Identification of small non-coding RNAs responsive to *GUN1* and *GUN5* related retrograde signals in *Arabidopsis thaliana*

Kristin Habermann[1] (iD), Bhavika Tiwari[1], Maria Krantz[2], Stephan O. Adler[2], Edda Klipp[2], M. Asif Arif[1,*] and Wolfgang Frank[1,*]

[1]*Plant Molecular Cell Biology, Department Biology I, Ludwig-Maximilians-Universität München, LMU Biocenter, Planegg-Martinsried 82152, Germany, and*
[2]*Department Biologie, Bereich Theoretische Biophysik, Humboldt-Universität Berlin, Berlin 10115, Germany*

SUMMARY

Chloroplast perturbations activate retrograde signalling pathways, causing dynamic changes of gene expression. Besides transcriptional control of gene expression, different classes of small non-coding RNAs (sRNAs) act in gene expression control, but comprehensive analyses regarding their role in retrograde signalling are lacking. We performed sRNA profiling in response to norflurazon (NF), which provokes retrograde signals, in *Arabidopsis thaliana* wild type (WT) and the two retrograde signalling mutants *gun1* and *gun5*. The RNA samples were also used for mRNA and long non-coding RNA profiling to link altered sRNA levels to changes in the expression of their cognate target RNAs. We identified 122 sRNAs from all known sRNA classes that were responsive to NF in the WT. Strikingly, 142 and 213 sRNAs were found to be differentially regulated in both mutants, indicating a retrograde control of these sRNAs. Concomitant with the changes in sRNA expression, we detected about 1500 differentially expressed mRNAs in the NF-treated WT and around 900 and 1400 mRNAs that were differentially regulated in the *gun1* and *gun5* mutants, with a high proportion (~30%) of genes encoding plastid proteins. Furthermore, around 20% of predicted miRNA targets code for plastid-localised proteins. Among the sRNA–target pairs, we identified pairs with an anti-correlated expression as well pairs showing other expressional relations, pointing to a role of sRNAs in balancing transcriptional changes upon retrograde signals. Based on the comprehensive changes in sRNA expression, we assume a considerable impact of sRNAs in retrograde-dependent transcriptional changes to adjust plastidic and nuclear gene expression.

Keywords: small non-coding RNA, non-coding RNA, gene regulation, retrograde signalling, *gun1*, *gun5*, *Arabidopsis thaliana*.

INTRODUCTION

Both mitochondria and chloroplasts are characteristic organelles of eukaryotes that have evolved through the endosymbiosis of distinct prokaryotic progenitors (Goksoyr, 1967). Cyanobacteria gave rise to plastids, and the majority of the endosymbiotic cyanobacterial genome was transferred into the nuclear DNA of the host organism. Consequently, most multiprotein complexes within the plastids are formed by organellar- and nuclear-encoded proteins, requiring a well-coordinated expression of both genomes (Zimorski *et al.*, 2014; Zhao *et al.*, 2019a). The nuclear gene expression is controlled by plastid-to-nucleus retrograde signalling (Kleine and Leister, 2016; Chan *et al.*, 2016), which is proposed to be mediated by several factors. For example, norflurazon (NF), a specific inhibitor of the enzyme phytoene desaturase, which produces β-carotenoids from phytoene, causes repression of photosynthesis-associated nuclear genes (*PhANGs*) (Woodson *et al.*, 2011). Carotenoids are part of the light-harvesting complexes and protect the cells from photooxidative damage (Kim and Apel, 2013). In the presence of NF the chloroplast suffers from photooxidation, leading to characteristic bleaching symptoms of the green plant tissues caused by the degradation of chlorophyll (Breitenbach *et al.*, 2001). Several decades ago *Arabidopsis thaliana* mutant screens were performed to identify factors which specifically block the expression of *PhANGs* under conditions of chloroplast developmental prevention (Susek *et al.*, 1993; Mochizuki *et al.*, 2001; Meskauskiene *et al.*, 2001; Larkin *et al.*, 2003; Gray *et al.*, 2003; Gutierrez-Nava *et al.*, 2004; Ball *et al.*,

2004; Rossel *et al.*, 2006; Saini *et al.*, 2011). Several GENOME UNCOUPLED (*gun*) mutants were identified with disturbed retrograde signalling leading to a de-repression of *PhANGs*. Interestingly, five different *gun* mutants, *gun2* to *gun6*, are affected in the tetrapyrrole biosynthesis pathway (TPB). The *gun5* mutant has a defective regulatory CHLH subunit of the magnesium-chelatase (Mochizuki *et al.*, 2001). The *gun4* mutation also affects the subunit of the magnesium-chelatase, leading to an increased efficiency. The *gun2*, *gun3* and *gun6* mutants are impaired in heme oxygenase, phytochromobilin synthase and Fe-chelatase, respectively (Woodson *et al.*, 2011; Woodson *et al.*, 2013). Based on these studies, it has been proposed that chloroplast metabolites may act as retrograde signals (Kakizaki *et al.*, 2009). The *gun1* mutant is not related to the remaining *gun* mutants since *GUN1* encodes a member of the chloroplast-localised pentatricopeptide repeat proteins, which usually act in post-transcriptional processes (Tadini *et al.*, 2016). The *gun1* mutant is able to perceive signals from the TPB, plastid gene expression and redox state, but the mode of action of *GUN1* in retrograde signalling remains unknown (Kleine and Leister, 2016). Microarray studies have been performed to compare transcriptional changes of *A. thaliana* wild type (WT) and *gun1* and *gun5* mutants in response to NF, revealing a strong correlation between the *gun1* and the *gun5* mutant because a large number of genes were consistently regulated in both mutants, including de-repression of *PhANGs*.

To date, all studies analysing gene expression in various retrograde signalling mutants focused on the analysis of protein-coding genes. However, it is well known that classes of non-coding RNAs (ncRNAs), including long ncRNAs (lncRNAs) as well as small ncRNAs (sRNAs), have important functions in diverse biological processes because they mainly act in the control of gene expression (Wang and Chekanova, 2017; Huang *et al.*, 2019).

LncRNAs with a size larger than 200 nucleotides were shown to have important functions in the control of gene expression (Wierzbicki *et al.*, 2008; Dinger *et al.*, 2009) and to exert their function by various mechanisms. One specific role of lncRNAs is the regulation of mRNA splicing, where they can either activate or inhibit specific splicing events (Ma *et al.*, 2014). They also mediate epigenetic modifications and act in microRNA (miRNA) target mimicry, where the lncRNA harbours a miRNA binding site, causing miRNA binding and sequestration (Franco-Zorrilla *et al.*, 2007; Swiezewski *et al.*, 2009; Heo and Sung, 2011). A specific gene regulatory class of ncRNA comprises sRNAs with a size of 20–24 nucleotides. They can interfere with nuclear transcription by regulating epigenetic modifications (Khraiwesh *et al.*, 2010; Bannister and Kouzarides, 2011; Holoch and Moazed, 2015) or they can act post-transcriptionally by targeting RNAs, mediating RNA cleavage or translational inhibition (Meister and Tuschl, 2004; Bartel,

2004; Kim, 2005). sRNAs can be divided into two classes on the basis of their origin: hairpin RNA (hpRNA) and small interfering RNA (siRNA) (Axtell, 2013a). One of the most important classes of hpRNA are miRNAs, which are processed from stem-loop transcripts by DICER-LIKE1 enzymes (Park *et al.*, 2002; Meyers *et al.*, 2008) and guided through ARGONAUTE1 and RNA-induced silencing complex to their target RNAs by sequence complementarity to mediate their cleavage or translational inhibition (Wierzbicki *et al.*, 2008; Voinnet, 2009). Until now only one recent study reported on a functional role of miRNAs in retrograde signalling (Fang *et al.*, 2018). It was shown that tocopherols positively regulate the accumulation of 3′-phosphoadenosine 5′-phosphate (PAP), which is an inhibitor of exonuclease 2 (XRN2), which negatively regulates mRNA and pri-miRNA levels by degradation of 5′ uncapped mRNA. Moreover, miR395 mediates cleavage of the mRNA encoding ATP sulfurylase (APS), the enzyme catalysing the initial step of PAP synthesis (Fang *et al.*, 2018).

Two other sRNA classes are formed from double stranded RNAs (dsRNAs), which are derived from endogenous transcripts and generate natural antisense transcript-derived siRNA (nat-siRNA) or trans-acting siRNA (ta-siRNA), based on their specific biogenesis pathways. Nat-siRNAs are generated from two genes encoding overlapping transcripts in antisense orientation, leading to the formation of dsRNA molecules (Borsani *et al.*, 2005). Nat-siRNAs are processed from these dsRNAs and mediate subsequent cleavage of one of the initial overlapping transcripts. According to their genomic location, NAT pairs can be distinguished into *cis*-NAT pairs, generated from opposing DNA strands within an identical genomic region, and *trans*-NAT pairs, produced from transcripts encoded by separated genomic regions (Lapidot and Pilpel, 2006; Yuan *et al.*, 2015). The first identified nat-siRNA was shown to have an important function in salt stress adaptation of *A. thaliana* (Borsani *et al.*, 2005), where it is involved in the regulation of proline biosynthesis. Unlike nat-siRNAs, ta-siRNA generation is triggered by miRNAs, since ta-siRNA precursor transcripts are cleaved in a miRNA-dependent manner and further processed into phased 21 nt ta-siRNA duplexes to control target RNAs (Chen, 2009). The role of sRNAs in retrograde signalling has not been analysed yet and information on the role of lncRNAs in retrograde control is completely lacking. To gain information whether these classes of ncRNA act in retrograde signalling, we made use of two well-characterised mutants affecting plastid-to-nucleus signalling events. *A. thaliana* *gun1* and *gun5* mutants were grown under standard conditions and in the presence of NF, and RNA expression profiles were compared to WT controls to identify functional sRNA–RNA target pairs that are modulated by retrograde signals.

RESULTS

De novo sRNA sequencing after norflurazon treatment

To identify sRNAs that may act in retrograde signalling pathways, seedlings of *A. thaliana* WT and the two retrograde signalling mutants *gun1* and *gun5* were treated for 4 days with 5 μM NF under continuous light (Figure S1a) and sRNA sequencing was performed from six independent biological replicates samples, yielding a minimum of 5 million reads per replicate. The length distribution of all sRNA reads was analysed and we observed an enrichment of reads with a length of 21 and 24 nt (Figure S1b–d). The 21 nt peak corresponds to an expected enrichment of miRNAs, ta-siRNAs and nat-siRNAs, whereas the 24 nt peak complies with enriched repeat-associated sRNAs. The ShortStack sRNA analysis software has been used to map the sRNA data set against different reference databases (Table S1).

DeSeq2 was used to calculate the differential expression (2-fold regulation and false discovery rate [FDR] $\leq$ 0.05) of sRNAs between the samples, with a special focus on sRNAs that were differentially expressed in NF-treated samples with respect to their untreated controls, and in NF-treated *gun* mutants compared to the NF-treated WT (Table S2). Specific sRNA clusters arising from different ncRNA classes were found to be differentially expressed (Figure 1). These classes include mature miRNAs, *cis*-nat-siRNAs and *trans*-nat-siRNA, as well as sRNAs derived from lncRNAs. Upon growth on normal media, we identified only a small number of differentially regulated sRNAs in the *gun* mutants as compared to the WT, whereas the number of differentially regulated sRNAs between the mutants and WT strongly increased upon NF treatment (Table S3).

NF treatment caused an increased number of differentially expressed sRNAs in WT and both *gun* mutants, indicating a considerable sRNA regulation by retrograde signals. Furthermore, we observed a higher number of differentially expressed sRNAs in both NF-treated *gun* mutants compared to NF-treated WT, pointing to a strong regulation of sRNAs that underlies specific retrograde signals in these mutants. Most of the changes affect miRNA and nat-siRNA expression levels, and we mainly focused on these sRNA classes with regard to their differential expression and further target analysis to predict the regulatory functions of these sRNAs (Figure 1 and Table S3).

Analysis of differentially expressed miRNAs

Because beside a recent analysis of tocopherol-responsive miRNAs (Fang *et al.*, 2018) little is known about the role of miRNAs in retrograde signalling, we analysed changes in miRNA expression in response to NF in *A. thaliana* WT and in the *gun1* and *gun5* mutants. The comparison of differentially expressed miRNAs between the samples is shown in a hierarchically clustered heatmap (Figure 2a). Only a low number of differentially expressed miRNAs was observed in the untreated mutants compared to the WT control. In the *gun1* mutant (*gun1*/WT), only six differentially expressed miRNAs were detected, and only five miRNAs were detected in the untreated *gun5* mutant compared to the WT control (*gun5*/WT).

We hypothesised that miRNAs can play a role in retrograde signalling that should be reflected by an enrichment of differentially expressed miRNAs after NF treatment. Indeed, we observed a remarkable increase in the number of differentially expressed miRNAs in response to NF treatment with a similar number of NF-responsive miRNAs in the three analysed genotypes (Figure 2b). In total, we observed 22 miRNAs to be differentially regulated in the NF-treated WT compared to the untreated control (WT NF/WT). Twenty-four miRNAs were differentially expressed in the NF-treated *gun1* compared to the untreated *gun1* mutant (*gun1* NF/*gun1*), and 18 miRNAs were differentially regulated in the NF-treated *gun5* mutant compared to the untreated *gun5* control (*gun5* NF/*gun5*).
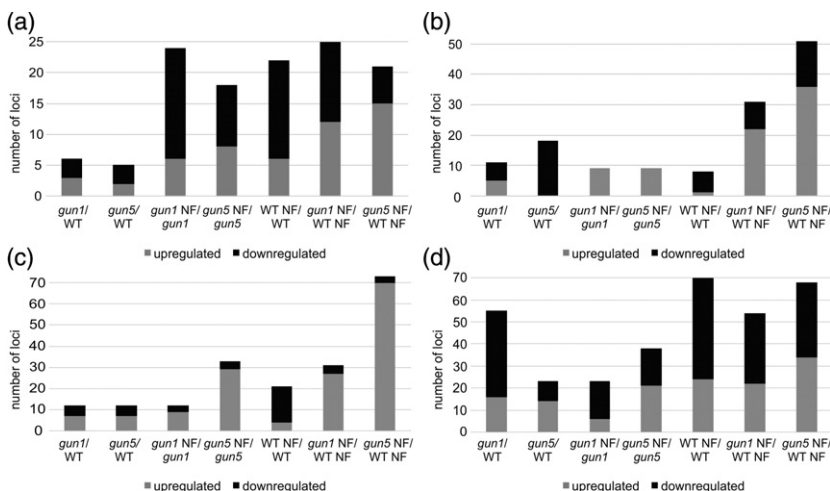

Figure 1. Differentially expressed sRNAs within the different samples. Overview of differentially regulated sRNAs between the different samples ($\log_2$(FC) $\leq$ −2 or $\geq$ +2; FDR $\leq$ 0.05) subdivided into specific sRNA classes. (a) miRNAs, (b) sRNAs derived from lncRNA, (c) *cis*-NAT pairs and (d) *trans*-NAT pairs. The up- and downregulation of the members of each class are depicted by grey (up) and black (down) partitions of the respective bars.

Figure 2. Behaviour of the differentially expressed miRNAs. (a) Hierarchically clustered (UPGMA) heat-map depicting miRNAs that are differentially regulated in at least one sample displaying normalised log$_2$(FC) values. (b) UpSet plot depicting the number of differentially expressed miRNAs in response to NF in WT (WT NF/WT) and both *gun* mutants (*gun1* NF/WT NF and *gun5* NF/WT NF).
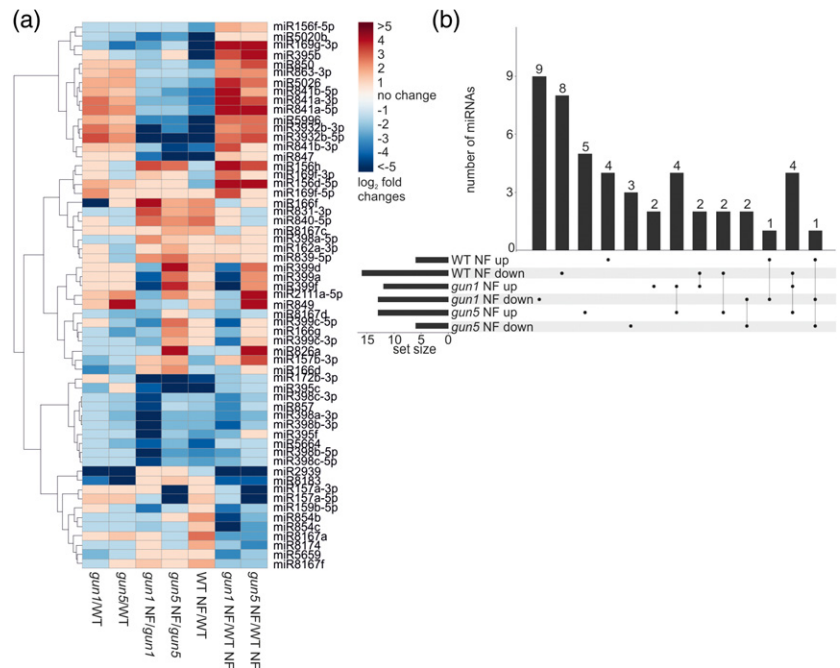
Interestingly, we further detected miRNAs that seem to be controlled by retrograde signals, as de-repressed miR-NAs were observed in the *gun1* and *gun5* mutants in response to NF treatment, which is reminiscent of the de-repression of *PhANGs* in these mutants. We focused on miRNAs with altered expression levels in the treated WT (WT NF/WT) and correlated them with differentially expressed miRNAs in the NF-treated *gun* mutants (Figure 2b). In response to NF, two miRNAs (miR169g-3p and miR5996) showed patterns of de-repression in both *gun* mutants similar to de-repressed *PhANGs*, and one miRNA (miR3932-5p) was de-repressed only in the NF-treated *gun5* mutant compared to the treated WT (*gun5* NF/WT NF). Furthermore, five miRNAs were downregulated in the treated WT (WT NF/WT) and were upregulated in at least one NF-treated *gun* mutant (*gun* NF/WT NF). We also iden-tified two miRNAs which seemed to be controlled by retro-grade signals in an opposite manner. These two miRNAs were found to be upregulated in the treated WT (WT NF/WT) and downregulated in at least one of the treated *gun* mutants (*gun* NF/WT NF). In addition, we found miRNAs which showed a specific regulation restricted to NF-treated *gun* mutants when compared to the treated WT. Two miR-NAs were found to be downregulated in both treated *gun* mutants (*gun* NF/WT NF). Moreover, nine miRNAs were specifically downregulated in the NF-treated *gun1* mutant compared to the treated WT (*gun1* NF/WT NF), and the expression of three miRNAs was reduced in the treated *gun5* mutant (*gun5* NF/WT NF). Furthermore, two miRNAs were upregulated in the treated *gun1* mutant and five miR-NAs were upregulated in the treated *gun5* mutant (*gun* NF/

WT NF). We also detected four upregulated miRNAs com-mon for both treated *gun* mutants (*gun* NF/WT NF).

Differentially regulated nat-siRNAs

To identify nat-siRNAs from our sRNA sequencing data we made use of different accessible databases (Table S1) com-prising experimentally validated and computationally pre-dicted *cis*- and *trans*-NAT pairs (Jin *et al.*, 2008; Zhang *et al.*, 2012; Yuan *et al.*, 2015).

We identified 12 various *cis*-NAT pairs producing differ-entially regulated nat-siRNA clusters in both untreated *gun1* and *gun5* mutants (*gun1*/WT and *gun5*/WT) (Fig-ure 1c). Besides this, 57 and 23 *trans*-NAT pairs were detected to produce differentially regulated nat-siRNA in the *gun1* and *gun5* mutants, respectively (Figure 1d).

Upon NF treatment we detected 21 *cis*-NAT pairs (Fig-ure S2a) and 70 *trans*-NAT pairs (Figure S2b) in the WT (WT NF/WT) producing differentially regulated nat-siRNA clusters from at least one transcript of these NAT pairs. In the treated *gun1* mutant (*gun1* NF/*gun1*), nat-siRNAs from 12 *cis*-NAT pairs were detected to be differentially expressed (Figure 1c) and 23 differentially regulated *trans*-NAT pairs producing nat-siRNA clusters were identified to be differentially regulated in the treated *gun1* mutant (*gun1* NF/*gun1*). In the NF-treated *gun5* mutant, we identi-fied 33 *cis*-NATs and 38 *trans*-NATs generating differen-tially expressed nat-siRNAs (*gun5* NF/*gun5*) (Figure 1c,d).

The overlap and co-regulation as well as specific expres-sion of the differentially expressed *cis*-NAT pairs and *trans*-NAT pairs producing differentially regulated nat-siRNA clusters were analysed between the samples and

are shown in an UpSet plot (Figure S2). We focused on the analysis of NF-responsive differentially expressed nat-siR-NAs in the WT to provide information on nat-siRNAs that are controlled by retrograde signals. Moreover, we compared NF-treated WT with both NF-treated *gun* mutants to identify NF-responsive nat-siRNA misregulation that is caused by the perturbed retrograde signals in these mutants. We detected 31 *cis*-NAT pairs and 54 *trans*-NAT pairs producing differentially expressed nat-siRNAs in the NF-treated *gun1* mutant (*gun1* NF/WT NF). In the NF-treated *gun5* mutant we detected 73 *cis*-NAT pairs and 68 *trans*-NAT pairs that produce differentially regulated nat-siRNA clusters (*gun5* NF/WT NF). For both *gun* mutants the majority of *cis*-derived nat-siRNAs were upregulated, whereas the majority of *trans*-derived nat-siRNAs were downregulated in these mutants (Figure S2). We identified five *cis*-NAT pairs to be downregulated in the treated WT (WT NF/WT) and upregulated in both treated *gun* mutants (*gun* NF/WT NF), thus representing the *gun*-specific de-repression of nuclear-encoded *PhANGs* (Figure S2a). We also detected one nat-siRNA cluster produced from a *cis*-NAT pair displaying an opposing expression pattern (upregulated in WT NF/WT and downregulated in both *gun* NF/WT NF). Within *trans*-derived nat-siRNAs we identified 19 sRNA clusters that were differentially regulated in response to NF in WT (WT NF/WT) and showed further differential regulation in response to NF in both *gun* mutants (*gun* NF/WT NF) (Figure S2b). Five of them resemble the *gun*-specific de-repression, since they were downregulated in the WT (WT NF/WT) and upregulated in both mutants (*gun* NF/WT NF). Eleven *trans*-derived nat-siRNAs displayed an opposite expression and were upregulated in the treated WT (WT NF/WT) and downregulated in both treated *gun* mutants (*gun* NF/WT NF). In addition, two differentially expressed nat-siRNA from *trans*-NAT pairs were downregulated within all three samples and another one was upregulated in the treated WT (WT NF/WT) and in the treated *gun5* mutant (*gun5* NF/WT NF) and downregulated in the treated *gun1* mutant (*gun1* NF/WT NF).

Other differentially regulated sRNA classes

Besides the differentially expressed miRNAs and NAT pairs we also found differentially expressed sRNAs produced from lncRNAs and phased siRNA (phasiRNAs) precursors (Figure S3 and Table S3). Similar to miRNAs and nat-siR-NAs, we detected only a small number of differentially regulated sRNA clusters derived from lncRNA precursors in the untreated genotypes (*gun*/WT). In total, 11 differentially expressed sRNA clusters produced from lncRNAs were noticed in the *gun1* mutant (*gun1*/WT) and all 18 differentially expressed sRNA clusters in the untreated *gun5* mutant were downregulated (*gun5*/WT).

When comparing the individual genotypes with and without NF treatment, we identified only a considerably small number of differentially expressed sRNAs. In the treated WT, eight sRNA clusters processed from lncRNA precursors were identified to be differentially expressed (WT NF/WT). For both treated *gun* mutants, we observed nine different upregulated sRNA clusters (*gun* NF/*gun*).

Comparing the NF-treated *gun* mutants with the NF-treated WT we noticed an increase in the number of differentially expressed sRNAs (Figure S3). Generally, we observed a higher number of upregulated sRNA clusters produced from lncRNA precursors in both treated *gun* mutants (*gun* NF/WT NF). Of 31 differentially expressed sRNA clusters, 22 were detected to be upregulated in the NF-treated *gun1* mutant (*gun1* NF/WT NF). For the treated *gun5* mutant, 36 out of 51 differentially expressed sRNA clusters were found to be upregulated (*gun5* NF/WT NF). We detected only one sRNA cluster derived from a lncRNA precursor that was downregulated in the treated WT (WT NF/WT) and upregulated in both NF-treated *gun* mutants (*gun* NF/WT NF). Two sRNA clusters were downregulated in the treated WT (WT NF/WT) and upregulated in the treated *gun1* mutant (*gun1* NF/WT NF). Another two sRNA clusters derived from lncRNA precursors were downregulated in the treated WT (WT NF/WT) and upregulated in the treated *gun5* mutant (*gun5* NF/WT NF). Furthermore, 11 sRNA clusters produced from lncRNA precursors were similarly regulated in both treated *gun* mutants (*gun* NF/WT NF), with six of them upregulated and five downregulated (Figure S3).

In addition to the lncRNA-derived sRNA clusters, we identified two differentially expressed phasiRNAs. One phasiRNA derived from locus AT1G63070 was 5.8-fold upregulated in the untreated *gun1* mutant (*gun1*/WT) and 4.1-fold upregulated in the NF-treated *gun1* mutant (*gun1* NF/WT NF). The second phasiRNA produced from the locus AT5G38850 was 2.6-fold downregulated in the treated WT (WT NF/WT) and 3.5-fold downregulated in the treated *gun1* mutant (*gun1* NF/*gun1*).

Analysis of lncRNA and mRNA in the *gun* mutants

Besides sRNA sequencing, we also sequenced mRNAs and lncRNAs to gain more information about NF-dependent regulation of lncRNAs and to examine the correlation of sRNAs with their targets.

The samples were mapped against the *A. thaliana* genome deposited in Araport11 (Table S4) and differential expression of mRNA and lncRNA between the samples (Tables S5 and S6) was calculated with Cuffdiff. Representative transcripts belonging to different RNA classes showing differential expression levels in the RNA sequencing data were selected for expression analyses by quantitative real-time PCR (qRT-PCR), which confirmed the mRNA and lncRNA sequencing data (Figure S4). We selected various genes which were detected to be differentially expressed in the treated WT as well as in both NF-treated *gun* mutants. Furthermore, we selected two transcripts each

that displayed a low, moderate and high abundance, respectively. In addition, we included one lncRNA that was found to be differentially regulated in all three samples.

Classification of differentially expressed ncRNAs detected via ribosomal depleted RNA sequencing

We identified differentially expressed transcripts belonging to distinct ncRNA classes (Tables S5 and S6), including lncRNAs, which may act in regulatory processes of gene expression, as well as tRNA, rRNA and small nucleolar RNA (snoRNA), which act in protein translation and splicing and usually have few regulatory functions. Under normal growth conditions we identified 10 differentially expressed ncRNAs in each of the *gun* mutants as compared to the untreated WT (Table 1 and Figure S5a). The number of differentially expressed ncRNAs increased upon NF treatment, indicating potential roles upon plastid perturbations that trigger retrograde signalling. In total, we identified 34 differentially expressed ncRNAs in the NF-treated WT compared to the untreated control (Table 1). In the NF-treated *gun1* and *gun5* mutants (*gun* NF/*gun*), we identified 32 and 70 differentially expressed ncRNAs, respectively. Interestingly, in the NF-treated *gun* mutants we observed 20 and 45 differentially expressed ncRNAs in the *gun1* and *gun5* mutants (*gun* NF/WT NF), respectively.

An UpSet plot (Figure S5b) depicts the distribution of differentially expressed ncRNAs between various samples (WT NF/WT, *gun1* NF/WT NF and *gun5* NF/WT NF). We identified two interesting lncRNAs (AT1G05562 and AT4G13495), which represent the classical *gun*-related expression as these show a downregulation in response to NF treatment in WT, but are upregulated in both NF-treated *gun* mutants. Furthermore, three lncRNAs (AT3G01835, AT5G07325 and AT5G07745) were identified to be upregulated in the NF-treated WT (WT NF/WT) and downregulated in the treated *gun1* mutant (*gun1* NF/WT NF).

Another interesting lncRNA (AT4G13495) was de-repressed in both NF-treated *gun* mutants with 7.2-fold and 3.4-fold upregulation in *gun1* and *gun5* mutants (*gun* NF/WT NF), respectively, whereas this lncRNA was highly downregulated (fold change [FC] of −10.5) in the treated WT (WT NF/WT). From our sRNA data we already detected sRNAs arising from this lncRNA and in agreement with the expression level of this lncRNA, the total sRNAs generated from this transcript were downregulated in the treated WT (FC of −2.8; WT NF/WT) and 3.4-fold upregulated in the treated *gun1* mutant (*gun1* NF/WT NF). Interestingly, this lncRNA overlaps with three individual miRNA precursors (miR5026, miR850 and miR863) in sense direction, suggesting that these miRNAs can be processed from the individual precursors as well as from the overlapping lncRNA. In line with this hypothesis, we observed a consistent differential expression of the lncRNA and the three individual miRNAs within the analysed samples (Table 2).

In addition, we identified two differentially regulated lncRNAs overlapping with mRNA transcripts in antisense that may act as precursors for the generation of nat-siRNAs. One lncRNA (AT1G05562) that may act as a natural antisense transcript was downregulated in the treated WT (FC of −3.7; WT NF/WT) and upregulated in the treated *gun1* and *gun5* mutants with a FC of 3.9 and 4.2 (*gun* NF/WT NF), respectively. This lncRNA transcript is able to overlap with an mRNA encoding an UDP-glucose transferase (AT1G05560). Furthermore, the overlapping mRNA transcript was downregulated in the treated WT (FC of −5.8; WT NF/WT) and upregulated in both treated mutants (FC of 3.6 for *gun1* NF/WT NF; FC of 3.1 for *gun5* NF/WT NF). We also detected differentially expressed sRNA clusters processed from this region in the sRNA sequencing data in the treated WT (FC of −4.4 for WT NF/WT) as well as in the treated *gun1* mutant (FC of 6.6 for *gun1* NF/WT NF). Thus, the regulation of nat-siRNAs correlates with the expression of the respective lncRNA–mRNA transcript pair and the differential expression seems to be regulated by specific retrograde signalling pathways.

Table 1 Overview of differentially expressed ncRNAs in response to NF in *A. thaliana* WT and *gun1* and *gun5* mutants

	gun1/ WT	*gun5*/ WT	WT NF/ WT	*gun1* NF/ *gun1*	*gun5* NF/ *gun5*	*gun1* NF/ WT NF	*gun5* NF/ WT NF
lncRNAs	6	5	15	13	34	11	20
snRNAs	0	0	3	0	8	1	5
snoRNAs	0	0	3	3	11	2	4
rRNAs	0	0	2	0	1	1	0
tRNAs	0	0	3	4	1	2	1
pseudogenes	4	5	5	6	9	2	10
transcript regions	0	0	2	5	3	1	4
MIR precursors	0	0	1	1	2	0	0
antisense RNAs	0	0	0	0	1	0	1
Total	10	10	34	32	70	20	45

Table 2 Expression data for the lncRNA AT4G13495 that overlaps in sense with the individual miRNA precursors miR5026, miR850 and miR863

ID	FC WT NF/WT	FDR	FC *gun1* NF/WT NF	FDR	FC *gun5* NF/WT NF	FDR
AT4G13495	−10.54	0.001	7.17	0.001	3.36	0.001
miR5026	−2.63	0.085	4.08	0.004	3.08	0.042
miR850	−2.58	0.077	2.61	0.068	3.81	0.007
miR863-5p	−1.7	0.561	1.29	0.941	2.34	0.467
miR863-3p	−1.51	0.538	2.7	0.025	2.61	0.045

In addition, we identified a *TAS3* precursor transcript (AT3G17185) that was downregulated in the NF-treated WT (FC of −2.9 for WT NF/WT) and de-repressed in the treated *gun5* mutant (FC of 2.6 for *gun5* NF/WT NF). Ta-siRNAs produced from the *TAS3* transcript control the expression of transcripts coding for auxin response factors such as *ARF2*, *ARF4* and *ETT*. However, we detected neither differentially expressed *TAS3*-derived ta-siRNAs nor differential expression of their cognate targets between the analysed samples.

Differentially regulated nuclear- and organellar-encoded mRNAs after NF treatment

In parallel to sRNA and lncRNA, we analysed the data obtained from the ribosomal depleted nuclear- (Figure 3) and organellar-encoded (Figure 4) RNA sequencing to identify protein-coding mRNAs that are regulated by retrograde signalling pathways. Furthermore, to categorise putative functions of differentially regulated RNAs after NF treatment, Gene Ontology (GO) enrichment terms were explored (Table S8 and Figure S6). We detected only a low number of differentially expressed genes (DEGs), with 212 and 165 differentially expressed transcripts in the untreated *gun1* and *gun5* mutants (*gun*/WT), respectively (Figure 3a). However, when we analysed differential gene expression in response to NF, we observed a remarkable increase in the number of DEGs (Figure 3b). We identified 1557 DEGs in the WT in response to NF (WT NF/WT). For both treated mutants compared to their respective untreated controls, we identified slightly lower numbers of DEGs. In total, 1361 DEGs were identified in the treated *gun1* mutant (*gun1* NF/*gun1*) and 1177 DEGs were detected in the treated *gun5* mutant (*gun5* NF/*gun5*). In addition, we compared mRNA expression between the NF-treated *gun* mutants and the NF-treated WT. We identified 905 DEGs in
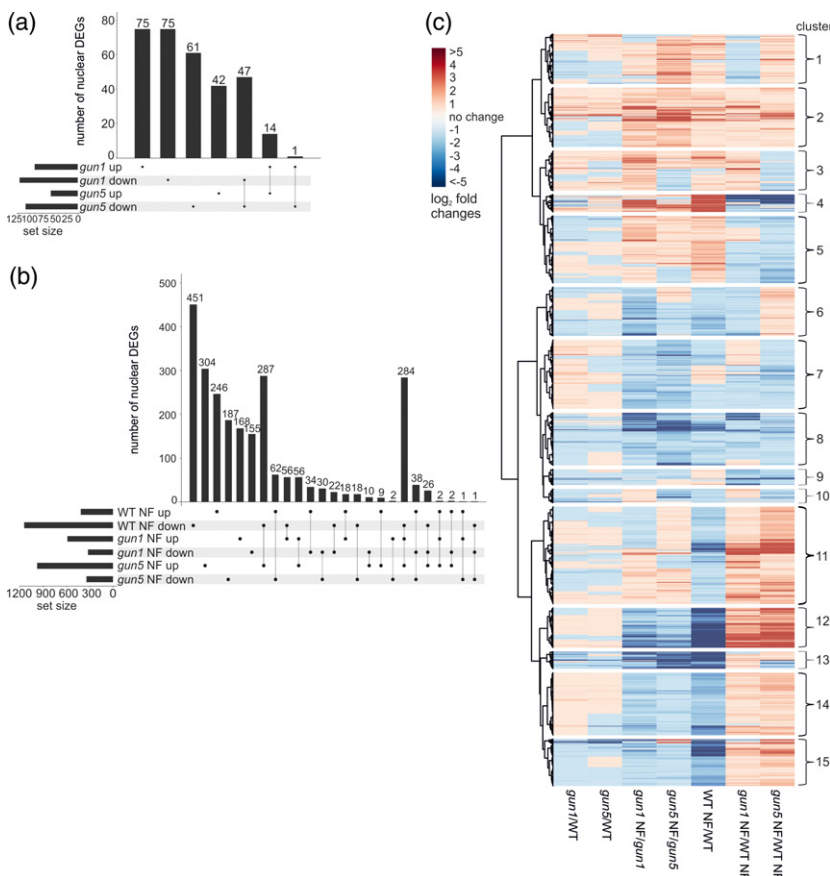

Figure 3. Distribution of nuclear DEGs in the untreated and NF-treated samples. (a) UpSet plot showing the distribution of differentially regulated mRNAs in the untreated *gun* mutants compared to the WT. (b) UpSet plot depicting the distribution of differentially regulated mRNAs in response to NF in WT (WT NF/WT) and both *gun* mutants (*gun1* NF/WT NF and *gun5* NF/WT NF). (c) Hierarchically clustered (UPGMA) heatmap of normalised log$_2$(FC) values from nuclear-encoded DEGs with 15 clusters based on co-expression patterns.

the NF-treated *gun1* mutant (*gun1* NF/WT NF) and 1319 DEGs in the treated *gun5* mutant (*gun5* NF/WT). We generated a hierarchically clustered heatmap from all 3352 nuclear-encoded mRNAs that were differentially regulated in at least one sample (Figure 3c). Based on the co-expression of DEGs we were able to separate 15 specific clusters of differentially regulated nuclear-encoded genes (Table S7). We identified 1557 DEGs in the treated WT (WT NF/WT) and 75% of the mRNAs were downregulated. As expected, the NF-treated *gun* mutants behaved in an opposite manner, as the majority of the RNAs were upregulated, with 65% and 75% upregulated DEGs in the treated *gun1* and *gun 5* mutants (*gun* NF/WT NF), respectively.

To identify the most interesting candidates regulated by retrograde signals, we analysed the overlap between the treated WT (WT NF/WT) and both treated *gun* mutants (*gun* NF/WT NF) to detect those genes that display a typical *gun*-related expression in both mutants (Figure 3b). We identified 284 DEGs in response to NF in WT (WT NF/WT) as well as in both *gun* mutants (*gun* NF/WT NF). These DEGs seem to be controlled by retrograde signalling pathways, because they are repressed by NF in the WT and de-repressed in the *gun* mutants. Furthermore, we detected 56 DEGs with a specific de-repression in the treated *gun1* mutant (*gun1* NF/WT NF) and another 287 DEGs

specifically de-repressed in the *gun5* mutant (*gun5* NF/WT NF). Most likely, the regulation of the genes requires specific retrograde signals, as we identified genes showing a specific de-repression restricted to only one of the *gun* mutants.

Besides the analysis of nuclear-encoded genes, we investigated organellar gene expression and studied the expression of genes encoded by the plastidic and mitochondrial genomes in the WT and both *gun* mutants in the absence or presence of NF. We generated two hierarchically clustered heatmaps for plastidic (Figure 4a) and mitochondrial (Figure 4b) genes that were differentially expressed in at least one of the samples. As expected, we only detected eight mitochondrial genes with differential expression in at least one of the samples, as NF treatment affects carotenoid biosynthesis in the plastids and should not directly affect mitochondrial gene expression. Furthermore, the low number of affected genes in the mitochondria indicates an insignificant crosstalk between plastids and mitochondria triggered by plastid-derived retrograde signals.

In contrast, we detected a considerable high number of differentially regulated plastid-encoded genes. Upon growth in the absence of NF, none of the plastid-encoded genes were differentially expressed in the *gun5* mutant
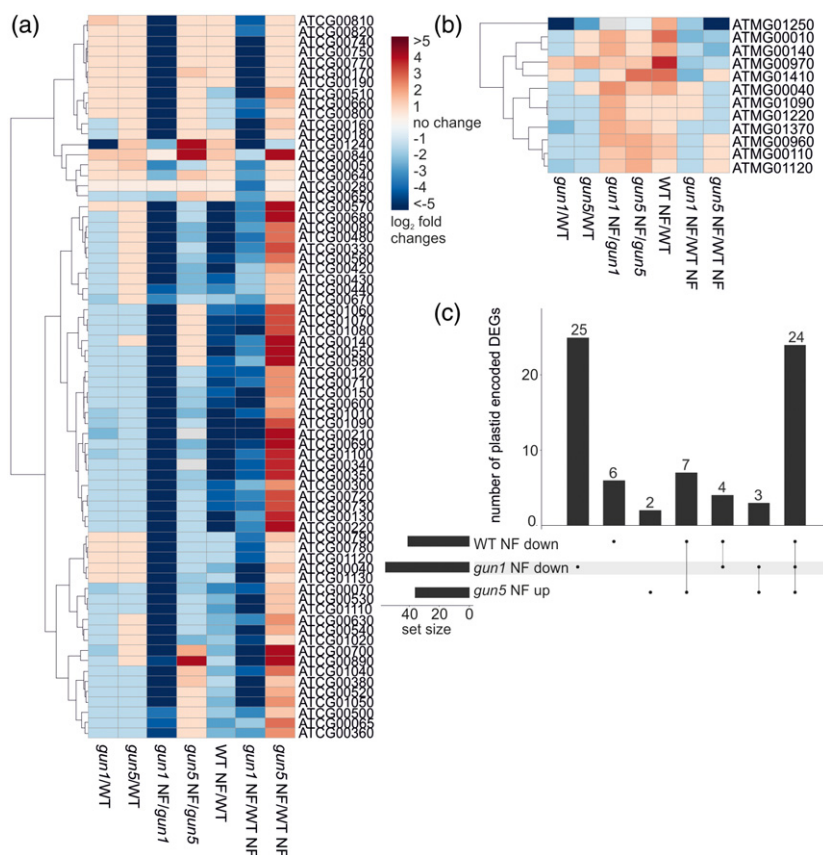

Figure 4. Distribution of differentially expressed plastidic and mitochondrial DEGs in the untreated and NF-treated samples. Hierarchically clustered (UPGMA) heatmap depicting (a) plastidic and (b) mitochondrial genes that are differentially expressed in at least one of the samples displaying normalised $\log_2$(FC) values. (c) UpSet plot depicting the expression of plastid-encoded DEGs detected in the NF-treated *gun* mutants (*gun1* NF/WT NF and *gun5* NF/WT NF) and in the NF-treated WT (WT NF/WT).

and only one plastid-encoded gene was differentially expressed in the *gun1* mutant (*gun*/WT). However, after NF treatment we detected 41, 56 and 36 differentially expressed plastid-encoded genes in the treated WT (WT NF/WT) and *gun1* and *gun5* mutants (*gun* NF/WT NF), respectively. Furthermore, we noticed a highly interesting phenomenon: Almost all plastid-encoded differentially expressed mRNAs were downregulated in the NF-treated *gun1* mutant (*gun1* NF/WT NF) and upregulated in the NF-treated *gun5* mutant (*gun5* NF/WT NF). Thus, based on the plastidic gene expression, both mutants respond in an almost completely opposed manner to NF treatment, suggesting specific perturbations in the NF-triggered organellar signalling pathways. We observed 27 differentially expressed plastid-encoded mRNAs in response to NF in both the *gun1* and the *gun5* mutant (Figure 4c) compared to the NF-treated WT, but they were regulated in an opposing manner: They were all downregulated in the treated *gun1* mutant, but upregulated in the treated *gun5* mutant.

miRNA target analysis

We performed miRNA target prediction with 'psRNATarget' using all protein-coding and non-coding transcripts from Araport11 to correlate the expression of miRNAs with putative target RNA transcripts (Table S9). For each predicted miRNA target, we considered its expression changes to subclassify the miRNA–RNA pairs. For the differentially regulated miRNAs, which were detected in WT NF/WT, *gun1* NF/WT NF and *gun5* NF/WT NF, we were able to predict 218 protein-coding targets as well as 16 non-coding target RNAs, and some of these can be targeted by several miRNAs. We generated a non-redundant list of miRNA targets and excluded transcripts with low fragments per kilobase of transcript per million reads (FPKM) values (≥ 5). Applying these parameters, we obtained 119 predicted miRNA targets that were categorised into three different classes based on their expression. It has to be noted that a specific miRNA–RNA pair can be grouped into different categories since the miRNA as well as the cognate RNA target can be differently regulated between the analysed samples. The first category comprises miRNA–RNA pairs that are 'unchanged' according to the FC of the RNA transcript (but not the miRNA) and includes 101 miRNA–RNA pairs. The second category contains seven miRNA–RNA pairs that show an anticorrelated expression pattern, and the third category encompasses 16 miRNA–RNA pairs
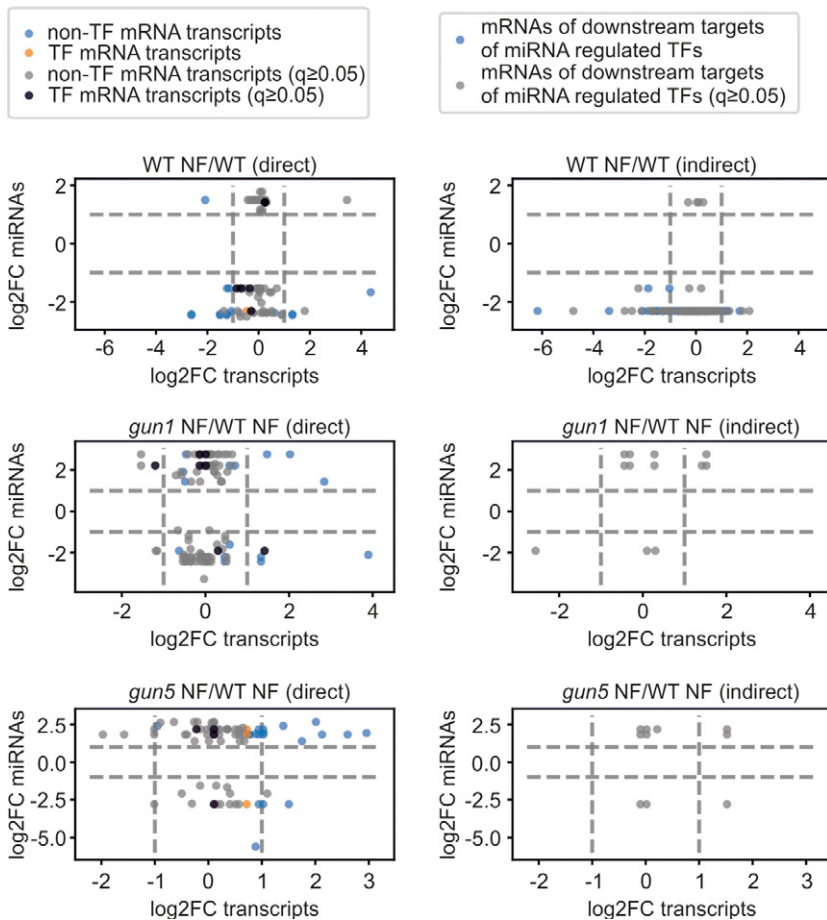

Figure 5. Scatter plots of differentially expressed miRNAs and their targets. Only miRNAs with FDR ≤ 0.05 were included. The direct plots (left panel) depict all differentially expressed miRNAs and their direct predicted target transcripts. MiRNA target transcripts encoding transcription factors are shown in orange (FDR ≤ 0.05) and black (FDR ≥ 0.05). MiRNA target transcripts encoding other proteins are shown in blue (FDR ≤ 0.05) and grey (FDR ≥ 0.05). The indirect plots (right panel) depict downstream targets of transcription factors that are miRNA-regulated. Here, the mRNAs of these downstream genes are plotted against the miRNAs controlling their respective transcription factor mRNAs. The blue dots correspond to FDR ≤ 0.05 and the grey dots to FDR ≥ 0.05.

where the miRNA and the predicted target show the same direction of their differential expression (both up- or both downregulated). Two different miRNAs together with at least two of their targets were validated by qRT-PCR, confirming their anticorrelated expression pattern (Figure S7a, b). Scatter plots (Figure 5) were created to show the distribution of the differentially regulated miRNAs and their correlating targets and were divided into 'direct' and 'indirect' scatter plots. The direct plots show the correlation of miRNAs and their cognate RNA targets either coding for transcription factors or coding for other proteins. The indirect scatter plots depict the expression of downstream genes that are controlled by miRNA-regulated transcription factors. From the direct scatter plot, it is obvious that most differentially expressed miRNA targets do not encode transcription factors. Nevertheless, we identified transcription factor transcripts which are controlled by miRNAs, and their effect on the transcription factor targets can be seen in the indirect plots. For example, the indirect plot shows many differentially expressed transcripts coding for transcription factors, which are controlled by miRNAs in the NF-treated WT (WT NF/WT).

We identified one miRNA–RNA target pair (Table S9) that has been shown to play a role in the acclimation to phosphate deficiency. MiR399a was downregulated in the treated *gun1* mutant (*gun1* NF/WT NF), whereas the expression of its target *PHO2* (AT2G33770), encoding a ubiquitin-conjugating E2 enzyme, remained unchanged in the treated *gun1* mutant (*gun1* NF/WT NF). MiR850 and its cognate target, encoding a chloroplast RNA-binding protein (AT1G09340), belong to the category of miRNA–target pairs showing the same expression (Table S9) since both were upregulated in the treated *gun5* mutant (*gun5* NF/WT NF). This chloroplast RNA-binding protein is necessary for the proper function of the chloroplast and mutations in this gene cause growth deficiency (Fettke *et al.*, 2011). Furthermore, we also identified miR157a-5p (FC of −7 in *gun5* NF/WT NF), displaying an anticorrelated expression to its target PHOTOSYSTEM II REACTION CENTRE PSB28 PROTEIN (AT4G28660), which is 2.9-fold upregulated (*gun5* NF/WT NF). PSB28 is highly conserved in photosynthetic eukaryotes and lack of *PSB28* results in a pale-green phenotype in rice, pointing to a role in the assembly of chlorophyll-containing proteins such as CP47 (Lu, 2016).

Nat-siRNA target analysis

We detected a larger number of differentially expressed sRNAs arising from predicted NAT pairs than from any other sRNA class in the treated WT (WT NF/WT) as well as in both NF-treated *gun* mutants (*gun* NF/WT NF). Filtering the differentially expressed nat-siRNAs with at least five normalised reads in one of six samples (WT, WT NF, *gun1*, *gun1* NF, *gun5* and *gun5* NF) led to a total number of 73 non-redundant *cis*-NATs and 193 non-redundant *trans*-NAT

pairs. These pairs were further analysed and we only selected the nat-siRNA producing transcript pairs with at least five normalised reads for one of the two overlapping transcripts. This reduced the number to 64 non-redundant *cis*-NAT and 40 non-redundant *trans*-NAT pairs (Table S10). The expression changes of two nat-siRNAs together with their overlapping transcripts in NF-treated WT and the NF-treated *gun5* mutant were confirmed by qRT-PCR (Figure S7c,d).

For many *trans*-NAT pairs, we observed that one of the transcripts was derived from a transposable element or a pre-tRNA, whereas the second overlapping transcript represented a protein-coding gene. Among these *trans*-NAT pairs, we only detected one overlapping transcript encoding a plastid-localised protein, suggesting a low impact of *trans*-NAT pairs in the adjustment of plastid and nuclear gene expression in response to NF. The *trans*-nat-siRNA generated from this pair was found to be downregulated in the treated WT (WT NF/WT) and upregulated in both treated *gun* mutants (*gun* NF/WT NF). The first overlapping transcript codes for the plastid-localised UDP-glucosyl transferase 75B2 (AT1G05530), which is able to bind UDP-glucose, important for cellulose and callose synthesis (Hong *et al.*, 2001). Its expression was unchanged in the treated WT (WT NF/WT) as well as in both treated *gun* mutants (*gun* NF/WT NF). The second overlapping transcript represented a lncRNA (AT1G05562) that was downregulated in the treated WT (WT NF/WT) and upregulated in both treated *gun* mutants (*gun* NF/WT NF).

Interestingly, out of 64 *cis*-NAT pairs that give rise to differentially regulated nat-siRNAs, we detected 31 individual transcripts which encode plastid proteins, indicating a considerable role of *cis*-NAT pairs in the direct control of genes coding for plastid proteins via NF-triggered retrograde signals. Moreover, within the *cis*-NAT pairs we identified 35 individual transcripts encoding nuclear-localised proteins, pointing to a large impact of these in the indirect adjustment of nuclear gene expression via nuclear regulatory proteins. One sRNA processed from a *cis*-NAT pair was detected to be downregulated in the treated WT (WT NF/WT). Interestingly, both overlapping transcripts were identified to encode plastid-localised proteins. The expression of the first overlapping transcript (AT1G29900), which codes for a subunit of carbamoyl phosphate synthetase, which is presumed to be necessary for the conversion of ornithine to citrulline in the arginine biosynthesis pathway (Molla-Morales *et al.*, 2011), was unchanged (WT NF/WT). In agreement with the expression of the nat-siRNAs, the second overlapping transcript (AT1G29910) was downregulated by NF in the WT (WT NF/WT). This transcript encodes a chlorophyll A/B-binding protein, which is the major protein of the light-harvesting complex and is required for absorbing light during photosynthesis.

Network analysis

In order to gain a comprehensive picture of the role of miRNAs in retrograde signalling and to analyse possible downstream effects, we investigated a miRNA–RNA-target network that also comprises related transcription factor to target gene connections. The results were combined in a complex interaction network (Data S1), since one miRNA can control many mRNAs encoding transcription factors, which in turn control several downstream genes, but also one miRNA target can be controlled by numerous miRNAs (Figure S8 and Table S12). Within the considered network, most miRNAs regulate just a small number of target transcripts (Figure S8a), but there are some miRNAs regulating up to 140 targets. In contrast, the majority of miRNA targets are regulated by only a few miRNAs, but there are still some targets that can be regulated by up to 15 miRNAs (Figure S8b). We observed that miRNAs controlling the highest number of targets mainly regulate mRNAs that do not encode transcription factors (Figure S8c), while the distribution of miRNA targets encoding transcription factors indicates that most miRNAs regulate only a small number of such targets, with the highest number being eight (Figure S8d). Some motifs are recurrent in the miRNA–RNA target network (Figure 6). We explored the network for different characteristic relations of regulatory linkage and behaviour. Here, we found simple expected patterns where a miRNA, miR157a-5p, was downregulated and its target

mRNA transcript encoding a plastid-localised protein (AT4G28660) in turn was upregulated, or vice versa (Figure 6a), but we also observed many miRNA targets that did not show any differential expression on the mRNA level, although corresponding miRNAs were differentially expressed. The effect of these miRNAs might be visible on the protein level due to inhibition of translation. If the target mRNA encodes a transcription factor, we should see the miRNA-dependent regulation in the expression of downstream targets of this transcription factor (Figure 6b), as reported before (Megraw *et al.*, 2016). As an example, the transcription factor AT4G36920 can act as an activator or repressor on its targets, and furthermore, the transcription factor can control other transcription factors or downstream targets like AT2G33380, which is a *CALEOSIN 3* transcript and important for stress responses (Sham *et al.*, 2015). In addition, the downstream transcription factors can also target other transcription factors or downstream targets, which increases the network complexity. Furthermore, this points to a sophisticated interaction between miRNAs and their targets, because miRNAs indirectly regulate genes encoding plastid proteins through the direct control of transcription factor mRNAs. Besides transcription factor mRNAs, many miRNAs are able to regulate transcripts of genes that do not encode transcription factors, but also these transcripts do not always show the expected behaviour. For instance, miR395c is predicted to control
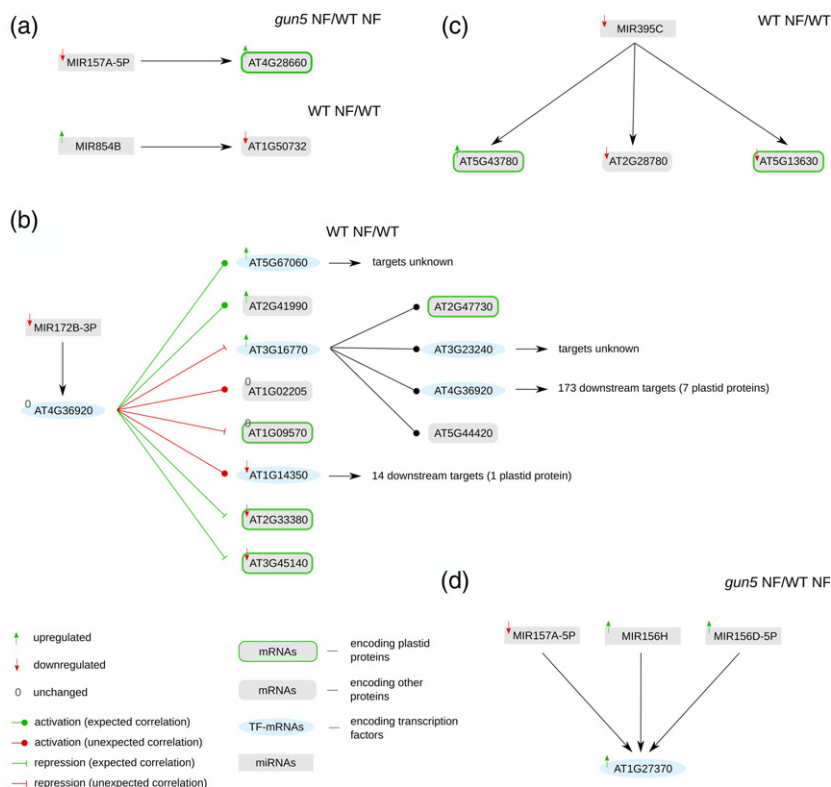

Figure 6. Illustration of different network motifs which we observed in the miRNA–RNA target network in connection with relative changes of RNA levels between treatments. (a) Examples of expected regulations where a miRNA and its target mRNA exhibit inversed differential expression. (b) The interaction between a miRNA regulating the mRNA of a transcription factor, and the interactions of this transcription factor with its downstream target genes. (c) A downregulated miRNA which regulates four different targets. (d) An example of three miRNAs that regulate a single mRNA. The whole network can be accessed through the supporting Data S1 in GML format.

four different mRNA targets, including three transcripts encoding plastid-localised proteins (Figure 6c). Further, we found examples where several miRNAs are able to control the transcript of a single transcription factor (Figure 6d). These cases show possible interactions between miRNAs and their targets and suggest a wide range of direct and indirect impacts of miRNAs to regulate gene expression. Nevertheless, behavioural predictions are impossible without additional information on the exact mode of action of each miRNA and the magnitude of its influence.

DISCUSSION

Until now, it is not known whether ncRNAs and sRNAs are regulated by retrograde signalling in response to NF treatment and how they contribute to the control of nuclear gene expression in response to plastid-derived signals. To better understand these biological processes, we combined sRNA sequencing with mRNA/lncRNA sequencing of *A. thaliana* WT seedlings and the two retrograde signalling mutants, *gun1* and *gun5*, to identify ncRNAs and mRNAs regulated by retrograde signals.

Generally, after NF treatment we detected nearly the same number of DEGs in all treated samples compared to the untreated WT. Further, we observed an overall tendency that more DEGs were downregulated than upregulated in response to NF treatment. In addition, we could observe an overrepresentation of DEGs encoding plastid-localised proteins in all three samples and detected more DEGs to be upregulated in the treated *gun*mutants compared to the treated WT.

Previous studies with different *gun* mutants were performed using *A. thaliana* microarrays lacking probes for ncRNAs (Strand *et al.*, 2003; Koussevitzky *et al.*, 2007; Woodson *et al.*, 2013). Koussevitzky *et al.* (2007) analysed changes in mRNA levels in WT (Col-0), *gun1* and *gun5* mutant seedlings grown on media with and without NF. About 43% of upregulated and 67% of downregulated DEGs in the present study overlap with those of Koussevitzky *et al.* (2007) in response to NF (Figure S9a,b). Generally, more downregulated DEGs and larger changes in the NF-treated *gun5* mutant than in the other mutant were identified in both studies. A good overlap of DEGs was found in both *gun* mutants. About 56% of the DEGs detected in *gun1-102* (*gun1* NF/WT NF) in our study were also detected in the treated *gun1-9* mutant by Koussevitzky *et al.* (2007) (Figure S9c), and about 50% of the DEGs identified in the treated *gun5* mutant (*gun5* NF/WT NF) in our study were also identified by Koussevitzky *et al.* (2007) (Figure S9d). However, in our data set we identified also 44% (*gun1* NF/WT NF) and 50% (*gun5* NF/WT NF) of DEGs that have not been shown to be controlled by *gun*-related retrograde signalling pathways before, which might be due to the differences between the two methods (RNA

sequencing versus microarrays) and the varying duration of the NF treatment between the studies (5 versus 4 days). Recently, another RNA sequencing study reported NF-responsive transcriptome changes in a different *gun1* mutant (*gun1-1*) (Richter *et al.*, 2020); 55% and 49% of the DEGs detected in the NF-treated *gun1* and *gun5* mutant compared to the NF-treated WT overlapped with DEGs in our study (Figure S9e,f). RNA sequencing was also performed in the *gun1-9* mutant grown in the presence of NF (Zhao *et al.*, 2019b) with an overlap of 65% of DEGs compared to our data (Figure S9g). Taken together, we observed a considerably high overlap with other transcriptome studies despite the differences between the studies regarding growth conditions, available mutants and analysis methods.

In our study, we observed an opposite regulation of differentially expressed plastid-encoded transcripts in both *gun* mutants, while the nuclear-encoded DEGs showed large overlap between the *gun1* and *gun5* mutants. Surprisingly, in response to NF all differentially expressed plastid-encoded transcripts were downregulated in the *gun1* mutant, whereas they were upregulated in the treated *gun5* mutant. These observations are in line with the model suggesting that plastid gene transcription is controlled by retrograde signalling networks, including sigma factors (SIG2 and SIG6) and plastid-encoded RNA polymerase (PEP), which might be crucial for proper plastid RNA transcription (Woodson *et al.*, 2013). It seems that *GUN1* activates PEP (Maruta *et al.*, 2015) and a perturbed PEP activation in the *gun1* mutant may prevent the upregulation of the plastid-encoded genes compared to WT upon NF treatment.

We identified an interesting lncRNA (AT4G13495) showing classical de-repression in both *gun* mutants (*gun* NF/WT NF) (Table 2). This lncRNA overlaps in sense direction with three different miRNA precursors (*MIR5026*, *MIR850* and *MIR863*) and all three miRNAs were differentially expressed in at least one treatment (WT NF/WT, *gun1* NF/WT NF and *gun5* NF/WT NF). We assume that all three miRNAs can be produced either from the three individual miRNA precursor transcripts or from the lncRNA. We did not find any predicted target for miR5026 according to the applied psRNATarget parameters. MiR850 was upregulated in the *gun5* mutant (*gun5* NF/WT NF), and two predicted cognate target RNAs, encoding a chloroplast RNA-binding protein (AT1G09340) and a threonine-tRNA ligase (AT2G04842), respectively, were upregulated as well. MiR863 targets the *SERRATE* transcript (AT2G27100), encoding an accessory protein essential for the miRNA biogenesis pathway, and thus may influence the regulation of several miRNAs (Meng *et al.*, 2012). MiR863 was upregulated in both treated *gun* mutants (*gun* NF/WT NF), but we did not detect significant changes of the *SERRATE* transcript in the two treated *gun* mutants.

Concerning the overall regulation of differentially expressed sRNAs belonging to different classes (miRNAs, nat-siRNAs and other sRNA producing loci), we detected more downregulated sRNAs in the treated WT (WT NF/WT), whereas both treated *gun* mutants exhibited a higher number of upregulated sRNAs (*gun* NF/WT NF). Principally, this suggests an increased sRNA processing in response to NF in both *gun* mutants, resembling the de-repression of nuclear-encoded genes, and we assume that these sRNAs might have an impact on retrograde-controlled nuclear gene expression. sRNAs are able to affect nuclear transcripts regulated by retrograde signals and they may regulate mRNA transcripts, affecting plastid-localised proteins. Among all sRNA classes, we observed in all treatments the highest number of differentially regulated sRNAs within the nat-siRNA class. Furthermore, all differentially regulated sRNAs have been associated to their corresponding putative differentially expressed RNA targets (Table S13) and we could detect high numbers of differentially regulated sRNAs. Besides an effect of sRNAs on their direct targets, we expect based on our network analyses a considerable indirect regulation by sRNAs through transcription factors (Data S1).

Interestingly, we found that miR169g-3p, a heat- and salt stress-responsive miRNA (Szyrajew *et al.*, 2017; Pegler *et al.*, 2019), is the most strongly downregulated miRNA in the treated WT (−151.6-fold; WT NF/WT), and the most strongly upregulated one in the treated *gun5* mutant (38.2-fold; *gun5* NF/WT NF). We did not find any predicted target for miR169g-3p according to our parameters.

Unexpectedly, after miRNA target prediction the expression of most of the targets was not anticorrelated to the expression changes of their cognate miRNA, leading us to conclude that miRNAs might not be involved in the expression of genes controlled by retrograde signalling pathways, or the expressional changes of miRNAs somehow balance transcriptional changes of their targets to maintain constant steady-state levels. Another possibility could be that they act as translational repressors and do not have a direct effect on the transcript abundance of their target RNAs. However, we predicted 20 miRNA targets coding for transcription factors and 22 targets encoding plastid-localised proteins to be targeted by 23 differentially regulated miRNAs. Thus, we assume that miRNAs may have important functions in the control of transcripts that code for regulatory proteins that are directly involved in transcriptional control and may contribute to the manifold changes of gene expression in response to retrograde signals. Further, nuclear transcripts that code for plastid-localised proteins are targets of miRNAs, suggesting that these specific miRNA–mRNA pairs can play an important role in the retrograde signalling pathway, and thus may contribute to the adjustment of plastidic and nuclear gene expression. One interesting case involves miR395b and miR395c, which target the mRNA for the magnesium-chelatase subunit *GUN5* (AT5G13630). In the NF-treated WT, both miRNAs and the target mRNA are downregulated compared to the untreated control, whereas in the treated *gun5* mutant both miRNAs and the target are upregulated. Even though the expression of this miRNA–mRNA pair is not anticorrelated, the enhanced miRNA levels may balance an increased transcription rate of the target mRNA to keep physiologically relevant steady-state levels. Magnesium-chelatase is required in the chlorophyll biosynthesis pathway, where it catalyses the insertion of Mg^{2+} into protoporphyrin IX, and the *gun5* mutant is characterised by a single nucleotide substitution resulting in a defective magnesium-chelatase. In the WT, the *GUN5* transcript level decreases in response to NF-triggered retrograde signalling, whereas the transcript level in the *gun5* mutant remains high and cannot be efficiently downregulated by the increased miRNA levels. The seven detected classical anticorrelated miRNA–mRNA pairs point to regulatory functions of specific miRNAs in the retrograde signalling pathway, because we assume efficient miRNA-mediated target cleavage followed by a reduced mRNA steady-state level. In this category of anticorrelated pairs, we identified the mRNA for the transcription factor SPL10, representing a validated target of miR157a, suggesting miR157 acts in retrograde signalling by affecting the levels of a transcriptional regulator and its downstream targets. Another anticorrelated predicted miRNA–mRNA pair is miR398, targeting the transcript of the multidrug and toxic compound extrusion (MATE) efflux protein (AT2G04050). We found miR398 to be downregulated in the treated WT compared to the untreated control, and the target was slightly upregulated. This MATE efflux protein belongs to a huge class of membrane proteins located in the plasma membrane and the chloroplast envelope membrane (Wang *et al.*, 2016) that are able to bind cytotoxic compounds like primary and secondary metabolites, xenobiotic organic cations (Omote *et al.*, 2006) and toxic substances such as pollutants and herbicides (Diener *et al.*, 2001) to eliminate them from the cell (Liu *et al.*, 2016). NF-triggered downregulation of miR398 most likely causes elevated transcript levels and increased levels of the encoded plasma membrane-located MATE efflux protein. We speculate that the regulated MATE efflux protein might be involved in the extrusion of the applied herbicide NF or the extrusion of toxic compounds accumulating within the cell in response to NF treatment. Another interesting target encodes a plastid protein that appeared to be upregulated in the NF-treated *gun5* mutant (*gun5* NF/WT NF) by decreased levels of the cognate miRNAs. *PSB28* (AT4G28660), targeted by miR157a, encodes a protein that is part of the photosystem II reaction centre and is suggested to function in the biogenesis and assembly of chlorophyll-containing proteins (Mabbitt *et al.*, 2014). NF treatment usually leads to the

downregulation of *PhANGs* and thus should cause decreased expression levels of the *PSB28* mRNA. However, miR157a contributes to the downregulation of *PSB28* mRNA levels post-transcriptionally and seems to be controlled by retrograde signals, as indicated by the misregulation of miR157a in the *gun5* mutant.

Fang *et al.* (2018) identified miR395 and miR398 to be important in retrograde signalling triggered by tocopherols, and we confirmed both miRNAs applying NF as another trigger of retrograde signalling. Additionally, the transcript encoding the enzyme APS, which catalyses the initial step in PAP synthesis (Klein and Papenbrock, 2004; Pornsiriwong *et al.*, 2017) was identified to be targeted by miR395 (Liang *et al.*, 2010). We found miR395b to be downregulated in the treated WT (WT NF/WT) and upregulated in the treated *gun5* mutant (*gun5* NF/WT NF). In the treated WT, reduced miR395b levels lead to increased APS transcript levels, causing elevated PAP synthesis, which acts as retrograde inhibitor of XRNs and should provoke elevated pri-miRNA and mature miRNA levels. Besides, Fang *et al.* (2018) detected the downregulation of miR398 in the WT after NF treatment, which is in line with our sRNA sequencing data. The COPPER/ZINC SUPEROXIDE DISMUTASE 2 (CSD2) was previously found to be a target of miR398 (Guan *et al.*, 2013). After heat stress, Fang *et al.* (2018) found miR398, PAP and tocopherol levels to be increased and CSD2 levels to be decreased in the WT, and they hypothesised that tocopherols and PAP are required for miR398 biogenesis under heat stress. The *CSD2* mRNA escaped our miRNA target prediction due to a considerably high number of mismatches within the miRNA binding site, causing a score value that was above our cut-off value. Still, we identified this miRNA as differentially expressed supporting the previous study by Fang *et al.* (2018).

Besides differentially expressed miRNAs, we identified an even higher number of differentially regulated nat-siR-NAs in the treated WT (WT NF/WT) and both *gun* mutants (*gun* NF/WT NF). Most of the overlapping transcripts encode nuclear or plastid proteins, suggesting that nat-siR-NAs have a considerable impact on the control of *PhANGs* encoding plastid proteins. For most of the *cis*-NAT pairs we observed similar correlations between RNA transcript and sRNA expression levels. For example, the levels of two overlapping transcripts (AT1G05560 and AT1G05562) and the related *cis*-nat-siRNA were decreased in the treated WT (WT NF/WT) and increased in both *gun* mutants (*gun* NF/WT NF). The gene AT1G05562 encodes an antisense lncRNA and overlaps with the gene AT1G05560, which codes for a UDP-glucose transferase. Another interesting *cis*-nat-siRNA and one of the overlapping transcripts encoding a chlorophyll binding protein (AT1G29930) were downregulated in the treated WT (WT NF/WT) and upregulated in the treated *gun5* mutant (*gun5* NF/WT NF),

whereas levels of the other overlapping transcript, coding for a nuclear RNA polymerase (AT1G29940), remained unchanged in both treatments.

Here, we could demonstrate that NF treatment and subsequent retrograde signals lead to comprehensive changes in the steady-state levels of non-coding sRNAs comprising all known sRNA classes. The majority of the identified differentially expressed sRNAs belong to the *cis*- and *trans*-nat-siRNAs, followed by miRNAs, representing the second most abundant class. Thus, we postulate that mainly these two sRNA classes act as important regulators of gene expression in retrograde signalling. We also identified a considerably high number of so far unknown nuclear-encoded DEGs and thus add to the knowledge about genes that are controlled by retrograde signalling. Finally, we were able to identify promising sRNA–RNA target pairs that may act in the adjustment of plastidic and nuclear gene expression in retrograde signalling pathways.

EXPERIMENTAL PROCEDURES

Plant material and growth conditions

Arabidopsis thaliana WT (*Col-0*) and the retrograde signalling mutants *gun1-102* and *gun5-1* were used in this study. *Gun1-102* (SAIL_290_D09) harbours a transfer DNA insertion within the AT2G31400 gene locus resulting in a loss-of-function allele (Tadini *et al.*, 2016). *Gun5-1* is an EMS mutant harbouring a point mutation within the gene AT5G13630 causing an Ala/Val substitution at residue 990 (A990V) resulting in deficient magnesium-protoporphyrin IX synthesis (Mochizuki *et al.*, 2001). Surface-sterilised seeds were incubated on ½ MS agar plates containing 1.5% sucrose. For treatments with NF, seeds were incubated on the same medium supplemented with 5 μM norflurazon (Sigma-Aldrich, Taufkirchen, Germany). After vernalisation (2 days at 4°C in darkness) the seeds were grown for 4 days under continuous light (115 μmol photons m^{-2} sec^{-1}) at 22°C. Whole plants were harvested and immediately frozen in liquid nitrogen and stored at −80°C until RNA isolation. All control experiments and norflurazon treatments were performed in three biological replicates for each genotype.

RNA isolation

The plant material was ground in liquid nitrogen and RNA isolation was performed using TRIzol reagent (Invitrogen) according to the manufacturer's protocol. RNA integrity was monitored by agarose gel electrophoresis and RNA concentration and purity were determined spectrophotometrically (260 nm/280 nm and 260 nm/230 nm absorbance ratios).

sRNA purification

For sRNA sequencing 30 μg of total RNA were separated by 15% PAGE for 2 h at 120 V. The sRNA fractions with sizes ranging from 18 to 29 nucleotides were excised from the gel and eluted in 0.3 M NaCl overnight at 4°C with rotation. Remaining gel pieces were removed using a Spin-X centrifuge tube (Sigma-Aldrich) and 1 μl GlycoBlue (15 mg ml^{-1}, Thermo Fisher), 25 μl sodium acetate (3 M, pH 5.0) and 625 μl ethanol were added to the 250 μl flow-through and samples were incubated for 4 h at −80°C. After centrifugation for 30 min with 17 000 ***g*** at 4°C the RNA pellet was

washed twice with 80% ethanol, dried and dissolved in nuclease-free water.

Quantitative RT-PCR

Prior to cDNA synthesis, RNA samples were subjected to DNase I digestion (NEB, Ipswich, MA, USA) to remove residual genomic DNA. Total RNA (2 μg) was incubated at 37°C for 30 min together with DNaseI (2 U; NEB). To inactivate the DNaseI, 2.5 μl 50 mM EDTA was added and samples were incubated at 65°C for 10 min. The RNA was denatured for 5 min at 65°C in the presence of 100 pmol of an oligo-dT23VN oligonucleotide and 10 mM dNTPs and transferred to ice. Subsequently, cDNA synthesis was performed for 1 h at 42°C using M-MuLV reverse transcriptase (200 U; NEB) followed by heat inactivation at 80°C for 5 min. To monitor successful cDNA synthesis, we performed RT-PCR using gene-specific primers for the gene *UBI7* (AT3G53090) (Table S11).

For each qRT-PCR we used cDNA equivalent to 20 ng μl^{-1} RNA with gene-specific primers and an EvaGreen qPCR mix. The samples were pre-heated for 2 min at 95°C and qRT-PCR cycling conditions were as follows: 12 sec at 95°C, 30 sec at 58°C and 15 sec at 72°C for 40 cycles. All qRT-PCRs were performed in three technical triplicates with the CFX Connect Real-Time PCR device (Bio-Rad, Feldkirchen, Germany). The Ct-values were used to calculate changes in gene expression by the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). The values were normalised to the housekeeping gene *UBI1* (AT4G36800). Oligonucleotide sequences of all gene-specific primers are listed in Table S11.

Stem-loop qRT-PCR

Stem-loop qRT-PCRs were used for sRNA quantification as described previously (Kramer, 2011). RNA from three independent biological replicates (300 ng) was used for cDNA synthesis. RNA was denatured for 5 min at 65°C together with 100 pmol of stem-loop oligonucleotides (Table S11) and 10 mM dNTPs. cDNA synthesis was performed for 5 min at 25°C and 20 min at 42°C using M-MuLV reverse transcriptase (200 U; NEB), followed by heat inactivation at 80°C for 5 min. RT-PCR for the gene *UBI7* (AT3G53090) served as control (Table S11).

RNA sequencing

For the generation of mRNA libraries, including poly(A)-tailed lncRNAs, 10 μg total RNA from each sample was vacuum-dried in the presence of RNAstable (Sigma-Aldrich). The libraries were prepared using the Next Ultra RNA Library Prep Kit (NEB) by Novogene (China). The samples were sequenced strand-specifically as 150 bp paired-end reads on a HiSeq-PE150 platform with at least 15 million read pairs per library.

sRNA libraries for each RNA sample were generated twice following two slightly modified protocols. The first set of libraries was generated from 5 μg total RNA with the NEBNext Multiplex Small RNA Library Prep Set for Illumina according to the manufacturer's instructions and 1 h of 3′ adapter ligation. The second set of sRNA libraries was prepared from purified sRNAs obtained from 30 μg of total RNA using the same kit as described above performing 3′ adapter ligation for 18 h. For both libraries, excessive non-ligated 3′ adapters were made inaccessible by converting them into dsRNA by hybridisation of complementary oligonucleotides. 5′ adapter ligation was carried out at 25°C for 1.5 h, reverse transcription was performed by using the ProtoScript II reverse transcriptase and libraries were amplified by 12 PCR cycles. The PCR products were separated by 6% PAGE for 2 h at

60 V. The cDNA library fractions with a size ranging from 138 to 150 nucleotides were excised from the gel and eluted overnight. The sRNA libraries were sequenced as 50 bp single-end reads on an Illumina HiSeq1500 sequencer with approximately 10 million reads per library.

Analysis of mRNA and lncRNA

The mRNA and lncRNA sequencing results were analysed with the open source and web-based platform GALAXY (Afgan et al., 2016). The FASTQ raw sequences were trimmed with the tool Trimmomatic to remove adapter sequences with their default parameters (Bolger et al., 2014). Tophat (Kim et al., 2013) was used to map the reads against the A. thaliana reference genome (https://www.arabidopsis.org/, release: TAIR10) with a maximum intron length parameter of 3000 nt. The transcripts were annotated in Araport11 (https://apps.araport.org/thalemine/dataCategories.do); we considered annotated ncRNAs longer than 200 bp as lncRNAs. Differential expression of transcripts was analysed by Cuffdiff (Trapnell et al., 2010) to normalise the sequencing depth of each library and to calculate FPKM values. The FDR was used as a statistic indicator to exclude type I errors or rather false positives. Transcripts having FDR ≤ 0.05 and log$_2$(FC) ≤ −1 and ≥ +1 were considered as DEGs. The package pheatmap (https://cran.r-project.org/web/packages/pheatmap/pheatmap.pdf) was used to generate hierarchically (UPGMA) clustered heatmaps of differentially expressed RNAs (Kolde, 2019).

Gene ontology terms

GO enrichment terms were analysed using the DAVID Bioinformatics Resources 6.8 (https://david.ncifcrf.gov) with default parameters (Huang da et al., 2009a,b) and results were visualised with the R package 'ggpubr' (Wickham, 2016).

Analysis of sRNA

NEBNext Kit adapter sequences were clipped from the sequencing reads using a custom script within GALAXY that identifies Illumina adapter sequences using a seed sequence of 10 nt. After adapter clipping FASTQ files of the raw reads with a length of 18–26 nt were loaded into the CLC Genomics Workbench 11.0.1 (Qiagen, Hilden, Germany) for further analyses. The ShortStack analysis package was used for advanced analysis of the sRNA sequences (Axtell, 2013b). The FASTQ files of the six biological replicates derived from each treatment were first mapped against the A. thaliana TAIR10 reference genome (https://www.arabidopsis.org/, release: TAIR10). The merged alignments were mapped against a file covering all A. thaliana mature miRNAs (http://www.mirbase.org/) and a second file comprising different RNA classes, namely nat-siRNAs, ta-siRNAs, phasiRNAs and lncRNAs. A nat-siRNA database (Table S1) was generated from previously annotated NAT pairs (Jin et al., 2008; Zhang et al., 2012; Yuan et al., 2015), phasiRNAs were taken from Howell et al. (Howell et al., 2007) and lncRNAs were downloaded from Araport11 (https://apps.araport.org/thalemine/dataCategories.do, release: Araport 11 Annotation). After mapping to the respective references, the individual raw reads for each replicate were used for normalisation and differential expression analysis based on a calculation with DeSeq2 (Love et al., 2014). The sRNAs were filtered by fold changes between ≤ −2 and ≥ +2. The significance of the differentially expressed sRNAs was evaluated with FDR ≤ 0.05. MiRNA target RNAs were identified using the psRNATarget (Dai et al., 2018) prediction V2 tool (http://plantgrn.noble.org/psRNATarget/) from protein-coding and non-coding transcripts present in Araport11. An expectation value of less than 2.5 was considered as a cut-off

for true miRNA targets, where mRNAs harbouring a lower number of mismatches to the reverse and complementary miRNA obtain lower score values.

Network analysis

Using the sRNA and mRNA sequencing data together with the miRNA target prediction, we assembled an interaction network of miRNAs and their putative targets. For network analysis, we used the python package networkX (Hagberg *et al.*, 2008). We further investigated the miRNA targets that were predicted as described above using the psRNATarget tool to identify all miRNA targets that encode transcription factors. For this, we compared all miRNA targets with a reference database, containing *A. thaliana* transcription factors, which was generated using publicly available data (http://atrm.cbi.pku.edu.cn). Furthermore, this reference database was extended by incorporating available information about whether the transcription factors act as activators or repressors of gene expression together with available information about the individual target genes of the transcription factors (https://agris-knowledgebase.org). The data obtained from the RNA sequencing experiments were then used to generate a network of miRNAs and their targets, differentiating between miRNA targets encoding transcription factors and targets encoding other proteins. For network analyses connected to measurements, only RNAs with a FDR less than 0.05 were considered, unless indicated otherwise. Network analyses were performed to determine relationships between miRNAs and their targets with special focus on miRNA targets encoding transcription factors. In these network analyses we considered the impact of miRNAs on the transcripts coding for transcription factors and the impact of the expression of transcription factor mRNAs on downstream genes, regulated by these transcription factors. If the change of a miRNA results in an expected change of an mRNA coding for a transcription factor, or at least downstream genes show expected transcriptional changes, we classified this behaviour as 'expected'. For example, if miRNA expression was reduced and the target mRNA encoding a transcription factor was upregulated, the transcription factor acted as an activator and downstream genes of this transcription factor were consequently also upregulated, this would be considered as 'expected' behaviour. Furthermore, the scatter plots presented for all differentially expressed miRNAs (FDR $\leq$ 0.05) were subdivided into two categories: plots depicting relations of miRNA and their direct target transcripts (direct) and plots depicting indirect relations comprising miRNAs that control mRNAs encoding transcription factors and their downstream target genes (indirect).

ACCESSION NUMBERS

ATG accession numbers: GUN1 (AT2G31400), GUN5 (AT5G13630). The raw Illumina sRNA and mRNA/lncRNA sequencing data are deposited in the NCBI SRA database with the ID PRJNA557616.

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

WF and MAA designed the research; KH performed the research with help from MAA and BT; KH, MAA, BT and WF analysed the data; network analysis was performed by SOA, MK and EK; and KH, MAA, MK, SOA and WF wrote the paper. All authors read and approved the final manuscript.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Appendix S1. Figure and table captions.

Figure S1. Effect of NF on the plants and sRNA size distribution.

Figure S2. Distribution of differentially expressed *cis*- and *trans*-nat-siRNA in response to NF.

Figure S3. Distribution of differentially expressed sRNAs derived from lncRNAs in response to NF.

Figure S4. Validation of transcript levels by qRT-PCR.

Figure S5. Distribution of differentially expressed ncRNAs in the untreated and treated samples.

Figure S6. GO term enrichment analysis.

Figure S7. Validation of sRNAs and their associated transcripts.

Figure S8. Distributions of miRNA–RNA target interaction numbers.

Figure S9. Comparison of differentially expressed mRNAs from Koussevitzky *et al.* (2007), Richter *et al.* (2020), Zhao *et al.* (2019) and our study.

Table S1. Reference sequences for the *cis*- and *trans*-NAT pairs.

Table S2. sRNA sequencing and mapping results for each independent biological replicate.

Table S3. Lists of all significant differentially expressed sRNA classes.

Table S4. RNA mapping results after the analysis with Tophat for each independent biological replicate.

Table S5. Overview of all significant differentially expressed mRNAs and other RNA classes.

Table S6. Accession numbers of all significant differentially expressed mRNAs and other RNA classes.

Table S7. List of the individual clusters of the mRNA heatmap.

Table S8. GO term enrichment analysis for significant DEGs.

Table S9. Lists of miRNA target prediction.

Table S10. Lists of *cis*- and *trans*-nat-siRNAs and their target correlation.

Table S11. Sequences of oligonucleotides used in this study.

Table S12. Distributions of miRNA–RNA target interaction numbers.

Table S13. Overview of all significant differentially expressed (DE) sRNAs and their corresponding differentially expressed mRNA target.

Data S1. MiRNA–RNA target network in GML format that can be accessed with the free software 'Gephi' (Bastian *et al.*, 2009).

REFERENCES

Afgan, E., Baker, D., van den Beek, M. *et al.* (2016) The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res.* **44**, W3–W10.

Axtell, M.J. (2013a) Classification and comparison of small RNAs from plants. *Annu. Rev. Plant Biol.* **64**, 137–159.

Axtell, M.J. (2013b) ShortStack: comprehensive annotation and quantification of small RNA genes. *RNA*, **19**, 740–751.

Ball, L., Accotto, G.P., Bechtold, U. *et al.* (2004) Evidence for a direct link between glutathione biosynthesis and stress defense gene expression in Arabidopsis. *Plant Cell*, **16**, 2448–2462.

Bannister, A.J. and Kouzarides, T. (2011) Regulation of chromatin by histone modifications. *Cell Res.* **21**, 381–395.

Bartel, D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, **116**, 281–297.

Bastian, M., Heymann, S. and Jacomy, M. (2009) Gephi: an open source software for exploring and manipulating networks. *International AAAI Conference on Weblogs and Social Media*.

Bolger, A.M., Lohse, M. and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114–2120.

Borsani, O., Zhu, J., Verslues, P.E., Sunkar, R. and Zhu, J.K. (2005) Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in Arabidopsis. *Cell*, **123**, 1279–1291.

Breitenbach, J., Zhu, C. and Sandmann, G. (2001) Bleaching herbicide norflurazon inhibits phytoene desaturase by competition with the cofactors. *J. Agric. Food Chem.* **49**, 5270–5272.

Chan, K.X., Phua, S.Y., Crisp, P., McQuinn, R. and Pogson, B.J. (2016) Learning the languages of the chloroplast: Retrograde signaling and beyond. *Annu. Rev. Plant Biol.* **67**, 25–53.

Chen, X. (2009) Small RNAs and their roles in plant development. *Annu. Rev. Cell Dev. Biol.* **25**, 21–44.

Dai, X., Zhuang, Z. and Zhao, P.X. (2018) psRNATarget: a plant small RNA target analysis server (2017 release). *Nucleic Acids Res.* **46**, W49–W54.

Diener, A.C., Gaxiola, R.A. and Fink, G.R. (2001) Arabidopsis ALF5, a multidrug efflux transporter gene family member, confers resistance to toxins. *Plant Cell*, **13**, 1625–1638.

Dinger, M.E., Pang, K.C., Mercer, T.R., Crowe, M.L., Grimmond, S.M. and Mattick, J.S. (2009) NRED: a database of long noncoding RNA expression. *Nucleic Acids Res.* **37**, D122–D126.

Fang, X., Zhao, G., Zhang, S., Li, Y., Gu, H., Li, Y., Zhao, Q. and Qi, Y. (2018) Chloroplast-to-nucleus signaling regulates microRNA biogenesis in Arabidopsis. *Dev Cell*, **48**, 371–382.

Fettke, J., Nunes-Nesi, A., Fernie, A.R. and Steup, M. (2011) Identification of a novel heteroglycan-interacting protein, HIP 1.3, from *Arabidopsis thaliana*. *J. Plant Physiol.* **168**, 1415–1425.

Franco-Zorrilla, J.M., Valli, A., Todesco, M., Mateos, I., Puga, M.I., Rubio-Somoza, I., Leyva, A., Weigel, D., Garcia, J.A. and Paz-Ares, J. (2007) Target mimicry provides a new mechanism for regulation of microRNA activity. *Nat. Genet.* **39**, 1033–1037.

Goksøyr, J. (1967) Evolution of eucaryotic cells. *Nature*, **214**, 1161.

Gray, J.C., Sullivan, J.A., Wang, J.H., Jerome, C.A. and Maclean, D. (2003) Coordination of plastid and nuclear gene expression. *Philos. Trans. R Soc. Lond B Biol. Sci.* **358**, 135–145.

Guan, Q., Lu, X., Zeng, H., Zhang, Y. and Zhu, J. (2013) Heat stress induction of miR398 triggers a regulatory loop that is critical for thermotolerance in Arabidopsis. *Plant J.* **74**, 840–851.

Gutierrez-Nava, M.D.L., Gillmor, C.S., Jimenez, L.F., Guevara-Garcia, A. and Leon, P. (2004) CHLOROPLAST BIOGENESIS genes act cell and noncell autonomously in early chloroplast development. *Plant Physiol.* **135**, 471–482.

Hagberg, A., Schult, D. and Swart, P. (2008). Exploring Network Structure, Dynamics, and Function using NetworkX. Proceedings of the 7th Python in Science conference (SciPy 2008). 11–15.

Heo, J.B. and Sung, S. (2011) Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. *Science*, **331**, 76–79.

Holoch, D. and Moazed, D. (2015) RNA-mediated epigenetic regulation of gene expression. *Nat. Rev. Genet.* **16**, 71–84.

Hong, Z., Zhang, Z., Olson, J.M. and Verma, D.P. (2001) A novel UDP-glucose transferase is part of the callose synthase complex and interacts with phragmoplastin at the forming cell plate. *Plant Cell*, **13**, 769–779.

Howell, M.D., Fahlgren, N., Chapman, E.J., Cumbie, J.S., Sullivan, C.M., Givan, S.A., Kasschau, K.D. and Carrington, J.C. (2007) Genome-wide analysis of the RNA-DEPENDENT RNA POLYMERASE6/DICER-LIKE4 pathway in Arabidopsis reveals dependency on miRNA- and tasiRNA-directed targeting. *Plant Cell*, **19**, 926–942.

Huang, C.Y., Wang, H., Hu, P., Hamby, R. and Jin, H. (2019) Small RNAs - Big players in plant-microbe interactions. *Cell Host Microbe*, **26**, 173–182.

Huang, D.W., Sherman, B.T. and Lempicki, R.A. (2009a) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13.

Huang, D.W., Sherman, B.T. and Lempicki, R.A. (2009b) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57.

Jin, H., Vacic, V., Girke, T., Lonardi, S. and Zhu, J.K. (2008) Small RNAs and the regulation of cis-natural antisense transcripts in Arabidopsis. *BMC Mol. Biol.* **9**(6). https://doi.org/10.1186/1471-2199-9-6

Kakizaki, T., Matsumura, H., Nakayama, K., Che, F.S., Terauchi, R. and Inaba, T. (2009) Coordination of plastid protein import and nuclear gene expression by plastid-to-nucleus retrograde signaling. *Plant Physiol.* **151**, 1339–1353.

Khraiwesh, B., Arif, M.A., Seumel, G.I., Ossowski, S., Weigel, D., Reski, R. and Frank, W. (2010) Transcriptional control of gene expression by microRNAs. *Cell*, **140**, 111–122.

Kim, C. and Apel, K. (2013) 1O2-mediated and EXECUTER-dependent retrograde plastid-to-nucleus signaling in norflurazon-treated seedlings of *Arabidopsis thaliana*. *Mol. Plant*, **6**, 1580–1591.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R. and Salzberg, S.L. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36.

Kim, V.N. (2005) Small RNAs: classification, biogenesis, and function. *Mol. Cells*, **19**, 1–15.

Klein, M. and Papenbrock, J. (2004) The multi-protein family of Arabidopsis sulphotransferases and their relatives in other plant species. *J. Exp. Bot.* **55**, 1809–1820.

Kleine, T. and Leister, D. (2016) Retrograde signaling: organelles go networking. *Biochim. Biophys. Acta*, **1857**, 1313–1325.

Kolde, R. (2019) Pretty Heatmaps. 1.0. 12 ed. https://cran.r-project.org/web/packages/pheatmap/pheatmap.pdf

Koussevitzky, S., Nott, A., Mockler, T.C., Hong, F., Sachetto-Martins, G., Surpin, M., Lim, J., Mittler, R. and Chory, J. (2007) Signals from chloroplasts converge to regulate nuclear gene expression. *Science*, **316**, 715–719.

Kramer, M.F. (2011) Stem-loop RT-qPCR for miRNAs. *Curr Protoc. Mol. Biol.* **95**(1), 15.10.1–15.10.15. https://doi.org/10.1002/0471142727.mb1510s95

Lapidot, M. and Pilpel, Y. (2006) Genome-wide natural antisense transcription: coupling its regulation to its different regulatory mechanisms. *EMBO Rep.* **7**, 1216–1222.

Larkin, R.M., Alonso, J.M., Ecker, J.R. and Chory, J. (2003) GUN4, a regulator of chlorophyll synthesis and intracellular signaling. *Science*, **299**, 902–906.

Liang, G., Yang, F. and Yu, D. (2010) MicroRNA395 mediates regulation of sulfate accumulation and allocation in *Arabidopsis thaliana*. *Plant J.* **62**, 1046–1057.

Liu, J., Li, Y., Wang, W., Gai, J. and Li, Y. (2016) Genome-wide analysis of MATE transporters and expression patterns of a subgroup of MATE genes in response to aluminum toxicity in soybean. *BMC Genom.* **17**, 223.

Livak, K.J. and Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*, **25**, 402–408.

Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550.

Lu, Y. (2016) Identification and roles of photosystem II assembly, stability, and repair factors in Arabidopsis. *Front. Plant Sci.* **7**, 168.

Ma, X., Shao, C., Jin, Y., Wang, H. and Meng, Y. (2014) Long non-coding RNAs: a novel endogenous source for the generation of Dicer-like 1-dependent small RNAs in *Arabidopsis thaliana*. *RNA Biol.* **11**, 373–390.

Mabbitt, P.D., Wilbanks, S.M. and Eaton-Rye, J.J. (2014) Structure and function of the hydrophilic Photosystem II assembly proteins: Psb27, Psb28 and Ycf48. *Plant Physiol. Biochem.* **81**, 96–107.

Maruta, T., Miyazaki, N., Nosaka, R. *et al.* (2015) A gain-of-function mutation of plastidic invertase alters nuclear gene expression with sucrose treatment partially via GENOMES UNCOUPLED1-mediated signaling. *New Phytol.* **206**, 1013–1023.

Megraw, M., Cumbie, J.S., Ivanchenko, M.G. and Filichkin, S.A. (2016) Small genetic circuits and microRNAs: big players in polymerase II transcriptional control in plants. *Plant Cell*, **28**, 286–303.

Meister, G. and Tuschl, T. (2004) Mechanisms of gene silencing by double-stranded RNA. *Nature*, **431**, 343–349.

Meng, Y., Shao, C., Ma, X., Wang, H. and Chen, M. (2012) Expression-based functional investigation of the organ-specific microRNAs in Arabidopsis. *PLoS One*, **7**, e50870.

Meskauskiene, R., Nater, M., Goslings, D., Kessler, F., op den Camp, R. and Apel, K. (2001) FLU: a negative regulator of chlorophyll biosynthesis in Arabidopsis thaliana. *Proc. Natl. Acad. Sci. USA*, **98**, 12826–21231.

Meyers, B.C., Axtell, M.J., Bartel, B. *et al.* (2008) Criteria for annotation of plant MicroRNAs. *Plant Cell*, **20**, 3186–3190.

Mochizuki, N., Brusslan, J.A., Larkin, R., Nagatani, A. and Chory, J. (2001) Arabidopsis genomes uncoupled 5 (GUN5) mutant reveals the involvement of Mg-chelatase H subunit in plastid-to-nucleus signal transduction. *Proc. Natl. Acad. Sci. USA*, **98**, 2053–2058.

Molla-Morales, A., Sarmiento-Manus, R., Robles, P., Quesada, V., Perez-Perez, J.M., Gonzalez-Bayon, R., Hannah, M.A., Willmitzer, L., Ponce, M.R. and Micol, J.L. (2011) Analysis of ven3 and ven6 reticulate mutants reveals the importance of arginine biosynthesis in Arabidopsis leaf development. *Plant J.* **65**, 335–345.

Omote, H., Hiasa, M., Matsumoto, T., Otsuka, M. and Moriyama, Y. (2006) The MATE proteins as fundamental transporters of metabolic and xenobiotic organic cations. *Trends Pharmacol. Sci.* **27**, 587–593.

Park, W., Li, J., Song, R., Messing, J. and Chen, X. (2002) CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana*. *Curr. Biol.* **12**, 1484–1495.

Pegler, J.L., Oultram, J.M.J., Grof, C.P.L. and Eamens, A.L. (2019) Profiling the abiotic stress responsive microRNA landscape of *Arabidopsis thaliana*. *Plants (Basel)*, **8**, 58.

Pornsiriwong, W., Estavillo, G.M., Chan, K.X. *et al.* (2017) A chloroplast retrograde signal, 3'-phosphoadenosine 5'-phosphate, acts as a secondary messenger in abscisic acid signaling in stomatal closure and germination. *Elife*, **6**, e23361.

Richter, A.S., Tohge, T., Fernie, A.R. and Grimm, B. (2020) The genomes uncoupled-dependent signalling pathway coordinates plastid biogenesis with the synthesis of anthocyanins. *Philos. Trans. R Soc. Lond. B Biol. Sci.* **375**, 20190403.

Rossel, J.B., Walter, P.B., Hendrickson, L., Chow, W.S., Poole, A., Mullineaux, P.M. and Pogson, B.J. (2006) A mutation affecting ASCORBATE PEROXIDASE 2 gene expression reveals a link between responses to high light and drought tolerance. *Plant Cell Environ.* **29**, 269–281.

Saini, G., Meskauskiene, R., Pijacka, W., Roszak, P., Sjogren, L.L., Clarke, A.K., Straus, M. and Apel, K. (2011) 'happy on norflurazon' (hon) mutations implicate perturbation of plastid homeostasis with activating stress acclimatization and changing nuclear gene expression in norflurazon-treated seedlings. *Plant J.* **65**, 690–702.

Sham, A., Moustafa, K., Al-Ameri, S., Al-Azzawi, A., Iratni, R. and Abuqamar, S. (2015) Identification of Arabidopsis candidate genes in response to biotic and abiotic stresses using comparative microarrays. *PLoS One*, **10**, e0125666.

Strand, A., Asami, T., Alonso, J., Ecker, J.R. and Chory, J. (2003) Chloroplast to nucleus communication triggered by accumulation of Mg-protoporphyrinIX. *Nature*, **421**, 79–83.

Susek, R.E., Ausubel, F.M. and Chory, J. (1993) Signal transduction mutants of Arabidopsis uncouple nuclear CAB and RBCS gene expression from chloroplast development. *Cell*, **74**, 787–799.

Swiezewski, S., Liu, F., Magusin, A. and Dean, C. (2009) Cold-induced silencing by long antisense transcripts of an Arabidopsis Polycomb target. *Nature*, **462**, 799–802.

Szyrajew, K., Bielewicz, D., Dolata, J., Wojcik, A.M., Nowak, K., Szczygiel-Sommer, A., Szweykowska-Kulinska, Z., Jarmolowski, A. and Gaj, M.D. (2017) MicroRNAs are intensively regulated during induction of somatic embryogenesis in Arabidopsis. *Front. Plant Sci.*, **8**, 18.

Tadini, L., Pesaresi, P., Kleine, T. *et al.* (2016) GUN1 controls accumulation of the plastid ribosomal protein S1 at the protein level and interacts with proteins involved in plastid protein homeostasis. *Plant Physiol.* **170**, 1817–1830.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J. and Pachter, L. (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515.

Voinnet, O. (2009) Origin, biogenesis, and activity of plant microRNAs. *Cell*, **136**, 669–687.

Wang, H.V. and Chekanova, J.A. (2017) Long Noncoding RNAs in Plants. *Adv. Exp. Med. Biol.* **1008**, 133–154.

Wang, L., Bei, X., Gao, J., Li, Y., Yan, Y. and Hu, Y. (2016) The similar and different evolutionary trends of MATE family occurred between rice and *Arabidopsis thaliana*. *BMC Plant Biol.* **16**, 207.

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

Wierzbicki, A.T., Haag, J.R. and Pikaard, C.S. (2008) Noncoding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. *Cell*, **135**, 635–648.

Woodson, J.D., Perez-Ruiz, J.M. and Chory, J. (2011) Heme synthesis by plastid ferrochelatase I regulates nuclear gene expression in plants. *Curr. Biol.* **21**, 897–903.

Woodson, J.D., Perez-Ruiz, J.M., Schmitz, R.J., Ecker, J.R. and Chory, J. (2013) Sigma factor-mediated plastid retrograde signals control nuclear gene expression. *Plant J.* **73**, 1–13.

Yuan, C., Wang, J., Harrison, A.P., Meng, X., Chen, D. and Chen, M. (2015) Genome-wide view of natural antisense transcripts in *Arabidopsis thaliana*. *DNA Res.* **22**, 233–243.

Zhang, X., Xia, J., Lii, Y.E. *et al.* (2012) Genome-wide analysis of plant nat-siRNAs reveals insights into their distribution, biogenesis and function. *Genome Biol.* **13**, R20.

Zhao, C., Wang, Y., Chan, K.X. *et al.* (2019a) Evolution of chloroplast retrograde signaling facilitates green plant adaptation to land. *Proc. Natl. Acad. Sci. USA*, **116**, 5015–5020.

Zhao, X., Huang, J. and Chory, J. (2019b) GUN1 interacts with MORF2 to regulate plastid RNA editing during retrograde signaling. *Proc. Natl. Acad. Sci. USA*, **116**, 10162–10167.

Zimorski, V., Ku, C., Martin, W.F. and Gould, S.B. (2014) Endosymbiotic theory for organelle origins. *Curr. Opin. Microbiol.* **22**, 38–48.