The handle [http://hdl.handle.net/1887/38624](http://hdl.handle.net/1887/38624) holds various files of this Leiden University dissertation.

**Author:** Tol, Niels van  
**Title:** Phenotypic engineering of photosynthesis related traits in Arabidopsis thaliana using genome interrogation  
**Issue Date:** 2016-03-24
Chapter 4

Chloroplast genome interrogation in
Arabidopsis seedlings

Niels van Tol, Vera Veltkamp, Johan Pinas,
Paul J.J. Hooykaas and Bert J. van der Zaal
Abstract

The large majority of core photosynthesis proteins in plants are encoded by nuclear genes, but a small portion has been retained in the plastid genome. Here, we report about the use of nuclear encoded, chloroplast targeted zinc finger artificial transcription factors (ZF-ATFs) to modulate the transcription patterns of chloroplast genes, a technique designated chloroplast genome interrogation. Using this system, we obtained evidence that ZF-ATFs can be translocated to chloroplasts, can induce phenotypic changes and can influence the operating light use efficiency of Photosystem II. Our data suggest that the distortion of chloroplast gene expression patterns might be a feasible approach to manipulate the efficiency of plant photosynthesis.
Introduction

Photosynthesis is the process that fixes solar energy as chemical energy. In green plant tissues it is conducted by specialized plastid organelles named chloroplasts, which harbor the core of the photosynthetic apparatus. Sunlight is absorbed by chlorophyll molecules that are associated with Photosystems I and II (PSI and PSII) that are anchored in the thylakoid membranes of chloroplasts, and catalyze the photoexcitation of electrons. The resulting linear electron transport leads to the photoreduction of NADP, and indirectly drives the synthesis of ATP through a pH gradient that is generated by the Cytochrome b$_6$f proton pump through chemiosmotic coupling. In the light, the energy rich compounds NADPH and ATP are used in the Calvin-Benson cycle for CO$_2$ fixation by the enzyme complex RuBisCo to yield a carbohydrate product that can be partitioned to different plant organs and used for various metabolic processes supporting plant growth and development.

During the domestication of photosynthetic bacteria as chloroplasts an estimated number of 4500 bacterial genes has been incorporated into the nuclear genomes of plants [1]. These genes have acquired eukaryotic gene expression signals and in many cases sequences encoding N-terminal signal peptides known as chloroplast transit peptides (CTPs), which mediate chloroplast import [2]. Remarkably, a small but significant portion of the bacterial genes has been retained within chloroplasts. In higher plants, these genes now reside on a single circular chromosome of 120-170 kb that is maintained in high copy numbers in the chloroplast stroma. Chloroplastic DNA therefore accounts for a very significant portion of the total cellular DNA, with up to 50 copies per chloroplast and up to 100 chloroplasts per cell in mature photosynthetic tissue [3]. The chloroplast genome of the model plant species *Arabidopsis thaliana* encodes 54 structural thylakoid membrane proteins, 31 proteins involved in the regulation of plastid gene expression and contains 45 tRNA and rRNA encoding genes [4].

The engineering of chloroplast genes has been designated as one of the targets for increasing the efficiency of photosynthesis in plants [5]. As many chloroplast encoded proteins have structural or catalytic roles in chloroplast function, the introduction of mutant alleles or orthologous genes from other photosynthetic organisms might result in more efficient thylakoid membrane function. For a limited number of plant species there are plastid transformation protocols available [6] to introduce gene constructs into target chloroplast genomic loci through homologous DNA recombination, but it is typically very tedious to generate homoplasmic plants with stably transgenic chloroplasts, especially in tissue in which the chloroplast genome is maintained in very high copy numbers. Another pitfall of this approach is that mutation of the core thylakoid membrane components usually results in impairment rather than gain of thylakoid membrane function. To our knowledge, no chloroplast mutants with enhanced photochemistry have been reported. We therefore
hypothesized that changing the stoichiometry rather than the structure of chloroplast encoded proteins might be an alternative and more promising approach to influence thylakoid membrane function. In order to change the stoichiometry of chloroplast encoded thylakoid membrane proteins, we considered artificial transcription factor (ATF)-mediated genome interrogation a suitable method.

The key principle of genome interrogation is based on the introduction of ATFs with low complexity DNA binding domains to induce large scale changes in gene expression patterns that might lead to different phenotypes of interest [7]. In our lab, we have successfully used ATFs with zinc fingers (ZFs) as DNA binding domains (ZF-ATFs) for genome interrogation experiments in Arabidopsis [8-10]. In our setup, the ZF-ATFs contained an array of three of the 16 different ZFs that can recognize a cognate 3 base pair (bp) consensus DNA sequence of 5’-GNN-3’ [11], with ‘N’ being any of the four bases. The ZF domains were fused to protein moieties that can either stimulate or repress transcription, such as the transcriptional activation domain of the herpes simplex VP16 protein or the EAR transcriptional repressor motif from Arabidopsis itself [12-16]. Gene constructs encoding these 3F-ATFs can be introduced into the nuclear plant genome through Agrobacterium tumefaciens-mediated floral dip transformation to obtain transgenic plants. The cognate 9 bp target site of each 3F-ATF will on average occur once in every 130,000 bp of double stranded DNA, and thus approximately 1000 times within the 130 Mbp Arabidopsis genome. In this way, the expression of nearby genes might be distorted by 3F-ATFs in trans and in a dominant manner, potentially leading to the differential expression of a large number of genes, which in turn might trigger novel phenotypes to arise.

In the present study we explored the use of ZF-ATFs in chloroplasts. To change chloroplast gene expression patterns it had to be taken into account that chloroplasts have also retained their own transcriptional and regulatory machinery [17], consisting of the phage-type nuclear-encoded RNA polymerase (NEP), which mostly transcribes plastid housekeeping genes, and the bacterial type plastid-encoded RNA polymerase (PEP), which mostly transcribes photosynthesis genes. The process of plastid gene expression is also tightly regulated through anterograde and retrograde signaling with the nucleus [18, 19]. In view of these considerations, we had to redesign our previously established genome interrogation setup in such a way that ZF-ATFs can function in a prokaryotic environment.

Here, we describe the construction of ZF-ATF expression cassettes that can be introduced into the nuclear plant genome using standard methods, and can result in ZF-ATF activity in chloroplasts. This system was tested by expressing chloroplast targeted fusions of the bacterial transcriptional activators CRP and LuxR to low complexity 2F DNA binding domains. We obtained evidence that a very small collection of ZF-ATFs already contained constructs that induced variation in the phenotype and operating light use efficiency of PSII reaction centers of Arabidopsis seedlings, indicating that manipulation of chloroplast gene expression patterns could further be explored as an option for the enhancement of plant photosynthesis.
Results

Design of the chloroplast genome interrogation system

Gene constructs encoding ZF-ATFs with novel features had to be designed for genome interrogation experiments in chloroplasts. Foremost, as described above, the expression of chloroplast genes is mediated by a system of polymerases and regulatory proteins that are of bacterial origin. As we had no guarantee that established modulators of eukaryotic gene expression could also function as such within a prokaryotic context, we decided to look for prokaryotic protein modules that could be direct activators of chloroplast gene expression. Firstly, we selected the *E.coli* Cyclic AMP Receptor Protein (CRP), which has been shown to activate *lac* gene expression in *E.coli* through a direct interaction with RNA polymerase [20]. CRP has also previously been used for genome interrogation experiments in *E.coli* [21]. For the present study we opted for the use of the C-terminal part of CRP consisting of amino acids 134-190 (designated CRPD2) [21], which lacks the cAMP binding domain and which is a more potent transcriptional activator than the full length CRP protein [21].

As a second option we selected the *Aliivibrio fischeri* protein LuxR, which is a regulator of *lux* gene promoters [22]. The C-terminal part of LuxR lacking the N-terminal amino acids 2-162 (designated LuxRΔN), was reported to contain the most critical amino acids for the interaction with RNA polymerase and to lead to inducer independent transcriptional activation activity in *A. fischeri* [23, 24]. Importantly, LuxRΔN was also shown to possess transcriptional transactivation activity in *E. coli* [25]. Based on the published characteristics of CRPD2 and LuxRΔN we thus hypothesized that both could be suitable modulators of PEP activity in chloroplasts without the requirement of other regulatory proteins.

As the concept of genome interrogation relies on generating relatively large changes in gene expression patterns we did not consider the use of 3Fs, as the 155 kbp chloroplast genome on average contains only one 3F binding site. Instead, we opted to make use of 2F domains, which have 6 bp DNA recognition sites that occur approximately 75 times in a typical chloroplast genome. Provided that the affinity of 2F domains for DNA is still high enough to allow for a preferential presence at these target sites, the activity of CRPD2 or LuxRΔN could lead to differential gene expression at many chloroplast genomic loci. In support of this idea, we have previously found that expression of different nuclear targeted 2F-ATFs can lead to transcriptional changes in Arabidopsis [9, 10]. More recently, we have found that salinity tolerance can be induced by a 2F-VP16 fusion (van Tol et al., manuscript in preparation). For chloroplast genome interrogation experiments, we thus decided to randomly select eight different 2Fs for ZF-ATF construction. These 2Fs were denoted as 2F1 through 8 (Table 2).

To avoid any inhibitory effects on the activity of the effector domains when translationally fused to 2F domains, we selected a flexible linker peptide optimized for LuxR activity [25] to function as a spacer between the 2F and effector domain modules. This linker consists
of five repeats of the peptide ARTQYSESM each separated by the amino acid G [25], and provides a distance of 150 Å between the 2Fs and the effector domains, which was determined to be optimal for LuxRΔN activity [25]. In order to achieve the translocation of ZF-ATFs into the stroma of chloroplasts, we chose to use the N-terminal chloroplast targeting peptide (CTP) of the FedA protein of Arabidopsis [26, 27]. This CTP has been shown to mediate the translocation of heterologous proteins into chloroplasts [28, 29]. We simultaneously chose to make use of the promoter of the FEDA gene to drive ZF-ATF expression.

An overview of the expression cassettes that were designed based on the considerations described is presented in Figure 1, and an overview of the amino acid sequences of the translational fusions encoded by the effector constructs is provided in Figure 2. CTP-mediated chloroplast translocation of ZF-ATFs was verified by confocal microscopy on the mesophyll tissue of Col-0 plants and second generation transformant (T2) plants harboring the construct pCTP-2F1-GFP. As expected, GFP signal could be detected in the chloroplasts of pCTP-2F1-GFP plants (Fig 3).

![Diagram of expression cassettes](image1)

**Fig. 1** Overview of expression cassettes that were generated for chloroplast genome interrogation experiments. Expression of all fusion constructs is under control of the promoter sequence of AtFEDA (pFEDA) and the NOS terminator sequence (tNOS). The names of the constructs are indicated on the left side of the panels. A) Control constructs encoding fusions of effectors (LuxRΔN or CRPD2) without DNA binding domains to an N-terminal linker (L) and with or without N-terminal chloroplast transit peptide (CTP). B) ZF-ATF encoding constructs consisting of fusions of eight different 2Fs (2Fn; n=1-8) to either LuxRΔN or CRPD2 through the linker sequence, either with or without CTP. C) Construct encoding a fusion of 2F1 and GFP through the linker sequence with an N-terminal CTP.
Characterization of primary transformants to establish the experimental setup

To investigate whether chloroplast genome interrogation can lead to phenotypic differences and altered photosynthetic performance, Arabidopsis Col-0 plants were transformed with T-DNA constructs encoding CTP-LuxRΔN and CTP-CRPD2 fusions without (Fig. 1A) or with 2Fs as DNA binding domains (Fig. 1B). A few dozen primary transformants were readily obtained for all constructs at approximately equal transformation efficiencies, indicating that none of the constructs had effects that are detrimental to embryonic development. We observed some variation in seedling phenotype, but this was randomly distributed and is thus likely to be due to kanamycin selection. When further cultivated on soil the primary transformants did not gain other conspicuous phenotypes that could be attributed to the expression constructs, nor to the presence of a CTP or of 2F domains. Based on these observations in the primary transformant stage we concluded that the expression of chloroplast targeted ZF-ATFs did not have any marked negative effects on chloroplast function.

Fig. 2 Overview of the amino acid (aa) sequences encoded by the ORFs of pCTP-Linker (A), pCTP-LuxR (B) and pCTP-CRP (C). An overview of the composition of these constructs is presented in Figure 1A. The linker, LuxRΔN and CRPD2 peptides are presented in green, blue and red, respectively. The insertion sites of the 2Fs are labelled ‘2F insert.’
To avoid the interference of phenotypic aberrations due to kanamycin selection with further quantitative experiments, we decided to assess the T2 progeny of the primary transformants without applying any antibiotic selection. To obtain a first indication of whether any of the expression cassettes had effects on chloroplast performance, we chose to harvest the seeds of 12 randomly chosen primary transformants representing a particular expression construct and to analyze a large number of seedlings consisting of an equal mixture of these 12 T2 lines. These mixtures where designated ‘T2-pools.’ As there are no practical options available to assess chloroplast genome wide transcription patterns we opted to use growth and photosynthesis as indicators of ZF-ATF induced changes in chloroplast performance.

Fig. 3 Confocal microscopy images of abaxial mesophyll tissue of the sixth leaf of Col-0 and pCTP-2F1-GFP (T2) plants at 25 dpg (20x magnification). Merged images are an overlay between 633 nm, GFP and bright field images.

Phenotypic and photosynthetic properties of mixed T2 populations
When sown on soil wild type Col-0 seedlings and T2-pool seedlings harboring the pCTP-Linker construct had germinated homogenously at 4 days post germination (dpg) and did not display conspicuous phenotypic variation (Fig. S1A). However, T2-pool seedlings harboring the constructs pCTP-CRP and pCTP-LuxR exhibited noticeable variation in size (Fig. S1A), suggesting that chloroplast targeting of CRPD2 and LuxRΔN without DNA binding domains can already influence seedling development. Among the seedlings harboring the constructs pCTP-2Fn-CRP and pCTP-2Fn-LuxR (Fig. S1B and Fig. S1C, respectively) the variation in growth and pigmentation was visibly more substantial than the variation induced by pCTP-CRP and pCTP-LuxR alone, indicating that ZF-ATFs can trigger phenotypic variation in Arabidopsis seedlings when targeted to chloroplasts.
To investigate whether LuxRNΔ and CRPD2 can have an effect on photosynthesis, we quantified the operating light use efficiency of PSII (φPSII) of populations of T2-pool seedlings by CF imaging. We used φPSII as a measure for photosynthesis because it can be quantified in a high throughput manner for large numbers of seedlings, and because there is a strong correlation between φPSII, linear electron transport rate and the rate of CO₂ fixation in plants [30]. As expected, seedlings harboring the construct pFEDA, which lacks an open reading frame, did not display changes in φPSII compared to Col-0 (Fig. 4). However, T2-pool seedlings harboring the constructs pCTP-LuxR and pCTP-CRP did display significant increases in φPSII compared to Col-0 (Fig. 4). These φPSII increases were not found for seedlings harboring the constructs pLuxR and pCRP (Fig. 4), indicating that LuxRNΔ and CRPD2 can only trigger φPSII increases when translocated to chloroplasts. For unknown reasons significant φPSII increases were also noted for seedlings harboring the constructs pCTP-Linker and pLinker, suggesting that the linker domain has an effect φPSII regardless of the presence of a CTP.

![Fig. 4](image-url)

**Fig. 4** Quantification of the operating light use efficiency of Photosystem II (φPSII) at 200 μmol m⁻² s⁻¹ of actinic light of Col-0 seedlings and seedlings harboring control constructs without putative chloroplast genome interrogation activity (7 dpg). The presented boxplots were generated from quadrant data obtained from approximately 250 seedlings (n=9). Asterisks (*) indicate significant differences with Col-0 (p < 0.05) determined by one-way ANOVA analysis.

To investigate whether ZF-ATF activity in chloroplasts can have an effect on photosynthesis, we quantified φPSII of T2-pool seedlings harboring the two series of ZF-ATF encoding constructs pCTP-2Fn-LuxR and pCTP-2Fn-CRP, respectively. There were significant increases in the φPSII of seedlings harboring the constructs pCTP-2F3-LuxR, pCTP-2F4-
CRP and pCTP-2F7-CRP compared to the respective empty vector control constructs lacking 2Fs as DNA binding domains (Fig. 5A and 5B). There was also a noticeable upward shift in the distribution of the φPSII data of T2-pool seedlings harboring the construct pCTP-2F1-LuxR which was not statistically significant, but which we considered to be of interest regardless. Altogether these data suggested that the activity of chloroplast targeted 2F1-LuxRΔN, 2F3-LuxRΔN, 2F4-CRPD2 and 2F7-CRPD2 fusions can lead to enhancement of φPSII.

Identification of individual 2F-LuxR lines with CTP-dependent φPSII increases

To corroborate the indications that ZF-ATFs can modulate φPSII and do so in a 2F dependent manner, we investigated the contribution of the individual lines that were used to generate T2-pools to the φPSII increases. In this way it could be confirmed that five out of eight independent pCTP-2F1-LuxR lines and three out of 12 independent pCTP-2F3-LuxR lines displayed significantly higher φPSII than the empty vector control pCTP-LuxR (Fig. 6). To examine whether the induction of these φPSII increases is dependent on chloroplast translocation of the ZF-ATFs, Col-0 plants were transformed with the constructs p2F1-LuxR and p2F3-LuxR, which encode the same ZF-ATFs but lack the CTP. Subsequently, the φPSII of populations of the T2 progeny of five randomly chosen primary transformants with these constructs was quantified. Seedlings harboring the constructs p2F1-LuxR and p2F3-LuxR displayed φPSII

---

**Fig. 5** Quantification of the operating light use efficiency of Photosystem II (φPSII) at 200 μmol m⁻² s⁻¹ of actinic light of seedlings harboring T-DNA constructs encoding chloroplast targeted 2F-LuxRΔN (A) or 2F-CRPD2 (B). Seedlings harboring constructs without 2Fs serve as negative controls. The presented boxplots were generated from quadrant data obtained from approximately 250 seedlings (n=9). Asterisks (*) indicate significant differences with the negative controls pLuxR and pCRP (p < 0.05) determined by one-way ANOVA analysis.
values which were either similar to or significantly lower than those of seedlings harboring the empty vector control construct pLuxR (Fig. 7A and B), indicating that 2F1-LuxRN and 2F3-LuxRN fusions can only induce increases in φPSII when translocated to chloroplasts.

We attempted to perform a similar analysis as described above for lines harboring the constructs pCTP-2F4-CRP and pCTP-2F7-CRP, but for unknown reasons the seed germination percentage of these lines declined steeply within a few months, making unbiased analysis of the seedling impossible.

Fig. 6 Determination of the contribution of individual T2 pCTP-2F-LuxR and pCTP-2F3-LuxR lines to the overall increase in φPSII of the complex mixture of these lines presented in Figure 5. The operating light use efficiency of Photosystem II (φPSII) was quantified at 200 μmol m⁻² s⁻¹ of actinic light. The presented boxplots were generated from quadrant data obtained from approximately 250 seedlings (n=9). Asterisks (*) indicate significant differences with pCTP-LuxR (p < 0.05) determined with a heteroscedastic T-test assuming unequal variance to account for transgene segregation.
Fig. 7 Quantification of the operating light use efficiency of Photosystem II (φPSII) at 200 μmol m⁻² s⁻¹ of actinic light of seedlings (11 dpg) harboring constructs encoding 2F1-LuxR (A) and 2F3-LuxR (B) fusions lacking a CTP (five independent lines each). Seedlings harboring the construct pLuxR serve as a negative control. The presented boxplots were generated from quadrant data obtained from approximately 250 seedlings (n=9). Asterisks (*) indicate significant differences with pLuxR (p < 0.05) determined with a heteroscedastic T-test assuming unequal variance to account for transgene segregation.

Plants harboring the constructs pCTP-2F1-LuxR and pCTP-2F3-LuxR display changes in growth and experience leaf damage at the later stages of development

All of the φPSII measurements described above were performed on young Arabidopsis seedlings. To examine the effect of the expression of chloroplast targeted 2F1-LuxRΔN and 2F3-LuxRΔN fusions on the growth of Arabidopsis plants in later stages of development, we quantified the rosette surface area (RSA) of segregation populations of T2 plants harboring the constructs pCTP-2F1-LuxR (lines 2, 8 and 10) and pCTP-2F3-LuxR (lines 1, 3 and 11). RSA was used as a non-destructive proxy for growth, because there is a strong correlation between RSA and biomass in Arabidopsis [31]. For both constructs, two out of three lines displayed a reduction in RSA throughout development compared to Col-0 (Fig. S2A), and one out of three displayed wild type growth, respectively (Fig. S2A). At 31 dpg, a substantial fraction of the T2 plants of five out the six lines displayed mild to severe leaf necrosis, ranging from small spots to partial or even complete death of mature leaves (Fig. S2B). Regardless, these observations did not indicate that 2F1-LuxR and 2F-LuxR have marked effects in the later stages of development.
Discussion

In this study, we have described the design of a novel system to perform chloroplast genome interrogation in Arabidopsis seedlings. This system was tested using two types of chloroplast targeted ZF-ATFs consisting of fusions of the bacterial transcriptional activators CRPD2 and LuxRΔN to arrays of 2Fs as DNA binding domains. Using a relatively small number of 2Fs we have found evidence that both types of chloroplast targeted ZF-ATFs can induce phenotypic variation of Arabidopsis seedlings and can modulate their φPSII.

Although chloroplast genome engineering is considered an option for manipulating plant photosynthesis [5], the transformation of chloroplasts and subsequent selection of homoplasmic plant lines is an inherently tedious procedure and has not yet been reported for any commercially important plant species. However, the chloroplast genome interrogation system that we have investigated in this study should allow for the in trans manipulation of gene expression patterns in all chloroplasts and all copies of the chloroplast genome simultaneously through the integration of a single artificial gene in the nuclear genome. Since nuclear transformation protocols have become available for numerous plant species, the chloroplast genome interrogation system could in principle readily be applied to commercially interesting plant species without the requirement of detailed a priori knowledge regarding their plastid biology.

Using φPSII as read-out we have gathered evidence that several 2F and effector domain combinations from a rather limited pool can already have effects on chloroplast performance. Even though 2Fs are expected to bind to DNA with much lower specificity and lower affinity than 3Fs [8] and recognize just 6 or 7 bp of DNA (the latter due to target site overlap [32]), our data indicate that they still have sufficiently distinct binding specificity and affinity for DNA to interact with the chloroplast genome in planta. The use of the promoter of the FedA gene combined with the sequence encoding its CTP is therefore likely to have led to a ZF-ATF protein concentration in the chloroplast stroma that is sufficiently high to allow for 2F-specific DNA interactions. In the present study we did not yet attempt to directly investigate ZF-ATF induced transcriptional changes in chloroplasts because this has a number of practical limitations. Our data therefore do not yet allow for any conclusive statements regarding the activities of the LuxRΔN or the CRPD2 domains. Our evidence in terms of φPSII indicates that they do modulate chloroplast transcription. As mentioned above, there was a marked decrease in the viability of seeds harboring pCTP-CRP constructs, forcing us to refrain from further analysis of these seedlings. Even though CRPD2 is seemingly detrimental to long term seed survival, two out of the eight T2-pools harboring pCTP-2Fn-CRP constructs still exhibited an increase in φPSII, which was comparable to the two out of eight T2-pools harboring pCTP-2Fn-LuxR constructs that were further analyzed. It also has to be taken into account that LuxRΔN still possesses some putative DNA binding activity [25], meaning
that it could potentially bind to DNA regardless of the presence of a 2F domain. This might also explain why the seedlings harboring the construct pCTP-LuxR have significantly higher \( \phi_{\text{PSII}} \) values than Col-0 seedlings.

The initial \( \phi_{\text{PSII}} \) quantifications were performed on complex populations consisting of T2 seedlings from multiple lines (T2-pools) that were segregating for the transgenes. As the observed increases in \( \phi_{\text{PSII}} \) were relatively small, this could have been attributable to a single or small number of lines with very pronounced increases in \( \phi_{\text{PSII}} \). We also noted that the ZF-ATF encoding constructs on average had a substantial effect on seedling growth and pigment composition (Fig. S1B and C), which could to some extent have masked otherwise higher average \( \phi_{\text{PSII}} \) values. To be able to attribute the \( \phi_{\text{PSII}} \) increases specifically to ZF-ATFs, further verification of the phenotype in individual T2 seedling populations showed to be essential. As described, the majority of 5 out of 8 independent pCTP-2F1-LuxR T2 lines exhibited a significantly higher \( \phi_{\text{PSII}} \) levels (Fig. 5A), and 3 out of 12 independent pCTP-2F3-LuxR lines did so as well (Fig. 5B). With these rather large fractions of lines exhibiting the phenotype, we consider the conclusion that both pCTP-2F1-LuxR and pCTP-2F3-LuxR can enhance \( \phi_{\text{PSII}} \) justified. Not all transgenic lines did necessarily display the phenotype, but this might very well be dependent on individual differences in transgene expression levels due to differences in copy number and/or insertion site(s).

Even though our data suggest that ZF-ATFs can induce increases in the \( \phi_{\text{PSII}} \) of Arabidopsis seedlings, we have not observed that this is translated into more growth (Fig. S2), possibly suggesting that the \( \phi_{\text{PSII}} \) increases do not lead to more efficient overall photosynthesis. We think that this is most likely explained by the fact that \( \phi_{\text{PSII}} \) increases induced at the level of chloroplast gene expression also require differential retrograde signalling to the nucleus [4, 18], as both growth and photosynthesis are performed and regulated at the cellular and tissue levels. We therefore consider it likely that the translation of chloroplast performance to more growth and more efficient overall photosynthesis requires additional differential expression of nuclear genes. There is also debate about whether or not photosynthetic efficiency is at all positively correlated with biomass accumulation [33]. Quite recently, a transplastomic tobacco mutant expressing a hybrid RuBisCo with a spectacularly high rate of CO\(_2\) fixation was generated [34], but these plants are rather small and albino, further suggesting that is difficult to synchronize chloroplast photosynthetic performance with overall plant performance.

In conclusion, based on basic knowledge of chloroplast biology and without requiring further \textit{a priori} knowledge of the chloroplast genome we have successfully designed a novel chloroplast genome interrogation system. Using a relatively small setup we have already found evidence that ZF-ATF mediated chloroplast genome interrogation can induce small but significant changes in the photosynthetic performance of chloroplasts. Altogether our work suggests that it would be worthwhile to further investigate chloroplast genome interrogation as a novel tool to enhance the photosynthetic performance of plants.
Materials and methods

Growth conditions and plant material
The Arabidopsis accession Columbia-0 (Col-0) was used as the wild type and as the background genotype for all transformations described below. All seeds were stratified for 3-4 days at 4 °C prior to the experiments. Soil grown seedlings and plants were cultivated in a climate controlled growth chamber at a constant temperature of 20 °C, 70% relative humidity, a light intensity of approximately 200 µmol m⁻² s⁻¹ of photosynthetically active radiation (PAR), and at a 12 h photoperiod (referred to as ‘standard growth conditions’). Primary floral dip transformants were first grown on selection medium in a climate controlled tissue culture chamber at a constant temperature of 20 °C, 50% relative humidity, a PAR light intensity of approximately 50 µmol m⁻² s⁻¹, and at a 16 h photoperiod, subsequently transferred to soil after approximately 3 weeks and were further cultivated at standard growth conditions.

Construction of Arabidopsis plant lines expressing ZF-ATFs
A library of plasmids containing DNA fragments encoding all 256 different 2Fs was previously constructed [35]. Eight different 2Fs consisting of two different ZFs were randomly selected from this library (Table 1). The DNA sequence of CRP was derived from NCBI (REFSEQ accession NC_000913.2). The amino acid (aa) sequence of CRPD2 was derived from Lee et al., 2008 [21]. The aa residues 134-136 (NLA) were also included, as these were reported to constitute a flexible hinge [36]. The DNA sequence of LuxR was derived from NCBI (accession M25752, version 1). The aa sequences of LuxRΔN and the flexible linker were derived from Volzing et al., 2011 [25]. The promoter of AtFEDA (At1g60950) including the 5'-UTR sequence and the sequences encoding the CTP and the first 8 amino acids of FedA was amplified by PCR from the genomic DNA of Col-0 using the forward primer pFEDA FW (Table 2) and the reverse primers pFEDA REV1 and pFEDA REV2 (Table 2), yielding a 2029 bp pFEDA fragment. The sequences of all DNA fragments obtained by the insertion of oligonucleotides or PCR products were verified by Sanger sequencing (Macrogen Europe, Amsterdam). The binary vector plasmid pRF [9] was used as the backbone for all cloning steps.

The RPS5A promoter sequence and the 3F-VP16 ORF [9] were removed, and the pFEDA fragment was subsequently ligated in, yielding the plasmid pFEDA. A 700 bp oligo DNA fragment (Fig. S3) encoding the flexible linker, LuxRAN and CRPD2 (codon optimized for Arabidopsis) was synthesized by the company ShineGene (Shanghai, China), and was ligated into pFEDA. Either one or both of the two effector encoding modules were subsequently removed, yielding the plasmids pCTP-CRP, pCTP-LuxR and pCTP-Linker, respectively. The eight randomly selected 2F fragments were each ligated into pCTP-CRP, pCTP-LuxR and pCTP-Linker as SfiI fragments, yielding the plasmids which were designated pCTP-2Fn-
CRP, pCTP-2Fn-LuxR and pCTP-2Fn-Linker, respectively. Using pFEDA as a template, a PCR product lacking the CTP was generated using the primer combination pFEDA FW and MASTAL REV (Table 2), and was ligated into pCTP-CRP, pCTP-LuxR and pCTP-Linker, yielding the plasmids which were named pCRP, pLuxR and pLinker, respectively. The DNA sequence encoding eGFP was amplified by PCR using the forward primer GFP FW and reverse primer GFP RV (Table 2), and ligated into pCTP-2F1-Linker, yielding the plasmid designated pCTP-2F1-GFP. Plasmid sequences are available upon request. Col-0 plants were transformed with each of the generated constructs separately using the floral dip method [37] as described previously [9].

**Table 1.** 2Fs that were randomly assembled for chloroplast genome interrogation binary vector construction.

<table>
<thead>
<tr>
<th>Name</th>
<th>5′-3′ DNA recognition sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>2F1</td>
<td>GTC-GGG</td>
</tr>
<tr>
<td>2F2</td>
<td>GGG-GGA</td>
</tr>
<tr>
<td>2F3</td>
<td>GGA-GAG</td>
</tr>
<tr>
<td>2F4</td>
<td>GAG-GAT</td>
</tr>
<tr>
<td>2F5</td>
<td>GGG-GTA</td>
</tr>
<tr>
<td>2F6</td>
<td>GAT-GTC</td>
</tr>
<tr>
<td>2F7</td>
<td>GCC-GCT</td>
</tr>
<tr>
<td>2F8</td>
<td>GGA-GCC</td>
</tr>
</tbody>
</table>

**Table 2.** Primers that were used for the construction of the library of chloroplast genome interrogation binary vector constructs.

<table>
<thead>
<tr>
<th>Name</th>
<th>5′-3′ DNA sequence (restriction site underlined)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pFEDA FW</td>
<td>GGTCGACTGCCCCATTACGAAAAGATTCCGATTTGG (SalI)</td>
</tr>
<tr>
<td>pFEDA REV1</td>
<td>CCTCTGAGGTAACTTGACCTTGAGCTGGATGCACTTGTATGAGC (XhoI)</td>
</tr>
<tr>
<td>pFEDA REV2</td>
<td>GGAGCTCAGGCTCTCGAGGATTACCTTGAGCTGACCTTG (SacI)</td>
</tr>
<tr>
<td>MASTAL REV</td>
<td>GCTCGAGAGCAGTGGAAGCCATTTTTTTTTT (XhoI)</td>
</tr>
<tr>
<td>GFP FW</td>
<td>GACTAGTGAGCAAGGGCGAGGAGCTGTTCCACCG (SpeI)</td>
</tr>
<tr>
<td>GFP RV</td>
<td>GGAGCTCTTATTGATACAGTCCGTCCATGCGG (SacI)</td>
</tr>
<tr>
<td>FEDA FW</td>
<td>CACGCCATTTCCACAAGC</td>
</tr>
</tbody>
</table>

**Confocal microscopy**

Confocal microscopy was performed on the abaxial tissue of the sixth leaf of Col-0 plants (25 dpg) and T2 plants harboring the construct pCTP-2F1-GFP using a Zeiss L5M5 Exciter (Zeiss, Jena, Germany) at 20x magnification. Excitation of the tissue was performed with a 488 nm laser. All leaves were imaged at the same laser power. The emission of chlorophyll fluorescence was collected with a 560 nm long pass filter and GFP fluorescence was collected with a 505 to 530 nm band pass filter. Merges between chlorophyll fluorescence, GFP fluorescence and bright field images were generated using ImageJ.
φPSII quantification of Arabidopsis seedlings
Approximately 250 T2 seeds from either T2-pools (as explained in the results section) or from individual T1 parents were sown on soil in pots with a diameter of 15.7 cm and height of 65 mm (Soparco, Condé-sur-Huisne, France). At one time point in development (varying between experiments depending on seedling growth rate), the operating light use efficiency of Photosystem II ($\Phi_{m}^{'}/F_{m}^{'}$ [30]; referred to as φPSII) was quantified using a CF Imager (Technologica, Essex, United Kingdom). The seedlings were exposed to 200 μmol m$^{-2}$ s$^{-1}$ actinic light in the CF imaging chamber (the same light intensity as in the growth chamber) for 1 min, after which $F_{m}^{'}$ and $F_{m}^{'}$ images were generated by exposure to a 6226 μmol m$^{-2}$ s$^{-1}$ saturating actinic light pulse. $\Phi_{m}^{'}/F_{m}^{'}$ images were constructed from the $F_{m}^{'}$ and $F_{m}^{'}$ images. Local differences in seedling density along the surface of individual pots led to variation in $\Phi_{m}^{'}/F_{m}^{'}$ images. To account for this variation, images containing data of 250 seedlings were subdivided into 9 equal and non-overlapping quadrants, each representing the average φPSII of approximately 28 seedlings. As the quadrants did not overlap and the independent seedlings grew in different parts of large pots, the quadrants were considered biological replicates (n=9). For the first comparative measurements of T2-pools (Fig. 4 and 5) the quadrant data were statistically analyzed by one-way ANOVA. For comparisons of individual T2 lines (Fig. 6 and 7) the heteroscedastic T-Test function of Microsoft Excel was used, assuming unequal variance between samples due to possible differences in transgene segregation. In all cases a $p$-value of 0.05 was used as a threshold for significance.

Growth analysis of pCTP-2F1-LuxR and pCTP-2F3-LuxR lines
For rosette surface area quantification, approximately 50 seeds of Col-0 and of independent pCTP-2F1-LuxR and pCTP-2F3-LuxR lines were sown in pots with a diameter of 15.7 cm and height of 65 mm (Soparco, Condé-sur-Huisne, France), and stratified at 4 °C for 3 days. Seedlings were transferred to 67 x 67 x 65 mm pots (Pöppelmann, Lohne, Germany) at 7dpg. Photos were taken of all trays from the top with a fixed digital camera (Canon EOS 1100D) from 10 dpg onwards and every two to five days. The surface area of each rosette was subsequently calculated in pixel$^{2}$ using the ‘Analyze Particles’ function of ImageJ, and then converted to mm$^{2}$ by multiplying this value by the mm$^{2}$/pixel$^{2}$ ratio of each RGB image.

Acknowledgements
We would like to thank Paul de Mooij for his help in the initial stages of the experimental design of this study. We would like to thank Prof. Dr. H.J.M. de Groot and Dr. A. Alia Matysik (Leiden Institute of Chemistry, Leiden University, the Netherlands) for stimulating discussions. This work was carried out within the research programme of BioSolar Cells, co-financed by the Dutch Ministry of Economic Affairs.
References


Fig. S1 An overview of the phenotypes of Arabidopsis seedlings (4 dpg) harboring empty vector control constructs (A), 2F-CRPD2 encoding constructs (B) and 2F-LuxR encoding constructs (C). An overview of the composition of the constructs is provided in Figure 1. Presented are populations of T2 seedlings that are segregating for the constructs, and have germinated from a complex mixture of seeds originating from 10-12 independent primary transformants (T1).
Fig. S2 Overview of growth of Col-0 plants, pCTP-2F1-LuxR plants (lines 2, 8 and 10; T2 generation; segregating for the expression construct) and pCTP-2F3-LuxR plants (lines 1, 3 and 11; T2 generation; segregating for the expression construct). A) Quantification of rosette surface area throughout development (n=33 for Col-0, n=14-18 for the transgenic lines). Error bars were not included for the sake of clarity of the figure. B) The rosette phenotypes of three representative individuals for each genotype at 31 dpg. The transgenic plants all display some degree of leaf damage or death, except for pCTP-2F3-LuxR,3 which was therefore not included.
Fig. S3 Overview of the 700 bp oligo DNA sequence that was synthesized for the construction of chloroplast genome interrogation expression cassettes. Sequences encoding the linker, LuxRΔN and CRP are presented in green, blue and red font, respectively. Restriction sites are underlined. Stop codons are labelled ‘STOP’.