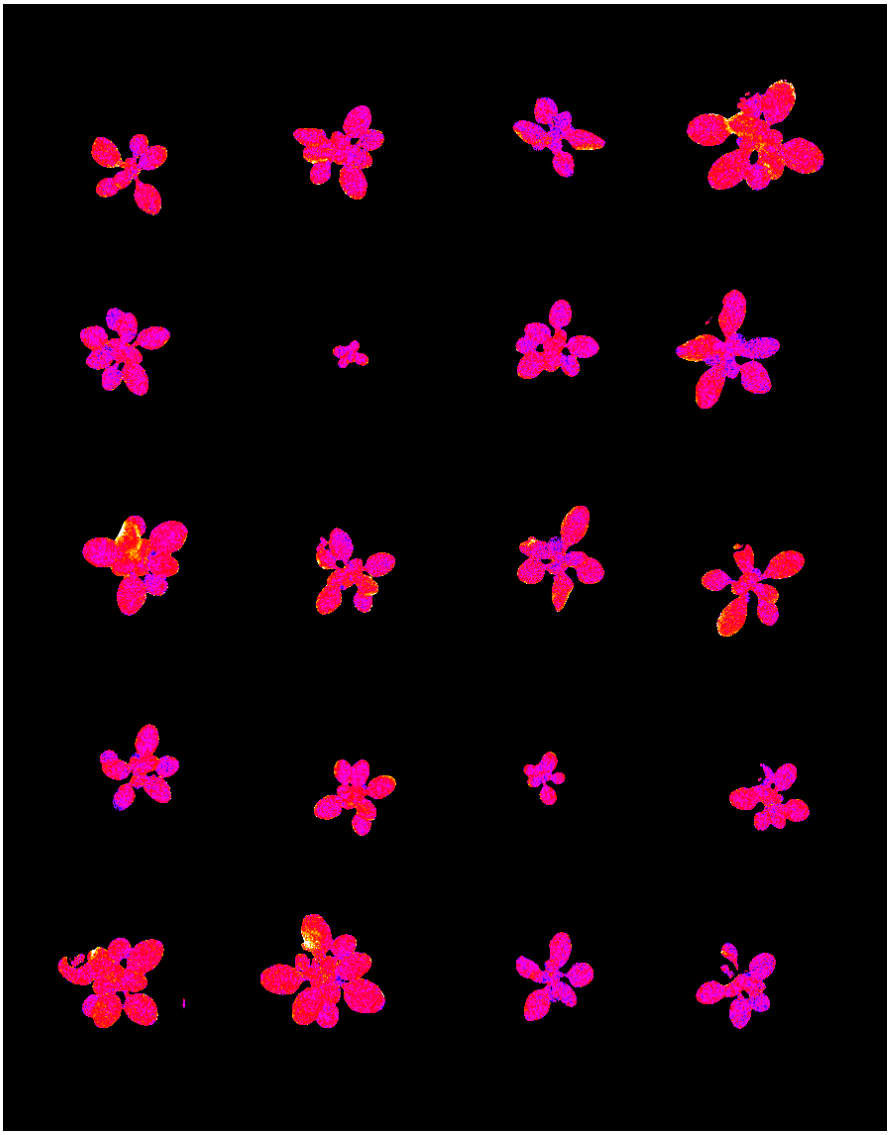
EXPLORING PLASMOTYPIC VARIATION IN ARABIDOPSIS THALIANA



Delfi Dorussen

Exploring Plasmotypic Variation in *Arabidopsis thaliana*

Delfi Dorussen
Registration number: 1047507
MSc Biology (Development & Adaptation) Major Thesis (36 ECTS)

Supervised by Tom Theeuwen MSc
Examined by René Boesten MSc and prof.dr. Mark Aarts
Department of Genetics, Wageningen University & Research

Table of Contents

Abstract

The role of the chloroplast and mitochondrial genomes in shaping plant phenotypes is often overlooked. While the cytoplasmic genomes are smaller than the nuclear genome, they encode proteins that are central to plant metabolism and photosynthesis. Furthermore, the organelles are in a prime position to detect changes in environmental conditions via changes in their redox state. The ability of plants to respond to these environmental changes is a key determinant of photosynthetic efficiency, affecting crop yields. However, until recently, it has been difficult to separate the effects of the nuclear and cytoplasmic genomes as they are inherited together from the maternal parent. The development of *Arabidopsis thaliana* cybrids, which possess the nuclear genome of one parent and the cytoplasmic genomes of the other parent, allows the effect of variation in these genomes to be studied separately. In this project, the contribution of the cytoplasmic genomes to natural variation in photosynthetic traits is shown. Data from high-throughput phenotyping studies was analysed, revealing significant contributions of the cytoplasmic genomes to the heritability of photosynthetic traits, including dynamic traits such as non-photochemical quenching (NPQ). A plasmotype with a significant effect on photosynthetic efficiency, Can-0, has been identified. Furthermore, possible genetic variants explaining the observed phenotypes have been identified. Overall, this indicates that natural variation in the cytoplasmic genomes may be an underutilised resource for fine-tuning the photosynthetic efficiency of plants.

Introduction

Current yield increases are not expected to meet the global demand for food by 2050 (Ray et al., 2013). The yield potential of a crop is defined as the product of the available photosynthetically active radiation, its interception efficiency, its photosynthetic efficiency and its partitioning efficiency (Monteith, 1977; Long et al., 2006). Partitioning and interception efficiency are both nearing their theoretical maxima, following large improvements in these traits during the Green Revolution, leading to significant yield increases (Zhu, Long and Ort, 2010). However, photosynthetic efficiency can still be improved. This room for improvement was shown by growing crops at elevated concentrations of atmospheric CO_2, resulting in an increase in photosynthetic efficiency and a concomitant increase in yield (Zhu, Long and Ort, 2010). This indicates that crop photosynthesis is currently an underutilised resource for improving yields.

Photosynthesis is a complex biological process, consisting of many interlocking reactions. As such, photosynthesis can be limited in a number of different ways, broadly categorised into limitation of ribulose-1,5-bisphosphate (RuBP) carboxylation, and limitation of RuBP regeneration (von Caemmerer, 2000). RuBP carboxylation depends on the activity of Rubisco, the enzyme that catalyses the carboxylation reaction, while RuBP regeneration depends on the light reactions and photosynthetic electron transport for the production of NADPH and ATP, and the activity of the enzymes in the Calvin-Benson-Bassham (CBB) Cycle (von Caemmerer, 2000). As the carboxylation activity of Rubisco is strongly linked with the CO_2 concentration in the chloroplasts, photosynthesis is further limited by conductance of CO_2 through the stomata and the mesophyll (Zhu, Long and Ort, 2010). Under ambient CO_2 concentrations, photosynthesis, measured as the rate of CO_2 fixation, can be limited by any one of these processes (Zhu, Long and Ort, 2010), indicating that RuBP carboxylation and regeneration both provide opportunities for the improvement of photosynthesis.

While photosynthesis is usually measured under uniform, steady-state conditions, plants are often subject to dynamic environmental conditions in the field. This includes changes in light intensity, for example due to cloud cover or shading from neighbouring plants, temperature, water availability, and the presence of pathogens and herbivores (Murchie et al., 2018; Matsubara, 2018). A plant's ability to respond to these changes is likely a key factor determining its productivity (Murchie et al., 2018). Considerable attention has been given to plant responses to fluctuating light conditions; photosynthesis is intimately linked to light

availability, and this is the most rapidly changing feature of the abiotic environment (Pearcy, 1990). Plants respond to changes in light intensity in a number of different ways but there is generally a lag between the change in irradiance and induction of photosynthesis, resulting in suboptimal photosynthetic efficiency (Pearcy, 1990). Thus, improving the efficiency with which plants respond to the dynamic light environment may provide a route for improving crop yields.

On a transition from low light to high light (such as a shade to sun transition), Rubisco is activated by Rubisco activase (Rca). Rca removes sugar phosphates (including RuBP and the sugar phosphates formed by Rubisco misfire reactions – xylulose-1,5-bisphosphate (XuBP) and D-glycero-2,3-pentodiulose-1,5-bisphosphate (PDBP)) from the Rubisco active site in an ATP-dependent manner (Mueller-Cajar, Stotz and Bracher, 2014). In many plant species, including *Arabidopsis thaliana*, two isoforms of Rca are produced: a long isoform (α) and a short isoform (β), whereby the long isoform can be directly redox-regulated by the ferredoxin-thioredoxin couple (Zhang and Portis, 1999; Mueller-Cajar, Stotz and Bracher, 2014). This ensures that Rca activity, and therefore Rubisco activity, is restricted to periods of high irradiance and thus high photosynthetic electron transport. Furthermore, Rca activity increases with an increased ATP/ADP ratio and Mg^{2+} concentration, both of which are increased in the chloroplast stroma under high irradiance (Lilley and Portis, 1997; Carmo-Silva et al., 2015; Hazra et al., 2015). The expression level of Rca and the relative abundance of the α and β isoforms have been proposed to explain differences in the speed of photosynthetic induction (Carmo-Silva et al., 2015). Indeed, the speed of photosynthetic induction appears to be closely linked to Rubisco activity; carbon assimilation following a shade-to-sun transition was primarily limited by the carboxylation of RuBP (for which the $v_{c,max}$ is often used as a proxy) in soybean, wheat, and rice (Soleh et al., 2017; Salter et al., 2019; Acevedo-Siaca et al., 2020). Additionally, constitutive activation of Rubisco in *A. thaliana* resulted in a faster increase in photosynthetic activity compared to wild type plants (Kaiser et al., 2016).

Carbon assimilation is further limited by the lag in stomatal opening, allowing CO_2 diffusion into the leaf, following a shade-to-sun transition (Lawson and Blatt, 2014). Stomatal opening is an order of magnitude slower than the response of the photosynthetic apparatus, suggesting that a faster stomatal response would have a significant impact on carbon assimilation (Pearcy, 1990; Matthews, Vialet-Chabrand and Lawson, 2020). It is unknown how fluctuations in irradiance are sensed and relayed to the stomatal guard cells; direct

sensing of red light, changes in internal CO_2 concentration, guard cell photosynthesis (producing sugars or ATP that may act as the signal), and the redox state of the plastoquinone pool have all been proposed (Matthews, Vialet-Chabrand and Lawson, 2020). Differences in guard cell morphology also affect the speed with which stomata open in response to light, with the dumbbell-shaped guard cells found in monocots allowing a faster response than the elliptical-shaped guard cells found in eudicots such as *A. thaliana* (McAusland et al., 2017).

The relatively fast response of photosynthetic electron transport to a change in irradiance compared to the slower induction of CO_2 fixation also results in an imbalance between the two phases of photosynthesis and can lead to an over-reduction of the electron transport chain upon a transition to high light. In particular, the D1 protein of the PSII complex is prone to damage and this damage can result in photoinhibition (Murchie and Ruban, 2020). One way that plants cope with this oxidative stress is through the induction of non-photochemical quenching (NPQ), which acts to dissipate the excess energy captured by PSII as heat (Murchie and Ruban, 2020). The energy-dependent form of NPQ, qE, is induced in response to changes in pH across the thylakoid membrane and involves the protein PsbS and the production of zeaxanthin via the xanthophyll cycle. Both PsbS and zeaxanthin are thought to promote aggregation of the light harvesting complexes surrounding PSII (LHCII), promoting energy dissipation (Murchie and Ruban, 2020). NPQ appears to be particularly important when plants are grown under fluctuating conditions; overexpression of *PsbS* in rice grown under fluctuating light conditions resulted in improved productivity (Hubbart et al., 2018). However, the slow relaxation of NPQ following a sun to shade transition, when the plant is no longer at risk of photoinhibition, can also limit biomass accumulation. Overexpression of *PsbS* and enzymes of the xanthophyll cycle increased the speed of NPQ relaxation in tobacco, improving the dry matter productivity by 15% (Kromdijk et al., 2016). However, similar results were not observed when the same genes were overexpressed in *A. thaliana*, indicating that different yield bottlenecks may be present in different species (Garcia-Molina and Leister, 2020).

An imbalance between photosynthetic electron transport and CO_2 assimilation may also occur due to other environmental fluctuations, such as temperature. Under low temperatures, the activity of metabolic enzymes generally decreases. Thus, the photosynthetic system may become quickly overexcited, leading to photoinhibition (Adam and Murthy, 2013). An investigation into natural variation in photosynthetic responses to low

temperatures in *A. thaliana* revealed that the low temperature response involves a number of proteins previously identified as being involved in the high light response (Prinzenberg et al., 2020). Overall, this indicates that environmental fluctuations are central to the regulation of photosynthesis.

Different methods have been proposed to improve plant photosynthesis. There are significant ongoing efforts to improve photosynthesis by genetic modification, for example by overexpressing rate-limiting enzymes in the Calvin-Benson-Bassham Cycle (Driever et al., 2017), genetically engineering C4 photosynthesis into rice (Kajala et al., 2011), or by introducing alternative photorespiratory pathways to improve Rubisco efficiency (South et al., 2019). However, genetic modification relies on prior knowledge of the mechanisms and pathways involved and so is limited by our current understanding of photosynthesis in particular, as well as plant physiology more broadly. Alternatively, existing natural variation in photosynthetic traits could be used (Flood, Harbinson and Aarts, 2011). Natural variation in photosynthesis has been identified both between and within species (Wullschleger, 1993), including crop species such as wheat and rice (Driever et al., 2014; Qu et al., 2017). However, in many cases, the genetic variation underlying the phenotypic variation has not been fully explored (Flood, Harbinson and Aarts, 2011). Assessing and dissecting the natural variation present in a population will allow a greater understanding of the genetic regulation of photosynthesis and potentially provide insight into the physiological mechanisms underpinning this process.

Even less is known about the contribution of natural variation in the cytoplasmic genomes to variation in photosynthetic traits. Plants have two cytoplasmic genomes, the mitochondrial and plastid genomes, derived from the genomes encoded by the organelles' endosymbiotic ancestors (Timmis et al., 2004). During the course of evolution, many genes have been transferred from the organelles to the nucleus via endosymbiotic gene transfer – for example, the *A. thaliana* chloroplast genome contains 87 genes, while its cyanobacterial ancestor likely had a genome of several thousand genes (Sato et al., 1999; Timmis et al., 2004). The core set of genes retained in the *A. thaliana* chloroplast genome encodes components of the photosynthetic apparatus (including subunits of photosystem I (PSI), photosystem II (PSII), ATP synthase, the NADH dehydrogenase-like (NDH) complex, and the large subunit of Rubisco (RbcL)) and the plastid's own genetic machinery (including tRNAs, rRNAs, ribosomal proteins, and an RNA polymerase) (Sato et al., 1999). The colocation for redox regulation of gene expression (CoRR) hypothesis has been proposed to

explain the retention of genes in the plastids and mitochondria (Allen, 2015). This hypothesis postulates that the retention of genes in the organelles allows their expression to be directly regulated in response to the local redox state, allowing for fine-tuning of the metabolic processes taking place within the organelles (Allen, 2015). In the chloroplasts, for example, expression of *psbA* (encoding a component of PSII) and *psaAB* (encoding a component of PSI) are coordinated in response to the plastoquinone redox state, allowing the stoichiometry of the two photosystems to be altered as necessary (Pfannschmidt et al., 1999; Allen, 2015). Overall, this indicates a close relationship between the metabolic processes taking place within the organelles and their genomes.

Due to the small size of the cytoplasmic genomes relative to the nuclear genome and the assumption that polymorphisms in the cytoplasmic DNA are overwhelmingly neutral (following the observation that the cytoplasmic genomes are generally under purifying selection), the role of the organelle genomes in environmental adaptation has typically been overlooked (Bock, Andrew and Rieseberg, 2014). However, there are indications that genetic variation in the cytoplasmic genome can indeed be adaptive. For example, the chloroplast *rbcL* gene, encoding the large subunit of Rubisco, was shown to be under positive selection in land plants (Kapralov and Filatov, 2007). In particular, residues affecting interactions with other Rubisco subunits or interactions with Rca were often under positive selection (Kapralov and Filatov, 2007). Furthermore, a mutation in *psbA*, encoding the D1 protein of PSII, was positively selected in the Ely ecotype of *A. thaliana* as it confers resistance to atrazine, a herbicide (El-Lithy et al., 2005). Cytoplasmic natural variation was investigated more systematically by Roux et al. (2016), who used repeated backcrossing to generate combinations of nuclear and cytoplasmic genomes from different *A. thaliana* accessions. This revealed contributions of both the cytoplasmic genome and interactions between the cytoplasmic and nuclear genomes to variation across a range of traits (Roux et al., 2016).

The contribution of the cytoplasmic genomes to variation specifically in photosynthetic traits in *A. thaliana* was further analysed by Flood et al. (2020) using a haploid induction system to produce cybrids, combining the nuclear and cytoplasmic genomes of different accessions. For example, the Ely plasmotype (cytoplasmic genomes) could be combined with the Columbia-0 (Col-0) nucleotype (nuclear genome) to produce the Ely_Col-0 cybrid. The cybrids were created by crossing a *GFP-tailswap* haploid inducer line with the desired plasmotype with an accession of the desired nucleotype. When the haploid inducer line is

the maternal parent, the maternal nuclear genome is expelled from the zygote while the cytoplasmic genome is retained along with the paternal nuclear genome. Subsequent genome duplication allows recovery of fully diploid lines. These cybrids are thus produced without repeated backcrossing, with one parent contributing the nuclear genome and the other parent contributing the cytoplasmic genomes. This approach was followed using 7 accessions (Bur, C24, Col-0, L*er*, Sha, Ws-4, and Ely), resulting in a panel of 49 cybrid lines (including self-cybrids, such as Col-0_Col-0). Phenotyping of these cybrid lines under stable conditions for photosynthetic traits, including the efficiency of PSII (ϕPSII) and the components of NPQ, showed that there was a contribution of the plasmotype to the measured phenotypes; the plasmotype accounted for 28% of the broad sense heritability (H^2) when the Ely plasmotype was included, and 2.9% when this plasmotype was excluded. The large additive effect of the Ely plasmotype is likely due to the *psbA* mutation, which reduces ϕPSII as well as conferring atrazine resistance. Nucleotype-plasmotype interactions were also found to contribute to H^2; this interaction accounted for 6.1% of H^2 when Ely was included and 5.2% when Ely was excluded. This is consistent with the fact that processes taking place in the organelles require proteins encoded by both the nuclear genome and proteins encoded by the cytoplasmic genomes; there is thus considerable room for epistatic interactions between the different genomes.

To investigate the role of the plasmotype in the response to dynamic environmental conditions, the cybrid panel was also phenotyped under fluctuating light conditions. This revealed additional contributions of the plasmotype and the nucleotype-plasmotype interaction to the H^2 of various photosynthetic traits, including qE and ϕNPQ. Overall, these results demonstrate that both additive and epistatic effects of the cytoplasmic genomes play a role in phenotypic variation in *A. thaliana*, especially under fluctuating environmental conditions.

However, the lines chosen for the analysis of Flood et al. (2020) are unlikely to represent the full scope of genetic diversity in *A. thaliana*. Analysis of the genome sequences from 1135 *A. thaliana* accessions revealed the existence of relict populations from the Iberian Peninsula, the Cape Verde Islands, the Canary Islands, Sicily and Lebanon. The genome sequences of these relict populations have a large number of pairwise differences with the non-relict Eurasian accessions, and with each other (Alonso-Blanco et al., 2016). Further sequencing of African *A. thaliana* accessions and comparison of these accessions to the known Eurasian lineages showed that Africa harbours the largest proportion of the genetic diversity

in *A. thaliana* (Durvasula et al., 2017). These relict accessions were not included in the original cybrid panel, suggesting that a large amount of the potential natural variation was not accounted for. Moreover, as no recombination takes place within the cytoplasmic genomes, mutations that arise in these genomes will only be inherited by direct maternal descendants. Variation may thus exist between the cytoplasmic genomes of otherwise closely related accessions. Therefore, a new cybrid panel has been constructed including a large number of accessions (relicts and non-relicts) to more accurately represent the geographic and genetic diversity present within *A. thaliana* (Fig 1). This panel consists of 60 different plasmotypes in 4 nucleotype backgrounds, resulting in 240 cybrid lines.

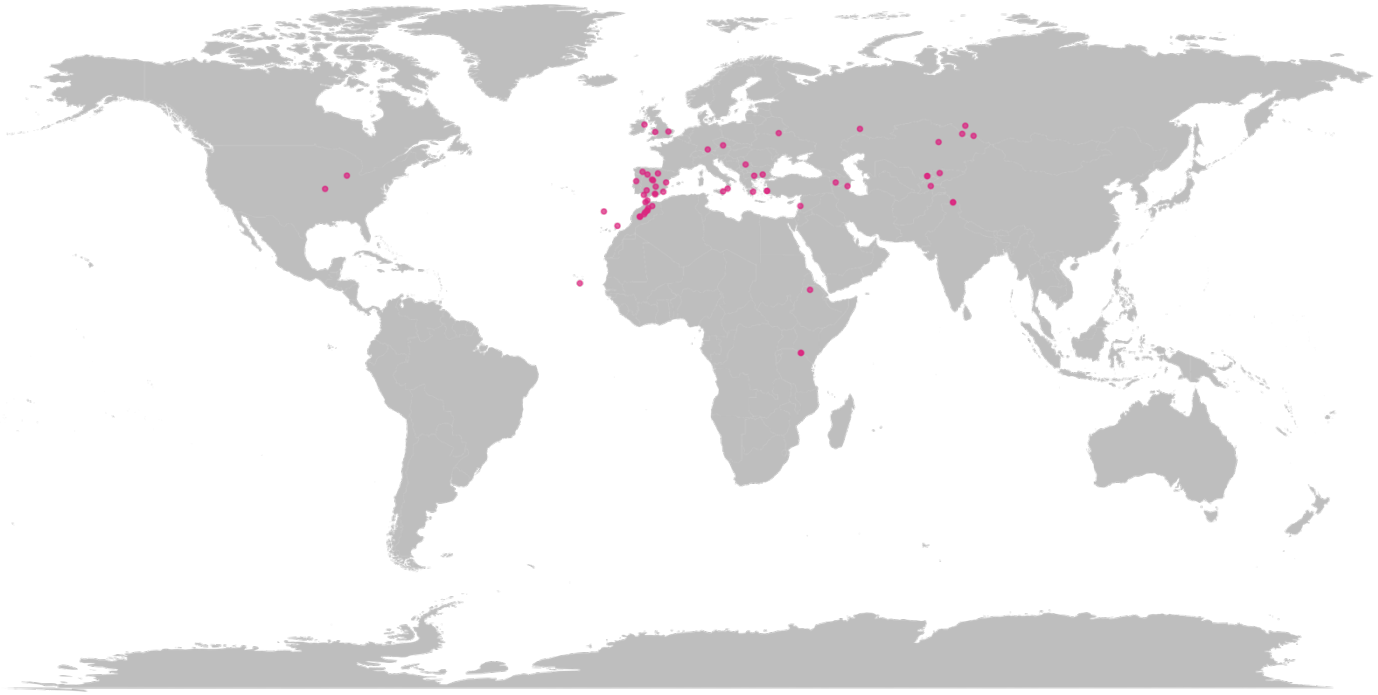


Figure 1: Geographic origin of the 60 progenitor lines of the cybrid panel.

The aim of this research project was to investigate the contribution of the plasmotype to variation in photosynthetic traits, particularly under dynamic environmental conditions. By analysing phenotypic data from the new cybrid panel, novel variation in photosynthetic traits was revealed that could be attributed to differences between the plasmotypes. Considering the impact of environmental fluctuations on the activity of the organelles and their genetic regulation, the response of the cybrids to fluctuating conditions such as variable light intensity and temperature was also examined. The second aim of this project was to improve

our understanding of the genetic and physiological basis for the variation in photosynthetic traits. This was explored by investigating the genetic differences between the plasmotypes in the cybrid panel and by performing further phenotyping studies. Thus, by coupling phenotypic variation to genetic variation we have started to dissect the mechanisms underlying the phenotypic differences between the cybrids.

Methods

Nomenclature

All cybrids are named as Plasmotype_Nucleotype. For example, the Agl-0_Bur cybrid has the Agl-0 plasmotype and the Bur nucleotype.

Statistical Analysis of Phenotypic Data

Phenovator Experiments

The cybrid panel was phenotyped using the Phenovator, a high-throughput phenotyping platform (Flood et al., 2016). The Phenovator performs chlorophyll fluorescence measurements for 1440 plants multiple times per day. The cybrid panel was screened using this system four times; once with only the Col-0 nucleotype, and three times with all four nucleotypes. The set-up for these experiments is described in more detail by Tijink (2021a).

Leaf area and PSII efficiency (ϕPSII) measurements were extracted from image files produced by the Phenovator using the TTI.exe software. Measurements from four separate experiments were subsequently compiled and outlier detection was performed on the basis of leaf area on day 15. The mean leaf area was calculated for each genotype and all plants with a leaf area less than the mean – 1.5 × standard deviation were removed. Furthermore, all genotypes that showed poor germination, resulting in stunted growth, were removed. These cybrids were not always removed by the outlier detection as this relied on the mean leaf area and standard deviation for each genotype. This included all cybrids with the Cvi nucleotype in one experiment and all of the Shah_Cvi cybrids. Plants with the Ely plasmotype were removed from all subsequent analyses as this plasmotype confers a particularly large reduction in ϕPSII (El-Lithy et al., 2005), thus obscuring smaller contributions of the plasmotype to the heritability of photosynthetic traits.

Linear mixed models were constructed for each ϕPSII measurement using the lme4 package in R (version 4.1.0) (Bates et al., 2015). The contribution of the nucleotype, plasmotype and the nucleotype-plasmotype interaction to each phenotype were determined using equation (1), allowing the variance component to be calculated for each term, using the restricted maximum likelihood approach.

(1)

$$\underline{Y} = \underline{Nucleotype} + \underline{Plasmotype} + \underline{Nucleotype \times Plasmotype} + \underline{Sowing\ Block} + \underline{Image\ Position} + \underline{Experiment} + \underline{\varepsilon}$$

The broad sense heritability (H^2) was calculated as the sum of the genetic variance components (Nucleotype, Plasmotype and Nucleotype × Plasmotype). The contribution of each genetic component to the H^2 was summarised as a percentage of the total H^2. Phenotypes with a H^2 lower than 5% were removed for all subsequent analyses as variation in these phenotypes cannot be accurately predicted based on genotype. A threshold of 5% was also used in the analyses of Flood et al. (2020).

The additive (across nucleotype) and interaction (within nucleotype) effects of the plasmotypes were calculated by introducing Nucleotype, Plasmotype and Nucleotype × Plasmotype into the model as fixed effects. This resulted in the model given in equation (2).

(2)

$$\underline{Y} = Nucleotype + Plasmotype + Nucleotype \times Plasmotype + \underline{Sowing\ Block} + \underline{Image\ Position} + \underline{Experiment} + \underline{\varepsilon}$$

The best linear unbiased estimation (BLUE) for each plasmotype and each nucleotype-plasmotype interaction were calculated using the emmeans package in R (Lenth, 2021).

A principal component analysis (PCA) was carried out with R, using the BLUEs for each plasmotype. Further PCAs were performed for each of the four nucleotypes using the BLUEs for the nucleotype-plasmotype interactions.

Outside Tunnel Experiments

Phenotypic data was also collected for the cybrid panel grown in outdoor conditions. The set-up for these experiments is described in Lawson (2020) and Fornaguera Espinosa (unpublished).

Growth measurements, RGB measurements and chlorophyll fluorescence parameters (ϕPSII, NPQ_t, ϕNO, ϕNPQ, qI_t, qE_t) from the two experiments were compiled and outlier detection was carried out on the basis of leaf area. The mean leaf area was calculated per genotype and plants with a leaf area less than the mean – 1.5 × standard deviation were

removed. As in the Phenovator experiments, plants with the Ely plasmotype were removed from all subsequent analyses.

Statistical modelling was carried out as with the data from the Phenovator experiments, using equation (3) for the estimation of variance components and the calculation of H^2 and equation (4) for the calculation of the BLUEs. Phenotypes with a H^2 lower than 5% were removed from all subsequent analyses.

(3)

$$\underline{Y} = \underline{Nucleotype} + \underline{Plasmotype} + \underline{Nucleotype \times Plasmotype} + \underline{Block} + \underline{\varepsilon}$$

(4)

$$\underline{Y} = Nucleotype + Plasmotype + Nucleotype \times Plasmotype + \underline{Block} + \underline{\varepsilon}$$

A principal component analysis (PCA) was carried out with R, using the BLUEs for each plasmotype. Further PCAs were performed for each of the four nucleotypes using the BLUEs for the nucleotype-plasmotype interactions. Following inspection of the PCA eigenvectors, it was concluded that the growth and RGB measurements were responsible for most of the variation in the data, rather than photosynthetic traits. The PCA was thus also performed without these phenotypes, allowing for better comparison between the different phenotyping experiments, for which these RGB measurements were not made.

DEPI Experiments

Further phenotyping screens were carried out using the Dynamic Environmental Photosynthetic Imaging (DEPI) system. The DEPI system allows all of the plants to be measured simultaneously, allowing repeated measurements under fluctuating light conditions (Cruz et al., 2016). The cybrid panel was phenotyped five times in the DEPI system.

Growth measurements and chlorophyll fluorescence parameters (Fv/Fm, NPQ, NPQ_t, ϕPSII, ϕNO, ϕNPQ, qE_{SV}, qE_t, qi, qI_t, qI) from five experiments were compiled and outlier detection was carried out on the dark-adapted yield of PSII photochemistry (Fv/Fm). The mean Fv/Fm for all plants was calculated and plants with an Fv/Fm less than the mean – 1.5 × standard deviation were removed. As in the Phenovator and tunnel experiments, plants with the Ely plasmotype were removed from all subsequent analyses.

Statistical modelling was carried out as with the data from the Phenovator experiments, using equation (5) for the estimation of variance components and the calculation of H^2 and equation (6) for the calculation of the BLUEs. Phenotypes with a H^2 lower than 5% were removed for all subsequent analyses.

(5)

$$\underline{Y} = \underline{Nucleotype} + \underline{Plasmotype} + \underline{Nucleotype \times Plasmotype} + \underline{Camera} + \underline{Experiment} + \underline{\varepsilon}$$

(6)

$$\underline{Y} = Nucleotype + Plasmotype + Nucleotype \times Plasmotype + \underline{Camera} + \underline{Experiment} + \underline{\varepsilon}$$

A principal component analysis (PCA) was carried out with R, using the BLUEs for each plasmotype. Further PCAs was performed for each of the four nucleotypes using the BLUEs for the nucleotype-plasmotype interactions. The growth measurements were excluded from these PCAs.

All plots were produced using ggplot2 in R (Wickham, 2016).

Genomic Data Analysis

Genotype Confirmation

Genomic analysis of the cybrids largely followed the workflow used by Flood et al. (2020). All of the cybrids were sequenced at 8X coverage using Illumina paired-end sequencing. The reads were trimmed using Cutadapt (Martin, 2011), removing the adaptor sequences and bases from the 5' and 3' ends with a Phred quality score below 20. Reads shorter than 75 bp were also removed. The reads were then mapped to the *A. thaliana* Col-0 reference genome, TAIR10.1, using speedseq and the Burrow-Wheeler aligner, BWA-MEM (Chiang et al., 2015; Li, 2013). The alignments were sorted and indexed using Samtools (Li et al., 2009). Duplicate reads were marked using GATK (Van der Auwera and O'Connor, 2020). Variant calling was performed using freebayes and the resulting variant call format (VCF) file was separated into three VCF files – one for each of the nuclear, mitochondrial and plastid genomes (Garrison and Marth, 2012). Distance matrices were calculated for each of the three genomes using PLINK (Chang et al., 2015), and the ape package in R was used to produce neighbour-joining trees (Paradis and Schliep, 2019). The cybrids with incorrect genotypes were removed from all statistical analyses.

The effect sizes of single nucleotide polymorphisms (SNPs) and indels in the mitochondrial and plastid genomes of 1544 *A. thaliana* accessions were predicted using SnpEff (Cingolani et al., 2012). This database was subsequently filtered to include only the 60 progenitor lines of the cybrid panel. Unique polymorphisms, occurring only in a genotype of interest, were identified from this database. Polymorphisms with a 'moderate' or 'high' score from SnpEff were considered to be potential causal polymorphisms underlying the cybrid's phenotype.

Plasmotype Association Studies

A Genome-Wide Association Studies (GWAS) method was used to test the association between genetic variants in the plastid genome and the phenotypes measured in the DEPI experiments. Due to the lack of recombination within the plastid genome, the interpretation of the GWAS results deviates from a standard GWAS performed on the nuclear genome. Unless genetic variants are shared between different plasmotypes, all of the SNPs within a plasmotype will be associated with the phenotype with the same probability. Thus, to avoid confusion with a nuclear GWAS, this test has been named a Plasmotype Association Study.

To carry out the Plasmotype Association Study, the BLUEs for each plasmotype from the DEPI phenotyping experiments were used alongside the VCF file for the plastid genome on the basis of 40X sequencing coverage. Binary PED files were produced from the plastid VCF file using PLINK (Chang et al., 2015). The BLUEs for each of the 1986 phenotypes were added manually to the .fam file. GEMMA was used to produce a centred relationship matrix for the plastid genomes of the 60 progenitor lines and to run a univariate association study (Zhou and Stephens, 2012). Significance thresholds at $\alpha = 0.05$ were determined in a number of different ways. Firstly, 1000 permutations of the plasmotype association study for ϕPSII at 55.0136 hours were produced. The maximum $-\log_{10}(p)$ value for each permutation was extracted and the observed LOD scores were ordered. The observed LOD score at the 950th permutation was taken as the threshold. Bonferroni thresholds were also calculated for the GWAS, on the basis of the total number of SNPs in the plastid genome ($p = \alpha/\sqrt{\#}$ of SNPs) and on the basis of the total number of plasmotypes ($p = \alpha/\sqrt{\#}$ of plasmotypes). The latter method was considered to be acceptable as no recombination takes place within the plastid genome and so all SNPs are linked.

Phenotyping of the Can-0 Plasmotype

Plant Material

A. thaliana cybrids with the Can-0 and Col-0 plasmotypes in the Bur, Col-0, Cvi, and Tanz nucleotypes were used. The T-DNA insertion lines *ndhM* and *ndhO*, mutants in the nuclear NDH genes *NdhM* and *NdhO*, were used as negative controls.

Seeds were pre-sown in petri dishes lined with damp filter paper and stored in a transport box lined with damp paper towel at 4°C in the dark for five days. Germination was triggered by moving the seeds to a climate chamber with 16 hours per day of light and a temperature of 22°C for one day. The seeds were then sown in 12 blocks of 20 plants in a randomised block design, produced using the agricolae package in R (De Mendiburu, 2021). An equal number of replicates for each genotype was sown. Each seed was sown onto a rockwool block measuring 4 × 4 × 4 cm soaked in Hyponex solution, and covered with a plastic square with a hole to allow the seedling to grow. Each plastic cover was secured to the rockwool block using a colourless pipette tip. The plants were grown in a climate chamber under 250 µmol/m^2/s light, 10 hours of light, 14 hours of dark, temperatures of 20°C during the day and 18°C during the night, and 70% humidity. The plants were watered twice a week with Hyponex.

Phenotyping

Phenotyping was carried out 19 days after sowing using the PlantScreen™ SC System. Plants were first placed in an opaque box for 20 to 30 minutes, allowing them to dark adapt. The plants were then placed in the PlantScreen™ SC System, where they were allowed to dark adapt for a further 5 minutes before a saturating pulse was applied. Fluctuations of high light (1000 µmol/m^2/s) and low light (200 µmol/m^2/s), each lasting 5 minutes, were then applied. A saturating pulse was applied at the end of each high light and low light period. The fluctuations lasted a total of 20 minutes. Following the last low light period, the lights were switched off and the post-illumination fluorescence rise (PIFR) was measured for 1 minute. The light intensity was then increased again to 1000 µmol/m^2/s for 5 minutes, after which the PIFR was measured again for 1 minute. Chlorophyll fluorescence measurements were taken regularly throughout this protocol. A schematic of the chlorophyll fluorescence measurements is shown in Fig 2. These measurements allowed the following photosynthetic parameters to be calculated for each of the high light and low light periods: Fv/Fm, ϕPSII, ϕNPQ and ϕNO (equations 7-10) (Klughammer and Schreiber, 2008).

(7)

$$\frac{Fv}{Fm} = \frac{Fm - Fo}{Fm}$$

(8)

$$\phi PSII = \frac{Fm' - F'}{Fm'}$$

(9)

$$\phi NPQ = \frac{F'}{Fm'} - \frac{F'}{Fm}$$

(10)

$$\phi NO = \frac{F'}{Fm}$$

The NADH dehydrogenase-like (NDH) complex activity was determined according to the method used by Otani, Yamamoto and Shikanai (2017), shown in equation 11. Fo' was taken as the minimum value of the first 10 fluorescence measurements in the dark, and Fp was taken as the maximum value of the next 20 fluorescence measurements in the dark.

(11)

$$NDH\ activity = \frac{Fp - Fo'}{Fm}$$

The following day (20 days after sowing) the plants were exposed to a full day of fluctuating light, with a minimum light intensity of 200 µmol/m^2/s and a maximum light intensity of 1000 µmol/m^2/s. They were then phenotyped again in the PlantScreen™ SC System, using the same protocol as before.

For each dataset (before and after the day of fluctuating light) an outlier detection was carried out on the basis of Fv/Fm. The mean Fv/Fm for all plants was calculated and plants with an Fv/Fm less than the mean – 1.5*standard deviation were removed. For each photosynthetic parameter calculated, a linear mixed model was used to estimate the BLUE for each genotype (equation 12), using the emmeans package in R (Lenth et al., 2021).

(12)

$$\underline{Y} = Genotype + \underline{Block} + \underline{\varepsilon}$$

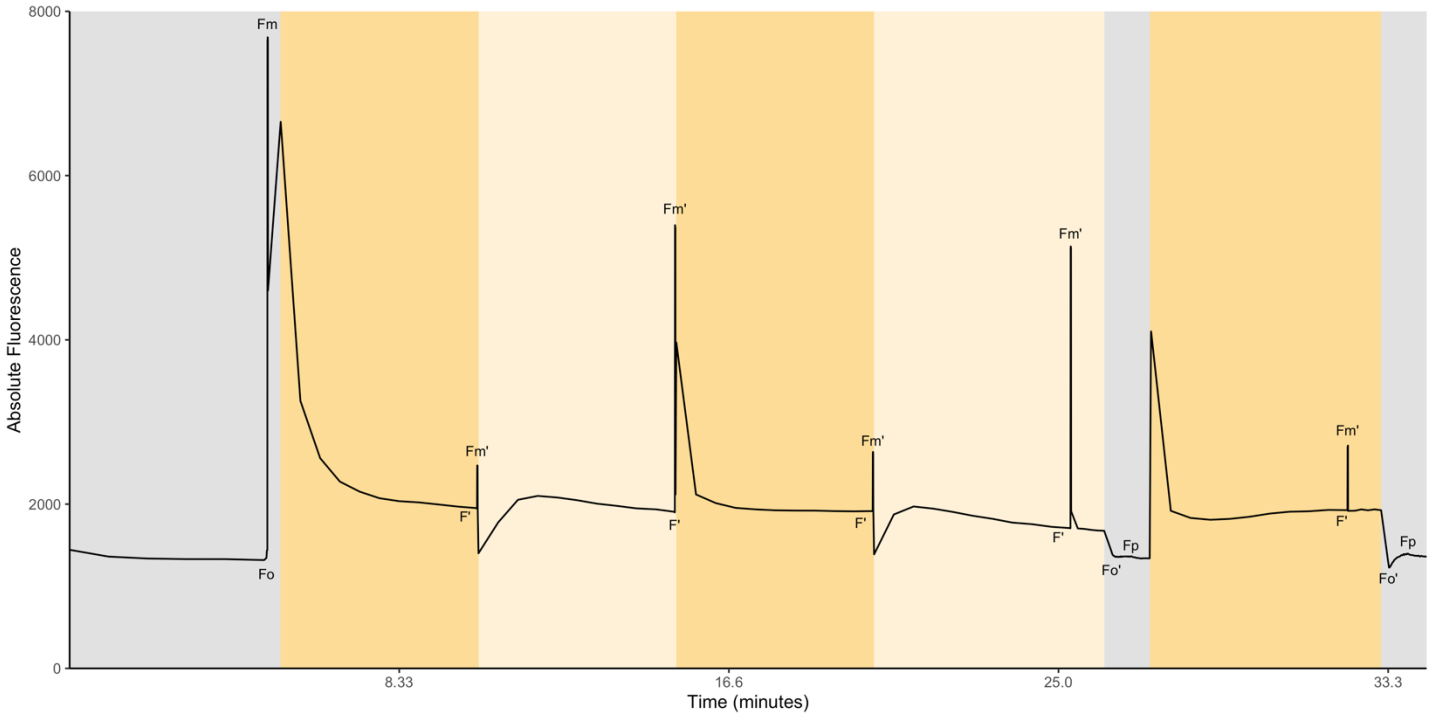


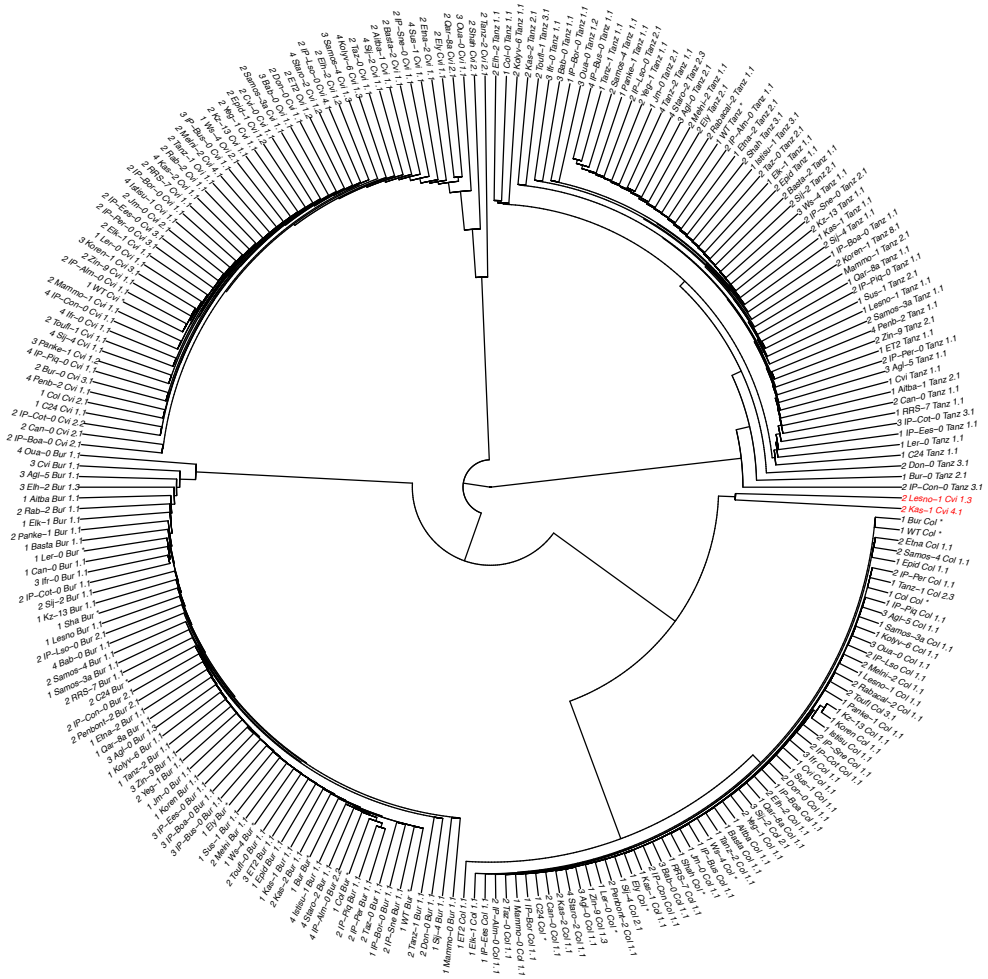
Figure 2: Chlorophyll fluorescence measurements made in the PlantScreen™ SC System. *The measuring protocol consists of 5 minutes dark adaptation (grey), followed by fluctuations of 1000 µmol/m²/s (dark yellow) and 200 µmol/m²/s (light yellow). Post-illumination fluorescence measurements were taken in the dark (grey) after the last low light period and the last high light period. Labels show the measurements used to estimate Fo, Fm, F', Fm', Fo' and Fp.*

Results

Confirmation of cybrid genotypes

To determine whether the cybrids had the expected genotypes, neighbour-joining analyses were carried out on the cybrids' genome sequences (Fig 3). The resulting neighbour-joining cladograms were inspected to identify cybrids that did not match their expected genotype. The cladogram on the basis of the nuclear genomes was used to identify cybrids that did not cluster with their expected nucleotype (Bur, Col-0, Cvi, or Tanz) (Fig 3A). Lesno-1_Cvi and Kas-1_Cvi did not cluster with other cybrids with the Cvi nucleotype; rather they are most closely related to the cybrids with the Col-0 nucleotype. This could be due to incorrect crosses during the production of the cybrids or due to incomplete genome elimination during the haploid induction step. The cladogram on the basis of the plastid genomes was used to identify cybrids that did not cluster with their expected plasmotype (Fig 3B). Rab-2_Bur, Panke-1_Bur, IP-Bor-0_Col-0, IP-Per-0_Col-0, IP-Piq-0_Col-0, Lesno-1_Cvi and Elk-1_Tanz all clustered with different plastid genomes than expected. This is likely due to crosses made with the incorrect maternal parent (the plasmotype donor). These eight cybrids were removed from all subsequent analyses.

A



B

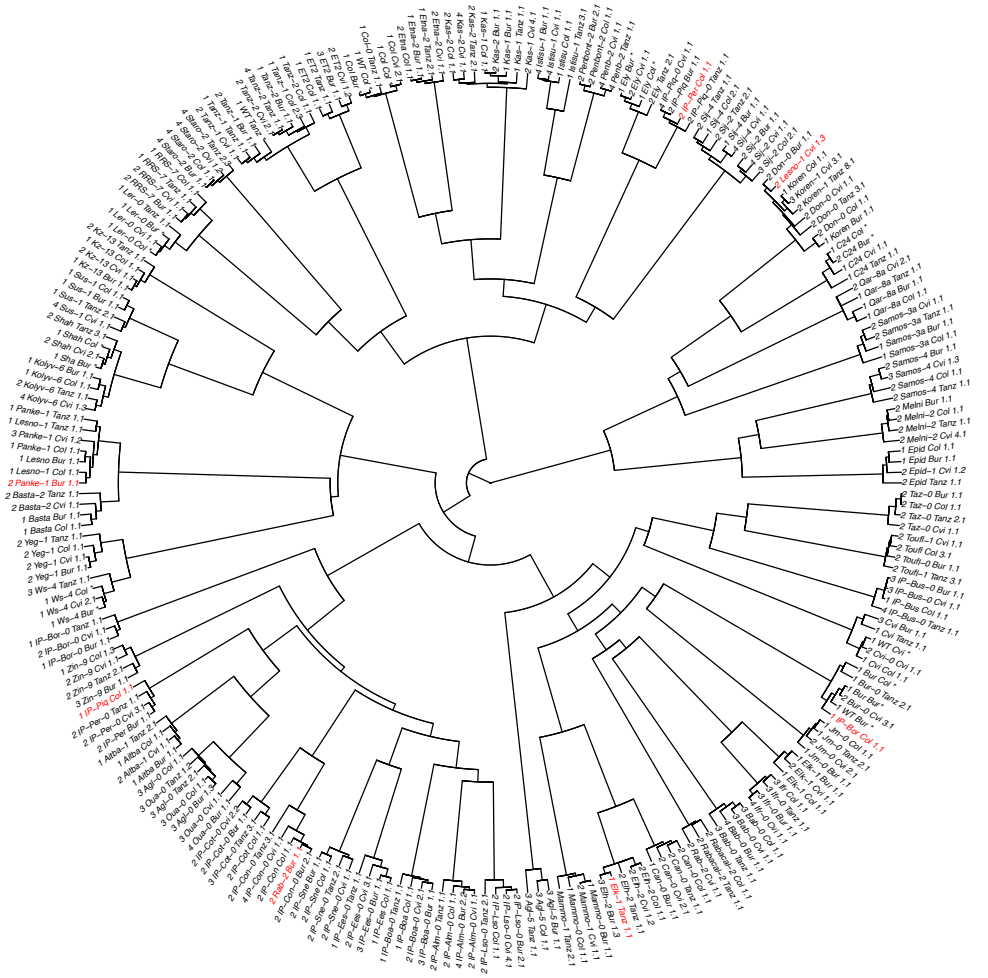


Figure 3: Neighbour-joining of cybrids to reveal incorrect genotypes.

(A) Neighbour-joining cladogram based on the nuclear genomes of the cybrid panel. Cybrids that clustered with a nucleotype that does not match their expected nucleotype are shown in red. (B) Neighbour-joining cladogram based on the plastid genomes of the cybrid panel. Cybrids that clustered with a plasmotype that does not match their expected plasmotype are shown in red.

Dissection of the heritability of photosynthetic traits

The contribution of the cytoplasmic genomes to the heritability of photosynthetic traits was investigated by screening the cybrid population over the course of six days in the Dynamic Environmental Photosynthetic Imaging (DEPI) system. The traits measured include the efficiency of Photosystem II (ϕPSII), non-photochemical quenching (NPQ) and its constituents (ϕNPQ, ϕNO, qE and qI). The plants received alternating days of stable and fluctuating light, and for the last two days the temperature was decreased (Fig 4C). The broad-sense heritability (H^2), the proportion of the variation explained by genotype, as well as the contribution of the plasmotype and the nucleotype-plasmotype interaction to the H^2, was estimated for each photosynthetic trait (Fig 4A-B). The mean contribution of the plasmotype and the nucleotype-plasmotype interaction was relatively low with 1.67% and 3.21% respectively. However, for particular phenotypes, the contribution of the plasmotype or the nucleotype-plasmotype interaction was considerably higher. The plasmotype accounted for up to 26.3% of H^2 (qE_{SV} at 7.0136 hours), whereas the nucleotype-plasmotype interaction accounted for up to 68.8% of H^2 (qE_{SV} at 56.0153 hours).

The biggest contributions of the plasmotype and nucleotype-plasmotype interaction were made to the H^2 of qE_{SV}, the rapid component of NPQ as calculated by the Stern-Volmer approach (Fig 4A-B). This indicates that the cytoplasmic genomes (and their interaction with the nuclear genome) are particularly relevant for dynamic components of photosynthesis. The plasmotype and nucleotype-plasmotype interaction make considerable contributions to the H^2 of qE_{SV} across the six days. For the plasmotype this ranges from a maximum of 15.73% on day 4 to a maximum of 26.3% on day 1 and for the nucleotype this ranges from a maximum of 25.25% on day 5 to a maximum of 68.8% on day 3. However, the overall H^2 of both qE parameters (qE_{SV} and qE_t) showed a different trend, increasing under fluctuating light conditions (day 1 to day 2, and day 3 to day 4) (Supplementary Fig 1A). For example, the average H^2 for qE_{SV} increased from 15% on day 3 to 44% on day 4. In contrast, the average H^2 for ϕNO, a measure of basal NPQ, increased from 49% on day 3 to 64% on day

4 (Supplementary Fig 1B). This suggests that phenotyping under fluctuating environmental conditions can reveal additional contributions of genetic variation to variation in photosynthetic traits, particularly those that are dynamic, such as qE.

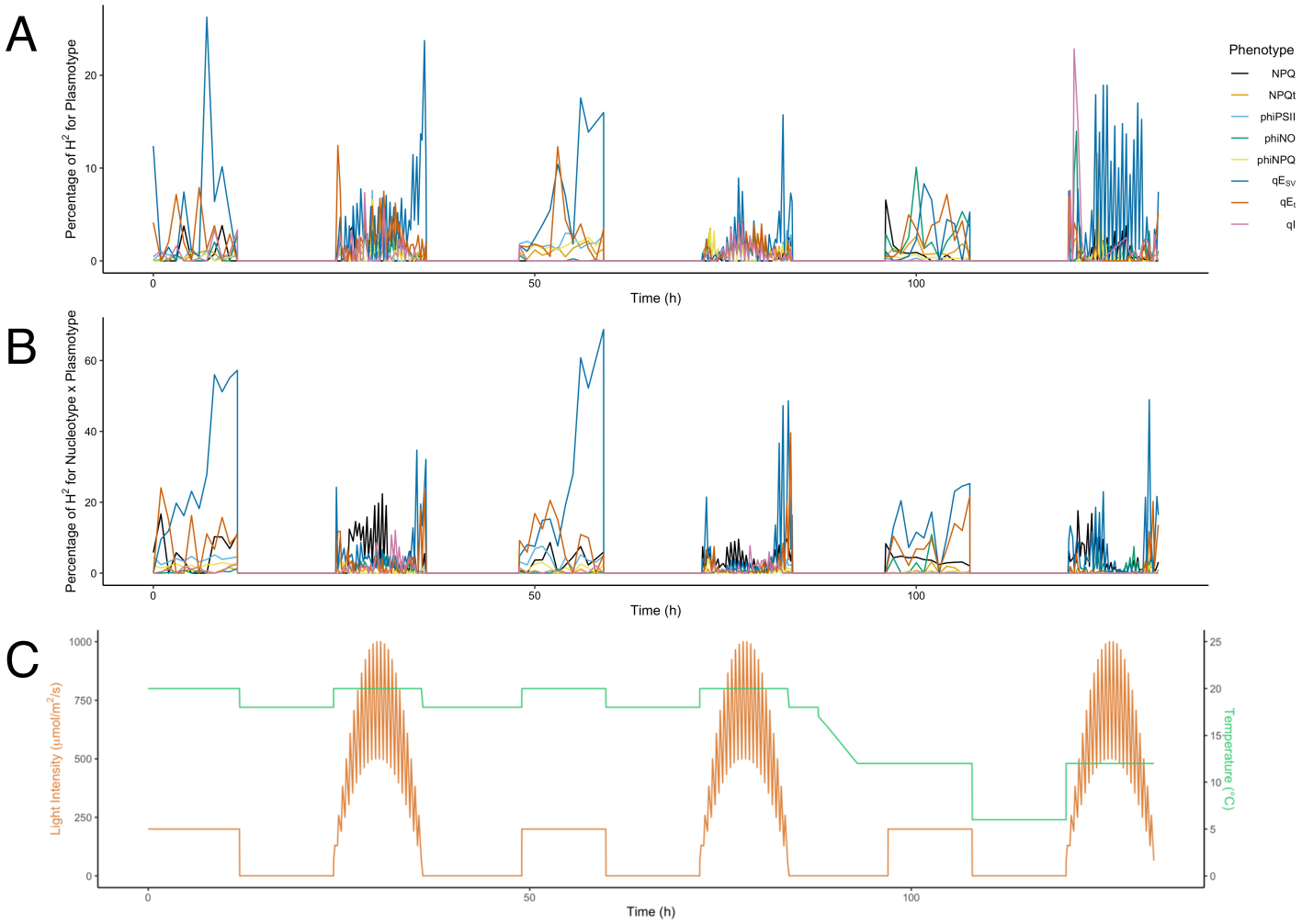


Figure 4: The contribution of the plasmotype and the nucleotype-plasmotype interactions to broad sense heritability (H^2) of photosynthetic phenotypes varies over time.

(A) Percentage of the H^2 for 8 phenotypes attributed to the plasmotype over the course of 6 days in the DEPI system.

(B) Percentage of the H^2 for 8 phenotypes attributed to the interaction between the nucleotype and the plasmotype over the course of 6 days in the DEPI system.

(C) Light intensity and temperature over the course of the 6 days. This consisted of alternating days of stable light intensity (200 µmol/m²/s) and fluctuating light intensity (up to 1000 µmol/m²/s). The temperature was decreased from 20°C/18°C to 12°C/6°C for days 5 and 6.

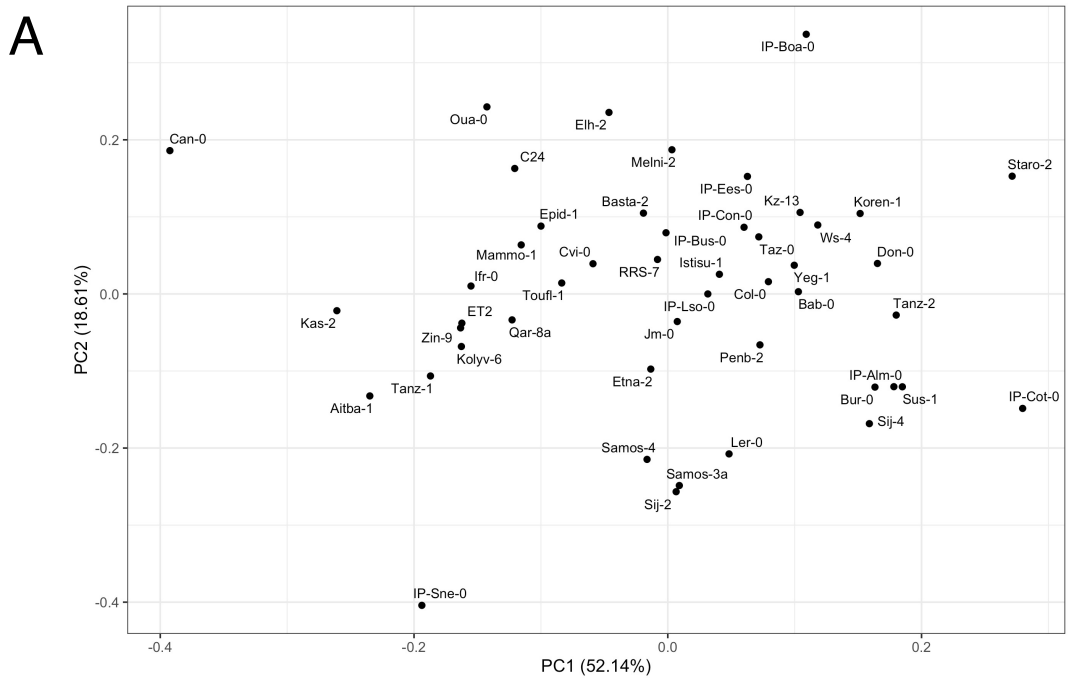
Identification of plasmotypes of interest by Principal Component Analysis

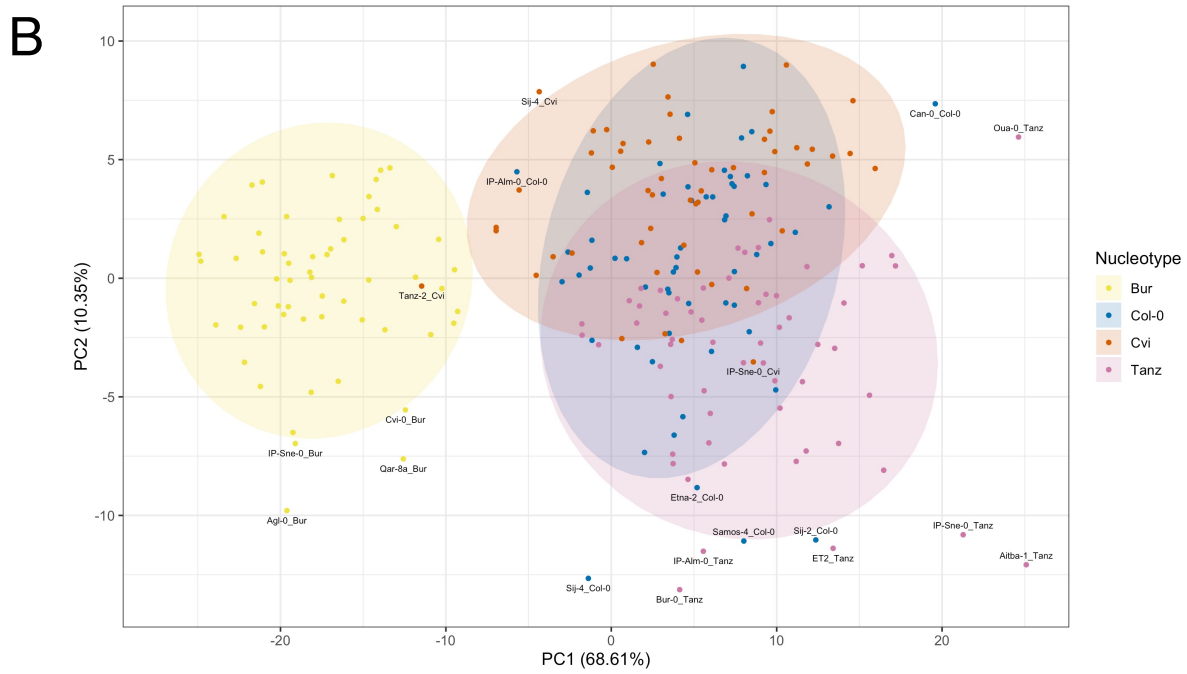
Phenotypic variation between the plasmotypes was explored in more detail using a principal component analysis (Fig 5A), using the BLUEs for each plasmotype for the 1707 phenotypes measured in the DEPI system. The first principal component explains 52.14% of the variation in the data and the second principal component explains 18.61% of the variation in the data. Therefore, the PCA biplot (Fig 5A) represents over 70% of the variation in the dataset, suggesting that this method can be used for dimensionality reduction of this dataset. Plasmotypes that were found on the outside of the cloud of points were considered potential candidates for further investigation, as these plasmotypes are the most phenotypically different from the 'average' plasmotype. The Can-0, IP-Sne-0, IP-Boa-0, IP-Cot-0 and Staro-2 plasmotypes took the most extreme values along the first two principal components, identifying them as plasmotypes of interest.

The loadings of the phenotypic measurements were also plotted to identify which measurements most strongly contributed to the two principal components (Supplementary Fig 2). This indicated that qE_t and NPQ_t under fluctuating light conditions explained most of the variation in the first principal component, and qE_t and NPQ_t under low temperatures explained most of the variation in the second principal component. Thus, Can-0, IP-Cot-0 and Staro-2 are expected to have the most extreme phenotypes under fluctuating light conditions, whereas IP-Sne-0 and IP-Boa-0 are expected to have the most extreme phenotypes under low temperature conditions.

As the nucleotype-plasmotype interaction was found to make widespread contributions to the heritability of photosynthetic traits (Fig 4B), a further PCA was carried out on the BLUEs for each nucleotype-plasmotype combination (Fig 5B). This analysis indicates that the nucleotype is also important to consider when examining phenotypic differences between the cybrids; the Bur nucleotype in particular separates from the other three (Col-0, Cvi and Tanz) along the first principal component. Moreover, the outlying nucleotype-plasmotype combinations (those that are outside of the 95% confidence level for the multivariate t-distribution of the given nucleotype) are different for the four nucleotypes. For example, the Oua plasmotype is only an outlier when combined with the Tanz nucleotype, whereas the Agl-0 plasmotype is only an outlier when combined with the Bur nucleotype. Furthermore, when looking at a single plasmotype, in this case Can-0 (Fig 5C), its position within each nucleotype distribution can vary. Can-0 is an outlier when combined with the Col-0 nucleotype, and is found on the edge of the multivariate t-distributions of the Bur and Cvi

nucleotypes. For the Tanz nucleotype, on the other hand, Can-0 is found more towards the middle of the distribution. This emphasises the role of epistatic interactions between the cytoplasmic and nuclear genomes in determining photosynthetic phenotypes.


A
IP-Boa-0
Oua-0
Elh-2
Can-0
C24
Melni-2
Staro-2
IP-Ees-0
Basta-2
Kz-13
Koren-1
Epid-1
IP-Con-0
IP-Bus-0
Ws-4
Mammo-1
Cvi-0
Taz-0
Don-0
Istisu-1
Ifr-0
RRS-7
Yeg-1
Toufl-1
Col-0
Bab-0
Tanz-2
ET2
IP-Lso-0
Kas-2
Qar-8a
Jm-0
Zin-9
Penb-2
Kolyv-6
IP-Alm-0
Etna-2
Tanz-1
Bur-0
Sus-1
IP-Cot-0
Aitba-1
Sij-4
Samos-4
Ler-0
Samos-3a
Sij-2
IP-Sne-0
PC2 (18.61%)
PC1 (52.14%)
0.2
0.0
-0.2
-0.4


B
Sij-4_Cvi
Can-0_Col-0
Oua-0_Tanz
IP-Alm-0_Col-0
Tanz-2_Cvi
IP-Sne-0_Cvi
Cvi-0_Bur
IP-Sne-0_Bur
Qar-8a_Bur
Etna-2_Col-0
Agl-0_Bur
Samos-4_Col-0
Sij-2_Col-0
IP-Sne-0_Tanz
IP-Alm-0_Tanz
ET2_Tanz
Aitba-1_Tanz
Sij-4_Col-0
Bur-0_Tanz
Nucleotype
Bur
Col-0
Cvi
Tanz
PC2 (10.35%)
PC1 (68.61%)
10
5
0
-5
-10
-20
-10
0
10
20

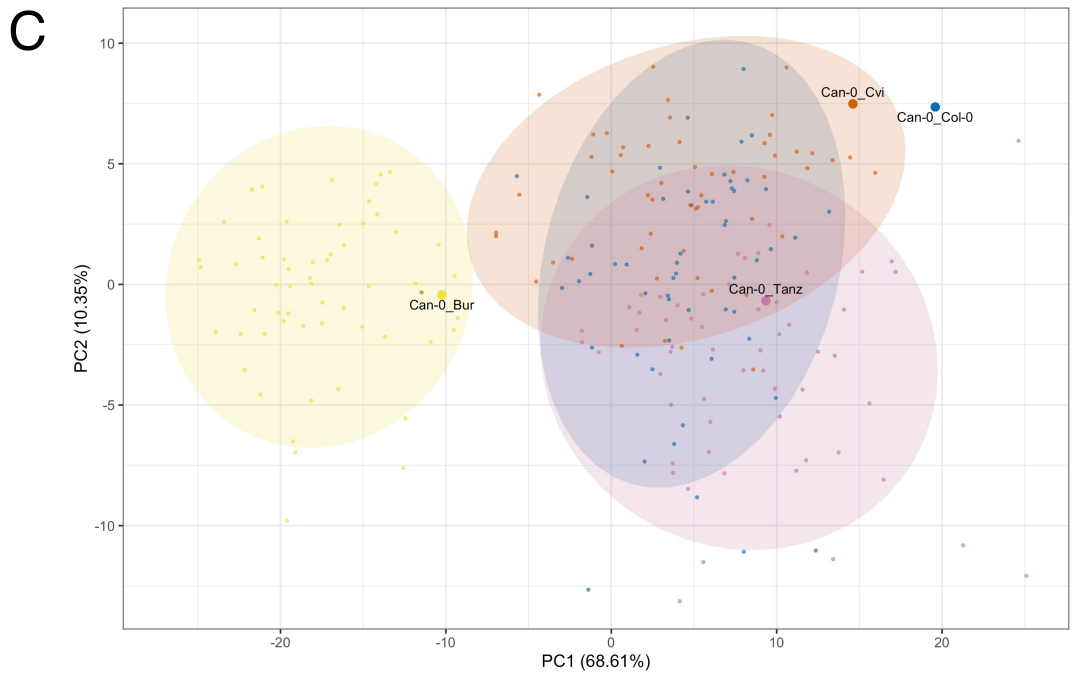

C
Can-0_Cvi
Can-0_Col-0
Can-0_Tanz
Can-0_Bur
PC2 (10.35%)
PC1 (68.61%)
10
5
0
-5
-10
-20
-10
0
10
20

Figure 5: Principal Component Analysis (PCA) reveals variation between plasmotypes and nucleotype-plasmotype combinations.

(A) PCA biplot of the BLUEs of each plasmotype for the 1707 phenotypes measured in the DEPI experiment. Each point is labelled with the name of the plasmotype donor.

(B) PCA biplot of the BLUEs of different plasmotype-nucleotype combinations for the 1707 phenotypes measured in the DEPI experiment. Each ellipse indicates the 95% confidence level for a multivariate t-distribution for the BLUEs of a given nucleotype – Bur (yellow), Col-0 (blue), Cvi (red), and Tanz (pink). Points outside of the 95% confidence interval for their corresponding nucleotype are labelled as Plasmotype_Nucleotype.

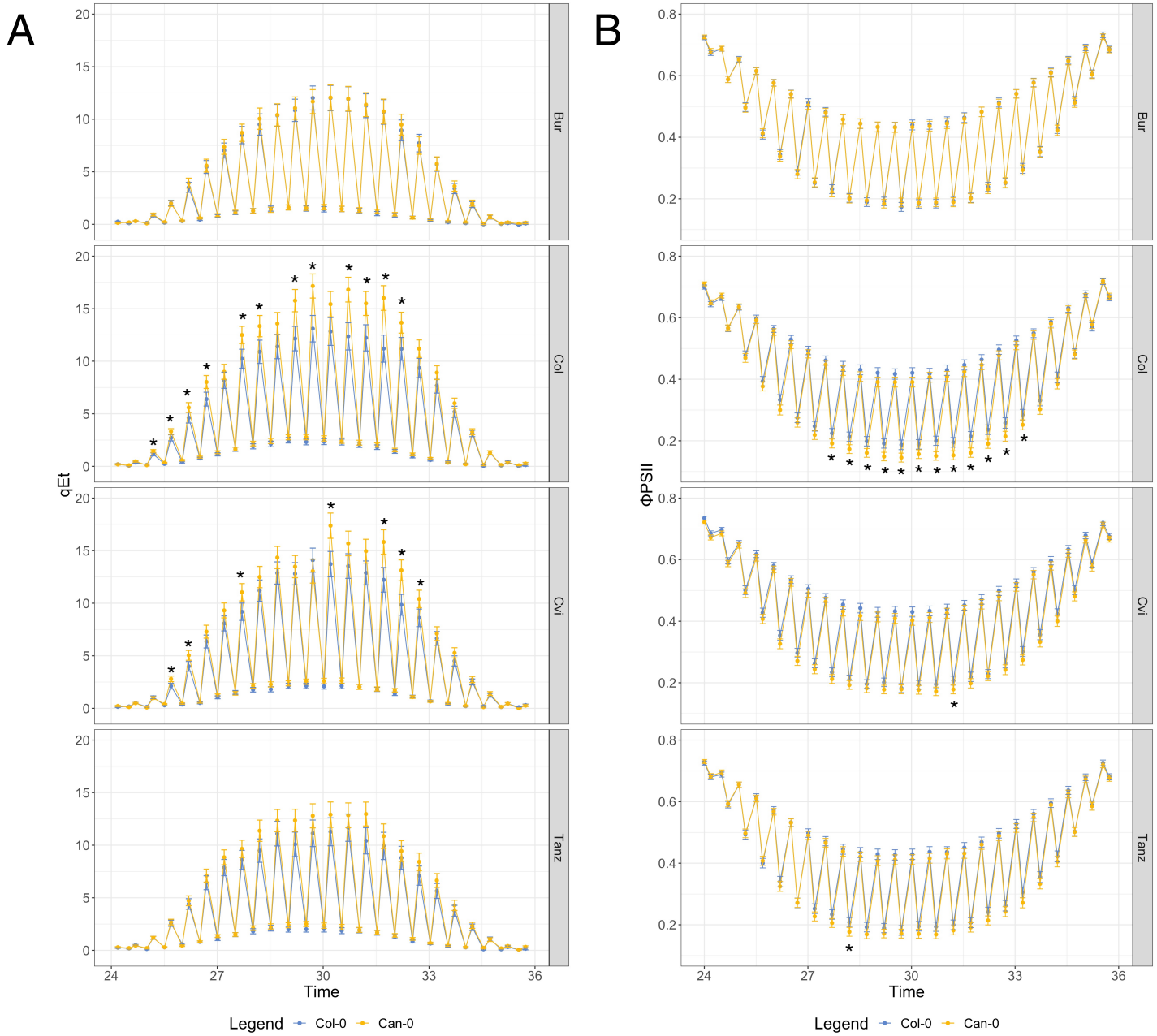
(C) PCA biplot for the BLUEs of different plasmotype-nucleotype combinations for the 1707 phenotypes measured in the DEPI experiment. Each ellipse indicates the 95% confidence level for a multivariate t-distribution for the BLUEs of a given nucleotype. The four cybrids with the Can-0 plasmotype are labelled.

The phenotype of the Can-0 plasmotype in more detail

Following the identification of Can-0 as a plasmotype of interest from the PCA, its phenotype was investigated in more detail. Time series were produced for several photosynthetic parameters measured over the course of the DEPI experiment, using the BLUEs for the Can-0 plasmotype (Fig 6A-B). On the second day, during which light fluctuations were applied (Fig 4C), the Can-0 plasmotype was associated with an increased qE_t when compared to the Col-0 plasmotype in the Col-0 nucleotype background (Fig 6A). qE_t is a measure of the rapid component of NPQ_t (Tietz et al., 2017). Significant differences in qE_t were observed between Can-0_Col-0 and Col-0_Col-0 ($\alpha = 0.05$) at several time points when the plants were exposed to high light intensities. Can-0_Cvi also had a significantly higher qE_t than Col-0_Cvi at several time points, also under high light intensity.

This higher qE_t was associated with a reduced efficiency of PSII (ϕPSII) for the Can-0 plasmotype in the Col-0 nucleotype background (Fig 6B). This suggests that the Can-0 plasmotype has an inhibitory effect on photosynthesis and that this effect is particularly evident in combination with the Col-0 nucleotype. In the Cvi nucleotype, however, the Can-0 plasmotype only had a significantly lower ϕPSII at one time point, suggesting that the effect on qE_t in the Cvi nucleotype is not large enough to substantially affect ϕPSII.

Furthermore, the measurements from a single time point indicate that the Can-0 plasmotype affects the partitioning of excitation energy in PSII between photochemistry (ϕPSII), inducible non-photochemical dissipation (ϕNPQ), and basal non-photochemical dissipation (ϕNO) (Fig 6C). Compared to the Col-0 plasmotype, ϕNPQ is increased with the Can-0 plasmotype while ϕPSII and ϕNO are decreased. Pairwise comparisons between the additive effects of the Can-0 plasmotype and the Col-0 plasmotype show that all three of these differences are significant ($\alpha = 0.05$). Overall, this indicates that the Can-0 plasmotype results in a stronger induction of NPQ in response to high light.


A
B
qEt
ΦPSII
Bur
Col
Cvi
Tanz
Time
Legend
Col-0
Can-0

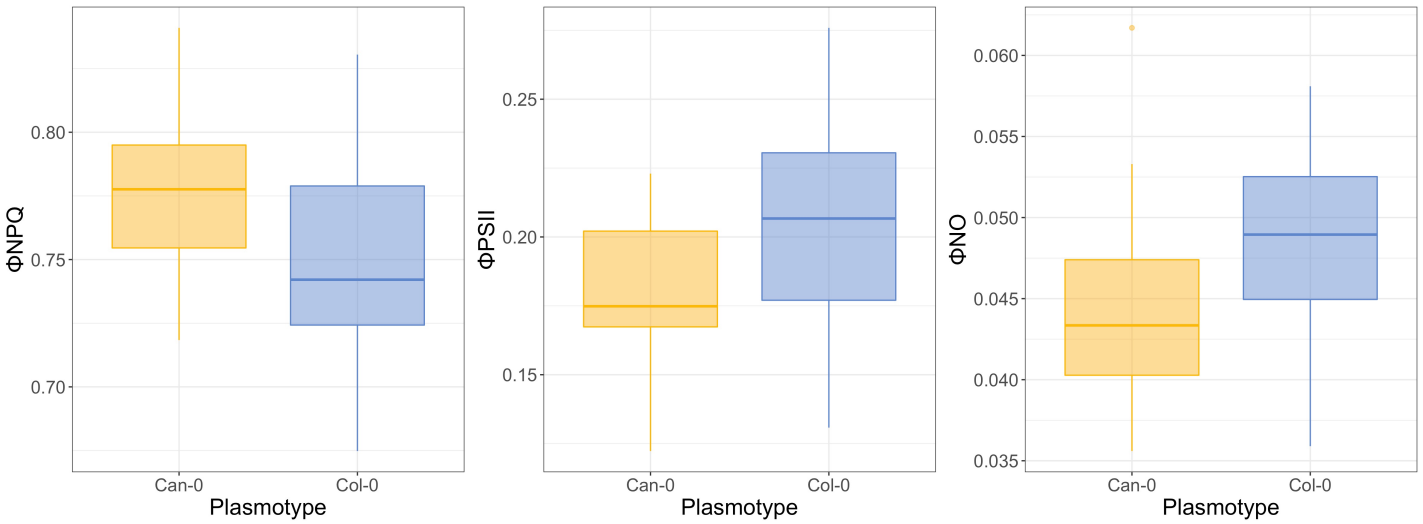

C
ΦNPQ
ΦPSII
ΦNO
Can-0
Col-0
Plasmotype

Figure 6: The Can-0 plasmotype affects NPQ and ϕPSII in the Col-0 and Cvi nucleotypes.
(A) Estimated qE_t ± SE for the Can-0 (yellow) and Col-0 (blue) plasmotypes in the four nucleotype backgrounds (Bur, Col-0, Cvi and Tanz) during the first day of fluctuating light in the DEPI system. Stars indicate a significant pairwise difference between the Can-0 plasmotype and the Col-0 plasmotype (α = 0.05).
(B) Estimated ϕPSII ± SE for the Can-0 (yellow) and Col-0 (blue) plasmotypes in the four nucleotype backgrounds during the first day of fluctuating light in the DEPI system. Stars indicate a significant pairwise difference between the Can-0 plasmotype and the Col-0 plasmotype (α = 0.05).
(C) Distribution of ϕNPQ, ϕPSII and ϕNO for the cybrids with the Can-0 (yellow) and Col-0 (blue) plasmotypes under high light, measured at 31.2125 hours in the DEPI system. n=16 for the Can-0 plasmotype; n=14 for the Col-0 plasmotype.

Identification of plasmotypes of interest by Plasmotype Association Studies
To identify single nucleotide polymorphisms (SNPs) associated with the phenotypes measured in the DEPI system, Genome-Wide Associated Studies (GWAS) were carried out (Fig 7). Unlike a conventional GWAS, a GWAS on the basis of the plastid (or mitochondrial) genome cannot be used to identify individual causal SNPs as no recombination takes place within these organellar genomes. As such, a group of SNPs exclusively found in a particular plasmotype will all be associated with the phenotype, even if only one SNP is causal to the phenotype. Thus, to distinguish it from a nuclear GWAS, we refer to this method as a Plasmotype Association Study. This method was first tested on a timepoint where the increased ϕPSII of the Bur-0 plasmotype (documented in Flood et al., 2020) could be seen (Fig 7A-B). Three different significance thresholds (α = 0.05) were determined for this GWAS: on the basis of 1000 permutations of the GWAS (Fig 7A), a Bonferroni threshold based on the number of SNPs in the plastid genome, and a Bonferroni threshold based on the number of plasmotypes. The unique SNPs associated with the Bur-0 plasmotype passed both Bonferroni thresholds but not the threshold determined by the permutation method, suggesting that this method may be too strict when applied to the plastid genome.

Further Plasmotype Association Studies were carried out on ϕPSII (Fig 7C) and ϕNPQ (Fig 7D) measurements taken during fluctuating light conditions. The two Bonferroni thresholds were applied to test for significance. In both cases, unique SNPs associated with the Can-0 plasmotype had the highest $-\log_{10}(p)$ values and passed both Bonferroni thresholds. The

Can-0 plasmotype was also identified through the PCA (Fig 5), suggesting that Plasmotype Association Studies and PCA can be used as complementary methods to identify plasmotypes of interest from the cybrid panel.

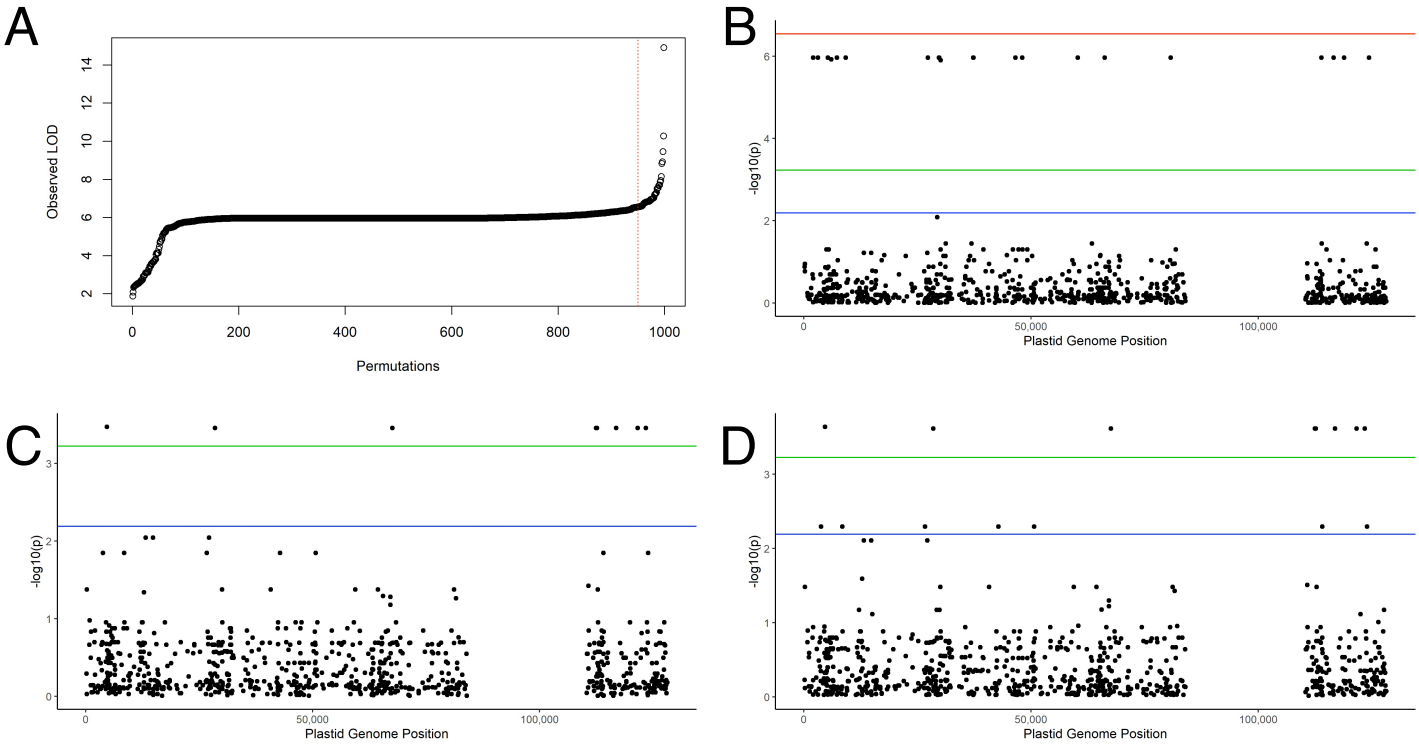


Figure 7: Plasmotype Association Studies can be used to identify plasmotypes of interest.
(A) The observed LOD scores for 1000 permutations of the Plasmotype Association Study for ϕPSII at 55.0136 hours. The red line indicates the position of the 950th permutation.
(B) Plasmotype Association Study on the BLUEs for each plasmotype for ϕPSII at 55.0136 hours, on the basis of the plastid genome. The single nucleotide polymorphisms (SNPs) with the highest -log$_{10}$(p) values correspond to the unique SNPs in the Bur-0 plasmotype. The -log$_{10}$(p) threshold on the basis of the 1000 permutations is shown in red.
(C) Plasmotype Association Study on the BLUEs for each plasmotype for ϕPSII at 31.2125 hours, on the basis of the plastid genome. The single nucleotide polymorphisms (SNPs) with the highest -log$_{10}$(p) values correspond to the unique SNPs in the Can-0 plasmotype.
(D) Plasmotype Association Study on the BLUEs for each plasmotype for ϕNPQ at 31.2125 hours, on the basis of the plastid genome. The single nucleotide polymorphisms (SNPs) with the highest -log$_{10}$(p) values correspond to the unique SNPs in the Bur-0 plasmotype.
The Bonferroni threshold (α = 0.05) based on the total number plastid single nucleotide polymorphisms (SNPs) is shown in green. The Bonferroni threshold (α = 0.05) based on the total number of plasmotypes is shown in blue.

Genetic polymorphisms in the Can-0 plasmotype

As the SNPs within a given plasmotype are linked, and so cannot be distinguished on the basis of their association with a phenotype, the nature of the SNPs must be examined in more detail. Three unique SNPs (not found in any of the other plasmotypes of the cybrid panel) with a 'moderate' or 'high' predicted effect were identified in the plastid genome of the Can-0 accession (Table 1). Two of these SNPs are in genes encoding subunits of the plastid NADH dehydrogenase-like (NDH) complex, NdhF and NdhD. The NDH complex contributes to cyclic electron transport around photosystem I (PSI) and pumps protons into the thylakoid lumen, increasing the ΔpH across the thylakoid membrane (Shikanai, 2014). The third SNP is in *rps15*, encoding a component of the small subunit of the plastid ribosome (Tiller and Bock, 2014). All three SNPs are missense mutations, although the isoleucine to leucine mutation in NdhD is very conservative. The serine to asparagine mutation in NdhF and the arginine to threonine mutation in Rps15 are both moderately conservative (Grantham, 1974). On the basis of these results, the causal SNP of the Can-0 plasmotype phenotype could not be conclusively identified.

Differences in the mitochondrial genome of Can-0 were also examined. No unique large deletions were observed in the mitochondrial genome and there were no unique SNPs with a 'moderate' or 'high' predicted effect. This suggests that the phenotype associated with the Can-0 plasmotype is due to its plastid genome, rather than its mitochondrial genome.

Position	Gene	Protein Function	Reference Amino Acid	Alternative Amino Acid	SnpEff Prediction
112442	*ndhF*	Subunit of the NDH complex[1]	Serine	Asparagine	Moderate
116924	*ndhD*	Subunit of the NDH complex[1]	Isoleucine	Leucine	Moderate
123438	*rps15*	Component of the small subunit of the plastid ribosome[2]	Arginine	Threonine	Moderate

Table 1: Unique single nucleotide polymorphisms (SNPs) in the plastid genome of the Can-0 accession.

[1] *Shikanai (2016)*

[2] *Tiller and Bock (2014)*

ϕPSII, ϕNPQ and NDH activity of the Can-0 plasmotype

Further phenotyping of cybrids with the Can-0 plasmotype was carried out to better understand its effect on the dynamic photosynthetic phenotypes characterised in the DEPI system (Fig 8). The effect of the Can-0 plasmotype was observed more strongly after one day of growth under fluctuating light; the Can-0 plasmotype resulted in a lower ϕPSII and a higher ϕNPQ than the Col-0 plasmotype in three out of the four nucleotypes (Fig 8A-B). However, only the difference in ϕPSII of the Can-0 and Col-0 plasmotypes in the Col-0 nucleotype was found to be statistically significant ($p = 0.0348$).

The effect of the Can-0 plasmotype appeared to be reversed in the Tanz nucleotype (higher ϕPSII and lower ϕNPQ), although none of the pairwise differences within this nucleotype were significant. This is consistent with the fact that the Can-0 plasmotype was not considered to be an outlier within the Tanz nucleotype distribution, as shown by PCA (Fig 5C), and again highlights the role of epistatic interactions between in the nuclear and cytoplasmic genomes in shaping photosynthetic phenotypes.

To determine whether the mutations in subunits of the NDH complex affected the function of the complex, NDH activity was also measured in cybrids with the Can-0 plasmotype (Fig 8C). As the NDH complex contributes to cyclic electron flow and thus returns electrons to the plastoquinone pool, NDH activity can be estimated by measuring the post-illumination fluorescence rise (PIFR) (Otani, Yamamoto and Shikanai, 2017). The lines *ndhM* and *ndhO*, with mutations in nuclear-encoded subunits of the NDH complex, were included as negative controls. The Can-0 plasmotype conferred an increased NDH activity in the Cvi nucleotype when compared to the Col-0 plasmotype, both before and after the day of fluctuating light ($p = 0.0006$ and 0.0372 respectively). However, no statistically significant differences between the Can-0 and Col-0 plasmotypes were observed in the Col-0 nucleotype, suggesting that the observed effects on ϕPSII and ϕNPQ may not be intrinsically linked to the activity of the NDH complex.

Differences in NDH activity were also observed between nucleotypes, with the highest activities observed in the Cvi and Tanz nucleotypes (Fig 8C). The NDH complex consists of 11 plastid-encoded subunits and 8 nuclear-encoded subunits, and interacts with a range of other nuclear-encoded proteins (Shikanai, 2016). Differences in NDH activity between nucleotypes may thus be explained by allelic variation in these nuclear-encoded proteins.

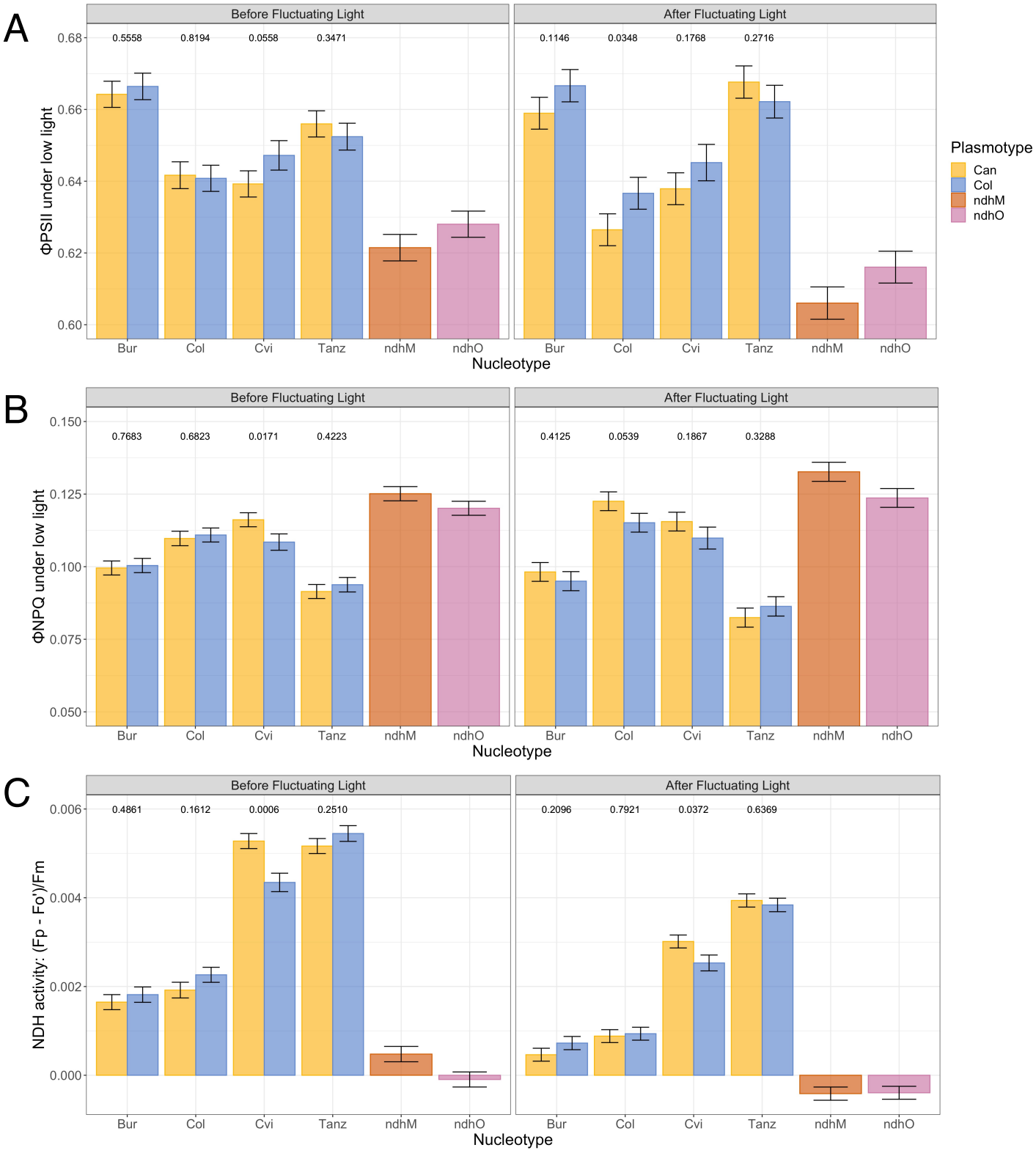


Figure 8: The Can-0 plasmotype affects ϕPSII, ϕNPQ and NDH activity.

(A) The BLUE for each genotype for ϕPSII measured under low light (200 µmol/m²/s), before and after a day of fluctuating light.

(B) The BLUE for each genotype for ϕNPQ measured under low light (200 µmol/m²/s), before and after a day of fluctuating light.

(C) The BLUE for each genotype for NDH activity measured after 200 µmol/m²/s, before and after a day of fluctuating light.
Error bars represent the BLUE ± SE. The p-values for each pairwise comparison within a nucleotype are shown above the bars. n=16-24 for each genotype.

Discussion

Trait heritability is dynamic in *A. thaliana*.

Phenotyping of the *A. thaliana* cybrid panel in the DEPI system allowed the heritability of photosynthetic traits to be calculated over time and in response to changing environmental conditions. In particular, the effect of fluctuating light and low temperature were observed (Fig 4, Supplementary Fig 1). While the heritability of certain traits, such as ϕNO, was found to be relatively constant over time, the heritability of others was more dynamic. The heritability of qE_t and qE_{SV}, two measures of qE, the rapid component of non-photochemical quenching, increased on the days with fluctuating light conditions (Supplementary Fig 1A).

Fluctuations in heritability are commonly associated with gene-environment (G×E) interactions (Visscher, Hill and Wray, 2008). For example, the heritability of environmentally-sensitive traits was found to decrease under drought conditions in barley (Chen et al., 2014). This framework can also be applied to the results presented here: qE is induced under high light conditions (Murchie and Ruban, 2020), and thus variation in genes controlling qE will only be apparent under these particular environmental conditions. This suggests that phenotyping under fluctuating environmental conditions is especially important for understanding the genetic mechanisms underlying dynamic components of photosynthesis, such as qE, whereas steady-state measurements are sufficient for understanding the basal properties of photosynthesis, such as ϕNO. As suggested by Flood et al. (2016), these fluctuations in heritability also have implications for plant breeding programmes; in order to efficiently select for these traits, plants should be phenotyped under dynamic environmental conditions when their heritability is high. This is particularly relevant for field crops which will experience dynamic environmental conditions, resulting in the induction of environmentally-sensitive processes such as NPQ.

Furthermore, the largest contributions of the plasmotype are observed for dynamic traits, particularly qE. While the average contribution of the plasmotype to the heritability of photosynthetic traits was relatively low, it explained up to 26.3% of the heritability of qE_{SV}. This effect was largely driven by the Can-0 plasmotype, which had the largest additive effect size for this phenotype. This suggests that genetic variation in the cytoplasmic genomes should not be overlooked when studying genetic variation in photosynthetic traits, particularly when considering dynamic traits. This is in line with the hypothesis that certain genes are

retained in the cytoplasmic genomes to allow their regulation in direct response to the metabolic state (for example, redox state) of the organelles (Allen, 2015).

The Can-0 plasmotype affects dynamic photosynthetic traits under particular conditions.

Due to the impact of the Can-0 plasmotype on the heritability of qE_{SV} and its identification as an extreme plasmotype from the PCA (Fig 5), the phenotype of the Can-0 plasmotype was investigated in more detail. This plasmotype confers an increase in qE and a decrease in ϕPSII under high light, particularly when combined with the Col-0 nucleotype (Fig 6). To better understand the mechanisms underlying this phenotype, the unique SNPs in the Can-0 cytoplasmic genomes were investigated, identifying a non-synonymous mutation in a subunit of the NDH complex as a candidate.

The NDH complex is involved in cyclic electron flow in the photosynthetic electron transport chain, accepting electrons from NADPH and returning them to the plastoquinone pool (Shikanai, 2014). The complex also pumps protons into the thylakoid lumen, increasing the ΔpH across the thylakoid membrane and so promoting the production of ATP (Strand, Fisher and Kramer, 2017). Several different physiological functions for the NDH complex have been suggested. Firstly, as the NDH complex oxidises NADPH and promotes the production of ATP, it adjusts the ratio of NADPH to ATP produced by the photosynthetic electron transport chain. This allows the ratio to more closely match the demands of the Calvin-Benson-Bassham (CBB) cycle, preventing the build-up of NADPH which could otherwise lead to the formation of reactive intermediates (Cruz et al., 2005). The NDH complex is considered to be particularly important in C4 photosynthesis for this reason; the carbon concentrating mechanism employed in C4 photosynthesis increases the demand for ATP, further exacerbating the mismatch in supply and demand between the electron transport chain and downstream processes (Ishikawa et al., 2016). Furthermore, cyclic electron flow is hypothesised to be more important under stress conditions such as low temperatures and fluctuating light, when the photosystems are more likely to become overexcited and photodamaged (Yamori, Makino and Shikanai, 2016; Strand, D'Andrea and Bock, 2019). Thus, an alteration in NDH activity is likely to have a larger effect under fluctuating environmental conditions, which is also when the phenotype of the Can-0 plasmotype is most prominent (Fig 6). Lastly, the proton pumping action of the NDH complex acidifies the thylakoid lumen. If the ΔpH across the thylakoid membrane is large enough, NPQ may be induced, promoting the dissipation of energy in the form of heat (Murchie and Ruban, 2020).

This would further protect the photosynthetic machinery under stress conditions. However, the slow relaxation of NPQ may reduce photosynthetic efficiency even under non-stress conditions, reducing plant productivity. Therefore, a change in the NDH activity may explain both the increase in NPQ and the decrease in ϕPSII observed in the Can-0 plasmotype.

Overall, this pointed to the SNP in *ndhF* as a likely candidate underpinning the phenotype of the Can-0 plasmotype. The NdhF subunit of the NDH complex is embedded in the thylakoid membrane and has been implicated in proton pumping (Laughlin et al., 2019; Strand, Fisher and Kramer, 2017). Furthermore, the NdhF subunit may be involved in the formation of the PSI-NDH supercomplex, stabilising the NDH complex under high light conditions (Otani, Kato and Shikanai, 2018; Peng and Shikanai, 2011). However, quantification of NDH activity in cybrids with the Can-0 plasmotype revealed no clear differences compared to cybrids with the Col-0 plasmotype, except in the Cvi nucleotype (Fig 8C). This may be because of the manner in which the NDH measurements were made; measurements in the PlantScreen™ SC System were carried out over the course of the day and involved the transportation of plants from the climate chamber to the greenhouse where the PlantScreen™ SC System is located. Thus, if NDH activity is a particularly dynamic trait, it may have been obscured by diurnal rhythms and/or acclimation responses of the plants. Therefore, it may be more accurate to measure NDH activity using a fixed camera system such as the DEPI platform, allowing a large number of plants to be measured simultaneously.

Alternatively, the phenotype of the Can-0 plasmotype may be due to a mutation in *rps15*, encoding a component of the small subunit of the plastid ribosome (Tiller and Bock, 2014). This mutation could explain the phenotype if it impairs translation in the chloroplasts. Impaired chloroplast translation makes plants more sensitive to photoinhibition as they are unable to quickly replace damaged components of the electron transport chain, such as the D1 subunit of PSII, leading to decreased photosynthetic efficiency (Sun and Zerges, 2015). A full knock-out of the Rps15 protein has been shown to decrease ϕPSII in tobacco and result in more severe growth phenotypes under stress conditions (Fleischmann et al., 2011). Furthermore, a reduction in chloroplast translation could explain the reduced green-ness of the cybrids with the Can-0 plasmotype observed when they were grown in outdoor conditions (Supplementary Fig 3).Therefore, neither mutation can be ruled out on the basis of the function of the proteins affected.

Both SNPs are unique to the Can-0 plasmotype, both within the 60 progenitor lines of the cybrid panel and within the 1544 *A. thaliana* accessions that have been sequenced. Thus, the two SNPs cannot be distinguished on the basis of comparisons between naturally occurring plasmotypes, as has been carried out for the putative causal SNPs of the Bur-0 plasmotype phenotype (Tijink, 2021a; Theeuwen, unpublished). Introduction of the Can-0 alleles of *ndhF* and *rps15* by genetic modification could be used instead to determine which SNP is causal to the phenotype. Recently, several methods have been developed to modify the plastid genome, including genome editing and the stable introduction of minicircles containing a gene of interest (Kang et al., 2021; Occhialini et al., 2021). These methods could be employed to further dissect the contribution of allelic variation in plastid genes to photosynthetic phenotypes.

Interestingly, the wild-type Can-0 accession also has reduced ϕPSII compared to the Col-0 accession (Flood, 2015), suggesting possible co-evolution of the nuclear and cytoplasmic genomes of Can-0. QTL analysis of a Can-0 × Col-0 RIL population revealed QTLs associated with decreased ϕPSII and chlorophyll reflectance in the Can-0 nuclear genome (Flood, 2015). The Can-0 accession originates from Lanzarote, one of the Canary Islands, where it would be exposed to high light intensities, high temperatures and low water availability (the climate of Lanzarote is classified as 'hot desert' in the Köppen-Geiger system (Peel, Finlayson and McMahon, 2007)). These conditions can increase plant oxidative stress, suggesting that Can-0's increased photoprotection may be adaptive, helping the Can-0 accession to survive in these harsh environmental conditions. The relative contribution of the nuclear and cytoplasmic genomes to the photosynthetic phenotype of the Can-0 accession could be investigated by comparing the WT Can-0 and Col-0 accessions to Can-0_Col-0 and Col-0_Can-0 cybrids. This could also reveal whether there are epistatic interactions between the Can-0 nuclear and cytoplasmic genomes (Fig 9), indicating the extent of cooperation between the genomes and whether they might have co-evolved. Similarly, the NDH activity of the WT Can-0 accession could be measured. This could provide further insight into the physiological function of the NDH complex, especially in adaptation to harsh environmental conditions. Furthermore, genotyping of other *A. thaliana* accessions from the Canary Islands could be used to assess whether the mutations in the Can-0 accession are adaptive.

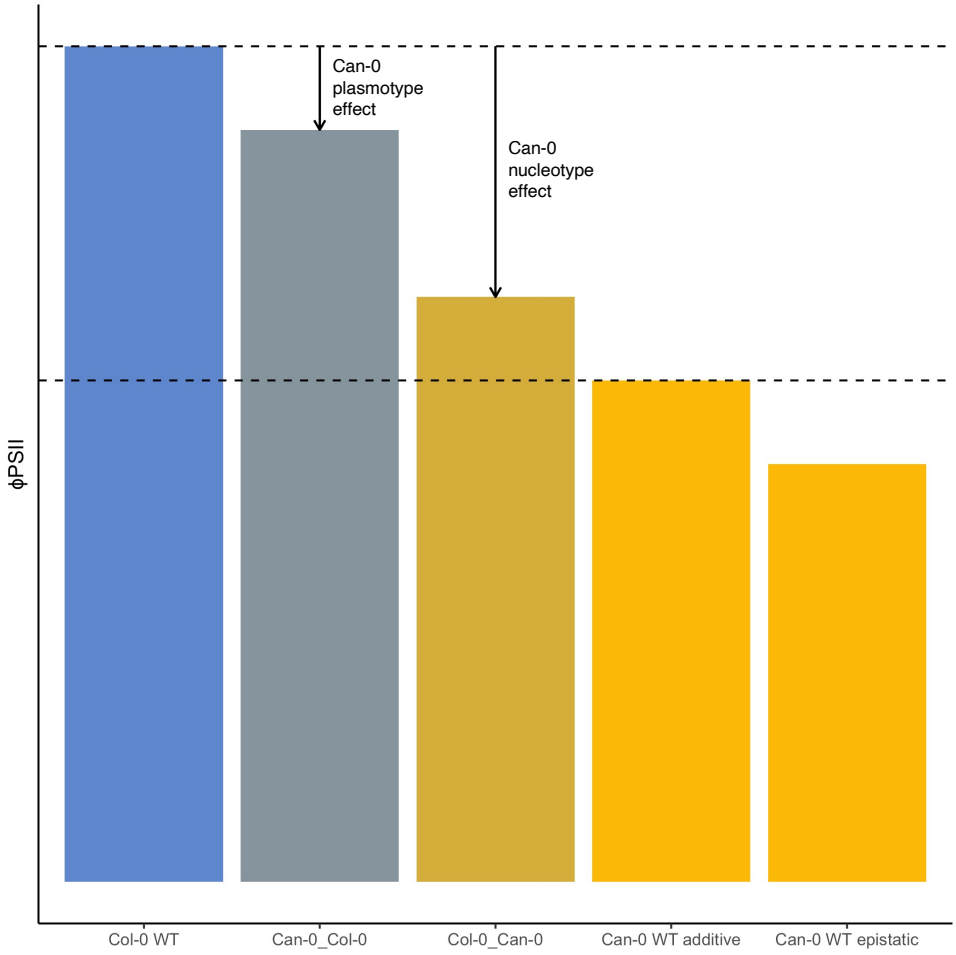


Figure 9: Determination of additive and epistatic effects in the Can-0 accession. *Hypothetical differences in ϕPSII for four different genotypes – Col-0 WT, the Can-0_Col-0 cybrid, the Col-0_Can-0 cybrid, and the Can-0 WT – are shown. If the Can-0 nucleotype and plasmotype are additive, the decrease in ϕPSII for the Can-0 WT compared to the Col-0 WT should be equal to the sum of the Can-0 plasmotype effect and the Can-0 nucleotype effect. If there is an epistatic interaction between the Can-0 nucleotype and plasmotype, the decrease in ϕPSII for the Can-0 WT is expected to be bigger than the sum of the nucleotype and plasmotype effects.*

Nucleotype-plasmotype interactions are important when considering photosynthetic phenotypes.

As well as the plasmotype alone, nucleotype-plasmotype interactions were found to contribute significantly to the heritability of photosynthetic traits, explaining up to 68.8% of H^2. Indeed, most plasmotype effects are not additive, but are observed in only one of the nucleotypes, or are observed in all but one of the nucleotypes (Fig 5B-C). This indicates the presence of widespread epistatic interactions between genes encoded in the nucleus and the organelles. This is perhaps unsurprising because the majority of organellar proteins are

encoded in the nucleus and therefore metabolic pathways in the organelles rely on proteins of different origins. Furthermore, multisubunit complexes in the organelles consist of subunits encoded by both the nuclear and cytoplasmic genomes: for example, the complexes making up the photosynthetic electron transport chain in the thylakoid membrane all consist of a combination of nuclear- and chloroplast-encoded proteins in *A. thaliana* (Allen et al., 2011). Supercomplexes also form between the complexes in the thylakoid membrane in a dynamic manner to regulate photosynthesis is response to environmental conditions, involving more direct interactions between nuclear- and chloroplast-encoded proteins (Minagawa, 2013). Thus, proteins of different origins are found in highly interconnected networks and there are ample possibilities for epistasis to occur. Cytonuclear epistasis has previously been identified, explaining differences in the *A. thaliana* metabolome and developmental traits (Joseph et al., 2013; Roux et al., 2016).

Cytonuclear epistasis is observed in cybrids with the Can-0 plasmotype (Fig 5C, Fig 6A-B, Fig 8). The Can-0 plasmotype was found to confer high qE under fluctuating light conditions in the Col-0 and Cvi nucleotype backgrounds, and low ϕPSII under fluctuating conditions in the Col-0 nucleotype background. Any effect of the Can-0 plasmotype thus appears to be hidden in the Bur-0 and Tanz nucleotypes. The Bur-0 and Tanz nucleotypes are associated with a lower qE and ϕNPQ (Fig 6A, Fig 8B) and a higher ϕPSII (Fig 6B, Fig 8A), potentially masking the Can-0 effect. This highlights that the nucleotype background in which the different plasmotypes are phenotyped should be taken into consideration when designing phenotyping experiments.

An extreme example of epistasis between the nuclear and cytoplasmic genomes comes in the form of cytoplasmic male sterility (CMS) – the development of male-sterile plants due to incompatibility between the nuclear and cytoplasmic genomes. In cases of CMS, the effect of the cytoplasm can often be counter-acted by restorer genes in the nucleus (Budar, Touzet and De Paepe, 2003). Thus, the degree of male fertility (shown by the rate of pollen abortion, for example) depends both on genes in the nucleus and in the organelles. As the chloroplasts and mitochondria are both central to plant metabolism, CMS has been explained as an inability to facilitate the energy-demanding process of pollen development (Budar, Touzet and De Paepe, 2003). Several of the plasmotypes used to make the new cybrid panel were known to induce CMS, including Shah, Kolyv-6 and Sus-1 (Gobron et al., 2013). These three plasmotypes were found to impair male fertility specifically in the Cvi nucleotype (Theeuwen, unpublished). Screening of the cybrid panel further identified the

Staro-2 plasmotype as an inducer of CMS, possibly as a result of large scale rearrangements of the mitochondrial genome (Tijink, 2021b). The investigation into the contribution of the cytoplasmic genomes to photosynthetic traits also identified Staro-2 as having a distinct phenotype (Fig 5A, Supplementary Fig 4A-B), associated with high ϕPSII (Supplementary Fig 5). This is in contrast to the changes in the mitochondrial genome, which are thought to decrease the efficiency of ATP production, suggesting that the increased photosynthetic efficiency of the Staro-2 plasmotype may act to compensate for the impaired energy supply. This would be in contrast to documented CMS mutants in *Nicotiana sylvestris*, where a mutation in the mitochondrial genome results in reduced photosynthesis (Sabar, De Paepe and de Kouchkovsky, 2000). Thus, investigation into the mechanism of CMS in Staro-2 could improve our understanding of the relationship between the mitochondria and the chloroplasts in the regulation of photosynthesis. However, further research is required to confirm that the photosynthetic phenotype of the Staro-2 plasmotype is indeed genetic in nature rather than a seed effect as initial investigations have proved inconclusive.

The contribution of the plasmotype is relatively small compared to the nucleotype. Overall, the average contribution of the plasmotype and the nucleotype-plasmotype interaction accounted for only ~5% of the total heritability of the photosynthetic traits investigated here. Accordingly, the vast majority of the variation between the cybrids can be explained by the nuclear genome. This is perhaps unsurprising as the nuclear genome of *A. thaliana* consists of 135 Mb, compared to only 155 kb for the chloroplast genome and 367 kb for the mitochondrial genome (AGI, 2000; Sato et al., 1999; Unseld et al., 1997). The cytoplasmic genomes thus make up approximately 0.4% of the total genome of *A. thaliana*. Therefore, contributions by the plasmotype alone to H^2 of more than 20% for particular phenotypes are very remarkable and deviate from our expectation that the nuclear genome is overwhelmingly responsible for natural phenotypic variation.

The PCA (Fig 5B-C) revealed the variation between nucleotypes in the measured photosynthetic traits. In particular, the Bur-0 nucleotype separated from the other three nucleotypes along the first principal component, approximately corresponding to differences in qE under fluctuating light conditions. Phenotyping of the Can-0 cybrids also showed that the Bur-0 nucleotype is associated with a high ϕPSII and low ϕNPQ (Fig 8A-B). QTL mapping of a Bur-0 × Col-0 recombinant inbred line (RIL) population has identified a QTL on chromosome 1 that could explain the high ϕPSII observed in Bur-0 (Flood et al., 2015). The

Tanz nucleotype also has a higher ϕPSII and a lower ϕNPQ compared to the Col-0 and Cvi nucleotypes (Fig 8A-B). This indicates that there is allelic variation in the nuclear genome that affects the regulation of photosynthesis which can be studied further using QTL mapping, for example.

Quantification of NDH activity also highlighted differences between the four nucleotypes (Fig 8C). In particular, the Cvi and Tanz nucleotypes have significantly higher NDH activity than the Bur-0 and Col-0 nucleotypes. Natural quantitative variation in the activity of the NDH complex has not previously been shown. Phenotyping of a larger selection of genotypes for NDH activity could reveal additional variation and provide a more detailed insight into the function and physiological relevance of the NDH complex. These measurements could be integrated into high-throughput phenotyping platforms that are already equipped to measure chlorophyll fluorescence.

Linking phenomics and (cytoplasmic) genomics still presents numerous challenges. While this this thesis aims to illustrate how high-throughput phenotyping data can be used to study genetic variation in photosynthesis, this process still requires fine-tuning to make it more efficient. I used several different methods to link differences in phenotype to differences in genotype, including Principal Component Analysis (Fig 5) and Genome-Wide Association Studies (Fig 7). GWAS for the cytoplasmic genomes cannot be interpreted in the same way as a GWAS performed on nuclear genomic data because no true recombination occurs in the cytoplasmic genomes. All of the unique SNPs/indels within a plasmotype are therefore completely linked and will be associated with the plasmotype's phenotype. A GWAS for the plastid genome, for example, can therefore be used to identify plasmotypes of interest (by identifying the origin of the SNPs associated with the phenotype), but cannot be used to identify causal SNPs underlying the phenotype unless there is sufficient overlap between the plasmotypes.

The SnpEff software was further used to narrow down the list of candidate SNPs. SnpEff predicts the effect of genetic variants, which are classified as 'modifier' (if the mutation is in an intergenic region), 'low' (synonymous mutations or intron variants), 'moderate' (non-synonymous mutations), and 'high' (frameshifts or loss/gain of stop codons). For this project, only variants with a 'moderate' or 'high' predicted effect were considered as candidates. However, this approach excludes variants that might affect the regulation of gene expression. Furthermore, in several cases it was difficult to explain the differences in

phenotype using the ‘moderate’ or ‘high’ effect variants alone. For example, from the PCA, it was concluded that the Tanz-2_Cvi cybrid was an outlier from the Cvi nucleotype population (Fig 5B). However, the only unique genetic variants in the Tanz-2 plasmotype are predicted to be ‘modifiers’. By including transcriptomics and/or proteomics into these analyses, the pathway from genotype to phenotype could be further elucidated.

These analyses are further complicated by the difficulty in replicating phenotypes measured in the DEPI system (in which only 4 replicates of each genotype were measured, divided over multiple experiments) in the PlantScreen™ SC System, which measures trays of plants one by one. Photosynthesis is a highly dynamic trait and the plasmotype is shown to particularly affect dynamic properties of photosynthesis such as NPQ. Therefore, a platform that is capable of measuring a large number of replicates simultaneously would allow even small effect sizes in dynamic traits to be captured.

Conclusions

In conclusion, the cytoplasmic genomes, as well as their interaction with the nuclear genome, make a considerable contribution to the heritability of dynamic photosynthetic traits. This is especially impressive given the relatively small size of the cytoplasmic genomes in comparison to the nuclear genome. Phenotyping of the expanded cybrid panel has revealed novel variation in photosynthesis, allowing the identification of Can-0 as a plasmotype of interest. This plasmotype confers a significant decrease in photosynthetic efficiency under fluctuating light, especially when combined with the Col-0 nucleotype. By examining the genetic basis of this phenotype in more detail, our understanding of the regulation of photosynthesis under dynamic conditions can be improved. This highlights the value of exploring natural variation in photosynthesis; it does not require prior knowledge of the underlying genetics, allowing novel mechanisms to be found. Improving our understanding of how genetic variation shapes photosynthetic traits can ultimately be used to improve the photosynthesis of crop plants. To do this, it is important that we invest in phenotyping platforms that allow us to accurately measure dynamic components of photosynthesis. Furthermore, we need to solidify the link between phenotypic variation and genotypic variation, for example by integrating transcriptomic/proteomic analyses into our research. By making these improvements, we can come closer to transitioning from the model plant, *A. thaliana*, to crops, with the goal of improving crop productivity.

Acknowledgements

I would like to thank the following people for helping to shape this project: Tom Theeuwen for his supervision throughout this project and for our many interesting discussions on photosynthesis, genetics and everything in between; René Boesten and Mark Aarts for acting as the examiners for this project, and for their useful comments, questions and feedback throughout; Jeremy Harbinson for his willingness to discuss and explain photosynthetic physiology; Jara Jauregui Beso for co-writing the script for the measurement of NDH activity; the staff at Unifarm for facilitating the Can-0 phenotyping experiments; and the Botanical Genetics group at Wageningen University & Research for welcoming me into the group for the past 6 months.

References

Acevedo-Siaca, L. G., Coe, R., Wang, Y., Kromdijk, J., Quick, W. P., Long, S. P. (2020). Variation in photosynthetic induction between rice accessions and its potential for improving productivity. *New Phytologist*, *227*(4), 1097-1108.

Adam, S., Murthy, S. D. S. (2013). Effect of Cold Stress on Photosynthesis of Plants and Possible Photoprotection Mechanisms. In: R. K. Gaur, P. Sharma, eds., *Approaches to Plant Stress and their Management*, 1st ed. New Delhi: Springer, 219-226.

Allen, J. F., de Paula, W. B. M., Puthiyaveetil, S., Nield, J. (2011). A structural phylogenetic map for chloroplast photosynthesis. *Trends in Plant Science*, *16*(12), 645-655.

Allen, J. F. (2015). Why chloroplasts and mitochondria retain their own genomes and genetic systems: Colocation for redox regulation of gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(33), 10231-10238.

Alonso-Blanco, C., Andrade, J., Becker, C., Bemm, F., Bergelson, J., et al. (2016). 1,135 Genomes Reveal the Global Pattern of Polymorphism in *Arabidopsis thaliana*. *Cell*, *166*(*2*), 481-491.

The Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, *408*, 796-815.

Bates, D., Mächler, M., Bolker, B., Walker, S. (2015). Fitting Linear Mixed-Effects Models using lme4. *Journal of Statistical Software*, *67*(1), 1-48.

Bock, D. G., Andrew, R. L., Rieseberg, L. H. (2014). On the adaptive value of cytoplasmic genomes in plants. *Molecular Ecology*, *23*(20), 4899-4911.

Budar, F., Touzet, P., De Paepe, R. (2003). The nucleo-mitochondrial conflict in cytoplasmic male sterilities revisited. *Genetica*, *117*(1), 3-16.

Carmo-Silva, E., Scales, J. C., Madgwick, P. J., Parry, M. A. J. (2015). Optimizing Rubisco and its regulation for greater resource use efficiency. *Plant, Cell and Environment*, *38*(9), 1817-1832.

Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*, *4*(1), s13742-015-0047-8.

Chen, D., Neumann, K., Friedel, S., Kilian, B., Chen, M., Altmann, T., Klukas, C. (2014). Dissecting the Phenotypic Components of Crop Plant Growth and Drought Responses Based on High-Throughput Image Analysis. *The Plant Cell*, *26*(12), 4636-4655.

Chiang, C., Layer, R. M., Faust, G. G., Lindberg, M. R., Rose, D. B., Garrison, E. P., Marth, G. T., Quinlan, A. R., Hall, I. M. (2015). SpeedSeq: ultra-fast personal genome analysis and interpretation. *Nature Methods*, *12*(10), 966-968.

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X., Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly*, *6*(2), 80-92.

Cruz, J. A., Avenson, T. J., Kanazawa, A., Takizawa, K., Edwards, G. E., Kramer, D. M. (2005). Plasticity in light reactions of photosynthesis for energy production and photoprotection. *Journal of Experimental Botany*, *56*(411), 395-406.

Cruz, J. A., Savage, L. J., Zegarac, R., Hall, C. C., Satoh-Cruz, M., Davis, G. A., Kent Kovac, W., Chen, J., Kramer, D. M. (2016). Dynamic Environmental Photosynthetic Imaging Reveals Emergent Phenotypes. *Cell Systems*, *2*(6), 365-377.

De Mendiburu, F. (2021). agricolae: Statistical Procedures for Agricultural Research. R package version 1.3-5. https://CRAN.R-project.org/package=agricolae.

Driever, S. M., Lawson, T., Andralojc, P. J., Raines, C. A., Parry, M. A. J. (2014). Natural Variation in photosynthetic capacity, growth, and yield in 64 field-grown wheat genotypes. *Journal of Experimental Botany*, *65*(17), 4959-4973.

Driever, S. M., Simkin, A. J., Alotaibi, S., Fisk, S. J., Madgwick, P. J., Sparks, C. A., Jones, H. D., Lawson, T., Parry, M. A. J., Raines, C. A. (2017). Increased SBPase activity improves photosynthesis and grain yield in wheat grown in greenhouse conditions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1730), 20160384.

Durvasula A., Fulgione A., Gutaker, R. M., Alacakaptan, I., Flood, P. J., Neto, C., Tsuchimatsu, T., Burbano, H. A., Picó, F. X., Alonso-Blanco, C., Hancock, A. M. (2017). African genomes illuminate the early history and transition to selfing in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(20), 5213-5218.

El-Lithy, M. E., Rodrigues, G. C., van Rensen, J. J. S., Snel, J. F. H., Dassen, H. J. H. A., Koornneef, M., Jansen, M. A. K., Aarts, M. G. M., Vreugdenhil, D. (2005). Altered photosynthetic performance of a natural *Arabidopsis* accession is associated with atrazine resistance. *Journal of Experimental Biology*, *56*(416), 1625-1634.

Fleischmann, T. T., Scharff, L. B., Alkatib, S., Hasdorf, S., Schöttler, M. A., Bock, R. (2011). Nonessential Plastid-Encoded Ribosomal Proteins in Tobacco: A Developmental Role for Plastid Translation and Implications for Reductive Genome Evolution. *The Plant Cell*, *23*(9), 3137-3155.

Flood, P. J., Harbinson, J., Aarts, M. G. M. (2011). Natural genetic variation in plant photosynthesis. *Trends in Plant Science*, *16*(6), 327-335.

Flood, P. J. (2015). *Natural genetic variation in Arabidopsis thaliana photosynthesis*. Wageningen University & Research.

Flood, P. J., Kruijer, W., Schnabel, S. K., van der Schoor, R., Jalink, H., Snel, J. F. H., Harbinson, J., Aarts, M. G. M. (2016). Phenomics for photosynthesis, growth and reflectance in *Arabidopsis thaliana* reveals circadian and long-term fluctuations in heritability. *Plant Methods*, *12*, 14.

Flood, P. J., Theeuwen, T. P. J. M., Schneeberger, K., Keizer, P., Kruijer, W., Severing, E., Kouklas, E., Hageman, J. A., Wijfjes, R., Calvo-Baltanas, V., Becker, F. F. M., Schnabel. S. K., Willems, L. A. J., Ligterink, W., van Arkel, J., Mumm, R., Gualberto, J. M., Savage, L., Kramer, D. M., Keurentjes, J. J. B., van Eeuwijk, F., Koornneef, M., Harbinson, J., Aarts, M. G. M., Wijnker, E. (2020). Reciprocal cybrids reveal how organellar genomes affect plant phenotypes. *Nature Plants*, *6*(1), 13-21.

Garcia-Molina, A., Leister, D. (2020). Accelerated relaxation of photoprotection impairs biomass accumulation in *Arabidopsis*. *Nature Plants*, *6*(1), 9-12.

Garrison, E., Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. Preprint at https://arxiv.org/abs/1207.3907.

Gobron, N., Waszczak, C., Simon, M., Hiard, S., Boivin, S., Charif, D., Ducamp, A., Wenes, E., Budar, F. (2013). A Cryptic Cytoplasmic Male Sterility Unveils a Possible Gynodioecious Past for *Arabidopsis thaliana*. *PLoS ONE*, *8*(4), e62450.

Grantham, R. (1974). Amino acid difference formula to help explain protein evolution. *Science*, *185*(4145), 862-864.

Hazra, S., Henderson, J. N., Liles, K., Hilton, M. T., Wachter, R. M. (2015). Regulation of Ribulose-1,5-bisphosphate Carboxylase/Oxygenase (Rubisco) Activase. *The Jounral of Biological Chemistry*, *290*(40), 24222-24236.

Hubbart, S., Smilie, R. A., Heatley, M., Swarup, R., Foo, C. C., Zhao, L., Murchie, E. H. (2018). Enhanced thylakoid photoprotection can increase yield and canopy radiation use efficiency in rice. *Communications Biology*, *1*, 22.

Ishikawa, N., Takabayashi, A., Noguchi, K., Tazoe, Y., Yamamoto, H., von Caemmerer, S., Sato, F., Endo, T. (2016). NDH-Mediated Cyclic Electron Flow Around Photosystem I is Crucial for C4 Photosynthesis. *Plant and Cell Physiology*, *57*(10), 2020-2028.

Joseph, B., Corwin, J. A., Li, B., Atwell, S., Kliebenstein, D. J. (2013). Cytoplasmic genetic variation and extensive cytonuclear interactions influence natural variation in the metabolome. *eLife*, *2*, e00776.

Kaiser, E., Morales, A., Harbinson, J., Heuvelink, E., Prinzenberg, A. E., Marcelis, L. F. M. (2016). Metabolic and diffusional limitations of photosynthesis in fluctuating irradiance in *Arabidopsis thaliana*. *Scientific Reports*, *6*(August), 31252.

Kajala, K., Covshoff, S., Karki, S., Woodfield, H., Tolley, B. J., Dionora, M. J. A., Mogul, R. T., Mabilangan, A. E., Danila, F. R., Hibberd, J. M., Quick, W. P. (2011). Strategies for engineering a two-celled photosynthetic pathway into rice. *Journal of Experimental Biology*, *62*(9), 3001-3010.

Kang, B-C., Bae, S-J., Lee, S., Lee, J. S., Kim, A., Lee, H., Baek, G., Seo, H., Kim, H., Kim, J-S. (2021). Chloroplast and mitochondrial DNA editing in plants. *Nature Plants*, *7*(7), 899-905.

Kapralov, M. V., Filatov, D. A. (2007). Widespread positive selection in the photosynthetic Rubisco enzyme. *BMC Evolutionary Biology*, *7*(April), 73.

Klughammer, C., Schreiber, U. (2008). Complementary PS II quantum yields calculated from simple fluorescence parameters measured by PAM fluorometry and the Saturation Pulse method. *PAM Application Notes*, *1*, 27-35.

Kromdijk, J., Glowacka, K., Leonelli, L., Gabilly, S. T., Iwai, M., Niyogi, K. K., Long, S. P. (2016). Improving photosynthesis and crop productivity by accelerating recovery from photoprotection. *Science*, *354*(6314), 857-861.

Laughlin, T. G., Bayne, A. N., Trempe, J-F., Savage, D. F., Davies, K. M. (2019). Structure of the complex I-like molecule NDH of oxygenic photosynthesis. *Nature*, *566*(7744), 411-414.

Lawson, A. (2020). *Revealing cyto-nuclear interactions through phenotypic variation: a study on cybrids of outdoor grown Arabidopsis thaliana*. Wageningen University & Research.

Lawson, T., Blatt, M. R. (2014). Stomatal Size, Speed, and Responsiveness Impact on Photosynthesis and Water Use Efficiency. *Plant Physiology*, *164*(4), 1556-1570.

Lenth, R. V. (2021). emmeans: Estimated Marginal Means, aka Lest-Squares Means. R package version 1.6.3. https://CRAN.R-project.org/package=emmeans.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format SAMtools. *Bioinformatics*, *25*(16), 2078-2079.

Li, H. (2013). Aligning sequencing reads, clone sequences and assembly contigs with BWA-MEM. Preprint at https://arxiv.org/abs/1303.3997.

Lilley, R. M., Portis, A. R. (1997). ATP Hydrolysis Activity and Polymerization State of Ribulose-1,5-Bisphosphate Carboxylase Oxygenase Activase (Do the Effects of Mg^{2+}, K^{+}, and Activase Concentrations Indicate a Functional Similarity to Actin?). *Plant Physiology*, *114*(2), 605-613.

Long, S. P., Zhu, X-G., Naidu, S. L., Ort, D. R. (2006). Can improvement in photosynthesis increase crop yields?. *Plant, Cell and Environment*, *29*(3), 315-330.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, *17*(1), 10-12.

Matsubara, S. (2018). Growing plants in fluctuating environments: why bother?. *Journal of Experimental Botany*, *69*(*20*), 4651-4654.

Matthews, J. S. A., Vialet-Chabrand, S., Lawson, T. (2020). Role of blue and red light in stomatal dynamic behaviour. *Journal of Experimental Biology*, *71*(7), 2253-2269.

McAusland, L., Vialet-Chabrand, S., Davey, P., Baker, N. R., Brendel, O., Lawson T. (2016). Effects of kinetics of light-induced stomatal responses on photosynthesis and water-use efficiency. *New Phytologist*, *211*(4), 1209-1220.

Minagawa, J. (2013). Dynamic reorganization of photosynthetic supercomplexes during environmental acclimation of photosynthesis. *Frontiers in Plant Science*, *4*, 513.

Monteith, J. L. (1977). Climate and the efficiency of crop production in Britain. *Philosophical Translations of the Royal Society B: Biological Sciences*, *281*(980), 277-294.

Mueller-Cajar, O., Stotz, M., Bracher, A. (2014). Maintaining photosynthetic CO_2 fixation via protein remodelling: the Rubisco activases. *Photosynthesis Research*, *119*(1-2), 191-201.

Murchie, E. H., Kefauver, S., Araus, J. L., Muller, O., Rascher, U., Flood, P. J., Lawson, T. (2018). Measuring the dynamic photosynthome. *Annals of Botany*, *112*(2), 207-220.

Murchie, E. H., Ruban, A. V. (2020). Dynamic non-photochemical quenching in plants: from molecular mechanism to productivity. *The Plant Journal*, *101*(4), 885-896.

Occhialini, A., Pfotenhauer, A. C., Li, L., Harbison, S. A., Lail, A. K., Burris, J. N., Piasecki, C., Piatek, A. A., Daniell, H., Stewart Jr, C. N., Lenaghan, S. C. (2021). Mini-synplastomes for plastid genetic engineering. *Plant Biotechnology Journal*, 1-14.

Otani, T., Yamamoto, H., Shikanai, T. (2017). Stromal Loop of Lhca6 is Responsible for the Linker Function Required for the NDH-PSI Supercomplex Formation. *Plant and Cell Physiology*, *58*(4), 851-861.

Otani, T., Kato, Y., Shikanai, T. (2018). Specific substitutions of light-harvesting complex I proteins associated with photosystem I are required for supercomplex formation with chloroplast NADH dehydrogenase-like complex. *The Plant Journal*, *94*, 122-130.

Paradis, E., Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, *35*, 526-528.

Pearcy, R. W. (1990). Sunflecks and Photosynthesis in Plant Canopies. *Annual Review of Plant Physiology and Plant Molecular Biology*, *41*, 421-453.

Peel, M. C., Finlayson, B. L., McMahon, T. A. (2007). Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth System Sciences*, *11*, 1633-1644.

Peng, L., Shikanai, T. (2011). Supercomplex Formation with Photosystem I Is Required for the Stabilization of the Chloroplast NADH Dehydrogenase-Like Complex in Arabidopsis. *Plant Physiology*, *155*(4), 1629-1639.

Pfannschmidt, T., Nilsson, A., Tullberg, A., Link, G., Allen, J. F. (1999). Direct transcriptional control of the chloroplast genes psbA and psaAB adjusts photosynthesis to light energy distribution in plants. *IUBMB Life*, *48*(3), 271-276.

Prinzenberg, A. E., Campos-Dominguez, L., Kruijer, W., Harbinson, J., Aarts, M. G. M. (2020). Natural variation in photosynthetic efficiency in *Arabidopsis thaliana* accessions under low temperature conditions. *Plant, Cell and Environment*, *43*(8), 2000-2013.

Qu, M., Zheng, G., Hamdani, S., Essemine, J., Song, Q., Wang, H., Chu, C., Sirault, X., Zhu, X-G. (2017). Leaf Photosynthetic Parameters Related to Biomass Accumulation in a Global Rice Diversity Survey. *Plant Physiology*, *175*(1), 248-258.

Ray, D. K., Mueller, N. D., West, P. C., Foley, J. A. (2013). Yield Trends Are Insufficient to Double Global Crop Production by 2050. *PLoS ONE*, *8*(6), e66428.

Roux, F., Mary-Huard, T., Barillot, E., Wenes, E., Botran, L., Durand, S., Villoutreix, R., Martin-Magniette, M-L., Camilleri, C., Budar, F. (2016). Cytonuclear interactions affect adaptive traits of the annual plant *Arabidopsis thaliana* in the field. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(13), 3687-3692.

Sabar, M., De Paepe, R., de Kouchkovsky, Y. (2000). Complex I Impairment, Respiratory Compensations, and Photosynthetic Decrease in Nuclear and Mitochondrial Male Sterile Mutants of *Nicotiana sylvestris*. *Plant Physiology*, *124*(3), 1239-1250.

Salter, W. T., Merchant, A. M., Richard, R. A., Trethowan, R., Buckley, T. N. (2019). Rate of photosynthetic induction in fluctuating light varies widely among genotypes of wheat. *Journal of Experimental Botany*, *70*(10), 2787-2796.

Sato, S., Nakamura, Y., Kaneko, T., Asamizu, E., Tabata, S. (1999). Complete Structure of the Chloroplast Genome of *Arabidopsis thaliana*. *DNA research*, *290*(6), 283-290.

Shikanai, T. (2014). Central role of cyclic electron transport around photosystem I in the regulation of photosynthesis. *Current Opinion in Biotechnology*, *26*, 25-30.

Shikanai, T. (2016). Chloroplast NDH: A different enzyme with a structure similar to that of respiratory NADH dehydrogenase. *Biochimica et Biophysica Acta*, *1857*(7), 1015-1022.

Soleh, M. A., Tanaka, Y., Kim, S. Y., Huber, S. C., Sakoda, K., Shiraiwa, T. (2017). Identification of large variation in the photosynthetic induction response among 37 soybean [*Glycine max* (L.) Merr.] genotypes that is not correlated with steady-state photosynthetic capacity. *Photosynthetic Research*, *131*(3), 305-315.

South, P. F., Cavanagh, A. P., Liu, H. W., Ort, D. R. (2019). Synthetic glycolate metabolism pathways stimulate crop growth and productivity in the field. *Science*. *363*(6422), eaat9077.

Strand, D. D., Fisher, N., Kramer, D. M. (2017). 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*(28), 11850-11860.

Strand, D. D., D'Andrea, L., Bock, R. (2019). Plastid NAD(P)H dehydrogenase-like complex: structure, function and evolutionary dynamics. *Biochemical Journal*, *476*(19), 2743-2756.

Sun, Y., Zerges, W. (2015). Translational regulation in chloroplasts for development and homeostasis. *Biochimica et Biophysica Acta*, *1847*(9), 809-820.

Tietz, S., Hall, C. C., Cruz, J. A., Kramer, D. M. (2017). $NPQ_{(T)}$: a chlorophyll fluorescence parameter for rapid estimation and imaging of non-photochemical quenching of excitons in photosystem-II-associated antenna complexes. *Plant, Cell and Environment*, *40*(8), 1243-1255.

Tijink, D. (2021a). *Zooming in on plasmotypic variation of Arabidopsis thaliana*. Wageningen University & Research.

Tijink, D. (2021b). *Investigating the cytoplasmic male sterility of the Staro-2 Arabidopsis thaliana accession*. Wageningen University & Research.

Tiller, N., Bock, R. (2014). The Translational Apparatus of Plastids and Its Role in Plant Development. *Molecular Plant*, *7*(7), 1105-1120.

Timmis, J. N., Ayliffe, M. A., Huang, C. Y., Martin, W. (2004). Endosymbiotic gene transfer: Organelle genomes forge eukaryotic chromosomes. *Nature Reviews Genetics*, *5*(2), 123-135.

Unseld, M., Marienfeld, J. R., Brandt, P., Brennicke, A. (1997). The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. *Nature Genetics*, *15*, 57-61.

Van der Auwera, G. A., O'Connor, B. D. (2020). *Genomics in the Cloud: Using Docker, GATK, and WDL in Terra*. Sebastopol: O'Reilly Media.

Visscher, P. M., Hill, W. G., Wray, N. R. (2008). Heritability in the genomics era – concepts and misconceptions. *Nature Reviews Genetics*, *9*(4), 255-266.

von Caemmerer, S. (2000). *Biochemical Models of Leaf Photosynthesis*. Collingwood: CSIRO.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer.

Wullschleger, S. D. (1993). Biochemical Limitiations to Carbon Assimilation in C3 Plants – A Retrospective Analysis of the A/C_i Curves from 109 Species. *Journal of Experimental Biology*, *44*(262), 907-920.

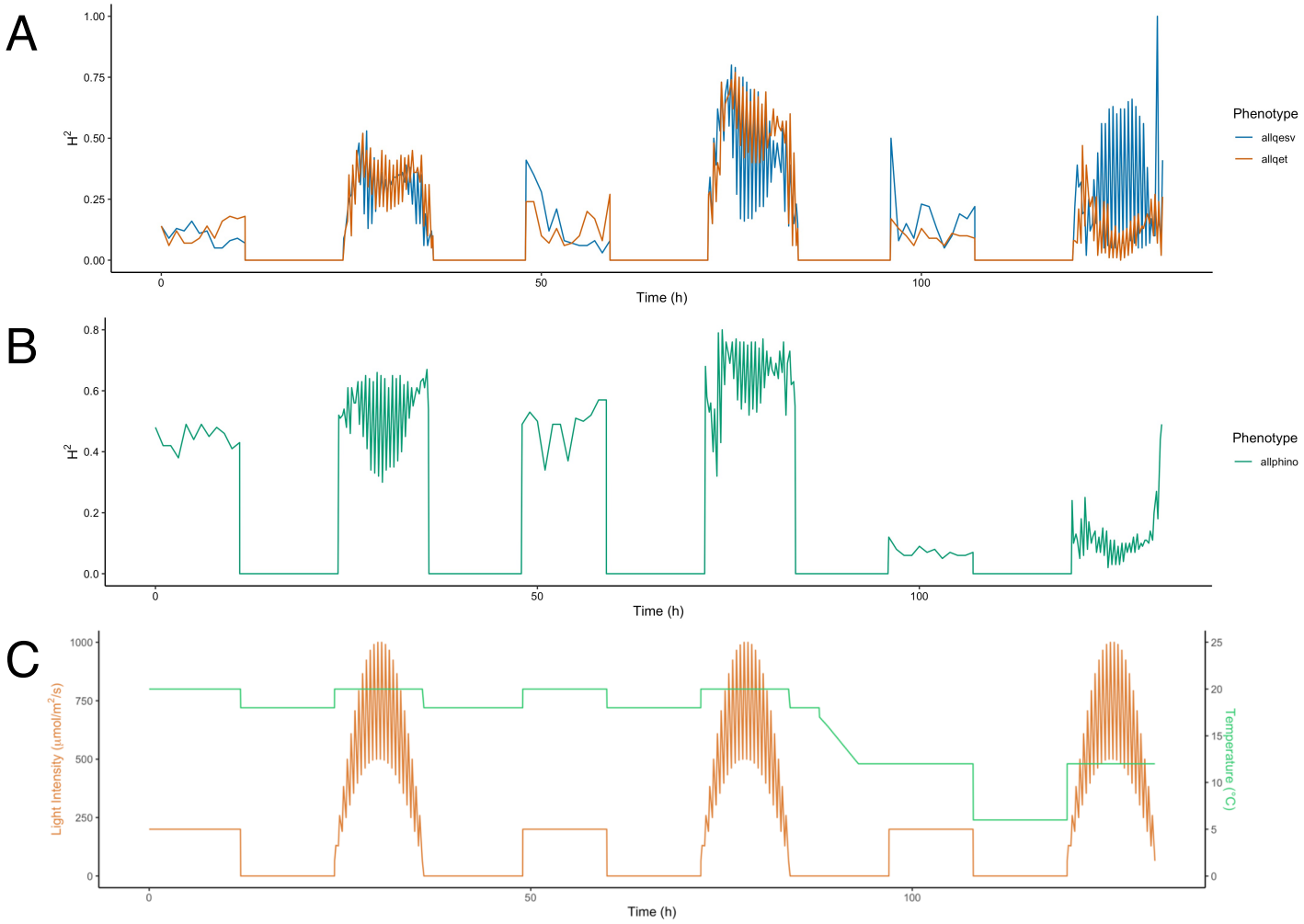
Yamori, W., Makino, A., Shikanai, T. (2016). A physiological role of cyclic electron transport around photosystem I in sustaining photosynthesis under fluctuating light in rice. *Scientific Reports*, *6*(July), 20147.

Zhang, N., Portis, A. R. (1999). Mechanism of light regulation of Rubisco: A specific role for the larger Rubisco activase isoform involving reductive activation by thioredoxin-f. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(16), 9438-9443.

Zhou, X., Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics*, *44*(7), 821-824.

Zhu, X., Long, S., Ort, D. (2010). Improving photosynthetic efficiency for greater yield. *Annual Review of Plant Biology*, *61*, 235-261.

Supplementary Figures

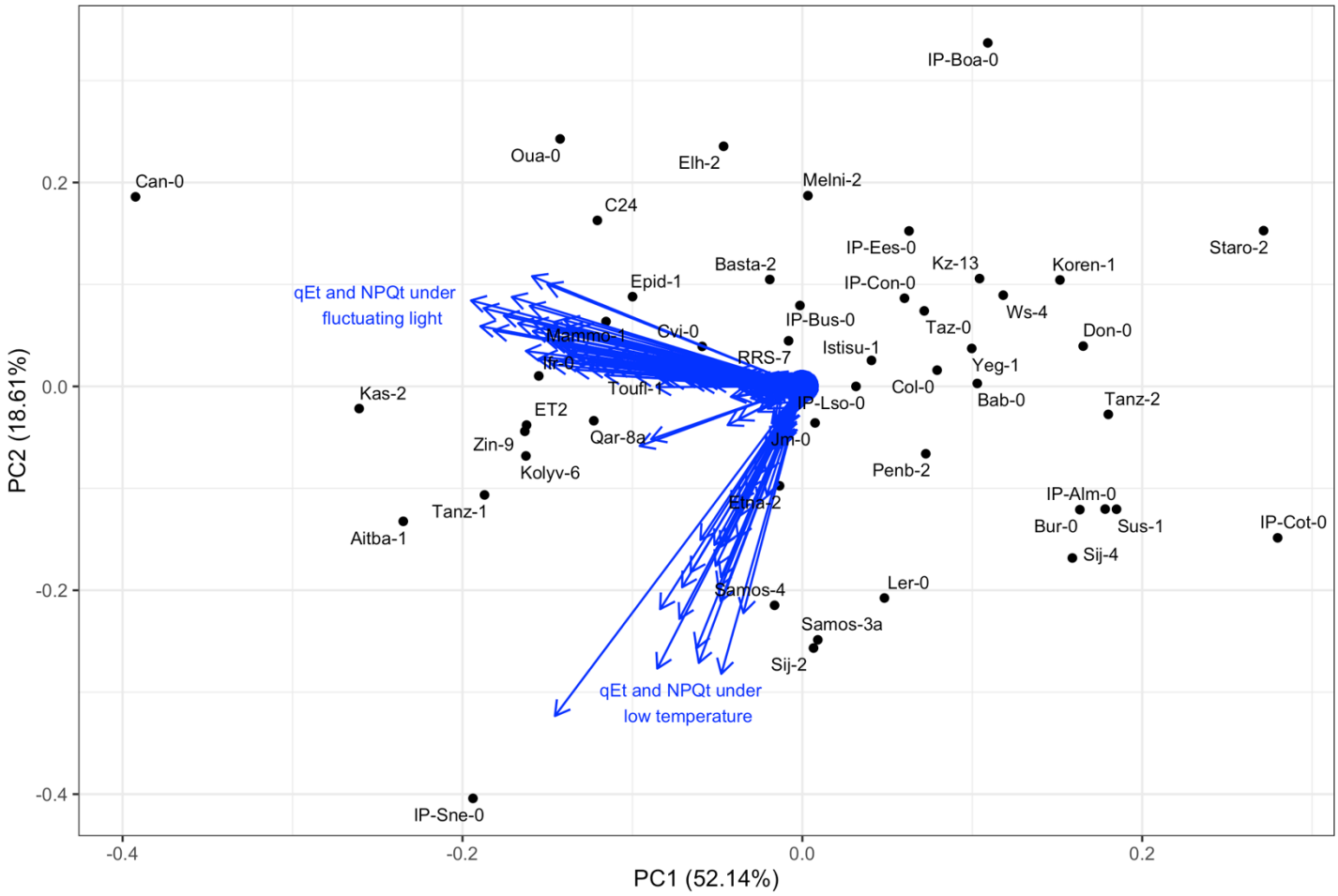


Supplementary Figure 1: The heritability of qE_{SV} and qE_t increases under fluctuating light conditions.

(A) The broad sense heritability (H^2) of qE_{SV} (blue) and qE_t (orange) over the course of the DEPI experiment.

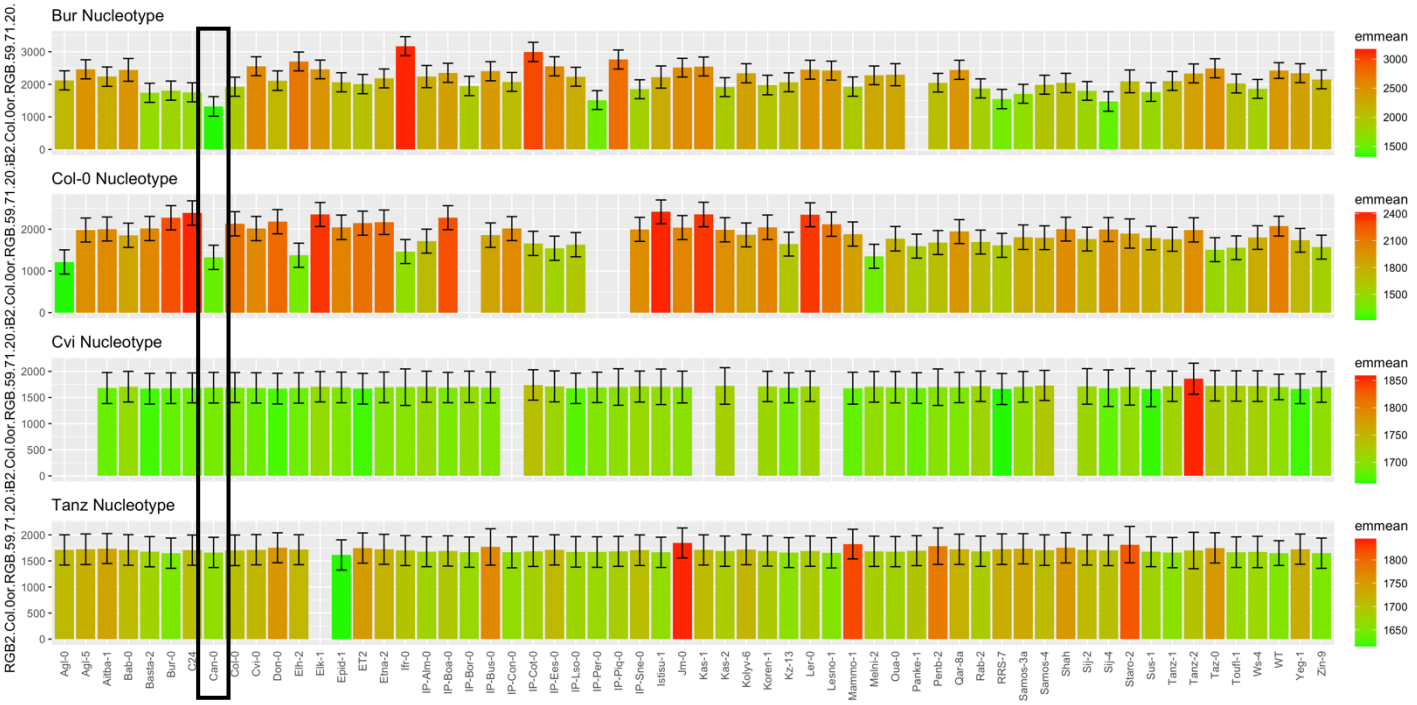
(B) The broad sense heritability (H^2) of ϕNO (green) over the course of the DEPI experiment.

(C) Light intensity and temperature over the course of the 6 days. This consisted of alternating days of stable light intensity (200 µmol/m²/s) and fluctuating light intensity (up to 1000 µmol/m²/s). The temperature was decreased from 20°C/18°C to 12°C/6°C for days 5 and 6.



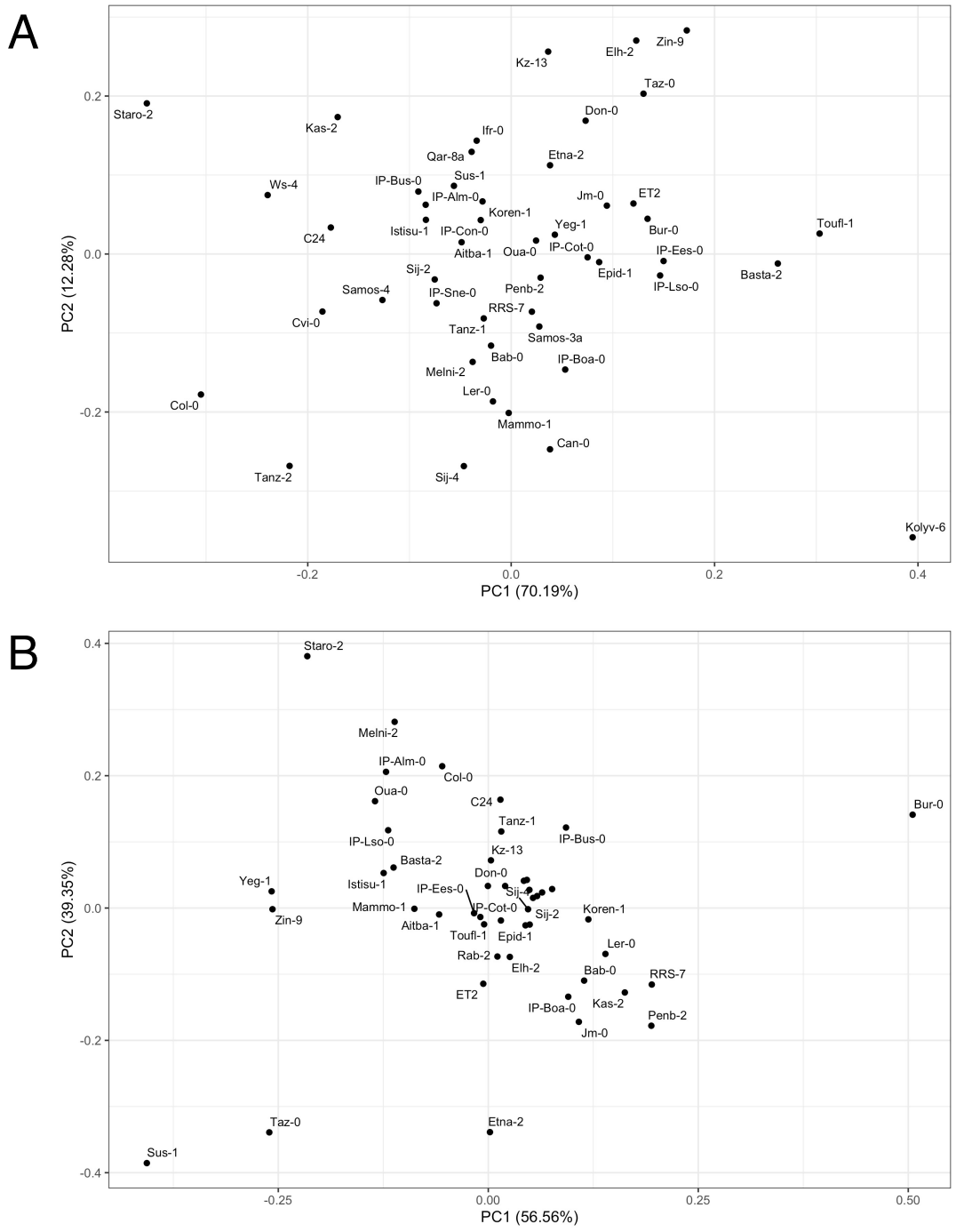
Supplementary Figure 2: qE_t and NPQ_t under fluctuating environmental conditions explain variation between the plasmotypes.

PCA biplot of the BLUEs of each plasmotype for the 1707 phenotypes measured in the DEPI experiment. Each point is labelled with the name of the plasmotype donor. The blue arrows indicate the loadings of the phenotypes in the PCA. Two groups of arrows are labelled, those that correspond to qE_t and NPQ_t under fluctuating light and those that correspond to qE_t and NPQ_t under low temperature.



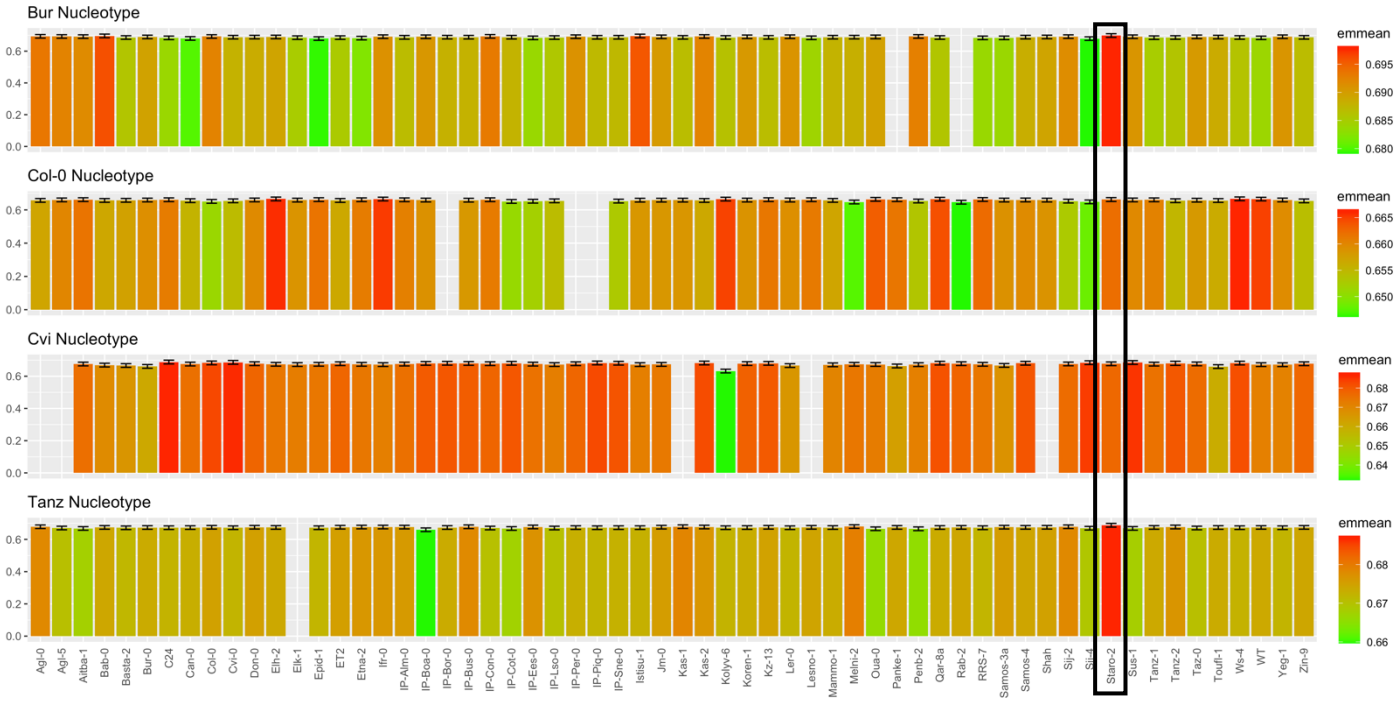
Supplementary Figure 3: The Can-0 plasmotype is associated with decreased green-ness of the leaves.

The BLUEs ± SE for each nucleotype-plasmotype combination for RGB 59.71.20 measured in the PlantScreen™ SC System for the cybrids grown in outdoor conditions. RGB 59.71.20 corresponds to the green-ness of the rosette. Colours correspond to the plasmotypes with the highest (red) and lowest (green) BLUEs within each nucleotype. The BLUEs for the Can-0 plasmotype are highlighted.



Supplementary Figure 4: Principal Component Analysis (PCA) reveals plasmotypes of interest from the Phenovator and outside tunnel experiments.

(A) PCA biplot of the BLUEs of each plasmotype for the ϕPSII measurements in the Phenovator. (B) PCA biplot of the BLUEs of each plasmotype for the photosynthetic parameters measured for the cybrids grown in outdoor conditions.



Supplementary Figure 5: The Staro-2 plasmotype is associated with high ϕPSII.

The BLUEs ± SE for each nucleotype-plasmotype combination for ϕPSII measured in the Phenovator. The BLUEs for ϕPSII at time 16.71 are shown. Colours correspond to the plasmotypes with the highest (red) and lowest (green) BLUEs within each nucleotype. The BLUEs for the Staro-2 plasmotype are highlighted.
