No-go decay dampens zinc finger protein mRNAs in vertebrates

Kota Ishibashi[1], Yuichi Shichino[2], Peixun Han[2], Toshifumi Inada[3], Seisuke Kimura[4,5], Shintaro Iwasaki[2,6], Yuichiro Mishima[1,2‡]

[1] Department of Frontier Life Sciences, Faculty of Life Sciences, Kyoto Sangyo University, Kita-ku, Kyoto 603-8555, Japan

[2] RNA Systems Biochemistry Laboratory, RIKEN Cluster for Pioneering Research, Wako, Saitama 351-0198, Japan

[3] Institute of Medical Science, The University of Tokyo, Minato-Ku, 108-8639, Japan

[4] Department of Industrial Life Sciences, Faculty of Life Sciences, Kyoto Sangyo University, Kita-ku, Kyoto 603-8555, Japan

[5] Center for Plant Sciences, Kyoto Sangyo University, Kita-ku, Kyoto 603-8555, Japan

[6] Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa, Chiba 277-8561, Japan

‡To whom correspondence should be addressed

Email: mishima@cc.kyoto-su.ac.jp

Running title: NGD dampens zinc finger protein mRNAs in vertebrates

Abstract

Quality control of translation is crucial for maintaining cellular and organismal homeostasis. Obstacles in translation elongation induce ribosome collision, which is monitored by multiple sensor mechanisms. The E3 ubiquitin ligase ZNF598 recognizes collided ribosomes, triggering ribosome-associated quality control (RQC) to rescue stalled ribosomes and no-go decay (NGD) to degrade stall-prone mRNAs in eukaryotes. However, the role of RQC and NGD in maintaining the translational homeostasis of endogenous mRNAs has remained largely unclear. In this study, we investigated endogenous substrate mRNAs of NGD using the maternal-to-zygotic transition (MZT) of zebrafish embryos as a model system. Combining RNA-Seq analysis of *znf598* mutant embryos with disome profiling and reporter assays, we revealed that mRNAs encoding the C2H2-type zinc finger domain (C2H2-ZF) frequently cause ribosome collisions and induce NGD during the MZT. Furthermore, we showed that mRNAs encoding C2H2-ZF are degraded by ZNF598 in human cells. Our results suggest that NGD maintains the quality of the translatome during protein evolution and animal development by mitigating the risk of ribosome collision at the abundantly present C2H2-ZF mRNA sequences.

(173 words)

Introduction

The translation of mRNA into protein is a fundamental and the costliest process in the cell (Hershey *et al*, 2019). To minimize the cost of translation and maintain proteome integrity, cells harness multiple quality control mechanisms that prevent faulty translation. The ribosome lies at the core of such translational quality control mechanisms; aberrant positioning, movements, and structure of ribosomes during mRNA translation serve as cues to trigger specific quality control pathways (D'Orazio & Green, 2021; Morris *et al*, 2021). The failure of translational quality control pathways produces harmful proteins and protein aggregates, leading to pathogenic outcomes (Stein *et al*, 2022; Choe *et al*, 2016; Lu, 2022).

Persistent stalling of the ribosome during the translation elongation process is detrimental to cells, as it perturbs ribosome traffic on mRNA and exposes the premature polypeptide to the intracellular environment. Ribosomes stall when encountering multiple translational challenges, including tRNA shortage, decoding defects, obstruction by RNA secondary structures, chemical damage to RNA, inhibitory nascent peptide sequences, and combinations thereof (Schuller & Green, 2018; Simms *et al*, 2014; Wilson *et al*, 2016). Eventually, the trailing ribosome catches up and collides with the leading stalled ribosome, forming an unusual ribosome dimer structure called a disome (or trisome if three ribosomes collide) (Matsuo *et al*, 2020; Ikeuchi *et al*, 2019b; Juszkiewicz *et al*, 2018; Tesina *et al*, 2020). Cellular machinery recognizes disomes in multiple ways to solve or mitigate translational problems (Meydan & Guydosh, 2021; Kim & Zaher, 2021).

In eukaryotes, the faulty translation leading to disome formation is terminated by splitting the stalled ribosome and degrading the premature nascent polypeptide by ribosome-

associated quality control (RQC) (Sitron & Brandman, 2020; Filbeck *et al*, 2022; Inada, 2020). In addition, no-go decay (NGD) degrades the mRNA on which ribosomes stall, preventing the accumulation of stall-prone mRNAs (Ikeuchi *et al*, 2019a; Simms *et al*, 2017; Doma & Parker, 2006). Although the coordination and order of RQC and NGD require further investigation, ubiquitination of the stalled ribosome in disome by ZNF598 (Hel2 in yeast), a conserved E3 ubiquitin ligase, is the most upstream step in both pathways (Sundaramoorthy *et al*, 2017; Garzia *et al*, 2017; Matsuo *et al*, 2017; Juszkiewicz & Hegde, 2017; Juszkiewicz *et al*, 2018). ZNF598/Hel2 ubiquitinates ribosomal small subunit proteins RPS10/eS10 and RPS20/uS10 at conserved lysine residues, followed by splitting of the stalled ribosome by the RQC-trigger (RQT) complex (also known as the ASC-1 complex; ASCC) (Juszkiewicz *et al*, 2020; Hashimoto *et al*, 2020; Best *et al*, 2023; Matsuo *et al*, 2020). Ubiquitination of the stalled ribosome by ZNF598/Hel2 is also required for inducing NGD in yeast and worms (Monem *et al*, 2023; Ikeuchi *et al*, 2019b; Tomomatsu *et al*, 2023). In these organisms, mRNA degradation via NGD is partly mediated by the ubiquitin-binding endonuclease Cue2/NONU-1 (Glover *et al*, 2020; D'Orazio *et al*, 2019; Tomomatsu *et al*, 2023). Alternatively, stall-prone mRNA might be degraded via GIGYF family proteins that bind to ZNF598/Hel2 (Veltri *et al*, 2022; Weber *et al*, 2020; Morita *et al*, 2012). Whether vertebrate NGD occurs coincidently with RQC remains controversial (Arthur *et al*, 2015; Juszkiewicz & Hegde, 2017), but Znf598 is required for stall-induced mRNA decay in zebrafish embryos (Mishima *et al*, 2022). These observations suggest a conserved mechanism acting on RQC and NGD in eukaryotes.

While the molecular mechanism of NGD is being unveiled, its endogenous substrate

mRNAs remain largely uncharacterized. The molecular analysis of NGD was expedited using artificial stall-inducing sequences such as a repeat of rare arginine codons and a poly(A) stretch encoding polylysine codons (Chen *et al*, 2010; Glover *et al*, 2020; D'Orazio *et al*, 2019; Tomomatsu *et al*, 2023; Arthur *et al*, 2015; Simms *et al*, 2017; Ikeuchi *et al*, 2019b). Nevertheless, such sequences tend to be avoided in the transcriptome (Arthur *et al*, 2015; Sharp & Li, 1987). It has been proposed that RQC prevents translation of the poly(A) tail that emerges by premature termination of transcription at the coding sequence or introns (Garzia *et al*, 2017; Juszkiewicz & Hegde, 2017), and NGD discards chemically damaged mRNAs prone to ribosome stalling (Yan *et al*, 2019; Simms *et al*, 2014). However, experimental demonstration of NGD acting on endogenous mRNAs is still challenging. Given that only a handful of endogenous NGD target mRNAs have been identified thus far (Tomomatsu *et al*, 2023; Arthur *et al*, 2015), the NGD substrate repertoire needs to be comprehensively determined. Indeed, widespread distributions of disomes in the transcriptome imply the broad impact of NGD on gene expression (Han *et al*, 2020; Zhao *et al*, 2021; Meydan & Guydosh, 2020; Arpat *et al*, 2020). This situation is akin to the early days of nonsense-mediated decay (NMD) research: NGD was once regarded as a quality control pathway of aberrant splicing but was later revealed to have a broader impact on gene expression (Karousis & Mühlemann, 2022). Analogically, comprehensive identification of endogenous NGD target mRNAs in a model organism is needed to understand the physiological role of NGD.

Early embryogenesis is a unique model system to study mRNA decay *in vivo*. After fertilization, thousands of maternal mRNAs stored in eggs are degraded via multiple mRNA

decay pathways. The massive decay of maternal mRNA and concurrent zygotic genome activation (ZGA) constitute a conserved developmental transition in metazoa called the maternal-to-zygotic transition (MZT) (Tadros & Lipshitz, 2009; Despic & Neugebauer, 2018; Yartseva & Giraldez, 2015; Lee *et al*, 2014). The molecular mechanisms underpinning the MZT have been extensively characterized in zebrafish, where codon-mediated decay and zygotically transcribed miR-430 play pivotal roles in maternal mRNA clearance (Giraldez *et al*, 2006; Mishima & Tomari, 2016; Bazzini *et al*, 2016). Combining a zebrafish mutant of *znf598* and a reporter mRNA containing a ribosome stall sequence of human cytomegalovirus (hCMV), we previously showed that ribosome stalling induces mRNA decay in a Znf598-dependent manner, demonstrating that NGD is active during the MZT (Mishima *et al*, 2022). Codon-mediated decay was unaffected by the loss of Znf598, implying that NGD target mRNAs are distinct from previously identified maternal mRNAs degraded during the MZT (Mishima *et al*, 2022). However, the contribution of NGD in shaping the mRNA expression patterns during the MZT has yet to be explored.

In this study, we investigated endogenous substrate mRNAs of NGD during the MZT using *znf598* mutant zebrafish embryos. Our RNA-Seq analysis showed that Znf598 downregulates mRNAs encoding the C2H2-type zinc finger domain (C2H2-ZF) at the onset of the MZT. We combined disome profiling and reporter analysis to characterize the molecular basis of C2H2-ZF sequences that induced ribosome stalling and Znf598-dependent NGD. We also investigated whether ZNF598 degrades mRNAs encoding C2H2-ZF in human cells. Based on these results, we propose the role of NGD in maintaining the quality of the translatome during protein evolution and animal development.

Results

Identification of Znf598-dependent NGD target mRNAs during zebrafish MZT

To identify endogenous mRNAs constitutively degraded via NGD under physiological conditions, we focused on maternal mRNA clearance in zebrafish (Despic & Neugebauer, 2018). We reasoned that maternal mRNAs degraded via NGD should be stabilized in maternal-zygotic *znf598* mutant embryos (MZ*znf598*) that developed at a comparable speed to wild-type embryos during the cleavage and gastrulation stages (Figure 1A) (Mishima *et al*, 2022). We performed RNA-Seq analysis with wild-type and MZ*znf598* embryos at one- and six hours postfertilization (hpf) and applied two criteria to identify NGD target maternal mRNAs following our previous studies of maternal mRNA stability (Mishima & Tomari, 2016) (Figure 1B: only uniquely mapped reads were analyzed in this study). First, the target mRNA amount should be decreased more than twofold (q-value <0.01) from 1 hpf to 6 hpf in wild-type embryos. Second, the target mRNA amount should be increased more than 1.5-fold (q-value <0.01) at 6 hpf in MZ*znf598* embryos compared to wild-type embryos. Sixty-one mRNAs passed these criteria and were considered candidates for NGD target maternal mRNA (Figure 1C and Supplemental Table 1). These mRNAs were distinct from unstable maternal mRNAs degraded by maternal or zygotic decay pathways (Mishima & Tomari, 2016) (Supplemental Figure 1A). Conversely, mRNA degradation by maternal and zygotic decay pathways, including degradation by codon-mediated decay and miR-430, was unaffected by the loss of Znf598, as suggested by our previous reporter analysis (Supplemental Figure 1B and C) (Mishima *et al*, 2022).

We performed qRT–PCR analysis of selected mRNAs in wild-type and

MZ*znf598* embryos to validate the RNA-Seq data (Figures 1D). In this validation, we included two rescue experiments to test the dependency of mRNA degradation on the E3 ubiquitin ligase activity of Znf598; MZ*znf598* embryos were injected with mRNAs encoding full-length Znf598 or a mutant lacking the RING domain (Figures 1E), which is essential for NGD (Mishima *et al*, 2022; Ikeuchi *et al*, 2019b). Among the five candidate mRNAs tested, *znf236* and *znf970* were degraded in a RING domain-dependent manner. In contrast, three mRNAs (*ap3s2*, *mapk14b*, and *piezo2*) were upregulated in MZ*znf598* embryos even with the full-length Znf598 rescue. Time course qRT–PCR analysis further confirmed that *znf236* and *znf970* mRNAs were degraded during the MZT in wild-type embryos, and degradation was suppressed in MZ*znf598* embryos (Figure 1F). Notably, *znf236* and *znf970* encode proteins with tandem C2H2-ZFs (Figure 1G).

We further focused on *znf236* mRNA to determine whether its degradation occurred via NGD. First, we blocked the translation of *znf236* mRNA by injecting a specific morpholino oligonucleotide (MO) that masked the initiation codon from the ribosome to confirm translation dependency (Nasevicius & Ekker, 2000; Doma & Parker, 2006). *znf236* mRNA was stabilized in the presence of a specific MO, confirming that its degradation required active translation (Figure 1H). Second, we performed the poly(A) test (PAT) assay for *znf236* mRNA in wild-type and MZ*znf598* embryos at 6 hpf (Figures 1I and J) (Kusov *et al*, 2001; Mishima & Tomari, 2016). *znf236* mRNA was not subjected to active deadenylation, as observed with the NGD reporter mRNA (Mishima *et al*, 2022). These data indicate that a few maternal mRNAs are degraded after fertilization depending exclusively on NGD by Znf598, as exemplified by the *znf236* and *znf970* mRNAs.

Specific C2H2 zinc finger domains induce ribosome collision and NGD

Using *znf236* mRNA as a model endogenous NGD target mRNA, we next analyzed ribosome movement on this mRNA that induced NGD (Figure 2A). We first examined the monosome and disome profiling data in wild-type zebrafish embryos at 4 hpf (Han *et al*, 2020) to analyze ribosome stalling and collisions on *znf236* mRNA (see below for details about footprint analysis). Although the read amount was not abundant, disome footprints were mapped to *znf236* mRNA (Figure 2B). Of those mapped, 30.5% (18/59) were clustered between codons 169 to 282 (denoted as the A site of the leading ribosome). In addition, a higher monosome footprint density was observed upstream of the end of the disome-clustered region (0.3 footprints/codon) compared to the remaining ORF (0.15 footprints/codon). These data suggested that ribosomes stalled and collided on *znf236* mRNA.

To validate the ribosome stall sites on *znf236* mRNA, we performed a tandem luciferase reporter assay that captured ribosome elongation defects in zebrafish embryos (Mishima *et al*, 2022) (Figure 2C). We first divided the entire ORF of *znf236* into four fragments (1-4) with comparable lengths. We found that fragment 1 most strongly inhibited ribosome elongation compared to the others (Figures 2A and D). Further analysis of fragment 1 revealed that two smaller fragments (fragments 1B and 1D) inhibited elongation (Figures 2A and 2E). Notably, fragment 1B corresponded to the disome-clustered region (Figure 2B). This fragment encodes four C2H2-ZFs, and disome footprints were distributed around three C2H2-ZFs, ZF6, ZF7, and ZF8 (Figure 2F). Consistent with disome footprints, a fragment encoding three tandem C2H2-ZFs (ZF678) was sufficient to induce ribosome stalling (Figure

2G).

To investigate whether ribosome stalling was induced by a specific C2H2-ZF or a combined sequence context of ZF6, ZF7, and ZF8, we inserted three repeats of each C2H2-ZF into the tandem luciferase reporter. This analysis revealed that ZF6 and ZF8 induced ribosome stalling, whereas ZF7 did not (Figure 2G). We further examined the number of ZF6 repeats required to cause ribosome stalling (Figure 2H). A single copy of ZF6 was insufficient to inhibit elongation at a detectable level, as was the negative control EGFP ORF. We observed that ribosome stalling was strengthened incrementally as the number of ZF6 repeats increased. These results indicate that tandem organization of specific C2H2-ZF is required to induce ribosome stalling in zebrafish embryos.

To analyze the relationship between the ribosome stalling activity of C2H2-ZF sequences and the induction of NGD, we inserted C2H2-ZF sequences at the 3′ end of the sfGFP ORF and measured their effects on mRNA stability by mRNA injection and qRT–PCR (Figure 2I). As reported previously, this assay detected Znf598-dependent NGD induced by the ribosome stall sequence of hCMV *gp48* uORF2 (Mishima *et al*, 2022; Matheisl *et al*, 2015) (Figure 2J). When we tested the effect of *znf236* ORF 1A-1D fragments on mRNA stability, fragment 1B destabilized the reporter mRNA in wild-type but not MZ*znf598* embryos (Figure 2J). Fragment 1D tended to destabilize the reporter mRNA by Znf598, consistent with the result of the dual luciferase assay (Figure 2E). In contrast, fragment 1C did not induce mRNA decay, and fragment 1A destabilized the reporter mRNA independently of Znf598. Furthermore, a tandem repeat of ZF6 was sufficient to recapitulate Znf598-dependent mRNA destabilization (Figure 2K). Injection of MO blocking the initiation codon

of GFP ORF (Mishima & Tomari, 2016; Mishima *et al*, 2022) into wild-type embryos stabilized the tandem ZF6 reporter mRNA to a level comparable to that in MZ*znf598* embryos (Figure 2K), demonstrating that the decay was translation dependent, as it was with endogenous *znf236* mRNA (Figure 1H). These results showed that the ribosome stall activity of *znf236* ORF fragments encoding C2H2-ZF coincided with the mRNA degradation activity of Znf598. We concluded that the specific C2H2-ZF sequences of *znf236* mRNA induced ribosome stalling and NGD in zebrafish embryos.

Znf598 dampens mRNAs encoding zinc finger domains at ZGA

Given that tandem C2H2-ZFs of *znf236* induced NGD, we wondered if C2H2-ZF genes are generally susceptible to NGD in zebrafish. Many C2H2-ZF genes contain multiple tandem C2H2-ZFs, with a median of eight C2H2-ZFs per protein in zebrafish. Indeed, our NGD target maternal mRNAs contained four C2H2-ZF genes, including *znf236* and *znf970,* all encoding more than ten tandem C2H2-ZFs (Supplemental Table 1). Tandem C2H2-ZF genes often increase their number by local gene duplication (Emerson & Thomas, 2009), and approximately one-third of C2H2-ZF genes are present on chromosome 4 in zebrafish (401/1115 C2H2-ZF genes) (Howe *et al*, 2013, 2016). While analyzing the expression pattern of C2H2-ZF genes, we noticed that many C2H2-ZF genes showed increased mRNA expression levels at 6 hpf (after ZGA) compared to 1 hpf (before ZGA) (Figure 3A). This trend was most evident in C2H2-ZF genes located on chromosome 4, as reported in a previous study (Supplemental Figure 2A) (White *et al*, 2017). Hence, a substantial fraction of C2H2-ZF mRNAs was transcribed at ZGA, and such zygotic mRNAs were not included

in our initial NGD target analysis focusing on maternal mRNAs.

Considering the zygotic expression of C2H2-ZF genes, we redefined our NGD target criteria to include both maternal and zygotic mRNAs. To this end, we analyzed mRNA levels in wild-type and MZ*znf598* mutant embryos at 6 hpf regardless of their expression levels at 1 hpf. We then identified mRNAs with upregulated or downregulated expression in MZ*znf598* mutant embryos at 6 hpf (q-value <0.01) and calculated the frequency of C2H2-ZF genes in each group (Figure 3B). Of all mRNAs expressed at 6 hpf, 5.8% encoded C2H2-ZF (882 C2H2-ZF genes out of 15245 protein-coding genes). In contrast, the ratio of C2H2-ZF genes increased to 23.1% in mRNAs upregulated in MZ*znf598* embryos (57 C2H2-ZF genes out of 247 protein-coding genes, Figure 3B and Supplemental Table 2). The C2H2-ZF genes were not enriched in downregulated mRNAs (18 out of 258 protein-coding genes). The enrichment in upregulated mRNAs was specific to C2H2-ZF because genes with another zinc finger domain, RING-type ZF (RING-ZF), were present at a comparable ratio in the three mRNA sets. Many upregulated C2H2-ZF mRNAs were derived from chromosome 4 (67%, 38 out of 57 upregulated C2H2-ZF genes) (Figure 3C), whereas C2H2-ZF mRNAs encoded outside chromosome 4, such as those in chromosomes 15, 16, and 22, were also upregulated in MZ*znf598* embryos (Figure 3C and Supplemental Figure 2B). As a result, C2H2-ZF mRNAs were globally upregulated in MZ*znf598* embryos (Figure 3D). qRT–PCR confirmed the downregulation of three zygotic C2H2-ZF mRNAs by Znf598 in a RING domain-dependent manner, as observed with maternal *znf236* and *znf970* mRNAs (Figure 3E). These data indicate that Znf598 reduces mRNAs encoding C2H2-ZF at the onset of ZGA in zebrafish embryos.

Ribosome stalling at C2H2-ZF sequences induces NGD

Next, we asked if the repression of C2H2-ZF mRNAs by Znf598 was mediated by ribosome collision at sequences encoding C2H2-ZF. To this end, we revisited our disome profiling data obtained at 4 hpf (Han *et al*, 2020). Since a subset of C2H2-ZF sequences is highly homologous at the nucleotide level due to local amplification (Howe *et al*, 2013; Emerson & Thomas, 2009), we reanalyzed the disome profiling data with a stringent mapping strategy so that only uniquely mapped reads containing up to a single mismatch were considered. As a result, we detected 35% fewer disome peaks in all protein-coding mRNAs in the new analysis (3311 to 2170 peaks) (Supplemental Figure 3A). This trend was more remarkable when we focused on C2H2-ZF genes (89% fewer disome peaks, 495 to 55). For example, disome peaks on *znf1128* mRNA were not detected in the new analysis, whereas disome footprints on *znf236* mRNA were maintained (Supplemental Figure 3B). Considering this technical limitation, we conservatively decided to use a stringent mapping strategy to exclude ambiguously mapped footprint reads from the analysis. We then sequenced two additional disome footprint libraries and merged the results with the previous two libraries to increase the footprint coverage (see methods). When we called disome peaks using the criteria in a previous study (Han *et al*, 2020), disome peaks were detected on 37 C2H2-ZF mRNAs expressed at 6 hpf. Although the number of peaks was likely to be underestimated due to the stringent mapping strategy, these 37 genes tended to be more upregulated in MZ*znf598* embryos than in the remaining C2H2-ZF genes (Supplemental Figure 4A).

The disome peak analysis aimed to detect strong and consistent ribosome collision

at a single codon (Han *et al*, 2020). Therefore, this analysis could not evaluate relatively broad disome clusters; indeed, the disome cluster of *znf236* mRNA was not detected as a peak when applying our disome peak criteria. To evaluate the frequency of ribosome stall events at C2H2-ZF sequences, we calculated a relative accumulation of disome footprints on C2H2-ZF sequences of a given ORF normalized by the total disome footprints formed on the entire ORF (Figure 3F, hereafter called disome accumulation score). This evaluation gives higher scores if disome footprints are clustered on C2H2-ZF sequences, whereas disome formation independent of C2H2-ZF sequences would be reflected as lower scores. Based on this score, we divided C2H2-ZF mRNAs generating ten or more disome footprints into three classes (high, middle, and low) (Supplemental Figure 4B). There was no bias of the disome accumulation score by the footprint read amount (Supplemental Figure 4C). We found that C2H2-ZF mRNAs in the high class were more upregulated in MZ*znf598* than those in the middle and low classes (Figure 3G). This result supports the model wherein disomes formed at C2H2-ZF sequences induced NGD in zebrafish embryos.

Ribosomes stall at positively charged C2H2-ZF sequences

Using the disome footprints derived from C2H2-ZF, we next tried to understand the sequence features of C2H2-ZF that caused ribosome collision. Approximately 85% of zebrafish C2H2-ZFs are 28 amino acids in length (8168/9622 C2H2-ZFs), with cysteine and histidine, which accommodate a zinc ion, located at invariant positions (Figure 3H) (Emerson & Thomas, 2009; Stubbs *et al*, 2011). First, we analyzed the A-site position of the leading ribosome in disome footprints to ask if ribosomes collided at any specific positions in the C2H2-ZF

sequence. This analysis revealed that disomes formed at various positions in the C2H2-ZF sequence (Figure 3I). For example, 94.1% of disome footprints in *znf1102* mRNA overlapped with C2H2-ZF sequences (disome accumulation score 0.941) and formed a major cluster between the fourth and seventh zinc finger domains (ZF4 to ZF7, Figure 3J). The A-site codon of the leading stalled ribosome in these disome footprints did not converge at a specific position in the C2H2-ZF sequence. Consistent with these results, disome occupancies at the codons on which A-, P-, and E-sites in the leading ribosome lay were not skewed drastically in C2H2-ZF sequences compared to entire ORF sequences; the difference of the occupancies was within a twofold (Supplemental Figure 4D). These analyses revealed that ribosomes stalled and collided in the C2H2-ZF sequences without strong positional or codon bias.

Electrostatic interactions between the negatively charged ribosome exit tunnel and a positively charged nascent polypeptide is known to cause ribosome stalling (Lu & Deutsch, 2008; Liutkute *et al*, 2020). Previous disome profiling studies consistently detected the presence of positively charged nascent polypeptides in the exit tunnel of the leading stalled ribosome (Han *et al*, 2020; Arpat *et al*, 2020; Zhao *et al*, 2021). Notably, a series of positively charged amino acids are conserved in C2H2-ZF, mediating interaction with negatively charged nucleic acids (Thomas & Schneider, 2011). Consistently, C2H2-ZF genes generally presented high isoelectric points (pI) than genes without C2H2-ZF (Supplemental Figure 4E). This trend become even stronger when we calculated pI with 8168 isolated C2H2-ZF sequences. Therefore, we hypothesized that ribosomes stalled when positively charged nascent polypeptides emerged while translating tandem C2H2-ZFs. To analyze the relationship between the ribosome stall site and nascent polypeptide more directly, we

calculated the net charge of the nascent polypeptide upstream of the A-site positions of the leading ribosome in disome footprints derived from C2H2-ZF sequences. We found that nascent polypeptides translated by the leading ribosomes that stalled at C2H2-ZF showed a higher mean net charge between 14 and 29 amino acids upstream of the A site than the randomly simulated nascent polypeptides derived from C2H2-ZF (Figure 3K). Since most C2H2-ZFs (~28 amino acids) are encoded as tandem repeats with fixed spacer sequences (Emerson & Thomas, 2009), the positively charged region could either be derived from the C2H2-ZF at which ribosome stalled or from an upstream C2H2-ZF, depending on the position of the leading stalled ribosome. These results suggest that ribosomes stalled at C2H2-ZF sequences when strongly positively charged nascent peptides emerged, and such stalling might induce ribosome collision and NGD in zebrafish.

ZNF598 degrades mRNAs encoding C2H2 zinc finger proteins in human cells

Having established the C2H2-ZF sequences as NGD targets in zebrafish, we wondered if this is a conserved phenomenon in vertebrates. To address this question, we analyzed mRNA half-life using BRIC-Seq in HEK293T cells expressing control shRNA or ZNF598-targeting shRNAs (Figure 4A) (Tani *et al*, 2012; Han *et al*, 2020). Among the 3070 transcripts with reliable half-life measurements, 169 transcripts encoded C2H2-ZF (5.5%) and 65 encoded RING-ZF (2.1%), which were comparable ratios to the genome background. We found that the half-life of transcripts encoding C2H2-ZF was significantly elongated in ZNF598 knockdown cells, whereas that for RING-ZF was not (Figure 4B). Similar to zebrafish chromosome 4, many C2H2-ZF mRNAs with increased half-lives in ZNF598 knockdown

cells were derived from chromosome 19, where the largest human C2H2-ZF gene cluster is located (Supplemental Figure 5A) (Emerson & Thomas, 2009). These data indicated that mRNAs encoding C2H2-ZF were destabilized by ZNF598 in human cells.

We then compared the BRIC-Seq results with the published disome profiling data from HEK293 cells (Han *et al*, 2020). First, we analyzed whether C2H2-ZF mRNAs with higher disome accumulation scores were preferentially stabilized in ZNF598 knockdown cells (Supplemental Figure 5B). Due to the limited number of C2H2-ZF mRNAs analyzed in BRIC-Seq, we divided C2H2-ZF mRNAs into two classes based on their disome accumulation scores (top and bottom). Although the small sample size of this analysis hampered the statistically significant conclusions, top-class mRNAs tended to have more increased half-lives following ZNF598 knockdown than bottom-class mRNAs. Second, we analyzed the disome position in human C2H2-ZF sequences (Supplemental Figure 5C and D). Similar to zebrafish, disomes that formed at C2H2-ZF sequences had no clear positional preference. Third, we analyzed the mean net charge of nascent polypeptides translated by the leading ribosomes stalled at human C2H2-ZF sequences (Supplemental Figure 5E). Nascent polypeptides translated by the ribosomes stalled at C2H2-ZF showed a high net charge; however, the high net charge region was more proximal to the decoding/peptidyl transferase centers in humans when compared to that in zebrafish. These observations imply the commonality and difference of the mechanisms that cause ribosome stalling at C2H2-ZF sequences and NGD in vertebrates.

Since C2H2-ZF mRNAs were degraded in a ZNF598-dependent manner in zebrafish and humans, we investigated the conservation of NGD target genes between the

two species. To address this point, we divided zebrafish C2H2-ZF genes into two classes based on Ensemble ortholog annotation as follows: orthologous C2H2-ZF genes that have annotated ortholog in the human genome (240 genes) and nonorthologous C2H2-ZF genes that lack annotated human orthologs (642 genes). We found that nonorthologous zebrafish C2H2-ZF genes were preferentially upregulated in MZ*znf598* embryos (Figure 4C). This result suggested that species-specific C2H2-ZF genes are a major source of NGD target sequences in zebrafish.

In HEK293T cells, however, mRNAs encoding orthologous (80 genes) and nonorthologous (89 genes) C2H2-ZF proteins were similarly destabilized in a ZNF598-dependent manner (Figure 4D). To further address whether the species-specific evolution of C2H2-ZF genes explains the susceptibility to NGD in humans, we focused on the role of mammalian C2H2-ZF genes in suppressing transposons and their remnant sequences via the KRAB domain (Jacobs *et al*, 2014; Yang *et al*, 2022; Wolf *et al*, 2020; Imbeault *et al*, 2017; Bruno *et al*, 2019). The KRAB domain-fused tandem C2H2-ZF proteins emerged during tetrapod evolution, and this domain is absent in zebrafish C2H2-ZF proteins (Emerson & Thomas, 2009; Imbeault *et al*, 2017). Indeed, approximately half of the human C2H2-ZF proteins analyzed by BRIC-Seq contained the KRAB domain (78 genes), most of which did not have orthologous genes in zebrafish (63 genes). Another half of the human C2H2-ZF proteins lacked the KRAB domain (91 genes), many of which had orthologous genes in zebrafish (65 genes). We found that the presence of the KRAB domain did not correlate with susceptibility to NGD in HEK293T cells (Supplemental Figure 5F). Overall, these results indicate that the C2H2-ZF sequence is a major target of NGD in vertebrates. However,

complex evolutionary events may underlie the generation of the C2H2-ZF sequences that induce ribosome collision and trigger NGD in each species.

Discussion

In this study, we systematically investigated NGD target mRNAs in zebrafish and humans. Our analysis identified mRNAs encoding C2H2-ZF as major targets of Znf598-dependent NGD, revealing a previously uncharacterized type of ribosome collision and its rescue by NGD in vertebrates.

Toward a comprehensive understanding of ribosome stalling and its quality control landscape, our study provided a unique example of ribosome collision events associated with a specific protein domain. Recent disome profiling in yeast and mammalian cells characterized fundamental features of ribosome collision events in eukaryotes (Han *et al*, 2020; Zhao *et al*, 2021; Meydan & Guydosh, 2020; Arpat *et al*, 2020). One crucial feature of ribosome collision is the high propensity of disome formation between or outside the protein domains (Arpat *et al*, 2020; Han *et al*, 2020), likely representing programmed ribosome slowdown assisting cotranslational protein folding (Kramer *et al*, 2019). Local codon and amino acid features associated with ribosome collision sites were also characterized extensively (Han *et al*, 2020; Zhao *et al*, 2021; Meydan & Guydosh, 2020; Arpat *et al*, 2020). In contrast, the characterization of ribosome collision events associated with wider protein domains or motifs is limited to the signal peptide (Arpat *et al*, 2020). Therefore, it is intriguing that our results identified C2H2-ZF, the most abundant DNA-binding domain in the eukaryote genome (Emerson & Thomas, 2009; Schmitz *et al*, 2016), as a collision-inducing sequence. The broad distribution of disomes in the C2H2-ZF sequence might hinder its identification as a typical collision site in previous disome analyses, which mainly focused on strong and consistent codon-wise collision sites (Han *et al*, 2020; Meydan & Guydosh,

2020; Zhao *et al*, 2021; Arpat *et al*, 2020). In this regard, a combined analysis of disome formation with its functional consequence, NGD, had clear merit for complementing ribosome collision events that were not addressed in previous studies.

How does the ribosome stall and collide while translating C2H2-ZF? Our data suggest that a clue to the stalling mechanism lies in the nascent polypeptide sequence. C2H2-ZF sequences generally have high pIs (Supplemental Figure 4E), and the nascent polypeptides associated with the ribosomes stalling at C2H2-ZF harbored a further positively charged region that is held within the ribosome exit tunnel (Figure 3J) (Liutkute *et al*, 2020). Therefore, translation of the tandem C2H2-ZF sequence sequentially exposes the positively charged nascent polypeptides to the ribosome exit tunnel and may synergistically hamper ribosome traverse. Considering the length of the nascent polypeptide that occupies the ribosome exit tunnel (~40 amino acids) and the requirement of three C2H2-ZF sequences (84 amino acids) to cause ribosome stalling (Figure 2H), it is also possible that the upper C2H2-ZF sequences contribute to stalling while folding co-translationally near or outside the exit of the tunnel. Indeed, the upper part of the bacterial ribosome exit tunnel can accommodate a folded C2H2-ZF of yeast ADR1a (Nilsson *et al*, 2015). Interaction between the domain-encoding nascent polypeptide and chaperons is another factor that could modulate ribosome stalling (Stein *et al*, 2019). Since not all tandem C2H2-ZFs caused ribosome collision and NGD, a combinatory effect of multiple interactions mediated by tandem C2H2-ZF may be required to cause prolonged ribosome stalling to levels sufficient for inducing ribosome collision and NGD. Sophisticated biochemical and structural analyses are required to understand the detailed stall mechanism that occurs at the tandem C2H2-ZF sequence.

The stall-prone nature of widespread C2H2-ZF suggests an evolutionary role of the cotranslational quality control system. Since ribosome stalling in C2H2-ZF sequences was detected by the collision sensor Znf598 and induced NGD, such stalling is qualitatively different from folding-assisting stalling and likely detrimental to the cell. Why does the widely shared protein domain harbor such a risk? The C2H2-ZF domain is the largest protein domain family in the eukaryote genome, most of which is present as tandem arrays in a protein (Emerson & Thomas, 2009; Stubbs *et al*, 2011). The evolution of tandem C2H2-ZF genes is highly dynamic, and species-specific amplification of the C2H2-ZF gene is frequently observed, as exemplified by those that occurred in zebrafish chromosome 4 and human chromosome 19 (Howe *et al*, 2013; Emerson & Thomas, 2009; Kauzlaric *et al*, 2017). Accordingly, the repertoire of C2H2-ZF has expanded rapidly and has a huge diversity in each species (Emerson & Thomas, 2009; Imbeault *et al*, 2017). On the one hand, the rapid evolution and diversification of the C2H2-ZF sequences contribute to expanding the DNA-binding capacity of the proteome, which regulates endogenous genes and transposable elements (Emerson & Thomas, 2009; Imbeault *et al*, 2017; Schmitges *et al*, 2016). On the other hand, it could drive the emergence of stall-prone C2H2-ZF sequences at a rate exceeding evolutionary pressure. In this regard, one possible role of NGD is maintaining the quality of the translatome during evolution by dampening the expression of transiently and frequently acquired stall-prone C2H2-ZF sequences.

At the individual organism level, our study revealed the role of NGD in controlling zygotically expressed C2H2-ZF mRNAs during development (White *et al*, 2017; Hadzhiev *et al*, 2019). Approximately one-third of zebrafish C2H2-ZF genes are encoded in the long

arm of chromosome 4, where retrotransposons, the microRNA miR-430, and other protein-coding genes, including C2H2-ZF genes, form clusters (Howe *et al*, 2016, 2013). Interestingly, miR-430 and C2H2-ZF genes share promoter features that allow transcription at ZGA (Hadzhiev *et al*, 2023). Therefore, the massive transcription of C2H2-ZF genes in chromosome 4, many of which encode stall-prone C2H2-ZF sequences, has a risk of overloading ribosomes and threatening translational homeostasis at the onset of ZGA. Such overloading could also occur in pathogenic developmental conditions at later stages. The loss of histone demethylase Kdm2aa causes melanoma formation and activates transcriptionally silenced chromosome 4 during juvenile stages, leading to the upregulation of C2H2-ZF mRNAs (Scahill *et al*, 2017). A plausible scenario is that NGD dampens stall-prone C2H2-ZF mRNAs in these situations, thereby helping maintain translational homeostasis during development (Figure 4E).

We showed that targeting the C2H2-ZF sequence by Znf598 is an evolutionally conserved phenomenon in vertebrates. However, the repertoire of NGD target C2H2-ZF genes differs between zebrafish and humans. In zebrafish, species-specific C2H2-ZF genes are downregulated by Znf598 (Figure 4C). This bias can be explained by zebrafish-specific amplification and diversification of C2H2-ZF sequences, such as that occurred in chromosome 4, generating stall-prone young C2H2-ZF domains that have not been optimized for translation. The number of C2H2-ZF genes correlates with retrotransposon amplification events; this mechanism represents an arms race of transcription-repressor C2H2-ZF genes and retrotransposons (Bruno *et al*, 2019; Thomas & Schneider, 2011). Therefore, it seems reasonable that stall-prone C2H2-ZFs emerged in rapidly evolving C2H2-ZF genes. Indeed,

we observed that mRNAs derived from human chromosome 19, where the largest C2H2-ZF cluster is located, tended to be stabilized in ZNF598 knockdown cells (Supplemental figure 5A). However, human ZNF598 destabilizes C2H2-ZF genes regardless of the presence of orthologous genes in zebrafish or the presence of the KRAB domain that mediates the transcriptional repression of retrotransposons (Figure 4D and Supplemental Figure 5F). Therefore, it was unclear what drove the emergence of stall-prone C2H2-ZF sequences in humans. A complex evolutionary history of C2H2-ZF genes in tetrapods makes it difficult to explain the susceptibility of C2H2-ZF sequences to NGD in humans. Further characterization of stall-prone C2H2-ZF in humans and other species would explain how stall-prone sequences emerge and disappear during the evolution of these widespread protein domains in the eukaryotic genome.

Finally, our study implies a broad impact of NGD on the expression of endogenous and artificial repetitive sequences. C2H2-ZF has been engineered as a platform for programmable DNA-binding modules, such as zinc finger nucleases (Klug, 2010). Elucidation of the ribosome stall mechanism by the C2H2-ZF sequence would enable the prediction of stall events and help design efficient DNA-binding modules by avoiding stall-prone C2H2-ZFs. We also speculate that other tandem repeat domains, especially those enriched with positively charged amino acids, harbor a risk of acquiring ribosome stall sequences. Further studies will elucidate the role of NGD and RQC in the regulation of mRNAs encoding tandem repeat domains.

Materials and Methods

Zebrafish

The zebrafish AB strain was used as a wild type. The zebrafish *znf598* mutant strain harboring an 11 bp deletion in exon 1 was described previously (Mishima *et al*, 2022). Fish were raised and maintained at 28.5°C under standard laboratory conditions according to the Animal Experiment Protocol (2023-37) at Kyoto Sangyo University. Fertilized eggs were obtained by natural breeding. Embryos were developed in system water at 28.5°C.

Bright-field images were acquired using a SteREO Lumar V12 microscope and an AxioCam MRc camera with ZEN software (Zeiss, Jena, Germany) using spline mode.

Microinjection

Myc-tagged Znf598 mRNAs and tandem luciferase mRNAs were transcribed from a linearized plasmid DNA template using the mMESSAGE mMACHINE SP6 Transcription Kit (Thermo Fisher Scientific) and purified using the RNeasy Mini Kit (QIAGEN). sfGFP mRNAs were transcribed from a linearized plasmid DNA template using the SP6-Scribe Standard RNA IVT Kit (CELLSCRIPT); the m^7G cap structure was added to the purified RNA using the ScriptCap m^7G Capping System (CELLSCRIPT), the poly(A) tail was added using the A-Plus Poly(A) Polymerase Tailing Kit (CELLSCRIPT), and the mRNA was purified using the RNeasy Mini Kit. Purified mRNAs were diluted in water to the following concentrations: tandem luciferase mRNAs and sfGFP reporter mRNAs: 50 ng/µl and Myc-Znf598 mRNAs: 200 ng/µl. Translation-blocking MOs for *znf236* (5′-

GCGGCATAGTTCCAGTTTATAGAGC-3′) and GFP (Mishima & Tomari, 2016), and a standard control MO were purchased from Gene Tools and diluted in distilled water to 100 μM. Microinjection was performed as described previously (Mishima *et al*, 2022; Mishima & Tomari, 2016) using an IM300 Microinjector (NARISHIGE). Approximately 1,000 pl of the solution was injected per embryo within 15 min of fertilization.

Plasmids

Plasmids for preparing dual luciferase mRNAs were constructed based on the pLuc plasmid (Mishima *et al*, 2022). DNA fragments encoding test sequences were prepared by PCR, annealing complementary oligonucleotides, or DNA fragment synthesis (IDT) and cloned between the EcoRI and XhoI sites of pLuc. The PCR primers are listed in Supplemental Table 3. To design DNA fragments encoding artificial zinc finger repeats (ZF6×2, 6×3, 7×3, 8×3), synonymous codon substitutions were introduced to avoid repeat sequences that prohibited DNA fragment synthesis. To construct plasmids for preparing sfGFP mRNAs, DNA fragments encoding test sequences were cloned between the EcoRI and EcoRV sites of the pCS2+neo-sfGFP-*suv39h1a* 3′ UTR (Mishima *et al*, 2022). Plasmids encoding pCS2+MT-Znf598 were described previously (Mishima *et al*, 2022).

RNA-Seq

RNA-Seq was performed as described previously (Mishima *et al*, 2022). Briefly, total RNA was extracted from 40–50 wild-type and MZ*znf598* zebrafish embryos at 1 and 6 hpf using TRI Reagent (Molecular Research Center). After DNaseI treatment and purification by the

RNeasy Mini Kit, RNA integrity was confirmed using an Agilent RNA 6000 Nano Chip (Agilent Technologies, USA). Total RNA samples were treated with a Ribo-Zero Gold rRNA Removal Kit (Human/Mouse/Rat) (Illumina) and used for RNA-Seq library preparation using an Illumina TruSeq Stranded Total RNA Library Prep Kit (Illumina, USA) with multiplexing. The pooled libraries were sequenced on NextSeq500 (Illumina) sequencing platforms with 76 bp single-end sequencing.

Ribosome footprint profiling

Part of the zebrafish monosome and disome footprint sequencing data at 4 hpf used in this study were published previously (Han *et al*, 2020) (GSE133392). We obtained two additional disome footprint libraries at 4 hpf, corresponding to the mosonome footprint libraries (Mishima *et al*, 2022) (GSE173604), to increase the depth of disome footprint reads. These additional libraries were prepared in parallel to the published libraries following the protocol described in (Mito *et al*, 2020).

Analysis of RNA-Seq data

Zebrafish RNA-Seq data were analyzed using Galaxy on public servers (Galaxy Community, 2022). Single-ended reads were preprocessed by fastp (Chen *et al*, 2018) (version 0.23.2+galaxy0) and mapped to the zebrafish genome reference GRCz11(danRer11) using STAR (Dobin *et al*, 2013) (version 2.7.8a+galaxy1) by allowing up to a single-nucleotide mismatch with --outFilterMultimapNmax 1 –outFilterMismatchNmax 1 options. Only uniquely mapped reads were subjected to downstream analysis. Mapped reads were counted

by featureCounts (Liao *et al*, 2014) (version 2.0.1+galaxy2), and differential expression was analyzed by edgeR (Robinson *et al*, 2010) (version 3.36.0+galaxy0). Only genes with Ensemble Canonical transcript annotation were analyzed as protein-coding genes in this study. Genes with at least 1 cpm were considered expressed at detectable levels and used for further analysis. To call differentially expressed genes, the FDR was set to 0.01.

Protein domain analysis

Protein domains in zebrafish and human amino acid sequences were detected using pfsearch (Bucher *et al*, 1996) (version 2.3.5) based on matrix profiles obtained from the PROSITE database (https://prosite.expasy.org) (Sigrist *et al*, 2013). The entry IDs of the matrix profiles used are as follows: PS50157 for the C2H2-ZF domain; PS50089 for the RING-ZF domain; and PS50805 for the KRAB domain.

Analysis of ribosome profiling data

Ribosome profiling data were processed as previously described (Iwasaki *et al*, 2016; Ingolia *et al*, 2012; Han *et al*, 2020) with minor changes. Briefly, reads were preprocessed using fastp (Chen *et al*, 2018; Dobin *et al*, 2013) and mapped to a zebrafish (GRCz11) or human (GRCh38/hg38) genome reference using STAR (Dobin *et al*, 2013) by allowing up to a single-nucleotide mismatch and allowing mapping to a single locus with --outFilterMultimapNmax 1 –outFilterMismatchNmax 1 options. Disome occupancy was calculated as the ratio of reads at given codons to the average reads per codon on all ORFs or on C2H2-ZF domains. The disome accumulation score was calculated as the ratio of reads

on C2H2-ZF domains in an ORF to the total number of reads on the entire ORF. Transcripts with at least 10 reads on the ORF were subjected to downstream analysis. Since most C2H2-ZF domains are 28 amino acids in length, disome occupancy and disome accumulation scores were calculated for C2H2-ZF domains 28 amino acids in length.

The net charge of a nascent polypeptide associated with the stalled leading ribosome was calculated with 6-amino-acid sliding windows using the Peptides package in R. A-site positions of the leading ribosomes in 28-amino-acid C2H2-ZF domains were used for the net charge calculation (experiment). For the control, the same number of A-site positions in 28-amino-acid C2H2-ZF domains was randomly chosen, and their net charge was calculated (simulation).

BRIC-Seq

Library preparation

BRIC-Seq was performed as described previously with modifications (Tani *et al*, 2012). The HEK293T cells expressing control and ZNF598 shRNA (Han *et al*, 2020) were seeded on 10 cm dishes and incubated in 10 ml DMEM with 150 μM bromouridine (BrU, FUJIFILM Wako Pure Chemical Corporation) for 24 h. Cells were then washed twice with 5 ml DMEM without BrU, incubated in 10 ml DMEM without BrU for 0, 2, 4, 6, 10, or 24 h, washed with 5 ml PBS, and lysed by adding 600 μl TRIzol reagent (Thermo Fisher Scientific) directly on the dishes. Total RNAs was purified using a Direct-zol RNA MiniPrep Plus Kit (Zymo Research).

BrU-containing spike-in mRNAs were synthesized by *in vitro* transcription. Template DNA fragments were PCR-amplified using psiCHECK2 plasmid (Promega) with 5′-TAATACGACTCACTATAGG-3′ and 5′-CACACAAAAAACCAACACACAG-3′ (Rluc); psiCHECK2 plasmid with 5′-ACTTAATACGACTCACTATAGGAAGCTTGGCATTCCGG-3′ and 5′-TGTATCTTATCATGTCTGCTCGAAG-3′ (Fluc); and pColdI-GFP plasmid (Tsuboyama *et al*, 2020) with 5′-TGACTAATACGACTCACTATAGGATCTGTAAAGCACGCCATATCG-3′ and 5′-TGGCAGGGATCTTAGATTCTG-3′ (GFP). *In vitro* transcription was performed using a T7-Scribe Standard RNA IVT Kit (CELLSCRIPT) in the presence of 1.2 mM BrUTP (Jena Bioscience) and 5 mM UTP. RNA was purified by RNAClean XP beads (Beckman Coulter).

For immunoprecipitation, 25 µl of Dynabeads M-280 Sheep anti-Mouse IgG (Thermo Fisher Scientific) was incubated with 2 mg of anti-BrdU antibody (BD Bioscience, 555627) for 3 h at 4°C and equilibrated with BRIC-Seq lysis buffer (lysis buffer without cycloheximide and chloramphenicol). Twenty micrograms of total RNA mixed with spike-in RNA mixture (1 ng Rluc, 0.2 ng Fluc, and 0.04 ng GFP) was incubated with the magnetic beads in BRIC-Seq lysis buffer for 2 h at 4°C, washed with 100 µl BRIC-Seq lysis buffer five times, and mixed with 200 µl TRIzol reagent. RNA was extracted by a Direct-zol RNA MicroPrep Kit and subjected to an RNA-Seq library preparation using a Ribo-Zero Gold rRNA Removal Kit (Human/Mouse/Rat) (Illumina) followed by a TruSeq Stranded Total RNA Kit (Illumina). Libraries were sequenced on a HiSeq 4000 (Illumina) with the single-end/50 nt-long read option.

Data analysis

Filtering by quality and adapter trimming were performed using fastp (Chen et al, 2018). Reads were then aligned to the noncoding RNAs using STAR 2.7.0a (Dobin *et al*, 2013) and excluded from analysis. The remaining reads were mapped to the human genome hg38 with the annotation of the GENCODE Human release 32 reference using STAR 2.7.0a with --outFilterMismatchNmax 1 --outFilterMultimapNmax 1 options. Read counts on each transcript were analyzed by featureCounts (Liao *et al*, 2014).

The fold changes in read counts compared to 0 h samples were calculated with the DESeq2 package (Love *et al*, 2014) and then renormalized to the mean change in three spike-ins. For the calculation of mRNA half-life, the $\log_2$-converted fold change in each time point was fitted to a linear model. Since the decay of transcripts sometimes slows at later time points, we calculated the half-life using all combinations of earlier samples (*i.e.*, 0-4 h, 0-6 h, 0-10 h, and 0-24 h) and adopted a fitting with the maximum R^2 value. Transcripts with $R^2 > 0.9$ and $0 <$ half-life < 24 h were used in downstream analyses.

<u>Ortholog analysis</u>

Ortholog analysis was performed based on Ensemble ortholog annotation (http://www.ensembl.org). For a gene in each species, the presence of an annotated orthologous gene in another species with any homology type (one to one, one to many, and many to many) was considered orthologous, whereas the absence of an annotated orthologous gene was considered nonorthologous.

qRT–PCR

qRT–PCR analysis of mRNAs was performed as previously described (Mishima *et al*, 2022). Briefly, total RNA was prepared using TRI Reagent (Molecular Research Center), and cDNA was synthesized using the PrimeScript RT reagent kit with gDNA eraser (TaKaRa). A random hexamer was used for cDNA synthesis. qRT–PCR was performed using SYBR Premix Ex TaqII (Tli RNaseH Plus) and the Thermal Cycler Dice Real Time System (TaKaRa). The data were normalized using 18S rRNA as a reference. The primer sequences are listed in Supplemental Table 3.

PAT assay

The PAT assay was performed as previously described (Mishima & Tomari, 2016). Briefly, 150 ng of total RNA was incubated with 75 U of yeast poly(A) polymerase (PAP) (Affymetrix) in the presence of GTP/ITP mix at 37°C for 60 min. cDNA was synthesized at 44°C for 15 min using the PrimeScript RT reagent kit with gDNA eraser (TaKaRa) and a y300 PAT universal C10 primer. PAT-PCR was performed using a 3′ UTR-specific forward primer and a PAT universal primer with GoTaq Green Master Mix (Promega). The PCR products were separated by 6% PAGE in 0.5× TBE. Gels were stained with GelRed (Biotium), and the signals were detected using an Amersham Imager 680 (Cytiva). The primer sequences are listed in Supplemental Table 3.

Luciferase assay

The luciferase assay was performed as previously described (Mishima *et al*, 2022). Briefly, 5–10 embryos were collected at 3 hpf and lysed in Passive Lysis Buffer (Promega) containing cOmplete Protease Inhibitor Cocktail (Sigma–Aldrich). Ten microliters of lysate containing one embryo was used for the assay. The luciferase activities were measured using the Dual-Glo Luciferase Assay System and a GloMax 20/20 luminometer (Promega). The intensity of Fluc activity was normalized to the intensity of Rluc activity. The normalized Fluc activity was further normalized to that of the control sample (no insert).

Western blotting

Proteins were detected as described previously (Mishima *et al*, 2022) using anti-Myc (MBL Life science, My3 mouse monoclonal, 1:4,000) and anti-α-tubulin-pAb HRP-DirecT (MBL Life science, PM054-7 rabbit polyclonal, 1: 10,000) antibodies. The signals were developed using Luminata Forte (Millipore) and detected using an Amersham Imager 680 (Cytiva).

Data availability

The following accession numbers for the sequence data reported in this paper are from the GEO database: GSE236143 (RNA-Seq in zebrafish), GSE236144 (additional disome profiling in zebrafish), and GSE233374 (BRIC-Seq).

Acknowledgments

This work was supported by the Japan Society for the Promotion of Science (JSPS) (JP23H02413 and JP22K19300 to YM and JP22K15042 to KI) and the Japan Agency for Medical Research and Development (AMED) (AMED PRIME, JP23gm6310017h0004) to YM. SI was supported by JSPS (JP23H02415), MEXT (JP20H05784), AMED (AMED-CREST, JP23gm1410001), and RIKEN (Pioneering Projects "Biology of Intracellular Environments"). YS was supported by JSPS (JP21K15023 and JP23K05648), MEXT (JP21H05734, 23H04268), and RIKEN (Pioneering Projects "Biology of Intracellular Environments"). SK was supported by JSPS (JP21H02513). PH was supported by the JSPS Research Fellowship for Young Scientists.

We thank our laboratory members for their discussions and critical comments on the project, Kimi Wakabayashi for fish maintenance, Kaori Kaminoyama and Tomoaki Sakamoto for RNA-Seq, and Mari Mito for technical assistance. pColdI-GFP was a kind gift from Kotaro Tsuboyama and Yukihide Tomari. DNA libraries for ribosome profiling were sequenced by the Vincent J. Coates Genomics Sequencing Laboratory in QB3 Genomics at UC Berkeley (RRID:SCR_022170) at UC Berkeley (RRID:SCR_022170), supported by the National Institutes for Health (NIH) Instrumentation Grant (S10 OD018174). Computational analysis was supported by the supercomputer HOKUSAI Sailing Ship in RIKEN.

The Galaxy server that was used for some calculations is in part funded by Collaborative Research Centre 992 Medical Epigenetics (DFG grant SFB 992/1 2012) and German Federal Ministry of Education and Research (BMBF grants 031 A538A/A538C RBC, 031L0101B/031L0101C de.NBI-epi, 031L0106 de.STAIR (de.NBI)).

Author contributions

Y.M. and K.I. designed the project. Y.M. and Y.S. performed the experiments. K.I., Y.S., P.H., S.I., and Y.M. analyzed the NGS data. S.K. provided the RNA-Seq data. T.I. provided experimental materials. Y.M. and K.I. drafted the manuscript. All the authors discussed the results and approved the manuscript for submission.

Conflicts of interest

The authors declare no conflicts of interest associated with this manuscript.

Figure legends

Figure 1. Identification of maternal mRNAs degraded by NGD in zebrafish

(A) Bright-field pictures of wild-type and MZ*znf598* embryos at the 8-cell and shield stages. (B) A scheme of the RNA-Seq experiment to identify maternal mRNAs degraded by Znf598. (C) A scatter plot showing the fold changes in mRNA expression at 6 hpf compared to 1 hpf in wild-type embryos (x-axis) and fold changes in mRNA expression in wild-type and MZ*znf598* embryos at 6 hpf (y-axis). Unstable maternal mRNAs (blue), mRNAs upregulated in MZ*znf598* at 6 hpf (orange), and NGD target candidates (red) are indicated. (D) qRT–PCR analysis of NGD target candidate mRNAs in wild-type (blue) and MZ*znf598* embryos (red) at 6 hpf relative to that at 2 hpf. The results with MZ*znf598* embryos rescued by injecting mRNAs encoding Myc-tagged full-length Znf598 (light blue) or Znf598 lacking the RING domain (pink) are also shown. *gstm* mRNA (miR-430 target) and *ddx4* mRNA (stable maternal mRNA) are shown as controls for NGD-independent maternal mRNAs. (E) Western blotting to detect the Myc-tagged Znf598 proteins at 6 hpf. Tubulin (Tub) was detected as a loading control. (F) Time course qRT–PCR analysis of *znf236* and *znf970* mRNAs in wild-type (blue) and MZ*znf598* (red) embryos. (G) Schematics of Znf236 and Znf970 proteins. C2H2-ZF are indicated in orange. (H) qRT–PCR analysis of *znf236* mRNA in the presence of control MO (blue) or translation-blocking MO (green). *trmt11* mRNA, a codon-mediated decay target, is shown as a control. An asterisk indicates $p < 0.05$ (unpaired Student's t-test, adjusted for multiple comparisons by Holm's method) (I) PAT assay of *znf236* mRNA in wild-type and MZ*znf598* embryos at 6 hpf. The lane labeled A0 shows the 3′ UTR fragment without a poly(A) tail. (J) Quantification of the PAT assay in (I). The averages of the two

experiments are shown. The graphs represent an average of three independent experiments in D, F, and H. The error bars show the standard deviation. Open circles in D and H show each data point.

Figure 2. Analysis of ribosome stalling and NGD-inducing sequences in *znf236* mRNA

(A) Schematics of the *znf236* ORF fragments analyzed in this study. C2H2-ZF are indicated in orange. Fragments that caused significant ribosome stalling in the tandem luciferase assay are indicated in red. (B) Distributions of monosome (upper, gray) and disome (lower, blue) footprints on the *znf236* ORF. (C) A scheme of the tandem ORF assay. The Rluc ORF (orange) and Fluc ORF (light green) are separated by two P2A translation skipping sequences (light blue). A sequence to be tested (shown as X in gray) is inserted between the two P2A sequences. (D) Results of the tandem ORF assay with *znf236* fragments 1-4. (E) Results of the tandem ORF assay with the *znf236* fragment 1A-1D. (F) A detail of the 1B fragment. Amino acid sequences of ZF6, ZF7, and ZF8 are shown below. A-site positions of the stalled leading ribosome and corresponding footprint amounts are indicated as a red gradient. (G) Results of the tandem ORF assay with ZF6, ZF7, and ZF8 repeats. (H) Results of the tandem ORF assay with ZF6, ZF7, and ZF8 repeats. (I) A scheme of the mRNA injection assay to measure the effect of a sequence to be tested (shown as X in gray) on mRNA stability. (J and K) qRT–PCR analysis of injected sfGFP reporter mRNAs at 6 hpf relative to 2 hpf in wild-type (blue), MZ*znf598* (red), and GFP MO-injected wild-type (green) embryos. The graphs in D, E, G, H, J, and K represent an average of three independent experiments. The error bars show the standard deviation. Open circles show each data point. In D, E, G, and H, relative

Fluc signals normalized by Rluc signals are shown. Values with no insert (-) were set to one. Asterisks in D, E, G, and H indicate $p < 0.05$ (Dunnett's test, compared to no insert). Asterisks in J indicate $p < 0.05$ (unpaired Student's t-test, adjusted for multiple comparisons by Holm's method). Asterisks in K indicate $p < 0.05$ (Dunnett's test, compared to wild type).

Figure 3. Analysis of ribosome stalling and NGD induced by C2H2-ZF sequences in zebrafish

(A) Violin plots showing the fold changes in mRNA expression at 6 hpf compared to 1 hpf in wild-type embryos. All genes that were expressed (gray), genes with C2H2-ZF (red), and genes with RING-ZF (purple) are shown. The p values were calculated by the Wilcoxon signed rank test. (B) Bar graph showing fractions of C2H2-ZF or RING-ZF genes among all genes (gray), genes upregulated in MZ*znf598* (orange), and genes downregulated in MZ*znf598* (blue) at 6 hpf. The p value was calculated by the chi-square test. (C) A pie chart showing chromosome distributions of C2H2-ZF genes upregulated in MZ*znf598* at 6 hpf. The numbers of C2H2-ZF genes are indicated below. (D) Cumulative distributions of fold changes in mRNA levels in MZ*znf598* embryos compared to wild-type embryos at 6 hpf. All genes (black), C2H2-ZF genes (red), and RING-ZF genes (purple) are shown. The x-axis shows the fold change, and the y-axis shows the cumulative fraction. The p values are shown on the right (Kolmogorov–Smirnov test). (E) qRT–PCR analysis of NGD target C2H2-ZF mRNAs in wild-type (blue) and MZ*znf598* embryos (red) at 6 hpf. The results with MZ*znf598* embryos rescued by injecting mRNAs encoding Myc-tagged full-length Znf598 (light blue) or Znf598 lacking the RING domain (pink) are also shown. *actb1* mRNA is shown as a

control. The chromosomal locations are indicated below the gene name. The graph represents an average of three independent experiments. The error bars show the standard deviation. Open circles show each data point. (F) A scheme of disome accumulation score analysis. (G) Cumulative distributions of fold changes in mRNA levels in MZ*znf598* embryos compared to wild-type embryos at 6 hpf. All genes (black) and C2H2-ZF genes with high (dark blue), middle (blue), and low (light blue) disome accumulation scores are shown. The x-axis shows the fold change, and the y-axis shows the cumulative fraction. The p values are shown on the right (Kolmogorov–Smirnov test). (H) A logo representation of amino acid diversity among 8168 C2H2-ZFs collected from zebrafish genes. (I) A histogram showing the A-site position of the leading ribosome in disome footprints relative to the C2H2-ZF sequence in zebrafish. (J) The distributions of disome footprints on the *znf1102* ORF (blue). A scheme of the Znf1102 protein is shown, with C2H2-ZF indicated in orange. The amino acid sequences of ZF4, ZF5, ZF6, and ZF7 are shown below. A-site positions of the stalled leading ribosome and corresponding footprint amounts are indicated as a red gradient. (K) Mean net charge in the nascent chain around ribosome collision sites on C2H2-ZF sequences in zebrafish. Nascent chain sequences determined by the A-site position of experimentally detected disome footprints (red) and those determined by the randomly chosen A-site position (blue) are shown.

Figure 4. Analysis of NGD induced by C2H2-ZF sequences in human cells

(A) A scheme of BRIC-Seq in HEK293T cells. (B) Cumulative distributions of fold changes in mRNA half-life in HEK293T cells, comparing ZNF598 knockdown cells and control cells.

All genes (black), C2H2-ZF genes (red), and RING-ZF genes (purple) are shown. The x-axis shows the fold change in mRNA half-life. (C) Cumulative distributions of fold changes in mRNA levels in MZ*znf598* embryos compared to wild-type embryos at 6 hpf. Conserved genes without C2H2-ZF (black), nonconserved genes without C2H2-ZF (gray), conserved genes with C2H2-ZF (orange), and nonconserved genes with C2H2-ZF (orange) are shown. The x-axis shows the fold change. (D) Cumulative distributions of fold changes in mRNA half-life in HEK293T cells, comparing ZNF598 knockdown cells and control cells. Conserved genes without C2H2-ZF (black), nonconserved genes without C2H2-ZF (gray), conserved genes with C2H2-ZF (orange), and nonconserved genes with C2H2-ZF (orange) are shown. The x-axis shows the fold change in mRNA half-life. (E) A model of NGD for dampening mRNAs encoding stall-prone tandem C2H2-ZF sequences. In B-D, the p values are shown on the left (Kolmogorov–Smirnov test).

Supplemental Figure 1. Additional information about maternal mRNA analysis of zebrafish embryos

(A) A Venn diagram of NGD target candidates (yellow), maternal decay mRNAs (pink), and zygotic decay mRNAs (blue). The numbers of genes in each category are shown. (B and C) Cumulative distributions of maternal decay mRNAs and zygotic decay mRNAs. The x-axis shows the fold change in mRNA expression at 6 hpf compared to 1 hpf in wild-type and MZ*znf598* embryos, and the y-axis shows the cumulative fraction. The p values calculated by the Kolmogorov–Smirnov test are indicated.

Supplemental Figure 2. Additional information about the expression patterns of C2H2-ZF mRNAs in zebrafish

(A) Chromosomal distributions of C2H2-ZF genes and their relative mRNA expression at 6 hpf compared to 1 hpf in wild-type embryos. The x-axis shows chromosome numbers, and the y-axis shows the fold change in mRNA levels at 6 hpf compared to 1 hpf. (B) Chromosomal distributions of C2H2-ZF genes and their relative mRNA expression in MZ*znf598* embryos compared to wild-type embryos at 6 hpf. The x-axis shows chromosome numbers, and the y-axis shows fold changes in mRNA levels. In A and B, the value of each C2H2-ZF gene is plotted as a dot. The box represents the interquartile range (IQR), with the median indicated as a thick horizontal line in the box. The whiskers represent the variation within 1.5 IQR outside the upper and lower quartiles.

Supplemental Figure 3. Evaluation of mapping strategies for disome footprints

(A) Venn diagrams of disome peaks detected in a previous study (Han *et al.*, 2020) and in this study (1 mismatch & unique map). The number of peaks detected in each study is shown. Left: disome peaks in all genes. Right: disome peaks in C2H2-ZF genes. (B) A comparison of disome footprint distributions on the *znf1128* ORF (left) and *znf236* ORF (right) between two studies.

Supplemental Figure 4. Additional information about disome profiling in zebrafish

(A) Cumulative distributions of fold changes of mRNA levels in MZ*znf598* embryos compared to wild-type embryos at 6 hpf. All genes (black), C2H2-ZF genes with disome

peaks (green), or without disome peaks (light green) are shown. The x-axis shows the fold change, and the y-axis shows the cumulative fraction. The p values are shown on the left (Kolmogorov-Smirnov test). (B) A histogram showing distributions of disome accumulation score on the C2H2-ZF. The x-axis shows the disome accumulation score, and the y-axis shows gene counts. (C) A scatter plot showing the disome accumulation score (x-axis) and total read numbers on each ORF (y-axis). (D) Disome occupancy at the A-, P-, or E-site codon of the leading stalled ribosome in all genes (x-axis) and C2H2 sequences (y-axis). *rho*, Spearman's rank correlation. (E) Violin plots showing the distribution of the predicted pI values in zebrafish genes or domains. pI values of genes without C2H2-ZF (blue), genes with C2H2-ZF (green), and isolated C2H2-ZF (red) are shown. The p values were calculated by the Wilcoxon signed rank test.

Supplemental Figure 5. Additional information about mRNA half-life and disome profiling in human cells.

(A) Chromosomal distributions of C2H2-ZF genes and their relative mRNA half-life in ZNF598 knockdown cells compared to control cells. The x-axis shows chromosome numbers, and the y-axis shows the fold change in mRNA half-life. The value of each C2H2-ZF gene is plotted as a dot. The box represents the IQR, with the median indicated as a thick horizontal line in the box. The whiskers represent the variation within 1.5 IQR outside the upper and lower quartiles. (B) Cumulative distributions of fold changes in mRNA levels in HEK293T cells, comparing ZNF598 knockdown cells and control cells. All genes (black) and C2H2-ZF genes with top (dark blue) or bottom (light blue) disome accumulation scores are shown.

The x-axis shows the fold change in mRNA half-life, and the y-axis shows the cumulative fraction. The p values are shown on the left (Kolmogorov–Smirnov test). (C) A logo representation of amino acid diversity among C2H2-ZF collected from human genes. (D) A histogram showing the A-site position of the leading stalled ribosome in disome footprints relative to the C2H2-ZF sequence in HEK293 cells. (E) Mean net charge in the nascent chains around ribosome collision sites on C2H2-ZF sequences in HEK293 cells. Nascent chain sequences determined by the A-site position of experimentally detected disome footprints (red) and those determined by the randomly chosen A-site position (blue) are shown. (F) Cumulative distributions of fold changes in mRNA levels in HEK293T cells, comparing ZNF598 knockdown cells and control cells. All genes (black), C2H2-ZF genes with the KRAB domain (green), or C2H2-ZF genes without the KRAB domain (pink) are shown. The x-axis shows the fold change in mRNA half-life, and the y-axis shows the cumulative fraction. The p values are shown on the left (Kolmogorov–Smirnov test).

References

Arpat AB, Liechti A, De Matos M, Dreos R, Janich P & Gatfield D (2020) Transcriptome-wide sites of collided ribosomes reveal principles of translational pausing. *Genome Res* 30: 985–999

Arthur LL, Pavlovic-Djuranovic S, Koutmou KS, Green R, Szczesny P & Djuranovic S (2015) Translational control by lysine-encoding A-rich sequences. *Sci Adv* 1: e1500154

Bazzini AA, del Viso F, Moreno-Mateos MA, Johnstone TG, Vejnar CE, Qin Y, Yao J, Khokha MK & Giraldez AJ (2016) Codon identity regulates mRNA stability and translation efficiency during the maternal-to-zygotic transition. *EMBO J* 35: 2087–2103

Best K, Ikeuchi K, Kater L, Best D, Musial J, Matsuo Y, Berninghausen O, Becker T, Inada T & Beckmann R (2023) Structural basis for clearing of ribosome collisions by the RQT complex. *Nat Commun* 14: 921

Bruno M, Mahgoub M & Macfarlan TS (2019) The Arms Race Between KRAB-Zinc Finger Proteins and Endogenous Retroelements and Its Impact on Mammals. *Annu Rev Genet* 53: 393–416

Bucher P, Karplus K, Moeri N & Hofmann K (1996) A flexible motif search technique based on generalized profiles. *Comput Chem* 20: 3–23

Chen L, Muhlrad D, Hauryliuk V, Cheng Z, Lim MK, Shyp V, Parker R & Song H (2010) Structure of the Dom34-Hbs1 complex and implications for no-go decay. *Nat Struct Mol Biol* 17: 1233–40

Chen S, Zhou Y, Chen Y & Gu J (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34: i884–i890

Choe Y-J, Park S-H, Hassemer T, Körner R, Vincenz-Donnelly L, Hayer-Hartl M & Hartl FU (2016) Failure of RQC machinery causes protein aggregation and proteotoxic stress. *Nature* 531: 191–5

Despic V & Neugebauer KM (2018) RNA tales - how embryos read and discard messages from mom. *J Cell Sci* 131: jcs201996

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M & Gingeras TR (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29: 15–21

Doma MK & Parker R (2006) Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. *Nature* 440: 561–4

D'Orazio KN & Green R (2021) Ribosome states signal RNA quality control. *Mol Cell*

D'Orazio KN, Wu CC-C, Sinha N, Loll-Krippleber R, Brown GW & Green R (2019) The endonuclease Cue2 cleaves mRNAs at stalled ribosomes during No Go Decay. *Elife* 8: 1–27

Emerson RO & Thomas JH (2009) Adaptive evolution in zinc finger transcription factors. *PLoS Genet* 5: e1000325

Filbeck S, Cerullo F, Pfeffer S & Joazeiro CAP (2022) Ribosome-associated quality-control mechanisms from bacteria to humans. *Mol Cell* 82: 1451–1466

Galaxy Community (2022) The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. *Nucleic Acids Res* 50: W345–W351

Garzia A, Jafarnejad SM, Meyer C, Chapat C, Gogakos T, Morozov P, Amiri M, Shapiro M, Molina H, Tuschl T, *et al* (2017) The E3 ubiquitin ligase and RNA-binding protein ZNF598 orchestrates ribosome quality control of premature polyadenylated mRNAs. *Nat Commun* 8: 16056

Giraldez AJ, Mishima Y, Rihel J, Grocock RJ, Van Dongen S, Inoue K, Enright AJ & Schier AF (2006) Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. *Science* 312: 75–9

Glover ML, Burroughs AM, Monem PC, Egelhofer TA, Pule MN, Aravind L & Arribere JA (2020) NONU-1 Encodes a Conserved Endonuclease Required for mRNA Translation Surveillance. *Cell Rep* 30: 4321-4331.e4

Hadzhiev Y, Qureshi HK, Wheatley L, Cooper L, Jasiulewicz A, Van Nguyen H, Wragg JW, Poovathumkadavil D, Conic S, Bajan S, *et al* (2019) A cell cycle-coordinated Polymerase II transcription compartment encompasses gene expression before global genome activation. *Nat Commun* 10: 691

Hadzhiev Y, Wheatley L, Cooper L, Ansaloni F, Whalley C, Chen Z, Finaurini S, Gustincich S, Sanges R, Burgess S, *et al* (2023) The miR-430 locus with extreme promoter density forms a transcription body during the minor wave of zygotic genome activation. *Dev Cell* 58: 155-170.e8

Han P, Shichino Y, Schneider-Poetsch T, Mito M, Hashimoto S, Udagawa T, Kohno K, Yoshida M, Mishima Y, Inada T, *et al* (2020) Genome-wide Survey of Ribosome Collision. *Cell Rep* 31: 107610

Hashimoto S, Sugiyama T, Yamazaki R, Nobuta R & Inada T (2020) Identification of a novel trigger complex that facilitates ribosome-associated quality control in mammalian cells. *Sci Rep* 10: 3422

Hershey JWB, Sonenberg N & Mathews MB (2019) Principles of translational control. *Cold Spring Harb Perspect Biol* 11

Howe K, Clark MD, Torroja CF, Torrance J, Berthelot C, Muffato M, Collins JE, Humphray S, McLaren K, Matthews L, *et al* (2013) The zebrafish reference genome sequence and its relationship to the human genome. *Nature* 496: 498–503

Howe K, Schiffer PH, Zielinski J, Wiehe T, Laird GK, Marioni JC, Soylemez O, Kondrashov F & Leptin M (2016) Structure and evolutionary history of a large family of NLR proteins in the zebrafish. *Open Biol* 6: 160009

Ikeuchi K, Izawa T & Inada T (2019a) Recent Progress on the Molecular Mechanism of Quality Controls Induced by Ribosome Stalling. *Front Genet* 9: 1–7

Ikeuchi K, Tesina P, Matsuo Y, Sugiyama T, Cheng J, Saeki Y, Tanaka K, Becker T, Beckmann R & Inada T (2019b) Collided ribosomes form a unique structural interface to induce Hel2-driven quality control pathways. *EMBO J* 38: e100276

Imbeault M, Helleboid P-Y & Trono D (2017) KRAB zinc-finger proteins contribute to the evolution of gene regulatory networks. *Nature* 543: 550–554

Inada T (2020) Quality controls induced by aberrant translation. *Nucleic Acids Res* 48: 1084–1096

Ingolia NT, Brar GA, Rouskin S, McGeachy AM & Weissman JS (2012) The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nat Protoc* 7: 1534–50

Iwasaki S, Floor SN & Ingolia NT (2016) Rocaglates convert DEAD-box protein eIF4A into a sequence-selective translational repressor. *Nature* 534: 558–61

Jacobs FMJ, Greenberg D, Nguyen N, Haeussler M, Ewing AD, Katzman S, Paten B, Salama SR & Haussler D (2014) An evolutionary arms race between KRAB zinc-finger genes ZNF91/93 and SVA/L1 retrotransposons. *Nature* 516: 242–245

Juszkiewicz S, Chandrasekaran V, Lin Z, Kraatz S, Ramakrishnan V & Hegde RS (2018) ZNF598 Is a Quality Control Sensor of Collided Ribosomes. *Mol Cell* 72: 469-481.e7

Juszkiewicz S & Hegde RS (2017) Initiation of Quality Control during Poly(A) Translation Requires Site-Specific Ribosome Ubiquitination. *Mol Cell* 65: 743-750.e4

Juszkiewicz S, Speldewinde SH, Wan L, Svejstrup JQ & Hegde RS (2020) The ASC-1 Complex Disassembles Collided Ribosomes. *Mol Cell* 79: 603-614.e8

Karousis ED & Mühlemann O (2022) The broader sense of nonsense. *Trends Biochem Sci* 47: 921–935

Kauzlaric A, Ecco G, Cassano M, Duc J, Imbeault M & Trono D (2017) The mouse genome displays highly dynamic populations of KRAB-zinc finger protein genes and related genetic units. *PLoS One* 12: e0173746

Kim KQ & Zaher HS (2022) Canary in a coal mine: collided ribosomes as sensors of cellular conditions. *Trends Biochem Sci* 47: 82–97

Klug A (2010) The discovery of zinc fingers and their applications in gene regulation and genome manipulation. *Annu Rev Biochem* 79: 213–31

Kramer G, Shiber A & Bukau B (2019) Mechanisms of Cotranslational Maturation of Newly Synthesized Proteins. *Annu Rev Biochem* 88: 337–364

Kusov YY, Shatirishvili G, Dzagurov G & Gauss-Müller V (2001) A new G-tailing method for the determination of the poly(A) tail length applied to hepatitis A virus RNA. *Nucleic Acids Res* 29: E57-7

Lee MT, Bonneau AR & Giraldez AJ (2014) Zygotic Genome Activation During the Maternal-to-Zygotic Transition. *Annu Rev Cell Dev Biol* 30: 581–613

Liao Y, Smyth GK & Shi W (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30: 923–30

Liutkute M, Samatova E & Rodnina M V (2020) Cotranslational Folding of Proteins on the Ribosome. *Biomolecules* 10

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

Lu B (2022) Translational regulation by ribosome-associated quality control in neurodegenerative disease, cancer, and viral infection. *Front Cell Dev Biol* 10

Lu J & Deutsch C (2008) Electrostatics in the Ribosomal Tunnel Modulate Chain Elongation Rates. *J Mol Biol* 384: 73–86

Matheisl S, Berninghausen O, Becker T & Beckmann R (2015) Structure of a human translation termination complex. *Nucleic Acids Res* 43: 8615–8626

Matsuo Y, Ikeuchi K, Saeki Y, Iwasaki S, Schmidt C, Udagawa T, Sato F, Tsuchiya H, Becker T, Tanaka K, *et al* (2017) Ubiquitination of stalled ribosome triggers ribosome-associated quality control. *Nat Commun* 8: 159

Matsuo Y, Tesina P, Nakajima S, Mizuno M, Endo A, Buschauer R, Cheng J, Shounai O, Ikeuchi K, Saeki Y, *et al* (2020) RQT complex dissociates ribosomes collided on endogenous RQC substrate SDD1. *Nat Struct Mol Biol* 27: 323–332

Meydan S & Guydosh NR (2020) Disome and Trisome Profiling Reveal Genome-wide

Targets of Ribosome Quality Control. *Mol Cell*: 1–15

Meydan S & Guydosh NR (2021) A cellular handbook for collided ribosomes: surveillance pathways and collision types. *Curr Genet* 67: 19–26

Mishima Y, Han P, Ishibashi K, Kimura S & Iwasaki S (2022) Ribosome slowdown triggers codon-mediated mRNA decay independently of ribosome quality control. *EMBO J* 41

Mishima Y & Tomari Y (2016) Codon Usage and 3′ UTR Length Determine Maternal mRNA Stability in Zebrafish. *Mol Cell* 61: 874–885

Mito M, Mishima Y & Iwasaki S (2020) Protocol for Disome Profiling to Survey Ribosome Collision in Humans and Zebrafish. *STAR Protoc* 1: 100168

Monem PC, Vidyasagar N, Piatt AL, Sehgal E & Arribere JA (2023) Ubiquitination of stalled ribosomes enables mRNA decay via HBS-1 and NONU-1 in vivo. *PLoS Genet* 19: e1010577

Morita M, Ler LW, Fabian MR, Siddiqui N, Mullin M, Henderson VC, Alain T, Fonseca BD, Karashchuk G, Bennett CF, *et al* (2012) A Novel 4EHP-GIGYF2 Translational Repressor Complex Is Essential for Mammalian Development. *Mol Cell Biol* 32: 3585–3593

Morris C, Cluet D & Ricci EP (2021) Ribosome dynamics and mRNA turnover, a complex relationship under constant cellular scrutiny. *Wiley Interdiscip Rev RNA* 12: e1658

Nasevicius A & Ekker SC (2000) Effective targeted gene 'knockdown' in zebrafish. *Nat Genet* 26: 216–20

Nilsson OB, Hedman R, Marino J, Wickles S, Bischoff L, Johansson M, Müller-Lucks A, Trovato F, Puglisi JD, O'Brien EP, *et al* (2015) Cotranslational Protein Folding inside the Ribosome Exit Tunnel. *Cell Rep* 12: 1533–40

Robinson MD, McCarthy DJ & Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26: 139–40

Scahill CM, Digby Z, Sealy IM, Wojciechowska S, White RJ, Collins JE, Stemple DL, Bartke T, Mathers ME, Patton EE, *et al* (2017) Loss of the chromatin modifier Kdm2aa causes BrafV600E-independent spontaneous melanoma in zebrafish. *PLoS Genet* 13: e1006959

Schmitges FW, Radovani E, Najafabadi HS, Barazandeh M, Campitelli LF, Yin Y, Jolma A, Zhong G, Guo H, Kanagalingam T, *et al* (2016) Multiparameter functional diversity of human C2H2 zinc finger proteins. *Genome Res* 26: 1742–1752

Schmitz JF, Zimmer F & Bornberg-Bauer E (2016) Mechanisms of transcription factor evolution in Metazoa. *Nucleic Acids Res* 44: 6287–97

Schuller AP & Green R (2018) Roadblocks and resolutions in eukaryotic translation. *Nat Rev Mol Cell Biol* 19: 526–541

Sharp PM & Li WH (1987) The codon Adaptation Index--a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res* 15: 1281–95

Sigrist CJA, de Castro E, Cerutti L, Cuche BA, Hulo N, Bridge A, Bougueleret L & Xenarios I (2013) New and continuing developments at PROSITE. *Nucleic Acids Res* 41: D344-7

Simms CL, Hudson BH, Mosior JW, Rangwala AS & Zaher HS (2014) An Active Role for the Ribosome in Determining the Fate of Oxidized mRNA. *Cell Rep* 9: 1256–1264

Simms CL, Yan LL & Zaher HS (2017) Ribosome Collision Is Critical for Quality Control during No-Go Decay. *Mol Cell* 68: 361-373.e5

Sitron CS & Brandman O (2020) Detection and Degradation of Stalled Nascent Chains via Ribosome-Associated Quality Control. *Annu Rev Biochem* 89: 417–442

Stein KC, Kriel A & Frydman J (2019) Nascent Polypeptide Domain Topology and Elongation Rate Direct the Cotranslational Hierarchy of Hsp70 and TRiC/CCT. *Mol Cell* 75: 1117-1130.e5

Stein KC, Morales-Polanco F, van der Lienden J, Rainbolt TK & Frydman J (2022) Ageing exacerbates ribosome pausing to disrupt cotranslational proteostasis. *Nature* 601: 637–642

Stubbs L, Sun Y & Caetano-Anolles D (2011) Function and Evolution of C2H2 Zinc Finger Arrays. *Subcell Biochem* 52: 75–94

Sundaramoorthy E, Leonard M, Mak R, Liao J, Fulzele A & Bennett EJ (2017) ZNF598 and RACK1 Regulate Mammalian Ribosome-Associated Quality Control Function by Mediating Regulatory 40S Ribosomal Ubiquitylation. *Mol Cell* 65: 751-760.e4

Tadros W & Lipshitz HD (2009) The maternal-to-zygotic transition: a play in two acts. *Development* 136: 3033–3042

Tani H, Mizutani R, Salam KA, Tano K, Ijiri K, Wakamatsu A, Isogai T, Suzuki Y & Akimitsu N (2012) Genome-wide determination of RNA stability reveals hundreds of short-lived noncoding transcripts in mammals. *Genome Res* 22: 947–956

Tesina P, Lessen LN, Buschauer R, Cheng J, Wu CC, Berninghausen O, Buskirk AR, Becker

T, Beckmann R & Green R (2020) Molecular mechanism of translational stalling by inhibitory codon combinations and poly(A) tracts. *EMBO J* 39: 1–17

Thomas JH & Schneider S (2011) Coevolution of retroelements and tandem zinc finger genes. *Genome Res* 21: 1800–1812

Tomomatsu S, Watanabe A, Tesina P, Hashimoto S, Ikeuchi K, Li S, Matsuo Y, Beckmann R & Inada T (2023) Two modes of Cue2-mediated mRNA cleavage with distinct substrate recognition initiate no-go decay. *Nucleic Acids Res* 51: 253–270

Tsuboyama K, Osaki T, Matsuura-Suzuki E, Kozuka-Hata H, Okada Y, Oyama M, Ikeuchi Y, Iwasaki S & Tomari Y (2020) A widespread family of heat-resistant obscure (Hero) proteins protect against protein instability and aggregation. *PLoS Biol* 18: e3000632

Veltri AJ, D'Orazio KN, Lessen LN, Loll-Krippleber R, Brown GW & Green R (2022) Distinct elongation stalls during translation are linked with distinct pathways for mRNA degradation. *Elife* 11

Weber R, Chung M-Y, Keskeny C, Zinnall U, Landthaler M, Valkov E, Izaurralde E & Igreja C (2020) 4EHP and GIGYF1/2 Mediate Translation-Coupled Messenger RNA Decay. *Cell Rep* 33: 108262

White RJ, Collins JE, Sealy IM, Wali N, Dooley CM, Digby Z, Stemple DL, Murphy DN, Billis K, Hourlier T, *et al* (2017) A high-resolution mRNA expression time course of embryonic development in zebrafish. *Elife* 6

Wilson DN, Arenz S & Beckmann R (2016) Translation regulation via nascent polypeptide-mediated ribosome stalling. *Curr Opin Struct Biol* 37: 123–133

Wolf G, de Iaco A, Sun M-A, Bruno M, Tinkham M, Hoang D, Mitra A, Ralls S, Trono D & Macfarlan TS (2020) KRAB-zinc finger protein gene expansion in response to active retrotransposons in the murine lineage. *Elife* 9: 1–22

Yan LL, Simms CL, McLoughlin F, Vierstra RD & Zaher HS (2019) Oxidation and alkylation stresses activate ribosome-quality control. *Nat Commun* 10: 5611

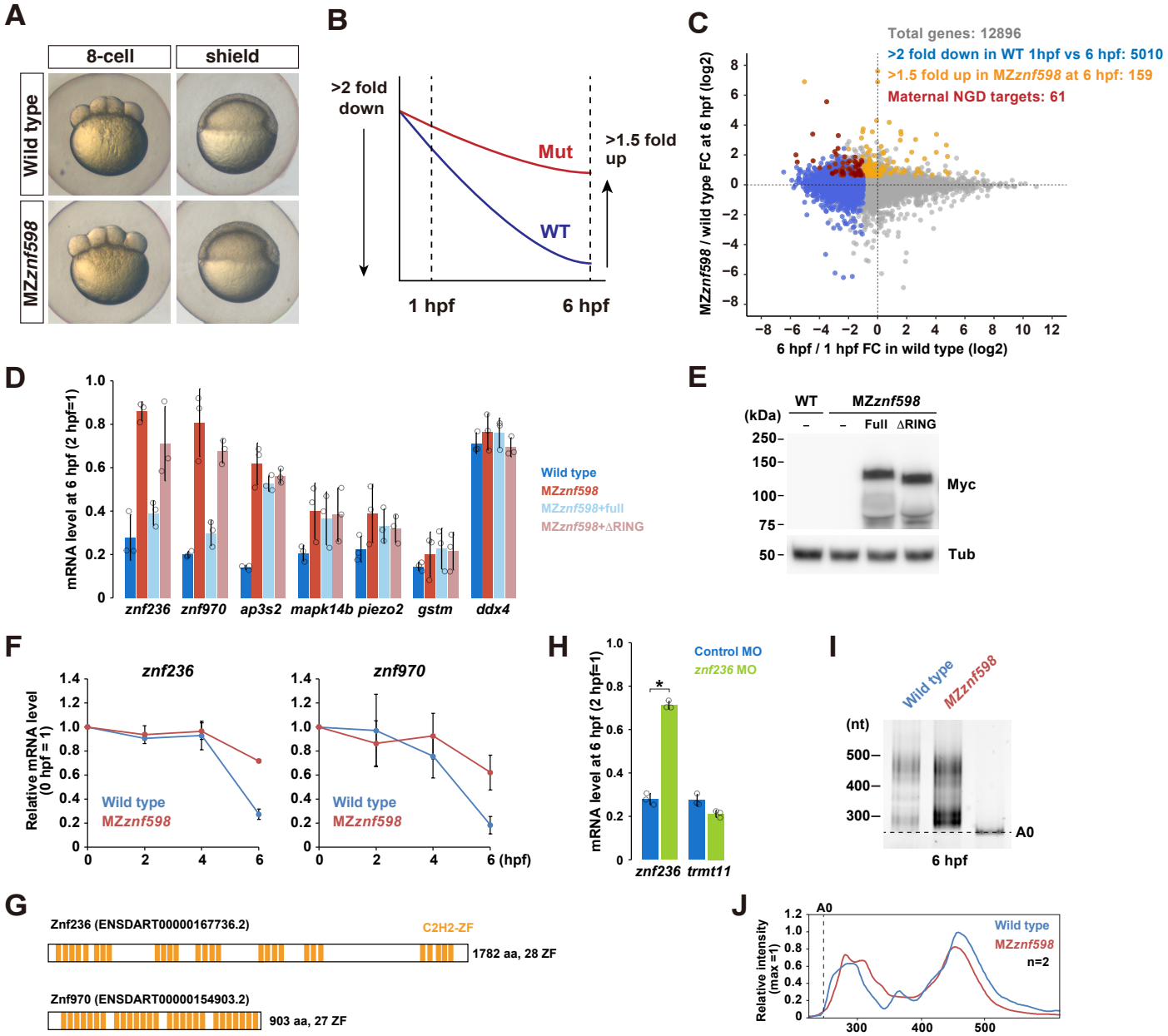
Yang B, Fang L, Gao Q, Xu C, Xu J, Chen Z-X, Wang Y & Yang P (2022) Species-specific KRAB-ZFPs function as repressors of retroviruses by targeting PBS regions. *Proc Natl Acad Sci U S A* 119: e2119415119

Yartseva V & Giraldez AJ (2015) The Maternal-to-Zygotic Transition During Vertebrate Development: A Model for Reprogramming. *Curr Top Dev Biol* 113: 191–232

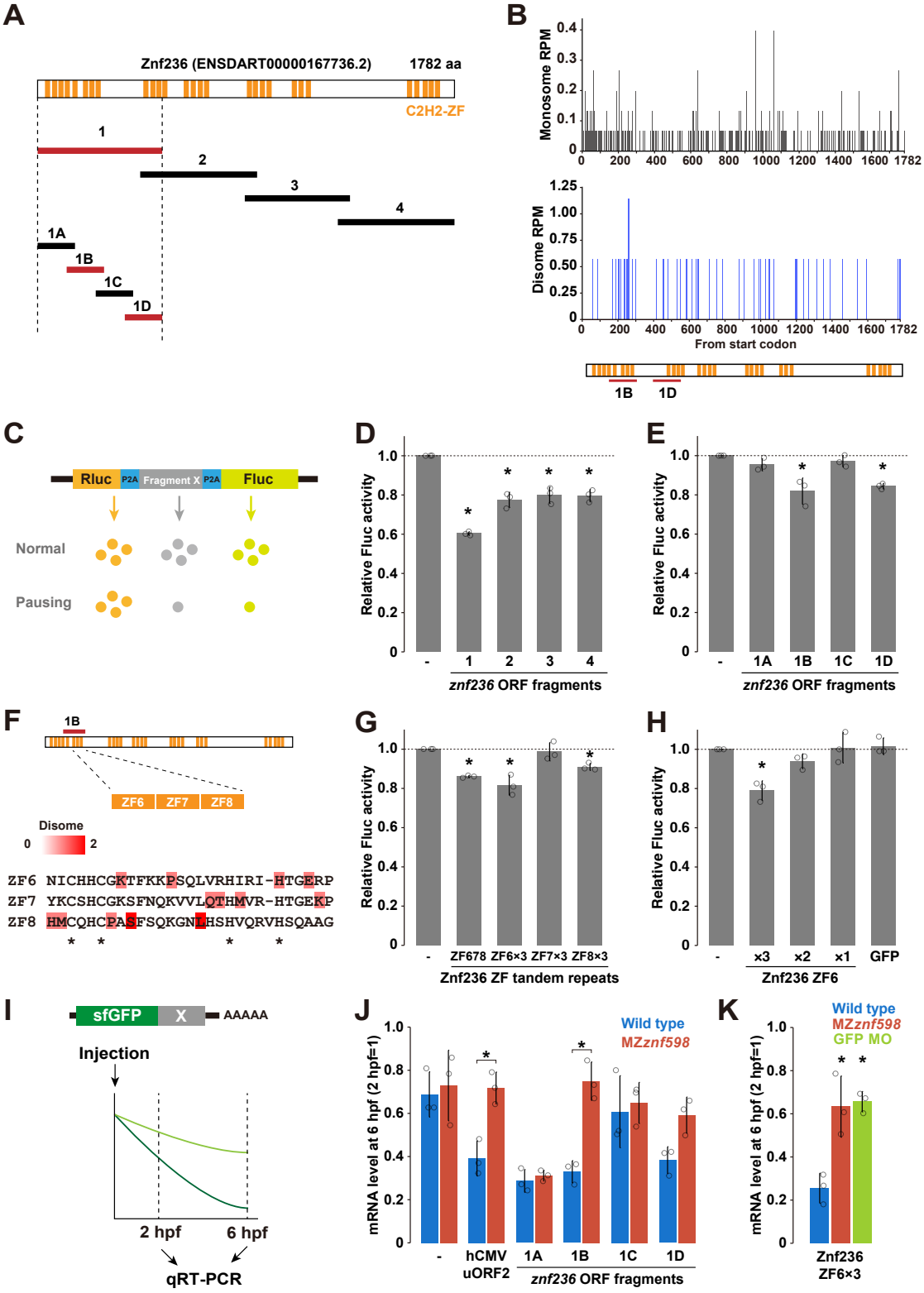
Zhao T, Chen YM, Li Y, Wang J, Chen S, Gao N & Qian W (2021) Disome-seq reveals widespread ribosome collisions that promote cotranslational protein folding. *Genome*

Biol 22

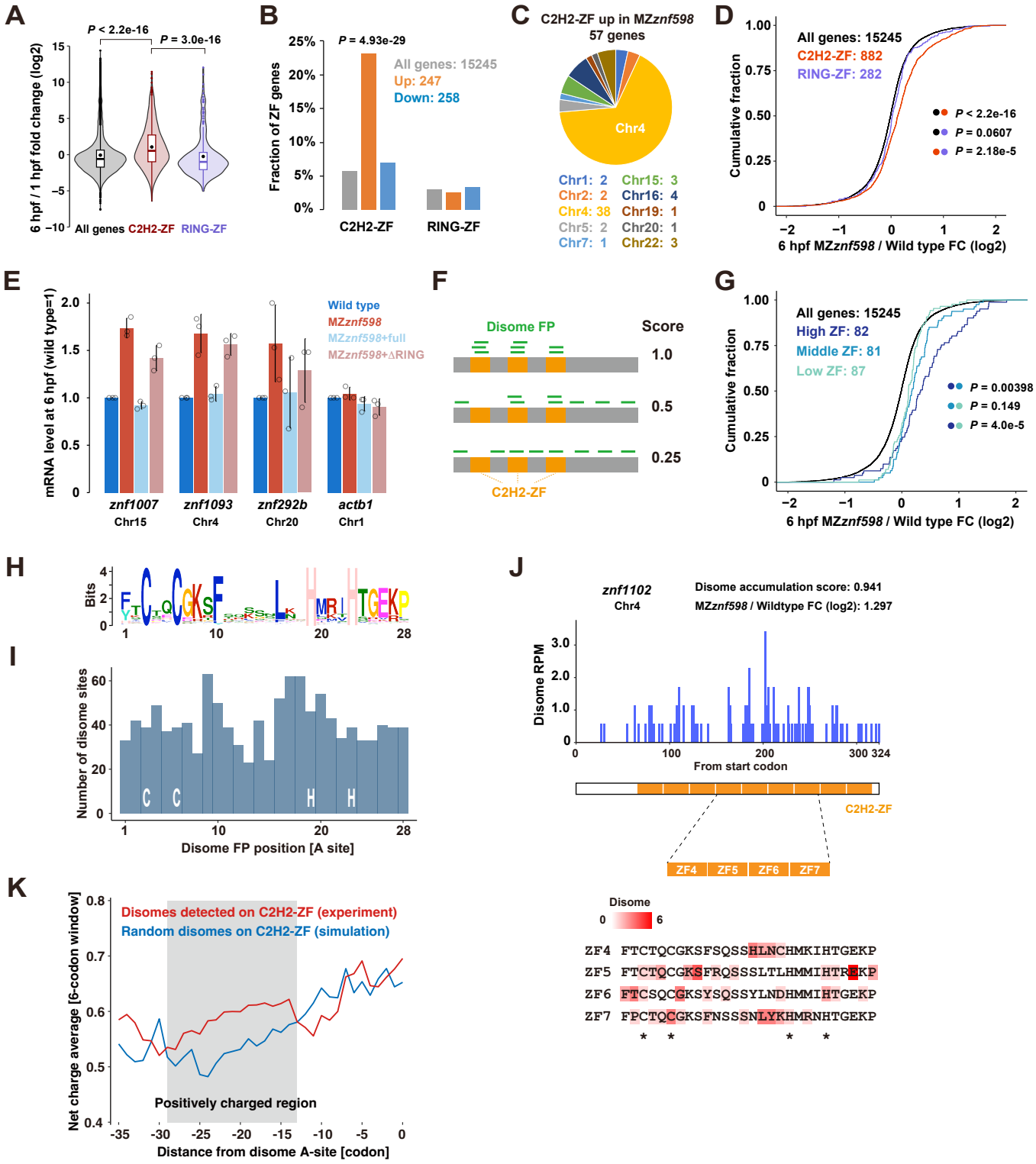
Ishibashi *et al.* Figure 1



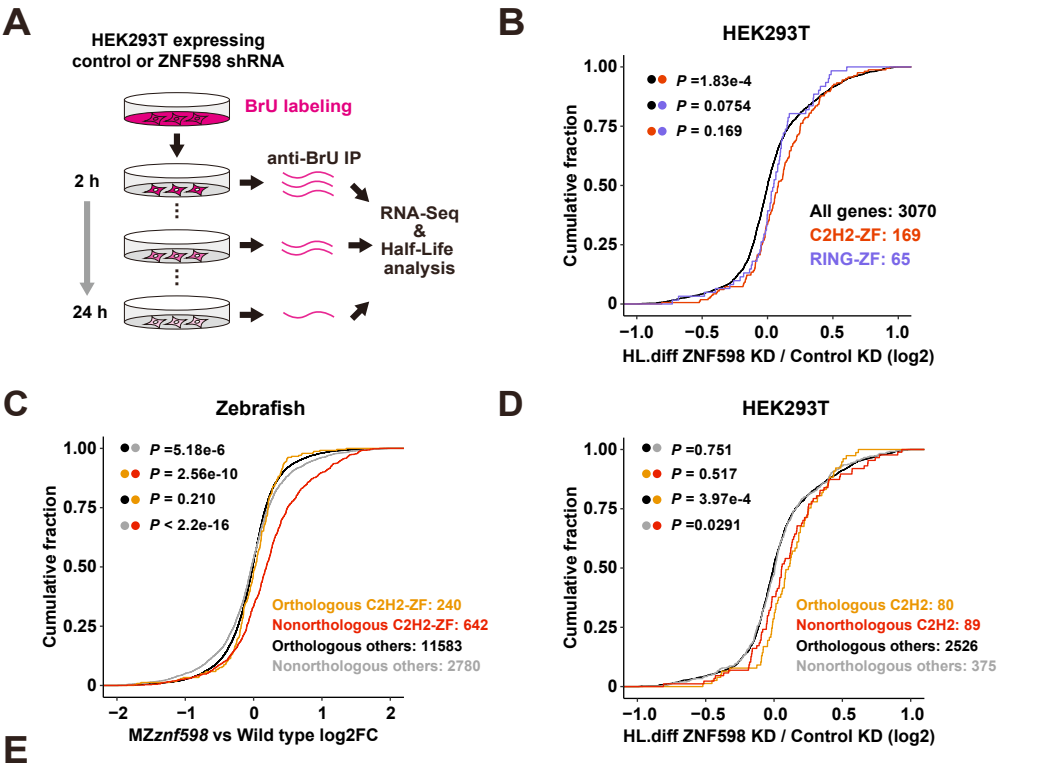
Ishibashi *et al.* Figure 2

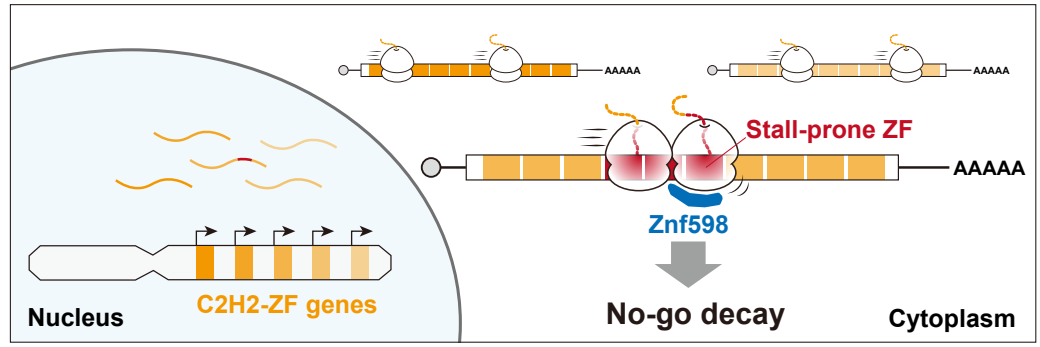


Ishibashi *et al.* Figure 3

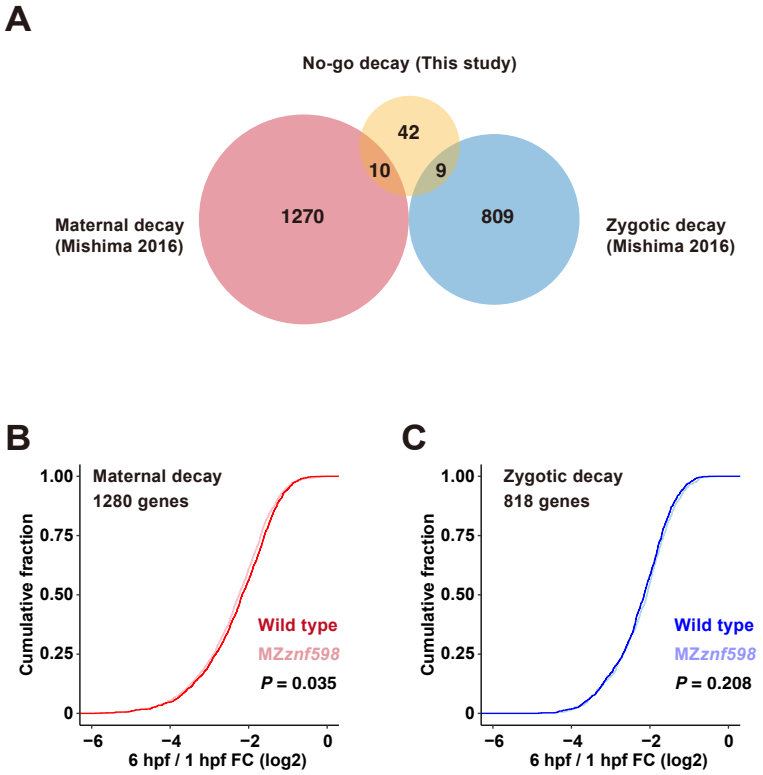


Ishibashi *et al.* Figure 4

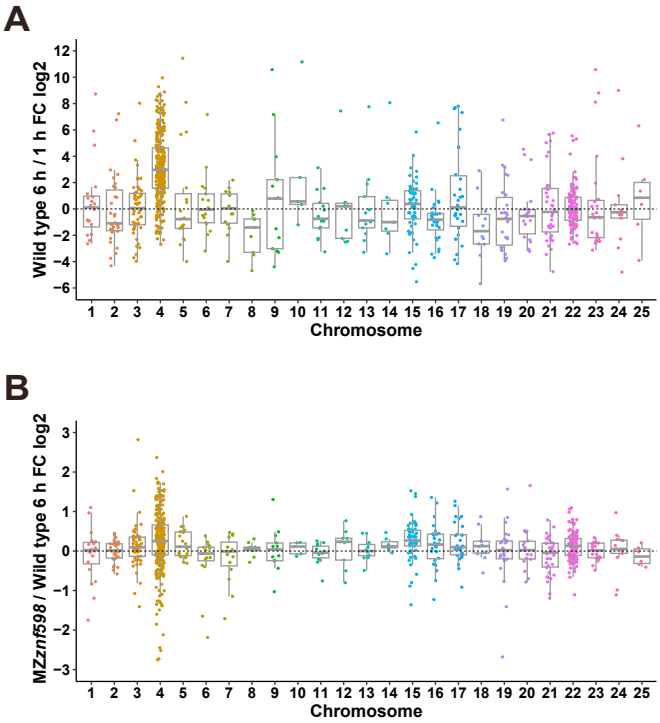




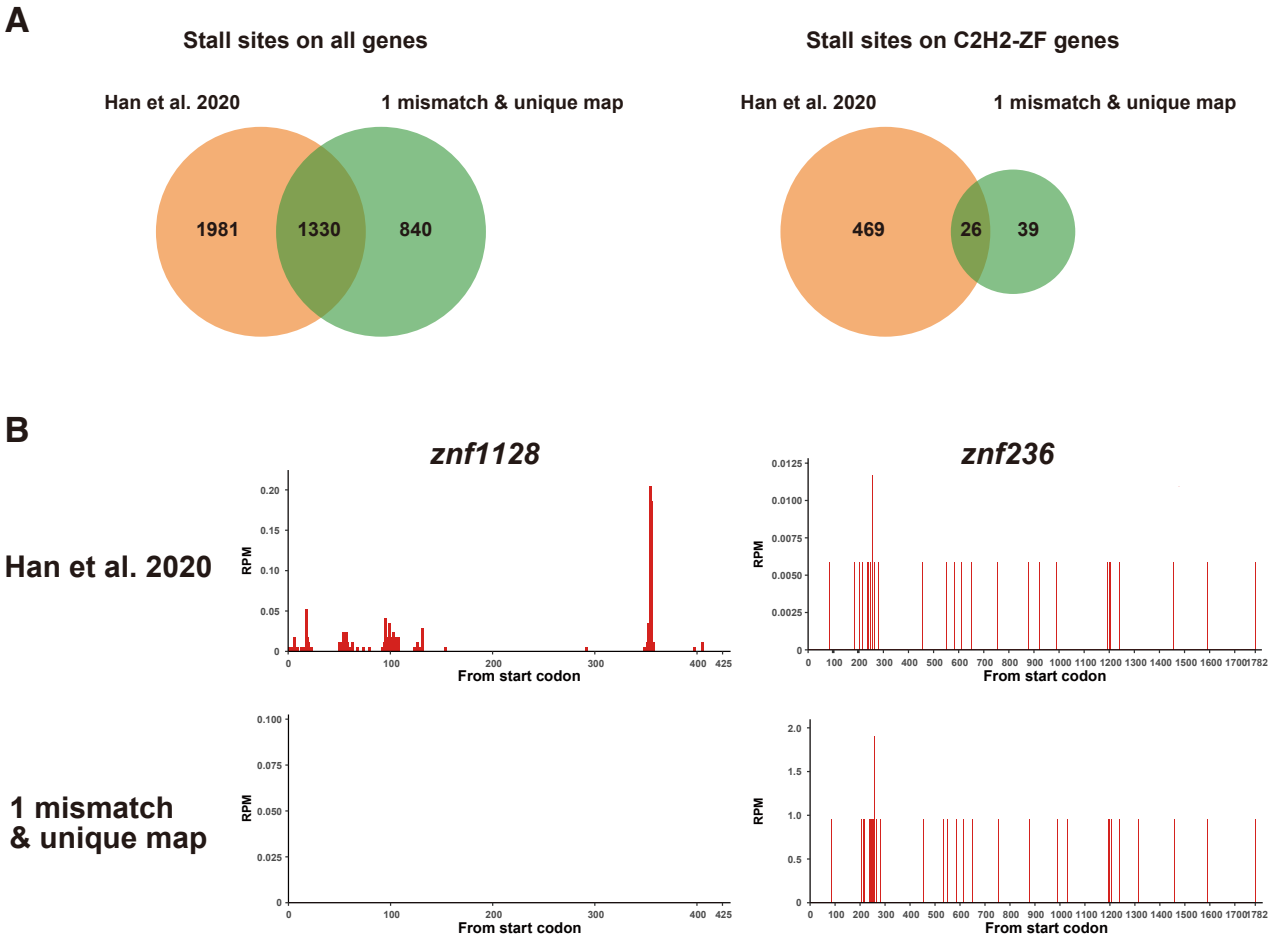
Ishibashi *et al*., Supplemental Figure 1



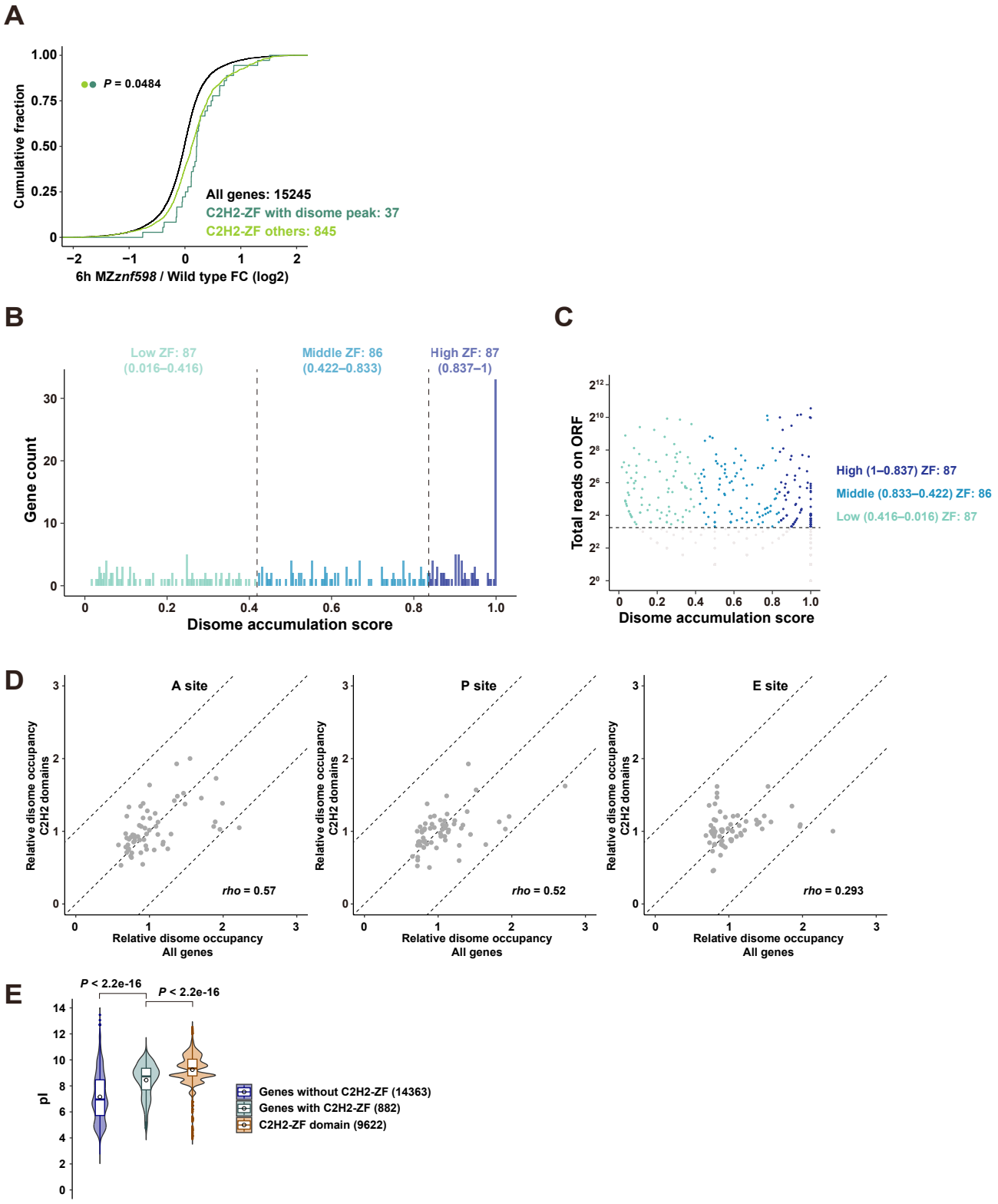
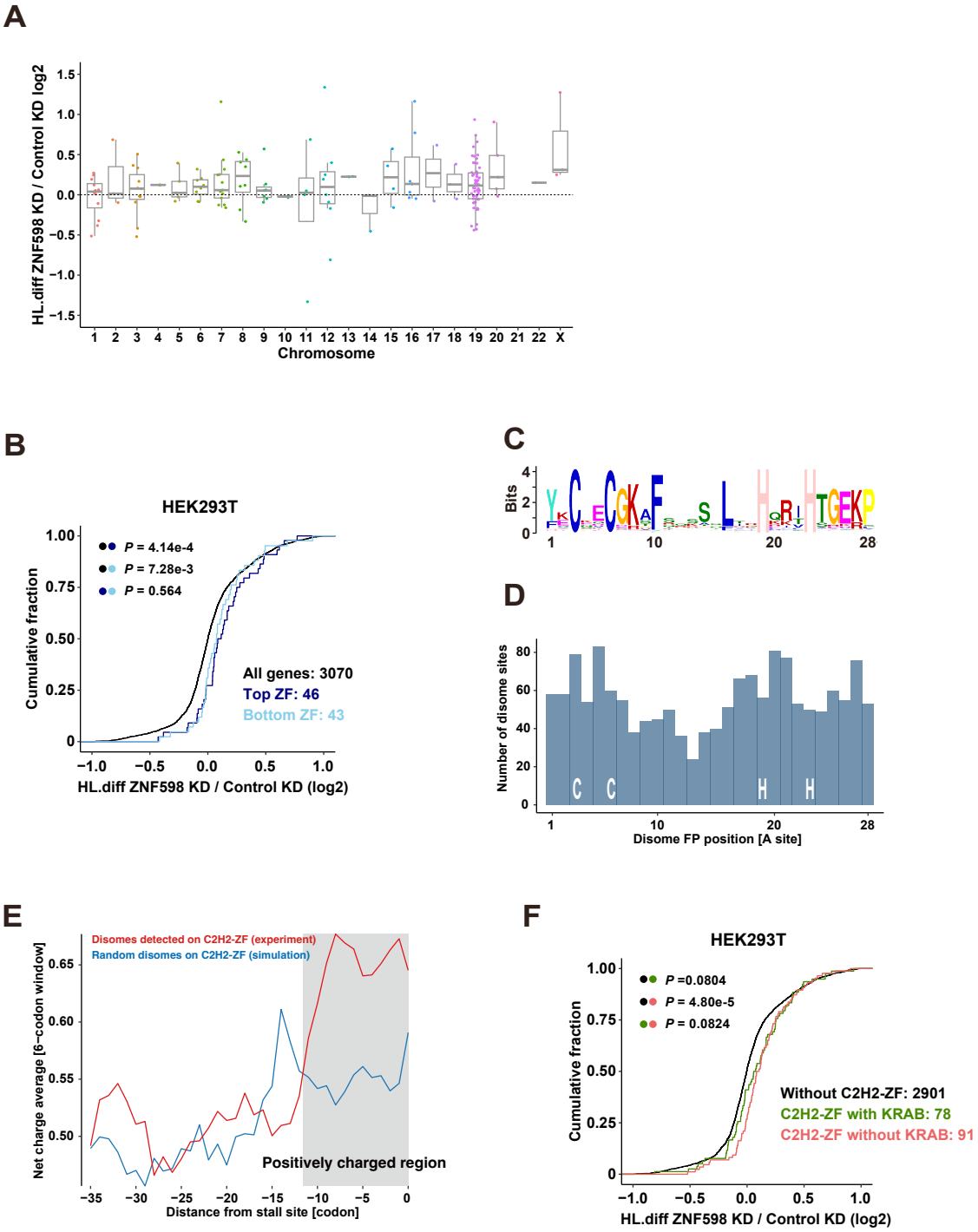
Ishibashi *et al*., Supplemental Figure 2



Ishibashi *et al.*, Supplemental Figure 3



Ishibashi *et al.*, Supplemental Figure 4



Ishibashi *et al*., Supplemental Figure 5