



WAGENINGEN
UNIVERSITY & RESEARCH

Reciprocal cybrids reveal how organellar genomes affect plant phenotypes

Nature Plants

Flood, Pádraic J.; Theeuwen, Tom P.J.M.; Schneeberger, Korbinian; Keizer, Paul; Kruijer, Willem et al
<https://doi.org/10.1038/s41477-019-0575-9>

This article is made publicly available in the institutional repository of Wageningen University and Research, under the terms of article 25fa of the Dutch Copyright Act, also known as the Amendment Taverne. This has been done with explicit consent by the author.

Article 25fa states that the author of a short scientific work funded either wholly or partially by Dutch public funds is entitled to make that work publicly available for no consideration following a reasonable period of time after the work was first published, provided that clear reference is made to the source of the first publication of the work.

This publication is distributed under The Association of Universities in the Netherlands (VSNU) 'Article 25fa implementation' project. In this project research outputs of researchers employed by Dutch Universities that comply with the legal requirements of Article 25fa of the Dutch Copyright Act are distributed online and free of cost or other barriers in institutional repositories. Research outputs are distributed six months after their first online publication in the original published version and with proper attribution to the source of the original publication.

You are permitted to download and use the publication for personal purposes. All rights remain with the author(s) and / or copyright owner(s) of this work. Any use of the publication or parts of it other than authorised under article 25fa of the Dutch Copyright act is prohibited. Wageningen University & Research and the author(s) of this publication shall not be held responsible or liable for any damages resulting from your (re)use of this publication.

For questions regarding the public availability of this article please contact openscience.library@wur.nl

Reciprocal cybrids reveal how organellar genomes affect plant phenotypes

Pádraic J. Flood^{1,2,3,10*}, Tom P. J. M. Theeuwen^{1,10*}, Korbinian Schneeberger³, Paul Keizer⁴, Willem Kruijer⁴, Edouard Severing³, Evangelos Kouklas¹, Jos A. Hageman⁴, Raúl Wijffjes⁵, Vanesa Calvo-Baltanas¹, Frank F. M. Becker¹, Sabine K. Schnabel⁴, Leo A. J. Willems⁶, Wilco Ligterink⁶, Jeroen van Arkel⁷, Roland Mumm⁷, José M. Gualberto⁸, Linda Savage⁹, David M. Kramer⁹, Joost J. B. Keurentjes¹, Fred van Eeuwijk⁴, Maarten Koornneef^{1,3}, Jeremy Harbinson², Mark G. M. Aarts¹ and Erik Wijnker^{1*}

Assessment of the impact of variation in chloroplast and mitochondrial DNA (collectively termed the plasmotype) on plant phenotypes is challenging due to the difficulty in separating their effect from nuclear-derived variation (the nucleotype). Haploid-inducer lines can be used as efficient plasmotype donors to generate new plasmotype-nucleotype combinations (cybrids)¹. We generated a panel comprising all possible cybrids of seven *Arabidopsis thaliana* accessions and extensively phenotyped these lines for 1,859 phenotypes under both stable and fluctuating conditions. We show that natural variation in the plasmotype results in both additive and epistatic effects across all phenotypic categories. Plasmotypes that induce more additive phenotypic changes also cause more epistatic effects, suggesting a possible common basis for both additive and epistatic effects. On average, epistatic interactions explained twice as much of the variance in phenotypes as additive plasmotype effects. The impact of plasmotypic variation was also more pronounced under fluctuating and stressful environmental conditions. Thus, the phenotypic impact of variation in plasmotypes is the outcome of multi-level nucleotype-plasmotype-environment interactions and, as such, the plasmotype is likely to serve as a reservoir of variation that is predominantly exposed under certain conditions. The production of cybrids using haploid inducers is a rapid and precise method for assessment of the phenotypic effects of natural variation in organellar genomes. It will facilitate efficient screening of unique nucleotype-plasmotype combinations to both improve our understanding of natural variation in these combinations and identify favourable combinations to enhance plant performance.

Chloroplasts and mitochondria play essential roles in metabolism, cellular homeostasis and environmental sensing^{2,3}. Their genomes contain only a limited set of genes whose functioning requires tight coordination with the nucleus through signalling pathways that modulate nuclear and organellar gene expression^{3,4}. Plasmotype variation

can be strongly additive, such as in the case of chloroplast-encoded herbicide tolerance⁵, or can manifest itself in complex cytonuclear interactions as non-additive, nonlinear effects (epistasis) such as found for secondary metabolites⁶. The phenotypic consequences of epistasis can be detected when a plasmotype causes phenotypic effects in combination with some, but not all, nuclear backgrounds. Recent studies suggest that cytonuclear epistasis is the main route through which variation in the plasmotype is expressed^{6–12}, and that additive effects are both rare and of marginal effect.

Plasmotypic variation is relevant from both an agricultural and an evolutionary perspective^{13–15}, but to understand or utilize it, it is necessary to separate nuclear from mitochondrial and chloroplastic effects. Reciprocal cross designs, where nucleotypes segregate in different plasmotypic backgrounds, have been used to identify plasmotype-specific quantitative trait loci^{6,10}, but are limited to just two plasmotypes. A larger number of plasmotypes can be studied using back-cross designs where plasmotypes are introgressed into different nuclear backgrounds^{11,16–18}, but back-crossing approaches are lengthy and any undetected nuclear introgressions may confound the results.

To precisely and rapidly address the contribution of organellar variation to plant phenotypes, we explored the use of a haploid-inducer (HI) line available in *Arabidopsis* (*GFP-tailswap*)^{1,19}. When pollinated with a wild-type (WT) plant, the *GFP-tailswap* nuclear genome is lost from the zygote through uniparental genome elimination. This generates haploid cybrid offspring with a paternally derived nuclear genome and maternally (*GFP-tailswap*) derived mitochondria and chloroplasts (Fig. 1). These haploid plants produce stable diploid (doubled haploid) offspring following genome duplication or restitutional meiosis¹⁹. We set out to test the use of this approach to investigate how plasmotypic variation affects plant phenotypes and to what extent this variation manifests itself as either additive variation or cytonuclear epistasis.

Seven different *Arabidopsis* accessions were selected for our experiment: six that represent a snapshot of natural variation (Bur,

¹Laboratory of Genetics, Wageningen University & Research, Wageningen, the Netherlands. ²Horticulture and Product Physiology, Wageningen University & Research, Wageningen, the Netherlands. ³Department of Plant Developmental Biology, Max Planck Institute for Plant Breeding Research, Cologne, Germany. ⁴Biometris, Wageningen University & Research, Wageningen, the Netherlands. ⁵Bioinformatics Group, Wageningen, the Netherlands. ⁶Laboratory of Plant Physiology, Wageningen University & Research, Wageningen, the Netherlands. ⁷Bioscience, Wageningen University & Research, Wageningen, the Netherlands. ⁸Institut de Biologie Moléculaire des Plantes, CNRS, Université de Strasbourg, Strasbourg, France. ⁹MSU-DOE Plant Research Lab, Michigan State University, East Lansing, MI, USA. ¹⁰These authors contributed equally: Pádraic J. Flood, Tom P.J.M. Theeuwen. *e-mail: flood@mpipz.mpg.de; tom.theeuwen@wur.nl; erik.wijnker@wur.nl

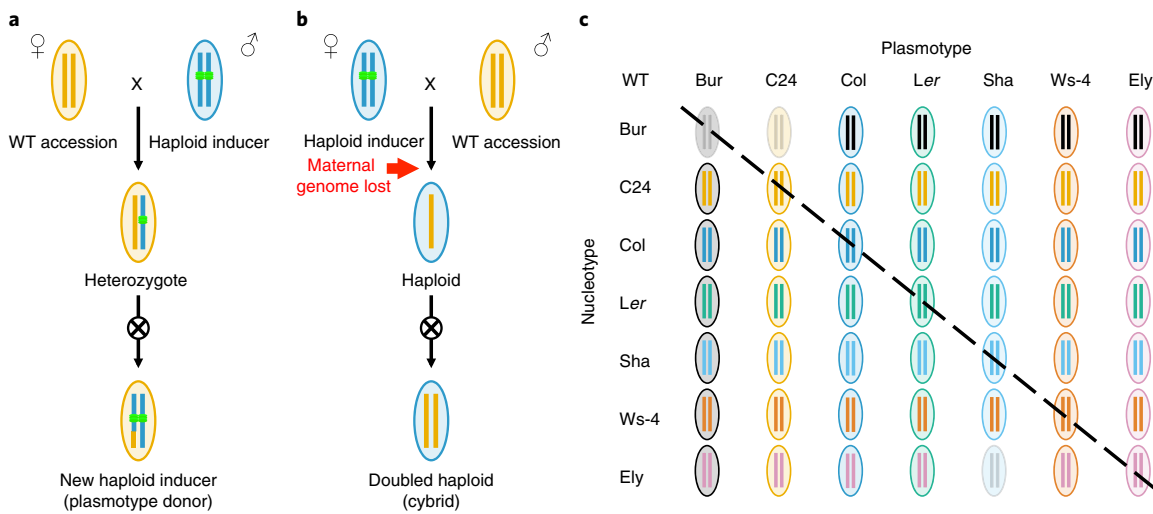


Fig. 1 | Generation of a cybrid test panel. **a**, Generation of a new haploid-inducer (HI) line with a new plasmotype. The HI expresses a GFP-tagged *CENH3/HRT12* in a *cenh3/htr12* mutant background. A cross between a WT (female) and a HI (male) results in a hybrid F1. A diploid F1 is selected in which no genome elimination has occurred. Self-fertilization generates an F2 population in the plasmotype of the WT mother, from which an F2 plant is selected that is homozygous for the *cenh3/htr12* mutation and carries the *GFP-tailswap* transgene. This F2 plant is a new HI line and can serve as a plasmotype donor when used as a female in crosses. Vertical bars represent the nucleotype and ovals represent the plasmotype. HI centromeres (indicated in green, signifying GFP-tagged CENH3/HRT12 proteins as encoded by the *GFP-tailswap* construct) cause uniparental genome elimination. **b**, HI lines can function as plasmotype donors when used as a female parent. In this case, uniparental genome elimination (red arrow) leads to a haploid offspring plant with the nucleotype of the WT male parent but the plasmotype of the HI mother. **c**, Full diallel of all nucleotype-plasmotype combinations for which cybrids were generated. The broken diagonal line highlights the WT nucleotype-plasmotype combinations generated by crossing WT plants and plasmotype donors with the plasmotype of the WT (self-cybrids). *Bur*^{Bur}, *Bur*^{C24} and *Ely*^{Sha} appear faded because they were not included in the phenotyping experiments but have subsequently been recreated.

C24, Col-0, *Ler*-0, Sha and WS-4) and *Ely*, an accession with a large-effect mutation in the chloroplast-encoded *psbA* gene²⁰. This mutation results in reduced photosystem II (PSII) efficiency^{20,21} and was included to evaluate the consequence of a strong plasmotype effect in our test panel. We first generated haploid inducers for all seven plasmotypes (Fig. 1a) and then used each inducer to generate cybrid offspring for all seven nucleotypes (Fig. 1b,c). Cybrid genotypes will henceforth be denoted as nucleotype^{plasmotype} (that is, *Ely*^{Bur} denotes a cybrid with *Ely* nucleotype and *Bur* plasmotype). Wild-type nucleotype-plasmotype combinations were also regenerated in this way (hereafter referred to as self-cybrids—that is, *Bur*^{Bur}, *C24*^{C24} and so on) for subsequent comparison to their WT progenitors. The genotypes of all haploid cybrids were verified by whole-genome resequencing. This led to the exclusion of *Bur*^{C24} and *Bur*^{Bur}, which were identified as containing the same nucleotypic de novo duplication of 200 kb, probably derived from a spontaneous duplication in a *Bur*-WT progenitor used in creating these cybrids (see Methods and Extended Data Fig. 1). With the exception of *Ely*^{Sha}, for which we obtained seeds at a later stage, we acquired doubled haploid seeds from all haploid cybrids resulting in a test panel of 46 cybrids and 7 WT progenitors. As with *Ely*^{Sha}, *Bur*^{C24} and *Bur*^{Bur} were subsequently recreated and the complete panel will be submitted to the European Arabidopsis Stock Centre (www.arabidopsis.info). To visualize the genetic variation between lines within our panel we generated neighbour-joining trees for the nuclear, mitochondrial and chloroplast genomes (Extended Data Fig. 2 and Supplementary Figs. 1–3). The nucleotypes were found to be approximately equidistant, while the *Ler*, *Ely* and *Col* plasmotypes appear to be more closely related to each other than to the other plasmotypes.

We phenotyped the cybrid panel under constant environmental conditions for absolute and relative growth rate, biomass accumulation, epinastic leaf movement, PSII efficiency (Φ_{PSII}), non-photochemical quenching (NPQ) and elements thereof (Φ_{NO} , Φ_{NPQ} , q_E

and q_1), a reflectance-based estimate of chlorophyll, flowering time, germination, pollen abortion and primary metabolites. To simulate the more variable conditions that are frequently encountered in the field, we also screened the panel under fluctuating light for all the above-mentioned photosynthesis-related phenotypes, and assayed germination rates under osmotic stress and after a controlled deterioration treatment. Counting individual metabolite concentrations and single time points in the time series separately, we collected in total 1,859 phenotypes (Supplementary Data 1 and Supplementary Table 4). To avoid over-representation of highly correlated and non-informative phenotypes, we selected a subset of 92 phenotypes (Methods and Supplementary Table 2) comprising 24 from constant-growth conditions, 32 from fluctuating or challenging environmental conditions and 36 primary metabolites for further analysis (Extended Data Fig. 3 and Supplementary Table 2).

Comparison of six self-cybrids to their genetically identical WT progenitors for these 92 phenotypes did not reveal significant phenotypic differences (Supplementary Table 1), from which we infer that uniparental genome elimination is a robust method for generation of cybrids. To determine the relative contributions of nucleotype, plasmotype and their interaction to the observed phenotypic variation, we estimated the fraction of broad-sense heritability (H^2 , also called repeatability²²) explained by each. Across the entire panel the average contribution to H^2 of nucleotype, plasmotype and nucleotype-plasmotype interaction was 65.9, 28.0 and 6.1%, respectively (Supplementary Tables 2 and 3 and Supplementary Data 2). Most of the plasmotype-derived additive variation was caused by the *Ely* plasmotype, arising from the *psbA* mutation. When this plasmotype was excluded from the analysis, the nucleotype, plasmotype and their interaction accounted for 91.9, 2.9 and 5.2% of genetic variation, respectively (Supplementary Tables 2 and 3 and Supplementary Data 2). Thus, while nucleotype-derived additive variation is the main genetic determinant of the cybrid phenotype,

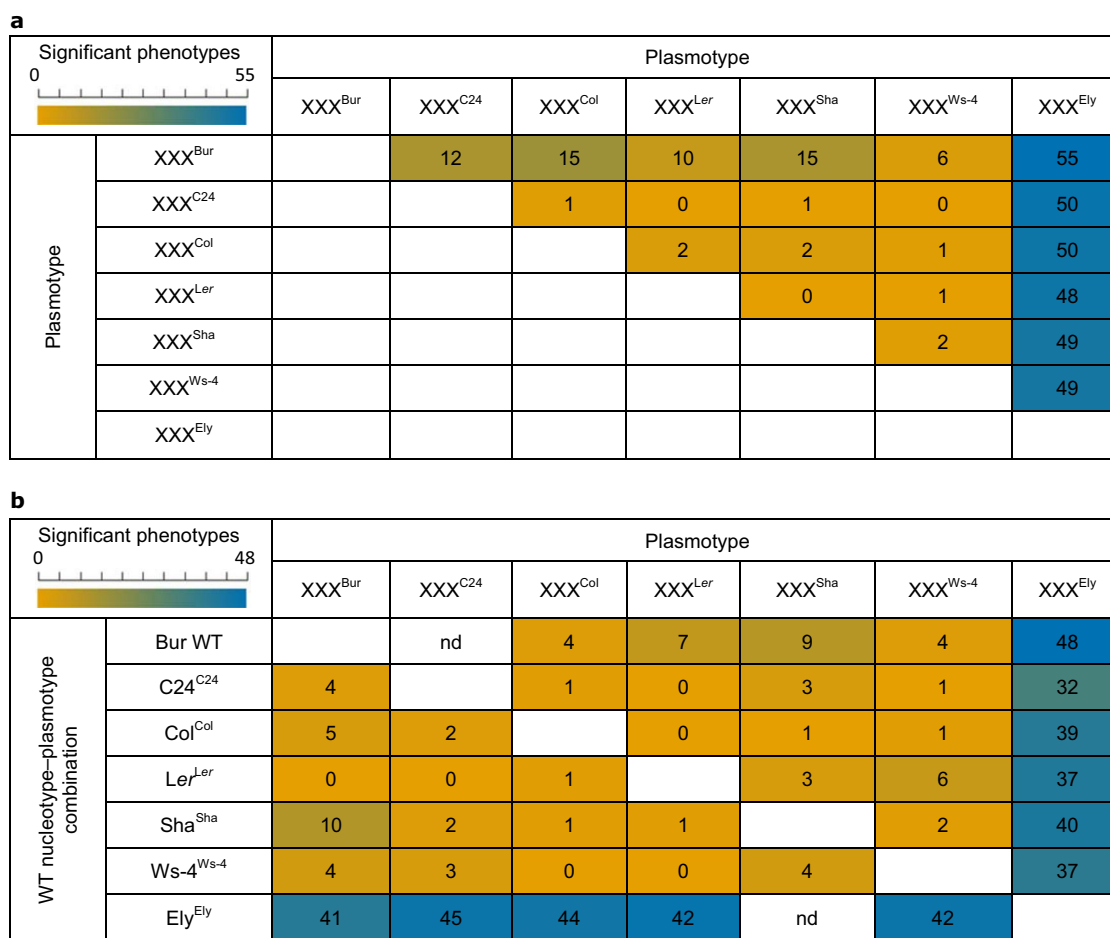


Fig. 2 | Significant plasmotype-induced effects in 92 phenotypes. a, Number of observed significant plasmotype additive effects when a specific plasmotype was changed for another, regardless of the nucleotype. Note that the replacement of Bur (top row) and Ely plasmotypes (far right column) resulted in the highest number of plasmotype additive effects. **b**, Number of observed significant epistatic effects in phenotypes between WT nucleotype-plasmotype combinations and cybrids with different plasmotypes. Rows indicate the number of significant effects when comparing self-cybrids to cybrids with identical nucleotype but non-native plasmotype. Columns indicate specific plasmotype changes. Note that changing the Ely plasmotype for another plasmotype (bottom row and far right column) resulted in many epistatic effects due to the large-effect mutation in the chloroplast-encoded *PsbA* gene of the Ely plasmotype. Similar effects, but of smaller magnitude, resulted from changing the Bur plasmotype (top row and first column). Post hoc tests were used with Hochberg's *P* value correction for **a**, and Dunnett's *P* value correction (with WT as control) for **b**. $\alpha = 0.05$; nd, not determined. For underlying *P* values and phenotypes see Supplementary Data 2. Orange cells indicate a low number of significant effects while blue cells denote a high number of significant effects.

variation caused by both plasmotype additive and epistatic effects resulted in substantial phenotypic differences.

Next we sought to assess whether there are general patterns in how specific nucleotypes and plasmotypes interact. To this end we first assessed which plasmotype changes would result in additive phenotypic changes. Plasmotype replacements involving the Ely plasmotype led to additive changes in, on average, 50 (out of 92) phenotypes across the 7 nucleotypes (Fig. 2a). Changes involving the Bur plasmotype led, on average, to 10 significant additive effects, 8 of which are photosynthesis-related (Supplementary Data 2). Other plasmotype changes showed, on average, one additive effect in predominantly non-photosynthetic phenotypes. Comparison of WT cytonuclear combinations with all their iso-nuclear cybrid lines also showed that plasmotype changes involving Ely and Bur plasmotypes show the most epistatic effects (on average 43 and 6, respectively; Fig. 2b). The number of epistatic effects resulting from the Bur plasmotype ranged between 0 (*Ler^{Ler}* versus *Ler^{Bur}*) and 10 (*Sha^{Sha}* versus *Sha^{Bur}*), indicating high variability. Plasmotype changes involving other plasmotypes showed more modest numbers of significant

epistatic effects, ranging from 0 to 6. Plasmotypes that resulted in more additive effects also caused more epistatic effects (Pearson correlation coefficient = 0.99, $P = 1.3 \times 10^{-5}$), suggesting a possible common cause (Extended Data Fig. 4).

Although the average total explained variance due to cytonuclear epistasis is only 5.2%, these interactions can have marked effects for specific phenotypes or in specific cybrids. Explained variance for certain phenotypes can be markedly higher—for example, for projected leaf area this was 12.3%, for hyponastic leaf movement 8.3% and for Φ_{NPQ} 17.8%. A strong epistatic effect in pollen abortion (43.5%) was due to relatively high pollen abortion in *Sha^{Sha}* (Fig. 3a), which we also observed in the *Sha* WT. The higher pollen abortion in its native nucleotype is surprising, and could indicate incomplete compensation due to the accumulation of deleterious variants—or perhaps to facilitate increased outcrossing. The only cybrid for which we initially failed to obtain seed was *Ely^{Sha}*. This haploid was recreated and pollinated with WT *Ely* pollen to increase the likelihood of seed set. The diploid offspring showed 45% of pollen abortion and, despite producing

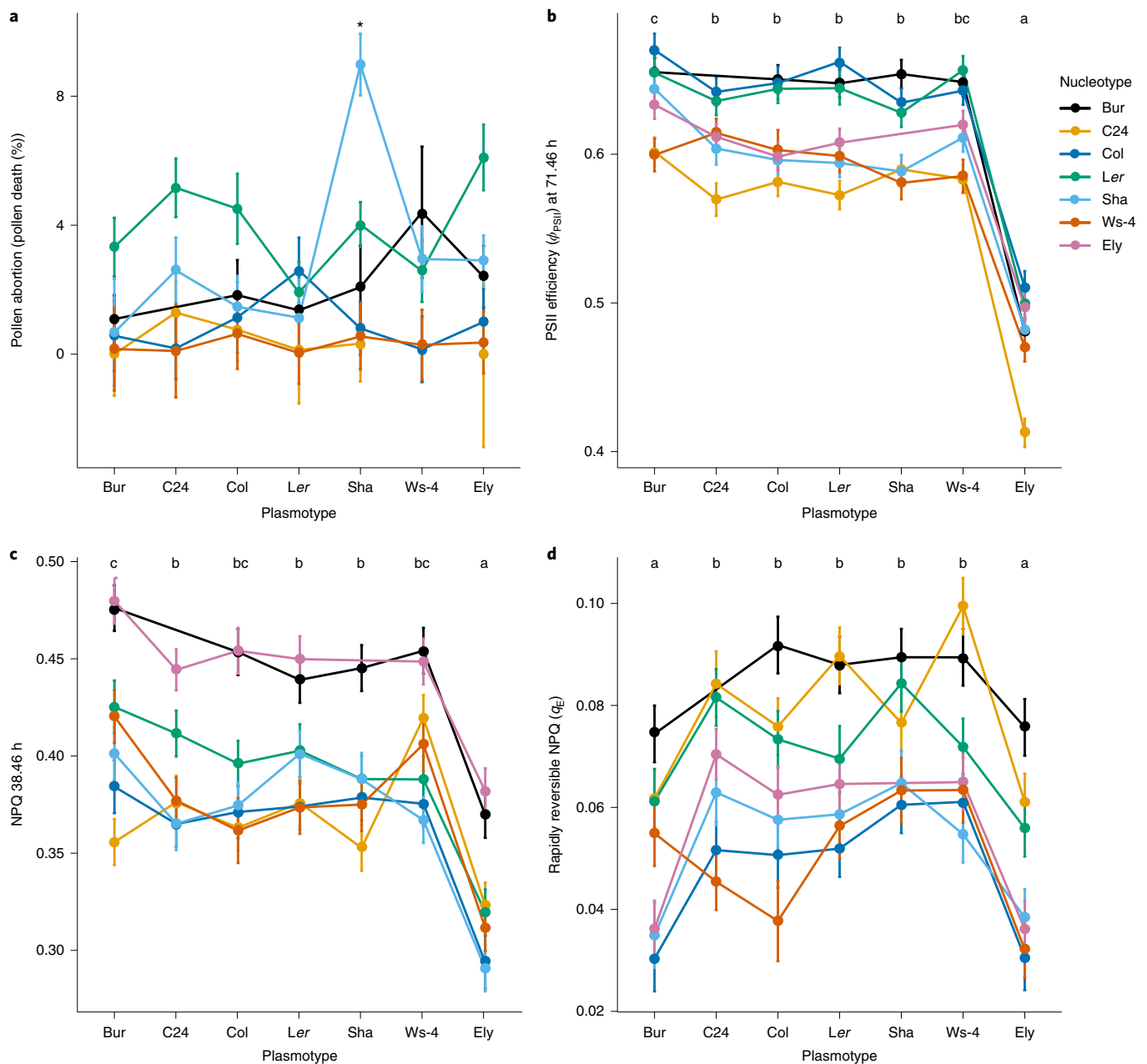


Fig. 3 | Plasmotype changes result in cytonuclear epistasis and also, in the case of cybrids with the Ely and Bur plasmotype, in additive effects. a, Pollen abortion, percentage of dead pollen out of 250. **b**, PSII efficiency (Φ_{PSII}) 71.46 h after the start of the experiment, after a full day under fluctuating light with a maximum difference between 500 and 100 $\mu\text{mol m}^{-2}\text{s}^{-1}$ (see Fig. 4c for light treatment). **c**, NPQ at 38.46 h after the start of the experiment, at 300 $\mu\text{mol m}^{-2}\text{s}^{-1}$ irradiance on a sigmoidal light curve starting at 65 $\mu\text{mol m}^{-2}\text{s}^{-1}$. **d**, The rapidly reversible component of NPQ, q_E , at 259 $\mu\text{mol m}^{-2}\text{s}^{-1}$ after a full day of fluctuating light with maximum difference between 500 and 100 $\mu\text{mol m}^{-2}\text{s}^{-1}$. Any deviation from a horizontal line represents a potential additive or epistatic effect. Error bars, s.e.m. The asterisk in **a** indicates a unique and significant difference between the Sha^{Sha} cybrid and other cybrids with Sha nucleotypes (epistasis) (Hochberg's test, n = minimum of four biologically independent plants; for precise P values, see Supplementary Data 1). The letters above **b–d** represent significant differences between plasmotypes regardless of the nucleotype (additivity) (Hochberg's test, n = 28 (4 biologically independent cybrids across 7 different nucleotypes); letters vary when $P < 0.05$). **b–d**, Plants were grown under 200 $\mu\text{mol m}^{-2}\text{s}^{-1}$ light intensity for 21 d before starting the experiment.

pollen, all plants were male sterile. This indicates that, in combination with the Ely plasmotype, the Sha plasmotype results in full cytoplasmic male sterility (Extended Data Fig. 5). In combination with the Sha plasmotype, pollen abortion across the seven nucleotypes ranged from near zero to 8.9% in the Sha^{Sha} self-cybrid and to full male sterility in Ely^{Sha}, exemplifying the degree to which epistasis can manifest itself.

Cybrids with the Ely plasmotype exhibited strong additive effects—all had a lower PSII efficiency (Φ_{PSII}) (Fig. 3b) and lower values for other photosynthesis-related phenotypes: NPQ, q_E and chlorophyll content (Fig. 3c and Supplementary Data 2). The reduced level of Φ_{PSII} is likely to be responsible for the concomitant reductions in biomass, growth rate and seed size and altered primary metabolite concentration (Supplementary Data 2). To test whether

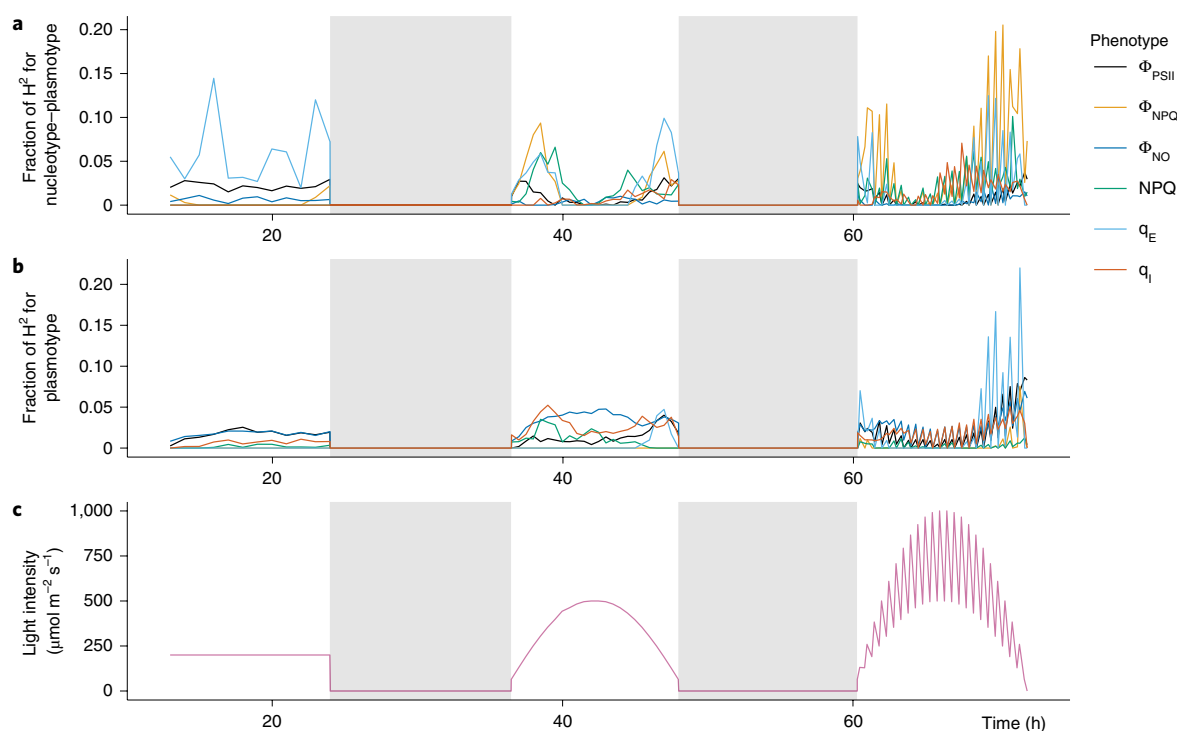


Fig. 4 | The fraction of explained genetic variation for photosynthesis phenotypes differs depending on light conditions. a, Fraction of H^2 for plasmotype epistatic effects. **b**, Fraction of H^2 for plasmotype additive effects. **c**, Light intensity for three consecutive days with growth under steady light (d1), sinusoidal light intensity (d2) and fluctuating light intensity (d3). Days are separated by nights (grey shaded areas). Note that the fraction of H^2 for different phenotypes changed markedly during d2 and 3. Some phenotypes are explained largely by additive effects (for example, q_E) and others by interaction (for example, Φ_{NPQ}). A replication of this experiment is shown in Extended Data Fig. 7.

additive effects could also be detected at the level of gene expression, we contrasted the transcriptome of Ely^{Ely} with that of hybrids Ely^{Ler} and Ely^{Bur}. We also compared the transcriptomes of Ler^{Ler}, Ler^{Bur} and Ler^{Ely} (Supplementary Data 3; for details see Extended Data Fig. 6 and Supplementary Table 5). Exchanging the Ely plasmotype with Ler or Bur, in either the Ler or Ely nuclear background, resulted in a consistent change in the expression of 40 genes, most of which were upregulated (Supplementary Table 6). A Gene Ontology-term analysis revealed that these genes are significantly enriched for those involved in photorespiration (GO:0009853) and in glycine and serine family amino acid metabolism (GO:0006544 and GO:0009069, respectively; Supplementary Data 3). This is in line with the low serine and glycine content of hybrids with the Ely plasmotype, which suggests reduced photorespiration (Supplementary Data 2)²³ and can be linked to lower overall photosynthetic activity.

The Ely plasmotype was deliberately included in our panel for its strong additive effect. In addition to Ely, we also observed strong additive effects from the Bur plasmotype, which are mainly restricted to photosynthetic parameters. Under normal conditions PSII efficiency is slightly increased by the Bur plasmotype (1.6%) but, under fluctuating light intensity this difference became more apparent (3.5% increase; Figs. 3b and 4 and Extended Data Fig. 7). This increase in Φ_{PSII} under fluctuating conditions resulted in a corresponding reduction in Φ_{NO} and Φ_{NPQ} of 7.3 and 2.2%, respectively. NPQ, q_E and q_I were also influenced by plasmotype, but the time points at which these differences occurred differed per phenotype (Fig. 4a,b). The Bur plasmotype increased NPQ, with the highest increase of 5.9% at the beginning of day 2 (38.46 h) (Fig. 3c), while the rapidly reversible component of NPQ, q_E , had a maximum reduction of 26.6% at the end of d3 (71.46 h) (Fig. 3d).

These photosynthesis-related phenotypes are likely to be due to chloroplast-derived variation. In support of a chloroplast origin

for this photosynthetic variation, measurements of mitochondrial respiration suggest that Bur is not an outlier and shows standard respiration rates (Extended Data Fig. 8). Based on DNA sequence coverage plots there are no obvious duplications or deletions in the mitochondrial or chloroplast sequence of Bur, thus we expect that altered expression or protein activity, as opposed to gene gain or loss, is driving Bur-derived phenotypes (Extended Data Fig. 9). We annotated the sequence variation of all plasmotypes using SnpEff²⁴. From this we found no large-effect mutations in the Bur mitochondria. There were, however, unique missense variants in the chloroplastic genes *MATURASE K (MATK)*, *NAD(P)H-QUINONE OXIDOREDUCTASE SUBUNIT 6 (NDHG)* and the chloroplast open reading frame 1 (*YCF1*), as well as a frameshift mutation in transfer RNA-lysine (*TRNK*) (Supplementary Data 4). Of these, *NDHG* is noteworthy because of its functions. It is part of the NAD(P) H-dehydrogenase-like complex (NDH) that is located within the thylakoid membrane and acts, amongst others, as a proton pump in cyclic electron flow around photosystem I and chlororespiration. NDH creates a pH differential that generates non-photochemical quenching^{25,26}. However, determination of whether the missense mutation in NDH underlies the observed phenotypic changes in the photosynthetic parameters would require further experimentation. In contrast to Ely, the plasmotype that spread in response to the use of herbicides, an anthropogenic selective pressure⁵, the Bur plasmotype represents a naturally occurring plasmotype that has an additive impact on key photosynthetic phenotypes.

Our experiments have shown that a clean, systematic exploration of plasmotypic variation in a plant species is feasible. To our knowledge, apart from the *cenh3* mutant used here there is only one other intraspecific haploid inducer available (the maize *ig* mutant) that can be used via the maternal line and thus replace the plasmotype^{27–29}. Current knowledge of *cenh3*-mediated uniparental

genome elimination should allow for the creation of maternal haploid inducers in a wider range of species³⁰. This would allow elite nucleotypes to be brought into new plasmotypic backgrounds to explore novel plasmotype–nucleotype combinations. Our data indicate that there is substantial variation for phenotypes such as NPQ and Φ_{PSII} that are important for plant productivity^{31–33}. Next to Ely we identified one new plasmotype (Bur) that significantly impacts photosynthesis in an additive manner. Expanding our panel would probably find further examples, suggesting that future research aiming to enhance crop photosynthesis should pay close attention to plasmotypic variation. Apart from studying natural variation, the use of haploid inducers as plasmotype donors could be used to transfer cytoplasmic male sterility, herbicide resistance or genetically engineered plasmotypes. Plant plasmotypes are notoriously difficult to modify genetically, although recently there have been some advances in this regard^{34–37}. The use of haploid inducers as plasmotype donors could further increase the accessibility of such modifications. Transformations could be undertaken in a compatible nucleotype and, once achieved, can be transferred to different nucleotypes thus amplifying the potential impact of successful plasmotype modifications.

Exploring the potential of plasmotypic variation via the use of haploid-inducer lines is not only promising for plant breeding, but also for understanding the role played by such variation in plant adaptation^{13,14}. Our results show that, despite considerable genetic divergence between the genotypes used in our panel, all cybrids were viable, and this in itself suggests a remarkable degree of conservation for the fundamental components of cytonuclear interactions. Although we did find clear additive effects of certain plasmotypes, the majority of plasmotype-derived variation manifested as epistasis in the traits we measured, which is in line with previous research in plants, animals and fungi^{6–8,11,18}. Also in line with studies of mitochondrial interactions in animals is the observation that phenotypic variation due to plasmotypic variation becomes more pronounced under fluctuating and stressful conditions^{18,38–40}. Both our results and previous work suggest that multi-level interactions (that is, nucleotype–plasmotype–environment) may be the primary mechanism by which plasmotypic variation is expressed. Thus, plasmotypic variation may act as an evolutionary capacitor providing novel phenotypes in specific genetic and environmental contexts. In our rapidly changing climate, such variation may be particularly important for both crops and wild species. To fully understand the impact and functional relevance of plasmotypic variation, future studies should expand both the number of plasmotypes and the range of environmental conditions assayed. The speed and precision with which new cybrids can be created make such research feasible.

Methods

Plant materials. Seven *Arabidopsis* accessions were chosen for the construction of a full nucleotype–plasmotype diallel. Ely (CS28631) is atrazine resistant due to a chloroplast-encoded mutation in *PsbA*, which leads to a modified D2 protein that greatly reduces PSII efficiency²⁰. Ws-4 (CS5390) was included for its unusual PSII phosphorylation dynamics⁴¹. Bur (CS76105) is commonly used in diversity panels and is a standard reference accession. Sha (CS76227) was selected based on its capacity to induce cytoplasmic male sterility in certain crosses⁴². The set was completed by the addition of *Ler* (CS76164), *Col* (CS76113) and *C24* (CS76106), three genotypes widely used in *Arabidopsis* research. *Col* is the reference for nuclear, mitochondrial and chloroplast sequences and at the start of this project *C24* was the reference for the mitochondrial sequence, hence its inclusion. The *GFP-tailswap* haploid inducer, which expresses a green fluorescent protein (GFP)-tagged CENTROMERE HISTONE 3 protein in a *cenH3/htr12* mutant background, is in a *Col* background¹⁹.

Generation of a nucleotype–plasmotype diallel. To generate new nucleotype–plasmotype combinations, plants of all seven accessions (Bur, *C24*, *Col*, Ely, *Ler*, Sha and Ws-4) were crossed as males to *GFP-tailswap*, resulting in all cybrids having the *Col* plasmotype. New HI lines were created by crossing the original *GFP-tailswap* line as a male to the six additional plasmotype mothers (Bur, *C24*, Ely, *Ler*, Sha and Ws-4). Genome elimination did not always occur, and some of the offspring were diploid F1 lines. These were selfed and F2 lines homozygous for

the *cenH3/htr12* mutation and carrying the *GFP-tailswap* were selected as new HI lines in different plasmotypic backgrounds (Fig. 1a). Plants of all seven accessions were then crossed as males to these new HI lines, and the haploids arising from these 49 crosses were identified based on their phenotype (as described in ref. 43). These haploid lines self-fertilized, following either somatic genome duplication or restitutional meiosis⁴⁹, and gave rise to doubled haploid offspring (Fig. 1b). The resulting 49 lines comprise a full diallel of 21 pairs of reciprocal nucleotype–plasmotype combinations (cybrids), as well as seven nucleotype–plasmotype combinations that have, in principle, the same nucleotype–plasmotype combinations as their WT progenitors (self-cybrids; Fig. 1c, diagonal line). All cybrids and WT accessions were propagated for one generation before use in further experiments, with the exception of Ely^{Sha} of which the original haploid died without setting seed and was recreated at a later stage by generating haploids that were pollinated with Ely WT plants to ensure seed set.

Genotype confirmation. To confirm that all cybrids in our panel are authentic, all 49 cybrids and 7 WT progenitors were whole-genome sequenced at the Max Planck Genome Centre Cologne (Germany) using Illumina HiSeq 2500 150-base pair (bp) paired-end sequencing. The cybrids were sequenced at 8.5× coverage, and the WT progenitors at 40× coverage. To remove erroneous bases, we performed adaptor and quality trimming using Cutadapt (v.1.18)⁵⁴. Sequences were clipped if they matched at least 90% of the total length of one of the adaptor sequences provided in the instruction manual of NEBNext Multiplex Oligos for Illumina (Index Primers Set 1). In addition, we trimmed bases from the 5' and 3' ends of reads if they had a phred score of 20 or lower. Reads <70 bp after trimming were discarded. Trimmed reads were aligned to a modified version of the Col-0 reference genome (TAIR10, European Nucleotide Accession no. GCA_000001735.2), which contains an improved assembly of the mitochondrial sequence (Genbank accession no. BK010421)⁴⁵ using bwa mem (v.0.7.10-r789)⁴⁶ with default parameters. The resulting alignment files were sorted and indexed using samtools (v.1.3.1)⁴⁷. Duplicate read pairs were marked using the MarkDuplicates tool of the GATK suite (v.4.0.2.1), using an optical duplicate pixel distance of 100 as recommended in the documentation of GATK when working with data from unpatterned Illumina flowcells. Variants were called using a workflow based on GATK Best Practices. Base quality scores of aligned reads were recalibrated using GATK BaseRecalibrator with default parameters, using a set of variants of a worldwide panel of 1,135 *Arabidopsis* accessions⁴⁸ (obtained from ftp://ftp.ensemblgenomes.org/pub/plants/release-37/vcf/arabidopsis_thaliana/) as known sites. Following base recalibration, variants were called in each sample using GATK HaplotypeCaller, allowing for a maximum of three alternative alleles at each site. Samples were then jointly genotyped using GATK GenomicsDBImport and GATK GenotypeGVCFs, with default parameters. This last step generated three different variant call format (VCF) files: one containing the calls of the nuclear genome, one containing calls of the mitochondrial genome and one containing calls of the chloroplast genome.

To remove probable false-positive calls, we filtered the call sets using two complementary approaches. First, we filtered the nuclear call set using GATK VariantRecalibrator and GATK ApplyVQSr (-truth-sensitivity-filter-level set at 99.9), using the set of variants called in the worldwide panel of 1,135 *Arabidopsis* accessions as a training and truth set (prior = 10.0). This step could not be performed for mitochondrial and chloroplast calls as these lack a golden truth set that can be used for recalibration. Second, we filtered variants based on their quality-by-depth score (QD). For the nuclear call set we used a QD score of 40, leaving 3.7 million single-nucleotide polymorphisms (SNPs), for the chloroplast call set a QD of 25, leaving 356 SNPs and for the mitochondrial call set a QD of 20, leaving 135 SNPs.

Forty-six cybrids were found to have the correct genotypes. With one line, Bur^{Ws-4}, there was a sample mix-up during library preparation with Sha^{Sha}, leading to two sequenced Sha^{Sha} samples and no sequenced Bur^{Ws-4} sample. Fortunately we did have a true Bur^{Ws-4} cybrid, which we confirmed based on both phenotype (Bur and Sha nucleotypes are phenotypically distinct from one another) and genotype, using KASP markers (see below) (Supplementary Table 8). Two other lines, *C24*^{*C24*} and Ws-4^{*Col*}, had a high number of heterozygous calls in their plasmotypes, with *C24*^{*C24*} being heterozygous with *C24*^{*Col*} and Ws-4^{*Col*} being heterozygous with Ws-4^{*Bur*}. Because in the creation of Ws-4^{*Col*} no plant was used with a Bur plasmotype, this suggested a sample mix-up. To confirm this, and that the putative event of cross-contamination had occurred in the laboratory, we designed seven KASP markers (LGC, <https://www.lgcgroup.com/>) to genotype all cybrids. These KASP markers are designed to be unique for the chloroplast allele of one accession, and target SNPs that were called as heterozygous in the sequence analysis (Supplementary Table 7). The KASP assay can distinguish between homozygous and heterozygous states. We assayed all seven KASP markers on all cybrids; for *C24*^{*C24*} and Ws-4^{*Col*} this included plants from the same seed batch as the plants used for sequencing, as well as direct offspring of the sequenced plants. All lines proved to be the genotypes predicted and no chloroplast heterozygosity was observed in any of the lines, including *C24*^{*C24*} and Ws-4^{*Col*} (Supplementary Table 8). Unfortunately, the Ely^{Sha} used for sequencing died before setting seed and, although it has since been recreated, it could not be included in our phenotypic analyses. We used the KASP marker for the Sha chloroplast, and confirmed it to be correct (Supplementary Table 8).

To check for any incomplete chromosome eliminations we calculated the read coverage for all cybrids, normalized per chromosome. We did not observe any remaining chromosomes, but we did detect a 200-kb duplication of chromosome 2 nuclear DNA in Bur^{C24} and the self-cybrid Bur^{Bur}. Because the same duplicated segment is present in two independent cybrid lines and is a duplication of Bur nuclear origin (based on sequence identity), we conclude that this segment resulted from a de novo duplication in one of the WT Bur lines used to generate these two cybrids. Following the exclusion of phenotyping data for Bur^{Bur} and Bur^{C24}, we limited our analyses to 46 rather than 49 cybrids. The parental lines were included in the screens to test for possible unforeseen effects of cybrid production (which involves a haploid growth stage). This brought the total number of phenotyped lines in this study to 53 (40 cybrids, 6 self-cybrids and 7 WT).

The functional effects of the chloroplastic and mitochondrial SNPs and INDELS were predicted using SnpEff²⁴. A SnpEff database was built using the genome, transcriptome and proteome as released in TAIR10.1. SNPs and INDELS were predicted on the filtered VCF, as mentioned above. In the analysis we considered only those variants with a 'HIGH' or 'MODERATE' impact.

Phenotyping. Cybrids were phenotypically assessed using different platforms. For details on the number of phenotypes per experiment, see Supplementary Table 4.

Growth, PSII efficiency (Φ_{PSII}), chlorophyll reflectance and leaf movement (all parameters at $n=24$) were screened in the Phenovator platform, a high-throughput phenotyping facility located in a climate-controlled growth chamber⁴⁹. This phenotyping platform measured the plants for: Φ_{PSII} using chlorophyll fluorescence, reflectance at 480, 532, 550, 570, 660, 700, 750 and 790 nm, and projected leaf area based on pixel counts of near-infrared images⁴⁹. The growth chamber was set to a 10/14 h day/night regime (at 20 and 18 °C, respectively), 200 $\mu\text{mol m}^{-2} \text{s}^{-1}$ irradiance and 70% relative humidity. The plants were grown on a rockwool substrate and irrigated daily with a nutrient solution as described in ref. ⁴⁹.

Growth ($n=24$) and, subsequently, above-ground biomass ($n=12$) were measured in another high-throughput phenotyping facility⁵⁰ where projected leaf area was measured three times per day with 14 fixed cameras (uEye Camera, IDS Imaging Development Systems). This growth chamber was set to a 10/14 h day/night regime (20 and 18 °C, respectively), 200 $\mu\text{mol m}^{-2} \text{s}^{-1}$ light and 70% relative humidity. Plants were grown on rockwool and irrigated weekly with a nutrient solution as described previously.

Non-fluctuating and fluctuating light treatments were performed in the dynamic environmental phenotype imager (DEPI) facility of Michigan State University ($n=4$)⁵¹. This facility is able to measure the chlorophyll fluorescence-derived photosynthetic parameters Φ_{PSII} , Φ_{NO} , Φ_{NPQ} , NPQ, q_E and q_I . Three-week-old plants were moved into the facility, where they were left to acclimatize for 24 h after which 3 d of phenotyping was performed under different light regimes. On the first day the plants were illuminated under a constant light intensity of 200 $\mu\text{mol m}^{-2} \text{s}^{-1}$. On the second day the plants were subjected to sinusoidal light treatment where the light intensity was initially low and was gradually increased to a maximum of 500 $\mu\text{mol m}^{-2} \text{s}^{-1}$, from which it was decreased back to 0. On the third day the plants underwent fluctuating light treatment ranging between 0 and 1,000 $\mu\text{mol m}^{-2} \text{s}^{-1}$ at short intervals (Fig. 4c). For the second experiment in the DEPI phenotyping facility, the experiment was extended by 2 d in which d4 replicated d2 and d5 replicated d2 (Supplementary Data 1 and Extended Data Fig. 7c). For further details, see ref. ⁵¹.

Bolting time and flowering time were measured for all cybrids ($n=10$) in a greenhouse experiment conducted in April 2017, with the exception of Ely nucleotype cybrids which needed vernalization and were not included in this experiment. Additional lighting was provided when natural light intensity fell below 685.5 $\mu\text{mol m}^{-2} \text{s}^{-1}$, and stopped when it reached 1,142.5 $\mu\text{mol m}^{-2} \text{s}^{-1}$, with a maximum of 16 h per day.

Seeds for the germination experiments were generated from two rounds of propagation. In the first round, seeds were sown in a growth chamber set to a regime of 10/14 h day/night (20 and 18 °C, respectively), 200 $\mu\text{mol m}^{-2} \text{s}^{-1}$ light intensity and 70% relative humidity. After 3 weeks they were moved to an illuminated cold room at 4 °C for 6 weeks of vernalization. After vernalization all plants ($n=8$) were moved to a temperature-controlled greenhouse (20 °C) for flowering and seed ripening. Exceptions to this were *Ler*^{Ely}, *Ler*^{W5-4}, and *Ely*^{W5-4}, for which no doubled haploid seed was available at the beginning of the first propagation round. *Ler*^{Ely} and *Ler*^{W5-4} were sown later, during the vernalization stage, and flowered at the same time as the vernalized plants. *Ely*^{W5-4} produced haploid seed at a later stage and could not be included in the first propagation round. Plants were grown in a temperature-controlled greenhouse set at 20 °C. In this round, only lines with the Ely nucleotype were vernalized. For the germination experiments, seeds were stratified on wet filter paper for 4 d at 4 °C before assay in the Germinator platform⁵² for seed size, germination rate and total germination percentage. Germination under osmotic stress was performed on filter paper with 125 mM NaCl. For the controlled deterioration treatment, seeds were incubated for 2.5, 5 or 7 d at 40 °C and 82% relative humidity, and subsequently assayed in the Germinator platform without stratification.

To assess pollen abortion, all cybrid lines and WT progenitors (except those with the Ely nucleotype) were grown simultaneously in a growth chamber (Percival) under controlled conditions (16/8 h light/dark cycle, 21/18 °C and

50–60% relative humidity). Pollen abortion was manually assessed for all genotypes using differential staining of aborted and non-aborted pollen grains⁵³. A total of three plants and three flowers per plant of each cybrid were collected on the same day, and submerged in a 13- μl drop of phenol-free Alexander staining solution placed on a glass slide with a glass coverslip (18 \times 18 mm). For each flower, 250 pollen grains were counted and the number of aborted pollen grains therein.

Oxygen consumption of seedlings was measured in 2 ml of deionized water with a liquid-phase Oxytherm oxygen electrode system (Hansatech Instruments) calibrated at the measurement temperature. Three-day-old seedlings (about 50 mg) were directly imbibed in the electrode chamber. Rates of oxygen consumption were measured after tissue addition and subtracted from the rates after the addition of 500 μM KCN. Results are the mean of at least five measurements. Measurements for different genotypes were performed on consecutive days and, to correct for daily variation, normalized to Col-0 samples that were run daily.

Metabolomics. Plant material for primary metabolite analysis was obtained from the Phenovator photosynthetic phenotyping experiment. Plants were harvested 26 d after sowing which, due to the 10-h photoperiod, was before bolting for all lines. Samples were frozen in liquid nitrogen, and samples of each genotype were subsequently combined into four pools each comprising material from approximately six replicates. Each pool was ground and homogenized before an aliquot was taken for further analysis. Reference samples for metabolic analysis were composed of material from all seven parents in equal amounts, and were then homogenized. The method used for the extraction of polar metabolites from *Arabidopsis* leaves was adapted from ref. ⁵⁴, as described in ref. ⁵⁵. Specific adjustments for *Arabidopsis* samples were made as follows; polar metabolite fractions were extracted from 100 mg of *Arabidopsis* leaf material (fresh weight, with maximum 5% deviation). After the extraction procedure, 100- μl aliquots of the polar phase were dried by vacuum centrifugation for 16 h. Derivatization was performed online, similarly to the method described in ref. ⁵⁴, and the derivatized samples were analysed by a gas chromatography–time-of-flight–mass spectrometry system comprising an Optic 3 high-performance injector (ATAS, GL Sciences) and an Agilent 6890 gas chromatograph (Agilent Technologies) coupled to a Pegasus III time-of-flight mass spectrometer (Leco Instruments). Two microlitres of each sample was introduced into the injector at 70 °C using 5% of the sample (split 20). The detector voltage was set to 1,750 V. All samples were analysed in random order in four separate batches. The systematic variation inadvertently introduced by working in batches was removed following analysis of covariance. In this model the batch number was used as a factor (four levels) and 'run number within a batch' as a covariate, since it is also expected that (some) variation will be introduced by the sample run order within each batch. For this the S2 method described in ref. ⁵⁶ was used to perform least-squares regression. After quality control and removal of metabolites with >20% missing data and $H^2 < 5\%$, we were left with data on 36 primary metabolites. Metabolites were identified based on the Level of Identification Standard of the Metabolomics Standards Initiative⁵⁷.

Transcriptome analysis. Using the same material described in Metabolomics, total RNA was extracted from six cybrids, three in a *Ler* and three in an Ely nuclear background: *Ler*^{Ler}, *Ler*^{Ely}, *Ler*^{Bur} and *Ely*^{Ler}, *Ely*^{Ely}, *Ely*^{Bur}, with three replicates per genotype, totalling 18 plants. Library preparation was done with a selection on 3'-polyadenylated tails to preferentially include nuclear messenger RNA. Read alignment was performed using TopHat⁵⁸. Any chloroplast and mitochondrial genes remaining were excluded from further analysis. The raw counts were normalized and analysed using the DESeq2 package in R⁵⁹. Genes for which the expression levels were significantly different between two cybrids were determined by comparison of two genotypes using the contrast function of DESeq2. *P* values were determined using the Wald test, and adjusted with Benjamini–Hochberg correction ($\alpha=0.05$). Gene Ontology enrichment analysis was performed using the default setting in g:profiler. The complete set of detected genes in each cybrid was used as a statistical background in the analysis⁶⁰.

Phenotypic data analysis. We used the self-cybrids as our baseline in phenotypic comparisons to control for any possible effects of cybrid creation, with the exception of Bur^{Bur} which was replaced in all analyses with Bur-WT. Raw data were directly analysed, except for time series data of growth and chlorophyll reflectance which were preprocessed as follows. Time series data were fitted with a smooth spline using the gam function from the mgcv package in R⁶¹. The fitted B-spline was subsequently used to derive curve parameters. These include area under the curve, slope under mean, first, second (median) and third quartiles, minimal and maximal slopes and the time point where the slope is at maximum. These parameters allow us to quantify not only plant size and growth rate, but also the dynamic properties of the growth curve—that is, growth occurring early, late or constantly over time. In addition, we calculated relative growth rate per time point by dividing growth rate relative to plant size⁴⁹. All raw and derived parameters were analysed by fitting either a linear mixed model or a linear model. The former was used when a random correction parameter was present and, when such random correction parameters were absent, a linear model was used. The models were analysed using the restricted maximum likelihood procedure for each relevant phenotype with the lme4 package in R⁶². Because each experiment had a different

design, several models were employed (Supplementary Table 4). The following model was generally used and, in some instances, random terms (underlined in equations (1 and 2)) were added:

$$Y = \text{Nucleotype} + \text{Plasmotype} + (\text{Nucleotype} \times \text{Plasmotype}) + \text{Block} + \varepsilon \quad (1)$$

For every model, normality and equal variances were checked. Next, for every phenotypic parameter we calculated significant difference for the plasmotype and interaction term of the model (equation (1)). This was done by analysis of variance in which Kenward–Roger approximation for degrees of freedom was used. As post hoc tests we used a two-sided Dunnett's test, where we tested whether a given cybrid was different from the self-cybrid control within one nucleotype. Two-sided Hochberg's post hoc tests were used when all pairwise comparisons were tested within one nucleotype (to test for epistasis) and across all nucleotypes (to test for additivity). The significance threshold for all post hoc tests was set at $\alpha = 0.05$. The contributions of nucleotype, plasmotype and interaction between the two were determined by estimation of the variance components in mixed models containing the same terms as in equation (1). However, the fixed terms were taken as random:

$$Y = \text{Nucleotype} + \text{Plasmotype} + (\text{Nucleotype} \times \text{Plasmotype}) + \text{Block} + \varepsilon$$

In this model the variance components were estimated by the VarCorr function from the lme4 package. Total variance was calculated by summing all the variance components, after which the fraction of explained variance for every term in the model was calculated. The broad-sense heritability (in our case, equal to repeatability²³) was determined by the three genetic components—nucleotype, plasmotype and their interaction. The fraction of broad-sense heritability explained by the separate genetic components was calculated subsequently.

In total we measured 1,859 phenotypes. After data processing, further analysis was conducted only on phenotypes with $H^2 > 5\%$, removing those that were non-informative, leaving us with a total of 1,782. Furthermore, to avoid bias in the results due to overly correlated data when stating summary statistics, we further subset the remaining 1,782 phenotypes (Supplementary Data 2). Using a threshold based purely on correlation would favour the inclusion of variation driven largely by nucleotype. Since the population is balanced, we subtracted the averages of the nucleotype values from the cybrid phenotype value to reveal the plasmotype effect per cybrid. From these we calculated the Pearson correlations for all phenotypes. This highlighted that the most uncorrelated phenotypes arose mainly from one experiment assessing photosynthetic parameters under fluctuating light. The unbiased selection of a subset of phenotypes would have resulted in the omission of several phenotypic categories. To present a balanced overview of all phenotypic categories, we manually selected a subset comprising the following phenotypes. For time series in which we scored for up to 25 d after germination, we selected the morning measurements from days 8, 13, 18 and 23. The time series analysis of fluctuating light was measured sequentially for 3 (first experiment, Fig. 4) and 5 d (replicate experiment, Extended Data Fig. 7), with plants each day subjected to a different treatment. Because these treatments reached their extremes at the middle and end of the day, and the results of replicate experiments were very similar, we selected time points at the middle and end of the day from only the first experiment. For the different seed treatments, we used the germination time until 50% of the seeds had germinated. In addition, we included biomass, leaf movement, seed size and flowering time as single phenotypes, and all 36 primary metabolites. This resulted in 92 phenotypes that were used to provide summary and test statistics (for a correlation plot of these, see Extended Data Fig. 3). All data on the 1,859 phenotypes, including summary and test statistics, are available in Supplementary Data 1 and Supplementary Table 3.

The correlation between plasmotype additive and epistatic effects was calculated with and without the Ely plasmotype. For both additive and epistatic effects, every significant change between plasmotypes, within one nucleotype background, was counted (Supplementary Data 2). Pearson correlation coefficients and accompanying *P* values were calculated using the ggpubr package in R.

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

Data availability

Sequencing and transcriptome data are available through the European Nucleotide Archive with the primary accession codes PRJEB29654 and PRJEB35324. The raw datasets are available through Dryad at <https://doi.org/10.5061/dryad.cz8w9g05>. The analysed datasets that support our findings are available as Supplementary Data. The associated raw data for Figs. 3 and 4 are provided in Supplementary Data 1, and the raw data for Fig. 2 are provided in Supplementary Data 2. Source data for Figs. 3 and Fig. 4 and Extended Data Figs. 1, 3, 4 and 7–9 are provided with the paper. The germplasm generated in this project will be available via the European Arabidopsis Stock Centre (www.arabidopsis.info).

Received: 31 October 2018; Accepted: 25 November 2019;
Published online: 13 January 2020

References

- Ravi, M. et al. A haploid genetics toolbox for *Arabidopsis thaliana*. *Nat. Commun.* **5**, 5334 (2014).
- Chan, K. X., Phua, S. Y., Crisp, P., McQuinn, R. & Pogson, B. J. Learning the languages of the chloroplast: retrograde signaling and beyond. *Annu. Rev. Plant Biol.* **67**, 25–53 (2016).
- Petrillo, E. et al. A chloroplast retrograde signal regulates nuclear alternative splicing. *Science* **344**, 427–430 (2014).
- Kleine, T. & Leister, D. Retrograde signaling: organelles go networking. *Biochim. Biophys. Acta Bioenerg.* **1857**, 1313–1325 (2016).
- Flood, P. J. et al. Whole-genome hitchhiking on an organelle mutation. *Curr. Biol.* **26**, 1306–1311 (2016).
- Joseph, B., Corwin, J. A., Li, B., Atwell, S. & Kliebenstein, D. J. Cytoplasmic genetic variation and extensive cytonuclear interactions influence natural variation in the metabolome. *eLife* **2**, e00776 (2013).
- Zeyl, C., Andreson, B. & Weninck, E. Nuclear-mitochondrial epistasis for fitness in *Saccharomyces cerevisiae*. *Evolution* **64**, 910–914 (2010).
- Montooth, K. L., Meiklejohn, C. D., Abt, D. N. & Rand, D. M. Mitochondrial-nuclear epistasis affects fitness within species but does not contribute to fixed incompatibilities between species of *Drosophila*. *Evolution* **64**, 3364–3379 (2010).
- Joseph, B. et al. Hierarchical nuclear and cytoplasmic genetic architectures for plant growth and defense within *Arabidopsis*. *Plant Cell* **25**, 1929–1945 (2013).
- Tang, Z. et al. Potential involvement of maternal cytoplasm in the regulation of flowering time via interaction with nuclear genes in maize. *Crop Science* **54**, 544–553 (2014).
- Roux, F. et al. Cytonuclear interactions affect adaptive traits of the annual plant *Arabidopsis thaliana* in the field. *Proc. Natl Acad. Sci. USA* **113**, 3687–3692 (2016).
- Mossman, J. A., Ge, J. Y., Navarro, F. & Rand, D. M. Mitochondrial DNA fitness depends on nuclear genetic background in *Drosophila*. *G3 (Bethesda)* **9**, 1175–1188 (2019).
- Dobler, R., Rogell, B., Budar, F. & Dowling, D. K. A meta-analysis of the strength and nature of cytoplasmic genetic effects. *J. Evol. Biol.* **27**, 2021–2034 (2014).
- Bock, D. G., Andrew, R. L. & Rieseberg, L. H. On the adaptive value of cytoplasmic genomes in plants. *Mol. Ecol.* **23**, 4899–4911 (2014).
- Levings, C. S. The Texas cytoplasm of maize: cytoplasmic male sterility and disease susceptibility. *Science* **250**, 942–947 (1990).
- Miclaus, M. et al. Maize cytolines unmask key nuclear genes that are under the control of retrograde signaling pathways in plants. *Genome Biol. Evol.* **8**, 3256–3270 (2016).
- Sambatti, J. B., Ortiz-Barrientos, D., Baack, E. J. & Rieseberg, L. H. Ecological selection maintains cytonuclear incompatibilities in hybridizing sunflowers. *Ecol. Lett.* **11**, 1082–1091 (2008).
- Dowling, D. K., Abiega, K. C. & Arnqvist, G. Temperature-specific outcomes of cytoplasmic-nuclear interactions on egg-to-adult development time in seed beetles. *Evolution* **61**, 194–201 (2007).
- Ravi, M. & Chan, S. W. L. Haploid plants produced by centromere-mediated genome elimination. *Nature* **464**, 615–618 (2010).
- El-Lithy, M. E. et al. Altered photosynthetic performance of a natural *Arabidopsis* accession is associated with atrazine resistance. *J. Exp. Bot.* **56**, 1625–1634 (2005).
- Flood, P. J. et al. Natural variation in phosphorylation of photosystem II proteins in *Arabidopsis thaliana*: is it caused by genetic variation in the STN kinases? *Philos. Trans. R. Soc. B* **369**, 20130499 (2014).
- Falconer, D. & Mackay, T. J. H. *Introduction to Quantitative Genetics* (Longmans Green, 1996).
- Somerville, C. R. & Ogren, W. L. Photorespiration mutants of *Arabidopsis thaliana* deficient in serine-glyoxylate aminotransferase activity. *Proc. Natl Acad. Sci. USA* **77**, 2684–2687 (1980).
- Cingolani, P. et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly* **6**, 80–92 (2012).
- Strand, D. D., Nicholas, F. & Kramer, D. M. The higher plant plastid NAD(P) H dehydrogenase-like complex (NDH) is a high efficiency proton pump that increases ATP production by cyclic electron flow. *Journal of Biological Chemistry* **292**, 11850–11860 (2017).
- Laughlin, T. G. et al. Structure of the complex I-like molecule NDH of oxygenic photosynthesis. *Nature* **566**, 411–414 (2019).
- Kermicle, J. L. Androgenesis conditioned by a mutation in maize. *Science* **166**, 1422–1424 (1969).
- Schneerman, M., Charbonneau, M. & Weber, D. A survey of ig-containing materials. *Maize Genet. Coop. Newsl.* **74**, 92–93 (2000).
- Houben, A., Sanej, M. & Pickering, R. Barley doubled-haploid production by uniparental chromosome elimination. *Plant Cell Tissue Organ Cult.* **104**, 321–327 (2010).
- Karimi-Ashtiyani, R. et al. Point mutation impairs centromeric CENH3 loading and induces haploid plants. *Proc. Natl Acad. Sci. USA* **112**, 11211–11216 (2015).

31. Kromdijk, J. et al. Improving photosynthesis and crop productivity by accelerating recovery from photoprotection. *Science* **354**, 857–861 (2016).
32. Flood, P. J., Harbinson, J. & Aarts, M. G. M. Natural genetic variation in plant photosynthesis. *Trends Plant Sci.* **16**, 327–335 (2011).
33. Murchie, E. H. et al. Measuring the dynamic photosynthome. *Ann. Bot.-London* **122**, 207–220 (2018).
34. Ruf, S. et al. High-efficiency generation of fertile transplastomic *Arabidopsis* plants. *Nat. Plants* **5**, 282–289 (2019).
35. Kwak, S.-Y. et al. Chloroplast-selective gene delivery and expression in planta using chitosan-complexed single-walled carbon nanotube carriers. *Nat. Nanotechnol.* **14**, 447–455 (2019).
36. Zhang, J. et al. Full crop protection from an insect pest by expression of long double-stranded RNAs in plastids. *Science* **347**, 991–994 (2015).
37. Jin, S. & Daniell, H. The engineered chloroplast genome just got smarter. *Trends Plant Sci.* **20**, 622–640 (2015).
38. Hoekstra, L. A., Siddiq, M. A. & Montooth, K. L. Pleiotropic effects of a mitochondrial–nuclear incompatibility depend upon the accelerating effect of temperature in *Drosophila*. *Genetics* **195**, 1129–1139 (2013).
39. Mossman, J. A., Biancani, L. M., Zhu, C.-T. & Rand, D. M. Mitonuclear epistasis for development time and its modification by diet in *Drosophila*. *Genetics* **203**, 463–484 (2016).
40. Hill, G. E. et al. Assessing the fitness consequences of mitonuclear interactions in natural populations. *Biol. Rev.* **94**, 1089–1104 (2019).
41. Yin, L. et al. Photosystem II function and dynamics in three widely used *Arabidopsis thaliana* accessions. *PLoS ONE* **7**, e46206 (2012).
42. Gobron, N. et al. A cryptic cytoplasmic male sterility unveils a possible gynodioecious past for *Arabidopsis thaliana*. *PLoS ONE* **8**, e62450 (2013).
43. Wijnker, E. et al. Hybrid recreation by reverse breeding in *Arabidopsis thaliana*. *Nat. Protoc.* **9**, 761–772 (2014).
44. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBO J.* **17**, 10–12 (2011).
45. Sloan, D. B., Wu, Z. & Sharbrough, J. Correction of persistent errors in *Arabidopsis* reference mitochondrial genomes. *Plant Cell* **30**, 525–527 (2018).
46. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Preprint at <https://arxiv.org/abs/1303.3997> (2013).
47. Li, H. et al. The sequence alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
48. The 1001 Genomes Consortium. 1,135 genomes reveal the global pattern of polymorphism in *Arabidopsis thaliana*. *Cell* **166**, 481–491 (2016).
49. Flood, P. J. et al. Phenomics for photosynthesis, growth and reflectance in *Arabidopsis thaliana* reveals circadian and long-term fluctuations in heritability. *Plant Methods* **12**, 1–14 (2016).
50. Kokorian, J., Polder, G., Keurentjes, J., Vreugdenhil, D. & Guzman, M. O. in *Proc. ImageJ User and Developer Conference, Luxembourg, 27–29 October 2010* (eds Jahnen, A. & Moll, C.) 178–182 (Centre de Recherche Public Henri Tudor, 2010).
51. Cruz, J. A. et al. Dynamic environmental photosynthetic imaging reveals emergent phenotypes. *Cell Systems* **2**, 365–377 (2016).
52. Joosen, R. V. L. et al. Germinator: a software package for high-throughput scoring and curve fitting of *Arabidopsis* seed germination. *Plant J.* **62**, 148–159 (2010).
53. Peterson, R., Slovin, J. P. & Chen, C. A simplified method for differential staining of aborted and non-aborted pollen grains. *Int. J. Plant Biol.* **1**, 66–69 (2010).
54. Lisec, J., Schauer, N., Kopka, J., Willmitzer, L. & Fernie, A. R. Gas chromatography mass spectrometry-based metabolite profiling in plants. *Nat. Prot.* **1**, 387–396 (2006).
55. Carreno-Quintero, N. et al. Untargeted metabolic quantitative trait loci analyses reveal a relationship between primary metabolism and potato tuber quality. *Plant Physiol.* **158**, 1306–1318 (2012).
56. Wehrens, R. et al. Improved batch correction in untargeted MS-based metabolomics. *Metabolomics* **12**, 88 (2016).
57. Sumner, L. W. et al. Proposed minimum reporting standards for chemical analysis. *Metabolomics* **3**, 211–221 (2007).
58. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-seq. *Bioinformatics* **25**, 1105–1111 (2009).
59. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
60. Reimand, J. et al. g:Profiler—a web server for functional interpretation of gene lists (2016 update). *Nucleic Acids Res.* **44**, W83–W89 (2016).
61. Wood, S. N., Pya, N. & Säfken, B. Smoothing parameter and model selection for general smooth models. *J. Am. Stat. Assoc.* **111**, 1548–1563 (2016).
62. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 48 (2015).

Acknowledgements

H. Blankestijn, J. van de Belt, D. Oberste-Lehn, E. Schijlen, C. Hanhart, J. ter Riele and S. Schop (Wageningen University & Research) are acknowledged for help with experiments; J. Klasen (Max Planck Institute for Plant Breeding Research), A. Languillaume and R. van Bezouw (Wageningen University & Research) for statistical advice; and D. Aanen (Wageningen University & Research) for helpful discussions. This work was, in part, supported by the Netherlands Organization for Scientific Research through ALW-TTI Green Genetics (P.J.F.) and ALWGS.2016.012 (T.P.J.M.T.). The European Molecular Biology Organization supported this work through grant no. ALTF 679–2013 (E.W.), and the European Community through the Marie-Curie Initial Training Network ‘COMREC’ project no. 606956 funded under FP7-PEOPLE (V.C.-B.). ZonMw Enabling Technology Hotels and the Consortium for Improving Plant Yield Enabling Technology Hotels provided funds for the metabolomics, RNA-seq and seed phenotyping. Work at Michigan State University for DEPI phenotyping was supported by the US Department of Energy, Chemical Sciences, Geosciences, and Biosciences Division, Basic Energy Sciences, Office of Science at the US Department of Energy (through grant no. DE-FG02–91ER20021).

Author contributions

P.J.F. and E.W. conceived and designed the study. T.P.J.M.T. designed and performed the statistical analysis with help from P.J.F., W.K. and F.v.E. P.J.F., T.P.J.M.T., E.K., F.F.M.B., L.A.J.W., V.C.-B., J.v.A., J.M.G. and L.S. performed experiments. P.J.F., T.P.J.M.T., K.S., P.K., E.S., J.A.H., S.K.S., R.W., W.L., R.M., F.v.E. and E.W. analysed data. D.M.K., J.J.B.K., M.K., J.H. and M.G.M.A. contributed to the interpretation of results. P.J.F., T.P.J.M.T. and E.W. wrote the paper with substantial contributions from M.K., J.H. and M.G.M.A. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41477-019-0575-9>.

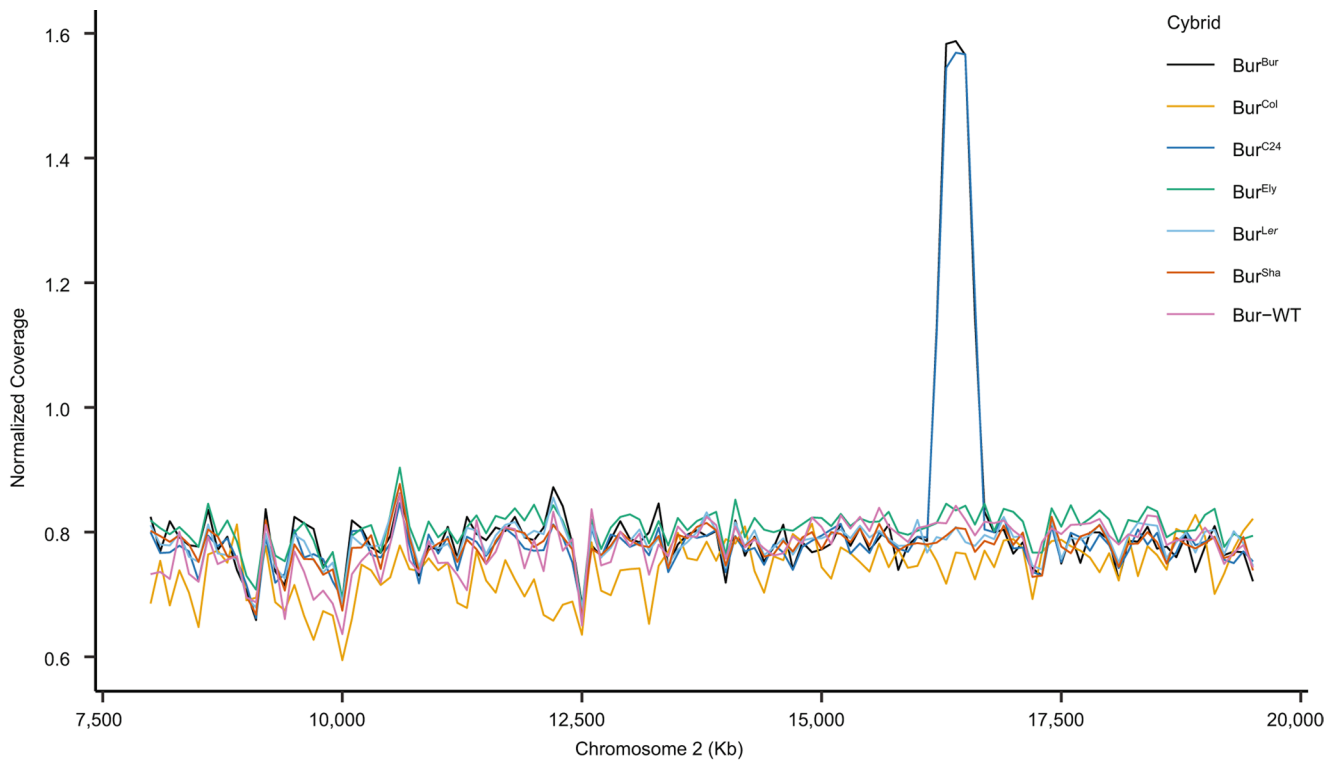
Supplementary information is available for this paper at <https://doi.org/10.1038/s41477-019-0575-9>.

Correspondence and requests for materials should be addressed to P.J.F., T.P.J.M.T. or E.W.

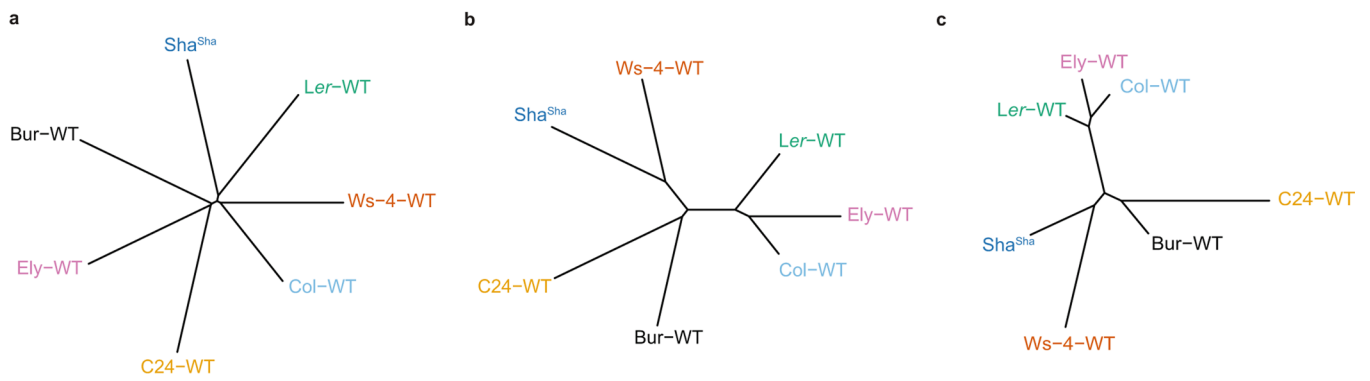
Reprints and permissions information is available at www.nature.com/reprints.

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

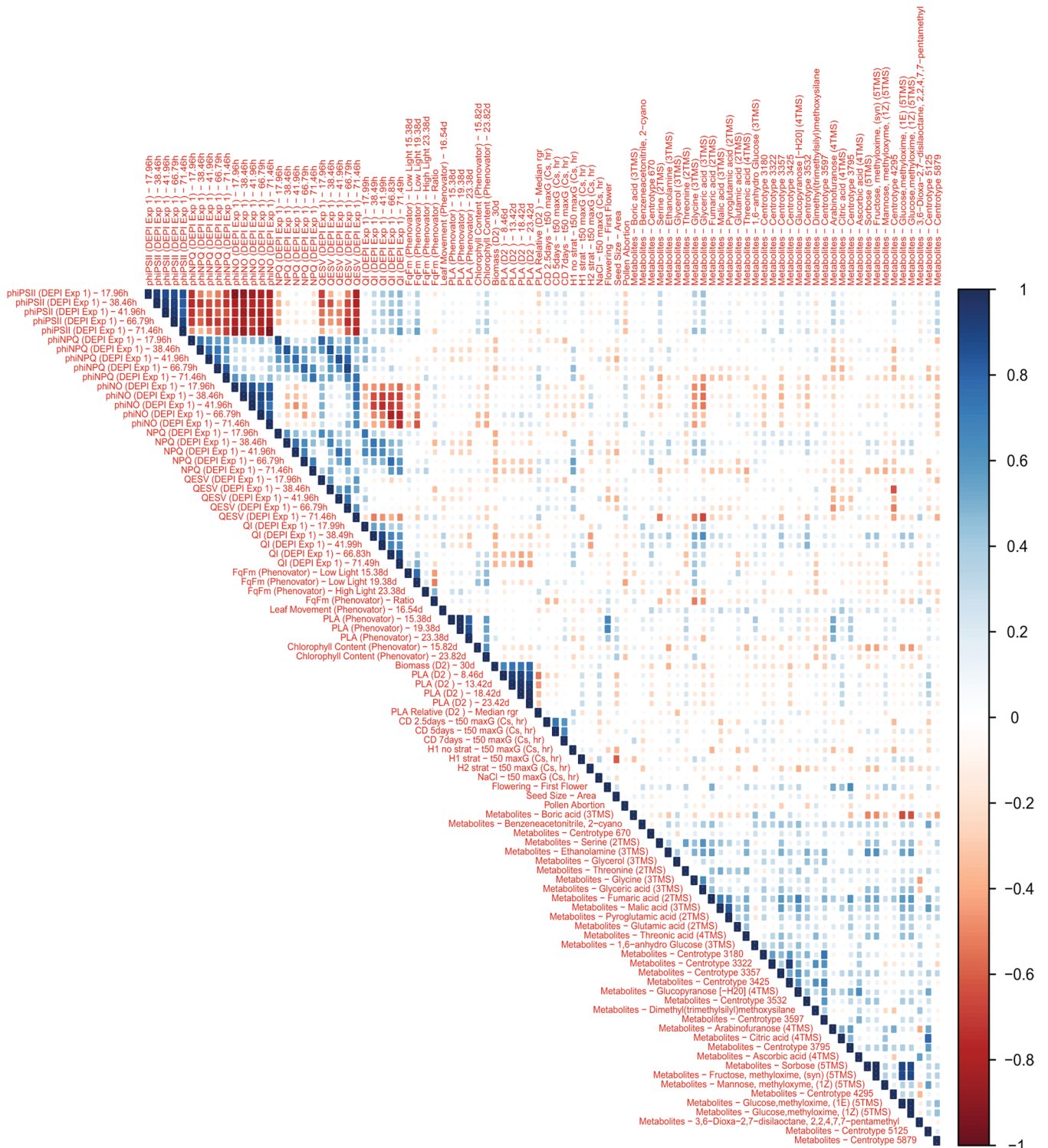
© The Author(s), under exclusive licence to Springer Nature Limited 2020



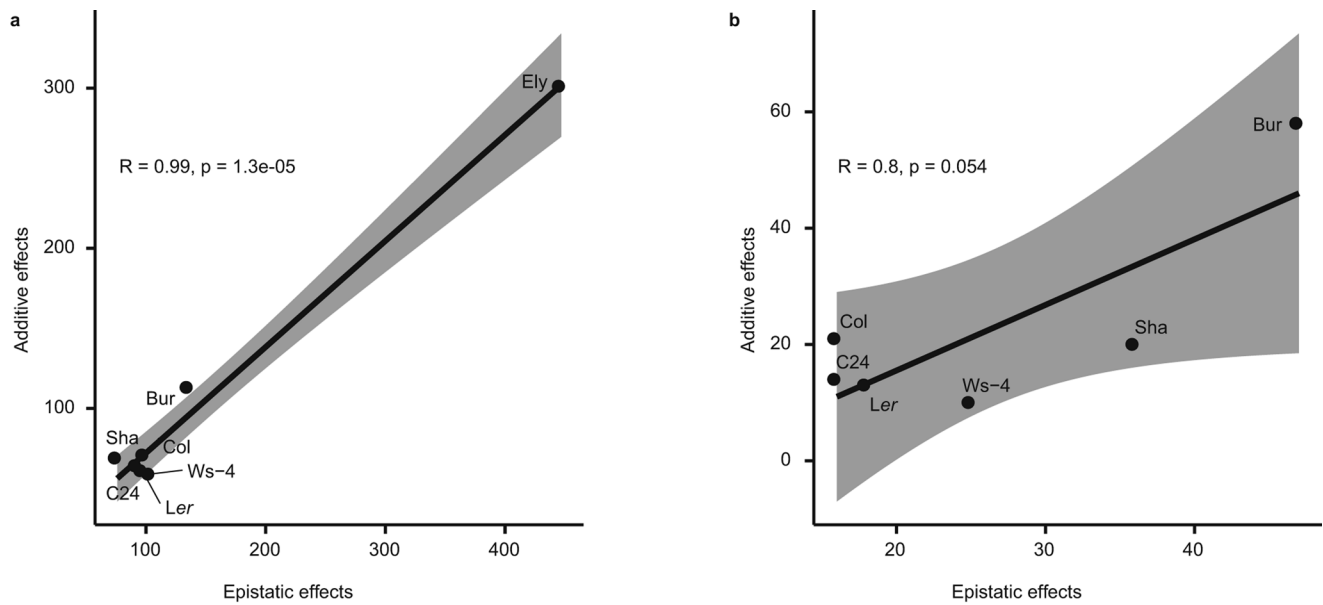
Extended Data Fig. 1 | Coverage plots reveal a duplication on chromosome 2 in two cybrid lines. Extended Data Fig. 1: Coverage plots reveal a duplication on chromosome 2 in two cybrid lines. This coverage plot shows the normalized read coverage at the lower end of the long arm of chromosome 2 for wild-type Bur (Bur-WT) and six cybrids with the Bur nucleotype (genotypes indicated as Nucleotype^{Plasmotype}). The coverage plot reveals the presence of a spontaneous nuclear DNA duplication in two cybrid lines (Bur^{Bur} and Bur^{C24}), presumably derived from their wild-type Bur progenitor. These lines were excluded from all further analyses.



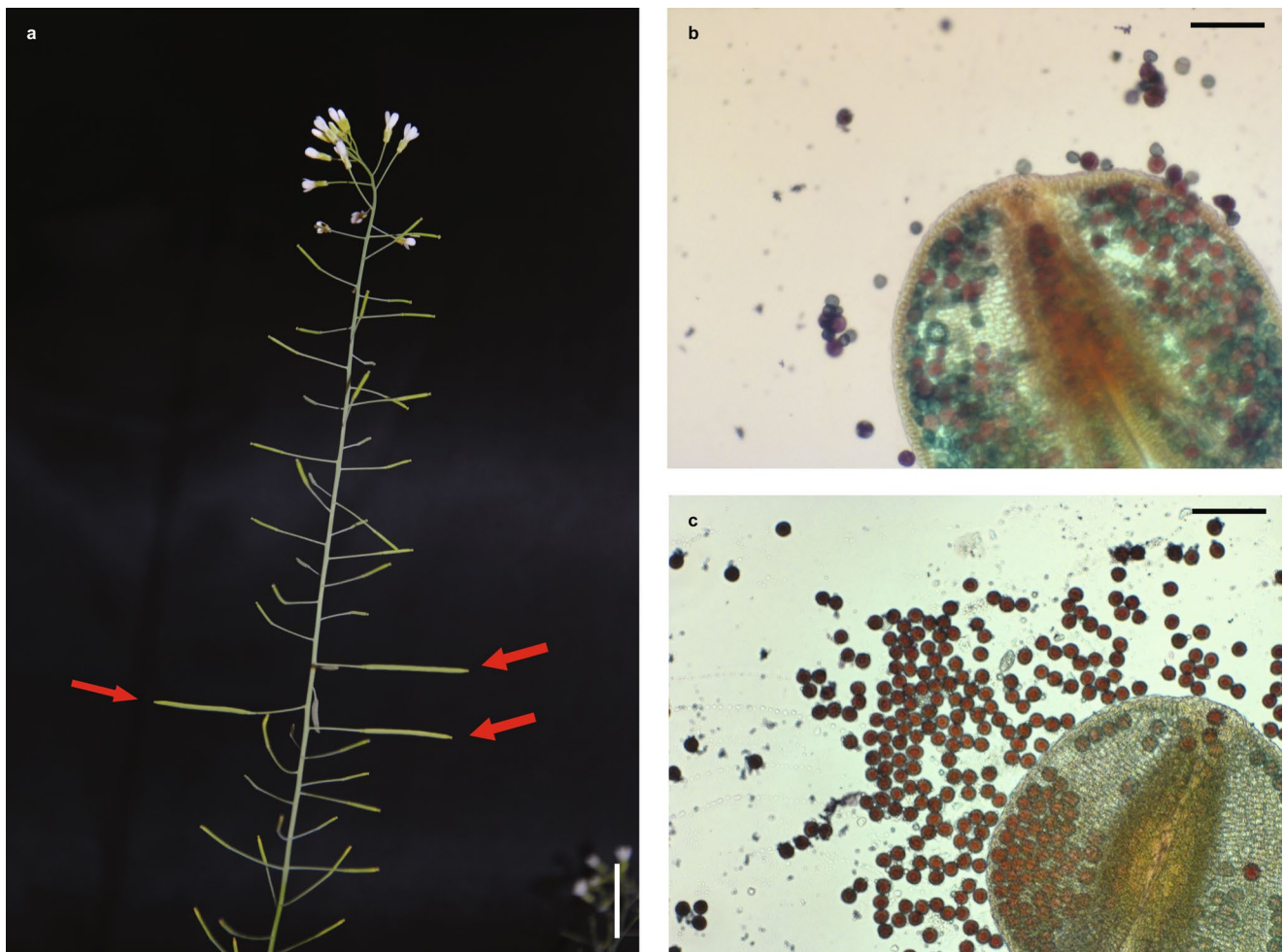
Extended Data Fig. 2 | Neighbor joining (NJ) trees based on SNPs and INDELS for nucleus (a), chloroplast (b) and mitochondria (c) for the seven *Arabidopsis thaliana* accessions.



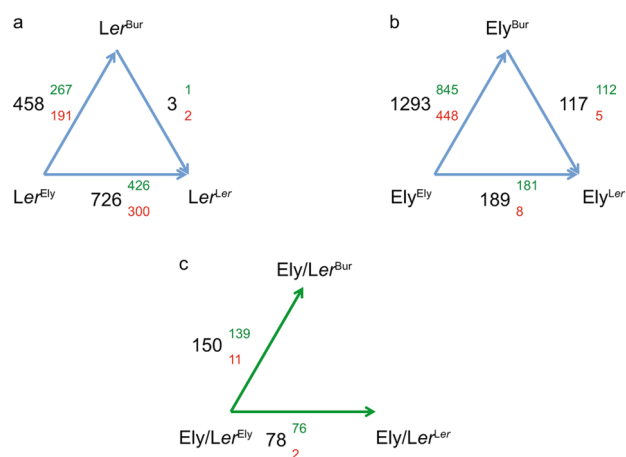
Extended Data Fig. 3 | Correlation of plasmotype effects amongst the subset of 92 phenotypes. For more information on the 92 phenotypes see Supplementary Tables 2 and 4. Dark blue has a correlation of 1 and dark red a correlation of -1. Correlation represents the Pearson correlation coefficient.



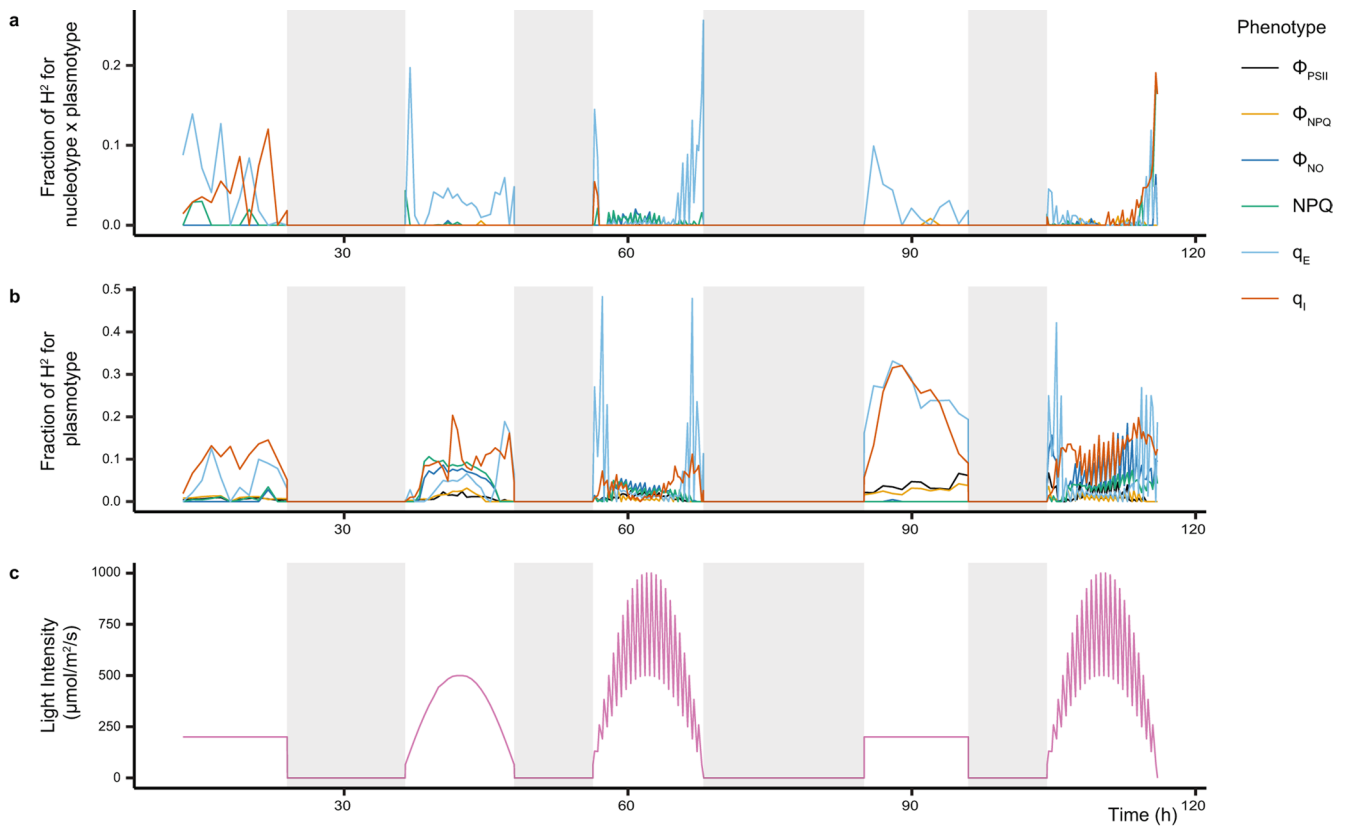
Extended Data Fig. 4 | Scatterplots showing the correlation between the number of plasmotype additive and plasmotype epistatic effects. a, Shows the correlation for additive and epistatic effects for all plasmotypes (including the Ely plasmotype; $n = 7$ plasmotypes) in every comparison, averaged over the nucleotypes, and counted over the 92 phenotypes. **b**, Shows the same correlation when excluding the Ely plasmotype ($n = 6$ plasmotypes). R is the Pearson correlation coefficient, the P value is based on a two-sided t-test and the shaded area shows the 95% confidence interval around the regression line.



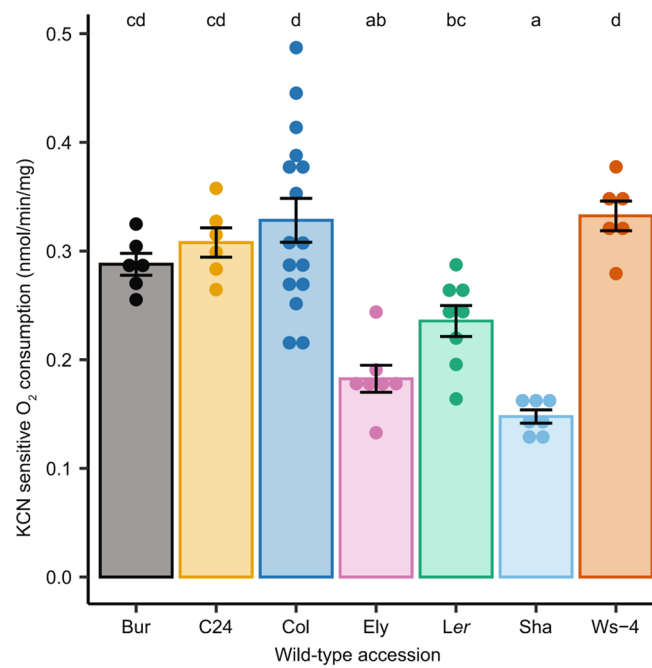
Extended Data Fig. 5 | Cytoplasmic male sterility in *Ely^{Sh}*. **a**, An *Ely^{Sh}* plant was pollinated on three open flowers using *Ely* wild-type pollen, which produced elongated siliques (indicated with red arrows), scale bar is 1 cm. **b**, Shows an anther and pollen of *Ely^{Sh}*, stained using Alexander stain. Pollen viability was assessed in 250 pollen per flower ($n=3$ flowers). Note the presence of a high percentage (45%) of greenish, almost colourless aborted pollen. Pollen with a red colour in this line are not able to fertilize ovules, as deduced from the male sterile phenotype of *Ely^{Sh}* (as shown in panel **a**), scale bar is 500 μm . **c**, Anther and pollen of *Ely* wildtype. Note that all pollen have a dark red colour, suggesting high viability. *Ely* wildtype is able to fertilize *Ely^{Sh}* (as shown in panel **a**), scale bar is 500 μm .



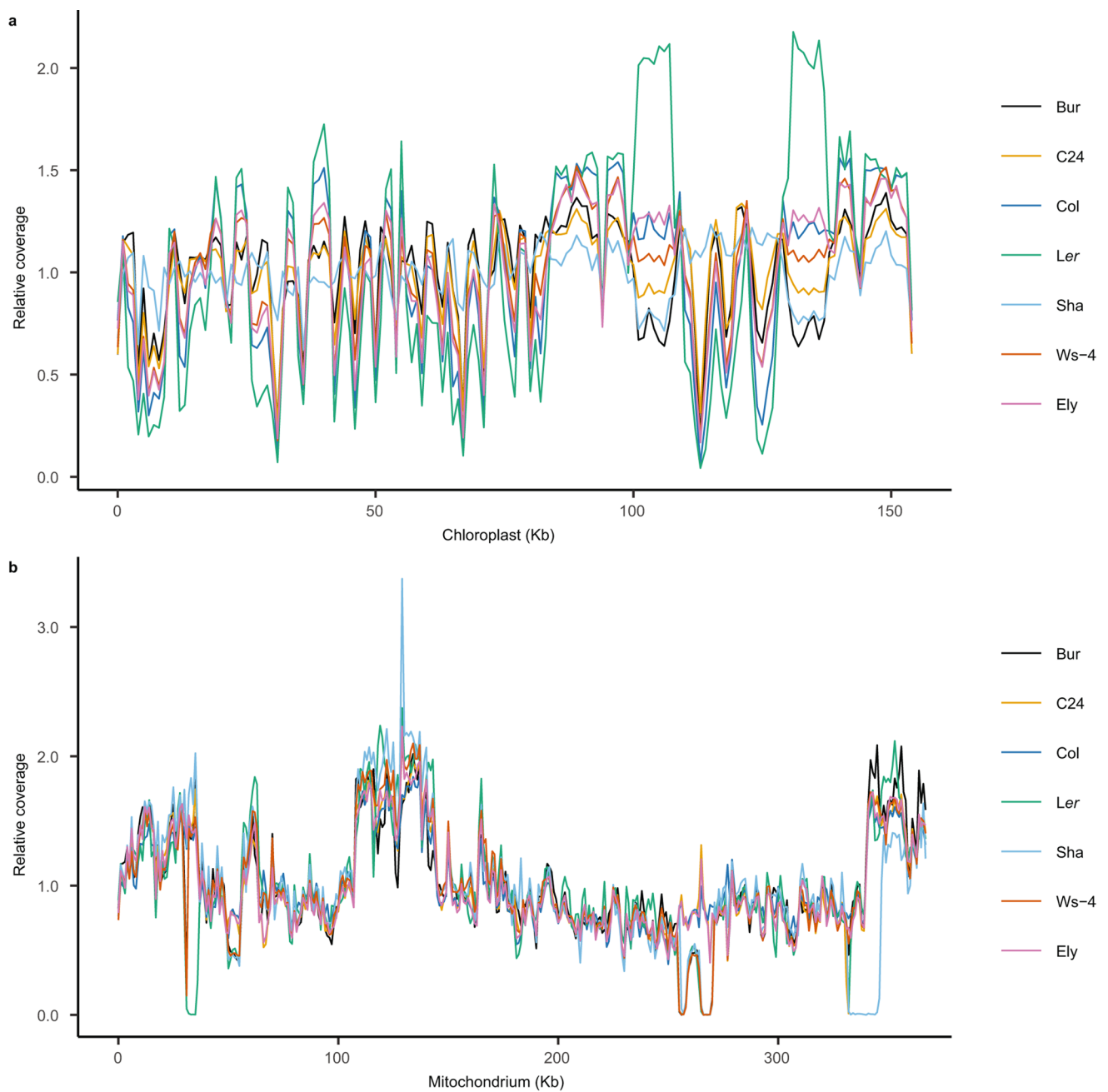
Extended Data Fig. 6 | Changes in gene expression between cybrids with a Ler nucleus (panel a), an Ely nucleus (panel b) and changes they have in common (panel c). Cybrid genotypes are indicated as Nucleotype^{Plasmotype}. The triangle in panel **a** shows cybrid comparisons with a Ler nuclear background (for plasmotypes Ely, Ler and Bur) and panel **b** shows cybrid comparisons with an Ely nuclear background for the same plasmotypes. Significantly differentially expressed (DE) genes between cybrid comparisons are indicated with black numbers. These DE genes are subdivided in upregulated genes (green numbers in superscript) and downregulated genes (red numbers in subscript), following the direction of the arrows between cybrids (that is the change from an Ely to a Ler plasmotype in a Ler nuclear background resulted in 726 DE genes, of which 426 were upregulated and 300 were downregulated). The green triangle in panel **c** shows what differentially expressed genes the comparisons in panels **a** and **b** have in common. For example, the Ler and Ely nuclear backgrounds show a common response of 78 DE genes when the Ely plasmotype is changed for a Ler plasmotype. The absence of one of the comparisons in this triangle is due to the absence of shared DE genes. The common effect of changing an Ely plasmotype for either Bur or Ler was derived by assessing what DE genes are similar along the axes in the green triangle **c**. These 78 and 150 genes have 40 shared DE genes (see Supplementary Table 6). For the raw data see Supplementary Data 3.



Extended Data Fig. 7 | The fraction of explained genetic variation (H^2) for changes in photosynthesis phenotypes (Φ_{PSII} , Φ_{NPQ} , Φ_{NO} , NPQ, q_E , q_i) in response to light conditions. **a, Shows the fraction of H^2 for epistatic interactions (nucleotype x plasmotype). **b**, Shows H^2 for plasmotype additive effects. **c**, Shows the light intensity for five consecutive days with growth under: steady light (day 1); in- and decreasing light intensity (day 2); fluctuating in- and decreasing light intensity (day 3); steady light (day 4) and fluctuating in- and decreasing light intensity (day 5). Days are separated by nights (shaded areas). The first three days of this experiment are identical to the light conditions of the experiment shown in Fig. 3.**



Extended Data Fig. 8 | KCN sensitive O₂ consumption of mitochondria in seedlings of wild-type accessions. Mitochondrial ATP-synthesis proceeds mainly through the phosphorylating cytochrome (KCN sensitive) pathway. KCN sensitive O₂ consumption by Bur does not differ significantly from C24, Col and Ws-4. Error bars represent the standard error of the mean (n = minimally 6 biologically independent samples). Letters indicate significant differences using the Tukey posthoc test, with an α = 0.05 threshold.



Extended Data Fig. 9 | Coverage plots reveal no unique deletions in Bur chloroplasts or mitochondria. Normalized read depth for the chloroplast (a) and mitochondrial (b) genome sequences were calculated in a sliding window of 1-kb. Because we observe no unique deletions or duplications in the Bur plasmotype that might be causal to the phenotypic effects observed in cybrids with the Bur plasmotype.

Reporting Summary

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

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a Confirmed

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

Our web collection on [statistics for biologists](#) may be useful.

Software and code

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

Data collection

No custom code was used in data collection. The automated imaging systems used are the GERMINATOR (version 1.0), Phenotyping XY system (version 3.1), uEye Cockpit (version 4.00.0000) and VisualPhenomics (version 5), all published or open source software. Please see material and methods for details.

Data analysis

No custom code was used in data analysis. During the analysis we used published or open source software. Images were analyzed using GERMINATOR (version 1.0), TTI-GG (version 3.2), VisualPhenomics (version 5) and ImageJ (version 1.8.0). Raw reads were processed through Cutadapt (version 1.18), bwa mem (version 0.7.10-r789), samtools (version 1.3.1), GATK suite (version 4.0.2.1). Subsequent analysis was done in Rstudio (version 1.0.153), using the packages lsmmeans (version 2.30-0), lme4 (version 1.1-21), lmerTest (version 3.1-0) and DeSeq2 (version 1.26.0), nlme (version 3.1-140), mgcv (version 1.8-28) and flux (0.3-0). GO enrichment was done using the g:CoCoo webtool. Graphs were made in Rstudio (1.0.153), using the packages ggplot2 (3.2.0), plyr (version 1.8.4), vcfR (version 1.8.0), ape (5.3), cluster (2.1.0), corrplot (version 0.84), Hmisc (version 4.2-0), tidyr (version 0.8.3), reshape2 (version 1.4.3), ggpubr (version 0.2.3) and ggrepel (version 0.8.1). Please see material and methods for further details.

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

Data

Policy information about [availability of data](#)

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

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Sequencing and transcriptome data are available through the European Nucleotide Archive with the primary accession codes PRJEB29654 and PRJEB35324. The raw datasets are available through Dryad via doi:10.5061/dryad.cz8w9gj05. The analysed datasets that support our findings are available as Supplementary Data. The associated raw data for Figs. 2 and 3 are provided in Supplementary Data 1, the raw data for Figure 2 is provided in Supplementary Data 2. The germplasm generated in this project will be available via the European Arabidopsis Stock Centre (www.arabidopsis.info).

Field-specific reporting

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

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

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

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

Sample size

Our study encompasses a reciprocal cybrid panel of 49 lines, seven lines of which (so called self-cybrids, as defined in the manuscript) served as controls. The seven wildtype progenitors of our panel were chosen to encompass a wide sample of natural genetic variation in Arabidopsis. Because the size of a reciprocal cybrid panel increases exponentially, we limited our choice to seven. As such, this choice was practically motivated. We anticipated that generating a panel of seven x seven was practically feasible.

Due to the explorative nature of this study, time and money was invested in measuring as many different phenotypes as possible with available resources. Because we did not know what to expect in terms of effect sizes (but anticipated effects to be relatively small), we did try to maximize the number of replicates within the different phenotyping platforms where possible. This usually meant that we grew the maximum number of replicates possible in a specific genotyping platform (i.e. DEPI system, Phenovator, D2). Sample sizes for all experiments are listed in Supplementary Table 4.

Data exclusions

Plants that died or did not grow after germination (i.e. did not establish well in experiments) were removed from the analysis. The loss of some plants during our large scale phenotyping experiments was anticipated beforehand, as well as the subsequent removal of those data from our analyses. Plant removal was done before genotype names were visible to us.

One cybrid Ely-Sha died before setting seed and was missing from all experiments, this is documented in the text. This line was later regenerated, and its (male-sterile) phenotype is described in the manuscript (e.g. in Extended Data Fig. 5). We also recovered the presence of a duplicated segment in two cybrid lines, lines Bur-Bur and Bur-C24. All data we collected from these lines were therefore excluded from further analysis and figures (i.e. after detection of the duplication, during the revision of our manuscript during the review process). The reasons for exclusion of these two lines from our data analysis are discussed in the manuscript, and the presence of the duplicated segment on chromosome two is shown in Extended Data Fig. 1.

Replication

Due to the explorative nature of this study, time and money was invested in measuring as many different phenotypes as possible with available resources. A main finding in our manuscript is the recovery of a main plasmotype effect of the Burren (Bur) plasmotype in PSII efficiency. This finding was detected in two different phenotyping platforms available to us (the Phenovator (Netherlands) and the DEPI system (USA)). In our revised manuscript, in which we emphasized the presence of this Bur plasmotype additive-effect, we included the replication of a phenotyping experiments in the DEPI system, in which we detected most Bur-plasmotype additive effects, confirming the presence of this Bur plasmotype main effect. This system assesses many, mainly photosynthetic parameters under fluctuating light conditions.

Because we had no prior idea of effect sizes, the replicate numbers used lines in our experiments were mostly limited by the phenotyping platforms used. In the case of the phenovator we grew 24 replicates per genotype from these we created 4 pools of 6 plants of which we used 3 pooled samples for metabolomics and transcriptomics. Replicates in the DEPI system were much more restricted (4 per line) as this system cannot accommodate for more plants. Further details of replicate numbers for each phenotype can be found in the text (summary in Supplementary Table. 4).

Randomization

All experiments conducted were done with completely randomized block designs.

Blinding

Names of seeds were visible to researchers to avoid mistakes during sowing, all experiments were set in automated phenotyping platforms, as such the researcher had no influence.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involvement in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Unique biological materials
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants

Methods

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Unique biological materials

Policy information about [availability of materials](#)

Obtaining unique materials