the seed parent. Data accoding to Schötz [6]. For details see therein, reviews in [7,12] and **SI Text**.

[4)] Class of inhertiance strength as determined by F. Schötz. For reviews see [7,10,12].

[5)] For details on the definition of these classes see Supplementary Text.

Table S12. Nucleoids per chloroplasts in plants with the strong plastome I-johSt vs. the weak plastomes VC1, V3g, and IV-atroSt[1)]

Variant/ plastome[2)]	# Cells analyzed	# Chloroplasts analyzed	# nucleoids counted	# nucleoids/chloroplast[3)]	One-way ANOVA *p*-value	Adjusted *p*-value[4)]
I-johSt	22	266	4,703	17.7 ± 6.8	0.53	n/a
VC1	22	435	7,953	18.3 ± 5.2	0.53	0.32
V3g	22	418	7,625	18.2 ± 4.9	0.53	0.32
IV-atroSt	22	322	5,839	18.1 ± 5.7	0.53	0.38

[1)] All analyses were performed in the constant nuclear genetic background of *O. elata* ssp. *hookeri* strain johansen Standard. For details on the lines see Table S4.

[2)] For details on the wild type or variant plastomes see Table S8.

[3)] Means ± standard deviations are given.

[4)] *p*-values obtained with two-tailed homoscedastic *t*-test followed by multiple testing correction according to Benjamini-Hochberg.

Table S13. Chloroplasts per cell in plants with the strong plastome I-johSt vs. the weak plastomes VC1, V3g, and IV-atroSt[1)]

Variant/ plastome[2)]	# Cells analyzed	# chloroplasts counted	# chloroplasts/ cell[3)]	One-way ANOVA *p*-value	Adjusted *p*-value[4)]
I-johSt	55	2,872	52.2 ± 10.3	0.93	n/a
VC1	55	2,920	53.1 ± 10.2	0.93	0.66
V3g	55	2,918	53.1 ± 9.2	0.93	0.66
IV-atroSt	55	2,935	53.4 ± 7.7	0.93	0.66

[1)] All analyses were performed in the constant nuclear genetic background of *O. elata* ssp. *hookeri* strain johansen Standard. For details on the lines see Table S4.

[2)] For details on the wild type or variant plastomes see Table S8.

[3)] Means ± standard deviations are given.

[4)] *p*-values obtained with two-tailed homoscedastic *t*-test followed by multiple testing correction according to Benjamini-Hochberg.

Table S14. Comparison of mesophyll cell chloroplasts with the strong plastome I-johSt vs. the weak ones VC1, V3g, and IV-atroSt[1)]

Variant/ plastome[2)]	# chloroplasts analyzed	Chloroplast volume index [length x width^2][3,5)]	One-way ANOVA *p*-value	Adjusted *p*-value[4)]	Chloroplast length [µm][3)]	One-way ANOVA *p*-value	Adjusted *p*-value[4)]	Chloroplast width [µm][3)]	One-way ANOVA *p*-value	Adjusted *p*-value[4)]
I-johSt	60	86.8 ± 27.7	0.51	n/a	5.4 ± 0.6	0.66	n/a	4.0 ± 0.5	0.25	n/a
VC1	60	78.7 ± 31.6	0.51	0.42	5.3 ± 0.6	0.66	0.35	3.8 ± 0.6	0.25	0.20
V3g	60	81.2 ± 31.0	0.51	0.46	5.3 ± 0.7	0.66	0.35	3.8 ± 0.5	0.25	0.41
IV-atroSt	60	85.1 ± 36.6	0.51	0.78	5.3 ± 0.7	0.66	0.35	3.9 ± 0.6	0.25	0.81

[1)] All analyses were performed in the constant nuclear genetic background of *O. elata* ssp. *hookeri* strain johansen Standard. For details on the lines see Table S4.

[2)] For details on the wild type or variant plastomes see Table S8.

[3)] Means ± standard deviations are given.

[4)] *p*-values obtained with two-tailed homoscedastic *t*-test followed by multiple testing correction according to Benjamini-Hochberg.

[5)] Chloroplast volume index calculated according to [60].

Table S15. Oligonucleotides used for real-time qPCR.

Name	Genome	Gene/locus	Sequence (5‘ - 3‘)	Primer efficiency	Amplification product [bp]	GenBank/EMBL accession number template
rbcL-gene-SH_for	plastid	*rbcL*	TGCAGTAGCTGCCGAATCTTCTACTG	1.97	122	AJ271079.4
OerbcLqPCR_rev2	plastid	*rbcL*	GATTCTCTTCTCCCGCAACAGGCT			EU262891.2
psbB_qPCR_for	plastid	*psbB*	GTGGAGACAAGGTATGTTCGTTATAC	1.92	125	AJ271079.4
psbB_qPCR_rev2	plastid	*psbB*	CCACACCTTCGTAACTCCAAA			EU262891.2
Oe_ndhI_qPCR-fw	plastid	*ndhI*	CTGACATTGGTAAACGACCCAAAGC	1.94	126	AJ271079.4
Oe_ndhI_qPCR-rev	plastid	*ndhI*	GTGGTAACTGCGTTGAGTATTGTCC			EU262891.2
mtM03_qPCR_for	mitochondrion	mtM03	TCGCTTTCTTCGTCCACCAG	1.88	184	n/a
mtM03_qPCR_rev	mitochondrion	mtM03	CGATTCGCTTAGTTCGGTAGGGT			
mtM04qTAG_for	mitochondrion	mtM04	TTATTTTTGCAGTGTGGCTCCTAA	1.87	118	n/a
mtM04qTAG_rev	mitochondrion	mtM04	CCCAGCGAGAACGCTACCT			
mtM06qTAG_for	mitochondrion	mtM06	GGAAGTCAAGCTAAAGCAGTCTAAGC	1.89	93	n/a
mtM06qTAG_rev	mitochondrion	mtM06	CTCTCATTAGCTCTCCTTTTTTTATATGC			
M02-1_qPCR-Fw	nucleus	M02	CACAAGCCTCCATCTTTACTCCGT	1.94	177	EU483117.1
M02-1_qPCR-Rev	nucleus	M02	GCTACCTTCTCCTTGGTAGGCG			
M19_qPCR-Fw	nucleus	M19	GTGAAGCTCTCCCCTGAGGT	1.92	82	EU447207.1
M19_qPCR-Rev	nucleus	M19	ATGAGACAGCGGGTTTGGC			
Mpgi02_qPCR-Fw	nucleus	*pigC*	CAGGTCTGTACTACATGTCGCTCT	1.97	141	GU176508.1
Mpgi02_qPCR-Rev	nucleus	*pigC*	ACCCATGAACCATTGCGAACT			

References

1 **Greiner S, Sobanski J, Bock R.** 2015. Why are most organelle genomes transmitted maternally? *Bioessays* **37**: 80-94.

2 **Grun P.** 1976. Cytoplasmic Genetics and Evolution New York: Columbia University Press.

3 **Hoekstra RF.** 1990. Evolution of uniparental inheritance of cytoplasmic DNA. In Smith MJ, Vida J, eds; *Organizational Constrains of the Dynamics of Evolution*. Manchester: Manchester University Press. p 269-78.

4 **Cosmides LM, Tooby J.** 1981. Cytoplasmic inheritance and intragenomic conflict. *J Theor Biol* **89**: 83-129.

5 **Eberhard WG.** 1980. Evolutionary consequences of intracellular organelle competition. *Q Rev Biol* **55**: 231-49.

6 **Schötz F.** 1968. Über die Plastidenkonkurrenz bei *Oenothera* II. *Biol Zentralbl* **87**: 33-61.

7 **Kirk JTO, Tilney-Bassett RAE.** 1978. The plastids. Their Chemistry, Structure, Growth and Inheritance Amsterdam, New York, Oxford: Elsevier.

8 **de Stordeur E.** 1997. Nonrandom partition of mitochondria in heteroplasmic *Drosophila*. *Heredity* **79**: 615-23.

9 **Ma H, O'Farrell PH.** 2016. Selfish drive can trump function when animal mitochondrial genomes compete. *Nat Genet* **48**: 798-802.

10 **Gillham NW.** 1978. Organelle Heredity New York: Raven Press.

11 **Greiner S, Rauwolf U, Meurer J, Herrmann RG.** 2011. The role of plastids in plant speciation. *Mol Ecol* **20**: 671-91.

12 **Cleland RE.** 1972. *Oenothera* - Cytogenetics and Evolution London, New York: Academic Press Inc.

13 **Kikuchi S, Bédard J, Hirano M, Hirabayashi Y, et al.** 2013. Uncovering the protein translocon at the chloroplast inner envelope membrane. *Science* **339**: 571-4.

14 **Chiu W-L, Stubbe W, Sears BB.** 1988. Plastid inheritance in *Oenothera*: organelle genome modifies the extent of biparental plastid transmission. *Curr Genet* **13**: 181-9.

15 **Salie MJ, Thelen JJ.** 2016. Regulation and structure of the heteromeric acetyl-CoA carboxylase. *Biochim Biophys Acta* **1861**: 1207-13.

16 **Stoike LL, Sears BB.** 1998. *Plastome mutator*-induced alterations arise in *Oenothera* chloroplast DNA through template slippage. *Genetics* **149**: 347-53.

17 **Sears BB, Stoike LL, Chiu WL.** 1996. Proliferation of direct repeats near the *Oenothera* chloroplast DNA origin of replication. *Mol Biol Evol* **13**: 850-63.

18 **Chiu W-L, Sears BB.** 1992. Electron microscopic localization of replication origins in *Oenothera* chloroplast DNA. *Mol Gen Genet* **232**: 33-9.

19 **Botella C, Jouhet J, Block MA.** 2017. Importance of phosphatidylcholine on the chloroplast surface. *Prog Lipid Res* **65**: 12-23.

20 **Osteryoung KW, Pyke KA.** 2014. Division and Dynamic Morphology of Plastids. *Annu Rev Plant Biol* **65**: 443-72.

21 **Barnard-Kubow KB, McCoy MA, Galloway LF.** 2017. Biparental chloroplast inheritance leads to rescue from cytonuclear incompatibility. *New Phytol* **213**: 1466-76.

22 **Nishimura Y, Stern DB.** 2010. Differential replication of two chloroplast genome forms in heteroplasmic *Chlamydomonas reinhardtii* gametes contributes to alternative inheritance patterns. *Genetics* **185**: 1167-81.

23 **Massouh A, Schubert J, Yaneva-Roder L, Ulbricht-Jones ES, et al.** 2016. Spontaneous chloroplast mutants mostly occur by replication slippage and show a biased pattern in the plastome of *Oenothera*. *Plant Cell* **28**: 911-29.

24 **Greiner S, Wang X, Rauwolf U, Silber MV, et al.** 2008. The complete nucleotide sequences of the five genetically distinct plastid genomes of *Oenothera*, subsection *Oenothera*: I. Sequence evaluation and plastome evolution. *Nucleic Acids Res* **36**: 2366-78.

25 **Rockenbach K, Havird JC, Monroe JG, Triant DA, et al.** 2016. Positive selection in rapidly

evolving plastid-nuclear enzyme complexes. *Genetics* **204**: 1507-22.

26 **Hirata N, Yonekura D, Yanagisawa S, Iba K.** 2004. Possible involvement of the 5'-flanking region and the 5'UTR of plastid *accD* gene in NEP-dependent transcription. *Plant Cell Physiol* **45**: 176–86.

27 **Hajdukiewicz PT, Allison LA, Maliga P.** 1997. The two RNA polymerases encoded by the nuclear and the plastid compartments transcribe distinct groups of genes in tobacco plastids. *EMBO J* **16**: 4041-8.

28 **Swiatecka-Hagenbruch M, Liere K, Börner T.** 2007. High diversity of plastidial promoters in *Arabidopsis thaliana*. *Mol Genet Genom* **277**: 725-34.

29 **Kozaki A, Mayumi K, Sasaki Y.** 2001. Thiol-disulfide exchange between nuclear-encoded and chloroplast-encoded subunits of pea acetyl-CoA carboxylase. *J Biol Chem* **276**: 39919-25.

30 **Dietrich W, Wagner WL, Raven PH.** 1997. Systematics of *Oenothera* section *Oenothera* subsection *Oenothera* (Onagraceae) Laramie: The American Society of Plant Taxonomists.

31 **Greiner S, Köhl K.** 2014. Growing evening primroses (*Oenothera*). *Front Plant Sci* **5**: 38.

32 **Harte C.** 1994. *Oenothera* - Contributions of a Plant to Biology Berlin, Heidelberg, New York: Springer.

33 **Schötz F.** 1954. Über Plastidenkonkurrenz bei *Oenothera*. *Planta* **43**: 182-240.

34 **Cleland RE.** 1935. Cyto-taxonomic studies on certain Oenotheras from California. *Proc Am Philos Soc* **75**: 339-429.

35 **Stubbe W.** 1989. *Oenothera* - An ideal system for studying the interaction of genome and plastome. *Plant Mol Biol Rep* **7**: 245-57.

36 **Kutzelnigg H, Stubbe W.** 1974. Investigation on plastome mutants in *Oenothera*: 1. General considerations. *Subcell Biochem* **3**: 73-89.

37 **Stubbe W, Herrmann RG.** 1982. Selection and maintenance of plastome mutants and interspecific genome/plastome hybrids from *Oenothera*. In Edelman M, Hallick RB, Chua N-H, eds; *Methods in Chloroplast Molecular Biology*. Amsterdam, New York, Oxford: Elsevier. p 149-65.

38 **Epp MD.** 1973. Nuclear gene-induced plastome mutations in *Oenothera hookeri*: I. Genetic analysis. *Genetics* **75**: 465-83.

39 **Chang T-L, Stoike LL, Zarka D, Schewe G, et al.** 1996. Characterization of primary lesions caused by the plastome mutator of *Oenothera*. *Curr Genet* **30**: 522-30.

40 **Chiu W-L, Johnson EM, Kaplan SA, Blasko K, et al.** 1990. *Oenothera* chloroplast DNA polymorphisms associated with plastome mutator activity. *Mol Gen Genet* **221**: 59-64.

41 **Sears BB, Sokalski MB.** 1991. The *Oenothera* plastome mutator: effect of UV irradiation and nitroso-methyl urea on mutation frequencies. *Mol Gen Genet* **229**: 245-52.

42 **Team RC.** 2015. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

43 **Rauwolf U, Golczyk H, Meurer J, Herrmann RG, et al.** 2008. Molecular marker systems for *Oenothera* genetics. *Genetics* **180**: 1289-306.

44 **Ågren JA, Greiner S, Johnson MTJ, Wright SI.** 2015. No evidence that sex and transposable elements drive genome size variation in evening primroses. *Evolution* **69**: 1053–62.

45 **Tillich M, Lehwark P, Ulbricht-Jones ES, Fischer A, et al.** 2017. GeSeq - versatile and accurate annotation of organelle genomes. *Nucleic Acids Res* **45**: W6-W11.

46 **Lehwark P, Greiner S.** 2018. GenBank 2 Sequin - a file converter preparing custom GenBank files for database submission. *bioRxiv*.

47 **Rice P, Longden I, Bleasby A.** 2000. EMBOSS: the European molecular biology open software suite. *Trends Genet* **16**: 276-7.

48 **Thompson JD, Higgins DG, Gibson TJ.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**: 4673-80.

49 **Lohse M, Drechsel O, Kahlau S, Bock R.** 2013. OrganellarGenomeDRAW - a suite of tools for

generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Res* **41**: W575–W81.

50 **Kearse M, Moir R, Wilson A, Stones-Havas S, et al.** 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**: 1647-9.

51 **Hunter SC, Ohlrogge JB.** 1998. Regulation of spinach chloroplast acetyl-CoA carboxylase. *Arch Biochem Biophys* **359**: 170-8.

52 **Thelen JJ, Ohlrogge JB.** 2002. The multisubunit acetyl-CoA carboxylase is strongly associated with the chloroplast envelope through non-ionic interactions to the carboxyltransferase subunits. *Arch Biochem Biophys* **400**: 245-57.

53 **Salem MA, Juppner J, Bajdzienko K, Giavalisco P.** 2016. Protocol: a fast, comprehensive and reproducible one-step extraction method for the rapid preparation of polar and semi-polar metabolites, lipids, proteins, starch and cell wall polymers from a single sample. *Plant Methods* **12**: 45.

54 **Hummel J, Segu S, Li Y, Irgang S, et al.** 2011. Ultra performance liquid chromatography and high resolution mass spectrometry for the analysis of plant lipids. *Front Plant Sci* **2**: 54.

55 **Tibshirani R.** 1996. Regression shrinkage and selection via the Lasso. *J Royal Stat Soc Ser B* **58**: 267-88.

56 **Porra RJ, Thompson WA, Kriedemann PE.** 1989. 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* **975**: 384-94.

57 **Rauwolf U, Golczyk H, Greiner S, Herrmann R.** 2010. Variable amounts of DNA related to the size of chloroplasts: III. Biochemical determinations of DNA amounts per organelle. *Mol Genet Genom* **283**: 35–47.

58 **Golczyk H, Greiner S, Wanner G, Weihe A, et al.** 2014. Chloroplast DNA in mature and senescing leaves: a reappraisal. *Plant Cell* **26**: 847-54.

59 **James TW, Jope C.** 1978. Visualization by fluorescence of chloroplast DNA in higher plants by means of the DNA-specific probe 4'6-diamidino-2-phenylindole. *J Cell Biol* **79**: 623-30.

60 **Butterfass T.** 1979. Patterns of chloroplast reproduction. A developmental approach to protoplasmic plant anatomy Vienna, New York: Springer.

61 **Johnson MTJ, Smith SD, Rausher MD.** 2009. Plant sex and the evolution of plant defenses against herbivores. *Proc Natl Acad Sci* **106**: 18079-84.

62 **Dotzek J.** 2016. Mitochondria in the genus *Oenothera* - Non-Mendelian inheritance patterns, in vitro structure and evolutionary dynamics [PhD Thesis]. University of Potsdam.

63 **Binder S, Brennicke A.** 1993. Transcription initiation sites in mitochondria of *Oenothera berteriana*. *J Biol Chem* **268** 7849-55.

64 **Greiner S, Wang X, Herrmann RG, Rauwolf U, et al.** 2008. The complete nucleotide sequences of the 5 genetically distinct plastid genomes of *Oenothera*, subsection *Oenothera*: II. A microevolutionary view using bioinformatics and formal genetic data. *Mol Biol Evol* **25**: 2019-30.

65 **Simon M, Durand S, Pluta N, Gobron N, et al.** 2016. Genomic conflicts that cause pollen mortality and raise reproductive barriers in *Arabidopsis thaliana*. *Genetics* **203**: 1353-67.

66 **Luo D, Xu H, Liu Z, Guo J, et al.** 2013. A detrimental mitochondrial-nuclear interaction causes cytoplasmic male sterility in rice. *Nat Genet* **45**: 573-7.

67 **Taylor RW, Turnbull DM.** 2005. Mitochondrial DNA mutations in human disease. *Nat Rev Genet* **6**: 389-402.

68 **Greiner S, Bock R.** 2013. Tuning a ménage à trois: co-evolution and co-adaptation of nuclear and organellar genomes in plants. *Bioessays* **35**: 354-65.

69 **Roux F, Mary-Huard T, Barillot E, Wenes E, et al.** 2016. Cytonuclear interactions affect adaptive traits of the annual plant *Arabidopsis thaliana* in the field. *Proc Natl Acad Sci* **113**: 3687-92.

70 **Ruf S, Kössel H, Bock R.** 1997. Targeted inactivation of a tobacco intron-containing open reading frame reveals a novel chloroplast-encoded photosystem I-related gene. *J Cell Biol* **1**: 95-102.

71 **Wolfson R, Higgins KG, Sears BB.** 1991. Evidence for replication slippage in the evolution of *Oenothera* chloroplast DNA. *Mol Biol Evol* **8**: 709-20.

72 **Williamson D.** 2002. The curious history of yeast mitochondrial DNA. *Nat Rev Genet* **3**: 475-81.

73 **MacAlpine DM, Kolesar J, Okamoto K, Butow RA, et al.** 2001. Replication and preferential inheritance of hypersuppressive petite mitochondrial DNA. *EMBO J* **20**: 1807-17.

74 **Nicholls TJ, Minczuk M.** 2014. In D-loop: 40 years of mitochondrial 7S DNA. *Exp Gerontol* **56**: 175-81.

75 **Lakshmanan LN, Gruber J, Halliwell B, Gunawan R.** 2015. Are mutagenic non D-loop direct repeat motifs in mitochondrial DNA under a negative selection pressure? *Nucleic Acids Res* **43**: 4098-108.

76 **Zoschke R, Liere K, Börner T.** 2007. From seedling to mature plant: *Arabidopsis* plastidial genome copy number, RNA accumulation and transcription are differentially regulated during leaf development. *Plant J* **50**: 710-22.

77 **Rast A, Heinz S, Nickelsen J.** 2015. Biogenesis of thylakoid membranes. *Biochim Biophys Acta* **1847**: 821-30.

78 **Glick RE, Sears BB.** 1994. Genetically programmed chloroplast dedifferentiation as a consequence of plastome-genome incompatibility in *Oenothera*. *Plant Physiol* **106**: 367-73.

79 **Kutzelnigg H, Meyer B, Schötz F.** 1975. Untersuchungen an Plastom-Mutanten von *Oenothera*: II. Überblick über die Ultrastruktur der mutierten Plastiden. *Biol Zentralbl* **94**: 513 – 26.

80 **Kutzelnigg H, Mayer B, Schötz F.** 1975. Untersuchung an Plastom-Mutanten von *Oenothera*: III. Vergleichende Untersuchung ultrastruktureller Charakterizierung der Mutanten. *Biol Zentralbl* **94**: 527-38.

81 **Pribil M, Labs M, Leister D.** 2014. Structure and dynamics of thylakoids in land plants. *J Exp Bot* **65**: 1955-72.

82 **Schötz F.** 1975. Untersuchungen über die Plastidenkonkurrenz bei *Oenothera*: V. Die Stabilität der Konkurrenzfähigkeit bei Verwendung von verschiedenartiger mutierter Testplastiden. *Biol Zentralbl* **95**: 17-26.

83 **Ohlrogge J, Browse J.** 1995. Lipid biosynthesis. *Plant Cell* **7**: 957-70.

84 **Ferro M, Brugière S, Salvi D, Seigneurin-Berny D, et al.** 2010. AT_CHLORO, a comprehensive chloroplast proteome database with subplastidial localization and curated information on envelope proteins. *Mol Cell Proteom* **9**: 1063-84.

85 **de Vries H.** 1913. Gruppenweise Artbildung - Unter spezieller Berücksichtigung der Gattung *Oenothera* Berlin: Gebrüder Borntraeger.

86 **Baerecke M.** 1944. Zur Genetik und Cytologie von *Oenothera ammophila* Focke, *Bauri* Boedijn, *Beckeri* Renner, *parviflora* L., *rubricaulis* Klebahn, *silesiaca* Renner. *Flora* **138**: 57-92.

87 **de Vries H.** 1919. *Oenothera rubrinervis*, a half mutant. *Bot Gaz* **67**: 1-26.

88 **de Vries H.** 1923. Über die Entstehung der *Oenothera Lamarckiana* mut. *Veluntina*. *Biol Centralbl* **43**: 213-24.

89 **de Vries H.** 1917. *Oenothera Lamarckiana* mut. *veluntina*. *Bot Gaz* **63**: 1-24.

90 **Heribert-Nilsson N.** 1912. Die Variabilität der *Oenothera Lamarckiana* und das Problem der Mutation. *Z indukt Abstamm Vererbungsl* **8**: 89-231.

91 **Renner O.** 1917. Versuche über die gametische Konstitution der Önotheren. *Z indukt Abstamm Vererbungsl* **18**: 121-294.

92 **Renner O.** 1943. Kurze Mitteilung über *Oenothera*: V. Zur Kenntnis von *O. silesiaca* n. sp., *parviflora* L., *ammophila* Focke, *rubricaulis* Kleb. *Flora* **36**: 325-35.

93 **de Vries H.** 1918. Mutations of *Oenothera suaveolens* DESF. *Genetics* **3**: 1-26.

94 **Stubbe W.** 1953. Genetische und zytologische Untersuchungen an verschiedenen Sippen von

Oenothera suaveolens. *Z indukt Abstamm Vererbungsl* **85**: 180-209.

95 **Hoeppener E, Renner O.** 1929. Genetische und zytologische Oenotherenstudien: I. Zur Kenntnis der *Oenothera ammophila* Focke. *Z indukt Abstamm Vererbungsl* **49**: 1-25.

96 **Renner O.** 1937. Wilde Oenotheren in Norddeutschland. *Flora* **31**: 182-226.

97 **Renner O.** 1950. Europäische Wildarten von *Oenothera*: II. *Ber Dtsch Bot Ges* **63**: 129-38.

98 **Renner O.** 1938. Über *Oenothera atrovirens* Sh. et Bartl. und über somatische Konversion im Erbgang des *cruciata*-Merkmals der Oenotheren. *Z indukt Abstamm Vererbungsl* **74**: 91-124.

99 **de Vries H.** 1901 - 1903. Die Mutationstheorie Leipzig: von Veit.

100 **Bartlett HH.** 1914. An account of the cruciate-flowered Oenotheras of the subgenus *Onagra*. *Am J Bot* **1**: 226-43.