Short title: Control of mRNA translation by *cis*-NATs

Control of cognate sense mRNA translation by *cis*-natural antisense RNAs

Jules Deforges[1], Rodrigo S Reis[1], Philippe Jacquet[1], Shaoline Sheppard[1], Veerendra P. Gadekar[2], Gene Hart-Smith[3], Andrea Tanzer[2], Ivo L Hofacker[2], Christian Iseli[4], Ioannis Xenarios[4], Yves Poirier[1]

[1]Department of Plant Molecular Biology, University of Lausanne, Biophore Building, CH-1015 Lausanne, Switzerland
[2]Institute of Theoretical Chemistry, University of Vienna, Wahringer Str 17, A-1090 Vienna, Austria
[3] School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, Australia
[4]Swiss Institute of Bioinformatics, Lausanne CH-1015, Switzerland

Footnotes:

Author Contributions

JD, RSR and YP conceived and designed the study. JD performed all experiments except for the establishment and performance of the protoplast transformation assay done by RSR. Bioinformatic analysis of the data was performed by JD and PJ, whereas proteomic analysis was performed by GHS. Analysis of the coding potential of *cis*-NATs was done by VPG, AT and ILH, whereas analysis of *cis*-NATs:sense mRNA configuration and main characteristics was done by SP and JD. Bioinformatic support was also provided by CI and IX. JD, RSR and YP wrote the paper. All authors read and approved the final manuscript. YP agrees to serve as the author responsible for contact and ensures communication.

This work was supported by a Sinergia grant (CRSII3_154471) from the Swiss National Fund to YP and ILH. The Swiss Institute of Bioinformatic (SIB) and Vital-IT Center for high-performance computing of the SIB are supported by University of Lausanne and the Swiss Federal Government.

Corresponding author: yves.poirier@unil.ch

One-sentence summary: A robust pipeline identifies and experimentally validates cis-natural antisense transcripts controlling cognate sense mRNA translation.

Abstract

Cis-Natural Antisense Transcripts (*cis*-NATs), which overlap protein coding genes and are transcribed from the opposite DNA strand, constitute an important group of non-coding RNAs. Whereas several examples of *cis*-NATs regulating the expression of their cognate sense gene are known, most *cis*-NATs function by altering the steady-state level or structure of mRNA via changes in transcription, mRNA stability or splicing, and very few cases involve the regulation of sense mRNA translation. This study was designed to systematically search for *cis*-NATs influencing cognate sense mRNA translation in *Arabidopsis thaliana*. Establishment of a pipeline relying on sequencing of total polyA$^+$ and polysomal RNA from Arabidopsis grown under various conditions (i.e., nutrient deprivation and phytohormone treatments) allowed the identification of 14 *cis*-NATs whose expression correlated either positively or negatively with cognate sense mRNA translation. Using a combination of *cis*-NAT stable over-expression in transgenic plants and transient expression in protoplasts, the impact of *cis*-NAT expression on mRNA translation was confirmed for 4 out of 5 tested *cis*-NAT:sense mRNA pairs. These results expand the number of *cis*-NATs known to regulate cognate sense mRNA translation and provide a foundation for future studies of their mode of action. Moreover, this study highlights the role of this class of non-coding RNAs in translation regulation.

Introduction

A large proportion of the genome of eukaryotes is transcribed into RNA that is not coding for proteins or house-keeping RNAs (e.g. tRNAs, ribosomal RNAs) (Djebali et al., 2012). Whereas first being considered as transcriptional noise, non-coding RNAs have emerged as major regulators of gene expression (Rinn and Chang, 2012; Bonasio and Shiekhattar, 2014; Chekanova, 2015; Ransohoff et al., 2018). Besides the well-characterized small RNAs that include short interfering RNAs (siRNAs) and micro RNAs (miRNAs), abundant long non-coding RNAs (lncRNAs) have been identified across a wide spectrum of organisms. LncRNAs are typically defined as capped and polyadenylated transcripts longer than 200 bases that do not contain conserved open reading frame capable of encoding proteins (Rinn and Chang, 2012; Bonasio and Shiekhattar, 2014; Chekanova, 2015; Ransohoff et al., 2018). However, recent studies have indicated that some lncRNAs could associate with ribosomes and, in some cases, generate small peptides (Ji et al., 2015; Hsu et al., 2016; Bazin et al., 2017). Whereas most of the lncRNAs identified by genome-wide studies have yet unknown functions, an increasing number has been shown to be involved in critical biological processes such as X chromosome inactivation in mammals (Brockdorff et al., 1992) or flowering in plants (Liu et al., 2010; Heo and Sung, 2011).

LncRNAs located in intergenic regions relative to coding genes are defined as long intergenic non-coding RNA (lincRNAs), whereas lncRNAs overlapping coding genes and transcribed from the opposite DNA strand are categorized as *cis*-Natural Antisense Transcripts (*cis*-NATs) (Rinn and Chang, 2012). In addition, lincRNAs able to bind target mRNAs by partial base-pair complementarity are defined as *trans*-Natural Antisense Transcripts (*trans*-NAT) (Lapidot and Pilpel, 2006). *Cis*-NATs are widespread in eukaryotes, with 20–70% of coding genes having an associated *cis*-NAT in *Saccharomyces cerevisiae*, *Drosophila melanogaster*, mice, human, and *Arabidopsis thaliana* (Faghihi and Wahlestedt, 2009; Liu et al., 2015). *Cis*-NATs can overlap completely with their cognate mRNAs or only at the 5' (head-to-head orientation) or the 3' end (tail-to-tail orientation).

Whereas various modes of action have been reported for lncRNAs impacting the regulation of target gene expression, the majority involves changes in the steady-state level or structure of mRNA via changes in transcription, mRNA stability or splicing (Rinn and Chang, 2012; Bonasio and Shiekhattar, 2014; Chekanova, 2015; Ransohoff et al., 2018). This is true for either lincRNAs or *cis*-NATs and applies to both animal and plant models. Well-characterized mechanisms by which lncRNAs affect gene transcription include recruitment of chromatin or transcription regulators and

4

displacement of transcriptional repressors. Examples of such mechanisms in animals includes inhibition of transcription via histone methylation by *HOTAIR* (Gupta et al., 2010) or DNA methylation by *pRNA* (Schmitz et al., 2010), stimulation of transcription by the recruitment of the activator PYGO2 by the lincRNA *PCGEM1* (Schmitz et al., 2010), and the displacement of the repressive glucocorticoid response element by the lincRNA *Gas5* (Kino et al., 2010). Similar mechanisms for lncRNAs in plants have been described, such as histone modification at the flowering locus *FLC* triggered by the *cis*-NAT *COOLAIR* (Liu et al., 2010) or the intronic lncRNA *COLDAIR* (Heo and Sung, 2011), as well as the transcriptional activation of pathogen-responsive gene *PR1* via the recruitment of a Mediator component by lincRNA *ELF18* (Seo et al., 2017). LincRNA can also interact with splicing factors to regulate alternative splicing, as described for lincRNA *MALAT1* in animals (Tripathi et al., 2010) and *ASCO* in plants (Bardou et al., 2014). Moreover, lincRNA can control mRNA stability via interaction with members of the *Staufen* double-stranded RNA (dsRNA)-binding proteins in animals (Gong and Maquat, 2011) or inhibition of microRNA action on mRNA degradation via target mimicry, as described for the Arabidopsis *IPS1* lincRNA involved in the response of plants to inorganic phosphate (Pi) deficiency (Franco-Zorilla et al., 2007). Another mechanism for the decrease in steady-state mRNA level associated with *cis*-NAT expression is the generation of siRNA via processing of double-stranded RNA generated by the overlapping *cis*-NAT with its cognate sense mRNA (Khorkova et al., 2014). However, considering the large number of potential *cis*-NAT:sense mRNA pairs in animal and plant genomes, relatively few examples of siRNA-mediated effects for *cis*-NATs have been described, indicating that this mechanism may be less frequently employed than initially thought.

In contrasts to the effects of lncRNA on transcription and mRNA levels, examples of modulation of mRNA translation by lncRNA are rather rare. Examples involving lincRNAs in human cell lines are the inhibition of translation of targets *CTNNB1* and *JUNB* via recruitment of the translational repressor Rck by the lincRNA-p21 (Yoon et al., 2012) and the inhibition of *c-Myc* translation by the recruitment of the eukaryotic initiation factor eIF4E by lncRNA *GAS5* (Hu et al., 2014). Repression of mRNA translation was also demonstrated for the *cis*-NAT of the *PU.1* gene, encoding a transcription factor in mammals (Ebralidze et al., 2008). Recently, three examples for the enhancement of translation by *cis*-NATs have been described. In rice (*Oryza sativa*), expression of the *cis*-NAT of the Pi exporter gene *PHO1;2* was shown to enhance the association of the cognate mRNA to polysomes, leading to the accumulation of PHO1;2 protein despite unchanged steady-state level of the corresponding mRNA (Jabnoune et al., 2013). In mice, *Uchl1* mRNA translation is enhanced by a *cis*-NAT that is exported to the cytoplasm upon inhibition of the Target of Rapamycin (TOR) pathway (Carrieri et al., 2012). Finally, the human regulator of megakaryocyte

differentiation *RBM15* is also translationally enhanced by a *cis*-NAT (Tran et al., 2016). Little is known about the mechanisms of action of the three translation enhancer *cis*-NATs reported so far. All three pairs are oriented in a head-to-head manner (5'-5'). For *RBM15*, the region of the antisense overlapping with the sense mRNA 5'UTR alone was found to be sufficient to enhance translation (Tran et al., 2016). In contrast, for *Uchl1*, two elements in the *cis*-NAT were found to be essential, namely the region overlapping with the 5' end of *Uchl1* mRNA and a non-overlapping inverted Short Interspersed Nuclear Element (SINE) B2 element, a class of retrotransposable repeat element (Carrieri et al., 2012). More recently, *cis*-NATs containing distinct SINE elements have been identified in mammals as potential translation enhancers (Schein et al., 2016), whereas expression of some ribosome-associated *cis*-NATs in plants were correlated with increased mRNA translation (Bazin et al., 2017).

The low number of *cis*-NATs experimentally validated to influence translation of the cognate mRNA might reflect the fact that most genome-wide studies of *cis*-NATs examined the correlation between steady-state level of mRNAs and the expression of *cis*-NATs, an approach that is not suitable for studying translation. In the present study, we took advantage of the polysome profiling method combined with strand-specific RNA sequencing to identify, in *A. thaliana* plants, *cis*-NATs whose expression level were associated with a change of cognate sense mRNA level, as well as translation across a range of experimental conditions. The impact of *cis*-NAT expression on cognate mRNA translation was further validated by expression of several *cis*-NATs in transgenic *A. thaliana* and/or by transient expression in protoplasts.

150
151

RESULTS

Experimental setup to identify *cis*-NATs associated with changes in mRNA level and mRNA translation

In order to identify *cis*-NATs impacting their cognate sense mRNA transcript level as well as mRNA translation, an experimental procedure was set up allowing the quantification of steady-state levels of coding and non-coding RNAs along with the determination of mRNA translation efficiency genome-wide in *A. thaliana* seedlings grown under various conditions. Whole seedlings grown in liquid cultures in the presence of a high (1 mM) or low (100 µM) concentration of Pi were analyzed, as well as roots and shoots from seedlings grown on agar-solidified medium supplemented with different phytohormones, namely auxin (indole acetic acid, IAA), abscisic acid (ABA), methyl-jasmonate (MeJA) or 1-aminocyclopropane-1-carboxylic acid (ACC), a precursor of ethylene. For each sample, steady-state levels of *cis*-NATs and mRNAs were determined by strand-specific sequencing of total polyA$^+$ RNA, whereas translation efficiency was assessed for the same sample by sequencing polysome-associated RNA purified by centrifugation through sucrose density gradients. Sequencing of each total or polysomal RNA sample yielded between 30 and 60 million paired-end reads. Three independent biological replicates were analyzed for each treatment, and a total of at least 120 million paired-end reads were obtained per condition.

The genes up- or down-regulated in response to the different treatments were identified by pairwise comparisons between hormone treated or low Pi samples and their corresponding controls. In response to low Pi, 2,991 protein-coding genes (according to the TAIR10 annotation) were significantly up-regulated with a fold change > 2 and adjusted p-value (adj.pval) < 0.1, and 2,149 were significantly down-regulated (Figure 1A, Supplemental Table S1-S2). Fewer genes were differentially expressed in response to the different hormone treatments (Supplemental Figure S1, Supplemental Table S1-S2). For example, upon auxin treatment, 377 and 120 protein-coding genes were up-regulated in roots and shoots, respectively (Supplemental Table S2). Untreated root and shoot tissues were also compared and their transcriptomes were dramatically different, as expected for two different organs, with 3,906 and 4,742 protein-coding genes significantly up- or down-regulated, respectively, in roots relative to shoots (Supplemental Table S2).

Quality assessment of the transcriptomic data was first performed by Gene Ontology (GO) term enrichment analyses for each set of up-regulated genes. This analysis confirmed the strong induction of marker genes associated with the different treatments. The genes up-regulated in response to low Pi were significantly enriched for the GO term “cellular response to Pi starvation” (GO:0016036 adj.pval=5.3×10^{-11}) (Supplemental Figure S2A). Among these, *Induced by Phosphate*

Figure 1

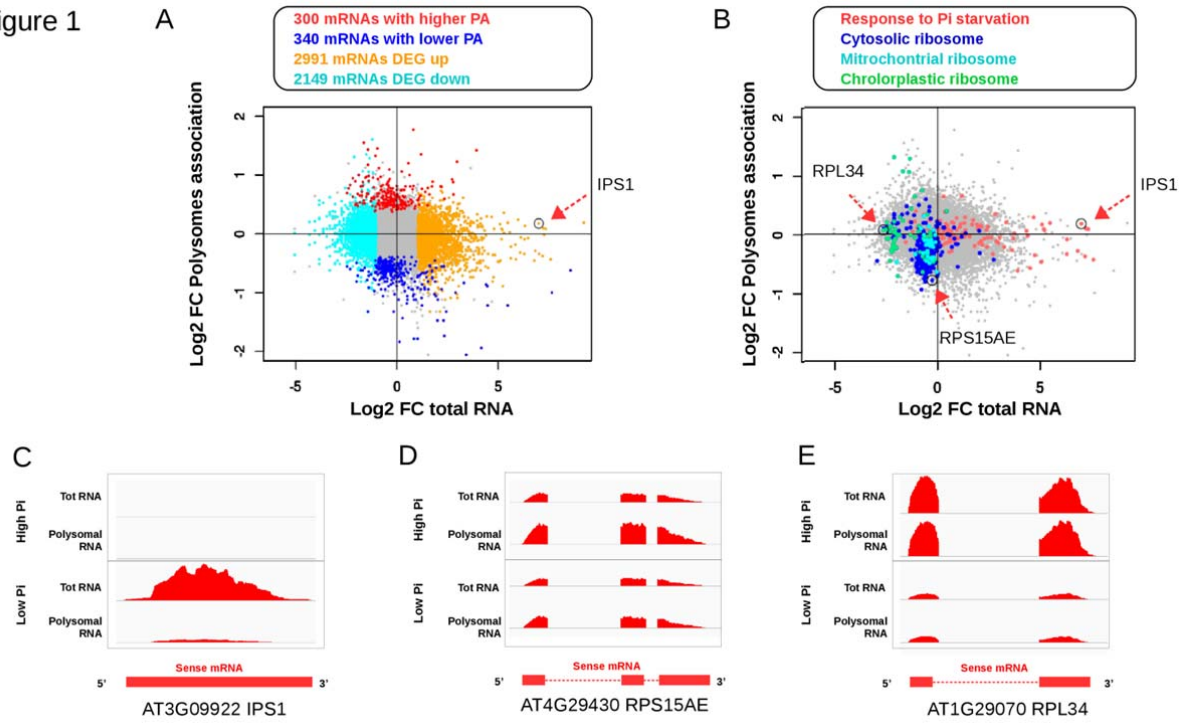

Figure 1. Steady state mRNA expression level and association with polysomes in response to grow of Arabidopsis to low Pi condition. A: Relation between log2-fold change of mRNA steady-state level (x axis) is plotted against the log2-fold change in polysome association (y axis). Coding genes significantly up- or down- regulated at the mRNA steady-state level are colored in yellow and cyan respectively, while mRNA significantly more or less associated with polysomes are colored in red and blue, respectively. The genes not showing any statistical difference are colored in grey. **B:** Same plot as A where genes associated with GO terms "Response to Pi starvation", "Cytosolic ribosome", "Mitochondrial ribosome", and "Chloroplastic ribosome" are colored in pink, dark blue, light blue and green, respectively. **C-E:** Normalized RNA-seq coverage plots for the *IPS1*, *RPS15AE* and *RPL34* genes. The two upper panels show the coverage plots for total mRNA and polysomal RNA from high Pi samples and the two lower panels correspond to low phosphate samples. The schematic exonic organization of each gene is represented by red boxes and lines below the plots.

Starvation 1 (*IPS1*), a known highly induced marker of Pi deficiency, was strongly over-expressed (fold change=127.9, Figure 1A, C). Similarly, the up-regulated genes were significantly enriched in GO terms "response to abscisic acid" (GO:0009737, adj.pval=$1.3x10^{-10}$), "response to auxin" (GO:0009733, adj.pval=$6.5x10^{-11}$), "jasmonic acid metabolic process" (GO:0009694, adj.pval=$5.6x10^{-11,}$) and "ethylene-activated signaling pathway" (GO:0009873, adj.pval=$6.4x10^{-8}$) in root samples treated with auxin, ABA, methyl jasmonate and ACC, respectively, compared to untreated roots (Supplemental Figure S2B-E).

The genes differentially expressed in response to low Pi and ABA treatments were further compared to previously published datasets (Supplemental Figure S3). Approximately 71% of the genes up-regulated in whole seedlings upon ABA treatment in the study of Song et al. (2016) were also up-regulated in ABA-treated roots and/or shoots in our dataset (941 genes out of 1,327). Similarly, 79% of genes up-regulated in the study of Bazin et al. (2017) (154 / 194) and 30.5% of genes up-regulated either in roots or shoot in Yuan et al. (2016) were common with the genes up-regulated under our low-Pi condition. The lower proportion of common differentially expressed genes with that reported by Yuan et al. (2016) might be explained by the differences in terms of

tissue analyzed (root versus whole seedlings) and growth conditions (e.g. liquid versus solid medium and different Pi concentrations).

Analysis of differential mRNA translation

Translation efficiency of coding genes can be estimated by measuring the proportion of mRNA molecules associated with polysomes relative to the amount of total RNA, as previously described (Mustroph et al., 2009; Juntawong et al., 2014). Using sequencing data from polysome-associated RNA, the ratio of polysome association (PA) was calculated for each gene by dividing the normalized readcount in the polysomal RNA fraction by the normalized readcount measured for total RNA steady-state level. The treated samples were compared to that in the corresponding control conditions and loci with a 30% increase or decrease in PA ratio and an adj.pval < 0.1 were considered differentially associated with polysomes, and thus potentially regulated translationally (Supplemental Table S1, S3). In response to growth under low-Pi conditions, 300 and 340 protein-coding genes were significantly more and less associated with polysomes, respectively, compared to that under high-Pi conditions (Figure 1A). GO enrichment analyses revealed that the coding genes with a lower association with polysomes in response to low Pi were strongly enriched for ribosomal proteins (GO:0022626cellular component "cytosolic ribosome", adj.pval=$2.26x10^{-11}$) (Figure 1B, Supplemental Figure S4A), such as the cytosolic ribosomal protein RPS15AE (Figure 1D). This finding was consistent with that of previous reports, where a similar down-regulation of the translation of the ribosomal proteins was observed by ribosome footprints in response to both Pi deficiency (Bazin et al., 2017) and hypoxia (Juntawong et al., 2014), validating both techniques for the analysis of mRNA translation. Of note, most of the genes constituting chloroplastic ribosomal proteins, such as RPL34 (Figure 1B, E) showed a strong decrease in mRNA steady-state level without significant change in PA. The genes encoding mitochondrial ribosomal proteins on the other hand were globally less associated with polysomes, similarly to those encoding cytosolic ribosomal proteins (Figure 1B).

Many genes were also found differentially associated with polysomes when comparing root and shoot tissues (Supplemental Figure S5, Supplemental Table S3). For example, 946 protein-coding genes were significantly more associated with polysomes in roots and 1,033 in shoots, in untreated samples. Interestingly, the strongest enrichment within the set of genes with higher PA in shoots corresponds to GO:0008380 "mRNA splicing" (adj.pval= $6x10^{-11}$) (Supplemental Figure S4B). *SERRATE* and *SR45*, for example, were strongly associated with polysomes in shoots and very poorly in roots (fold change PA = 5.3 and 4.2 for *SERRATE* and *SR45*, respectively), despite similar steady-state levels of mRNA in both tissues (Supplemental Figure S5B-D). Both SERRATE and

SR45 have been experimentally validated to participate in splicing, with an additional role for SERRATE in microRNA processing (Laubinger et al., 2008; Zhang et al., 2017).

De novo identification of *cis*-NATs

Cis-NATs expressed in response to the different treatments were identified using the pipeline described in the Materials and Methods (Figure 2A). In this pipeline, pairs of protein-coding genes having mRNAs that may overlap in a sense-antisense fashion are not included as *bona fide cis*-NATs. *De novo* transcriptome annotations corresponding to each of the 12 experimental conditions analyzed were merged, and, after comparison to the TAIR10.31 annotation (Berardini et al., 2015), a novel set of 4,411 *cis*-NATs were identified. Approximately 9% (374) of these *cis*-NATs were recently annotated in the Araport11 database (Cheng et al., 2017). We then used the FEELnc tool (Wucher et al., 2017) (see Materials and Methods and Supplemental Materials and Methods for details) to determine *in silico* the coding potential of all the newly identified *cis*-NATs. The large majority of these *cis*-NATs (98.5%) were lacking coding potential and only 63 were predicted to be potentially coding. This prediction of coding potential was well supported by our experimental polysome profiling data since the *cis*-NATs predicted to be coding were significantly more associated with polysomes than seen for the non-coding *cis*-NATs (Figure 2B). A similar difference was observed when comparing protein-coding and non-coding genes annotated in the TAIR10 database.

Exploring the conservation across plant genomes of the peptides encoded by the 63 *cis*-NATs with high coding potential, we identified a group of 10 peptides that were well conserved amongst plant genomes, and a second group of nine peptides conserved amongst Brassicaceae species only (Supplemental Figure S6). The remaining 44 predicted coding *cis*-NATs were poorly or not conserved. Seven of the *cis*-NATs encoding conserved peptides (Group I or II) were recently annotated as (putative) protein-coding genes in the ARAPORT11 database but not in TAIR10. The transcripts encoding these evolutionary conserved peptides should thus likely be considered as ~~novel~~ protein-coding genes.

Expression of the identified *cis*-NATs was well supported by published epigenetic profiling data from Jégu T. et al. (2017). The predicted transcription start sites of the *cis*-NATs were strongly enriched for the activating histone mark H3K9Ac as well as micrococcal S7 nuclease (Mnase) footprints to the same extent as that in TAIR10 protein-coding loci, confirming the transcriptionally active state of the promoter regions of the *cis*-NATs (Supplemental Figure S7). Moreover, 60% of

Figure 2

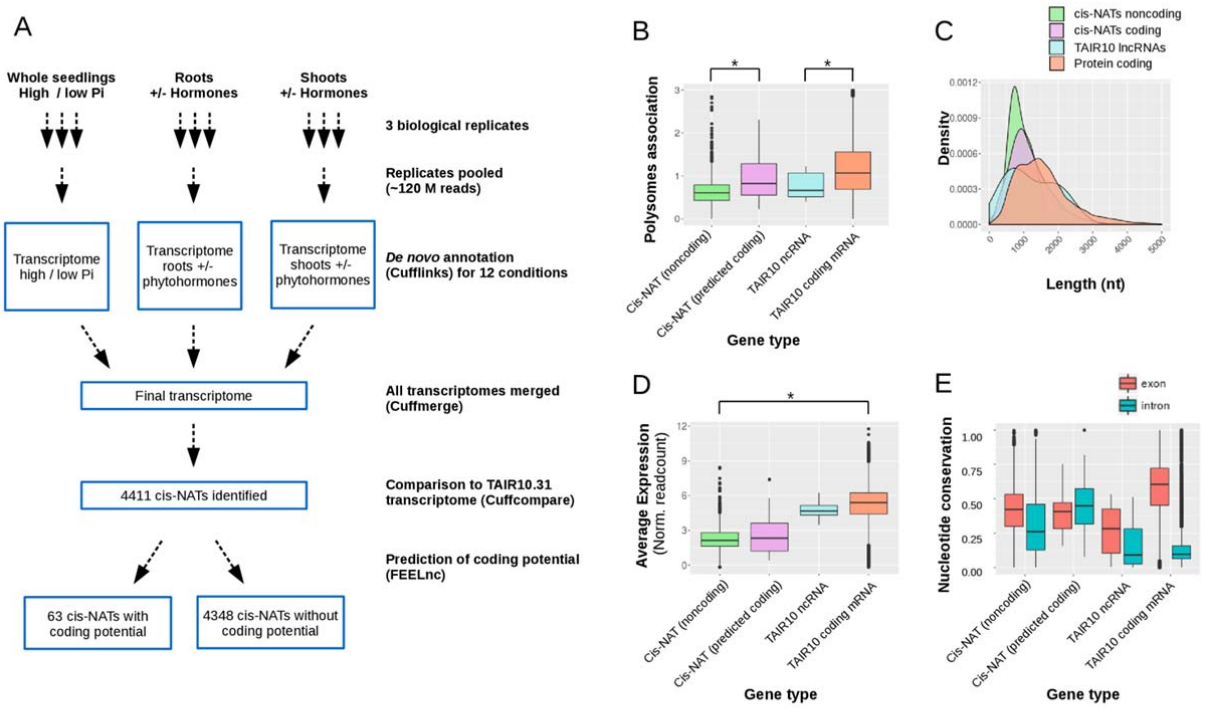

Figure 2: Identification and characterization of *cis*-NATs. A: Schematic diagram of the pipeline used for *de novo cis*-NAT identification from the 12 different experimental conditions. **B:** Boxplot comparing polysome association between *cis*-NATs predicted noncoding (green), or with coding potential (pink), ncRNA (cyan) and protein coding genes (salmon) annotated in TAIR10 database. **C, D:** Plots comparing transcript length (C) and RNA steady-state-level (D) between *cis*-NATs predicted noncoding (green), or with coding potential (pink), ncRNA (cyan) and protein coding genes (salmon) annotated in TAIR10 database. **E:** Boxplots comparing the nucleotide conservation across 20 angiosperm genomes within exonic and intronic regions of the four categories of transcripts listed above.

the *cis*-NATs detected in our dataset overlapped with *cis*-NATs previously identified in at least one of the three datasets used for comparison: the PlncDB database (Jin et al., 2013; Wang et al., 2014)• as well as the work of Yuan et al. (2016) and Bazin et al. (2017)• (Supplemental Figure S8).

Cis-NATs were on average shorter, expressed at a lower level and had a weaker genomic sequence conservation (PHASTcons score) compared to that of TAIR10 annotated non-coding RNA and protein-coding mRNAs (Figure 2C-E), consistent with previous reports of *cis*-NATs in plants and other eukaryotes (Wang et al., 2005; Khorkova et al., 2014; Yuan et al., 2015). Furthermore, the polysome association value of *cis*-NATs was significantly lower (0.64) compared to that of mRNAs (1.19), but similar to that of the non-coding transcripts annotated in the TAIR10 database (0.54) (Figure 2B).

To validate our pipeline of identification of differentially expressed *cis*-NATs and protein-coding genes, we analyzed by RT-qPCR the expression level of six protein-coding genes and six *cis*-NATs predicted to be up- or down-regulated in response to phosphate starvation. For the 12 genes analyzed, the RT-qPCR results showed a significant increase or decrease in RNA steady-state level in agreement with the RNAseq data (Supplemental Figure S9).

289

290 **_Cis_-NATs associated with changes in steady-state level of their cognate mRNA**

291 To identify *cis*-NATs potentially regulating the expression of their cognate mRNA at the
292 transcriptional or translational level, we looked for correlation between expression levels of all
293 available *cis*-NATs (e.g. *cis*-NATs identified in this study and those included in TAIR10.31, total of
294 4,846 *cis*-NATs) and cognate sense mRNA steady-state levels or polysome association across the
295 different experimental conditions analyzed, comparing hormone treated samples with untreated
296 controls for root and shoot tissues as well as seedlings grown under low- or high-Pi conditions.
297 Untreated root and shoot tissues were also compared to identify *cis*-NATs potentially regulating
298 tissue-specific gene expression. All *cis*-NAT:sense mRNA pairs were put into four categories
299 considering their region of overlap, namely overlap in the 5' end, 3' end, completely included within
300 the sense region, or *cis*-NATs that extend beyond the 5' and 3' region of the coding sense
301 (overhanging) (Supplemental Figure S10).

302

303 Analysis for potential effects of *cis*-NAT expression on steady-state sense mRNA level was
304 performed. For each pairwise comparison, the *cis*-NAT:mRNA pairs were considered correlated if
305 both the *cis*-NAT and the cognate mRNA were differentially expressed, with a fold change of at
306 least 2 and a FDR < 0.1. A total of 1,310 *cis*-NATs, including 67 annotated in TAIR10, were
307 differentially expressed in at least one condition (Supplemental Table S1). For 107 of these loci,
308 steady-state level of the cognate mRNA was positively correlated to *cis*-NAT expression (Table 1,
309 Supplemental Table S4). For example, both the mRNA and the *cis*-NAT of the locus AT2G37580
310 were significantly up-regulated upon ABA treatment in shoots (FC= 2.13, adj.pval=0.029 for the
311 mRNA and FC=2.06, adj.pval=$1.3x10^{-2}$ for the *cis*-NAT) (Figure 3A). We also found 41 pairs with
312 a negative correlation, such as AT1G68940, whose mRNA was up-regulated in response to low-Pi
313 conditions (FC=2.16, adj.pval=$7.9x10^{-3}$), whereas the *cis*-NAT was down-regulated (FC= 0.39,
314 adj.pval=$7.2x10^{-2}$) (Figure 3B). Pearson correlation coefficient between *cis*-NAT and mRNA
315 expression was calculated across the 12 experimental conditions analyzed, taking advantage of the
316 whole experimental dataset to identify *cis*-NATs with stronger positive and negative expression
317 correlation with their cognate sense mRNA. This analysis revealed that 86 *cis*-NAT:sense mRNA
318 pairs out of 107 had a positive correlation coefficient higher than 0.4, whereas 27 *cis*-NAT:sense
319 mRNA pairs out of 41 had a negative correlation coefficient lower than -0.4, across all 12
320 experimental conditions (Figure 3E-F, Table 1).

321

322 **Identification of putative translation regulator *cis*-NATs**

Figure 3

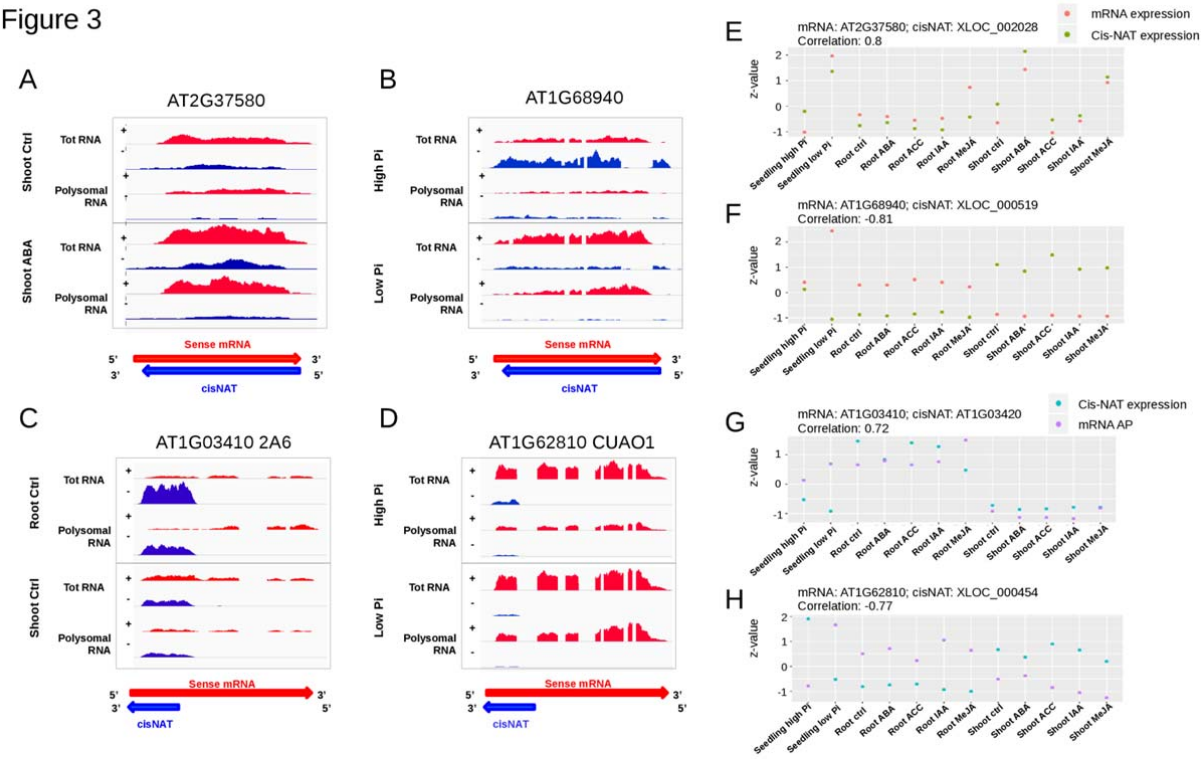

Figure 3: Correlations between expression of *cis*-NATs and changes in steady-state level or poylsome association of the cognate sense mRNA. A-D: Coverage plots showing the density of RNA-seq reads per position at AT2G37580, AT1G68940, AT1G03410 and *CuAO1* loci. The red and blue areas represent the density of reads mapping to the sense mRNA and *cis*-NATs, respectively. For each experimental condition, the upper part corresponds to total RNA-seq reads and the lower part to polysomal RNA-seq reads. The red and blue arrows below indicate the *cis*-NAT-mRNA pair orientation. **E-H:** Correlation plots showing the steady-state level of the coding mRNA (red dots) and *cis*-NAT (green dots) for AT2G37580 and AT1G68940 loci (E and F, respectively) or the steady-state level of the *cis*-NAT (cyan dots) and the association with polysomes of the cognate sense mRNA (purple dots) for AT1G03410 and *CuAO1* loci (G and H, respectively). The Z-score of normalized read counts calculated from the 12 experimental conditions is represented on the Y-axis. Pearson correlation coefficients between the 2 variables shown in each plot are indicated on top of the plots.

In order to identify *cis*-NATs influencing the translation of their cognate sense mRNA, we looked for *cis*-NAT:sense mRNA pairs where the *cis*-NAT was differentially expressed (fold change > 2 and adj.pval < 0.1) and the sense mRNA that was differentially associated with polysomes (at least 30% increase or decrease, adj.pval < 0.1) in response to treatment. A finer filtering step was also performed using additional criteria such as the size of the overlapping region or the relative level of expression between *cis*-NAT and mRNA (see Material and Methods for further details). A total of eight *cis*-NAT:sense mRNA pairs were identified for which *cis*-NAT differential expression was positively correlated to mRNA differential association with polysomes in at least one pairwise comparison (Supplemental Table S4). For example, AT1G03410 mRNA was more associated with polysomes (FC=1.79, adj.pval=0.01) when the *cis*-NAT was more expressed (FC=2.71, ajd.pval=$5.6x10^{-18}$) in untreated roots samples compared to untreated shoots (Figure 3C). A total of six pairs showed a negative correlation between *cis*-NAT expression and cognate mRNA association with polysomes (Supplemental Table S4), including the *CuAO1* locus whose sense mRNA was more associated with polysomes (FC=1.68, adj.pval=0.05) when the *cis*-NAT was less abundant (FC=0.29, adj.pval=$1.3x10^{-5}$) in Pi-deficient seedlings (Figure 3D). The expression of three out of the eight *cis*-NAT:sense mRNA pairs with a positive correlation and three out of the six pairs with negative correlation had a Pearson correlation coefficient > 0.4 and < -0.4, respectively,

with polysome association of their cognate mRNAs across the 12 experimental conditions (Figure 3G-H, Table 1, Supplemental Table S4).

The *cis*-NATs identified that positively or negatively correlated with sense mRNA steady-state transcript level or polysome association were further analyzed for relation with miRNAs, i.e. presence of miRNA precursor sequence, miRNA target sequence, and potential as a microRNA target mimic (see Material and Methods). Out of the 4,846 *cis*-NATs analyzed, 14% (682) were predicted to contain at least one miRNA binding site (Supplemental Table 1), including 7 and 14 *cis*-NATs negatively and positively correlated to cognate mRNA steady-state level, respectively (Supplemental Table 1). Two *cis*-NATs positively correlated with mRNA polysome association were predicted to contain miRNA binding sites, but none of the *cis*-NATs with a negative correlation. Only seven *cis*-NATs were predicted as miRNA precursors and 69 contained potential miRNA target mimic sites, including two *cis*-NATs positively correlated with cognate mRNA expression and one *cis*-NAT positively correlated with mRNA polysome association (Supplemental Table S1). No *cis*-NAT negatively correlated to mRNA expression or polysome association contained a putative miRNA target mimic site.

We also took advantage of 40 publicly available small RNA datasets to analyze the *cis*-NATs in relation to siRNAs. We identified 24,119,910 small reads between 18 and 28 nucleotides long mapping to TAIR10 reference genome. Of those, 666,181 mapped to *cis*-NAT loci and were considered as *cis*-NAT-siRNAs. Most of them were 21 and 24 nucleotides long (Supplemental Figure S11A) and the overlapping region of *cis*-NATs showed a significantly higher density in small RNAs compared to that of non-overlapping regions (Supplemental Figure S11B), in agreement with previous reports (Zhou et al., 2009; Yuan et al., 2015). We identified 1336 potential siRNA precursor *cis*-NATs, with at least five small reads mapping to the overlapping region and a read density at least two-fold higher in the overlapping region than that in the non-overlapping region. From this set of 1,336 *cis*-NATs, 25 belonged to the group of putative transcription enhancers (representing 23% of the 107 candidates), 10 to the group of putative transcription inhibitors (representing 24% of the 41 candidates), one to the group of putative translation enhancers (representing 12.5% of the 8 candidate) and none to the group of putative translation repressors (Supplemental Table S1).

We also looked for the presence of transposable elements or inverted repeats within the *cis*-NATs identified (see Material and Methods) (Supplemental Table S1). Approximately 10.5% of the *cis*-

NATs (i.e. 508) contained at least part of transposable element sequences, including one positively correlated and one negatively correlated to mRNA translation. Transposable element sequences were found in 15 out of the 107 *cis*-NATs correlated with cognate mRNA steady-state level. Similarly, we found that 121 *cis*-NATs contained inverted repeats, including one in a *cis*-NAT positively correlated with cognate mRNA translation and one in a *cis*-NATs negatively correlated with cognate mRNA steady-state level.

Experimental validation of *cis*-NAT regulation of cognate sense mRNA translation

A positive or negative correlation between *cis*-NAT expression and cognate mRNA association with polysomes could indicate translation enhancement or repression of the mRNA by the *cis*-NAT. To experimentally validate such a potential translation regulation activity, two *cis*-NATs with potential translation enhancer activity and two with putative translation repressor activity were cloned after the CaMV35S promoter and used to transform *A. thaliana* to produce transgenic lines over-expressing the *cis*-NATs in *trans*. Two independent transgenic lines were selected for each *cis*-NAT construct with a robust (>10-fold) over-expression of the transgenic *cis*-NAT compared to the steady-state level of the endogenous *cis*-NAT in wild-type lines but without a significant change of steady-state level of the endogenous cognate mRNA (Supplemental Figure S12). Polysome association of each cognate sense mRNA was analyzed by sucrose density gradient in the lines over-expressing the *cis*-NATs compared to that in a control line transformed with an empty vector. The distribution of mRNAs along the sucrose density gradient was determined by RT-qPCR (Figure 4B and E, 5B and E, and Supplemental Figure S13). In order to quantify the changes in terms of association with polysomes in a more robust manner, the proportion of mRNA present either in fractions containing free mRNA or monosomes (fractions 1 to 3) versus fractions containing polysomal mRNAs (fractions 4 to 6) was calculated for each of the eight independent biological replicates (Figure 4C and F, 5C and F). This analysis revealed that over-expression of the *cis*-NAT associated with the *CuAO1* locus was associated with a decrease in translation of the cognate mRNA (Figure 4B, C), in agreement with the negative correlation between *cis*-NAT steady-state level and mRNA polysome association (Figure 4A). In contrast, polysome association of the mRNA of locus At1g54260, encoding a potential transcription factor, was not significantly changed by the over-expression of its *cis*-NAT in *trans* (Figure 4E, F), despite the positive correlation (Figure 4D).

Similar analysis performed for two *cis*-NATs that displayed positive correlation with their cognate mRNA translation (Figure 5A, D) showed that lines over-expressing the *cis*-NAT to locus At3g26240 showed a reproducible shift of the its cognate mRNA towards the heavy polysome

Figure 4

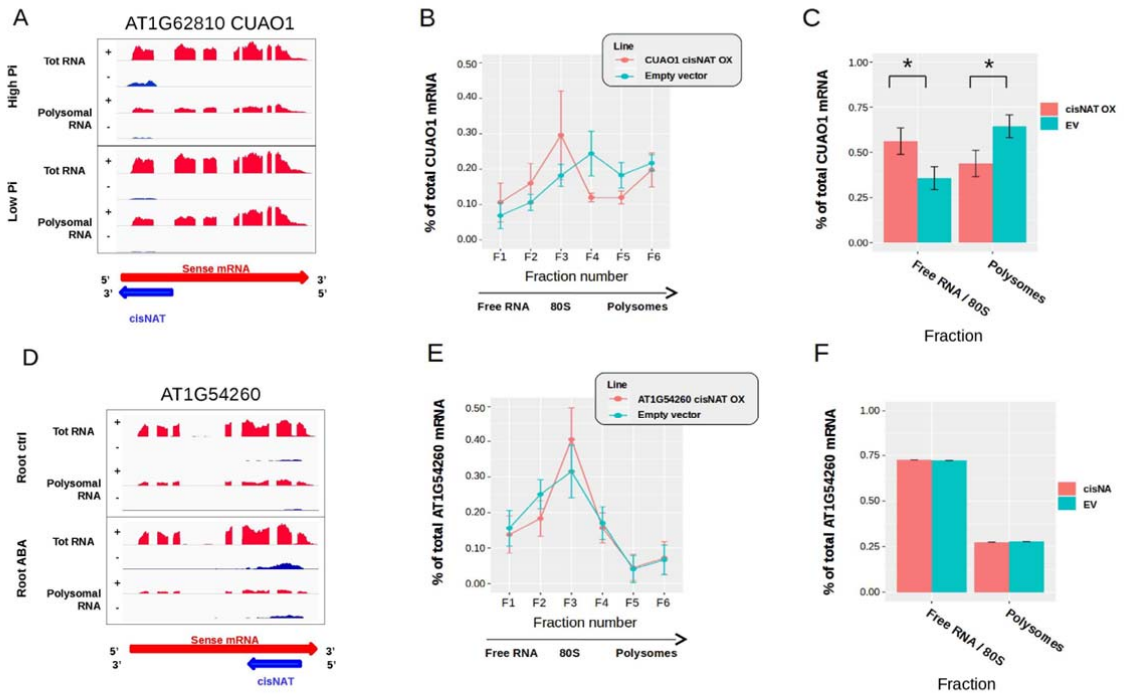

Figure 4: Expression of putative translation repressor *cis*-NATs in transgenic *A. thaliana*. **A and D:** Coverage plots showing the density of RNA-seq reads per position for the *CuAO1* (A) and AT1G54260 (D) loci, with the red and blue areas representing the sense mRNA and *cis*-NATs, respectively. **B and E:** Polysome profiles showing the proportion of endogenous mRNA in each of the 6 fractions of the sucrose gradient for transgenic lines over-expressing the *cis*-NAT (red) versus control lines transformed with an empty vector (turquoise) for the *CuAO1* (B) and AT1G54260 (E) sense mRNA-*cis*-NAT pair. **C and F:** Proportion of mRNA present in the first three fractions (free RNA and monosomes) and in the last three fractions (polysomes) of the gradient, in transgenic lines over-expressing *cis*-NAT (red) and in control lines transformed with an empty vector (turquoise) for the *CuAO1* (C) and AT1G54260 (F) sense mRNA-*cis*-NAT pair. Data in B, C, E and F represent the average between 8 independent biological replicates obtained with 2 independent transgenic lines. The error bars represents the confidence intervals with alpha=0.05. The significant differences (Student test with pvalue <0.05) are indicated by a star.

fractions in eight independent biological replicates, indicating a stimulatory effect of *cis*-NAT expression on translation (Figure 5B, C). In contrast, over-expression of the *cis*-NAT to locus *WRKY45* (AT3G01970) did not change significantly the polysome profile of its cognate sense mRNAs (Figure 5E, F).

A protoplast co-transformation system was developed to independently validate the effects of *cis*-NAT expression on the translation of their cognate sense mRNA. Protoplasts were transformed with a plasmid containing a sense-coding gene fused to NanoLuc luciferase (Nluc), in the presence or absence of a plasmid expressing the cognate *cis*-NAT. The sense-Nluc vectors contained also an independent expression cassette for the firefly luciferase (Fluc), used as an internal transformation and loading control (see Materials and Methods for further details). The ratio Nluc:Fluc activity was used to assess the effect of each selected *cis*-NAT on its sense-encoded protein accumulation. Increasing the molar amounts of *cis*-NAT_{CuAO1} resulted in a corresponding decrease in the expression of the CuAO1-Nluc fusion protein as detected by a decrease in the Nluc:Fluc ratio (Figure 6A). Importantly, this inhibitory effect was not observed at the transcript level (Supplemental Figure S14). Similarly, increasing amounts of *cis*-$NAT_{AT1G54260}$ resulted in reduction

Figure 5

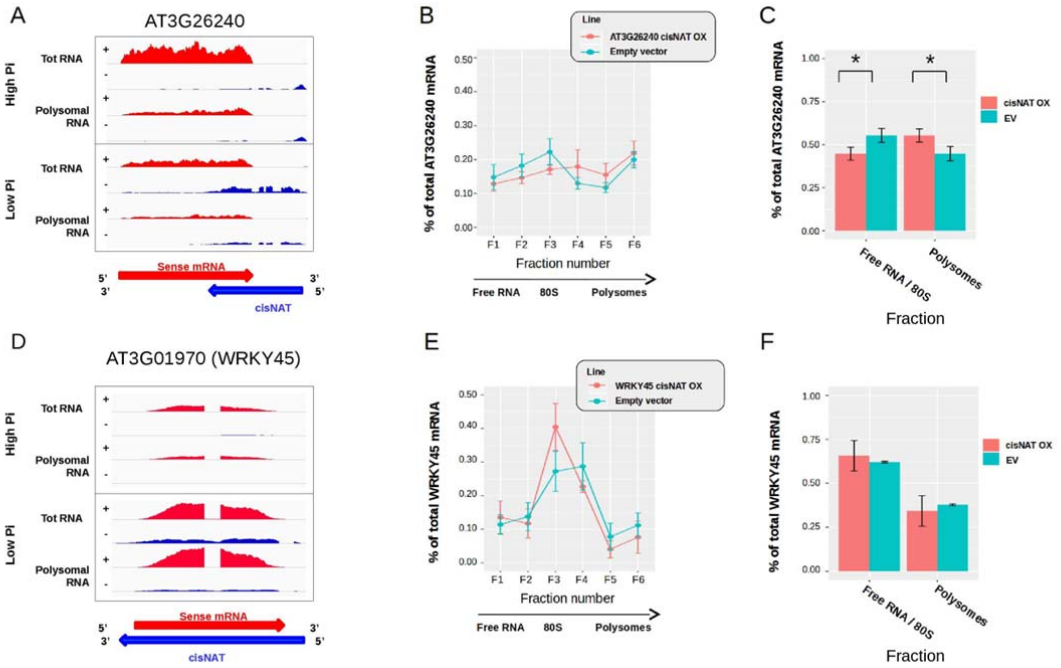

Figure 5: Expression of putative translation activator *cis*-NATs in transgenic *A. thaliana*. **A and D:** Coverage plots showing the density of RNA-seq reads per position for the AT3G26240 (A) and *WRKY45* (D) loci, with the red and blue areas representing the sense mRNA and *cis*-NATs, respectively. **B and E:** Polysome profiles showing the proportion of endogenous mRNA in each of the 6 fractions of the sucrose gradient for transgenic lines over-expressing the *cis*-NAT (red) versus control lines transformed with an empty vector (turquoise) for the AT3G26240 (B) and *WRKY45* (E) sense mRNA-*cis*-NAT pair. **C and F:** Proportion of mRNA present in the first three fractions (free RNA and monosomes) and in the last three fractions (polysomes) of the gradient, in transgenic lines over-expressing *cis*-NAT (red) and in control lines transformed with an empty vector (turquoise) for the AT3G26240 (C) and *WRKY45* (F) sense mRNA-*cis*-NAT pair. Data in B, C, E and F represent the average between 8 independent biological replicates obtained with 2 independent transgenic lines. The error bars represents the confidence intervals with alpha=0.05. The significant differences (Student test with pvalue <0.05) are indicated by a star.

in expression of the AT1G54260-Nluc fusion protein (Figure 6B) without an effect on transcript level (Supplemental Figure S14), although the effect was less pronounced than the one observed for the *cis*-NAT$_{CuAO1}$-sense construct pair. These later results are in contrast to the lack of significant effect observed in stable transgenic lines overexpressing *cis*-NAT$_{AT1G54260}$ and analyzed by quantification of polysomal mRNA in sucrose gradient fractions (Figure 4E and F). Such discrepancy may reflect the higher variability of the transgenic-polysomal approach or the higher sensitivity of the protoplast transformation method. To further validate the specificity of the effects of *cis*-NAT on translation of the corresponding sense mRNA, *cis*-NAT$_{CuAO1}$ was co-transformed with the AT1G54260-Nluc construct, and *cis*-NAT$_{AT1G54260}$ was co-transformed with the *CuAO1*-Nluc construct. In this *cis*-NAT swap experiment, no effect on Nluc:Fluc ratio were observed (Figure 6C and 6D), revealing that the inhibitory effect of *cis*-NAT$_{AT1G54260}$ and *cis*-NAT$_{CuAO1}$ on translation was specific to their cognate sense genes.

The protoplast system was also used to test the two *cis*-NATs acting as potential translational enhancers that were previously analyzed in transgenic plants, namely the *cis*-NATs to AT3G26240 and *WRKY45*, as well as an additional third candidate, *2A6* (AT1G03410) (Figure 7). There was a significant increase in Nluc:Fluc ratio upon addition of increasing amount of *cis*-NAT$_{AT3G26240}$ to its

Figure 6

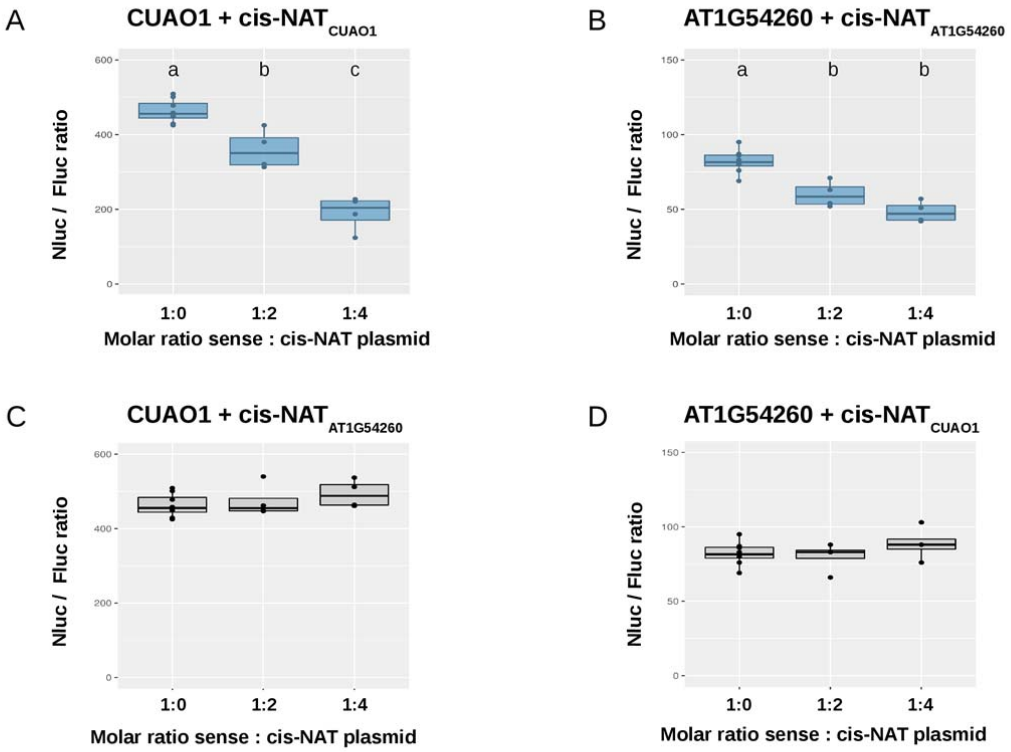

Figure 6. **Transient expression of putative translation repressor *cis*-NATs in protoplasts.** Arabidopsis leaf protoplasts were co-transformed with a plasmid combining a sense mRNA-NanoLuc luciferase (Nluc) fusion and a firefly luciferase (Fluc) along with various molar ratios of an independent plasmid for expression of a *cis*-NAT. The ratio of Nluc over Fluc activity is plotted for each combination of sense and *cis*-NAT plasmids. (A) Co-expression of *CuAO1*-Nluc fusion with its cognate *cis*-NAT. (B) Co-expression of AT1G54260-Nluc with its cognate *cis*-NAT. (C) Co-expression of *CuAO1*-Nluc with the *cis*-NAT to AT1G54260. (D) Co-expression of AT1G54260-Nluc with the *cis*-NAT to *CuAO1*. Statistical significant differences (t-test, p-value < 0.05; four biological replicates) between treatments are indicated distinct letters above the bars.

cognate sense construct, without a corresponding increase in transcript levels (Figure 7A and Supplemental Figure S14), confirming the translational enhancement of this *cis*-NAT. Similar to the results obtained with stable transgenic plants (Figure 5E, F), there was no significant effect of addition of *cis*-NAT$_{WRKY45}$ on the expression of the *WRKY45*-Nluc construct (Figure 7C). Although we did not generate transgenic lines to test the effect of *cis*-NAT$_{2A6}$ on the translation of its cognate sense mRNA, protoplast analysis revealed an increasing Nluc:Fluc ratio associated with the addition of *cis*NAT$_{2A6}$ to its corresponding sense 2A6-Nluc construct, without changes in mRNA levels (Figure 7B and Supplemental Figure S14), revealing a similar translational enhancement effect as those observed with *cis*-NAT$_{AT3G26240}$. A swap experiment performed between the *cis*-NAT$_{AT3G26240}$ and *cis*-NAT$_{2A6}$ showed no enhancement effect on unrelated sense-Nluc fusion (Figure 7D-E), confirming that the stimulatory effect of *cis*-NAT$_{AT3G26240}$ and *cis*-NAT$_{2A6}$ on translation was specific to their cognate sense genes.

Over-expression of the putative translation inhibitor *cis*-NAT$_{CuAO1}$ in transgenic lines probably resulted in lower levels of endogenous CuAO1 protein although we were not able to detect and quantify reliably CuAO1 protein by targeted mass spectrometry using N^{15}-labeled plants (Hart-

Figure 7

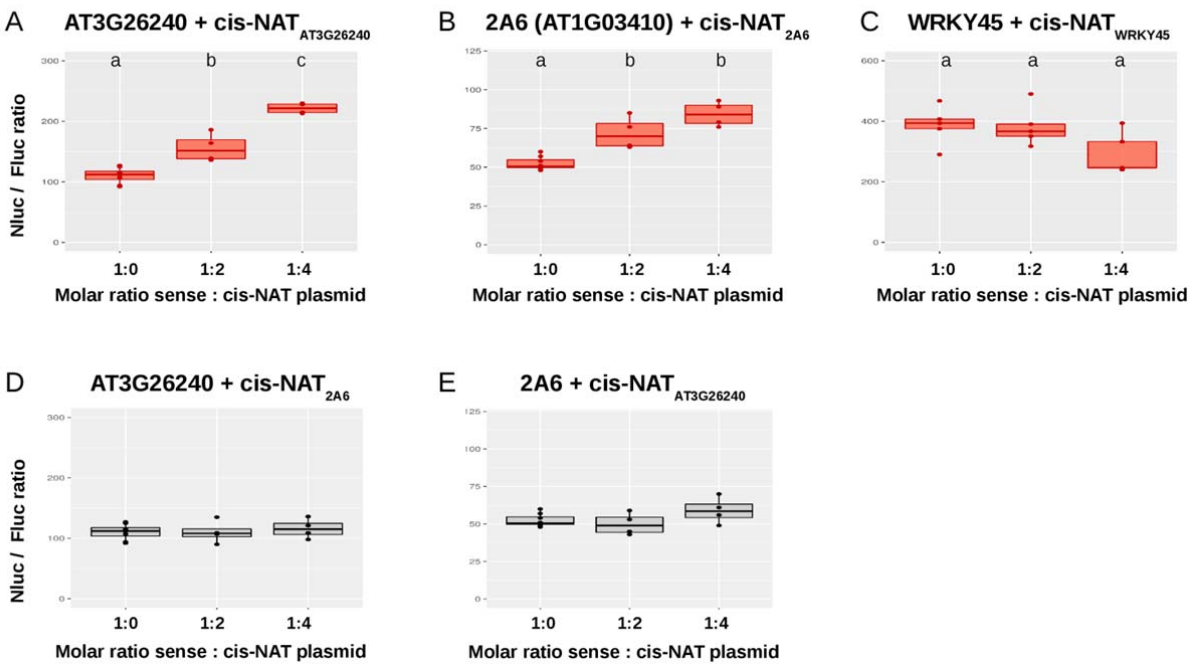

Figure 7. **Transient expression of putative translation activator *cis*-NATs in protoplasts.** Arabidopsis leaf protoplasts were co-transformed with a plasmid combining a sense mRNA-NanoLuc luciferase (Nluc) fusion and a firefly luciferase (Fluc) along with various molar ratios of an independent plasmid for expression of a *cis*-NAT. The ratio of Nluc over Fluc activity is plotted for each combination of sense and *cis*-NAT plasmids. (A) Co-expression of AT3G26240-Nluc fusion with its cognate *cis*-NAT. (B) Co-expression of *2A6* (AT1G03410)-Nluc with its cognate *cis*-NAT. (C) Co-expression of *WRKY45* (AT3G01970)-Nluc with its cognate *cis*-NAT. (D) Co-expression of AT3G26240-Nluc with the *cis*-NAT to *2A6*. (E) Co-expression of *2A6*-Nluc with the *cis*-NAT to AT3G26240. Statistical significant differences (t-test, p-value < 0.05; four biological replicates) between treatments are indicated distinct letters above the bars.

Smith et al., 2017). Since Arabidopsis CuAO1 knock-down mutants were shown to be impaired in NO production induced by polyamines (Wimalasekera et al., 2011), we undertook to quantify NO production upon spermidine treatment in two independent *cis*-NAT_{CuAO1} over-expressing transgenic lines along with a CuAO1 T-DNA knock-down mutant. NO production was strongly impaired in CuAO1 knock-down mutant, in agreement with the previous results of Wimalasekera et al. (2011), but also in the *cis*-NAT_{CuAO1} over-expressing line #1 (Supplemental Figure S15). The second *cis*-NAT over-expressing line also showed a reproducible reduction in NO production compared to that in the Col0 control although the associated p-value was above 0.05 (0.13).

468

21

Discussion

Out of a total set of 4,846 *cis*-NATs identified in this study or annotated in TAIR10, 157 (3.24%) were found to have a potential to regulate the expression of their cognate sense mRNA based on positive or negative correlations with either the steady-state mRNA level or polysome association. The great majority of those potential regulatory *cis*-NATs (147 out of 157; Table 1) were associated with changes in the transcript level of the cognate sense mRNA, with a stronger bias towards concordant expression (107 out of 147). This bias towards concordance is somewhat surprising since negative effect of *cis*-NAT expression on steady-state mRNA level is more commonly described in the literature than positive effects (Khorkova et al., 2014). It is possible that phenotypes associated with the disruption of *cis*-NATs with discordant expression pattern may be more apparent than for concordant expression pattern. Furthermore, co-expression of *cis*-NAT and cognate sense mRNA could, in many cases, be simply a consequence of local changes in chromatin state encompassing a whole locus that would equally affect the access of the transcription machinery to both the sense and antisense promoters, and thus would not be associated with a regulatory mechanism for controlling sense mRNA expression. However, numerous examples exist in the literature showing that increased expression of a *cis*-NAT may negatively affect sense mRNA steady-state level via changes in histone marks localized primarily at the promoters of the sense genes (Khorkova et al., 2014). Whereas fewer examples of similar local effect only on the promoter activity of the sense genes have been described for *cis*-NATs having concordant expression pattern (Mondal et al., 2010), more examples may be found through a more systematic analysis of this group of *cis*-NATs in Arabidopsis.

Non-coding RNAs, and particularly lincRNAs, can regulate the expression of coding mRNA by either masking a miRNA binding site via base pairing or by acting as a miRNA mimic (Wang et al., 2013; Cho and Paszkowski, 2017). At least one example has been described in animals for a *cis*-NAT masking a miRNA binding site present in the cognate sense gene (Faghihi et al., 2010). In the present study, 600 sense genes associated with a *cis*-NAT were found to be a potential target for a microRNA and 69 *cis*-NATs were found to contain a sequence that could act as a miRNA mimic (Supplemental Table S1). Of those, only two *cis*-NATs were found in the group associated with changes in steady-state mRNA level (both concordant) and only one associated with changes in mRNA polysome association, namely *cis*-NAT$_{WRKY45}$. However, the effect of *cis*-NAT$_{WRKY45}$ in its cognate sense RNA could not be experimentally validated either in transgenic plants nor in protoplasts (Figure 5D and 7C). Thus, whereas it is possible that some plant *cis*-NATs may function in miRNA masking or as a miRNA mimic, it would not appear to be common.

Overlap between *cis*-NATs and their cognate sense mRNAs can potentially generate siRNAs leading to gene silencing. There are several examples of *cis*-NATs down-regulating the cognate mRNA level via a siRNA-mediated silencing pathway, including in Arabidopsis (Held et al., 2008; Ron et al., 2010). It is thus possible that some of the *cis*-NATs identified in this study are potential transcription inhibitors and the associated *cis*-NAT-siRNAs may act through a siRNA pathway. There is also an example in Arabidopsis where *cis*-NAT expression leads to an increase in cognate sense mRNA transcript via the generation of *cis*-NAT-siRNAs that inhibit the action of a microRNA targeting the same cognate sense mRNA, thus leading to an increase in sense mRNA level (Gao et al., 2015). Such a mechanism could potentially apply to three genes having *cis*-NATs identified as potential transcription enhancers and associated with *cis*-NAT-siRNAs, namely At1g23090, At2g44430 and At2g45850, which could be targeted by miR826a, miR838 and miR837, respectively. In contrast, the only gene with a *cis*-NAT generating *cis*-NAT-siRNAs that belongs to the group of potential translation regulators does not harbor miRNA targets.

In silico analysis of the set of *cis*-NATs identified a small group of 63 *cis*-NATs that had a higher coding potential and that were more associated with polysomes than were other non-coding RNAs. Further analysis revealed that 19 of those *cis*-NATs could encode polypetides that were conserved either mainly in Brassicaceae (Group II, 9 *cis*-NATs) or more broadly in plants (Group I, 10 *cis*-NATs) (Supplemental Figure S6). Only one out of these 63 *cis*-NATs was positively correlated with change in sense mRNA steady-state level (*cis*-$NAT_{AT1G69260}$), but none were correlated with changes in mRNA translation (Supplemental Table S1). Association of mRNA with polysomes does not directly show if the RNA is being actively translated into a polypetide. However, analysis of RNA translation by ribosome footprint in both plants and animals have revealed that a remarkably broad spectrum of RNAs previously thought to be non-coding are actively being translated by ribosomes (Aspden et al., 2014; Ji et al., 2015; Hsu et al., 2016; Bazin et al., 2017). Recent analysis of non-coding RNAs in Arabidopsis roots revealed that 568 out of 1,676 *cis*-NATs had ribosome footprints consistent with translation (Bazin et al., 2017). Interestingly, translation of a small ORF present in a tasiRNA was shown to enhance tasiRNA production (Bazin et al., 2017), suggesting that whereas many lincRNAs and *cis*-NATs may indeed be translated into peptides, the act of translation itself rather than the specific sequence of the polypeptide may, in some cases, be the predominant mechanism of regulation. In that context, four *cis*-NATs found by Bazin et al. (2017) to have ribosome footprints are included in the group of 19 *cis*-NATs with coding potential that are well conserved in plants (Supplemental Figure S6). These four *cis*-NATs may be good candidates for transcripts coding for biologically active polypeptides.

In contrast to numerous reports of the effects of *cis*-NAT expression on sense gene transcription and/or transcript stability, very few examples of *cis*-NATs affecting the translation of their cognate sense mRNA have been described. Repression of mRNA translation by a *cis*-NAT has only been described for the *PU.1* gene in mouse (Ebralidze et al., 2008), whereas *cis*-NATs enhancing cognate mRNA translation have been reported for the rice *PHO1.2*, the mouse *Uchl1* and the human *RBM15* genes (Carrieri et al., 2012; Jabnoune et al., 2013; Tran et al., 2016). A major goal of this work has thus been to systematically explore the role of *cis*-NAT expression on translation of their cognate mRNA in Arabidopsis. A total of 14 candidate *cis*-NATs with putative repressive or stimulatory effects on cognate mRNA translation were found, which is 10-fold less compared to the number of *cis*-NATs with effects on mRNA steady-state level. Analysis of the configuration of the *cis*-NATs relative to the sense mRNA showed that *cis*-NATs associated with either translational stimulation or repression had a higher proportion of head-to-head configuration compared to that of all other *cis*-NATs (Supplemental Figure S10), although the low number of *cis*-NATs associated with translation makes this distinction not statistically significant. Furthermore, some of the Arabidopsis *cis*-NATs that were experimentally confirmed to affect translation have other configurations, such as tail-to-tail (AT1G54260, AT3G26240).

The effects of *cis*-NAT expression on cognate mRNA translation were experimentally tested by either stable transformation in plants and/or transient expression in leaf protoplasts for 5 of the 14 candidates. Four of those were validated, namely two *cis*-NATs mediating translational repression (*CuAO1* and AT1G54260) and two *cis*-NATs mediating translational stimulation (AT3G26240 and AT1G03410). These results highlight the robustness of the experimental pipeline used to identify the candidates.

CuOA1 encodes a copper amine oxidase involved in the catabolism of polyamines (Wimalasekera et al., 2011). *CuAO1* has been shown to be involved in the generation of nitric oxide, a key signaling molecule involved in a wide range of functions in plants, including seed germination, root development and ABA-induced stomatal closure (Besson-Bard et al., 2008). The Arabidopsis *cuao1-1* T-DNA knock-down mutant shows reduced production of NO after treatment with spermidine (Wimalasekera et al., 2011). Several stress conditions are known to induce NO synthesis, including phosphate deficiency (Sun et al., 2016). Although proteomic experiments could not reliably quantify the amount of CuAO1 protein in transgenic lines overexpressing *cis*-NAT_{CuAO1}, the same lines did show reduced NO production to levels similar to that of the *cuao1-1* mutant, supporting an inhibitory effect of *cis*-NAT_{CuAO1} expression on CuAO1 production (Supplemental Figure S15).

AT1g54260 harbors a highly conserved central globular domain (GH1) present in the linker histone H1, proteins that perform important functions on chromatin structure and influencing accessibility of trans-acting factors to DNA (Hergeth and Schneider, 2015; Kotlinski et al., 2017). The GH1 domain is known to bind DNA and the AT1G54260 protein belongs to the winged helix family of DNA binding proteins. Beside histones H1, proteins containing GH1 domain have been shown to binds to DNA, including at the telomeres, and potentially act at various level in the regulation of chromatin structure (Zhou et al., 2016). The *cis*-NAT to AT1G54260 is up-regulated by both ABA and low Pi (Supplemental Table S1). Modulation of AT1G54260 protein synthesis via expression of its *cis*-NAT could thus have broad impact on chromatin structure and gene regulation under various stress.

AT3G26240 encodes a protein of unknown function. AT1g03410 (*2A6*) encodes a protein containing a domain associated with oxoglutarate and iron-dependent dioxygenase. In plants, enzymes containing this domain catalyze the formation of plant hormones, such as ethylene, gibberellins, anthocyanidins and pigments such as flavones. The *cis*-$NAT_{AT1G03410}$ is of particular interest since it corresponds to a retroelement of the *Sadhu* family (Rangwala and Richards, 2010). The stimulatory activity of the *cis*-NAT of the mouse *Uchl1* on translation was shown to be dependent on the SINEB2 retroelement (Carrieri et al., 2012). *Sadhu* retroelements resemble SINEs in their structure, except that they do not contain similarity to known non-coding RNAs, such as 5SrRNA or tRNAs (the SINEB2 element is derived from a tRNA) (Weiner, 2002). Whereas the *cis*-NATs of both *Uchl1* and AT1g03410 are in the head-to-head configuration, the SINEB2 element of *cis*-NAT_{Uchl1} is located at the non-overlapping 3' end of the *cis*-NAT and *cis*-$NAT_{AT1G03410}$ is almost completely overlapping with the 5'UTR region of the sense mRNA except for the last 56 nucleotides (Supplemental Figure S16) (Carrieri et al., 2012). Whether or not SINEB2 elements and *Sadhu* retrotransposon stimulates mRNA translation by a similar mechanism remains to be determined.

The success of the validation methods relying on stable expression of *cis*-NATs in transgenic plants or transient expression in protoplast reveals that the effect of *cis*-NAT expression on sense mRNA translation can occur in *trans*. This implies that the *cis*-NAT produced from a distinct locus must be sufficiently stable to locate and anneal to its target mRNA and recruit or sequester factors that affect translation. This may, however, not always be the case, since the effects of some *cis*-NATs on mRNA transcript level have been found to occur only in *cis* and not in *trans* (Fedak et al., 2016; Rosa et al., 2016). Thus experimental validation of some *cis*-NATs for regulation of sense mRNA

610 translation may, in some cases, require other methods working in *cis*, such as precise mutation of
611 the *cis*-NAT locus by CRISPR/Cas9.
612
613 In conclusion, the experimental pipeline described in this work identified and validated a number of
614 novel *cis*-NATs in Arabidopsis that influence cognate sense mRNA translation. Although the
615 proportion of *cis*-NATs associated with changes in mRNA translation was relatively low compared
616 to the total number of *cis*-NATs expressed in the genome, it is likely that more candidates will be
617 found when plants are grown under different experimental conditions that lead to greater spectrum
618 of *cis*-NAT expression. Considering that a broad range of mechanisms have been identified for the
619 effect of lincRNAs and *cis*-NATs on transcriptional regulation, it is likely that the mechanisms
620 through which *cis*-NATs enhance or repress translation will also be quite diverse.
621

Materials and Methods

Plant materials

A. thaliana seeds (Col0) were germinated in half-strength Murashige and Skoog (MS) liquid medium containing 1 mM (high) or 100 uM (low) Pi. On day 5 and 6 after germination, the medium was replaced to maintain a constant level of Pi. On day 7, whole seedlings were harvested and used for total RNA extraction and polysome profiling. *A. thaliana* seeds were also germinated on agar-solidified half-strength MS medium for 10 days, after which the seedlings were flooded with a solution of half-strength MS containing 5 µM IAA, 10 µM ABA, 10 µM MeJA, 10 µM ACC, or no hormone for the untreated control. After 3 h of incubation, roots and shoots were split and harvested separately. For each of the 12 experimental conditions, 3 independent biological replicates were carried out at different times.

Total and polysomal RNA extraction

Plant samples (whole seedlings, roots or shoots) were flash frozen and ground in a mortar and pestle, and the polysomes were extracted essentially as described in Mustroph et al. (2009) with minor modifications (see Supplemental Materials and Methods).

Library preparation and RNA sequencing

From each total and polysomal RNA sample, strand specific libraries were prepared using the TruSeq Stranded Total RNA kit (Illumina) and polyA$^+$ RNAs were selected according to manufacturer's instructions. The libraries were sequenced on a HiSeq 2500 Illumina sequencer and about 30 million of paired-end reads per sample were obtained. In total, about 120 million reads were obtained for each of the 12 experimental conditions.

Identification of *cis*-NATs and analysis of their coding potential

To identify *cis*-NATs, the paired-end reads from the 3 replicates were pooled together and uniquely mapped to the TAIR10 genome using Hisat2 (Kim et al., 2015). For each of the 12 conditions, the transcriptome was determined *de novo* with Cufflinks (Trapnell et al., 2010), using the TAIR10.31 annotation as guide. The 12 annotation files obtained were merged using the Cuffmerge tool (Trapnell et al., 2010). This transcriptome was then compared to TAIR10.31 using Cuffcompare (Trapnell et al., 2010), and transcripts antisense to TAIR10.31 coding genes (class_code_x) were considered as putative *cis*-NATs. The readcount for each TAIR10.31 protein coding gene and each identified *cis*-NATs was determined using HTSeq-count (mode Union) and the identified *cis*-NATs with a ratio read count *cis*-NAT / coding gene < 0.01 were discarded as false positives likely due to imperfect strand specificity of the library preparation protocol (99.9%).

658

659 The “FEELnc codpot” module from FEELnc (version 0.01) (Wucher et al., 2017) was used to
660 identify *cis*-NATs that could potentially be coding for polypeptides (see Supplemental Material and
661 Methods).
662

663 **Characterization of *cis*-NATs**

664 Basic features such as length or GC content of transcripts, average steady-state levels or polysome
665 association were determined for each *cis*-NAT using custom functions written in Python. To
666 analyze the nucleotide conservation, PHASTcons scores where extracted from the 20 angiosperm
667 genome alignment as described by Hupalo et al. (2013). For each transcript, the average
668 PHASTcons score was calculated for exonic and intronic sequences. The presence of inverted
669 repeats was determined using the einverted program (EMBOSS;
670 http://emboss.bioinformatics.nl/cgi-bin/emboss/einverted) using default parameters. The presence
671 of miRNA binding sites within *cis*-NATs and coding transcripts was determined using
672 psRNATarget server (http://plantgrn.noble.org/psRNATarget/) with an expectation <= 3 and
673 unpaired energy (UPE) <= 25. Potential miRNA precursors were identified by comparing the
674 cDNA sequences of *cis*-NATs against a database of miRNA hairpins downloaded from miRBase
675 (http://www.mirbase.org).
676

677 The presence of potential miRNA target mimic sites was determined using custom python functions
678 following the rules edicted in Wu et al. (2013), namely: (i) perfect nucleotide pairing was required
679 at the second to eighth positions of miRNA sequence, (ii) bulges were only permitted at the 5' end
680 ninth to 12th positions of miRNA sequence, and (iii) should be composed of only three nucleotides.
681 No more than 3 mismatches or G/U pairs were allowed in pairing regions (not considering the
682 bulge).
683

684 Analysis of siRNAs that could be generated by *cis*-NATs was essentially performed according to
685 the method described by Yuan et al. (2015) using the Arabidopsis small RNA dataset available on
686 GEO. Briefly, the small reads 18–28 nucleotides long were mapped to the TAIR10 reference
687 genome using bowtie. For each *cis*-NAT locus, the length and density in small RNAs was
688 calculated for overlapping and non-overlapping regions by dividing the number of mapped small
689 reads by the length of the region using custom scripts and the python library pysam.
690

691 The presence of transposable elements within *cis*-NAT transcripts was determined by comparing
692 the *cis*-NATs sequences against a database containing all transposable elements annotated in
693 TAIR10 using Blastn with a cutoff of evalue=1e-12 and percent identify > 50.
694
695 **Quantification of TAIR10 and identified loci and identification of DEG**
696 For each experimental condition and biological replicate, the read count of TAIR10 as well as
697 identified loci was determined with HTSeq-count (mode Union) (Anders et al., 2015), and
698 normalized with DESeq2 (Love et al., 2014). A gene was considered differentially expressed
699 comparing two conditions if the adj.pval was < 0.1 and the fold change > 2 or < 0.5.
700
701 **Validation for DEG by RT-qPCR**
702 *A. thaliana* seedlings were grown in liquid cultures in the presence of a high or low concentration of
703 inorganic phosphate as described above in the "Plant materials" section. Total RNA was extracted
704 from whole seedlings with Trizol following manufacturer's instructions. One microgram of RNA
705 was then used for reverse-transcription using the M-MLV Reverse Transcriptase (Promega) and
706 oligo $d(T)_{15}$ as primer using manufacturer's instructions. RT-qPCR analysis to measure mRNA
707 steady-state level was completed using SYBR select Master Mix (Applied Biosystems) with a
708 primer set specific of the gene of interest as well as a primers specific of *ACT2* gene used as
709 reference. Log2 fold changes were calculated by the ΔΔCt method.
710
711 **Determination of polysome association (PA) ratio**
712 To estimate the translation efficiency for each gene, the polysomes association (AP) ratio was
713 determined using Xtail package (Xiao et al., 2016), which calculates the ratio between read count
714 from polysomal RNA sample and total RNA sample. Genes with a Xtail adj.pval < 0.1 and at least a
715 30% increase or decrease of the AP ratio were considered differentially associated with polysomes.
716
717 **Identification *cis*-NATs influencing steady-state level or polysome association of cognate sense**
718 **mRNA**
719 The candidate regulatory *cis*-NATs were identified by pairwise comparisons between whole
720 seedlings grown under high- or low-Pi conditions, roots or shoots treated with phytohormones and
721 appropriate untreated controls, as well as between untreated root and shoot tissues, using a series of
722 criteria. Only the pairs coding gene / *cis*-NAT overlapping by at least 50 nucleotides and with a
723 normalized read count for both coding gene and *cis*-NAT > 20 were considered. A *cis*-NAT was
724 considered positively correlated to its cognate coding mRNA expression if both *cis*-NAT and
725 coding mRNA were either up-regulated or down-regulated (fold change > 2 and adj.pval < 0.1)

between the two conditions compared. It was considered negatively correlated if one partner was up-regulated whereas the other was down-regulated (fold change > 2 and adj.pval < 0.1) between the two conditions compared. To identify the putative translation regulatory *cis*-NATs, only the pairs for which the coding gene was differentially translated with fold change > 1.3 and adj.pval < 0.1 between the two conditions compared, and with fold-change of mRNA steady-state level < 3 were kept. From these pairs, the *cis*-NATs had to be differentially expressed, with fold change > 2 and adj.pval < 0.1 and the ratio readcount *cis*-NAT / readcount coding gene had to be above 0.2, in at least one condition. The *cis*-NATs up-regulated when their cognate mRNA was more associated with polysomes were considered as putative translation enhancers, whereas *cis*-NATs up-regulated when their cognate mRNA was less associated with polysomes were considered as putative translation repressors.

Pearson correlation coefficient between mRNA and *cis*-NAT steady-state level was also calculated across the 12 experimental conditions analyzed for each candidate pair with a positive or negative correlation between *cis*-NAT and mRNA expression. Similarly, the correlation between PA ratio, and *cis*-NAT steady-state level was also calculated across the 12 experimental conditions for each translation regulator *cis*-NAT candidate. The candidate pairs with a correlation factor > 0.4 or < -0.4 were considered as the most robust candidates.

Creation and analysis of transgenic lines over-expressing putative translation regulatory *cis*-NATs

To create transgenic plants over-expressing the candidate translation regulator *cis*-NATs, the genomic sequence encompassing the transcribed region was cloned into the vector pFAST-R02 (Shimada et al., 2010), in the correct orientation, to allow synthesis of the *cis*-NAT transcript under the control of the cauliflower mosaic virus 35S promoter. The constructs were introduced into *A. thaliana* by *Agrobacterium tumefaciens*-mediated transformation using floral dipping (Clough and Bent, 1998). Transgenic lines over-expressing the different *cis*-NAT constructs or transformed with empty vector were grown for 10 days on agar-solidified half-strength MS medium containing Basta as a selection marker. Whole seedlings were crushed in liquid nitrogen and total RNA was extracted using standard procedure.

To purify polysomes, 10-day-old seedlings were ground into powder in liquid nitrogen and 2 volumes of polysome extraction buffer were added. The mixture was incubated for 15 minutes on ice, centrifuged at 16,000 *g* to pellet debris and 200 µL of supernatant were loaded on top of 5 mL sucrose gradients. After 75 min of centrifugation at 55,000 rpm in a SW55 rotor (Beckman) at 4°C,

761 the gradients were collected and split into 6 fractions. For each fraction, 500 µL was transferred into
762 a new eppendorf tube and RNA was extracted with 1 mL of Trizol according to manufacturer's
763 instructions. An additional step of acetate ammonium precipitation with ethanol and washing was
764 added to remove remaining salt and phenol traces. For each sample, 300 ng of RNA was then used
765 for reverse transcription using the M-MLV Reverse Transcriptase (Promega) and oligo $d(T)_{15}$ as
766 primer using manufacturer's instructions. To analyze *WRKY45* mRNA, due to the full overlap
767 between *cis*-NAT and mRNA, the reverse transcription was performed in the same conditions but
768 using a mix of reverse primers specific to *WRKY45* and *ACT2* mRNA instead of oligo $d(T)_{15}$. RT-
769 qPCR analysis to quantify the relative amount of endogenous mRNA in each fraction was
770 performed with a primer set specific for the gene of interest as well as a primer specific for the
771 *ACT2* gene used as reference. The results are presented as relative proportion of endogenous mRNA
772 in each fraction of the gradient, as described in Faye et al. (2014). The average of 8 independent
773 biological replicates obtained with 2 independent transgenic lines is reported. To be able to quantify
774 in a more robust manner the changes in terms of polysome association, the sum of the proportions
775 of mRNA in fractions 1–3 and fractions 4–6 were calculated to compare the proportion of mRNA
776 not or poorly translated, e.g. free mRNA (fraction 1 and 2) or associated with monosomes (fraction
777 3), versus the proportion of mRNA efficiently translated, e.g. associated with low (fraction 4) or
778 high polysomes (fractions 5–6).
779
780 **Transient translation assays in Arabidopsis protoplasts**
781 Plasmids used for protoplast transformation were assembled using *Bsa*I-based Golden Gate cloning,
782 and the final constructs contained a recombination site for Gateway™ cloning. A Gateway™
783 destination vector, for cloning and expression of sense-coding genes, included a C-terminal in-
784 frame fusion with a foot-and-mouth disease virus (FMDV) 2A peptide, followed by fusion with
785 NanoLuc® luciferase (Nluc) (plasmid nLucFlucGW, GenBank MH552885). Additionally, an
786 independent expression cassette driving firefly luciferase (Fluc) was also included in this vector.
787 Another Gateway™ destination vector, for cloning and expression of antisense noncoding genes,
788 was produced without any fusion or additional expression cassette (plasmid RHIP1pGW, GenBank
789 MH552886). Both Gateway™ destination vectors expressed the cloned gene, sense or antisense,
790 under control of the same promoter (1.1 kbp genomic sequence upstream of AT4G26410) and
791 terminator (250 bp downstream of AT5G59720). Genomic sequences for sense-coding genes (from
792 5'UTR to last codon, without STOP) and antisense-noncoding genes were cloned via Gateway™
793 cloning into their corresponding vector.
794

31

795 Protoplasts were produced and transformed essentially as described by Yoo et al. (2007) with minor
796 modifications (see Supplemental Material and Methods). Protoplasts were harvested by
797 centrifugation at 6,000 *g* for 1 min, and resuspended in 1X Passive Lysis Buffer (Promega, E1941).
798 The lysate was cleared by centrifugation and used for luminescence quantification using Nano-
799 Glo® Dual-Luciferase® Reporter Assay System (Promega, N1610), according to the manufacture's
800 instructions. Luminescence values for Nluc fused to sense-coding gene were normalized against
801 Fluc to control for loading and transfection efficiency. Statistical significant differences (Student's
802 *t*-test, p-value < 0.05) in ratio Nluc:Fluc were used to assess the effect of antisense noncoding gene
803 co-expression.
804

805 **Quantification of NO production**
806 NO production was quantified in 10-day-old seedlings treated with spermidine following the
807 procedure described in Wimalakasera et al. (2011). Briefly, 5–6 seedlings were equilibrated in 3 mL
808 of MES buffer (30 mM MES, 0.1 mM $CaCl_2$, 1 mM KCl) for 2 h. Then 4,5-diaminofluorescein
809 diacetate and spermidine or DMSO was added to the medium. After 30 min incubation at 24°C
810 under light with shaking, 100 µL of medium was transferred to 96 well plate and fluorescence was
811 quantified. Eight independent biological replicates were analyzed and the florescence was
812 normalized by mg of fresh weight of spermidine-treated seedlings or untreated control.
813

814 **Accession numbers**
815 The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO)
816 database, https//www.ncbi.nlm.nih.gov/geo (accession no. GSE116553). The processed data tables
817 (Supplemental Table S1 and S4) are included as additional files for this article. The sequence of
818 created plasmids used in this study has been submitted to GenBank, MH552885 and MH552886.
819

820 **Supplemental Data**
821 The following supplemental materials are available.
822 **Supplemental Figure S1.** Steady state mRNA expression level and association with polysomes in
823 response to treatment with phytohormones.
824 **Supplemental Figure S2.** GO terms enriched in the set of genes up-regulated in plants grown in
825 low Pi conditions or treated with various phytohormones.
826 **Supplemental Figure S3.** Analysis of the degree of overlap between independent studies analyzing
827 gene expression in response of low Pi or ABA treatment.

Supplemental Figure S4. GO terms enriched in the set of genes with changes in polysome association.

Supplemental Figure S5. Steady state mRNA expression level and association with polysomes in untreated roots compared to shoots.

Supplemental Figure S6. Evolutionary conservation of *cis*-NAT encoded peptides.

Supplemental Figure S7. Analysis of histone acetylation and nucleosome occupancy near the transcription start site of *cis*-NATs.

Supplemental Figure S8. Analysis of the degree of overlap in *cis*-NATs identified in distinct studies.

Supplemental Figure S9. Validation of differentially expressed genes by RT-qPCR.

Supplemental Figure S10. Proportion of the different types of orientation for the *cis*-NAT – sense mRNA pairs.

Supplemental Figure S11. Analysis of *cis*-NAT-siRNAs.

Supplemental Figure S12. Quantification of the endogenous cognate mRNA in *cis*-NAT overexpressing lines.

Supplemental Figure S13. Polysome profile.

Supplemental Figure S14. Levels of sense mRNA-NanoLuc luciferase (Nluc) fusion transcripts in transiently transformed protoplasts.

Supplemental Figure S15. Quantification in NO production upon spermidine treatment.

Supplemental Figure S16. Organization of the *cis*-NAT:mRNA pair at AT1G03410 locus.

Supplemental Table S1. Summary of features associated with each transcript

Supplemental Table S2. Genes differentially expressed in various conditions

Supplemental Table S3. Number of mRNAs differentially associated with polysomes.

Supplemental Table S4. RNAseq and polysome profiling data relative to putative transcription or translation regulatory cis-NATs.

Supplemental Materials and Methods

Acknowledgements

We thank the University of Lausanne Genomic Technology Facility for high-throughput sequencing support, the members of the Sinergia consortium, Jacques Rougemont (EPFL) and Bulak Arpat (SIB) for useful discussions, as well as Aidan Tay for help with statistical analysis. No conflict of interest declared.

Tables

Table 1. Number of *cis*-NATs correlated with cognate gene steady-state mRNA expression or association with polysomes. Number of mRNA / *cis*-NAT pairs with either a positive or negative correlation between *cis*-NAT and cognate gene steady-state mRNA expression (second and third columns), and number of pairs with positive or negative correlation between *cis*-NAT expression and cognate gene mRNA polysome association (PA) (fourth and fifth columns). The experimental conditions compared are indicated in the first column where root and shoot tissues are indicated with the prefix R and S, respectively, and untreated control conditions indicated with the suffix ctrl. The figures in brackets show the number of those pairs with a Pearson correlation coefficient > 0.4 or < -0.4 across the 12 experimental correlations.

Treatment	Positive correlation *cis*-NAT / mRNA expression	Negative correlation *cis*-NAT / mRNA expression	Positive correlation *cis*-NAT expression / mRNA PA	Negative correlation *cis*-NAT expression / mRNA PA
Low / high Pi	40 (33)	10 (3)	4 (1)	1 (1)
RIAA / Rctrl	2 (2)	0	0	0
RABA / Rctrl	17 (13)	0	1 (0)	1 (0)
RMeJA / Rctrl	6 (4)	0	0	0
SABA / Sctrl	10 (7)	1 (0)	0	0
SMeJA / Sctrl	3 (3)	1 (0)	0	0
Rctrl / Sctrl	47 (41)	29 (24)	3 (2)	4 (2)
Total (unique)	**107 (86)**	**41 (27)**	**8 (3)**	**6 (3)**

Figure legends

Figure 1. Steady state mRNA expression level and association with polysomes in response to growth of Arabidopsis under low-Pi conditions. A, Relation between log2-fold change of mRNA steady-state level (x axis) is plotted against the log2-fold change in polysome association (y axis). Coding genes significantly up- or down-regulated at the mRNA steady-state level are colored in yellow and cyan, respectively, whereas mRNA significantly more or less associated with polysomes

are colored in red and blue, respectively. The genes not showing any statistical difference are colored in grey. B, Same plot as A where genes associated with GO terms "Response to Pi starvation", "Cytosolic ribosome", "Mitochondrial ribosome", and "Chloroplastic ribosome" are colored in pink, dark blue, light blue and green, respectively. C to E, Normalized RNA-seq coverage plots for the *IPS1*, *RPS15AE* and *RPL34* genes. The two upper panels show the coverage plots for total mRNA and polysomal RNA from high Pi samples and the two lower panels correspond to low phosphate samples. The schematic exonic organization of each gene is represented by red boxes and lines below the plots.

Figure 2: Identification and characterization of *cis*-NATs. A, Schematic diagram of the pipeline used for *de novo cis*-NAT identification from the 12 different experimental conditions. B, Boxplot comparing polysome association of *cis*-NATs predicted to be noncoding (green) or coding (pink), ncRNA (cyan) and protein-coding genes (salmon) annotated in TAIR10 database. C and D, Plots comparing transcript length (C) and RNA steady-state-level (D) of *cis*-NATs predicted to be noncoding (green) or coding (pink), ncRNA (cyan) and protein-coding genes (salmon) annotated in TAIR10 database. E, Boxplots comparing the nucleotide conservation across 20 angiosperm genomes within exonic and intronic regions of the four categories of transcripts listed above.

Figure 3: Correlations between expression of *cis*-NATs and changes in steady-state level or poylsome association of the cognate sense mRNA. A to D, Coverage plots showing the density of RNA-seq reads per position at AT2G37580, AT1G68940, AT1G03410 and *CuAO1* loci. The red and blue areas represent the density of reads mapping to the sense mRNA and *cis*-NATs, respectively. For each experimental condition, the upper part corresponds to total RNA-seq reads and the lower part to polysomal RNA-seq reads. The red and blue arrows below indicate the *cis*-NAT-mRNA pair orientation. E to H, Correlation plots showing the steady-state level of the coding mRNA (red dots) and *cis*-NAT (green dots) for AT2G37580 and AT1G68940 loci (E and F, respectively) or the steady-state level of the *cis*-NAT (cyan dots) and the association with polysomes of the cognate sense mRNA (purple dots) for AT1G03410 and *CuAO1* loci (G and H, respectively). The Z-score of normalized read counts calculated from the 12 experimental conditions is represented on the y-axis. Pearson correlation coefficients between the two variables shown in each plot are indicated on top of the plots.

Figure 4: Expression of putative translation repressor *cis*-NATs in transgenic *A. thaliana.* A and D, Coverage plots showing the density of RNA-seq reads per position for the *CuAO1* (A) and AT1G54260 (D) loci, with the red and blue areas representing the sense mRNA and *cis*-NATs, respectively. B and E, Polysome profiles showing the proportion of endogenous mRNA in each of

the six fractions of the sucrose gradient for transgenic lines over-expressing the *cis*-NAT (red) versus that in control lines transformed with an empty vector (turquoise) for the *CuAO1* (B) and AT1G54260 (E) sense mRNA-*cis*-NAT pair. C and F, Proportion of mRNA present in the first three fractions (free RNA and monosomes) and in the last three fractions (polysomes) of the gradient. Determinations were in transgenic lines over-expressing *cis*-NAT (red) and in control lines transformed with an empty vector (turquoise) for the *CuAO1* (C) and AT1G54260 (F) sense mRNA-*cis*-NAT pair. Data in B, C, E and F represent the average of 8 independent biological replicates obtained with 2 independent transgenic lines. The error bars represent the confidence intervals with alpha=0.05. Asterisks indicate significant differences (Student's *t*-test with p-value <0.05).

Figure 5: Expression of putative translation activator *cis*-NATs in transgenic *A. thaliana.* A and D, Coverage plots showing the density of RNA-seq reads per position for the AT3G26240 (A) and *WRKY45* (D) loci, with the red and blue areas representing the sense mRNA and *cis*-NATs, respectively. B and E, Polysome profiles showing the proportion of endogenous mRNA in each of the six fractions of the sucrose gradient for transgenic lines over-expressing the *cis*-NAT (red) versus that in control lines transformed with an empty vector (turquoise) for the AT3G26240 (B) and *WRKY45* (E) sense mRNA-*cis*-NAT pair. C and F, Proportion of mRNA present in the first three fractions (free RNA and monosomes) and in the last three fractions (polysomes) of the gradient, in transgenic lines over-expressing *cis*-NAT (red) and in control lines transformed with an empty vector (turquoise) for the AT3G26240 (C) and *WRKY45* (F) sense mRNA-*cis*-NAT pair. Data in B, C, E and F represent the average of 8 independent biological replicates obtained with 2 independent transgenic lines. The error bars represent the confidence intervals with alpha=0.05. Asterisks indicate significant differences (Student's *t*-test with p-value <0.05).

Figure 6. **Transient expression of putative translation repressor *cis*-NATs in protoplasts.** Arabidopsis leaf protoplasts were co-transformed with a plasmid combining a sense mRNA-NanoLuc luciferase (Nluc) fusion and a firefly luciferase (Fluc) along with various molar ratios of an independent plasmid for expression of a *cis*-NAT. The ratio of Nluc over Fluc activity is plotted for each combination of sense and *cis*-NAT plasmids. A, Co-expression of *CuAO1*-Nluc fusion with its cognate *cis*-NAT. B, Co-expression of AT1G54260-Nluc with its cognate *cis*-NAT. C, Co-expression of *CuAO1*-Nluc with the *cis*-NAT to AT1G54260. D, Co-expression of AT1G54260-Nluc with the *cis*-NAT to *CuAO1*. Statistically significant differences (Student's *t*-test, p-value < 0.05; four biological replicates) between treatments are indicated by distinct letters above the boxes.

Figure 7. **Transient expression of putative translation activator *cis*-NATs in protoplasts.** Arabidopsis leaf protoplasts were co-transformed with a plasmid combining a sense mRNA-NanoLuc luciferase (Nluc) fusion and a firefly luciferase (Fluc) along with various molar ratios of an independent plasmid for expression of a *cis*-NAT. The ratio of Nluc over Fluc activity is plotted for each combination of sense and *cis*-NAT plasmids. A, Co-expression of AT3G26240-Nluc fusion with its cognate *cis*-NAT. B, Co-expression of *2A6* (AT1G03410)-Nluc with its cognate *cis*-NAT. C, Co-expression of *WRKY45* (AT3G01970)-Nluc with its cognate *cis*-NAT. D, Co-expression of AT3G26240-Nluc with the *cis*-NAT to *2A6*. E, Co-expression of *2A6*-Nluc with the *cis*-NAT to AT3G26240. Statistically significant differences (Student's *t*-test, p-value < 0.05; four biological replicates) between treatments are indicated by distinct letters above the boxes.

37

966

967

38

Parsed Citations

Anders, S., Pyl, P. and Huber, W. (2015) HTSeq – A Python framework to work with high-throughput sequencing data. Bioinformatics 31: 166-169.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Aspden, J.L., Eyre-Walker, Y.C., Philips, R.J., Amin, U., Mumtaz, M.A.S., Brocard, M. and Couso, J.P. (2014) Extensive translation of small Open Reading Frames revealed by Poly-Ribo-Seq. Elife 3: e03528.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Bardou, F., Ariel, F., Simpson, C.G., Romero-Barrios, N., Laporte, P., Balzergue, S., Brown, J.W.S. and Crespi, M. (2014) Long noncoding RNA modulates alternative splicing regulators in Arabidopsis. Dev. Cell 30: 166-176.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Bazin, J., Baerenfaller, K., Gosai, S.J., Gregory, B.D., Crespi, M. and Bailey-Serres, J. (2017) Global analysis of ribosome-associated noncoding RNAs unveils new modes of translational regulation. Proc. Natl. Acad. Sci. USA 114: E10018-E10027.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Berardini, T.Z., Reiser, L., Li, D.H., Mezheritsky, Y., Muller, R., Strait, E. and Huala, E. (2015) The Arabidopsis information resource: Making and mining the "gold standard" annotated reference plant genome. Genesis 53: 474-485.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Besson-Bard, A., Pugin, A. and Wendehenne, D. (2008) New insights into nitric oxide signaling in plants. Annu. Rev. Plant Biol. 59: 21-39.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Bonasio, R. and Shiekhattar, R. (2014) Regulation of transcription by long noncoding RNAs. Ann Rev Genet 48: 433-455.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Brockdorff, N., Ashworth, A., Kay, G.F., McCabe, V.M., Norris, D.P., Cooper, P.J., Swift, S. and Rastan, S. (1992) The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. Cell 71: 515-526.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Carrieri, C., Cimatti, L., Biagioli, M., Beugnet, A., Zucchelli, S., Fedele, S., Pesce, E., Ferrer, I., Collavin, L., Santoro, C., et al. (2012) Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. Nature 491: 454-457.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Chekanova, J.A. (2015) Long non-coding RNAs and their functions in plants. Curr. Opin. Plant Biol. 27: 207-216.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Cheng, C.Y., Krishnakumar, V., Chan, A.P., Thibaud-Nissen, F., Schobel, S. and Town, C.D. (2017) Araport11: a complete reannotation of the Arabidopsis thaliana reference genome. Plant J. 89: 789-804.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Cho, J. and Paszkowski, J. (2017) Regulation of rice root development by a retrotransposon acting as a microRNA sponge. Elife 6: e30038.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Clough, S.J. and Bent, A.F. (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J. 6: 735-743.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Djebali, S., Davis, C.A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., et al. (2012) Landscape of transcription in human cells. Nature 489: 101-108.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Ebralidze, A.K., Guibal, F.C., Steidl, U., Zhang, P., Lee, S., Bartholdy, B., Jorda, M.A., Petkova, V., Rosenbauer, F., Huang, G., et al. (2008) PU.1 expression is modulated by the balance of functional sense and antisense RNAs regulated by a shared cis-regulatory element. Genes Dev. 22: 2085-2092.
Pubmed: Author and Title

Google Scholar: Author Only Title Only Author and Title

Faghihi, M.A. and Wahlestedt, C. (2009) Regulatory roles of natural antisense transcripts. Nat. Rev. Mol. Cell Biol. 10: 637-643.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Faghihi, M.A., Zhang, M., Huang, J., Modarresi, F., Van der Brug, M.P., Nalls, M.A., Cookson, M.R., St-Laurent, G., III and Wahlestedt, C. (2010) Evidence for natural antisense transcript-mediated inhibition of microRNA function. Genome Biol. 11: R56.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Faye, M.D., Graber, T.E. and Holcik, M. (2014) Assessment of selective mRNA translation in mammalian cells by polysome profiling. J Vis Exp 28: e52295.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Fedak, H., Palusinska, M., Krzyczmonik, K., Brzezniak, L., Yatusevich, R., Pietras, Z., Kaczanowski, S. and Swiezewski, S. (2016) Control of seed dormancy in Arabidopsis by a cis-acting noncoding antisense transcript. Proc. Natl. Acad. Sci. USA 113: E7846-E7855.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Franco-Zorilla, 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.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Gao, W., Liu, W.W., Zhao, M. and Li, W.X. (2015) NERF encodes a RING E3 ligase important for drought resistance and enhances the expression of its antisense gene NFYA5 in Arabidopsis. Nucleic Acids Res. 43: 607-617.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Gong, C.G. and Maquat, L.E. (2011) lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3 ' UTRs via Alu elements. Nature 470: 284-288.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Gupta, R.A., Shah, N., Wang, K.C., Kim, J., Horlings, H.M., Wong, D.J., Tsai, M.C., Hung, T., Argani, P., Rinn, J.L., et al. (2010) Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature 464: 1071-1076.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Hart-Smith, G., Reis, R.S., Waterhouse, P.M. and Wilkins, M.R. (2017) Improved Quantitative Plant Proteomics via the Combination of Targeted and Untargeted Data Acquisition. Front Plant Sci 8:
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Held, M.A., Penning, B., Brandt, A.S., Kessans, S.A., Yong, W.D., Scofield, S.R. and Carpita, N.C. (2008) Small-interfering RNAs from natural antisense transcripts derived from a cellulose synthase gene modulate cell wall biosynthesis in barley. Proc. Natl. Acad. Sci. USA 105: 20534-20539.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Heo, J.B. and Sung, S. (2011) Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. Science 331: 76-79.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Hergeth, S.P. and Schneider, R. (2015) The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle. EMBO Rep. 16: 1439-1453.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Hsu, P.Y., Calviello, L., Wu, H.-Y.L., Li, F.-W., Rothfels, C.J., Ohler, U. and Benfey, P.N. (2016) Super-resolution ribosome profiling reveals unannotated translation events in Arabidopsis. Proc. Natl. Acad. Sci. USA 113: E7126-E7135.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Hu, G.Z., Lou, Z.K. and Gupta, M. (2014) The long non-coding RNA GAS5 cooperates with the eukaryotic translation initiation factor 4E to regulate c-Myc translation. Plos One 9: e107016.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Hupalo, D. and Kern, A.D. (2013) Conservation and functional element discovery in 20 angiosperm plant genomes. Mol. Biol. Evol. 30: 1729-1744.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Jabnoune, M., Secco, D., Lecampion, C., Robaglia, C., Shu, Q. and Poirier, Y. (2013) A rice cis-natural antisense RNA acts as a translational enhancer for its cognate mRNA and contributes to phosphate homeostasis and plant fitness. Plant Cell 25: 4166-4182.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Jegu, T., Veluchamy, A., Ramirez-Prado, J.S., Rizzi-Paillet, C., Perez, M., Lhomme, A., Latrasse, D., Coleno, E., Vicaire, S., Legras, S., et al. (2017) The Arabidopsis SWI/SNF protein BAF60 mediates seedling growth control by modulating DNA accessibility. Genome Biol. 18:
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Ji, Z., Song, R., Regev, A. and Struhl, K. (2015) Many lncRNAs, 5 ' UTRs, and pseudogenes are translated and some are likely to express functional proteins. Elife 4: e08890.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Jin, J., Liu, J., Wang, H., Wong, L. and Chua, N.-H. (2013) PLncDB: plant long non-coding RNA database. Bioinformatics 29: 1068-1071.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Juntawong, P., Girke, T., Bazin, J. and Bailey-Serres, J. (2014) Translational dynamics revealed by genome-wide profiling of ribosome footprints in Arabidopsis. Proc. Natl. Acad. Sci. USA 111: E203-E212.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Khorkova, O., Myers, A.J., Hsiao, J. and Wahlestedt, C. (2014) Natural antisense transcripts. Hum. Mol. Genet. 23: R54-R63.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Kim, D., Landmead, B. and Salzberg, S.L. (2015) HISAT: a fast spliced aligner with low memory requirements. Nat. Methods 12: 357-U121.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Kino, T., Hurt, D.E., Ichijo, T., Nader, N. and Chrousos, G.P. (2010) Noncoding RNA Gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. Science Signal 3: ra8.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Kotlinski, M., Knizewski, L., Muszewska, A., Rutowicz, K., Lirski, M., Schmidt, A., Baroux, C., Ginalski, K. and Jerzmanowski, A. (2017) Phylogeny-based systematization of Arabidopsis proteins with histone H1 globular domain. Plant Physiol. 174: 27-34.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Lapidot, M. and Pilpel, Y. (2006) Genome-wide natural antisense transcription: coupling its regulation to its different regulatory mechanisms. EMBO Rep. 7: 1216-1222.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Laubinger, S., Sachsenberg, T., Zeller, G., Busch, W., Lohmann, J.U., Raescht, G. and Weigel, D. (2008) Dual roles of the nuclear cap-binding complex and SERRATE in pre-mRNA splicing and microRNA processing in Arabidopsis thaliana. Proc. Natl. Acad. Sci. USA 105: 8795-8800.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Liu, F., Marquardt, S., Lister, C., Swiezewski, S. and Dean, C. (2010) Targeted 3 ' processing of antisense transcripts triggers Arabidopsis FLC chromatin silencing. Science 327: 94-97.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Liu, J., Wang, H. and Chua, N.-H. (2015) Long noncoding RNA transcriptome of plants. Plant Biotechnol. J. 13: 319-328.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Love, M., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15: 550.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Mondal, T., Rasmussen, M., Pandey, G.K., Isaksson, A. and Kanduri, C. (2010) Characterization of the RNA content of chromatin. Genome Res. 20: 899-907.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Mustroph, A., Zanetti, M.E., Jang, C.J.H., Holtan, H.E., Repetti, P.P., Galbraith, D.W., Girke, T. and Bailey-Serres, J. (2009) Profiling

translatomes of discrete cell populations resolves altered cellular priorities during hypoxia in Arabidopsis. Proc. Natl. Acad. Sci. USA 106: 18843-18848.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Rangwala, S.H. and Richards, E.J. (2010) The structure, organization and radiation of Sadhu non-long terminal repeat retroelements in Arabidopsis species. Mobile DNA 1: 10.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Ransohoff, J.D., Wei, Y. and Khavari, P.A. (2018) The functions and unique features of long intergenic non-coding RNA. Nature Rev Mol Cell Biol 19: 143-157.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Rinn, J.L. and Chang, H.Y. (2012) Genome regulation by long noncoding RNAs. Annu. Rev. Biochem. 81: 145-166.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Ron, M., Saez, M.A., Williams, L.E., Fletcher, J.C. and McCormick, S. (2010) Proper regulation of a sperm-specific cis-nat-siRNA is essential for double fertilization in Arabidopsis. Genes Dev. 24: 1010-1021.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Rosa, S., Duncan, S. and Dean, C. (2016) Mutually exclusive sense-antisense transcription at FLC facilitates environmentally induced gene repression. Nat Commun 7: 13031.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Schein, A., Zucchelli, S., Kauppinen, S., Gustincich, S. and Carninci, P. (2016) Identification of antisense long noncoding RNAs that function as SINEUPs in human cells. Scientific Reports 6: 33605.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Schmitz, K.M., Mayer, C., Postepska, A. and Grummt, I. (2010) Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes. Genes Dev. 24: 2264-2269.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Seo, J.S., Sun, H.X., Park, B.S., Huang, C.H., Yeh, S.D., Jung, C. and Chua, N.H. (2017) ELF18-INDUCED LONG-NONCODING RNA associates with Mediator to enhance expression of innate immune response genes in Arabidopsis. Plant Cell 29: 1024-1038.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Shimada, T.L., Shimada, T. and Hara-Nishimura, I. (2010) A rapid and non-destructive screenable marker, FAST, for identifying transformed seeds of Arabidopsis thaliana. Plant J. 61: 519-528.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Song, L., Huang, S.S.C., Wise, A., Castanon, R., Nery, J.R., Chen, H.M., Watanabe, M., Thomas, J., Bar-Joseph, Z. and Ecker, J.R. (2016) A transcription factor hierarchy defines an environmental stress response network. Science 354: aag1550.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Sun, H., Bi, Y., Tao, J., Huang, S., Hou, M., Xue, R., Liang, Z., Gu, P., Yoneyama, K., Xie, X., et al. (2016) Strigolactones are required for nitric oxide to induce root elongation in response to nitrogen and phosphate deficiencies in rice. Plant Cell Environ. 39: 1473-1484.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Tran, N.T., Su, H.R., Khodadadi-Jamayran, A., Lin, S., Zhang, L., Zhou, D.W., Pawlik, K.M., Townes, T.M., Chen, Y.B., Mulloy, J.C., et al. (2016) The AS-RBM15 lncRNA enhances RBM15 protein translation during megakaryocyte differentiation. EMBO Rep. 17: 887-900.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

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.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Tripathi, V., Ellis, J.D., Shen, Z., Song, D.Y., Pan, Q., Watt, A.T., Freier, S.M., Bennett, C.F., Sharma, A., Bubulya, P.A., et al. (2010) The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. Mol. Cell 39: 925-938.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Wang, H., Chung, P.J., Liu, J., Jang, I.-C., Kean, M.J., Xu, J. and Chua, N.-H. (2014) Genome-wide identification of long noncoding natural antisense transcripts and their responses to light in Arabidopsis. Genome Res. 24: 444-453.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Wang, X.J., Gaasterland, T. and Chua, N.H. (2005) Genome-wide prediction and identification of cis-natural antisense transcripts in Arabidopsis thaliana. Genome Biol. 6: R30.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Wang, Y., Xu, Z., Jiang, J., Xu, C., Kang, J., Xiao, L., Wu, M., Xiong, J., Guo, X. and Liu, H. (2013) Endogenous miRNA sponge lincRNA-RoR regulates Oct4, Nanog, and Sox2 in human embryonic stem cell self-renewal. Dev. Cell 25: 69-80.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Weiner, A.M. (2002) SINEs and LINEs: the art of biting the hand that feeds you. Curr. Opin. Cell Biol. 14: 343-350.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Wimalasekera, R., Villar, C., Begum, T. and Scherer, G.F.E. (2011) COPPER AMINE OXIDASE1 (CuAO1) of Arabidopsis thaliana contributes to abscisic acid- and polyamine-induced nitric oxide biosynthesis and abscisic acid signal transduction. Mol Plant 4: 663-678.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Wu, H.-J., Wang, Z.-M., Wang, M. and Wang, X.-J. (2013) Widespread long noncoding RNAs as endogenous target mimics for microRNAs in plants. Plant Physiol. 161: 1875-1884.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Wucher, V., Legeai, F., Hedan, B., Rizk, G., Lagoutte, L., Leeb, T., Jagannathan, V., Cadieu, E., David, A., Lohi, H., et al. (2017) FEELnc: a tool for long non-coding RNA annotation and its application to the dog transcriptome. Nucleic Acids Res. 45: e57.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Xiao, Z.T., Zou, Q., Liu, Y. and Yang, X.R. (2016) Genome-wide assessment of differential translations with ribosome profiling data. Nat Comm 7: 11194.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Yoo, S.-D., Cho, Y.-H. and Sheen, J. (2007) Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. Nat Protoc 2: 1565-1572.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Yoon, J.H., Abdelmohsen, K., Srikantan, S., Yang, X.L., Martindale, J.L., De, S., Huarte, M., Zhan, M., Becker, K.G. and Gorospe, M. (2012) LincRNA-p21 suppresses target mRNA translation. Mol. Cell 47: 648-655.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Yuan, C.H., Wang, J.J., Harrison, A.P., Meng, X.W., Chen, D.J. and Chen, M. (2015) Genome-wide view of natural antisense transcripts in Arabidopsis thaliana. DNA Res. 22: 233-243.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Yuan, J.P., Zhang, Y., Dong, J.S., Sun, Y.Z., Lim, B.L., Liu, D. and Lu, Z.J. (2016) Systematic characterization of novel lncRNAs responding to phosphate starvation in Arabidopsis thaliana. BMC Genom 17: 655.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Zhang, X.-N., Shi, Y., Powers, J.J., Gowda, N.B., Zhang, C., Ibrahim, H.M.M., Ball, H.B., Chen, S.L., Lu, H. and Mount, S.M. (2017) Transcriptome analyses reveal SR45 to be a neutral splicing regulator and a suppressor of innate immunity in Arabidopsis thaliana. BMC Genom 18: 772.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Zhou, X.F., Sunkar, R., Jin, H.L., Zhu, J.K. and Zhang, W.X. (2009) Genome-wide identification and analysis of small RNAs originated from natural antisense transcripts in Oryza sativa. Genome Res. 19: 70-78.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Zhou, Y., Hartwig, B., James, G.V., Schneeberger, K. and Turck, F. (2016) Complementary activities of TELOMERE REPEAT BINDING proteins and polycomb group complexes in transcriptional regulation of target genes. Plant Cell 28: 87-101.