Endogenous miRNA in the green alga *Chlamydomonas* regulates gene expression through CDS-targeting.

Betty Y-W. Chung*[1], Michael J. Deery[2], Arnoud J. Groen[2], Julie Howard[2] and David Baulcombe*[1].

*** Corresponding Authors: bcy23@cam.ac.uk, dcb40@cam.ac.uk**

[1] Department of Plant Sciences, University of Cambridge, Cambridge, CB2 3EA, United Kingdom

[2] Cambridge System Biology Centre and Department of Biochemistry, University of Cambridge, CB2 1GA, United Kingdom

Keywords miRNA, ribosome profiling, translation, Chlamydomonas, mRNA turnover

Word count: 4133 words

Abstract

MicroRNAs regulate gene expression as part of the RNA-induced silencing complex, for which the sequence identity of the miRNA provides the specificity to the target messenger RNA and the result is target repression. The mode of repression can be through target cleavage, RNA destabilization and/or decreased translational efficiency. Here, we provide a comprehensive global analysis of the evolutionarily distant unicellular green alga *Chlamydomonas reinhardtii* to quantify the effects of miRNA on protein synthesis and RNA abundance. We show that, similar to metazoan steady-state systems, endogenous miRNAs in *Chlamydomonas* can regulate gene-expression both by destabilization of the mRNA and by translational repression. However, unlike metazoan miRNA where target site utilization localizes mainly to 3'UTRs, in *Chlamydomonas* utilized target sites lie predominantly within coding regions. These results demonstrate the evolutionarily conserved mode of action for miRNAs, but details of the mechanism diverge between plant and metazoan kingdoms.

Introduction

MicroRNAs (miRNA) are 21-24 nucleotide RNAs present in many eukaryotes that guide the silencing effector Argonaute (AGO) protein to target mRNAs via a base pairing process[1]. The AGO complex either catalyzes endonucleolytic cleavage or promotes translation repression and/or accelerated decay of this target mRNA[2]. There is overwhelming evidence both *in vivo* and *in vitro* that translation repression plays a major role[3–7]. However, there has been controversy about which of these three mechanisms is more significant *in vivo,* especially when effects of miRNA on endogenous genes cannot be faithfully represented by reporter systems in which, at least in metazoans, the observed repression vastly exceeds that typically observed for endogenous mRNAs[8,9].

Recent *in vivo* studies in mammalian cells provide support for endogenous mRNA destabilization over translation repression as the dominant effect of miRNA under steady-state conditions[9]. However an inducible zebrafish embryo system in which

miR430 is only expressed two hours post fertilization, reveals that translation repression occurs prior to accelerated mRNA decay[7]. This conclusion was further supported by findings in mouse liver, primary macrophages and primary B and T cells [8].

In contrast to the metazoan systems, there is a lack of comprehensive studies on the endogenous effects of miRNAs in plants and the question remains as to whether miRNA modulates by translation repression and/or promoting mRNA turnover. In plants miRNA-mediated gene regulation does occur[10–12] but, unlike metazoan systems, the targets can be in the coding sequence as well as 3'UTR and the mechanism may involve endonucleolytic cleavage rather than accelerated decay or translation inhibition[13,14]. Most plant studies, however, are based on individual miRNAs or reporter assays that may not be informative about endogenous mRNA systems[9,15,16]. We therefore utilized the unicellular green alga *Chlamydomonas reinhardtii*, for which we have previously discovered and characterized its miRNAs[17] and generated *DCL3* mutants[18].

Chlamydomonas is a particularly amenable experimental system because its unicellularity reduces complications with tissue-specific effects. Similar to higher plants, the machinery for miRNA-mediated translation regulation is also functional in *Chlamydomonas*, where the seed-region rule utilised by the metazoan system is adequate for translation repression, at least within reporter systems[19]. In this present study, we utilized two silencing mutants raised from our previous forward genetic screen at *dcl3*[18] and *ago3* (Chung *et. al.* 2017 in preparation). The *dcl3-1* mutant results in almost complete loss of miRNA as well as 21-nt small interfering (si)RNAs whereas *ago3-25* is defective in AGO3 that binds to mRNA and is required for translation repression in the reporter system[20]. Neither mutant had obvious growth differences or morphological abnormalities under normal conditions[18]. Any effect seen in both *dcl3-1* and *ago3-25* on gene expression is likely, therefore, to be direct rather than an indirect secondary consequence of metabolic changes due to loss of miRNA-mediated regulation.

Here, through a combination of ribosome profiling, parallel RNA-Seq, sRNA-Seq and quantitative proteomics at mid-log phase of the *dcl3-1* mutant and its corresponding complemented strain we have demonstrated that, in contrast to the metazoan system, the primary effect of miRNA in *Chlamydomonas* is through interaction with CDS regions instead of 3' UTRs. However, similar to the metazoan system, miRNA in *Chlamydomonas reinhardtii* can also modulate gene expression via means of translational repression and mRNA turnover. Finally, and perhaps the most striking observation is that the translation apparatus itself is differentially regulated at the level of translation efficiency but not RNA abundance in the presence of the miRNA machinery.

Results and Discussion

Loss of DCL3 function does not affect the genome-wide RNA or translation profile.

To explore the possibility that DCL3-dependent miRNA or siRNA regulates gene expression by either promoting mRNA turnover or through interfering with translation, we applied ribosome profiling, parallel RNA-Seq and quantitative N15 proteomics to biological triplicates of the vegetative mid-log phase *dcl3-1* mutant and its corresponding complemented derivative (abbreviated as C*dcl3*) carrying a wild type *DCL3* allele introduced into the mutant strain. The experimental protocol is summarized in supplementary Figure 1 and supplementary Figure 2 illustrates the high degree of reproducibility between biological repeats in these data.

The slightly smaller footprint size of plant/algae ribosomes leads to differences in the phasing patterns compared to mammalian ribosome profiling studies[21]. In both the complemented strain C*dcl3* and the *dcl3* mutant, the 5' end of the 27-nt ribosome protected fragments (RPFs), mapped predominantly to the second codon position; in contrast and, as expected, RNA-Seq reads were uniformly distributed at all three codon positions (Figures 1A and B). The RPF 5' end position distributions at start and stop codons were also similar in the *dcl3-1* and C*dcl3* strains (Figures 1C and D respectively) in that there was a sharp 27-nt peak on the start codon (reflecting the rate-limiting initiation step of translation) and a sharp 28-nt peak on the stop codon (reflecting the conformation change from an elongating ribosome to a terminating ribosome, Supplementary figure 3B)[21]. In contrast, the RNA-seq reads are not limited to coding regions (Figures 1E, F and Supplementary Figures 3B).

The validity of these data was further confirmed with the analysis of *DCL3*. There were multiple *DCL3* mRNA reads from three replicate samples of the C*dcl3* strain that were restricted to the open reading frame in the RPF datasets. In *dcl3-1* the reads were from the region on the 5' side of the mutagenic *DCL3* insertion (Supplementary Figures 3C). Finally, Ribosome protected fragments (RPF), RNA abundance (RA), and translational efficiencies (TE) for expressed genes are well correlated between *dcl3-1* and C*dcl3* (R^2 = 0.95, 0.97 and 0.98 for TE, RPF and RNA, respectively, Supplementary Figure 3E). From these data, we conclude that any global effect of DCL3 on the translatome is minor but we could not rule out quantitative effects on a subset of RNAs.

To explore this possibility, we refined our analysis by dividing the mRNA profiles into those with or without predicted targets of the DCL3-dependent miRNAs. The first stage in this analysis was to re-evaluate the miRNA precursors in *C. reinhardtii* that we had previously identified as being both coding and non-coding RNAs. Now, however, with the use of the RPF data to identify translated open reading frames, we find that all miRNAs in this alga derive from introns or the exons (3'UTR or coding) of mRNAs. Supplementary Figure 4 and Table 2 is an updated summary of the 42 miRNA precursors in *C. reinhardtii* described in Valli *et. al.* 2016[18].

Our subsequent analysis differentiated mRNAs with miRNA targets in the 5' UTR, CDS and 3' UTR from those without targets. The CDS regions were defined by the R software Bioconductor package – riboSeqR - that utilizes the triplet periodicity of ribosome profiling for the *de novo* inference of AUG-initiated coding sequences that

are supported by RPFs[21] and we used the seed-sequence rule to identify miRNA target motifs[22,23]. This rule requires base-pairing of the first 8 nucleotides of miRNA and it is supported by direct assay of miRNA targeting and structural studies of human AGO2[24] and by experimental tests in higher plants[25] and *C. reinhardtii*[19].

To identify the miRNA-target mRNAs we first filtered for the 19 most-abundant *DCL3*-dependent miRNAs in our sRNA-Seq data (Supplementary Figure 5; see also Materials and Methods). We then applied the TargetScan prediction algorithm[22,23] to the mRNAs with RPF-validated ORFs. This criterion meant that the TargetScan algorithm was applied to 13,073 expressed transcripts (out of 17,741 annotated transcripts) of which 2,439 do not contain any predicted 8mer miRNA target sites. Of all the predicted target sites, a larger proportion (70%) are located in the CDS (Figure 2A) compared to UTRs (10% for 5'UTR and 36% for 3'UTR). This distribution is likely, at least in part, a reflection of greater length of the CDS compared to UTR regions. Using a more stringent miRNA targeting rule did not have a large change on these numbers: about half of the mRNA seed sequence targets also have >50% sequence complementarity to the relevant miRNA in the sequences upstream of the 3' eight nucleotides (Figure 2B).

Next, we excluded the RNAs with predicted target sites in more than one region (5'UTR/CDS/3'UTR) because for these it would have not been possible to differentiate the effects of miRNA acting in the different regions. In addition, we also excluded mRNAs with miRNA precursors because they are unstable in the presence of DCL3 as a consequence of miRNA processing (see supplementary Figure 4 and[18]). Following application of these filters our further analysis was based on 292 mRNAs with 5' UTR targets, 5,205 with CDS targets, 1,262 with targets in the 3' UTR and the 2,439 without predicted targets.

Similar to studies by the Bartel and Giraldez groups[7–9] we plotted cumulative distributions of differential translation efficiency, total RPF and RA for target and non-target mRNAs in the *dcl3-1* mutant and C*dcl3* to assess the miRNA-mediated effects of DCL3 (Figure 3A and B). Differential TE is computed as $(RPF_C/RNA_C)/(RPF_{dcl3}/RNA_{dcl3})$. The analysis revealed that, similar to the analysis of zebrafish[7], the major effects of Dicer loss of function (*dcl3-1* vs C*dcl3*) were on mRNAs containing target sites within the CDS and the effect is more significant in the RPF than the RNA data, contributing to its significant but small effect in TE. The effects were evident as a shift to increased RNA abundance for mRNAs with target sites in *dcl3-1* and they are consistent with the canonical role of miRNAs as negative regulators.

The difference in *dcl3-1* versus C*dcl3* was greater in transcripts with CDS rather than UTR target sites and this effects appears to be dosage dependent, where mRNAs with four or more CDS targets were affected to a greater extent than those with fewer target sites (Figure 3C). However, this dosage-dependent effect was not observed for mRNAs containing target sites in the UTRs (Supplementary Figure 6A). Furthermore, these effects are also consistent at the protein level for mRNAs with supportive proteomics data (Supplementary Figure 6B).

As the key AGO in *Chlamydomonas* known to be associated with miRNA is AGO3 which mediates translational repression in a reporter system[20], we also performed

ribosome profiling as well as corresponding RNA-seq on an *AGO3* mutant (*ago3-25*), raised from the same forward genetic screen as *dcl3-1*[18], as well as the corresponding parental strain and the wild type cc-1883 (Chung *et. al.* 2017, in preparation) in order to further validate whether these effects are truly due to the miRNA machinery. Supporting this, we also observed the dosage-dependent effect only for mRNAs containing target sites within the CDS in the *ago3-25* mutant background (Figure 3D and Supplementary Figure 6A).

The global effect of mRNA repression is not likely due to target RNA cleavage as there are only 85 potential CDS target sites (83 mRNAs) complying with the plant targeting rule in *Chlamydomonas*[17]. Moreover, of these potential CDS cleavage site mRNAs, only 18/83 were expressed in our dataset, albeit at very low level (Supplementary Figure 6C). We also investigated potential targets for expressed miRNA where the base-pairing is between positions 2-15 (allowing one mismatch) and, similar to the plant-rule potential targets, there were very few candidates (47 in total), of which only 31 are expressed in our dataset and the expression level for all 31 mRNAs is low (Supplementary Figure 6C). Thus, well expressed genes are unlikely to be cleaved under steady-state conditions, consistent with the lack of phenotype for both *dcl3-1* and *ago3-25* mutants. A recent degradome study is also consistent with there being minimal miRNA target site cleavage in *Chlamydomonas.* The study involved miR-910, an miRNA also expressed in our sample, that cleaved only two mRNAs upon salt-stress[26]. The endogenous miRNA-mediated RNA down-regulation by CDS-targeted miRNA is not, therefore, likely to be mainly through target cleavage.

Finally, we tested the effect of miRNA abundance on TE, RPF and RA by focusing on the most abundant miRNA in our corresponding sRNA-Seq datasets: miR-C89 (Figure 3E, F and supplementary Figure 5; 5'UTR and protein data excluded due to small sample size). MiR-C89 correlated with a larger shift in TE and RA than other miRNAs consistent with magnitude of the effect being influenced by miRNA abundance.

From these findings we conclude that, similar to metazoan systems[8,9], *Chlamydomonas* miRNA generally fine tunes gene expression through an effect on both RNA abundance and translation efficiency (Figure 3). The global effect on translation efficiency was significant although smaller than the effect on RNA abundance (Figures 3A and B), as in metazoans[9]. Unlike metazoans, however, the primary targets of miRNAs in *Chlamydomonas* are in the CDS instead of 3'UTRs (Figure 3). This difference may reflect differences between *Chlamydomonas* and metazoans in the ways in which miRNAs may influence elongating ribosomes.

Translation efficiency of 80S ribosomal proteins is higher in the DCL3 mutant.

Our finding that miRNA targeting in *Chlamydomonas* is influenced by miRNA abundance and the number of target sites (Figure 3) implies that some mRNAs may be affected more than others. Therefore, to detect possible changes in individual mRNAs, we plotted the *dcl3-1* versus C*dcl3* differences in TE and RA for all mRNAs with CDS-exclusive target sites (Figure 4). Using the *dcl3-1* mutation as a benchmark (log2FC(TE) = 0.7 and log2FC(RNA) = 1.18), individual RNAs that are negatively

regulated by miRNAs would distribute in field A of this figure if TE is affected (i.e. log2FC(TE) $\leq$ -0.7, yellow shaded area), field C if RA is affected but not TE (i.e. log2FC (RA) $\leq$ -1.18, -0.7 $\leq$ log2FC(TE) $\leq$ 0.7, purple shaded area) and in field B if there was a double effect on both TE and RA (log2FC(RA) $\leq$ -1.18, log2FC(TE) $\leq$ -0.7, red shaded area). Corresponding positive regulation would be indicated by distribution in fields A', B' and C' respectively (Figure 4A).

The distribution of mRNA in this plot is consistent with a higher degree of negative rather than positive regulation on a few mRNAs: there were 32 and 16 targets in A and A' respectively, 3 and 0 in B and B', and 15 and 3 in C and C'. From this analysis we conclude that there may be up to 32 mRNAs that are subject to strong translational regulation by miRNAs (from the A and B fields), 15 subject to strong regulation of RNA abundance (from the B and C*dcl3* fields) and 3 subject to strong regulation at both levels. The RNA-Seq and RPF data for *DCL3* mRNA and selected miRNA targets including rpL14 and Cre16.g67520 from field A are presented in Figure 4 C-E.

To assess whether the mRNAs in field C could either be miRNA targets or they could have DCL3 cleavage sites we monitored their level in *ago3-25* and the wild-type (Supplementary Figure 6D). Repression of RNAs that are targeted by DCL3 would be relieved in *dcl3-1* but not *ago3-25* whereas those that are targeted by miRNAs would be depressed in both mutants.

The data are consistent with miRNA targeting for most of the field C RNAs of Figure 4 because their repression was relieved in both mutants although Cre15.g643503.t1.1 was an exception (perhaps related to it having an unusually long CDS - 7884 nt, cf. average CDS length for expressed genes = 2429 nt; Supplementary Figure 7D). We therefore conclude that the RA effect we observe is genuinely directed by the miRNA-AGO complex. Further, in order to distinguish whether reduced expression in C*dcl3* relative to *dcl3-1* was a global effect or merely due to a small number of strongly repressed genes (i.e. fields A, B, C, A', B' and C' of Figure 4A), we repeated the analysis with the strongly repressed candidates excluded and found a similar pattern of global mRNA repression as with all mRNAs (Supplementary Figure 8A). Similarly, with the targets of miR-C89 the repression of TE or RA primarily results from small changes in the expression of many genes (Supplementary Figure 8A and B).

It is striking that mRNAs subject to either strong translational or RNA stability regulation (i.e. field A and C) are enriched with those encoding RNA-interacting proteins (e.g. translation, transcription and rRNA processing) (Supplementary Table 3). Of the mRNAs subject to translational regulation a gene ontology analysis revealed the enriched pathway of "translation and ribosome" with the mRNAs for 80S ribosomal proteins being particularly prominent (Figure 4A and Supplementary Table 3). These candidates also contribute to the outlier group for TE and RPF but not RA in the cumulative distributions for transcripts with supporting proteomic data (Supplementary Figure 6B). Furthermore, the same enrichment is also observed in the *ago3-25* mutant (Figure 4E and Supplementary Figure 7C). However, we do not observe enrichment for this pathway in previously published mammalian datasets[9] of miR-233 knockout cultured neutrophils compared with wild-type culture neutrophils, and HeLa cells after transfection with miR-1 or miR-155 (Supplementary Figure 8).

The enrichment of "translation and ribosome" function in fields A and C of Figure 4A and E is specific for 80S ribosomal proteins; the nucleus-encoded 70S ribosomal proteins for both chloroplasts and mitochondria were an internal control and cluster around the 0-fold change axis for both TE and RNA (Figure 4A and E). It is likely therefore that the specific effect for the 80S factors reflects the targeting specificity of miRNAs in *Chlamydomonas* or that it is a compensatory mechanism for the loss of a layer of regulation in the *dcl3-1* and *ago3-25* mutants.

It is possible that the distribution of ribosomes on the mRNA would be affected by absence of miRNAs (see Figures 4B and C for example rpL14 and Cre16.g675200). However, we did not observe any significant correlation between the position of the miRNA target sites and the distribution of RPF or RNA reads for the mRNAs of fields A and C of Figure 4A either individually or through a global analysis of multiple RNAs. In contrast, in the mRNA for *DCL3* there was an effect: the RPFs in the C*dcl3* sample extended to the stop codon and the RNA-Seq reads covered the full length mRNA whereas, in *dcl3-1,* the RPF and RNA-Seq data were more sparse than in C*dcl3* and they stopped at the site of the mutagenic *hyg* insert (Figure 4D and Supplementary 3C). Clearly, from this *DCL3* analysis, the RPF and RNA-Seq data can reflect both the quantitative and qualitative aspects of ribosome distribution and RNA accumulation.

We hypothesized that CDS-targeting of the miRNA-AGO complex should result in road-blocking of elongating ribosomes, resulting in ribosome pile-up and/or drop-off 5' and 3' end of miRNA target sites respectively. However we did not observe any significant changes in RPF density around miRNA target sites, indicating that RISC does not induce ribosome pileup within CDS regions. Presumably the efficient RNA helicase activity of the ribosomes is able to overcome the steric hindrance by the RISC in *Chlamydomona* [27,28]. There may, however, be a transient effect on ribosome translocation. Having now identified these RNAs with the greatest effect on TE and RNA we will be able to explore the factors affecting the two modes of RNA regulation and the conditions under which miRNAs have the greatest effect on their mRNA targets.

Materials and Methods

Culturing and harvesting *Chlamydomonas*

Three independent fresh single colonies of *Chlamydomonas reinhardtii* cells were sub-cultured as biological triplicates. Cells where grown in 50 ml Tris-acetate-phosphate (TAP) medium at 23 ºC in baffled flasks on a rotatory shaker (140 rpm) under constant illumination with white light (70 µE m^2 sec^{-1}) to mid-log phase (OD_{750} ~ 0.6), followed by inoculation into 750 ml TAP in 2 L baffled flasks at OD_{750} = 0.2. These were cultured in the same conditions until mid-log phase prior to harvesting by filtering off the media, after which the cell paste was immediately flash frozen and pulverized in liquid nitrogen with 5 mL of pre-frozen buffer (20 mM Tris-Cl pH 7.5, 140 mM KCl, 5 mM $MgCl_2$, 10 µg/ml cycloheximide, 100 µg/mL chloramphenicol, 0.05 mM DTT, 0.5% NP40, 1% Triton X-100 and 5% sucrose). The frozen powder was gradually thawed on ice and clarified by centrifugation for 30 min at 4700 rpm at 4 ºC followed by adjustment of A_{254} = 10 before further treatment, or snap frozen in liquid nitrogen and stored at −80 ºC. The extraction efficiency was monitored by polysome profiling (Supplementary Figure 3F). The flash freezing method was preferred as methods involving pretreatment with translational inhibitors such a cycloheximide or chloramphenicol can introduce various biases, in particular in artificially enhancing the initiation peak of the profile[29], which we also observed in *Chlamydomonas reinhardtii* when we compared flash-freezing with cycloheximide pretreatment (Supplementary Figure 3G).

Metabolic labelling and LC-MS/MS

For metabolic labelling, ammonia chloride (14N) was replaced with ammonia chloride-15N (Cambridge Isotope Laboratories Inc) in the TAP media used to maintain *dcl3-1*. There were no obvious differences in growth rates between algae maintained in N14 and N15. *dcl3-1*-N15 and *Complement*-N14 were mixed equally prior to protein extraction via TCA-acetone precipitation followed by resuspension in resuspension buffer (8 M urea, 500 mM NaCl, 10 mM Tris-Cl pH 8, 5 mM DTT) and resolved in 1.5 mm 10% bis-tris Novex Gel (Thermo Fisher Scientific Inc, Waltham, MA, USA). The experiment was performed in biological triplicate.

1D gel bands (12 per lane) were transferred into a 96-well PCR plate. The bands were cut into 1 mm^2 pieces, de-stained, reduced (DTT), alkylated (iodoacetamide) and subjected to enzymatic digestion with trypsin overnight at 37 °C. After digestion, the supernatant was pipetted into a sample vial and loaded onto an autosampler for automated LC-MS/MS analysis.

All LC-MS/MS experiments were performed using a Dionex Ultimate 3000 RSLC nanoUPLC (Thermo Fisher Scientific Inc, Waltham, MA, USA) system and a QExactive Orbitrap mass spectrometer (Thermo Fisher Scientific Inc, Waltham, MA, USA). Separation of peptides was performed by reverse-phase chromatography at a flow rate of 300 nL/min and a Thermo Scientific reverse-phase nano Easy-spray column (Thermo Scientific PepMap C18, 2 µm particle size, 100 Å pore size, 75 µm i.d. x 50 cm length). Peptides were loaded onto a pre-column (Thermo Scientific PepMap 100 C18, 5 µm particle size, 100 Å pore size, 300 µm i.d. x 5 mm length) from the Ultimate 3000 autosampler with 0.1% formic acid for 3 min at a flow rate of 10 µL/min. After this period, the column valve was switched to allow elution of peptides from the pre-column onto the analytical column. Solvent A was water +

0.1% formic acid and solvent B was 80% acetonitrile, 20% water + 0.1% formic acid. The linear gradient employed was 2-40% B in 30 min (total run time including a high organic wash step and requilibration was 60 min).

The LC eluant was sprayed into the mass spectrometer by means of an Easy-Spray source (Thermo Fisher Scientific Inc.). All *m/z* values of eluting ions were measured in an Orbitrap mass analyzer, set at a resolution of 70000 and was scanned between *m/z* 380-1500. Data dependent scans (Top 20) were employed to automatically isolate and generate fragment ions by higher energy collisional dissociation (HCD, NCE:25%) in the HCD collision cell and measurement of the resulting fragment ions was performed in the Orbitrap analyser, set at a resolution of 17500. Singly charged ions and ions with unassigned charge states were excluded from being selected for MS/MS and a dynamic exclusion window of 20 s was employed.

Protein identification and relative quantitation

Data were recorded using Xcalibur™ software version 2.1 (Thermo Fisher Scientific, San Jose, CA). Files were converted from .raw to .mzXML using MSConvert and then .mzXML files to .mgf using the in-house software iSPY[30,31]. The .mgf files were submitted to the Mascot search algorithm. The following parameters were employed: carbamidomethyl as a fixed modification, and oxidation on methionine (M) residues and phosphorylation on serine (S), threonine (T), and tyrosine (Y) residues as variable modifications; 20 ppm for peptide tolerance, 0.1 Da of MS/MS tolerance; a maximum of two missed cleavages, a peptide charges of +2, +3, or +4; and selection of a decoy database. Mascot .dat output files were imported into iSPY for 14N/15N quantitation and analysed through Percolator for improved identification[32]. The 14N and 15N peptide isotopic peaks from the MS1 dataset were used to compare the theoretical mass difference between the heavy and light peptides, and the typical isotopic distribution patterns. Only unique peptides with a posterior error probability (PEP-value) of ≤ 0.05 were considered for further analysis. Spectra were merged into peptides and proteins based on their median intensity in MS1, meaning the more intense the signal of the spectrum, the more weight it added to quantitation. The statistical programming environment R was used to process iSPY output files to check for the 15N incorporation rate and to confirm that the data were normally distributed. After normalization, only peptides detected in at least two biological replicates, with a fold change ≥ 1.5 and a *p*-value ≤ 0.05 were considered for further analysis. Relative protein expression values were computed as ($Protein_{C}/Protein_{dcl3}$) using the average of the triplicates for all follow-up analysis.

Nuclease footprinting

Lysates (200 µL) were slowly thawed on ice and treated with 6000 units RNase I (Thermo Fisher Scientific Inc.). in a thermo-mixer at 28 ºC, 400 rpm for 30 min. The reaction was stopped by mixing the digest reaction with 120 units of SUPERase-In RNase inhibitor (Thermo Fisher Scientific Inc.) followed by centrifugation for 2 min at 14000 rpm at 4 ºC to further clarify any remaining debris. The supernatant was layered onto a 1 M sucrose cushion prepared in *Chlamydomonas* polysome buffer, and RNA were purified as described in Ingolia *et. al.*[33]. Polysome integrity for the lysate and digestion conditions were assessed via polysome profiling (Supplementary Figure 3F).

Ribosome profiling and RNA-Seq

The methodologies were largely based on the protocols of Ingolia *et. al.* and Guo *et. al.*[9,33] with modifications (i) mRNA for corresponding RNA-Seq was enriched by removal of rRNA using the ribo-zero kit (plant seed and root kit), (ii) RNA-Seq size selection was in parallel with ribosome profiling (i.e. between 26 and 34 nt), and (iii) for ribosome profiling, ribosomal RNA contamination was removed by two rounds of treatment with duplex specific nuclease (DSN) for 30 min as described in (Chung *et. al.* 2015).

Preparation for sRNA libraries

Small RNA from total RNA samples used for RNA-Seq were size excluded in 15% TBU gel for miRNA enrichment (Thermos Scientific). The sRNA were further prepared according to the NEXTflex small RNA-Seq kit v2 (Bio Scientific), followed by sequencing on the NextSeq500 platform.

Computational analysis of ribosome profiling and RNA-Seq data

After removal of adaptor sequences, Illumina sequencing reads were mapped to the reference transcriptome (Phytozome 281) or miRNA precursor sequences described in Valli *et. al.* 2016[18] using bowtie-1 and processed as described in Chung *et. al.* 2015[21]. Only mRNAs with more than 50 RPF reads of size 27 or 28 nt uniquely mapped to more than 10 positions were considered. Corresponding RNA-Seq reads within coding regions *de novo* defined by ribosome profiling were extracted for differential RA as well as TE analysis using riboSeqR as described in Chung *et. al.* 2015[21]. Further filtering was applied for fold change analyses where mRNAs were only considered if they had (i) at least 10 normalised RPF and 10 normalised RNA counts, and (ii) the sum of all RPF or RNA counts over the three biological replicates for both *dcl3-1* and complement combined is at least 200. Normalisation was based on BaysSeq output [34]. Cumulative distributions for TE, RPF and RA fold changes were calculated based on the average of all three replicates. Differential analyses for the mouse data in Guo *et. al.* 2010[9] were obtained from the Gene expression Omnibus in NCBI (accession:GSE220001 and GSE21992).

Target prediction

Target prediction was done using TargetScan[23] using the same transcriptome input as for the ribosome profiling analysis. As there are no conserved sites available due to lack of miRNA data from the green algae phylum, we could not calculate context and scores; thus we only utilized the part of the software to detect all possible miRNA target sites. Further, as the efficacy between 8mer-A1 and 8mer-m8 sites are similar, we combined both types of target sites in the 8mer prediction, similar to Guo *et. al.* 2010 and Agarwal *et. al.* 2015[9,23]. Target prediction based on the plant rule was performed via TAPIR[35].

The list of miRNA used was based on the 19 *DCL3*-dependent miRNAs expressed based on the sRNA data, where the average reads within the complement is greater than 400 and the average ratio of complement to *dcl3-1* reads is greater than 150. The selected *DCL3*-dependent miRNA used are: chromosome_5_3227666_3227753_+ (miR-C89), chromosome_6_6776108_6776193_+ (miR-cluster20399), chromosome_13_2001067_2001197_- (miR-cluster 7085), chromosome_10_3399870_3399999_- (miR9897), chromosome_13_3152367_3152452_- (miR-C112), chromosome_6_3067368_3067456_+ (miR1162),

chromosome_12_6402226_6402307_- (miR1157),
chromosome_9_6365928_6366014_- (miR912),
chromosome_7_4386252_4386309_- , chromosome_17_6144120_6144204_+ (miR-cluster12551), chromosome_1_7070552_7070605_-,
chromosome_16_185088_185174_-(miR1169),
chromosome_2_8349161_8349264_+ , chromosome_2_9129508_9129593_- miR-cluster14712), chromosome_7_5926395_5926482_+ (miR-C59),
chromosome_14_3218783_3218866_- (miR910),
chromosome_6_7063792_7063881_- (miR1152),
chromosome_4_3100624_3100751_+ (miR1153) and
chromosome_1_5106349_5106475_+ (miR-C82). The miRNA precursor sequence used for mapping was based on Valli *et. al.* (2016). Only 8mer sites were utilized, and 8mer complementarity was verified via extraction of target sites followed by miRNA complementarity assessment using the Vienna RNA package program RNAduplex. The level of 3' complementarity was similarly investigated where nt 9 to 21 of the target site 3' of the seed region was extracted and the level of complementarity assessed with RNAduplex.

Acknowledgements

We thank J. Barlow for technical assistance and media preparation; T. J. Hardcastle and B. Santos for technical bioinformatic support; A. Valli for providing the silencing mutants and, A. Molnar and A. E. Firth for discussions. This work was supported by a Balzan Prize award and the European Research Council Advanced Investigator Grant ERC-2013-AdG 340642 TRIBE. B.Y.W.C. was supported by an EMBO long-term postdoctoral fellowship and a Sir Henry Wellcome Fellowship [096082]. D.C.B. is the Royal Society Edward Penley Abraham Research Professor.

Author contributions

B.Y.W.C. and D.C.B. conceived and designed the research. B.Y.W.C performed the experiments and analysed the data. M.J.D., A.J.G. and J.H. performed all the LC-MS/MS sample processing and iSPY analysis. B.Y.W.C. and D.C.B. wrote the manuscript.

Conflict of interest

The authors declared that they have no conflict of interest.

References

1. Bartel, D. P. MicroRNAs: target recognition and regulatory functions. *Cell* **136,** 215–33 (2009).
2. Ameres, S. L. & Zamore, P. D. Diversifying microRNA sequence and function. *Nat. Rev. Mol. Cell Biol.* **14,** 475–88 (2013).
3. Izaurralde, E. A role for eIF4AII in microRNA-mediated mRNA silencing. *Nat. Struct. Mol. Biol.* **20,** 543–5 (2013).
4. Filipowicz, W., Bhattacharyya, S. N. & Sonenberg, N. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat. Rev. Genet.* **9,** 102–14 (2008).
5. Iwakawa, H. oki & Tomari, Y. The Functions of MicroRNAs: mRNA Decay and Translational Repression. *Trends Cell Biol.* **25,** 651–665 (2015).
6. Djuranovic, S., Nahvi, A. & Green, R. miRNA-mediated gene silencing by translational repression followed by mRNA deadenylation and decay. *Science* **336,** 237–40 (2012).
7. Bazzini, A. a, Lee, M. T. & Giraldez, A. J. Ribosome profiling shows that miR-430 reduces translation before causing mRNA decay in zebrafish. *Science* **336,** 233–7 (2012).
8. Eichhorn, S. W. *et al.* MRNA Destabilization Is the dominant effect of mammalian microRNAs by the time substantial repression ensues. *Mol. Cell* (2014). doi:10.1016/j.molcel.2014.08.028
9. Guo, H., Ingolia, N. T., Weissman, J. S. & Bartel, D. P. Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* **466,** 835–40 (2010).
10. Brodersen, P. & Voinnet, O. Target Recognition and Mode of Action. *Nat. Rev. Mol. Cell Biol.* **10,** 141–148 (2009).
11. Reis, R. S., Hart-smith, G., Eamens, A. L., Wilkins, M. R. & Waterhouse, P. M. Gene regulation by translational inhibition is determined by Dicer partnering proteins. *Nat. Plants* **1,** 1–6 (2015).
12. Li, S. *et al.* MicroRNAs inhibit the translation of target mRNAs on the endoplasmic reticulum in arabidopsis. *Cell* **153,** 562–574 (2013).
13. Brodersen, P. *et al.* Widespread translational inhibition by plant miRNAs and siRNAs. TL - 320. *Science* **320 VN-,** 1185–1190 (2008).
14. Iwakawa, H. & Tomari, Y. Molecular Insights into microRNA-Mediated Translational Repression in Plants. *Mol. Cell* **52,** 591–601 (2013).
15. Baek, D. *et al.* The impact of microRNAs on protein output. *Nature* **455,** 64–71 (2008).
16. Hendrickson, D. G. *et al.* Concordant regulation of translation and mRNA abundance for hundreds of targets of a human microRNA. *PLoS Biol.* **7,** 25–29 (2009).
17. Molnar, a, Schwach, F., Studholme, D. J., Thuenemann, E. C. & Baulcombe, D. C. miRNAs control gene expression in the single-cell alga Chlamydomonas reinhardtii. *Nature* **447,** 1126–1129 (2007).
18. Valli, A. A. *et al.* Most microRNAs in the single-cell alga Chlamydomonas reinhardtii are produced by Dicer-like 3-mediated cleavage of introns and untranslated regions of coding RNAs. *Genome Res.* **26,** 519–529 (2016).
19. Yamasaki, T. *et al.* Complementarity to an miRNA seed region is sufficient to induce moderate repression of a target transcript in the unicellular green alga Chlamydomonas reinhardtii. *Plant J.* **76,** 1045–1056 (2013).
20. Yamasaki, T., Kim, E. J., Cerutti, H. & Ohama, T. Argonaute3 is a key player

in miRNA-mediated target cleavage and translational repression in Chlamydomonas. *Plant J.* **85,** 258–268 (2016).
21. Chung, B. Y. *et al.* The use of duplex-specific nuclease in ribosome profiling and a user-friendly software package for Ribo-seq data analysis. *Rna* **21,** 1731–1745 (2015).
22. Lewis, B. P., Shih, I., Jones-Rhoades, M. W., Bartel, D. P. & Burge, C. B. Prediction of mammalian microRNA targets. *Cell* **115,** 787–98 (2003).
23. Agarwal, V., Bell, G. W., Nam, J. W. & Bartel, D. P. Predicting effective microRNA target sites in mammalian mRNAs. *Elife* **4,** (2015).
24. Schirle, N. T., Sheu-Gruttadauria, J. & MacRae, I. J. Structural basis for microRNA targeting. *Science (80-.).* **346,** 608–613 (2014).
25. Mallory, A. C. *et al.* MicroRNA control of PHABULOSA in leaf development: importance of pairing to the microRNA 5' region. *EMBO J.* **23,** 3356–64 (2004).
26. Gao, X. *et al.* MicroRNAs modulate adaption to multiple abiotic stresses in Chlamydomonas reinhardtii. *Sci. Rep.* **6,** 38228 (2016).
27. Korostelev, A., Trakhanov, S., Laurberg, M. & Noller, H. F. Crystal Structure of a 70S Ribosome-tRNA Complex Reveals Functional Interactions and Rearrangements. *Cell* **126,** 1065–1077 (2006).
28. Qu, X. *et al.* The ribosome uses two active mechanisms to unwind messenger RNA during translation. *Nature* **475,** 118–121 (2011).
29. Gerashchenko, M. V. & Gladyshev, V. N. Translation inhibitors cause abnormalities in ribosome profiling experiments. *Nucleic Acids Res.* **42,** (2014).
30. Gutteridge, A. *et al.* Nutrient control of eukaryote cell growth: a systems biology study in yeast. *BMC Biol.* **8,** 68 (2010).
31. Marondedze, C. *et al.* A Quantitative Phosphoproteome Analysis of cGMP-Dependent Cellular Responses in Arabidopsis thaliana. *Mol. Plant* **9,** 621–623 (2016).
32. Brosch, M., Yu, L., Hubbard, T. & Choudhary, J. Accurate and sensitive peptide identification with mascot percolator. *J. Proteome Res.* **8,** 3176–3181 (2009).
33. Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. S. & Weissman, J. S. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* **324,** 218–23 (2009).
34. Hardcastle, T. J. & Kelly, K. A. baySeq: Empirical Bayesian methods for identifying differential expression in sequence count data. *BMC Bioinformatics* **11,** 422 (2010).
35. Bonnet, E., He, Y., Billiau, K. & van de Peer, Y. TAPIR, a web server for the prediction of plant microRNA targets, including target mimics. *Bioinformatics* **26,** 1566–1568 (2010).

Figure Legends

Figure 1. Ribosome profiling data.
(A, B) Mapping the 5' ends of ribosome protected fragments (RPFs) and corresponding RNA-Seq respectively, as a function of read size class (nt), within nucleus-encoded coding ORFs. Red, green and blue bars indicate the proportion of reads that map to codon positions 0, 1 and 2 (respectively).
(C, D) 5' end positions of 27-nt RPFs relative to start and stop codons (nt). Reads

were derived from strain C*dcl3* and *dcl3-1* (respectively) and summed over all transcripts. Phasing is indicated using the same colors as in panels A and B.
(E, F) 5' end positions of all RNA-seq reads relative to start and stop codons (nt). Reads were derived from strain C*dcl3* and *dcl3-1* (respectively) and summed over all transcripts. Phasing is indicated using the same colors as in panels A and B.

Figure 2. Distribution of 8mer target sites.
(A) Venn diagram showing number of transcripts predicted to be targeted with the 8mer rule.
(B) Proportion of 8mer target sites that also have at least 50% complementarity from nucleotides 11-21 of the miRNA

Figure 3. miRNA downregulates gene expression primarily through mRNA destabilization by CDS targeting.
(A) Cumulative distributions of ΔTE (left), ΔRPF (middle) and ΔRA (right) $\log_2$ fold changes in *dcl3-1* relative to C*dcl3*. Colors correspond to genes containing predicted 8mer miRNA target sites exclusively in the 5'UTR (orange), CDS (green), 3'UTR (blue), or no targets (black).
(B) Bar graph of differences between area under cumulative distribution of mRNA containing target sites and non-target containing mRNAs (5'UTR, CDS and 3'UTR in orange, green and blue, respectively). Significance (K.S. test) of the differences are indicated above each bar; p-values less than or equal to 0.01 are highlighted in red.
(C-D) Bar graph of differences between area under cumulative distribution of mRNA containing 1 (red), 2 (blue), 3 (purple) or 4 or more (green) CDS-exclusive target sites and non-target containing mRNAs. Significance (K.S. test) of the differences are indicated above each bar; p-values less than or equal to 0.01 are highlighted in red.
(E) Normalised miRNA abundances of C*dcl3* (in three biological replicates).
(F) Cumulative distributions (top) and significance (bottom; the red dotted line indicates p-value of 0.01) of ΔTE (left), ΔRPF (middle) and ΔRA (right) $\log_2$ fold changes for mRNAs containing miR-C89 target sites exclusively within the CDS (green) or 3'UTR (blue) (sample sizes 141 and 25, respectively). 5'UTR-exclusive targets were omitted due to low sample size.

Figure 4. Effects of miRNAs on TE and RA.
(A) Correspondence between TE and RA fold-changes between *dcl3-1* and C*dcl3* for nuclear-encoded genes containing miRNA target sites exclusively within the CDS (except DCL3, which was included as a marker). 80S, chloroplast and mitochondria ribosomal proteins are in orange, green and red, respectively.
(B-C) Histograms of 5' end positions of normalized RPF (colored, left-axis) and RNA-Seq (grey, right-axis) 27-nt reads mapped to genes with high differential TE: ribosomal proteins rpL14 and Cre16.g675200. The top (green title) and bottom (red title) graphs are derived from either C*dcl3* or *dcl3*-1, respectively. The colored horizontal line indicates the riboSeqR *de novo*-defined ORF; positions of potential miRNA target sites are annotated.
(D) Histogram of 5' end positions of normalized RPF (colored, left-axis) and RNA-Seq (grey, right-axis) 27-nt reads mapped to DCL3 transcripts. The blue horizontal line indicates the CDS (612-12,830 nt). The schematic below the plot shows the domain organization of DCL3 which contains two DEAD/DEAH box helicase domains (light and dark red boxes), a helicase domain (purple box), a proline-rich domain (orange box) and two ribonuclease III domains a and b (light and dark green

boxes, respectively). The thick grey line and the corresponding red arrow below indicate the hygromycin insertion site (nt 10,193).
(E) Correspondence between TE and RA fold-changes between *ago3-25* and wild type CC-1883 for nuclear-encoded genes containing miRNA target sites exclusively within the CDS. Nuclear-encoded 80S and chloroplast ribosomal proteins are in orange and green, respectively. Mitochondrial ribosomal proteins are not shown due to low level of detection in the dataset.

Supplementary Figure 1. Experimental workflow
Three independent single colonies from freshly streaked *Chlamydomonas dcl3-1* (green) or complement (blue) were inoculated into 50 mL of TAP media and grown until OD750 = 0.6 (mid-log phase). 0.25 mL of each culture was used for N15 incorporation for whole cell proteomics and the remaining culture was used to sub-culture 750 mL of TAP for ribosome profiling.

Supplementary Figure 2. Reproducibility of TE, ribosome profiling, RNA-Seq and N15 Proteomics
(A)-(E) Correspondence between biological triplicates for C*dcl3*, *dcl3-1* and replicates for wt, Parent and *ago3-25*, respectively.

Supplementary Table 1: Number of reads mapping to nuclear-encoded transcripts for each library (Phytozome 281).

Supplementary Figure 3: Generation of precise ribosome profiling data:
(A) Histogram of positions for all biological triplicates to which the 5' ends of ribosome profile footprints (RPFs) and corresponding RNA-Seq reads map, respectively, as a function of read size class (nt), for reads mapping to the interior region of nuclear-encoded coding ORFs. Red, green and blue bars indicate the proportion of reads that map to codon positions 0, 1 and 2 (respectively).
(B) Histogram of 5' end positions of 27 and 28-nt RPFs and RNA-seq (all sizes) relative to start and stop codons for all biological triplicates. Reads were derived from C*dcl3* or *dcl3-1* (respectively) and summed over all transcripts. Phasing is indicated using the same colors as in supplementary figure 3A.
(C) 27-nt reads mapped to DCL3 transcripts in all biological triplicates. The blue horizontal line indicates the CDS (612-12,830 nt). The schematic below the plot shows the domain organisation of DCL3 which contains two DEAD/DEAH box helicase domains (light and dark red boxes), a Helicase domain (purple box), a proline-rich domain (orange box) and two Ribonuclease III domains a and b (light and dark green boxes, respectively). The thin grey line and the corresponding red arrow indicates the Hygromycin insertion site (nt 10,193).
(D) 27-nt reads mapped to AGO3 transcripts in all biological replicates. The red horizontal line indicates the CDS. The schematic below the plot shows the domain organisation of AGO3 which contains N-terminal domain (blue), L1 and L2 (light and dark yellow, respectively), PAZ domain (purple), MID domain (green) and the PIWI domain (grey) (Chung *et. al.* 2017 submitted).
(E) Correlation of TE, RPF and RNA (averaged over biological repeats) between *dcl3-1* and C*dcl3* for all expressed genes. Blue lines represent a perfect correlation. Spearman correlation coefficients are indicated in bottom right corners.
(F) Typical polysome profile of undigested (black) and digested (green) lysates in this study.

(E) Histogram of 5' end positions of 27-nt RPFs relative to start and stop codons for C*dcl3* culture pre-treated with 100 µg/ml cycloheximide for 5 min prior to harvesting for ribosome profiling. Phasing is indicated using the same colors as in supplementary figure 3A.

Supplementary Table 2: Re-annotation of miRNA precursor-containing mRNAs.

Supplementary Figure 4. DCL3-dependent processing of miRNA down-regulates translation efficiency

(A) Scatter plot of $\log_2$ fold changes of all mRNAs for TE and RA fold-changes between *dcl3-1* and C*dcl3*. New annotation for precursor-containing transcripts: yellow circle = precursor-containing CDS, white circles = precursor-containing 3'UTRs, orange circles = precursor-containing introns, white circles with red outlines = transcripts previously annotated as non-coding transcripts but which are in fact coding and contain a miRNA precursor in the 3'UTR. Pearson correlation = 0.072 and 0.486 for intron- and exon-containing transcripts respectively.
(B)-(E) Histogram of normalised 5' end positions of 27-nt RPFs relative to start and stop codons (colour) and corresponding RNA-seq reads (grey) for miRNA-precursor containing transcripts. Reads were derived from the complement or DCL3 mutant (top and bottom in biological triplicates, respectively) and summed over all transcripts.

Supplementary Figure 5:

(A) Bar graph of differences between area under cumulative distribution of mRNA containing 1 (red), 2 (blue), 3 (purple) or 4 or more (green) 5'UTR, CDS or 3'UTR-exclusive target sites and non-target containing mRNAs. Significance (K.S. test) of the differences are indicated above each bar; p-values less than or equal to 0.01 are highlighted in red. Due to lower sequencing coverage, fewer mRNAs in the *ago3-25* and Parental strains passed the detection threshold, thus 5' and 3' UTR-exclusive targets containing three target sites were combined with those containing four or more target sites.
(B) Cumulative *dcl3-1* relative to C*dcl3* $\log_2$ fold-change distributions of ΔTE, ΔRPF, ΔRA and ΔProtein for genes with both NGS and proteomic support and with 0 (black), 1 (red), 2-3 (blue) or 4 or more (green) target sites. K.S. p-values are shown in the table below.
(C) Correlation of TE, RPF and RNA (averaged over biological repeats) between *dcl3-1* and C*dcl3* for all expressed mRNAs (black), and for mRNAs with predicted target sites (orange). Top, middle and bottom panels represents, respectively, all CDS-exclusive 8mer targets, CDS-exclusive targets with a higher degree of pairing (i.e. 2-15 nt base pairing, allowing 1 mismatch), and CDS-exclusive targets based on the plant rule. The blue lines represent a perfect correlation.
(D) Normalized mRNA expression in wild-type and *ago3-25*. Candidates are from box C of Figure 4A organized based on degree of repression in C*dcl3* relative to *dcl3-1*. Normalized average expression and standard deviation are based on biological replicates.

Supplementary Figure 6: miRNA quantification

Absolute (top) and relative (bottom) quantification for all known positive-strand miRNA reads detected in all corresponding sRNA-seq libraries. Sequencing and miRNA alignment statistics for each library are in the table below.

Supplementary table 3

Lists of mRNAs lying within boxes A, A', B, B', C and C' (Figure 4A) and their respective annotations. Annotations associated with the 80S translation machinery are highlighted in green, and other RNA binding proteins in red. Messenger RNAs with detectable protein in the N15 proteomics data are highlighted in blue.

Supplementary Figure 7

(A) Top panels show cumulative distributions of ΔTE (left), ΔRPF (middle) and ΔRA (right) $\log_2$ fold changes for CDS-exclusive targets in *dcl3-1* relative to C*dcl3*. Colors indicate CDS-exclusive targets (green), CDS-exclusive targets excluding strongly differentially expressed mRNAs (i.e. mRNAs in box A, A', B, B', C and C' in Figure 4A) (purple), and mRNAs without target sites (black). The bar graphs in the bottom panel show differences between areas under the corresponding cumulative distributions of target-site containing mRNA and non targets. Significance (K.S. test) of the differences are indicated above each bar; p-values less than or equal to 0.01 are highlighted in red.

(B) Scatter plot of $\log_2$ fold changes of all CDS-exclusive targets for ΔTE and ΔRA between *dcl3-1* and C*dcl3*. miR-C89 CDS-exclusive targets are highlighted in purple.

(C) Scatter plot of $\log_2$ fold changes of all CDS-exclusive targets for ΔTE and ΔRA between *ago3-25* and the Parental strain. 80S and chloroplast ribosomal proteins are in orange and green, respectively.

Supplementary Figure 8

Correspondence between ΔTE and ΔRA $\log_2$ fold-changes after deleting miR-233 in mouse neutrophil cells (A), after introducing miR-1 to HEK293 cells (B), and after introducing miR-155 to HEK293 cells (C). Fold-change data were obtained from[9].

Figure 1

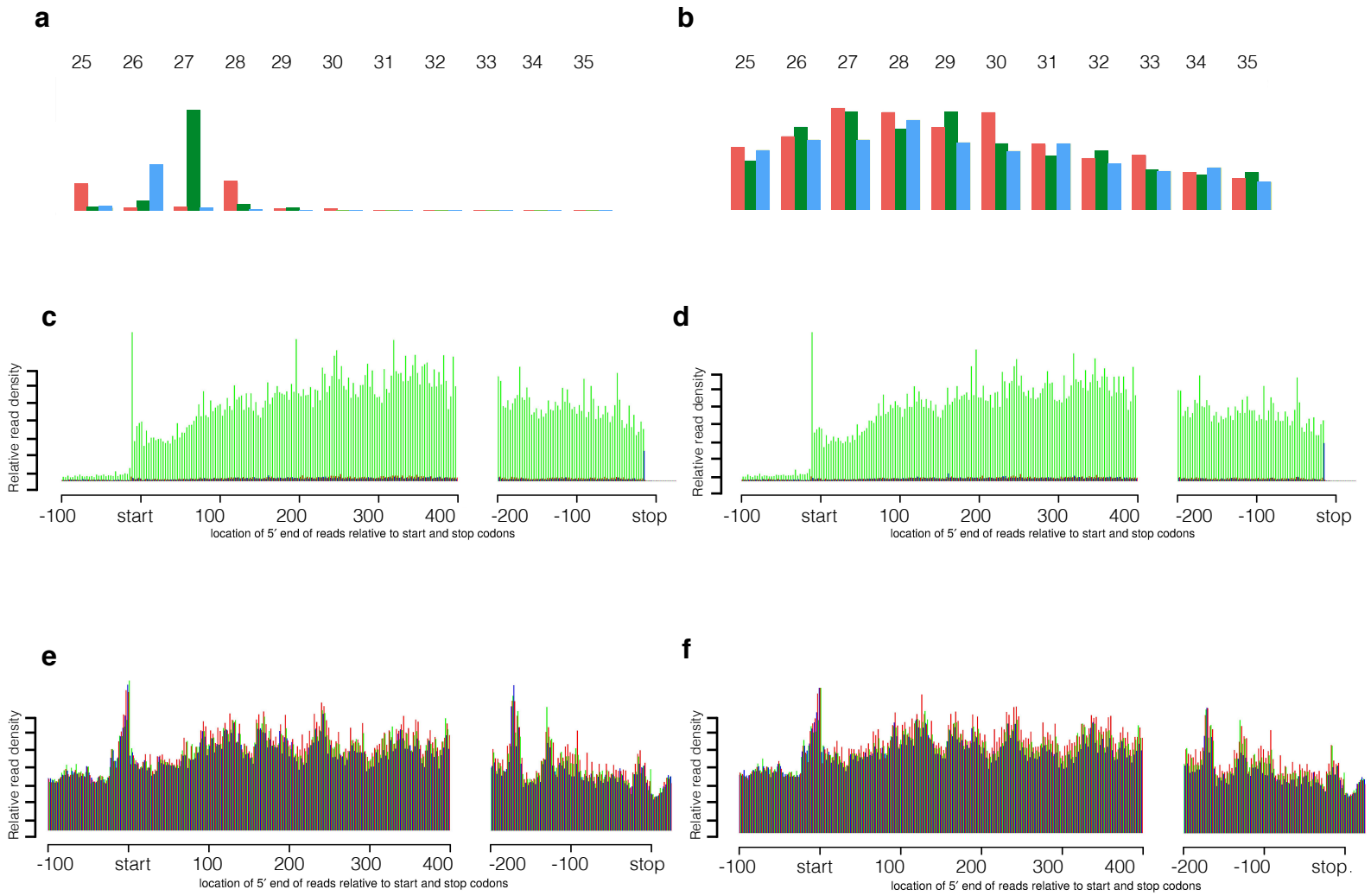

a
25 26 27 28 29 30 31 32 33 34 35
b
25 26 27 28 29 30 31 32 33 34 35
c
Relative read density
-100 start 100 200 300 400 -200 -100 stop
location of 5′ end of reads relative to start and stop codons
d
Relative read density
-100 start 100 200 300 400 -200 -100 stop
location of 5′ end of reads relative to start and stop codons
e
Relative read density
-100 start 100 200 300 400 -200 -100 stop
location of 5′ end of reads relative to start and stop codons
f
Relative read density
-100 start 100 200 300 400 -200 -100 stop
location of 5′ end of reads relative to start and stop codons

Figure 2

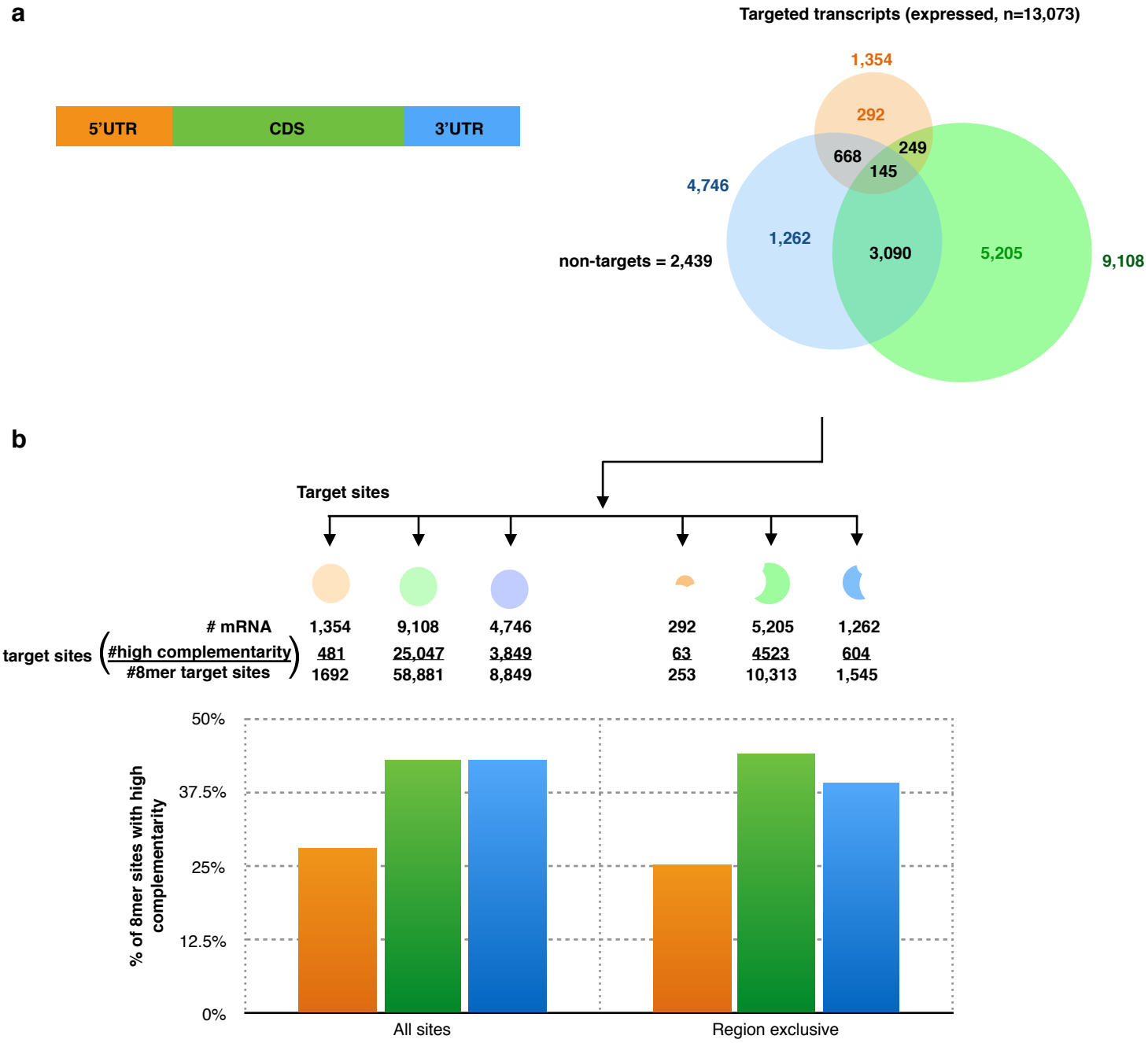
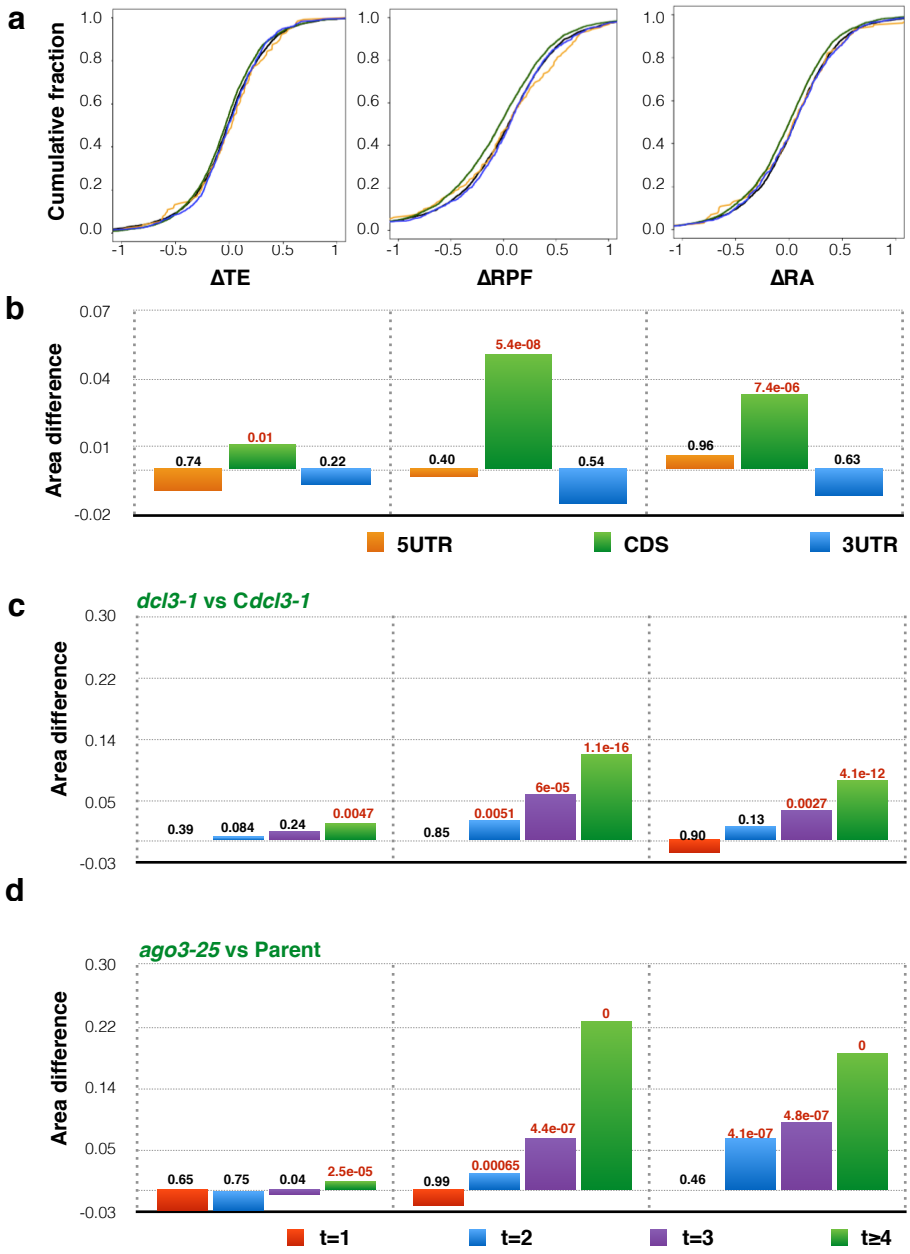

a
Targeted transcripts (expressed, n=13,073)
5'UTR
CDS
3'UTR
1,354
292
668
249
145
4,746
1,262
3,090
5,205
9,108
non-targets = 2,439
b
Target sites
mRNA 1,354 9,108 4,746 292 5,205 1,262
target sites (#high complementarity / #8mer target sites)
481 / 1692
25,047 / 58,881
3,849 / 8,849
63 / 253
4523 / 10,313
604 / 1,545
% of 8mer sites with high complementarity
50%
37.5%
25%
12.5%
0%
All sites
Region exclusive

Figure 3


a
Cumulative fraction
1.0
0.8
0.6
0.4
0.2
0.0
-1
-0.5
0.0
0.5
1
ΔTE
ΔRPF
ΔRA
b
Area difference
0.07
0.04
0.01
-0.02
0.74
0.01
0.22
0.40
5.4e-08
0.54
0.96
7.4e-06
0.63
5UTR
CDS
3UTR
c
dcl3-1 vs Cdcl3-1
Area difference
0.30
0.22
0.14
0.05
-0.03
0.39
0.084
0.24
0.0047
0.85
0.0051
6e-05
1.1e-16
0.90
0.13
0.0027
4.1e-12
d
ago3-25 vs Parent
Area difference
0.30
0.22
0.14
0.05
-0.03
0.65
0.75
0.04
2.5e-05
0.99
0.00065
4.4e-07
0
0.46
4.1e-07
4.8e-07
0
t=1
t=2
t=3
t≥4

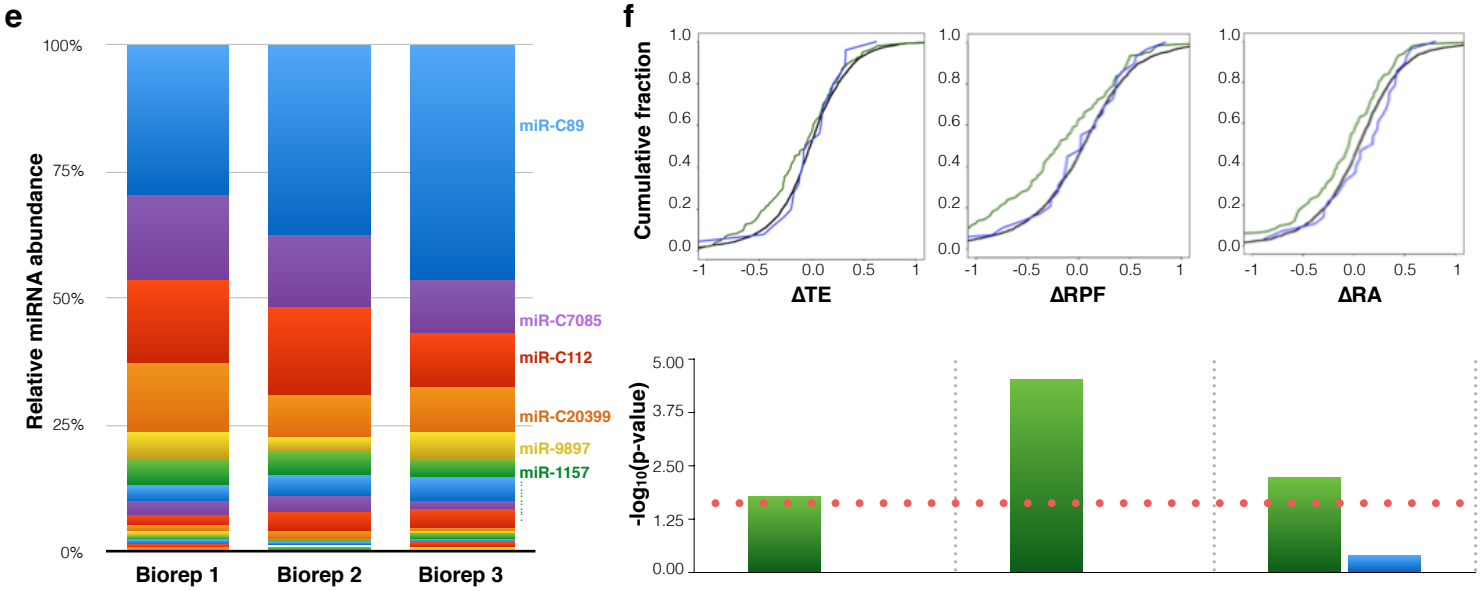
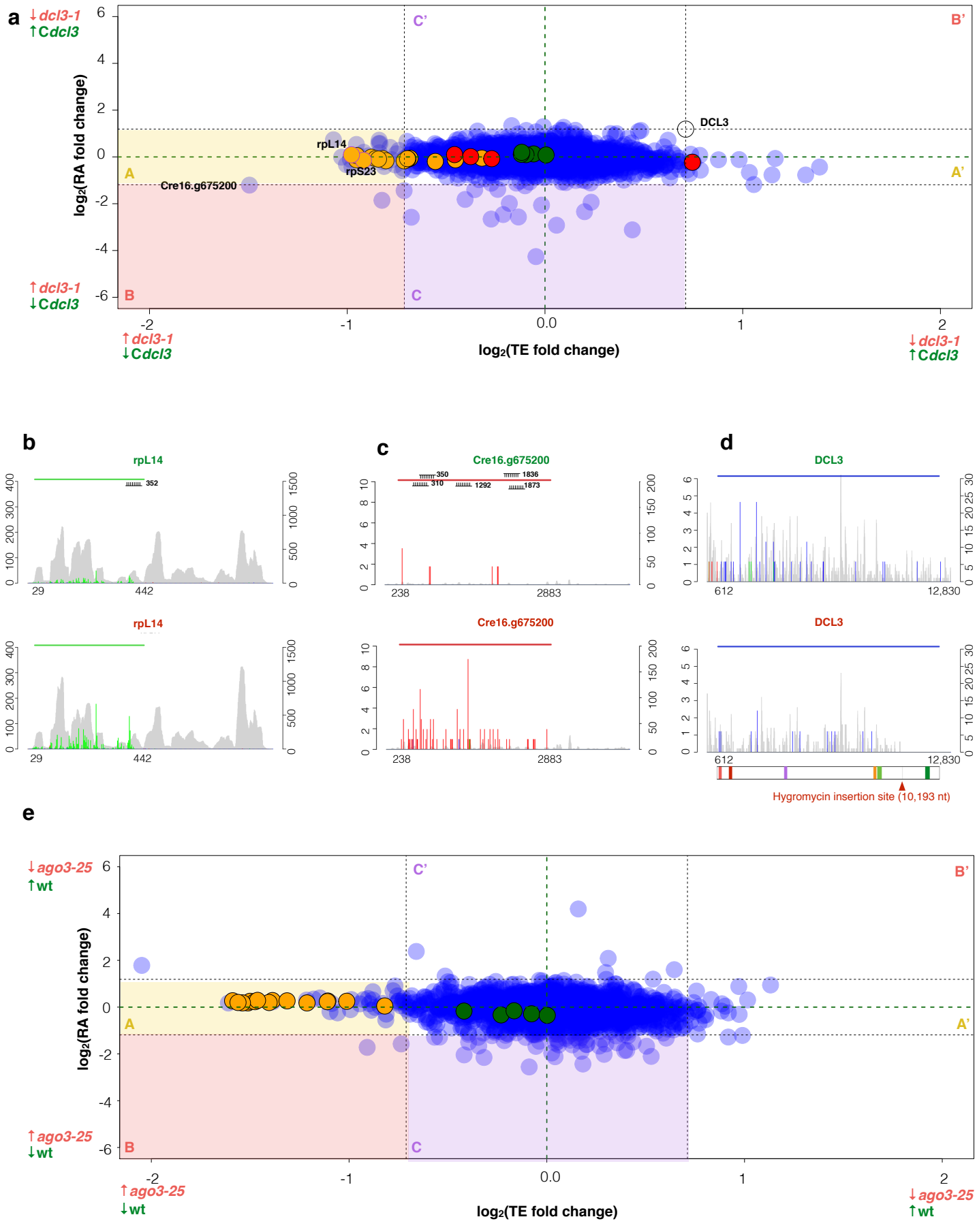

e
100%
75%
50%
25%
0%
Relative miRNA abundance
miR-C89
miR-C7085
miR-C112
miR-C20399
miR-9897
miR-1157
Biorep 1
Biorep 2
Biorep 3
f
Cumulative fraction
1.0
0.8
0.6
0.4
0.2
0.0
-1
-0.5
0.0
0.5
1
ΔTE
ΔRPF
ΔRA
-log10(p-value)
5.00
3.75
2.50
1.25
0.00

Figure 4

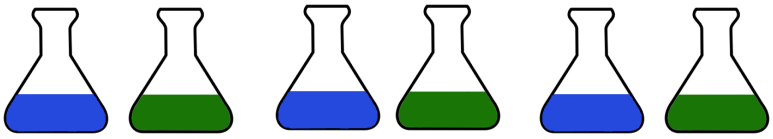

a
↓dcl3-1
↑Cdcl3
B'
C'
A
A'
B
C
DCL3
rpL14
rpS23
Cre16.g675200
log2(RA fold change)
log2(TE fold change)
↑dcl3-1
↓Cdcl3
↓dcl3-1
↑Cdcl3
b
rpL14
c
Cre16.g675200
d
DCL3
Hygromycin insertion site (10,193 nt)
e
↓ago3-25
↑wt
↑ago3-25
↓wt

Supplementary Figure 1

Inoculate three freshly streaked single colony to 50 ml of media

Ribosome profiling and RNA-Seq

Grow till OD750=0.6 (mid-log phase). Sub-culture into 750ml media



OD750=0.6

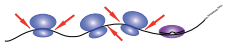
filter cells > snap freeze in liquid nitrogen

Pulverise in liquid nitrogen with buffer containing cycloheximide

Thaw powder on ice

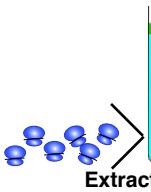
Clarify lysate, discard pellet and proceed with supernatants

RNase1 digest



Total RNA extraction

8,000 rpm, 18hr, 4C, 35% sucrose cushion



rRNA subtraction by ribo-zero

Extract RNA from pellet

Alkaline hydrolysis

Size selection (26-34nt)

3' adaptor ligation

5' adaptor ligation

RT-PCR

PCR

2X DSN treatment to remove rRNA

Pool libraries into one tube, sequence all in four lanes of HiSeq2000

N15 quantitative proteomics

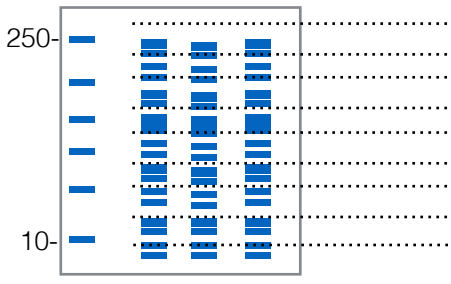
Inoculate 0.25ml into 20ml of either 14N or 15N media

Subculture 0.25ml OD750=0.6, 5 more times for full N15 incorporation

Mix equal number of cells (Complement N14 + Mutant N15)

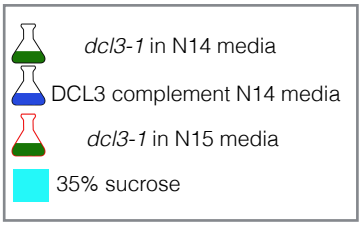
TCA-acetone protein extraction

Denaturing protein gel > Stain > Divide lanes into 16 size fractions (between 250 and 10 kDa)

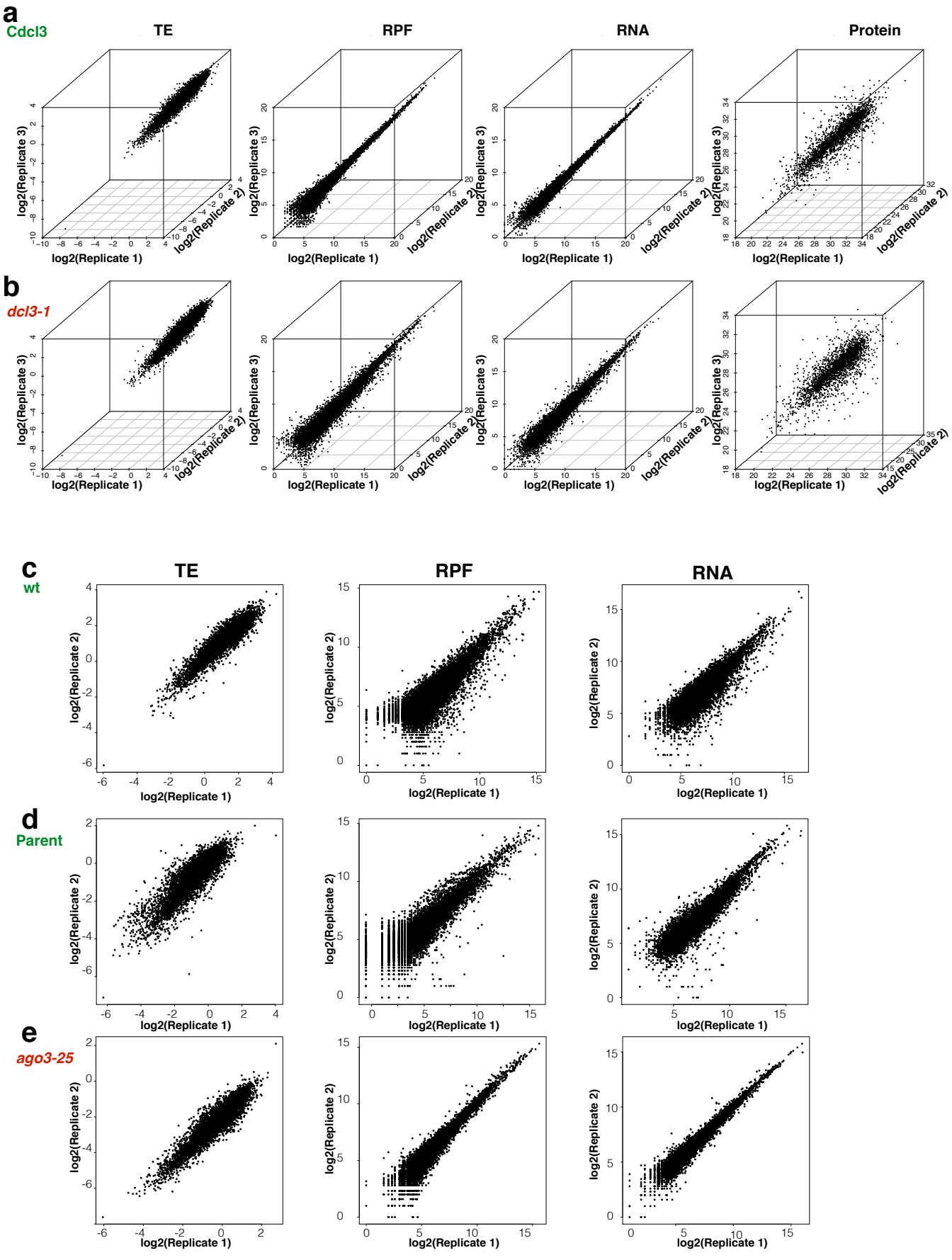


Trypsin digest > LC-MS etc...

Figure legends



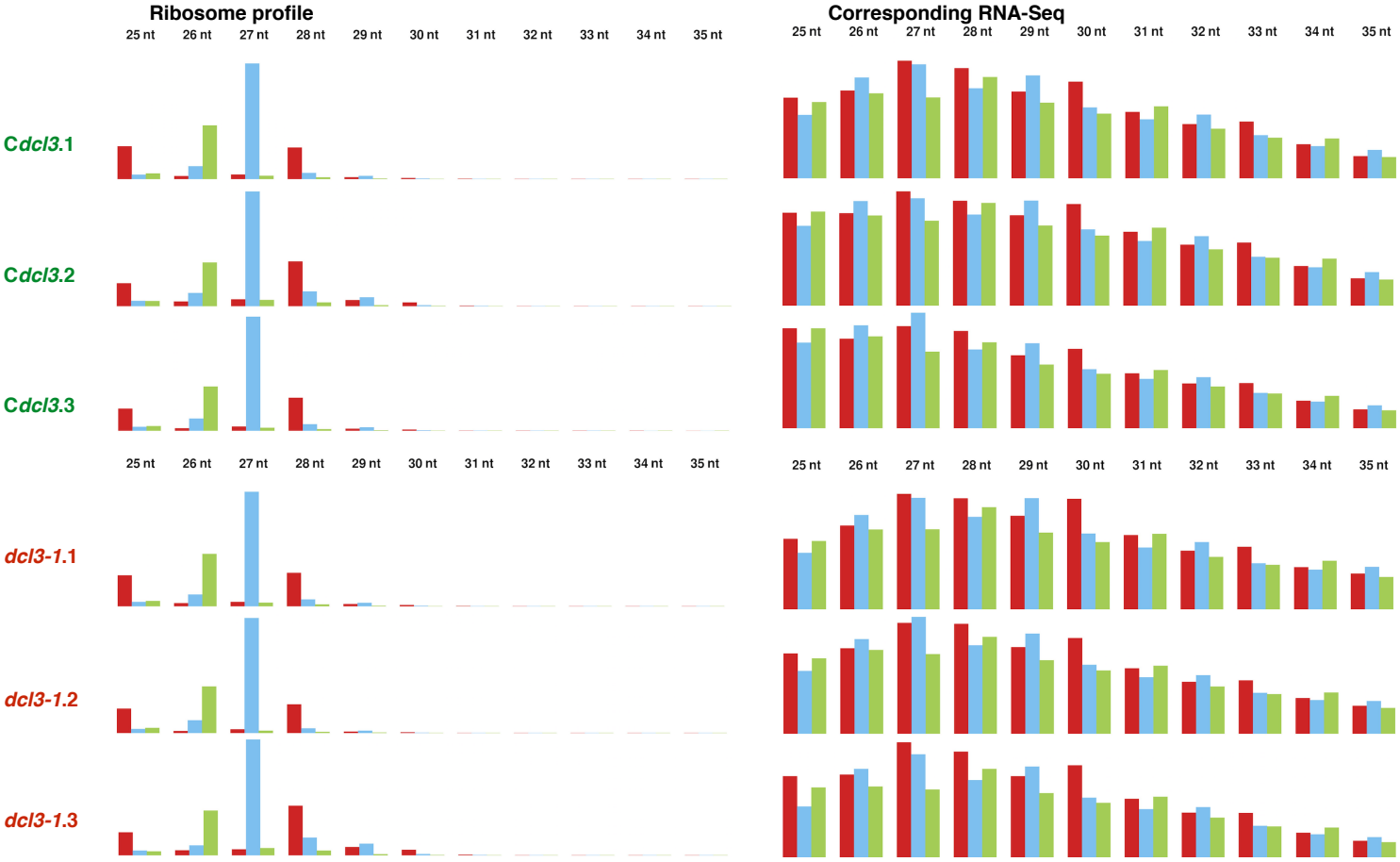
Supplementary Figure 2



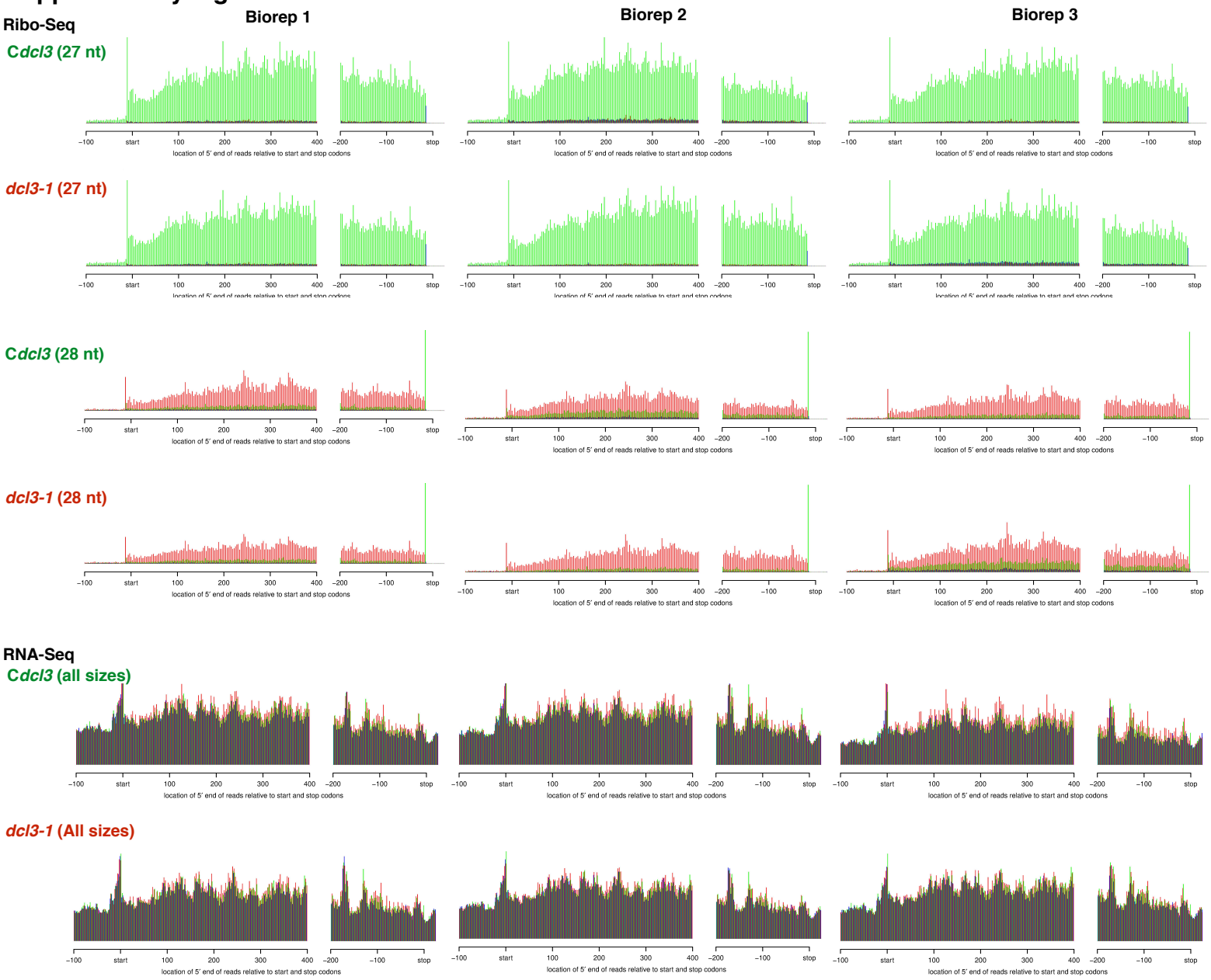
Supplementary table 1

		biorep 1	biorep 2	biorep 3
RiboSeq	Complement	368,613	291,373	444,768
	DCL3	515,717	461,013	590,953
RNA-seq	Complement	908,865	1,114,867	1,223,427
	DCL3	1,166,393	1,183,071	679,375

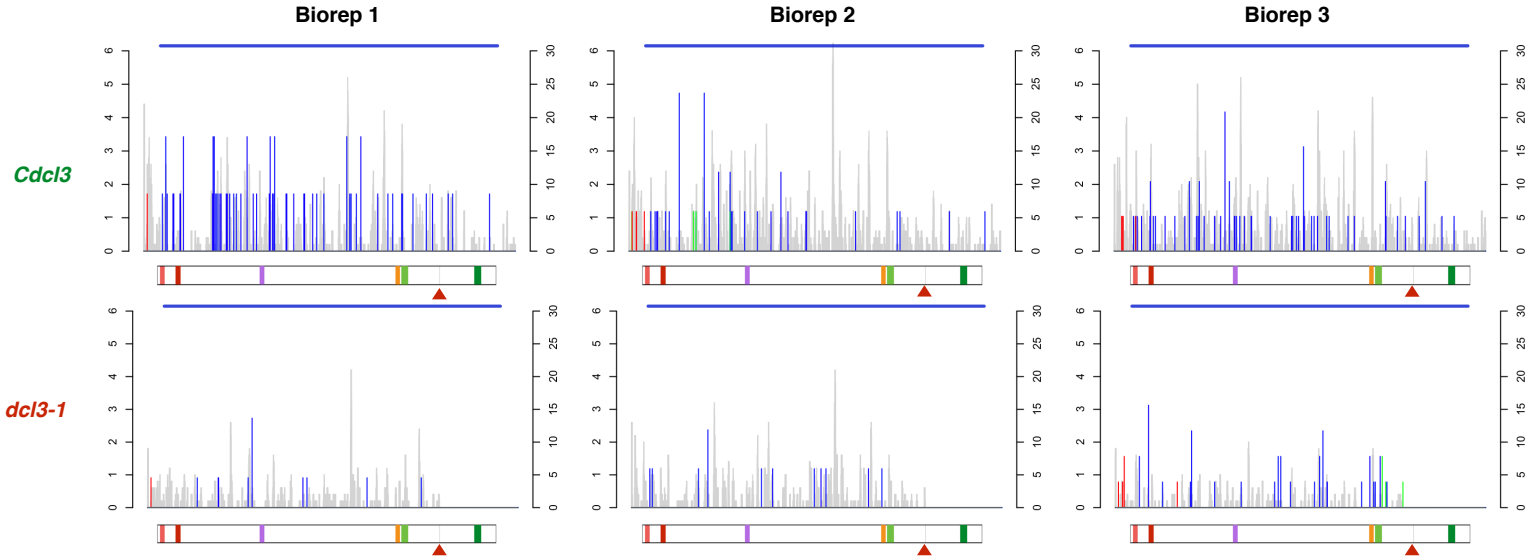
Supplementary Figure 3a



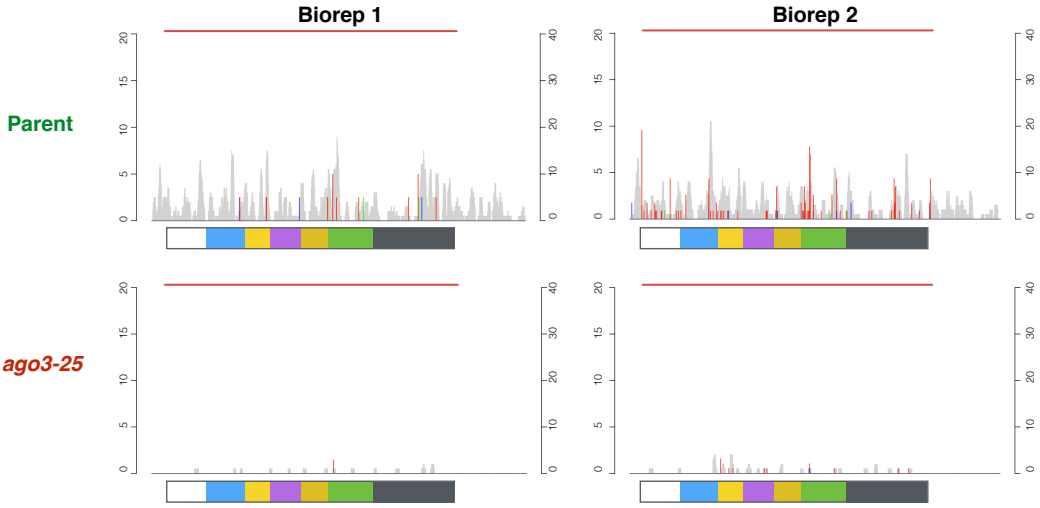
Supplementary Figure 3b



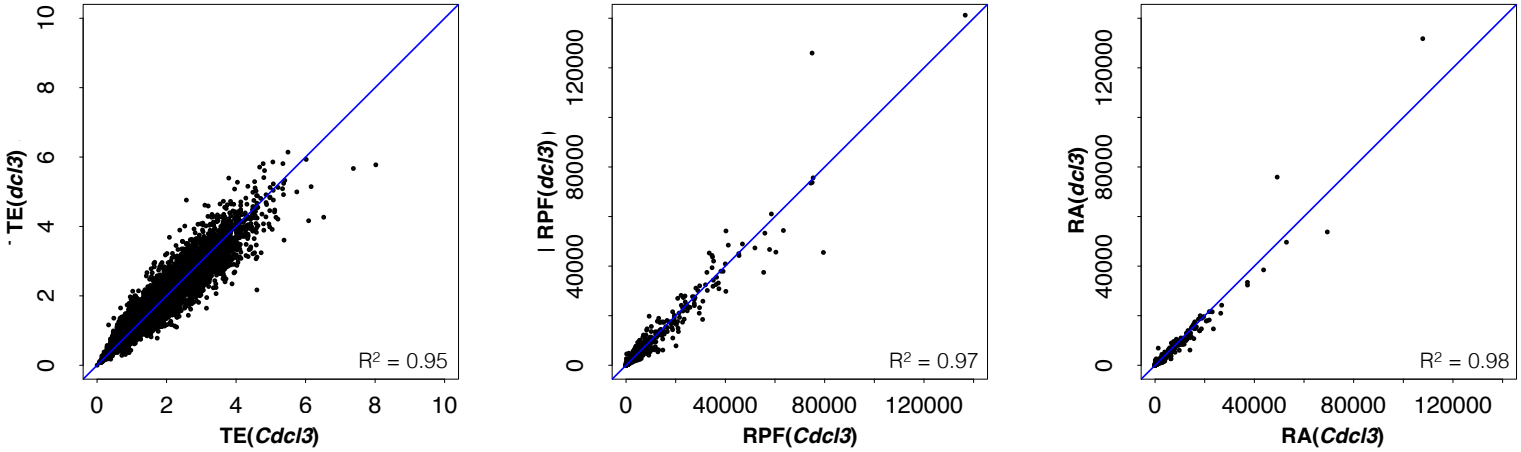
Supplementary Figure 3c



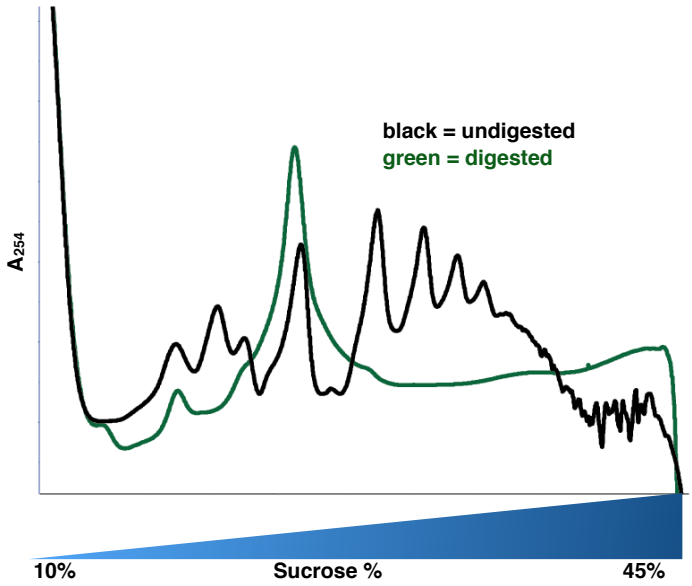
Supplementary Figure 3d



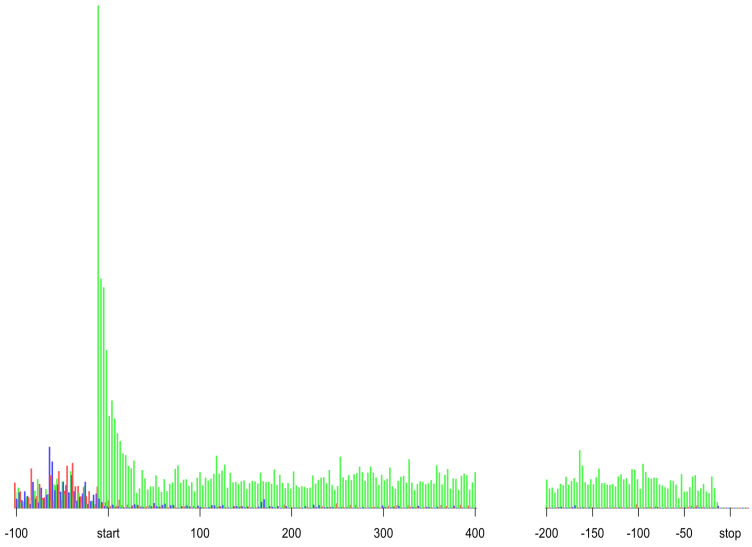
Supplementary Figure 3e



Supplementary Figure 3f



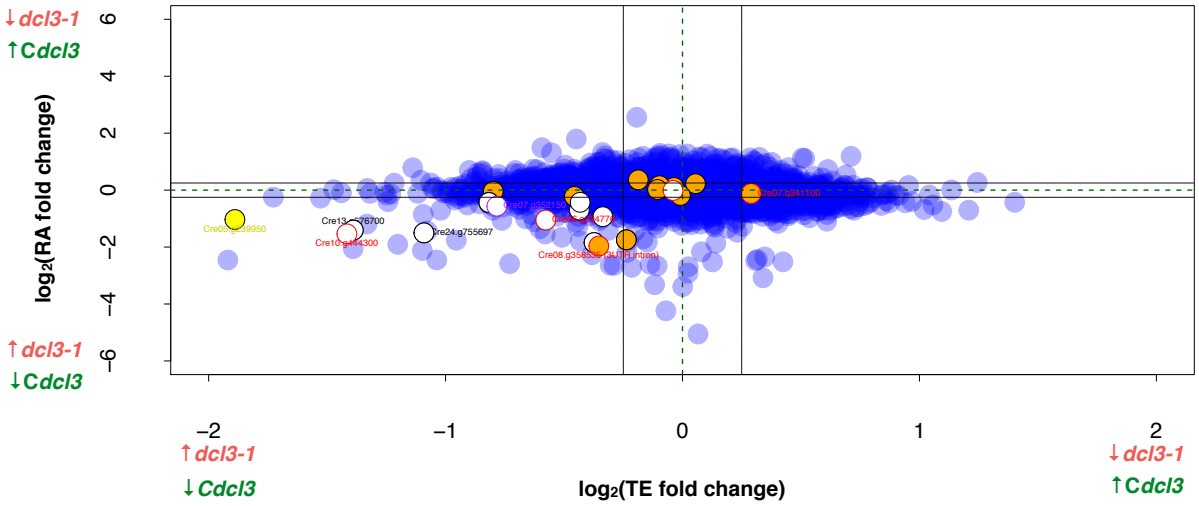
Supplementary Figure 3g



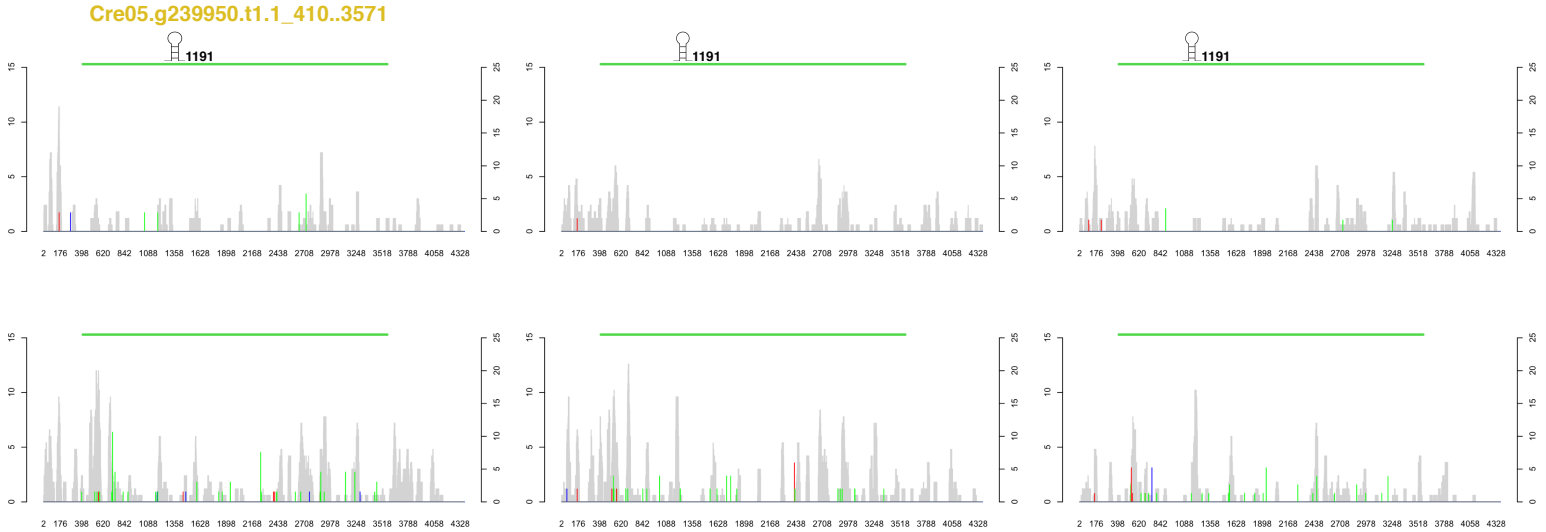
Supplementary Table 2

CDS-exons	CDS-Introns	3'UTR-exon	3'UTR-intron
Cre05.g239950	Cre04.g229050	Cre14.g623850	Cre08.g358535
	Cre01.g035500	Cre02.g089850	
	Cre04.g225700	Cre02.g143427	
	Cre06.g274550	Cre02.g143527	
	Cre06.g296983	Cre03.g195950	
	Cre07.g328400	Cre05.g242301	
	Cre07.g354150	Cre10.g465000	
	Cre12.g537671	Cre13.g576700	
	Cre14.g629200	Cre13.g585175	
	Cre01.g035500	Cre13.g585200	
	Cre02.g143327	Cre16.g694950	
	Cre04.g217925	Cre02.g143527	
	Cre04.g229050	Cre24.g755697	
	Cre08.g358537	Cre02.g143527	
	Cre09.g406983	Cre13.g576700	
	Cre16.g647602	Cre16.g694950	
	Cre07.g341100	Cre24.g755697	
		Cre07.g352150	
		Cre06.g294776	
		Cre10.g444300	
		Cre01.g051050	

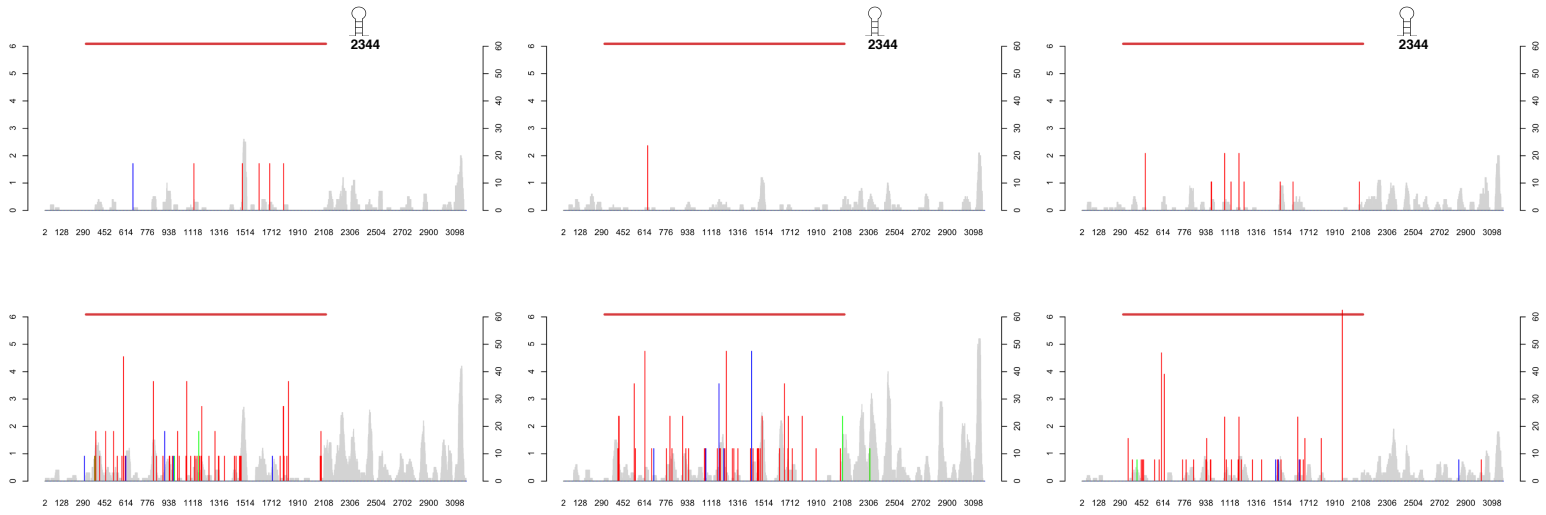
Supplementary Figure 4a



Supplementary Figure 4b

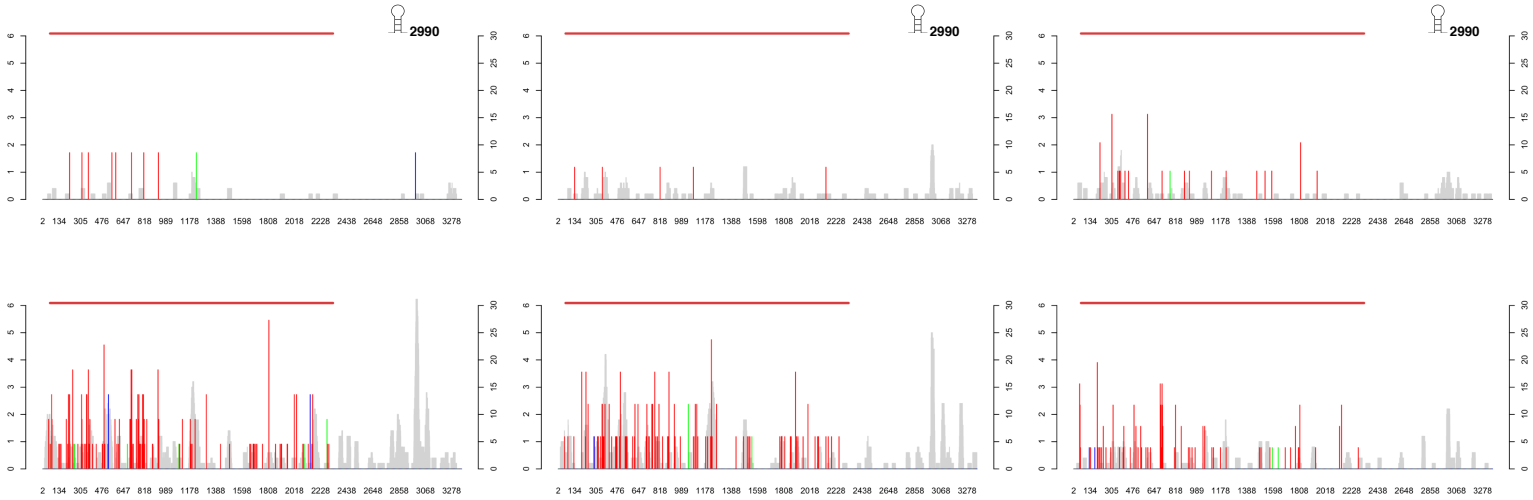


Supplementary Figure 4c



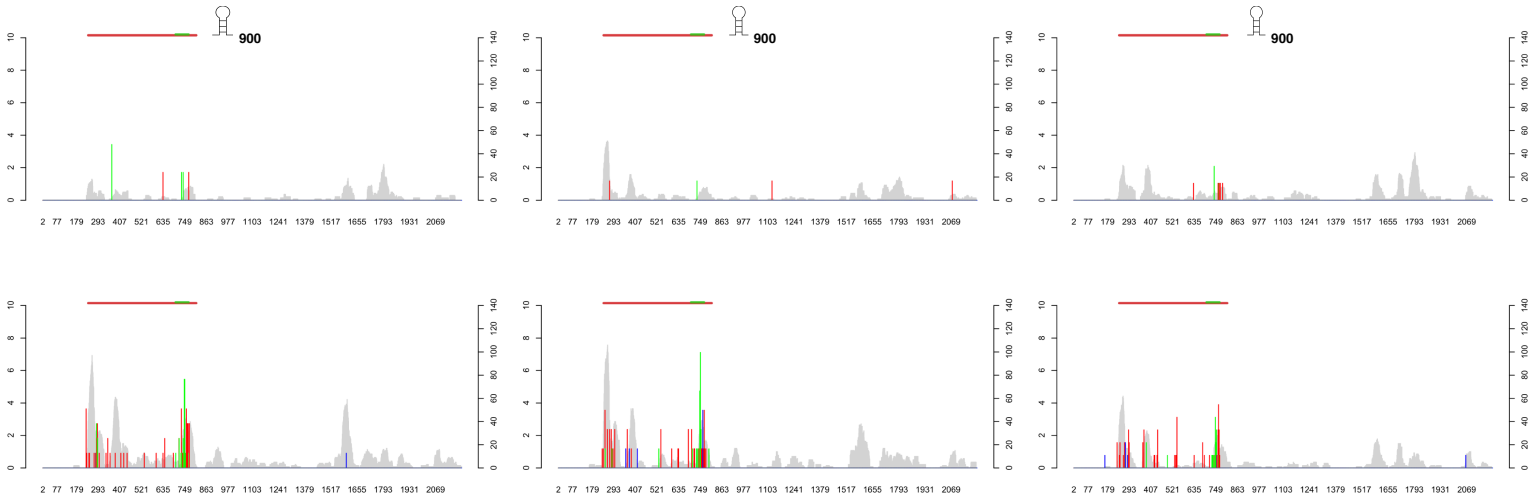
Supplementary Figure 4d

Cre24.g755697.t1.1_136..2328

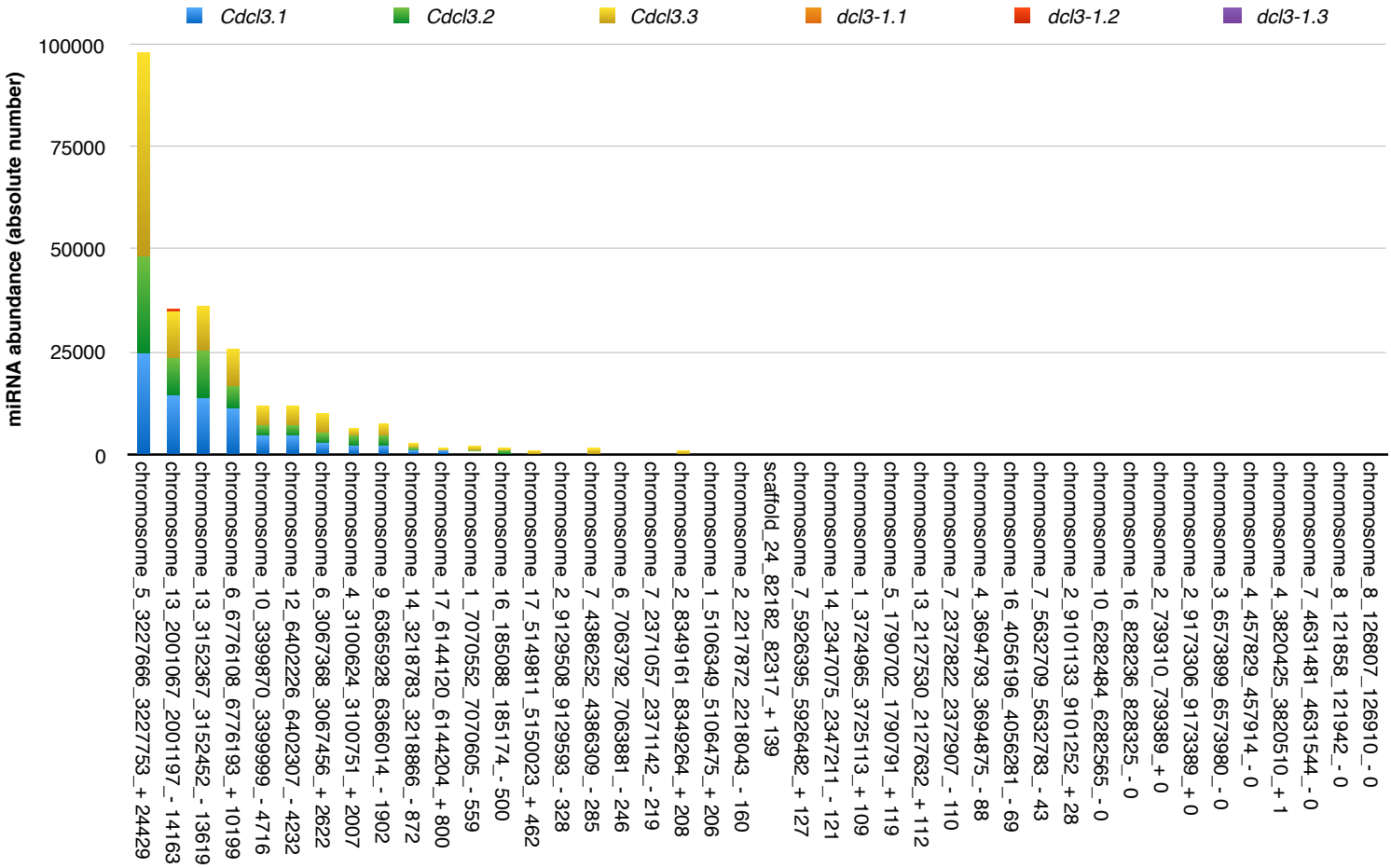


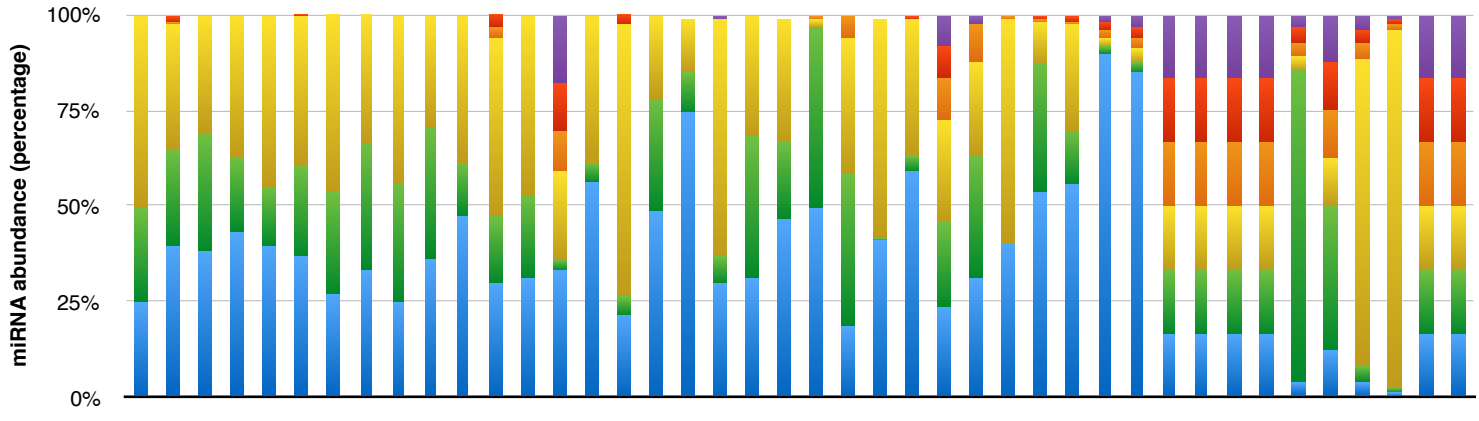
Supplementary Figure 4e

Cre10.g444300.t1.2_241..810



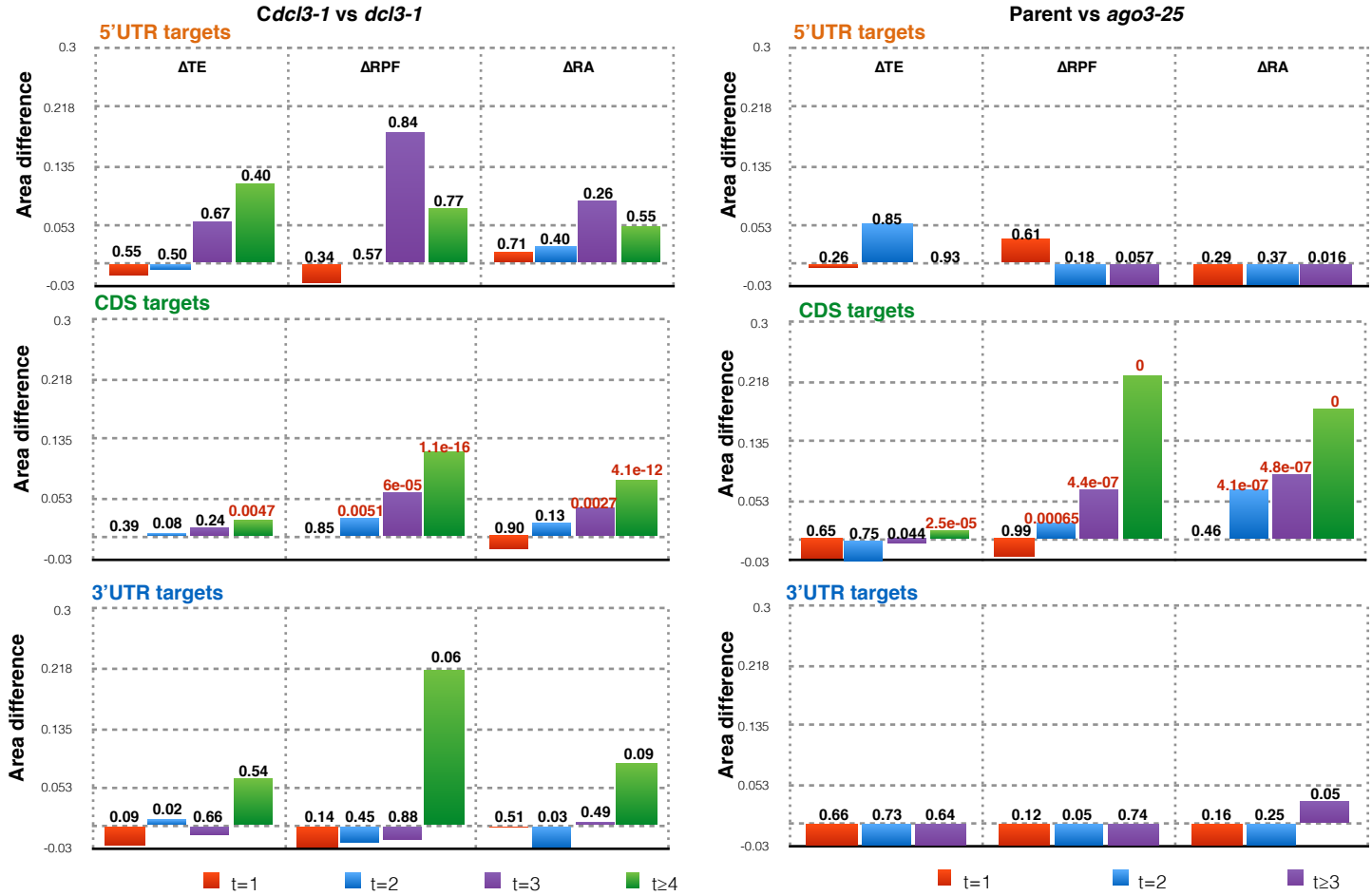
Supplementary Figure 5



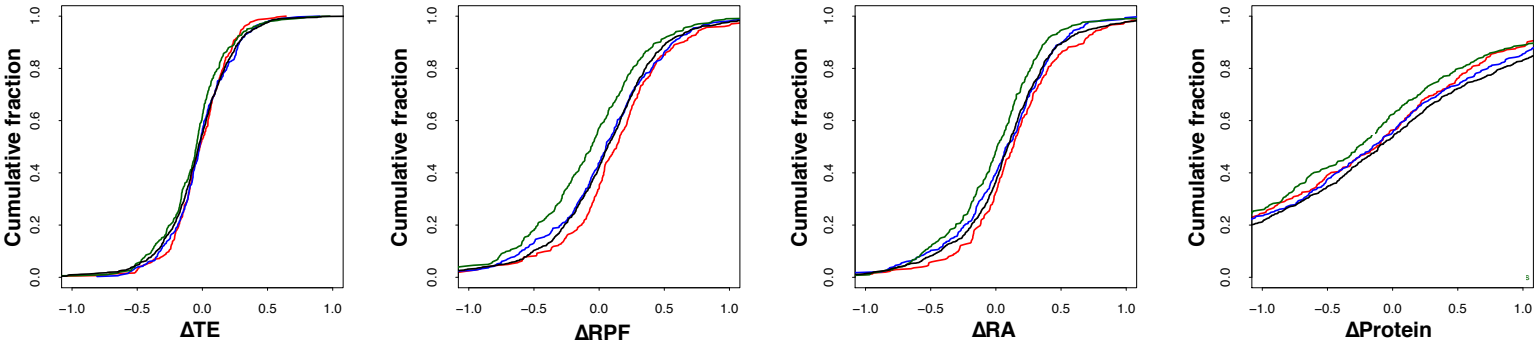


Library size	C*dcl3.1*	C*dcl3.2*	C*dcl3.3*	*dcl3-1.1*	*dcl3-1.2*	*dcl3-1.3*
trimmed reads	6,470,317	5,031,324	6,722,863	5,461,747	5,282,454	5,746,937
miRNA reads	153,598	112,892	179,859	2332	2017	2219

Supplementary Figure 6a

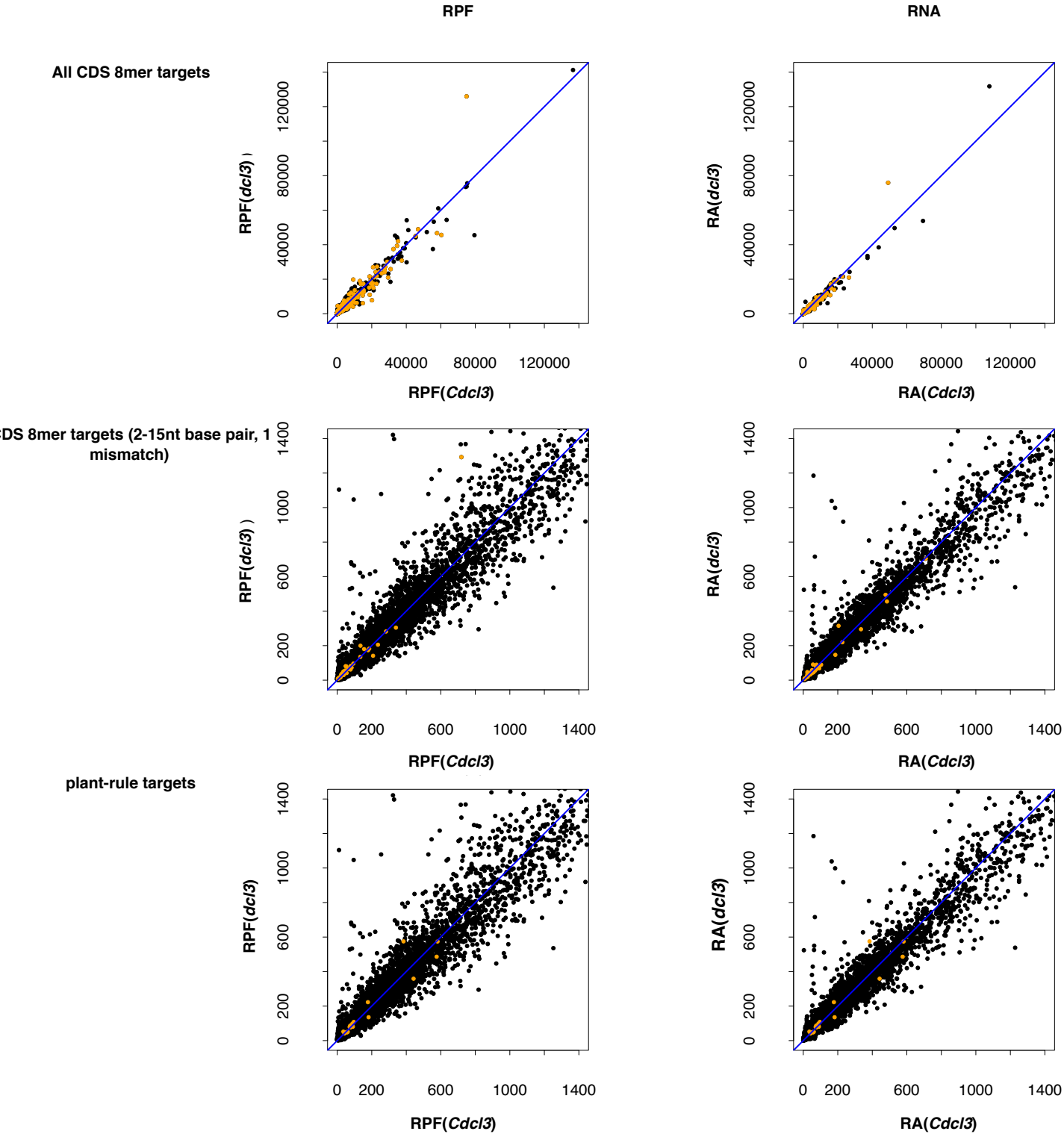


Supplementary Figure 6b

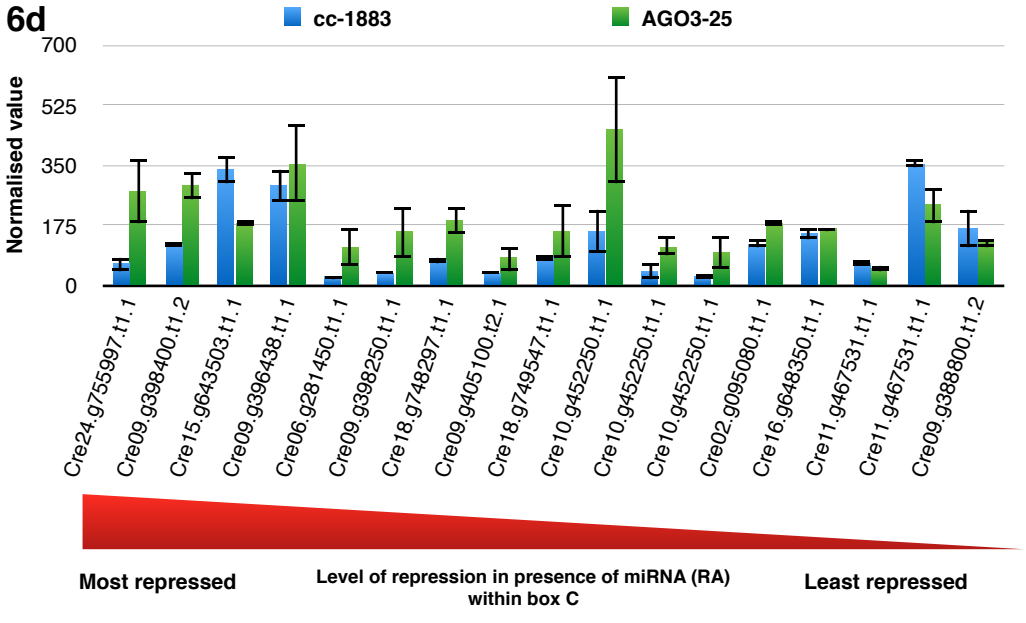


p-value with non-target (n=911) comparisons	TE	RPF	RNA	Protein
1 target site (n = 471)	0.16	0.013	0.26	0.12
2 target sites (n = 297)	0.57	0.73	0.49	0.85
≥4 target sites (n = 187)	0.05	5.3E-05	0.0026	0.043

Supplementary Figure 6c



Supplementary Figure 6d



Supplementary table 3

A (32)	
Cre01.g034600.t1.2	n/a
Cre03.g165215.t1.1	ubiquitin-like modifier-activating enzyme ATG7
Cre04.g222700.t1.2	Elongation factor 3
Cre09.g408950.t1.1	Autophagy-specific Gee 2, isoform A
Cre09.g397450.t1.1	Dimethylaniline Mnooxygenase
Cre12.g504200.t1.2	Ribosomal protein S23, component of cytosolic 80S ribosome and 40S small subunit
Cre17.g701200.t1.2	Ribosomal protein L14, component of cytosolic 80S ribosome and 60S large subunit
Cre08.g359450.t1.2	D-Alanine Ligase
Cre02.g093850.t1.1	Ras supressor protein (contains leucine-rich repeats)
Cre11.g468353.t1.1	SF14-voltage-gated potassium channel
Cre04.g214503.t1.1	Ribosomal protein S12, component of cytosolic 80S ribosome and 40S small subunit
Cre02.g091100.t1.2	Ribosomal protein L15, component of cytosolic 80S ribosome and 60S large subunit
Cre02.g106600.t1.2	Ribosomal protein S19, component of cytosolic 80S ribosome and 40S small subunit
Cre12.g498900.t1.2	Ribosomal protein S7, component of cytosolic 80S ribosome and 40S small subunit
Cre06.g299450.t1.2	n/a
Cre06.g280800.t1.2	Nuclear auto antigenic sperm protein
Cre07.g349950.t1.1	Transcription initiation factor RFIID subunit 6
Cre12.g521200.t1.2	DNA replication factor C complex subunit 1
Cre17.g720300.t1.2	Non-specific serine/threonine protein kinase
Cre06.g272950.t1.1	Ribosomal protein S18, component of cytosolic 80S ribosome and 40S small subunit
Cre08.g385800.t1.1	n/a
Cre03.g174900.t1.1	SARM1 (protein binding)
Cre12.g494050.t1.2	Ribosomal protein L9, component of cytosolic 80S ribosome and 60S large subunit
Cre09.g399141.t1.1	MFS transporter, ACS family, solute carrier family 17
Cre11.g467560.t1.1	TPR repeat containing protein
Cre16.g660750.t1.1	coiled-coil and C2 domain-containing protein 2A
Cre07.g357850.t1.2	Ribosomal protein L22, component of cytosolic 80S ribosome and 60S large subunit
Cre01.g040850.t1.2	G Protein-coupled receptor-related protein
Cre01.g036800.t1.1	Diacylglycerol kinase
Cre16.g661588.t1.1	FAST Leu-Rich Domain-containing protein
Cre07.g348550.t1.1	CGI-141-related/Lipase containing protin
Cre01.g023550.t1.1	Flagellar Associated Protein, putative outer arm dynein light chain

B (3)	
Cre16.g675200.t1.1	n/a
Cre12.g541400.t1.2	Las17-binding protein actin regulator (Ysc84)
Cre16.g677920.t1.2	n/a

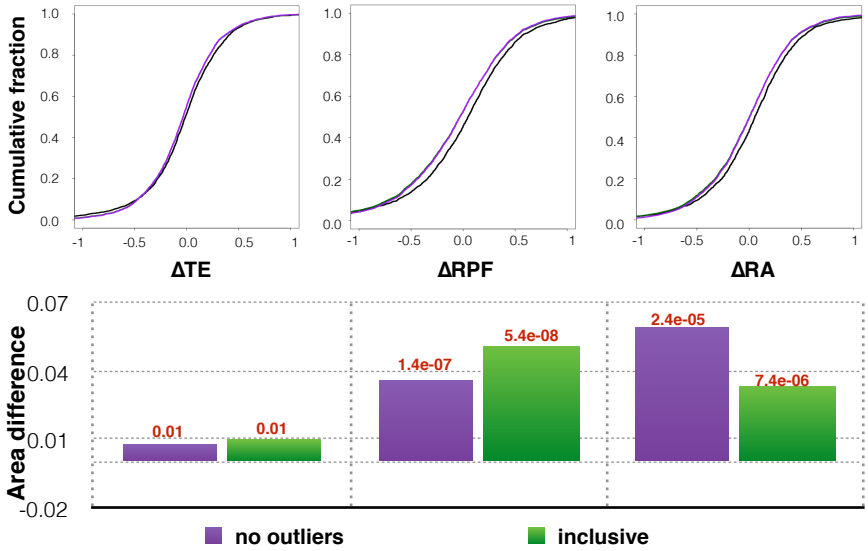
Supplementary table3

C (15)	
Cre02.g095080.t1.1	Major vault protein
Cre03.g155950.t1.2	n/a
Cre06.g281450.t1.1	Scavenger receptor cysteine rich (SRCR) protein
Cre09.g388800.t1.2	Glutamate dehydrogenase
Cre09.g396438.t1.1	n/a
Cre09.g398250.t1.1	n/a
Cre09.g398400.t1.2	Transient receptor potential ion channel protein
Cre09.g405100.t2.1	n/a
Cre10.g452250.t1.1	n/a
Cre11.g467531.t1.1	Flagellar Associated Protein
Cre15.g643503.t1.1	n/a
Cre16.g648350.t1.1	Proline Oxidase
Cre18.g748297.t1.1	n/a
Cre18.g749547.t1.1	n/a
Cre24.g755997.t1.1	Cell wall protein pherophorin-C18

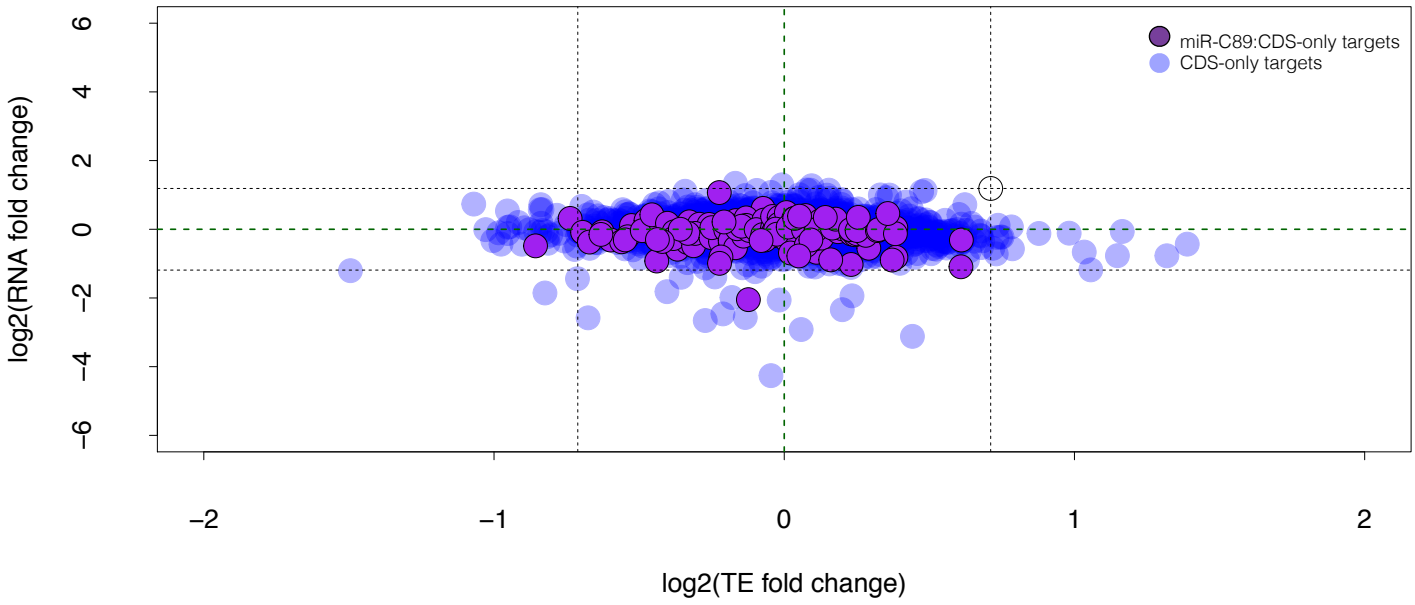
A' (16)	
Cre16.g668050.t1.1	Aspartyl protease (Asp_protease_2)
Cre06.g265850.t1.1	Tail-specific/C-terminal processing peptidase protease
Cre03.g191950.t1.2	RimM N-terminal domain (RimM)
Cre14.g614950.t1.2	Putative mitochondrial ribosomal protein S2, imported to mitochondria
Cre12.g554300.t1.1	Sodium:solute symporter
Cre15.g639050.t1.1	Zinc finger MYND domain containing protein 10
Cre14.g626800.t1.1	n/a
Cre14.g610663.t1.1	n/a
Cre06.g271950.t1.2	General vesicular transport factor P115
Cre07.g318300.t1.1	CAMP-dependent protein kinase regulatory chain
Cre17.g739850.t1.2	n/a
Cre10.g429200.t1.1	RuBisCO methyltransferase
Cre17.g703450.t1.1	n/a
Cre14.g623439.t1.1	Pyroglutamyl-peptidase I
Cre14.g626800.t1.1	n/a
Cre14.g622150.t1.1	n/a

C' (3)	n/a
Cre43.g760497.t1.1	n/a
Cre17.g734200.t1.2	L,L-diaminopimelate aminotransferase
Cre17.g734100.t1.2	n/a

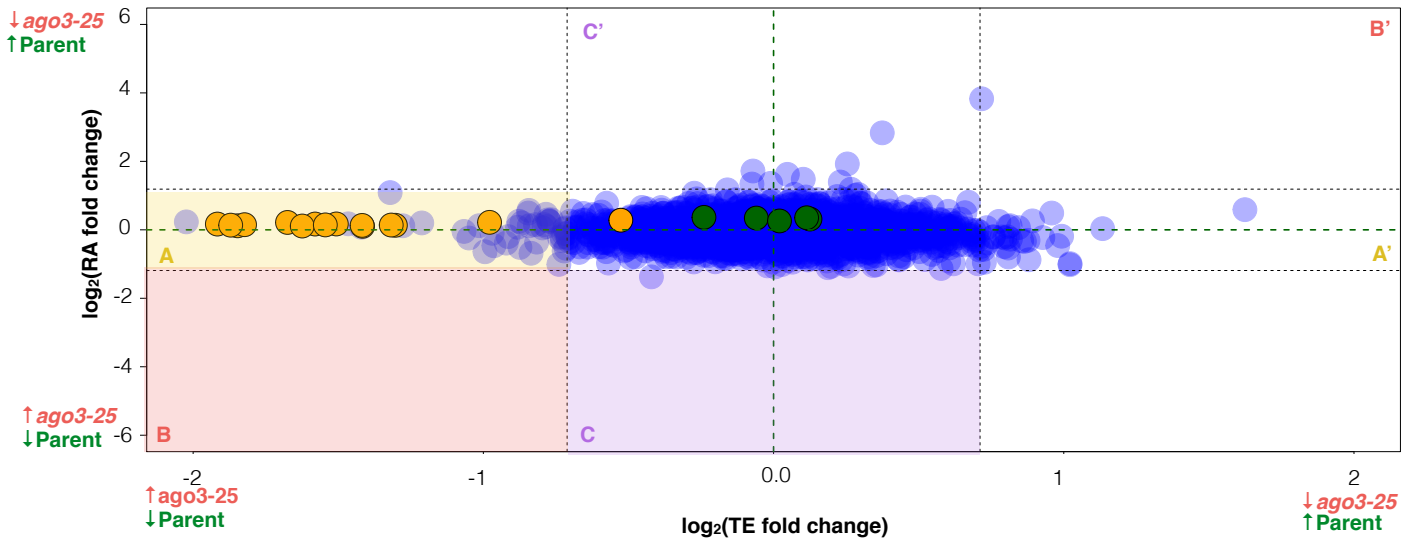
Supplementary Figure 7a



Supplementary Figure 7b



Supplementary Figure 7c



Supplementary Figure 8