Translation and codon usage regulate Argonaute slicer activity to trigger small RNA biogenesis

Germano Cecere (germano.cecere@pasteur.fr)
Institut Pasteur https://orcid.org/0000-0001-5912-3417
Meetali Singh
Institut Pasteur https://orcid.org/0000-0002-5783-073X
Eric Cornes
Institut Pasteur https://orcid.org/0000-0002-2938-3733
Blaise Li
Institut Pasteur https://orcid.org/0000-0003-3080-1899
Piergiuseppe Quarato
Institut Pasteur https://orcid.org/0000-0003-3839-3383
Loan Bourdon
Institut Pasteur https://orcid.org/0000-0001-7502-753X
Florent Dingli
Institute Curie https://orcid.org/0000-0002-7715-2446
Damarys Loew
Institute Curie https://orcid.org/0000-0002-9111-8842
Simone Procaccia
Institut Pasteur https://orcid.org/0000-0001-5791-3079

Article

Keywords: Argonaute, CSR-1, small-RNA, C. elegans, 22G-RNA biogenesis, translation

Posted Date: October 23rd, 2020

DOI: https://doi.org/10.21203/rs.3.rs-87655/v1

Version of Record: A version of this preprint was published at Nature Communications on June 9th, 2021.
See the published version at https://doi.org/10.1038/s41467-021-23615-w.

Abstract

In the *Caenorhabditis elegans* germline, thousands of mRNAs are concomitantly expressed with antisense 22G-RNAs, which are loaded into the Argonaute CSR-1. Despite their essential functions for animal fertility and embryonic development, how CSR-1 22G-RNAs are produced remains unknown. Here, we show that CSR-1 slicer activity is primarily involved in triggering the synthesis of small RNAs on the coding sequences of germline mRNAs and post-transcriptionally regulates a fraction of targets. CSR-1-cleaved mRNAs prime the RNA-dependent RNA polymerase, EGO-1, to synthesize 22G-RNAs in phase with ribosome translation in the cytoplasm, in contrast to other 22G-RNAs mostly synthesized in germ granules. Moreover, codon optimality and efficient translation antagonize CSR-1 slicing and 22G-RNAs biogenesis. We propose that codon usage differences encoded into mRNA sequences might be a conserved strategy in eukaryotes to regulate small RNA biogenesis and Argonaute targeting.

Main Text

In animals, small RNAs that are expressed in the germline and transmitted to embryo act as a defense mechanism to repress foreign RNAs such as viruses, transposons and other repetitive elements (REs). These small RNAs are essential for fertility and genome integrity[1,2]. Their function is controlled by the conserved family of Argonaute proteins (AGOs), which loads the small RNAs and function to repress complementary mRNA targets through their endonuclease activity or by recruiting other effector silencing proteins[3–6]. The *C. elegans* germline contains a complex small RNA regulatory network, with different classes of small RNAs, multiple AGO effectors and diverse biogenesis pathways[7]. One of the most abundant class of endogenous small RNAs in the germline is the 22G-RNAs, which are single-stranded antisense small RNAs produced by RNA-dependent RNA polymerase (RdRPs) as part of an amplification system to silence target transcripts (reviewed in[7]). The production of 22G-RNAs targeting REs is triggered by over 15,000 PIWI-interacting RNAs (piRNAs or 21U-RNAs) and loaded by Worm-specific Argonautes (WAGOs) to silence REs[8–11]. 22G-RNAs are also produced from the majority of germline-expressed mRNAs by the RdRP EGO-1 and loaded into the Argonaute CSR-1[12,13]. In contrast to the 22G-RNAs antisense to REs, which can be triggered in response to piRNAs, the primary trigger for generating CSR-1 22G-RNAs and why many germline mRNAs become targeted by CSR-1 is still unknown (Extended Data Fig. 1).

Given that the *C. elegans* piRNAs can trigger the silencing of their targets by imperfect complementarity, and therefore can potentially target germline-expressed mRNAs[14–16], the targeting by CSR-1 22G-RNAs can function as an anti-silencing mechanism to protect germline mRNAs from piRNAs silencing[13,17,18]. The anti-silencing function of CSR-1 was established with single-copy transgenes[15,17,18]. However, germline mRNAs remain protected from piRNAs silencing even in the absence of CSR-1[16], and sequence-encoded features of germline mRNAs have also been proposed to prevent piRNA silencing[14,16]. To what extent endogenous germline-expressed genes are regulated by the antagonistic functions of CSR-1 and piRNA pathways remains elusive (Extended Data Fig. 1).

In addition, CSR-1 has been proposed to directly regulate the expression of its germline targets. Of the Argonautes that load 22G-RNAs, only CSR-1 has demonstrated slicer activity on target mRNA *in vitro*[19]. Worms lacking CSR-1 protein display downregulation of germline CSR-1 targets[13,20,21], those expressing a CSR-1 catalytic mutant protein show upregulation of its germline target genes[22]. Therefore, the gene regulatory functions of germline CSR-1 22G-RNAs remain incompletely understood (Extended Data Fig. 1).

Many germline Argonautes, including CSR-1 and PIWI, and proteins involved in 22G-RNA biogenesis, including RdRPs, localize to perinuclear condensates called germ granules[13,23]. These germ granules are thought to be the site for the biogenesis of all germline 22G-RNAs. However, disruption of specific germ granule, the mutator foci, which participates in piRNA-dependent 22G-RNA accumulation, has no apparent effect on CSR-1 22G-RNAs[24,25]. Moreover, the type of RNA template used by the EGO-1 RdRP to generate CSR-1 22G-RNAs also remains mysterious. During exogenous RNAi, the addition of polyUG to cleaved mRNA targets by RDE-3 recruits RdRPs EGO-1 and RRF-1 to synthesize 22G-RNAs[26,27]. However, RDE-3 is not required to generate CSR-1 22G-RNAs[23,27]. Thus, whether CSR-1 22G-RNAs are synthesized in germ granules on a specific type of RNA substrates remains to be elucidated.

In the current study, we elucidate CSR-1 catalytic activity-dependent and independent germline gene regulation and decipher the rules governing CSR-1 22G-RNA biogenesis. We demonstrate that the slicer activity of CSR-1 triggers the biogenesis of 22G-RNAs on the coding sequence of germline mRNAs. We establish that CSR-1 22G-RNAs are synthesized on an actively translated mRNA template in the cytosol, independent of germ granules. Overall, this study establishes that translation and codon usage dictate CSR-1 slicer activity on a target mRNA to regulate small RNA biogenesis and functions.

Results

Defects in CSR-1 catalytic activity mainly impact 22G-RNA abundance

Worms lacking CSR-1 show downregulation[13,20,21] whereas those expressing CSR-1 protein in which the catalytic DDH motif was mutated to ADH show upregulation of some targets[22]. The global impact of CSR-1 mutations on gene expression might depend on the developmental context and might be biased by developmental defects[13,28]. Indeed, we observed differences during oogenesis in the adult *csr-1* catalytic mutant (*csr-1* ADH) and knockout (*csr-1* KO) worms marked by a delayed onset of oocyte production and increased accumulation of oocytes in germline at more advanced stages compared to wild-type (WT) (Extended Data Fig. 2a-c). To overcome this limitation, we developed a sorting strategy to obtain a synchronized population of WT and first-generation homozygotes for *csr-1* KO or CSR-1 ADH strains using COPAS biosorter. Using this strategy, we enriched for larval stage late L4 worms which is characterized by a closed vulva and absence of oocytes and lacking the germline developmental abnormality (Extended Data Fig. 2d-e).

Next, to precisely evaluate the role of CSR-1, we measured small RNA accumulation (sRNA-seq), transcription (GRO-seq), mRNA stability (RNA-seq), and translation (Ribo-seq) in WT and mutant worms. In addition, to assess the direct effect of CSR-1 22G-RNAs on these processes, we sequenced the small RNAs bound to immunoprecipitated CSR-1 from similarly sorted late L4 worms to precisely identify the CSR-1 targets at the same developmental stage. We detected a total of 4803 genes with antisense 22G-RNAs loaded into CSR-1 (IP over input ≥ 2-fold enrichment and RPM ≥ 1 in each replicate of CSR-1 IP) (Supplemental Table 1). The CSR-1 catalytic mutant displayed a global loss of 22G-RNAs for the majority of CSR-1 targets (Fig. 1a, c). However, only 7.7% (n = 119) of CSR-1 targets with > 2-fold reduction of 22G-RNAs (n = 1536) showed increased mRNA levels (Fig. 1b), indicating that most mRNA targets are not destabilized by CSR slicer activity. The increase in mRNA and translational levels of the targets correlated with 22G-RNA levels in CSR-1 IPs in a dose-dependent manner (Fig. 1d, e) in agreement with a previous report[22], but their transcription was unaffected (Fig. 1f). Therefore, we conclude that CSR-1 slices a subset of target mRNAs having abundant 22G-RNAs. Moreover, we found that the targets that are post-transcriptionally regulated by CSR-1 are enriched for mRNAs encoding CSR-1 interacting proteins, identified by mass spectrometry (MS/MS) (enrichment factor 6.1, $p < 1.6e^{-28}$) as direct, RNA-independent interactions (Extended Data Fig. 2g-i). Thus, CSR-1 slicer activity negatively regulates the expression of its own interactors, including CSR-1, suggesting a negative feedback loop.

Overall, these results suggest that the main role of CSR-1 catalytic activity is to control the accumulation of 22G-RNAs. In addition, CSR-1 post-transcriptionally regulates a small fraction of CSR-1 targets that have highly abundant 22G-RNAs.

CSR-1 protects a subset of oogenic enriched targets from piRNA-mediated transcriptional silencing

Similar to CSR-1 ADH worms, CSR-1 KO worms displayed a loss of 22G-RNAs as well as an upregulation of a subset of target mRNAs characterized by high abundance of 22G-RNAs (Extended Data Fig. 3a-d). However, the level of upregulation of CSR-1 target mRNAs was significantly lower in the *csr-1* KO compared to the *csr-1* ADH, possibly due to decreased transcription (Extended Data Fig. 3e). Indeed, we found that a subset of target genes displayed downregulated transcription and reduced mRNA levels in the KO compared to WT, but these were unaffected in the CSR-1 ADH (Fig. 1g). The majority of these genes (53%) were enriched for oogenic mRNAs (see Supplemental Table 1 for gene list). Given that CSR-1 is proposed to protect germline transcripts from piRNA-mediated silencing, we hypothesized that in the *csr-1* KO, piRNAs can trigger the loading of 22G-RNAs into the nuclear Argonaute HRDE-1 resulting in the reduced transcription of this subset of CSR-1 targets. Indeed, HRDE-1 loads 22G-RNAs from transcriptionally downregulated CSR-1 targets in the *csr-1* KO (Fig. 1h). Overall, these data suggest a non-catalytic role of the CSR-1 protein in protecting a subset of oogenic targets from piRNA-mediated HRDE-1 transcriptional silencing.

CSR-1 catalytic activity is required for biogenesis of 22G-RNAs on the coding sequence of target mRNAs

The global reduction of CSR-1-bound 22G-RNAs observed in CSR-1 mutants suggests that CSR-1 catalytic activity is required for 22G-RNA loading or biogenesis. Immunoprecipitation with WT or CSR-1 ADH proteins did not show any loss of binding of 22G-RNAs (Extended Data Fig. 4a), suggesting that catalytic inactive CSR-1 can still bind the 22G-RNAs produced in the mutant. We then investigated the distribution of CSR-1-bound 22G-RNAs along the target gene bodies. We found that 22G-RNAs are synthesized only at 3' untranslated region (3'UTR) of their target mRNAs in the *csr-1* catalytic mutant, indicating that the RdRP fails to synthesize 22G-RNAs on the coding sequence (CDS) (Fig. 2a-c). The CSR-1 catalytic mutation does not impair the loading of these 22G-RNAs generated from 3'UTR (Fig. 2a, c and Extended Data Fig. 4b). However, it fails to produce and load small RNAs from CDS (Fig. 2a-c and Extended Data Fig. 4b).

The RdRP EGO-1 has been proposed to exclusively synthesize CSR-1-bound 22G-RNAs[12,13,23]. To understand whether the small RNAs produced on the 3'UTR in the absence of CSR-1 protein or its catalytic activity are synthesized by EGO-1, we efficiently depleted CSR-1 using an auxin-induced degradation system, combined with *ego-1* RNAi knockdown (Extended Data Fig. 4c-e). First, we confirmed that CSR-1 22G-RNAs were depleted on CDS and enriched on 3'UTR upon auxin-induced CSR-1 depletion (Fig. 2d, e), suggesting that the effect on CSR-1 22G-RNAs was not caused by a gain-of-function mutation in CSR-1 ADH. Next, we observed decreased 22G-RNAs from both CDS as well as 3'UTR upon *ego-1* RNAi knockdown (Fig. 2d, f, and Extended Data Fig. 4f, g), implying that EGO-1 is exclusively responsible for the synthesis of the CSR-1 22G-RNAs in both WT and the *csr-1* mutants. However, the catalytic activity of CSR-1 is required to efficiently generate EGO-1-dependent 22G-RNAs along the coding sequences of target mRNAs.

Finally, we tested whether the restored expression of CSR-1 is sufficient to generate EGO-1-dependent 22G-RNAs on the gene body. For this purpose, we depleted CSR-1 by auxin-induced degradation for 38 h after hatching (0 h recovery) and then reintroduced CSR-1 by recovering expression for 5 and 10 hours (Extended Data Fig. 4h). As expected, the depletion of CSR-1 caused a loss of 22G-RNA accumulation on the CDS (Fig. 2g and Extended Data Fig. 4h - see 0 h recovery). However, upon reintroduction of CSR-1 expression (5 and 10 h recovery), we observed a steady increase of 22G-RNAs, mainly on the CDS (Fig. 2g, h). The lack of complete recovery of 22G-RNAs could be due to the accumulation of germline defects as a result of CSR-1 depletion during the initial period of germline development.

Overall, these data demonstrate that EGO-1 can be recruited on the 3'UTR of target mRNAs and initiate the production of 22G-RNAs. However, CSR-1-mediated slicing of mRNAs is required to template the production of small RNAs on the gene body.

Biogenesis of CSR-1 22G-RNAs and the regulation of their targets occurs in the cytosol

PIWI and RNAi biogenesis factors are known to localize in perinuclear condensates, called germ granules, and these germ granules have been proposed to be the site for biogenesis of 22G-RNAs[24,29–31]. CSR-1

and EGO-1 localize in both cytosol and the germ granules[13], suggesting that the biogenesis of CSR-1 22G-RNAs might also occur in these organelles. To test this possibility, we used RNAi to simultaneously deplete four core components of germ granules (*pgl-1, pgl-3, glh-1*, and *glh-4*)[32], called P granules (Extended Data Fig. 5a). This treatment was sufficient to disrupt the localization of the GLH-1, PIWI and CSR-1 in germ granules (Fig. 3a). However, the cytosolic localization of CSR-1 remained unaffected (Fig. 3a).

Next, we evaluated the effects of loss of germ granule localization of PIWI and CSR-1 on 22G-RNA biogenesis. As expected, piRNA-dependent 22G-RNAs were globally depleted upon P granule RNAi treatment (Fig. 3c). Surprisingly, CSR-1 22G-RNAs were unaffected upon P granule RNAi treatment, despite the loss of perinuclear CSR-1 germ granule localization (Fig. 3b, c). Furthermore, CSR-1 targets were not upregulated upon P granule RNAi (RNA-seq data from[33] (Fig. 3d and Extended Data Fig. 5a). These results highlight that CSR-1 22G-RNA biogenesis occurs in the cytosol independently of germ granules.

Translating mRNAs serve as the template for 22G-RNA biogenesis

Our data so far suggest that CSR-1 22G-RNAs are generated in the cytosol. Consistent with CSR-1 localization in the cytosol and germ granules, we identified ribosomal and ribosomal-associated proteins, which are enriched in the cytosol, and germ granule components in our IP-MS/MS as direct CSR-1 interactors (Fig. 4a and Barucci et.al[34]). Moreover, CSR-1 ADH showed reduced co-purification of ribosomal proteins and increased co-purification of germ granule components, compared to CSR-1 WT (Fig. 4b) and localizes primarily in germ granules (Extended Data Fig. 5b).

Based on these data, we hypothesized that 22G-RNAs are synthesized in the cytosol, using translating mRNAs as templates. To test this hypothesis, we mapped the distance between the start of the 29-nucleotide Ribosomal Protected Fragments (RPF)[35] and the 5' end of CSR-1 22G-RNAs (Fig. 4c). We observed periodicity of 3 nucleotides in phase with ribosomes (Fig. 4d), indicating that the synthesis of CSR-1 22G-RNAs occurs on mRNA templates engaged in translation. In contrast, the HRDE-1 loaded 22G-RNAs of P granule dependent piRNA targets (Supplemental Table 1) did not show phasing with ribosomes as observed due to a lack of 3 nucleotide periodicity (Fig. 4c and d), in agreement with the fact that P granules are devoid of translating mRNAs[36,37].

Altogether these results suggest that CSR-1 cleaves actively translating mRNAs, which become the template for EGO-1-mediated synthesis of 22G-RNAs on the coding sequence of mRNA targets.

mRNA translation antagonizes CSR-1 22G-RNA biogenesis

EGO-1 mediated synthesis of CSR-1 22G-RNAs does not occur on every germline mRNA at similar levels, and we found that the levels of 22G-RNA are independent of the levels of the mRNA template (Extended Data Fig. 6a). Given our observations that actively translating mRNAs serve as the template for CSR-1

22G-RNAs, we hypothesized that the translation efficiency (TE) of germline mRNAs impacts CSR-1 22G-RNA biogenesis. To test this hypothesis, we calculated the TE of CSR-1 targets using the Ribo-seq and RNA-seq data from WT worms at the late L4 stage. We observed that levels of CSR-1 associated 22G-RNAs produced from a target mRNA were inversely correlated with their TE (Fig. 5a), suggesting that translation antagonize the biogenesis of CSR-1 22G-RNAs.

Codon usage and the availability of the tRNA pool influence TE[38,39]. Therefore, we investigated whether these mechanisms affect the biogenesis of CSR-1 22G-RNAs. We determined optimal and non-optimal codons using our experimental data from Late L4 staged worms. First, we calculated the normalized average relative synonymous codon usage (RSCU) for genes for different categories of high or low TE (Fig. 5b). Codons showing enrichment in genes with high TE ($\log_2 TE \geq 3$) were considered optimal codons, and the ones under-represented were considered non-optimal codons (Fig. 5b). We confirmed that our classification of optimal/non-optimal codons correlated with tRNA copy number (Fig. 5d, Extended Data Fig. 6b, d) and tRNA pool available in the late L4 worm population (44 h) as measured by GROseq (Fig. 5e, Extended Data Fig. 6c, e). We noticed that for codons with no tRNA cognates and requiring tRNA binding by wobble pairing, all optimal codons end with C, and non-optimal with U. Translation elongation is lower for those ending with a U[40].

We then evaluated the codon usage of CSR1 targets by comparing their normalized average RSCU to highly translated mRNAs. We found that non-optimal codons were enriched and optimal codons were depleted in CSR-1 targets, suggesting that this might be an encoded feature of mRNA targets influencing the priming of 22G-RNA synthesis (Fig. 5c). Non-optimal codons are known to promote ribosome stalling[41–43]. To map differences in 22G-RNA biogenesis on sequences with optimal or non-optimal codons, we divided RPFs into two categories based on the presence of either an optimal or non-optimal codon at the A and P sites of the ribosome and then mapped the distance between 5’ of 22G-RNAs and RPFs. We observed a peak for the 5' end of 22G-RNAs downstream of RPF (29th position) when the A and P sites of the ribosomes are occupied by a non-optimal codon contrary to when optimal codons are present on A and P sites which show no bias (Fig. 5f). This result suggests that the 22G-RNA production is preferentially initiated downstream of ribosomes especially occupying bad codons, by CSR-1 mediated slicing and recruitment of EGO-1.

Altogether, these observations suggest that translation and ribosome position dictate the production of CSR-1 22G-RNAs.

Increasing the translation efficiency of CSR-1 target impairs CSR-1 22G-RNA biogenesis and function.

To determine whether non-optimal codons directly affect TE and CSR-1 22G-RNA biogenesis, we altered the coding potential of a CSR-1 target. We examined *klp-7*, which has the second-highest abundance of 22G-RNAs loaded by CSR-1 and is post-transcriptionally regulated by CSR-1. KLP-7 is a kinesin-13 microtubule depolymerase and is required for spindle organization and chromosome segregation[44]. Overexpression of KLP-7 in the *csr-1* mutant has been shown to cause microtubule assembly defects[22].

klp-7 showed enrichment of non-optimal codons and depletion of optimal codons similarly to other CSR-1 targets (Extended Data Fig. 7a). We optimized the codon usage in *klp-7* by incorporating exclusively synonymous optimal codons (Extended Data Fig. 7a). We used CRISPR-Cas9 to replace endogenous *klp-7* isoform b with the modified *klp-7* codon-optimized (*klp-7*_co) to avoid disrupting potential UTR-mediated regulation.

To ascertain whether codon optimization of *klp-7* affected the TE, we performed RNA-seq and Ribo-seq from synchronized and sorted late L4 population (44 h). Indeed, we detected a 2-fold increase in the TE of *klp-7* mRNA in the *klp-7*_co strain compared to WT (Fig. 6a). The TE of other CSR-1 targets remained unaffected in *klp-7*_co strain, indicating that the effects observed are specific to *klp-7* mRNA (Fig. 6a). In addition, KLP-7 protein levels were increased in two independent lines of *klp-7*_co compared to WT, consistent with increased translation (Extended Data Fig. 7c). We then measured the level of 22G-RNAs antisense to *klp-7* mRNA in the *klp-7*_co strain compared to WT, and we observed a 1.4-fold decrease in 22G-RNAs (Fig. 6a). The levels of 22G-RNAs for other CSR-1 targets remained unaffected (Fig. 6a). Further, the significant decrease in 22G-RNAs on the optimized *klp-7*_co allele was observed in exons 3–6 and was accompanied by an increase in ribo-seq peak height at those positions (Fig. 6b). The *klp-7*_co strain also showed increased *klp-7* mRNA level compared to WT (Fig. 6a), and we confirmed this result by RT-qPCR (Extended Data Fig. 7b). These results suggest that CSR-1 targeting and regulation is impaired on *klp-7_co* mRNA. To validate this, we performed *csr-1* RNAi and showed increased *klp-7* mRNA levels in the WT strain but not in the *klp-7*_co strain (Fig. 6c), suggesting that CSR-1 activity is reduced at *klp-7*_co mRNA. The increased levels of *klp-7* mRNA correlated with a reduction in brood size (Extended Data Fig. 7d) and higher embryonic lethality at 25 °C in *klp-7*_co strain compared to WT (Extended Data Fig. 7e), indicating the physiological relevance of *klp-7* mRNA targeting by CSR-1. Finally, to rule out any difference in the production of either 22G-RNAs or mRNA levels due to possible developmental defects between *klp-7*_co and WT strain, we generated a heterozygote strain of *klp-7*_co with a fluorescent GFP marker on the balancer chromosome. We sorted heterozygote GFP positive worms with one copy of modified *klp-7_co* and one copy of WT *klp-7* each and performed RNA-seq and sRNA-sEq. We observed similar results with a 1.8- fold increase in mRNA levels for *klp-7*_co compared to the WT *klp-7* copy and a 1.25-fold decrease in 22G-RNA levels (Extended Data Fig. 7f-g). These results demonstrate that CSR-1 22G-RNA biogenesis and activity is reduced on mRNA templates with optimized codons and increased translation.

Altogether, these results suggest efficiently translating ribosomes block the access of CSR-1 to the mRNA template and thereby hamper 22G-RNA production and, in turn, affects gene regulation by CSR-1 during germline development.

Discussion

In this study, we have determined the rules governing germline mRNA targeting by CSR-1 and addressed the long-standing paradox of CSR-1 function as anti-silencer or a slicer. We show a significant fraction of the slicing activity of CSR-1 is directed towards the production of 22G-RNAs on the coding sequence of its

targets. We demonstrate that only a fraction of CSR-1 slicing activity regulates targets post-transcriptionally in the germline. Our observations also suggest a catalytic-independent function of CSR-1 in preventing piRNAdependent chromatin silencing. Specifically, we showed that in the absence of CSR-1 protein, a subset of CSR-1 target genes is misrouted into the piRNA pathway, which represses their expression at transcriptional levels through the nuclear Argonaute HRDE-1. Therefore, in addition to the post-transcriptional regulation of germline mRNAs [22], CSR-1 can also license the transcription of germline genes (Fig. 7).

We further dissected the mechanism of CSR-1 22G-RNA production. We demonstrated that the synthesis of 22G-RNAs occurs in the cytosol on translating mRNA templates, whereas germ granules are dispensable for CSR-1 regulation. We propose that a low translation efficiency favors the CSR-1 slicer activity on the target mRNA occupied by ribosomes and initiates 22G-RNA biogenesis by priming RdRP EGO-1 activity. Finally, we have determined how CSR-1 can preferentially target some germline mRNAs. We discovered that incorporation or avoidance of non-optimal codons is a strategy adopted by germline mRNAs to be differentially regulated by CSR-1 22G-RNAs (Fig. 7).

Overall, this study highlights the codon dependence and translational efficiency of mRNAs in the germline for the regulation of CSR-1- 22G-RNAs biogenesis and, in turn, gene expression of the targets, which could have a significant bearing on germline gene regulation not just in worms but across species.

Biogenesis of CSR-1 22G-RNAs

In this study, we have established that the Argonaute CSR-1 slices target mRNAs to trigger the generation of RdRP-dependent 22G-RNAs on the gene body. We propose that CSR-1 slicer activity is required to generate new 3'-OH ends along the gene transcript to facilitate the initiation of RdRP (EGO-1) synthesis towards the 5' end of the mRNA target. This is consistent with previous *in vitro* RdRP analysis showing that non-polyadenylated 3'-OH ends of RNAs served as better substrates for 22G-RNA synthesis[19], suggesting that the cleavage of RNA may be vital for the processivity of RdRPs. Based on these results, we speculate that no primary small RNAs are required to generate CSR-1 22G-RNAs along the mRNA sequence. Instead, CSR-1 catalytic activity triggers the synthesis of 22G-RNAs by the RdRP EGO-1, starting from the 3'UTR of the target transcripts. Even if the catalytic activity of CSR-1 is required to generate 22G-RNAs along the gene body of target transcripts, it is still unknown what triggers the recruitment of EGO-1 on the 3'UTR. Primary small RNAs, which are yet to be identified, might prime the activity of EGO-1. Alternatively, EGO-1 might produce low levels of 22GRNAs from the polyadenylated tail of mRNAs instead of cleaved 3'OH end products. Thus, these low levels of 22G-RNAs, which are then loaded into CSR-1 can initiate the production of 22G-RNAs along the gene body. RNA binding proteins and/or other unknown factors together with specific sequences in the 3’UTR might also recruit and initiate EGO-1-dependent 22G-RNAs from the 3'UTR of selected mRNAs.

The role of translation and codon usage in CSR-1 22G-RNA biogenesis

Most germline-expressed mRNAs are targeted by CSR-1 to promote EGO-1 dependent 22G-RNAs, but they generate different levels of 22G-RNAs. We found that CSR-1 targets in adult germlines are mainly actively translating mRNAs, and 22G-RNAs are synthesized in phase with ribosomes. Therefore, the CSR-1-22G-RNA biogenesis machinery likely needs to cope with the presence of ribosomes on the target transcripts. We show that non-optimal codons in germline mRNAs enhance the capacity of CSR-1 to prime the synthesis of EGO-1-dependent 22G-RNAs along the gene body. The use of non-optimal codons by CSR-1 targets and priming of 22G-RNAs at stalled positions is a way to cope with the ribosomal presence on the target transcripts. Therefore, sequences that promote ribosome stalling promote targeting by CSR-1. We demonstrate that substituting non-optimal codons with optimal codons is sufficient to allow germline mRNAs to escape CSR-1-dependent regulation. There is increasing evidence that the translation machinery associates with the Argonautes and small RNA biogenesis factors. Recently it was shown that ribosome movement on translating mRNAs resolves mRNA structure to provide accessibility to Argonaute AGO2 downstream of ribosome and promote AGO2-target interaction[45,46]. Another report showed that RNAi can occur co-translationally with an accumulation of ribosomes upstream of the dsRNA targeted region[47]. Ribosomes have been shown to coordinate with piRNA biogenesis factors in mouse testes to achieve endonucleolytic cleavage of non-repetitive long RNAs to produce pachytene piRNAs[48]. In plants 22-nt siRNAs can repress translation leading to induction of transitive small RNA amplification by RNA-dependent RNA polymerase 6 (RDR6)[49]. Another recent report in plants showed that microRNA targeting recruits a double strand RNA binding protein which induces ribosome stalling and the ribosome stalling enhances generation of secondary small RNAs[50]. Therefore, we propose that the regulation of small RNA biogenesis by ribosome occupancy and codon usage of the target transcript might be a general strategy adopted across evolution.

Granule vs Cytosolic functions of CSR-1

We found that the slicer activity of CSR-1 and 22G-RNA biogenesis at germline mRNA targets are independent of germ granules. This raises the question on the function of CSR-1 in germline granules. CSR1 might be enriched in germ granules of adult gonads to prevent CSR-1 slicer activity on the majority of germline mRNAs. Indeed, only 7.7% of CSR-1-dependent 22G-RNA targets are significantly regulated by CSR-1 slicer activity in adults. Moreover, the majority of these targets are CSR1interacting proteins suggesting a negative feedback regulation of the CSR-1 pathway. This is in contrast with the recently described function of the maternally delivered CSR-1 in the embryo, which exclusively localizes in the cytosol of somatic blastomere where it cleaves and clears hundreds of maternal mRNA targets[51]. Therefore, we propose that CSR-1 slicer activity on mRNA targets is partially suppressed in the germline by titrating away a part of CSR-1 in germ granules and primarily serves to generate interacting small RNAs in the cytosol that fully operates intra-generationally in the embryo. This also explains why despite targeting almost all germline genes, CSR-1 catalytic activity regulates the expression of only a few in the germline. In addition, CSR-1 localization in the germ granule might serve to antagonize piRNA-dependent targeting on germline mRNAs and therefore license those transcripts to be translated in the cytosol.

Indeed, we have shown that most of the piRNA-dependent 22G-RNAs are generated in germ granules, and we propose that the competition between CSR-1 and PIWI might occur in germ granules.

Declarations

ACKNOWLEDGEMENTS

We would like to thank all the members of the Cecere laboratory, Manish Grover, Sudarshan Gadadhar, and Angela Anderson (Life Science Editors) for the helpful discussions on the paper. We thank Micheline Fromont for her help to set up Ribosome profiling. We thank Celine Didier for technical assistance. We thank the Heng-Chi Lee lab, Miska lab, Desai lab, Updike lab and Kennedy lab for sharing strains and reagents. Some strains were provided by the CGC, funded by the NIH Office of Research Infrastructure Programs (P40 OD010440).

This project has received funding from the Institut Pasteur, the CNRS, and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement No ERC-StG- 679243. MS and EC were supported by the

Pasteur-Roux-Cantarini Postdoctoral Fellowship program. PQ was supported by Ligue Nationale Contre le Cancer (SFB19032). FD and DL have received funding from Région Ile-de-France and Fondation pour la Recherche Médicale grants to support this study.

AUTHOR CONTRIBUTIONS

GC and MS identified and developed the core questions addressed in the project. MS performed most of the experiments and analyzed the results together with GC. EC and LB generated all the lines used in this study with the help of MS and PQ. EC performed the imaging experiments, RNA-seq for CSR-1 mutants and IP-sRNA-seq of HRDE-1 in CSR-1 KO. PQ performed GRO-seq for CSR-1 mutants. BL performed all the bioinformatics analysis. SP contributed for distance mapping analysis of 22G-RNA reads and Ribo-seq reads with BL. FD and DL performed MS/MS experiments and analyzed the data together with MS. MS and GC wrote the paper with the contribution of EC and PQ.

COMPETING INTERESTS

All the authors declare no competing interests.

DATA AND MATERIALS AVAILABILITY

All sequencing data (GRO-seq, RNA-seq, and sRNA-seq from total lysate or IP experiments, Ribo-seq) are available at the Gene Expression Omnibus (GEO) under accession code GSE155077 (secure token for reviewers: oncnuaccfbktfsd). The MS proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD012557 and PXD020293 (Username: reviewer04729@ebi.ac.uk, Password: M1ngh5jC). All other data supporting the findings of

this study are available from the corresponding author on request. The custom scripts generated for this study are available from the corresponding author on request. Some of the data analysis pipelines are available at https://gitlab.pasteur.fr/bli/bioinfo_utils

References

1. Ghildiyal, M. & Zamore, P. D. Small silencing RNAs: an expanding universe. *Nat Rev Genet* **10**, 94–108 (2009).
2. Okamura, K. & Lai, E. C. Endogenous small interfering RNAs in animals. *Nat Rev Mol Cell Biology* **9**, 673–8 (2008).
3. Hammond, S. M., Bernstein, E., Beach, D. & Hannon, G. J. An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. *Nature* **404**, 293–296 (2000).
4. Williams, R. W. & Rubin, G. M. ARGONAUTE1 is required for efficient RNA interference in Drosophila embryos. *Proc National Acad Sci* **99**, 6889–6894 (2002).
5. Tabara, H. *et al.* The rde-1 Gene, RNA Interference, and Transposon Silencing in C. elegans. *Cell* **99**, 123–132 (1999).
6. Grishok, A. *et al.* Genes and Mechanisms Related to RNA Interference Regulate Expression of the Small Temporal RNAs that Control C. elegans Developmental Timing. *Cell* **106**, 23–34 (2001).
7. Billi, A. C., Fischer, S. E. J. & Kim, J. K. Endogenous RNAi pathways in C. elegans. *Wormbook* 1–49 (2014) doi:10.1895/wormbook.1.170.1.
8. Das, P. P. *et al.* Piwi and piRNAs Act Upstream of an Endogenous siRNA Pathway to Suppress Tc3 Transposon Mobility in the Caenorhabditis elegans Germline. *Mol Cell* **31**, 79–90 (2008).
9. Bagijn, M. P. *et al.* Function, Targets, and Evolution of Caenorhabditis elegans piRNAs. *Science* **337**, 574–578 (2012).
10. Batista, P. J. *et al.* PRG-1 and 21U-RNAs Interact to Form the piRNA Complex Required for Fertility in C. elegans. *Mol Cell* **31**, 67–78 (2008).
11. Lee, H.-C. *et al.* C. elegans piRNAs Mediate the Genome-wide Surveillance of Germline Transcripts. *Cell* **150**, 78–87 (2012).
12. Maniar, J. M. & Fire, A. Z. EGO-1, a C. elegans RdRP, modulates gene expression via production of mRNA-templated short antisense RNAs. *Curr Biology Cb* **21**, 449–59 (2011).
13. Claycomb, J. M. *et al.* The Argonaute CSR-1 and Its 22G-RNA Cofactors Are Required for Holocentric Chromosome Segregation. *Cell* **139**, 123–34 (2009).
14. Seth, M. *et al.* The Coding Regions of Germline mRNAs Confer Sensitivity to Argonaute Regulation in C. elegans. *Cell Reports* **22**, 2254–2264 (2018).
15. Shen, E.-Z. *et al.* Identification of piRNA Binding Sites Reveals the Argonaute Regulatory Landscape of the C. elegans Germline. *Cell* **172**, 937–951.e18 (2018).
16. Zhang, D. *et al.* The piRNA targeting rules and the resistance to piRNA silencing in endogenous genes. *Science* **359**, 587–592 (2018).

17. Seth, M. *et al.* The C. elegans CSR-1 Argonaute Pathway Counteracts Epigenetic Silencing to Promote Germline Gene Expression. *Developmental Cell* **27**, 656–63 (2013).
18. Wedeles, C. J., Wu, M. Z. & Claycomb, J. M. Protection of Germline Gene Expression by the C. elegans Argonaute CSR-1. *Developmental Cell* **27**, 664–71 (2013).
19. Aoki, K., Moriguchi, H., Yoshioka, T., Okawa, K. & Tabara, H. In vitro analyses of the production and activity of secondary small interfering RNAs in C. elegans. *The EMBO Journal* **26**, 5007–5019 (2007).
20. Cecere, G., Hoersch, S., O'Keeffe, S., Sachidanandam, R. & Grishok, A. Global effects of the CSR-1 RNA interference pathway on the transcriptional landscape. *Nature Structural and Molecular Biology* **21**, nsmb.2801 (2014).
21. Conine, C. C. *et al.* Argonautes Promote Male Fertility and Provide a Paternal Memory of Germline Gene Expression in C. elegans. *Cell* **155**, 1532–1544 (2013).
22. Gerson-Gurwitz, A. *et al.* A Small RNA-Catalytic Argonaute Pathway Tunes Germline Transcript Levels to Ensure Embryonic Divisions. *Cell* **165**, 396–409 (2016).
23. Gu, W. *et al.* Distinct Argonaute-Mediated 22G-RNA Pathways Direct Genome Surveillance in the C. elegans Germline. *Molecular Cell* **36**, 231 244 (2009).
24. Phillips, C. M., Montgomery, T. A., Breen, P. C. & Ruvkun, G. MUT-16 promotes formation of perinuclear mutator foci required for RNA silencing in the C. elegans germline. *Gene Dev* **26**, 1433–44 (2012).
25. Zhang, C. *et al.* mut-16 and other mutator class genes modulate 22G and 26G siRNA pathways in Caenorhabditis elegans. *Proc National Acad Sci* **108**, 1201–1208 (2011).
26. Preston, M. A. *et al.* Unbiased screen of RNA tailing activities reveals a poly(UG) polymerase. *Nat Methods* **16**, 437–445 (2019).
27. Shukla, A. *et al.* poly(UG)-tailed RNAs in genome protection and epigenetic inheritance. *Nature* 1–6 (2020) doi:10.1038/s41586-020-2323-8.
28. Yigit, E. *et al.* Analysis of the C. elegans Argonaute Family Reveals that Distinct Argonautes Act Sequentially during RNAi. *Cell* **127**, 747–757 (2006).
29. Kim, J. K. *et al.* Functional Genomic Analysis of RNA Interference in C. elegans. *Science* **308**, 1164–1167 (2005).
30. Tops, B. B. J. *et al.* RDE-2 interacts with MUT-7 to mediate RNA interference in Caenorhabditis elegans. *Nucleic Acids Res* **33**, 347–355 (2005).
31. Chen, C.-C. G. *et al.* A Member of the Polymerase β Nucleotidyltransferase Superfamily Is Required for RNA Interference in C. elegans. *Curr Biol* **15**, 378–383 (2005).
32. Updike, D. L., Knutson, A. K., Egelhofer, T. A., Campbell, A. C. & Strome, S. Germ-granule components prevent somatic development in the C. elegans germline. *Curr Biology Cb* **24**, 970–5 (2014).
33. Campbell, A. C. & Updike, D. L. CSR-1 and P granules suppress sperm-specific transcription in the C. elegans germline. *Development* **142**, 1745–1755 (2015).

34. Barucci, G. *et al.* Small-RNA-mediated transgenerational silencing of histone genes impairs fertility in piRNA mutants. *Nat Cell Biol* **22**, 1–11 (2020).
35. Aeschimann, F., Xiong, J., Arnold, A., Dieterich, C. & Großhans, H. Transcriptome-wide measurement of ribosomal occupancy by ribosome profiling. *Methods* **85**, 75–89 (2015).
36. Schisa, J. A., Pitt, J. N. & Priess, J. R. Analysis of RNA associated with P granules in germ cells of C. elegans adults. *Dev Camb Engl* **128**, 1287–98 (2001).
37. Lee, C.-Y. S. *et al.* Recruitment of mRNAs to P granules by condensation with intrinsically-disordered proteins. *Elife* **9**, e52896 (2020).
38. Pechmann, S. & Frydman, J. Evolutionary conservation of codon optimality reveals hidden signatures of cotranslational folding. *Nat Struct Mol Biol* **20**, 237–43 (2012).
39. Presnyak, V. *et al.* Codon optimality is a major determinant of mRNA stability. *Cell* **160**, 1111–24 (2015).
40. Stadler, M. & Fire, A. Wobble base-pairing slows in vivo translation elongation in metazoans. *RNA* **17**, 2063–2073 (2011).
41. Duret, L. tRNA gene number and codon usage in the C. elegans genome are co-adapted for optimal translation of highly expressed genes. *Trends Genet* **16**, 287–289 (2000).
42. Novoa, E. M. & Pouplana, L. R. de. Speeding with control: codon usage, tRNAs, and ribosomes. *Trends Genet* **28**, 574–581 (2012).
43. Tuller, T. *et al.* An Evolutionarily Conserved Mechanism for Controlling the Efficiency of Protein Translation. *Cell* **141**, 344–354 (2010).
44. Gigant, E. *et al.* Inhibition of ectopic microtubule assembly by the kinesin-13 KLP-7MCAK prevents chromosome segregation and cytokinesis defects in oocytes. *Dev Camb Engl* **144**, 1674–1686 (2017).
45. Ruijtenberg, S. *et al.* mRNA structural dynamics shape Argonaute-target interactions. *Nat Struct Mol Biol* **27**, 790–801 (2020).
46. Małecka, E. M. & Woodson, S. A. Ribosomes clear the way for siRNA targeting. *Nat Struct Mol Biol* **27**, 775–777 (2020).
47. Pule, M. N., Glover, M. L., Fire, A. Z. & Arribere, J. A. Ribosome clearance during RNA interference. *Rna* **25**, rna.070813.119 (2019).
48. Sun, Y. *et al.* Ribosomes guide pachytene piRNA formation on long intergenic piRNA precursors. *Nat Cell Biol* **22**, 200–212 (2020).
49. Wu, H. *et al.* Plant 22-nt siRNAs mediate translational repression and stress adaptation. *Nature* 1–5 (2020) doi:10.1038/s41586-020-2231-y.
50. Iwakawa, H. *et al.* Ribosome stalling caused by the Argonaute-miRNA-SGS3 complex regulates production of secondary siRNA biogenesis in plants. *BioRxiv* (2020) doi:10.1101/2020.09.10.288902.

51. Quarato, P. *et al.* Argonaute catalytic activity is required for maternal mRNA clearance in embryos. *Biorxiv* 2020.02.03.919597 (2020) doi:10.1101/2020.02.03.919597.
52. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
53. Reis, M. dos, Savva, R. & Wernisch, L. Solving the riddle of codon usage preferences: a test for translational selection. *Nucleic Acids Res* **32**, 5036–5044 (2004).
54. Cecere, G., Zheng, G. X. Y., Mansisidor, A. R., Klymko, K. E. & Grishok, A. Promoters Recognized by Forkhead Proteins Exist for Individual 21U-RNAs. *Mol Cell* **47**, 734–745 (2012).
55. Holdorf, A. D. *et al.* WormCat: An Online Tool for Annotation and Visualization of Caenorhabditis elegans Genome-Scale Data. *Genetics* **214**, 279–294 (2019).
56. Brenner, S. The genetics of Caenorhabditis elegans. *Genetics* **77**, 71–94 (1974).
57. Paix, A., Folkmann, A., Rasoloson, D. & Seydoux, G. High Efficiency, Homology-Directed Genome Editing in Caenorhabditis elegans Using CRISPR-Cas9 Ribonucleoprotein Complexes. *Genetics* **201**, 47–54 (2015).
58. Dokshin, G. A., Ghanta, K. S., Piscopo, K. M. & Mello, C. C. Robust Genome Editing with Short Single-Stranded and Long, Partially Single-Stranded DNA Donors in Caenorhabditis elegans. *Genetics* **210**, 781–787 (2018).
59. Kamath, R. S. & Ahringer, J. Genome-wide RNAi screening in Caenorhabditis elegans. *Methods* **30**, 313–321 (2003).
60. Koster, J. & Rahmann, S. Snakemake–a scalable bioinformatics workflow engine. *Bioinformatics* **28**, 2520–2522 (2012).
61. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *Embnet J* **17**, 10–12 (2011).
62. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357–359 (2012).
63. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
64. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
65. Neph, S. *et al.* BEDOPS: high-performance genomic feature operations. *Bioinformatics* **28**, 1919–1920 (2012).
66. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2013).
67. Ramírez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**, W160–W165 (2016).
68. Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods* **17**, 261–272 (2020).

69. Puigbò, P., Bravo, I. G. & Garcia-Vallve, S. CAIcal: A combined set of tools to assess codon usage adaptation. *Biol Direct* **3**, 38 (2008).
70. Chan, P. P. & Lowe, T. M. Gene Prediction, Methods and Protocols. *Methods Mol Biology Clifton N J* **1962**, 1–14 (2019).
71. Cock, P. J. A. *et al.* Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **25**, 1422–1423 (2009).
72. Poullet, P., Carpentier, S. & Barillot, E. myProMS, a web server for management and validation of mass spectrometry-based proteomic data. *Proteomics* **7**, 2553–2556 (2007).
73. Valot, B., Langella, O., Nano, E. & Zivy, M. MassChroQ: A versatile tool for mass spectrometry quantification. *Proteomics* **11**, 3572–3577 (2011).
74. Vizcaíno, J. A. *et al.* 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res* **44**, 11033–11033 (2016).

Figures

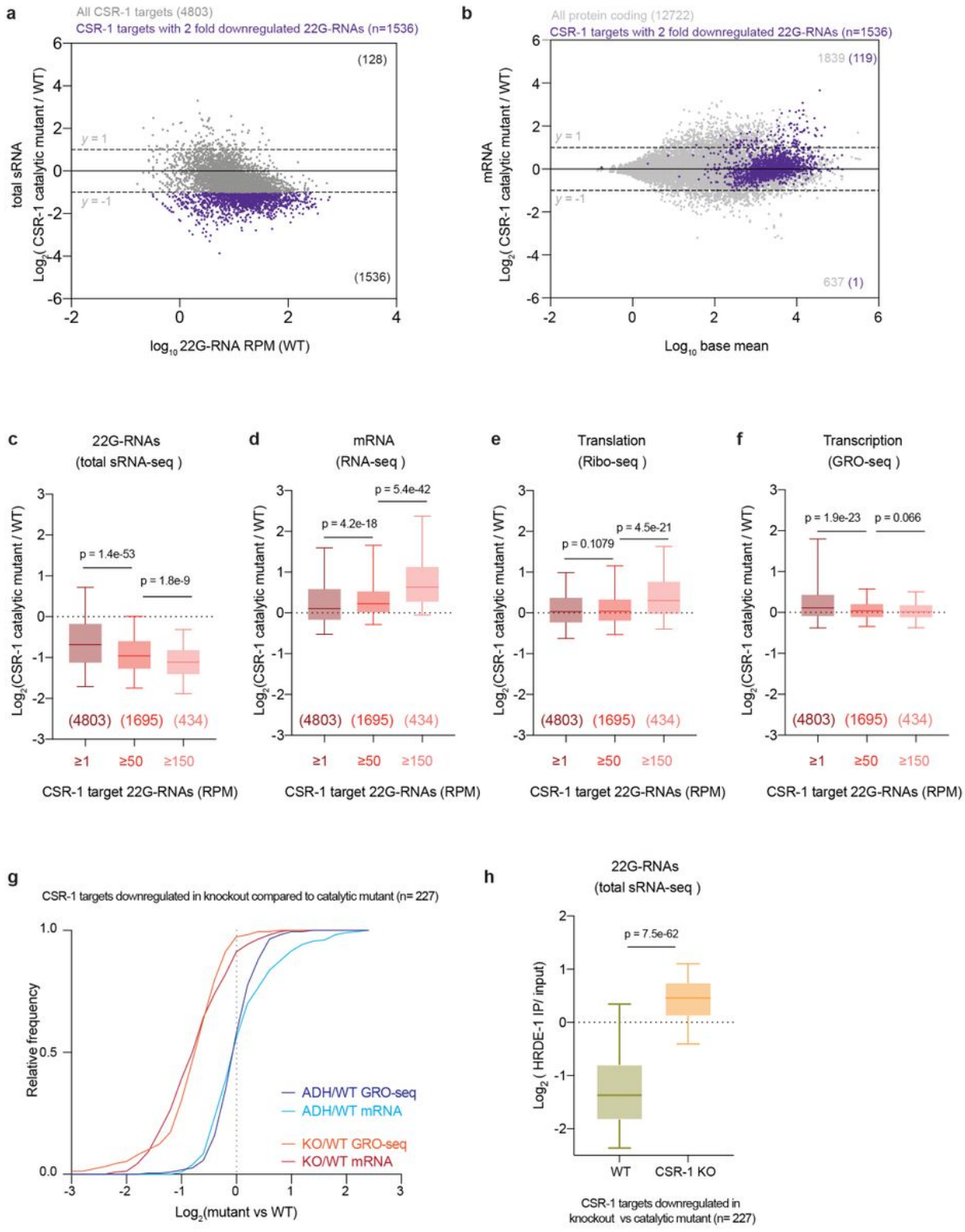

Figure 1

Defects in CSR-1 catalytic activity mainly impacts 22G-RNA abundance. a, MA-plot showing total 22G-RNA log2 fold change between CSR-1 ADH (catalytic mutant) and WT. The number in parenthesis indicates the number of misregulated genes ≥ 2-fold. The average from two biologically independent replicates is shown. b, MA-plot showing mRNA log2 fold change between CSR-1 ADH and WT. Genes with 22G-RNAs with 2-fold downregulation in CSR-1 catalytic mutant compared to WT are highlighted in

purple. The average from two biologically independent replicates is shown, with “base mean” computed using DESeq252. c-f, Box plots showing the log2 fold change of total 22G-RNAs (sRNA-seq), n = 2 biologically independent experiments c; or mRNAs (RNA-seq), n = 2 biologically independent experiments d; mRNAs engaged in translation (Ribo-seq), n = 3 biologically independent experiments e; and nascent RNAs (GRO-seq), n = 2 biologically independent experiments f, in CSR-1 ADH compared to WT strain. The distribution for the CSR-1 targets with 22G-RNA in CSR-1 IP ≥ 1 RPM, ≥ 50 RPM, or ≥ 150 RPM is shown (gene list in Supplemental Table 1). Box plots display median (line), first and third quartiles (box), and 5th /95th percentile value (whiskers). g, Cumulative frequency distribution for CSR-1 targets downregulated in CSR-1 KO compared to the CSR-1 ADH in GRO-seq (gene list in Supplemental Table 1). The comparison shows GRO-seq (p = 1.6e-49) and RNA-seq (p = 4.2e-37) for CSR-1 KO or CSR-1 ADH compared to WT. h, Box plots as in (c) showing the log2 fold change of 22G-RNAs (sRNA-seq) in HRDE-1 IPs compared to input in WT, CSR-1 KO (n = 2 biologically independent experiments).

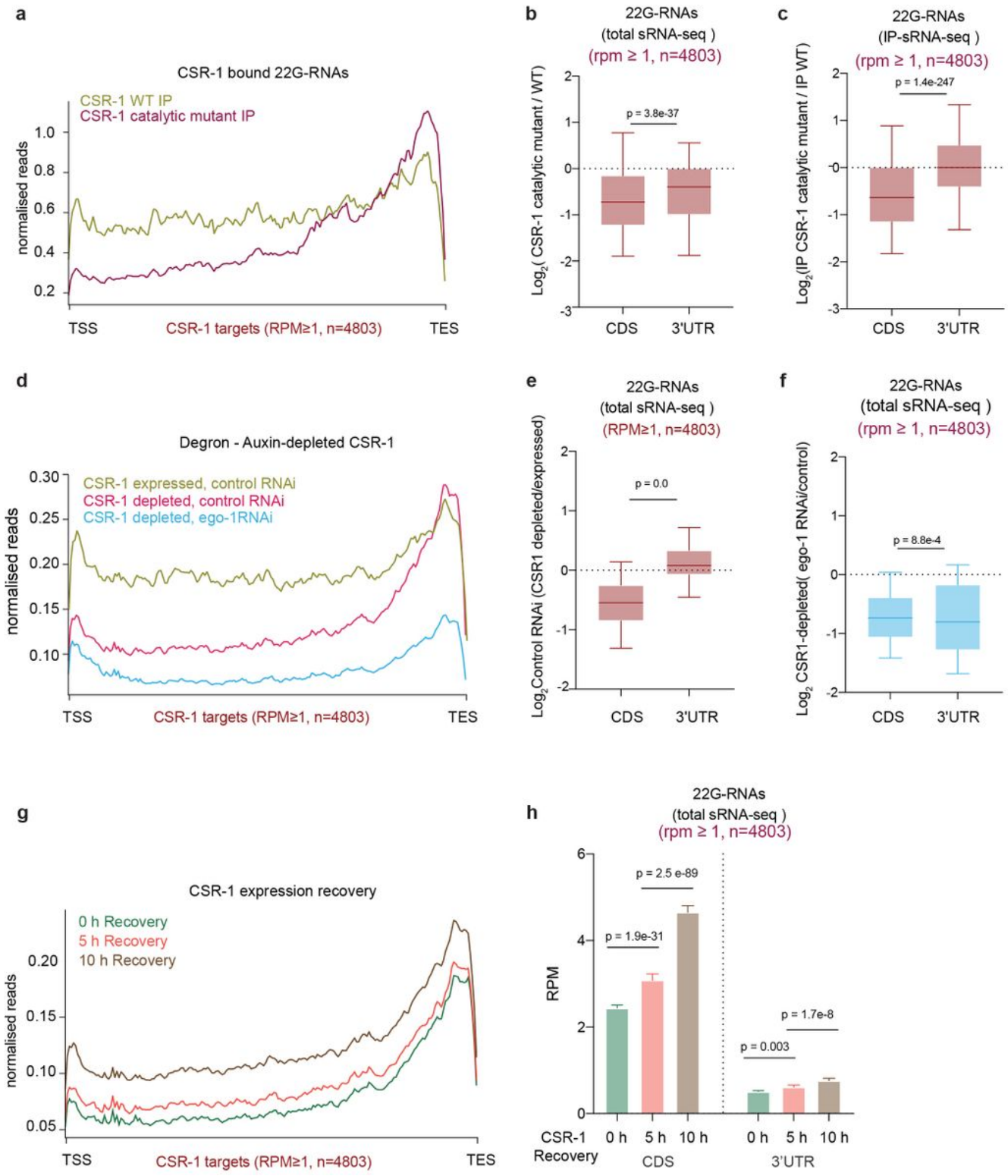

Figure 2

CSR-1 catalytic activity is required for the biogenesis of 22G-RNAs on the coding sequence. a, Metaprofile analysis showing the distribution of normalized 22G-RNA (sRNA-seq) reads (RPM) along all CSR-1 targets (22G-RNA ≥1 RPM) in WT CSR-1 or CSR-1 ADH immunoprecipitation (IP). TSS indicates the transcriptional start site, TES indicates the transcriptional termination site. b, Box plots showing the log2 fold change of the amount of total 22G-RNA generated from CDS and 3' UTR of CSR-1 targets (22G-RNA

≥1 RPM) in CSR-1 ADH compared to WT. c, Box plots showing the log2 fold change of the amount of 22G-RNA generated from CDS and 3’ UTR of CSR-1 targets (22G-RNA ≥1 RPM) bound in CSR-1 ADH IP compared to WT CSR-1 IP. Box plots display median (line), first and third quartiles (box), and 5th /95th percentile value (whiskers). Two-tailed P values were calculated using Mann–Whitney–Wilcoxon tests. n = 2 biologically independent experiments. d, Metaprofile analysis as in (a) showing the distribution of normalized total 22G-RNA reads (RPM) across CSR-1 targets (22G-RNA ≥1 RPM) upon ego-1 RNAi and Control RNAi treated in Auxin-depleted CSR-1 degron background (CSR-1 depleted) and degron control (CSR-1 expressed). e, Box-plot as in (b) showing the log2 fold change in the amount of 22G-RNA generated from CDS and 3’ UTR of CSR-1 targets (22G-RNA ≥1 RPM) in Auxin-depleted CSR-1 compared to non-depleted CSR-1 degron control in control RNAi background. f, Box-plot showing the log2 fold change in the amount of 22G-RNA generated from CDS and 3’ UTR of CSR-1 targets (22G-RNA ≥1 RPM) in ego-1 RNAi compared to control RNAi treated in Auxin-depleted CSR-1 degron background, n = 2 biologically independent experiments. g, Metaprofile analysis as in (a) showing the distribution of normalized total 22G-RNA reads (RPM) across CSR-1 targets (22G-RNA ≥1 RPM) after depletion of CSR-1, in the CSR-1 degron strain for 38 h by growing on auxin containing plates and recovery of CSR-1 expression by transferring on plates without auxin for 0 h, 5 h, and 10 h. h, Bar graph representing the data in (F) showing the median RPM of 22G-RNAs with 95 % confidence interval generated from CDS and 3’ UTR of CSR-1 targets (22G-RNA ≥1 RPM) for CSR-1 expression recovered for 0, 5 or 10 h. Two-tailed P values were calculated using Mann–Whitney–Wilcoxon tests, n = 2 biologically independent experiments.

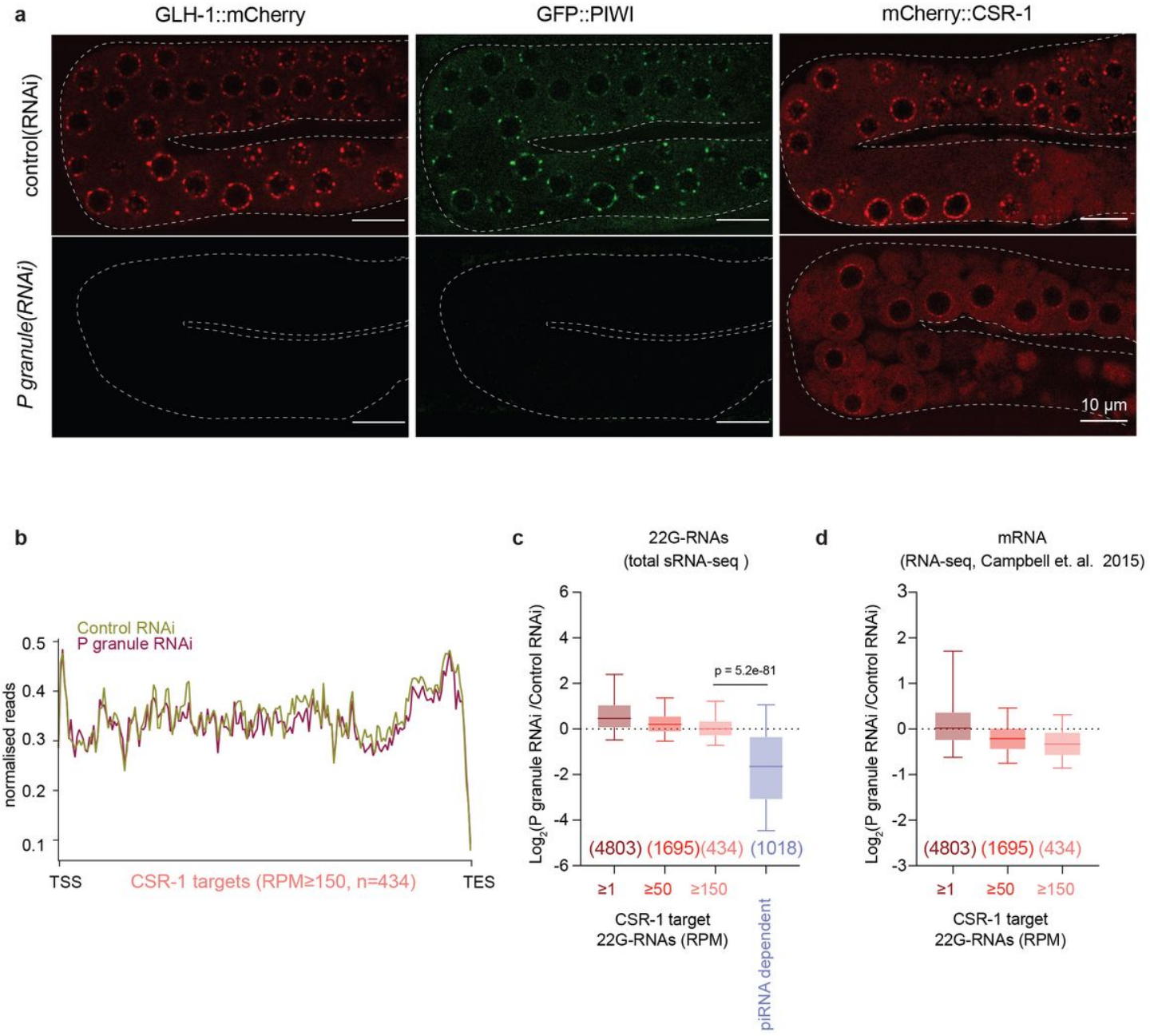

Figure 3

Biogenesis of CSR-1 22G-RNAs and the regulation of their targets do not require germ granule localization. a, Fluorescent images of animals expressing P granule marker GLH-1:mCherry, GFP:PIWI or mCherry:CSR-1 treated with either control RNAi or P granule RNAi (pgl-1, pgl-3, glh-1, and glh-4). CSR-1 is localized in both cytosol and P granule, and upon RNAi treatment, P granule localization of CSR-1 is lost while maintaining cytosolic localization. b, Metaprofile analysis showing the distribution of normalized 22G-RNA (sRNA-seq) reads (RPM) across CSR-1 targets (22G-RNA ≥ 150 RPM) upon control RNAi and P granule RNAi. TSS indicates the transcriptional start site, TES indicates the transcriptional termination site. c, Box plots showing the log2 fold change of total 22G-RNA (sRNA-seq) upon P granule RNAi compared to control RNAi. The distribution for the CSR-1 targets with 22G-RNA in CSR-1 IP ≥ 1 RPM, ≥ 50 RPM, or ≥ 150 RPM and piRNA dependent 22G-RNA target genes34 are shown. Box plots display

median (line), first and third quartiles (box), and 5th /95th percentile value (whiskers). Two-tailed P values were calculated using Mann–Whitney–Wilcoxon tests, n = 2 biologically independent experiments. d, Box plots showing the log2 fold change of mRNA expression (RNA-seq from Campbell et.al.33) upon P granule RNAi compared to control RNAi. The distribution for the CSR-1 targets with 22G-RNA in CSR-1 IP ≥ 1 RPM, ≥ 50 RPM, or ≥ 150 RPM is shown.

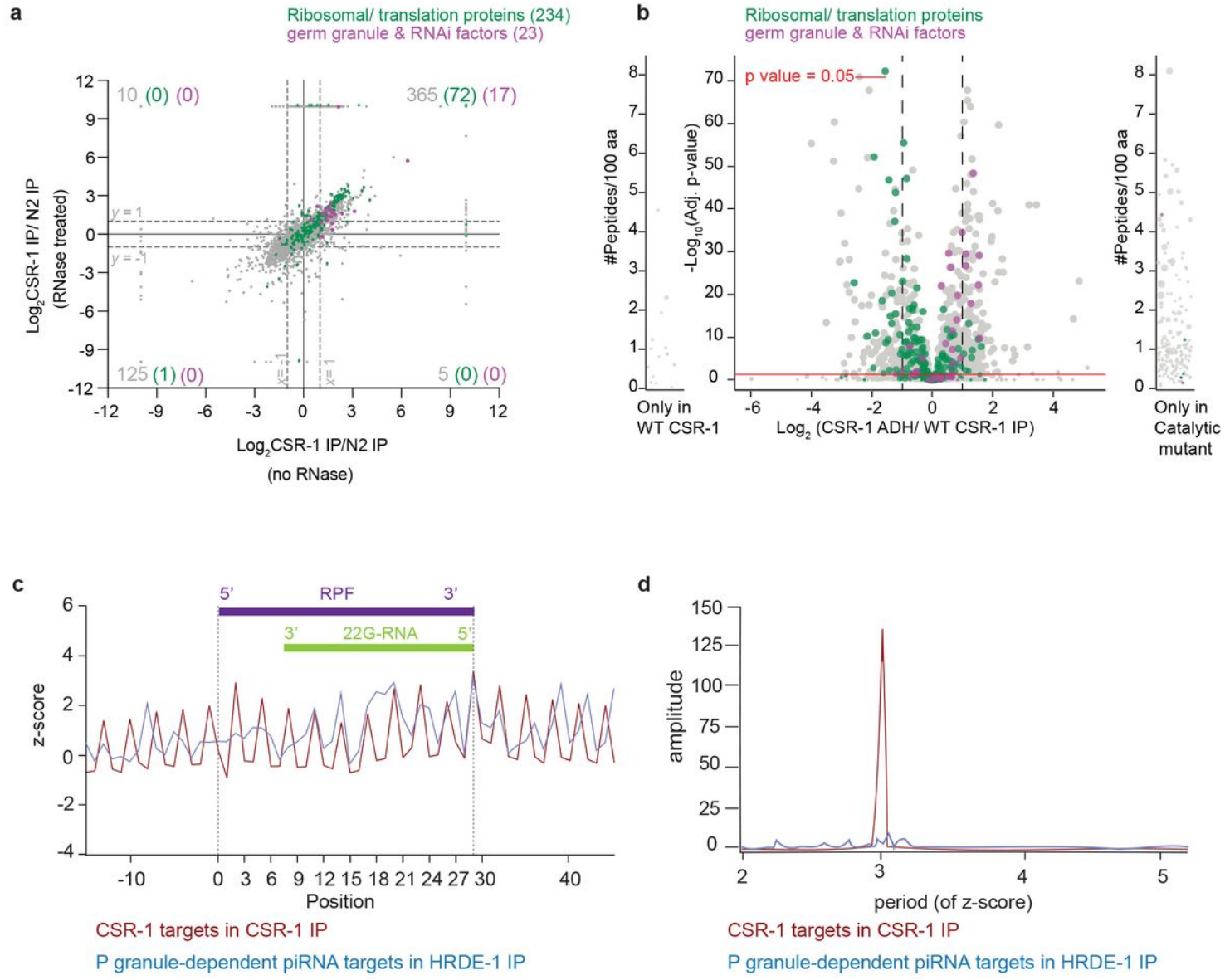

Figure 4

CSR-1 22G-RNAs are synthesized concomitantly with mRNA translation. a, Scatter plot comparing the log2 fold changes in CSR-1 interactors (IP-MS/MS) to control IPs performed in WT strain in the absence of RNase treatment34 (x-axis) to the IPs performed after RNase treatment (Supplemental Table 2). Ribosomal proteins and translation regulators are highlighted in green, and germ granule proteins, including RNAi factors, are highlighted in magenta. Number in grey refers to all interactors with log2 fold change of ≥1 and p-value ≤ 0.05 for each quadrant. The number in parenthesis is for ribosomal and translation associated proteins enriched, and granule & RNAi factors. n = 4 biologically independent

experiments. b, Volcano plot showing log2 fold change in enrichment values and corresponding significance levels for proteins co-purifying with CSR-1 ADH compared to WT CSR-1 (Supplemental Table 3). Ribosomal proteins and translation regulators are highlighted in green. Germ granule proteins, including RNAi factors, are highlighted in magenta. The size of the dots is proportional to the number of peptides used for the quantification. The linear model was used to compute the protein quantification ratio, and the red horizontal line indicates the two-tailed p-value = 0.05. n = 4 biologically independent experiments. c, Plot showing the z-score for read density for the of 5’ terminus of 22G-RNAs for CSR-1 targets (RPM≥1) in CSR-1 IP and P granule dependent piRNA targets in HRDE-1 IP (gene list in Supplemental Table 1) relative to the start of 29-nt long Ribosomal protected fragments (RPF). d, Periodogram based on Fourier transform for read-density around RPF 5’ start position showing periodicity of CSR-1 22G-RNAs phasing with RPFs. P granule dependent piRNA targets in HRDE-1 IP was used as control.

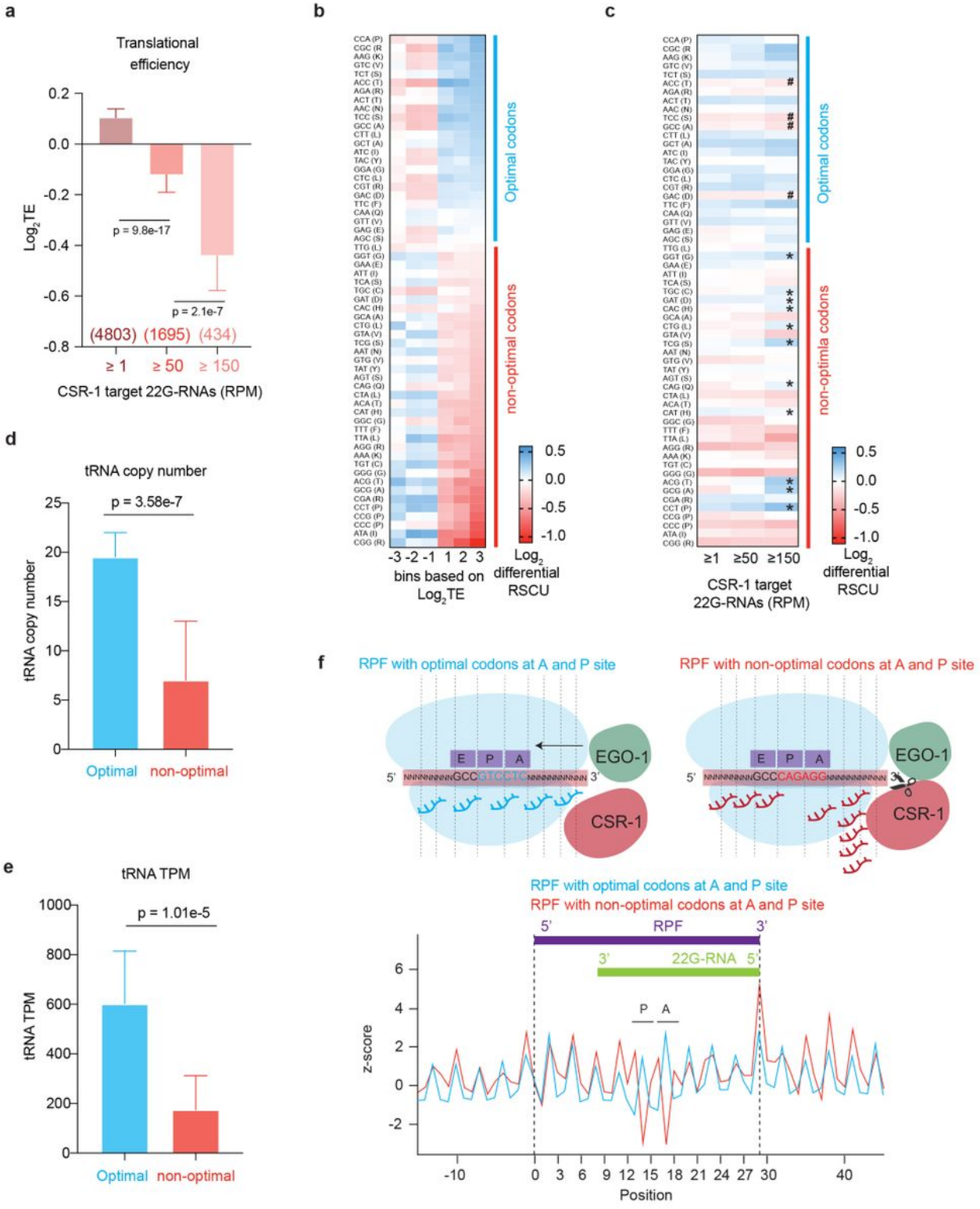

Figure 5

Highly translated mRNAs and optimal codons negatively correlate with CSR-1 22G RNA abundance. a, Translation efficiency Log2 (RPF TPM/ mRNA TPM) for CSR-1 targets in WT strain. The distribution for the CSR-1 targets 22G-RNA in CSR-1 IP ≥ 1 RPM, ≥ 50 RPM, or ≥ 150 RPM is shown. Bars represent median with 95 % confidence interval. Two-tailed P values were calculated using Mann–Whitney–Wilcoxon tests, n = 2 biologically independent experiments. b, heat map showing log2 fold change in

Relative Synonymous Codon Usage (RSCU) for all protein-coding genes categorized by increasing translational efficiency compared to genes showing neutral translational efficiency of 1, as explained in methods. Codons are arranged in order of decreasing RSCU (top to bottom) in the category of log2 TE ≥3. The blue line highlights optimal codons in genes with high TE, and the red line highlights non-optimal codons. c, heatmap similar to (b) showing log2 fold change in Relative synonymous codon usage for all CSR-1 targets (≥ 1, ≥ 50, and ≥150 RPM of 22G-RNA) compared to genes showing neutral translational efficiency as explained in methods. '*' marks over-used non-optimal codons by CSR-1 targets and '#' marks under-used optimal codons. d, Bar graph showing the median copy numbers for tRNAs for optimal or non-optimal codons with a 95 % confidence interval, e, Bar graph showing the median TPMs for tRNAs from the GRO-seq dataset for WT strain at the late l4 larval stage (44h) for optimal or non-optimal codons with a 95 % confidence interval. For codons with missing cognate tRNA, values have been adjusted by considering tRNA copy numbers and TPMs for tRNA recognizing these codons by wobble base pairing 53. (see Extended Data Fig. 5b, c for non-adjusted values). Two-tailed P values were calculated using Mann–Whitney–Wilcoxon tests, n = 2 biologically independent experiments. f, Plot showing the z-score for the read density for the of 5' terminus of 22G-RNAs for CSR-1 targets (RPM≥150) relative to the start of 29-nt long Ribosomal protected fragments (RPF) with either optimal or non-optimal codons at their P and A site. The scheme shows possible initiation of 22G-RNA biogenesis after a slicing event downstream of RPF with bad codons at the A and P site.

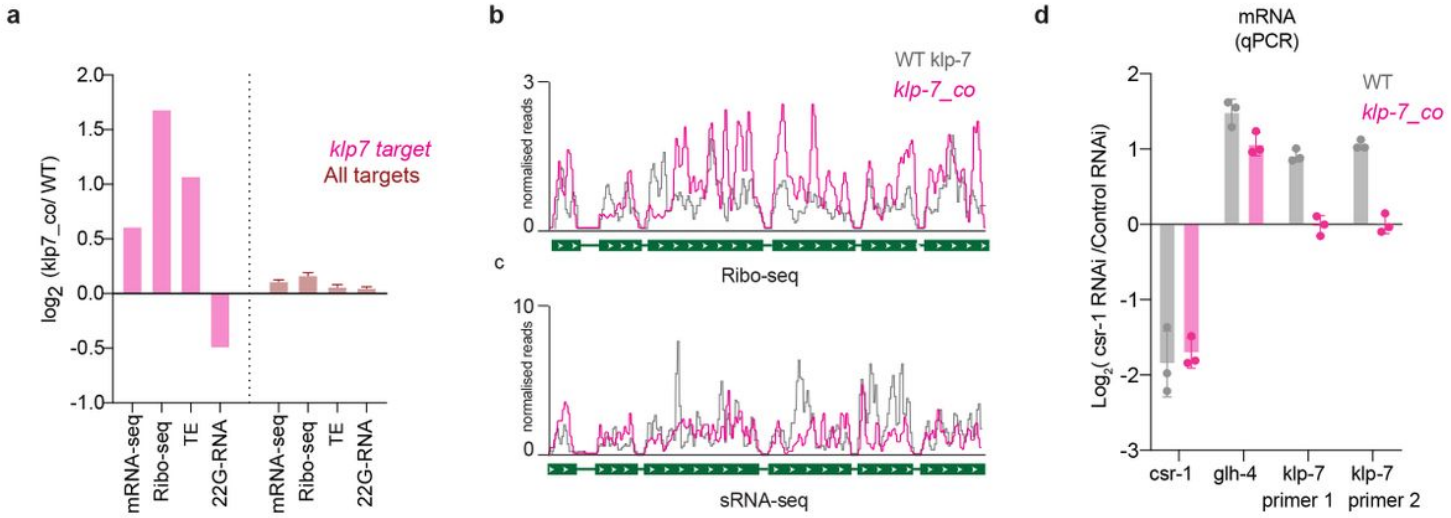

Figure 6

Increasing optimal codon usage of CSR-1 target decreases small RNA production. a, Plot showing log2 fold change for normalized reads for mRNAs from RNA-seq, RPF from Ribo-seq, and 22G-RNAs from sRNA-seq and differential Translational efficiency for klp-7 (top CSR-1 target) and all CSR-1 targets (RPM≥1, n=4803) for the strain with codon-optimized klp-7 (klp7_co) compared to WT strain. n = 2 biologically independent experiments. b-c, A genomic view of klp-7 showing of Ribo-seq (b) and 22G-RNAs (c), normalized reads for the strain with codon-optimized klp-7 (klp7_co) compared to WT strain. d, log2 fold change in expression of csr-1, glh-4, and klp-7 upon csr-1 RNAi compared to control RNAi by

qPCR in the WT strain and klp-7_co (strain with klp-7 codon optimization). n = 3 biologically independent experiments.

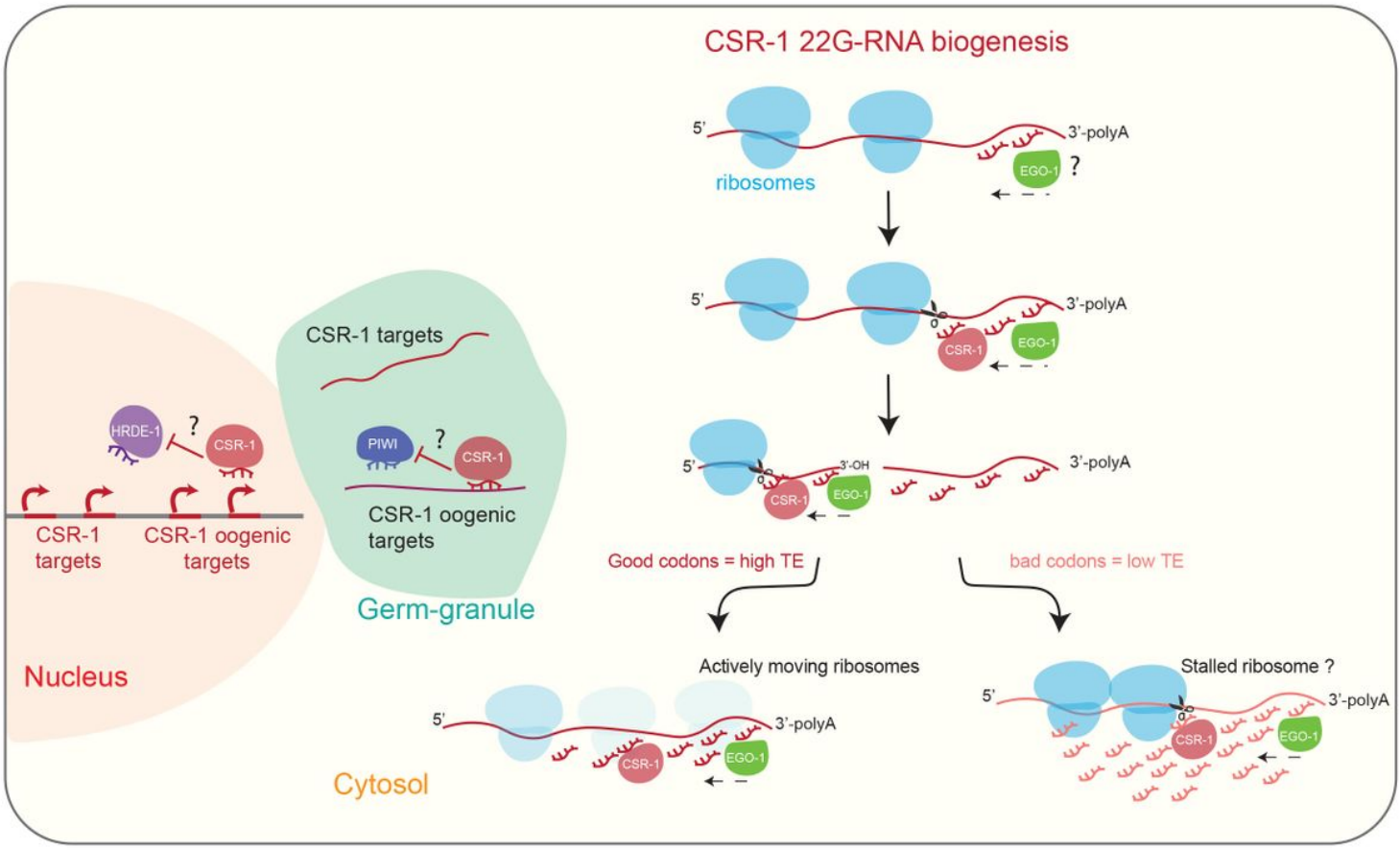

Figure 7

Model illustrating biogenesis of CSR-1-22G-RNA in the cytosol. CSR-1 targets most of the germline expressed genes. CSR-1 22G-RNAs are produced from mRNAs engaged in translation in the cytosol. We propose EGO-1 initiate 22G-RNA biogenesis at the 3'UTR on every actively translating mRNAs or by being recruited on specific 3'UTR sequences by yet unknown mechanism. However, to produce 22G-RNAs on coding sequence, CSR-1 slicing activity is required on the mRNA template. Codon usage and translation efficiency antagonistically regulate levels of 22G-RNAs production on different CSR-1 targets. We propose that CSR-1 can interact with ribosomes and the slicing event is more biased downstream of a possible stalled ribosome occupying a bad codon site. CSR-1 slicer activity can regulate gene-expression of few top targets, which further depends on the 22G-RNA levels. Additionally, CSR-1 can protect a set of mainly oogenic genes from piRNA mediated HRDE-1 transcriptional silencing, in a catalytic independent manner.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementaltables.xlsx

- ExtendedData1.jpg
- ExtendedData2.jpg
- ExtendedData3.jpg
- ExtendedData4.jpg
- ExtendedData5.jpg
- ExtendedData6.jpg
- ExtendedData7.jpg
- Supplementaltables.xlsx