Supplemental Methods

Zebrafish maintenance

Danio rerio (zebrafish) embryos were obtained through natural mating of 5-18 month old adults of the TU-AB stain and maintained at 28°C. Fish lines were maintained according to the International Association for Assessment and Accreditation of Laboratory Animal Care research guidelines, and protocols were approved by the Yale University Institutional Animal Care and Use Committee (IACUC).

Early zebrafish embryo transcriptome

Timecourse sample collection and RNA-seq library preparation

For the large developmental gene expression time course, synchronously developing embryos were collected every 30 minutes (between 0 hours per fertilization (hpf) and 5 hpf) and at 6 hpf and 8 hpf. Ten embryos were collected for each time point in duplicate. To allow for appropriate normalization of fold-changes, yeast total RNA (roughly 5% of estimated zebrafish total RNA) was spiked into TRIzol (Invitrogen, USA) prior to RNA extraction according to manufacturer's instructions (Invitrogen, USA). After extraction, RNA was subjected to poly(A)-selected RNA-seq library preparation (TruSeq mRNA, Illumina, USA) and ribosomal RNA-depleted total RNA-seq library preparation (TruSeq Total RNA, Illumina, USA) according to manufacturer's instructions. For Pol II inhibition, α-amanitin was obtained from Sigma-Aldrich and resuspended in nuclease-free water. Dechorionated embryos were injected with 0.2ng of α-amanitin at one-cell stage.

RNA-seq analysis: Developmental modes and their contribution to decay

Differential expression of genes

Raw reads were mapped to the zebrafish Zv9 and yeast R64-1-1 genomes using STAR (Dobin et al, 2013) version 2.4.2a with the following non-default parameters: *--alignEndsType --Local --outFilterMultimapNmax 100 --seedSearchStartLmax 30 --sjdbScore 10*. Genomic sequence indices for STAR were built including exon-junction coordinates from Ensembl r78 (Aken et al, 2017). Read counts per gene were computed by summing the total number of reads overlapping at least 10 nucleotides the gene annotation, including only uniquely mapped reads in the genome. Per gene annotation was obtained by concatenating all Ensembl isoforms together. To determine significantly over- and under-expressed genes, gene read counts were compared using DESeq2 (Love et al, 2014). Genes below 1 count in all replicates in either condition were excluded from the analysis. Total gene counts for all yeast genes were summed and the ratios between conditions were input as Factors to DESeq2. To get DE genes, counts for all Ensembl genes were input to the *results* function with *pAdjustMethod="fdr"* and *independentFiltering=FALSE* options.

Decay mode of genes

To categorize decay genes in *Maternal, Zygotic: miR-430 dependent* and *Zygotic: miR-430 independent* modes, conditions 2 hpf, 2 hour (h) α-amanitin, 6 hpf, 6 h α-amanitin and 6 h LNA miR-430 were used. First, decaying genes were selected as significantly higher at 2 hpf vs. 6 hpf and not significantly higher at 6hpf vs. 6h α-amanitin. Second, among decaying genes, *Zygotic: miR-430 dependent* were selected as significantly higher at 6h LNA miR-430 vs. 6 hpf. Remaining genes were then split into two modes. If only 6 h α-amanitin was significantly higher than 6 hpf, genes were classified as *Zygotic: miR-430 independent*. If only 2 h α-amanitin was significantly higher than 6 h α-amanitin, genes were classified as *Maternal*. If a gene could be classified in both lists, the most significant mode was prioritized, i.e. the lowest *P* determined the mode.

Contribution of each mode to gene decay

To calculate the contribution of each mode, *Maternal, Zygotic: miR-430 dependent* and *Zygotic: miR-430 independent*, $\log_2$ fold-changes as determined by DESeq2 for *Maternal mode*, (2 h α-amanitin vs. 6 h α-amanitin), *Zygotic: miR-430 dependent* (6 hpf vs. 6 h LNA miR-430) and *Zygotic: miR-430 independent* (6 hpf vs. 6 h α-amanitin) were used. Because this analysis used the 3 fold-changes corresponding to each mode and was not limited to using a single statistical test

as it was the case to determine the main mode, fold-changes could have a biologically non-expected sign. These fold-changes were set to not contribute: the three ratios above for *Maternal, Zygotic: miR-430 dependent* and *Zygotic: miR-430 independent* mode fold-changes that were positive, negative and negative, respectively, were set to zero. Because the α-amanitin treatment is also blocking *miR-430* expression, the *Maternal* and *Zygotic: miR-430 independent* modes were first normalized together to 1 and then within *Zygotic* modes, both *miR-430 dependent* and *independent* fold-changes were normalized.

In situ hybridization

In situ hybridization was performed as in Thisse et al, 2008 with 20ng of DIG-labeled RNA probe per 200µL hybridization reaction. RNA probes were synthesized from T7 promoter polymerase overhang on the reverse oligo. Oligo sequences used to amplify transcript regions are listed below (*in situ* probes):

org_Forward: `TGACTGACCAGTGCAACTACG`

org_Reverse: `AAACAGCAAATCGAGAAGCAA`

trip10_Forward: `ATGGACTGGGGAACTGAGCTTTG`

trip10_Reverse: `GAGAACCATAGAGTCATTCCTCG`

dnajc5ga_Forward: `ATCGCTGTACGCTTTCAAGG`

dnajc5ga_Reverse: `AAAACCCACTTCCCTTCTGG`

For *in situ* hybridization, greater than 20 embryos were analyzed for each condition and all displayed comparable level of staining at equal stain development time.

RESA

RESA reporter library construction

The transcriptome-based reporter library was generated with a previously constructed Illumina RNA sequencing library from Bazzini et al, 2012 by overlap-extension PCR with primers mapping to the SP6 promoter and downstream of the SV40 polyadenylation site (Yartseva et al, 2017). Five separate extension reactions were performed (5 cycles of 98C, 61C, and 72C), then pooled and used as a template for further amplification (48 separate PCR reactions of 10 cycles each). PCR products were pooled and concentrated by MiniElute PCR Purification Kit (Qiagen), followed by PAGE purification. Following *in vitro* transcription, reporter mRNA was injected into ~75 embryos at 100 ng/µL (1 nl). Total RNA was extracted from embryos using TRIzol reagent according to manufacturer's instructions. Reverse transcription was performed with a reporter-specific primer (`CATCAATGTATCTTATCATGTCTGGATC`; SuperScript III (Invitrogen)). The final Illumina library was prepared using the following primers: 5'-`ATGATACGGCGACCACCGACAGGTTCAGAGTTCTACAGTCCGACGATC`-3' and 5'- `CAAGCAGAAGACGGCATACGAGAT`(barcode)`GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT`-3'. PCR reactions were pooled (5 reactions (17 cycles each; Phusion (NEB))) and purified by PAGE. The targeted library was prepared as described previously (Yartseva et al, 2017).

T1 5' (Sp6 promoter, GFP coding sequence, and 5' Illumina adaptor): `GCTTGATTTAGGTGACACTATAGAATACAAGCTACTTG`
`TTCTTTTTGCAGGATCCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAG`
`TTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCAC`
`CCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCT`
`ACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAG`
`CTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCA`
`GAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCG`
`ACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAG`

TTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTCCGGACTCAGATCTAAGCTGAACCCTCCTGATGAGAGTGGCCCCGGCTG
CATGAGCTGCAAGTGTGTGCTCTCCTGACTCGAGGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT

T2 3' (Illumina adaptor and SV40 polyadenylation signal): AGATCGGAAGAGCACACGTCTGAACTCCAGTCACTCTAGAACTATAG
TGAGTCGTATTACGTAGATCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAA
TTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAG
GTGTGGGAGGTTTTTTAATTC

Reporter validation by qRT-PCR

Validation sequences were cloned into PCS2+ vector downstream of GFP using XbaI and XhoI restriction sites. RNA was *in vitro* synthesized using Sp6 mMessage Machine (ThermoFisher AM1340) from NotI linearized plasmids. Zebrafish embryos were injected with 4pg of each reporter and dsRED control mRNAs and collected at 2hpf and 8hpf. Total RNA (250ng) was reverse transcribed using SuperScript III Kit (Invitrogen #18080-051) with random hexamers. cDNA was diluted 1:20 and used in 10µL reaction (5µL SYBR Green master mix (ThermoFisher #4472908), 0.5µL 10µM Forward and Reverse primer mix, 1µL 1:20 diluted cDNA, 3.5µL water). A common forward primer complementary to GFP - CATG GTCCTGCTGGAGTTCGTGAC and a reporter-specific reverse primer (UACA - GTACATGAGACTCAATCACTGCTC, tdrd7 - TTCGAAAC ACCATGGATGCTTCTC) were designed. dsRED was amplified with Forward GAAGGGCGAGATCCACAAG and Reverse GGACTTGAA CTCCACCAGGTA primers. Biological and technical triplicates were performed for each sample and relative expression with ΔΔCT method was measured using ViiA 7 Software v1.2.2 with dsRED mRNA as a reference control.

Genetic deletion of the trip10 3'-UTR region

CRISPR-mediated mutagenesis was performed as described in Moreno-Mateos et al, 2015. Standard PCR with Taq enzyme was used for genotyping with For-CATGAAAGCCACGTCGATAA and Rev-CAAATGCAAAACAAACACTCG and annealing temperature of 61°C. A deletion of 223 nt was found: CTATGAACTGCGATTAATAGTTTGTTGAAATCTAAATAATATAATTTATTGAG TAAATAAGCGCTTGTATATTAAATAAACATGTATGTAAGATGTATCTGCACACGTTTGCTATTTCCGATAGGTCCACAACAATGCAGAGCGTGTTCGC GATGGCGAAAAACTATTTTAGGCTACTTATGAATATGTGGAATAATATATTTTAGATGTTACGACACCACTA.

RESA profiles

The Illumina TruSeq index adaptor sequence was first trimmed from raw reads (i) for transcriptomic library by aligning its sequence, requiring 100% match of the first five base pairs and a minimum global alignment score of 60 (Matches: 5, Mismatches: −4, Gap opening: −7, Gap extension: −7, Cost-free ends gaps) or (ii) for targeted library using Skewer (Jiang et al, 2014). Reads were then mapped to the zebrafish Zv9 genome using STAR (Dobin et al, 2013) version 2.4.2a with the following non-default parameters: *--alignEndsType EndToEnd --outFilterMultimapNmax 100 --seedSearchStartLmax 30 --sjdbScore 2*. Genomic sequence indices for STAR were built including exon-junction coordinates from Ensembl r78 (Aken et al, 2017).

Individual RESA profiles for each transcript (transcriptomic library) or UTR (targeted library) are positional read coverage. Coverage was calculated using pooled biological replicates by incrementing each position in the profile from first to the last overlapping positions of reads that were sense to the transcript/UTR. For the targeted library, sequenced with paired-end reads, both reads were merged into a single fragment, and every position of this fragment was used to increment the coverage. Profiles were then normalized to CPM (Count Per Million) using total counts of all transcript/ UTR profiles per sample.

FIRE analysis

Regulatory categories of transcriptome segments

To evaluate the regulatory effect of short RNA elements, RESA profiles were split into fixed-length sequence segments (30 and 100 nucleotides long for the transcriptomic and targeted libraries respectively) with 33% overlaps (Oikonomou et al, 2014). For every segment and each library, a score was calculated by averaging over the length of the segment the log-transformed frequency of read counts observed at each base. When comparing between two libraries at different time points (e.g. 2hpf and 6hpf) or treatments (e.g. wild type and LNA-treated), the difference of the scores of a given segment in the two libraries is transformed into a Z-score from the distribution of the all differences (e.g. *z(WT_2hpf; WT_6hpf)* is the Z-score of the difference between wild type at 2hpf and 6hpf). Z-scores can be used to evenly compare differences between different pairs of libraries.

To determine whether a given sequence segment is modulated by either a maternal, zygotic, *miR-430*, or stabilizing regulatory effect, the following strategy was used. For each segment, changes in sequencing coverage between the following five pairs of libraries were considered: (1) wild type at 2 hpf and 6 hpf (2) α-amanitin treated at 6 hpf and wild type at 6 hpf (3) LNA-430 treated at 6 hpf and wild type at 6 hpf (4) α-amanitin treated at 6 hpf and wild type at 2 hpf (5) LNA-430 treated at 6 hpf and wild type at 2 hpf. For each pair, Z-scores are computed for every segment as described above. Functional categories were then assigned to each sequence segment based on the values of these Z-scores.

For the destabilizing regulatory categories, only segments with a sequencing score at 2 hpf above threshold (segment score > 2) were considered to ensure that the sequence segment was present at the early stage. The maternal category was defined as: $z(WT\_2hpf; WT\_6hpf) < -z_{th}$ *and* $z(WT\_2hpf; \alpha Am\_6hpf) < -z_{th}$ *and* $z(WT\_2hpf; LNA\_6hpf) < -z_{th}$, for a permissive threshold of $z_{th}=0.5$. This pattern corresponds to a decaying number of reads at 6 hpf with respect to 2 hpf for the wild type that is not alleviated by α-amanatin or LNA-430 treatments. The zygotic category was defined as: $z(WT\_2hpf; WT\_6hpf) < -z_{th}$ *and* $z(WT\_6hpf; \alpha Am\_6hpf) > z_{th}$ *and* $z(WT\_2hpf; LNA\_6hpf) < -z_{th}$, which corresponds to a decaying number of reads at 6 hpf for the wild type that is alleviated by α-amanatin but not by LNA-430 treatment. The miR-430 category was defined as: $z(WT\_2hpf; WT\_6hpf) < -z_{th}$ *and* $z(WT\_6hpf; \alpha Am\_6hpf) > z_{th}$ *and* $z(WT\_6hpf; LNA\_6hpf) > z_{th}$, which corresponds to a decaying number of reads at 6 hpf for the wild type that is alleviated by both α-amanatin and LNA-430 treatments. Corresponding categories were defined for stabilizing modes of regulation by inverting the inequality signs in the above boolean statements, for example maternally stable segments are defined as: $z(WT\_2hpf; WT\_6hpf) > z_{th}$ *and* $z(WT\_2hpf; \alpha Am\_6hpf) > z_{th}$ *and* $z(WT\_2hpf; LNA\_6hpf) > z_{th}$.

Given this scheme, a particular sequence segment can be assigned to multiple categories. To resolve this discrepancy *P*-values associated with each category were calculated. The three Z-scores used to define a category correspond to three *P*-values combined using Fisher's method. Each segment was assigned to the category with lowest *P*-value. This process was repeated for both sets of replicates of the libraries and final categories were assigned to the sequence segment if the categories for the two replicates were both decaying or were both stabilizing. Benjamini-Hochberg corrected q-values were then calculated for each sequence and the number of segments assigned to various categories shown in Supplemental Figure S1A (q-value<0.05).

Post-transcriptional regulatory element discovery

The identification of distinct stabilizing and destabilizing categories of sequences enables *de novo* discovery of short motifs that are significantly informative of the different modes of regulation. Here, an updated version of FIRE, a computational framework for the discovery of regulatory elements (Elemento et al, 2007), was used to systematically explore the space of linear RNA elements within tens of thousands of sequences that fall into multiple categories simultaneously.

N-fold cross-validation option was introduced in FIRE, whereby the dataset is partitioned into N sets, N-1 parts are used as the training set for motif discovery and 1 set is left aside as a test set to evaluate the results. From the motifs

discovered on the training set, only ones that exhibit significant mutual information in the test set (Z-score>2) and have a pattern of over- and under-representation across categories that is similar to the representation pattern in the training set (correlation>0.5) were kept. The motif discovery was repeated N times with each of the N sets used once as the test set. The resulting motifs were combined and similar motifs were filtered out by means of sequence similarity and conditional information (Elemento et al, 2007; Oikonomou et al, 2014). The introduction of cross-validation avoids data over-fitting and produces a shorter, but more robust list of regulatory elements.

Groups of RNA *trans*-acting factors have binding sites that are highly similar with respect to nucleotide composition and are significantly shorter than the binding sites of transcription factors. In combination with the greedy nature of motif discovery algorithms this often results in degenerate and often uninterpretable motif representations when analyzing large libraries for linear RNA regulatory elements. Here, the average degeneracy allowed for the regular expression of each motif was capped, so that it is optimized for maximal mutual information while sustaining a well-defined sequence profile.

FIRE with zebrafish specific files, so that motif discovery can be applied to identify DNA and RNA regulatory elements for zebrafish, expression data, and updated release of FIRE with related files can be downloaded (see Data access section).

The categorized sequence segments from the RESA libraries are further analyzed for motif discovery. A more permissive list of sequences (P<0.05; 20,111 segments for the transcriptomic library, 3,090 and 2,642 for the targeted total RNA and poly(A) selected library respectively) were split into the categories described above (stabilizing and destabilizing maternal, zygotic and miR-430). Within each category, segments are sorted based on their P and further grouped into equally sized bins (900 and 400 segments per bin for transcriptomic and targeted libraries respectively). *De novo* motif discovery with 3-fold cross-validation on these categories was performed. FIRE revealed a total of 45 RNA motifs for the transcriptomic library and 38 motifs for the targeted library that were significantly informative of the observed regulatory categories. The most significant of these results are shown in Figure 3.

Over representation (yellow) and under representation (blue) patterns are shown for each discovered motif within each category of sequence segments. Mutual information values, Z-scores associated with a randomization-based statistical test, and robustness scores from a three-fold jackknifing test are also shown. The jack-knife test was run 10 times. Each time one-third of the library was randomly removed and the statistical significance of the MI value of the motif was reassessed. For each of the 10 tests, the remaining two-thirds of the library was shuffled 10,000 times and the motif was deemed significant if its MI was greater than all 10,000 MI scores from the randomized sets (Elemento et al, 2007).

Reporter validation

z*c3h18* oligos:

wt-F: `tcgagCTGCCTCCTCCACCTCCTCCTCTAGATCCTCCAGt`

wt-R: `ctagaCTGGAGGATCTAGAGGAGGAGGTGGAGGAGGCAGc`

mut-F: `tcgagCTGCGTCGTCCACGTCGTCCTCTAGATCCTCCAGt`

mut-R: `ctagaCTGGAGGATCTAGAGGACGACGTGGACGACGCAGc`

ccdc22 oligos:

pC, pU F: `cctgactcgactcgaccctacacgacgctcttccgatctGGCTCCATCTGCGTCAGCCCTCCATCCGAAATAATCCTCATCACGTCAGCAACTTTCAGCTCTTTATTGAG`

pC, pU R: `tcactatagttctaggagttcagacgtgtgctcttccgatctATTTGTGTCATTTTGATGCAAAAAACTCTTGTGTTAAAACCCTCAATAAAGAGCTGAAAGTTGCTGAC`

pC-3x, pU F: `cctgactcgactcgaccctacacgacgctcttccgatctGGCTCCATCTGCGTCCTCCgACCTCCGgCCTCCATCCGAAATAATCCTCATCACGTCAGCAACTTTCAGCTCTTTA`

pC-3x, pU R: `tcactatagttctaggagttcagacgtgtgctcttccgatctATTTGTGTCATTTTGATGCAAAAAACTCTTGTGTTAAAACCCTCAATAAAGAGCTGAAAGTTGCTGACGTGAT`

pC, pU-3x F: cctgactcgactcgaccctacacgacgctcttccgatctGGCTCCATCTGCGTCAGCCCTCCATCCGAAATAATCCTCATCACG TCAGCAACTTTCAGCTCTTTATTGAGGGTTT

pC, pU-3x R: tcactatagttctaggagttcagacgtgtgctcttccgatctATTTGTGTCATTTTGATGCAAAAAACTAAAAAACTAAAAAAT GTGTTAAAACCCTCAATAAAGAGCTGAAAGT

Forward and reverse oligos were annealed and cloned into pCS2-GFP (XhoI-XbaI), followed by linearization with NotI and *in vitro* transcription from Sp6 promoter (mMessage mMachine (Ambion)). mRNA was co-injected with mRNA encoding dsRed as described previously.

Hire-PAT reporter assay

Fam116b (ARE motif)

Forward oligo: 5'-CTCGAGTTTTCCCATGATGAAATTAGG-3'

Reverse oligo: 5'-TCTAGACCGATGCCATAGTGTTCTTG-3'

A 107 nucleotide fragment of the 3'-UTR of *fam116* was cloned into pCS2-GFP (XhoI-XbaI), followed by linearization with NotI and *in vitro* transcription from Sp6 promoter (mMessage mMachine (Ambion)). mRNA was co-injected with mRNA encoding dsRed as described previously. Total RNA was extracted from 6 hpf embryos using TRIzol reagent according to manufacturer instructions. One microgram of total RNA was used for the poly(A) tail length assay according to manufacturer instructions (Affymetrix Poly(A) Tail-Length Assay Kit). Poly(G/I) tailed cDNA was amplified using the a *fam116*-specific forward primer (5'-CTGAAGTGCACAGAAGAACCAC-3') and a FAM-labeled universal reverse primer (FAM-GGT AATACGACTCACTATAGCGAGACCCCCCCCCCTT). Fragment analysis was performed in a 3730XL 96-Capillary Genetic Analyzer and output was analyzed using GeneMarker software (Softgenetics).

Interactome capture

Interactome capture protocol

Wild type zebrafish embryos were irradiated at 4 hpf with UV at 254 nm for 4 minutes in a Spectrolinker™ XL-1000 UV crosslinker. During this time, the embryos were covered with water in 6-well plate on ice and at a distance from the UV source of ~9 cm. Immediately after cross-linking, the embryos were collected in 1.5 ml tubes and flash frozen in liquid nitrogen. α-amanitin-injected embryos were processed in the same way. The –UV control group of embryos were not UV-irradiated prior to collection. In total, 3 replicates of +UV wild-type embryos, +UV α-amanitin-injected embryos, and –UV embryos were processed. Each replicate contains around 4,000 embryos. Interactome capture was conducted as described in Castello et al, 2012. Embryos were lysed in Lysis/Binding Buffer (20 mM Tris-HCl (pH 7.5), 500 mM LiCl, 0.5% LiDS, 1 mM EDTA, 5 mM DTT) and the poly(A)$^+$ mRNAs with the associated cross-linked proteins were pulled-down with oligo d(T)$_{25}$ Magnetic Beads (NEB). Next, the pellets of oligo d(T)$_{25}$ beads were sequentially washed with buffers with decreasing concentration of LiDS and LiCl (Wash Buffer I (20 mM Tris-HCl (pH 7.5), 500 mM LiCl, 0.1% LiDS, 1 mM EDTA, 5 mM DTT); Wash Buffer II (20 mM Tris-HCl (pH 7.5), 500 mM LiCl, 1 mM EDTA, 5 mM DTT); Low Salt Buffer (20 mM Tris-HCl (pH 7.5), 200 mM LiCl, 1 mM EDTA; Elution Buffer (20 mM Tris-HCl (pH 7.5), 1 mM EDTA)). After washes, the pellets of beads were frozen, awaiting further processing. Crosslinked material was rinsed three times with wash buffer (150 mM NaCl, 50 mM Tris, pH 7.4) and then denaturated for 30 minutes with urea (8 M) in 0.1 M Tris, pH 7.4, 1 mM DTT before alkylation and on-bead pre-digestion with Endoproteinase LysC (Wako Chemicals USA, Inc.). After incubation for 3 hours, samples were diluted 4-fold with ammonium bicarbonate (25 mM) and further digested with trypsin (Promega) overnight. Digestions were stopped by addition of trifluoroacetic acid (TFA, 1 µL), and the resulting peptides were loaded and desalted on C18 Stage Tips.

LC-MS/MS analysis

Peptides were eluted from C18 Stage Tips with 60 µL of elution buffer (80% acetonitrile and 0.1% formic acid), and samples were dried down to 5 µL in a vacuum centrifuge. Peptides were then subjected to reversed phase chromatography on an Easy nLC 1000 system (Thermo Fisher Scientific) using a 50-cm column (New Objective) with an inner diameter of 75 µm, packed in-house with 1.9 µm C18 resin (Dr. Maisch GmbH). Peptides were eluted with an acetonitrile gradient (5–30% for 95 min at a constant flow rate of 250 nL/min) and directly electrosprayed into a mass spectrometer (Q Exactive; Thermo Fisher Scientific). Mass spectra were acquired on the spectrometer in a data-dependent mode to automatically switch between full scan MS and up to 10 data-dependent MS/MS scans. The maximum injection time for full scans was 20 ms, with a target value of 3,000,000 at a resolution of 70,000 at m/z = 200. The ten most intense multiple charged ions ($z \geq 2$) from the survey scan were selected with an isolation width of 3Th and fragmented with higher energy collision dissociation (HCD) with normalized collision energies of 25. Target values for MS/MS were set to 100,000 with a maximum injection time of 160 ms at a resolution of 17,500 at m/z = 200. To avoid repetitive sequencing, the dynamic exclusion of sequenced peptides was set to 35 s.

MS data analysis

MS and MS/MS spectra were analyzed using MaxQuant (version 1.3.8.1), utilizing its integrated ANDROMEDA search algorithms. Scoring of peptides for identification was carried out with an initial allowed mass deviation of the precursor ion of up to 6 ppm for the search for peptides with a minimum length of six amino acids. The allowed fragment mass deviation was 20 ppm. The false discovery rate (FDR) was set to 0.01 for proteins and peptides. Peak lists were searched against a local database for human proteome. Maximum missed cleavages were set to 2. The search included carbamidomethylation of cysteines as a fixed modification and methionine oxidation and N-terminal acetylation as variable modifications. The final list of proteins was curated to remove duplicates and retain only proteins with at least two peptides and at least one of them unique.

Whole-embryo lysate proteomics

4 hpf embryos were collected in triplicate in deyolking buffer (55 mM NaCl, 1.8 mM KCl, 1.25 mM $NaHCO_3$) and pipetted up and down to disrupt the embryos. Cells were pelleted by centrifuging 30 seconds at 300 x g. The pellet of cells was then washed twice with wash buffer (110 mM NaCl, 3.5 mM KCl, 2.7 mM $CaCl_2$, 10 mM Tris-HCl pH8.5), followed by flash freezing in liquid nitrogen. Samples were further processed and analyzed by mass spectrometry as described earlier for the interactome capture.

Mass spectrometry data

All raw mass spectrometry data files from this study are available at ProteomeXchange repositories (see Data access section).

Gene Ontology enrichment

Enrichment in the molecular function category was tested by comparing the proteins found in the interactome against all the proteins identified in the whole-embryo lysate of wild-type and α-manitin-injected embryos at 4 hpf. Gene Ontology annotation for these proteins was downloaded from Ensembl using BioMart. For each GO term, the number of proteins identified in the interactome with this GO term was normalized by the total number of proteins in all GO terms. The same calculation was conducted with the proteins in the input. The odds ratio for a given GO term is the result of dividing the value obtained from interactome proteins by value obtain from the input. Only GO terms present in both samples and with a $\log_2$ odds ratio >2.5 or < -2.5 were retained for further analysis.

RBP binding: Cross-linking Immunoprecipitation

RBP cloning

To clone each selected RBP into pCS2+, specific primers were designed using Primer3 (Rozen & Skaletsky, 1998) that also incorporate the corresponding restriction enzyme. The 5' primers also contain the sequence encoding the FLAG tag in frame with the downstream RBP. Next, the cDNA of each RBP was retrotranscribed from total RNA isolated from zebrafish embryos at mixed stages using SuperScript™ III Reverse Transcriptase (Invitrogen) and amplified with Phusion® High-Fidelity DNA Polymerase (NEB). Next, each PCR amplicon was cut with the appropriate set of restriction enzymes and ligated to a linearized pCS2+ vector with T4 DNA ligase (NEB). Finally, the resulting clones were linearized and *in vitro* transcribed to capped mRNA using the mMESSAGE mMACHINE™ SP6 Transcription Kit (Ambion). Oligos are listed in Supplemental Table 5.

iCLIP library cloning

The iCLIP protocol described in Huppertz et al, 2014 has been modified and adapted to zebrafish.

Sample collection

For each given RBP, zebrafish embryos at 1-cell were injected with 1 nl of the corresponding capped mRNA at 0.1 µg/µL. Sphere stage embryos (~4 hpf) were cross-linked with UV at 254 nm for 4 minutes. Plates of embryos were deposited over a tray of ice during cross-linking and immediately collected into 1.5 ml eppendorf tubes, flash frozen in liquid nitrogen, and stored at -80°C until processing. Around 200 embryos were collected per replicate.

Lysis and pull-down

Embryo samples were lysed with 1 ml of lysis buffer (50 mM Tris–HCl, pH 7.4, 100 mM NaCl, 1% Igepal CA-630, 0.5% sodium deoxycholate, 0.1% SDS, 1/100 volume of Protease Inhibitor Cocktail Set III (Calbiochem), 1/1000 volume of SUPERase•In™ RNase Inhibitor (Ambion)), and vortexed. 4 µL of Turbo DNAse (Ambion) were added to each tube, followed by 10 µL of RNAse I (Invitrogen) diluted to 1:200 in PBS. Samples were incubated in the thermomixer for 3 minutes at 37ºC and 1,100 rpm. Immediately after, 5 µL of RNase Inhibitor were added, and the samples were incubated on ice for at least 3 minutes. Next, samples were centrifuged at 4ºC for 10 minutes at 14,000 rpm. In the meantime, ANTI-FLAG® M2 Magnetic Beads (catalogue #M8823, Millipore-Sigma) were prepared at a ratio of 20 µL of beads per 200 embryos. Beads were washed 3 times with 800 µL of lysis buffer, avoiding bubbles. Next, the supernatant was transferred to the pellet of antibody with beads and incubated for 2 hours at 4ºC in an orbital rocker. For Ddx6 iCLIP, Ddx6 was pulled-down using a specific antibody (Abcam, ab40684) and Dynabeads™ Protein G (Invitrogen) in three replicates of ~1,000 embryos each. After, pull-down, beads were washed twice with High Salt buffer (50 mM Tris–HCl, pH 7.4, 1 M NaCl, 1 mM EDTA, 1% Igepal CA-630, 0.1% SDS, 0.5% sodium deoxycholate) and 3 times with PNK buffer (20 mM Tris–HCl, pH 7.4, 10 mM $MgCl_2$, 0.2% Tween-20), 800 µL each time.

5' labeling of RNA fragments

4 µL of the labeling reaction (0.2 µL PNK, 0.4 µL gamma-^{32}P-ATP, 0.4 µL 10× PNK buffer, 2.8 µL H2O, 0.2 µL RNase inhibitor) were added to the pellet of washed beads prior incubation of 5 minutes at 37ºC and 1,100 rpm in the thermomixer. The pellets were washed one time with 500 µL PNK buffer.

NUPAGE electrophoresis and membrane transfer

30 µL of NUPAGE sample loading buffer with reducing agent were added to the beads and incubated 10 minutes at 80 ºC, shaking at 1,110 rpm. Next, the supernatant was loaded into a precast NUPAGE 4-12% gradient gel in the Bolt system (Invitrogen). The electophoresis ran for 50 minutes at 180 V. After the run, the riboprotein complexes from the gel were transferred to a nitrocellulose membrane using the iBlot system (Invitrogen). After the transfer, the membrane was exposed to a storage phosphorimager plate for 3 hours at -20 ºC.

Proteinase K digestion

The membrane sections containing the ribo-protein complexes according to the scan of the storage phosphorimager screen were cut and added to a 2 ml tube with 200 µL of Proteinase K mix (Tris-HCl 100 mM, NaCl 50 mM, EDTA 10 mM, Proteinase K). The samples were incubated 20 minutes at 37 ºC and 1,100 rpm and an additional 200 µL of Proteinase K buffer with 7 M urea were added, followed by an additional incubation of 20 minutes at 37 ºC and 1,100 rpm. The RNA fragments were extracted from the membrane with 400 µL of phenol-chloroform and vigorous shaking for 10 minutes at room temperature. The liquid was transferred to a new tube and centrifuged 10 minutes at maximum 14,000 rpm. The aqueous phase was collected and the RNA was precipitated with 40 µL of sodium acetate 3 M, 0.7 µL glycogen, and 1 ml of 96% ethanol.

3'-end repair (dephosporylation)

The repair reaction mix (2 µL of 10× PNK pH 6.5 buffer, 0.5 µL polynucleotide kinase (NEB), 0.5 µL RNase Inhibitor, 15 µL water) was added to the purified pellet of RNA and incubated for 20 minutes at 37ºC, followed by RNA precipitation.

3' pre-adenylated adaptor ligation

The RNA was resuspended with of 10 µL water plus the ligation reaction mix (20 pmoles pre-adenylated DNA adapter, 2 µL 10× T4 RNA Ligase 2 truncated K227Q reaction buffer (NEB), 1 µL 200 U T4 RNA Ligase 2 truncated K227Q (NEB), 1 µL of RNaseOUT (Invitrogen), 3 µL of 50% PEG8000 and 4 µL of water). Ligation reactions were incubated at 16 °C overnight. Each sample within a replicate of the same RBP was ligated with an adaptor with different barcodes to allow sample pooling in the subsequent steps. The sequence of the 3' adaptor used was: rApp-NNcacaACTGTAGGCACCATCAAT-ddC, where lowercase represents one of the custom barcodes.

Size fractionation

To remove the unligated adaptor, all replicate samples were combined and loaded in a 15% denaturing polyacrylamide-urea gel. After running, the gel was exposed in a storage phosporimager screen and kept at -20 °C. Next, gel slices containing the ligated RNA were cut and RNA was extracted and precipitated.

Retrotranscription

The purified RNA was resuspended in 8 µL of water with 1 µL RT iCLIP primer (0.5 pmol/µL of the oligo -(Phos) NNNNAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGC-(SpC18)-CACTCA-(SpC18)-TTCAGACGTGTGCTCTTCCGATCTATTGATGGTGCCTACAG), 1 µL dNTP mix (10 mM) and incubated for 5 minutes at 70 ºC. Next the retrotranscription reaction mix was added: 4 µL H2O, 4 µL 5× First Strand Buffer, 1 µL 0.1 M DTT, 0.5 µL RNaseOUT, 0.5 µL SuperScript III (Invitrogen)). Reactions were incubated 5 minutes at 25 ºC, 20 minutes at 42 ºC, 40 minutes at 50 ºC, 5 minutes 80 ºC. After retrotranscription, the samples were loaded in a 15% denaturing polyacrilamide-urea gel and size selected. Next, the cDNA was extracted from the gel and precipitated.

Recircularization and PCR amplification

cDNA was resuspended in 15 µL of water, 2 µL CircLigase Reaction buffer 10×, 1 µL ATP, 1 µL $MnCl_2$ 50 Mm, and 1 µL of Circligase ssDNA ligase (Epicentre). Reactions were incubated 2 hours at 60 ºC. After recircularization, the libraries were amplified using primers compatible with Illumina adaptors (Forward oligo `AATGATACGGCGACCACCGAGA TCTACAC`, reverse oligo `CAAGCAGAAGACGGCATACGAGATcgtgatGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT`, where lower case represents one of the barcodes) and PCR products were purified in a 10% non-denaturing polyacrylamide gel. Libraries were sequenced at the Yale Center for Genome Analysis using HiSeq 2500 platform. Raw reads are publicly accessible in the Sequence Read Archive (see Data access section).

Label-transfer assay

Zebrafish embryos were injected at one-cell stage with *in vitro* transcribed mRNA encoding for the indicated RBPs. Embryos were collected at 4 hpf and processed following the same procedure described for iCLIP samples, including UV-cross linking, RNA digestion, FLAG pull-down, radiolabeling, protein gel and transfer. The only modification in this case was that a higher concentration of RNAseI (Invitrogen) was used (1:500 final dilution). After protein transfer, membranes were exposed to a phosphorimager storage screen to detect the radioactive signal from the RNA. After exposure, Western blot was performed on the membranes to detect FLAG-tagged RBPs. ANTI-FLAG® antibody produced in rabbit (catalogue #F7425, Millipore-Sigma) was used as a primary antibody in the Western blot.

Khsrp iCLIP with endogenous antibody

iCLIP experiment as described in Huppertz et al, 2014 was followed with minor changes. Embryos at sphere stage (4 hpf) were collected and irradiated with 254 nm UV light to induce crosslinking (embryos were snap frozen and stored in batches to yield a total of 1,000 embryos/condition). Frozen embryos were then thawed and homogenized on ice in lysis buffer. An affinity-purified rabbit polyclonal antibody raised against zebrafish KHSRP (generated by YenZym Antibodies, LLC) was used to isolate RNA-protein complexes. Briefly, 200 µL Protein G Dynabeads were added to 50 µg antibody in lysis buffer. Beads were incubated for an hour, washed three times with lysis buffer, and added to the lysates. As a control, a parallel experiment was performed without the addition of antibody. Subsequent steps in the iCLIP protocol were performed as described above. Twenty cycles of amplification of cDNA were used for final library construction. Barcoded PCR-amplified libraries were size-selected on a 6% TBE gel and combined for Illumina sequencing. Both the KHSRP pull-down and the no-antibody control were performed in triplicate.

iCLIP analysis: Binding profiles

Demultiplexing samples

After sequencing the iCLIP libairies, raw reads were composed as NNNN-insert-NN-barcode(4-mer)-adapter where (i) the 6N (NNNN+NN) are random nucleotides composing the Unique Molecular Identifier (UMI), (ii) the 4-mer barcode is the in-house barcode used to demultiplex samples and (iii) adapter is the 3' Illumina adapter. To demultiplex samples barcoded with in-house barcode, reads were first trimmed of Illumina adapter by aligning its sequence, requiring 100% match of the first five base pairs and a minimum global alignment score of 80 (Matches: 5, Mismatches: −4, Gap opening: −7, Gap extension: −7, Cost-free ends gaps). Only trimmed reads were then demultiplexed, allowing a single mismatch with the expected barcodes. The 5' 4N and 3' 2N composing the UMI were then clipped and added to the read name for downstream analysis. Reads with UMI containing unidentified nucleotide(s) were discarded.

Mapping

Demultiplexed reads were mapped to the zebrafish Zv9 genome using STAR (Dobin et al, 2013) version 2.4.2a with the following non-default parameters: *--alignEndsType EndToEnd --outFilterMultimapNmax 100 --seedSearchStartLmax 15 --sjdbScore 2*. Genomic sequence indices for STAR were built including exon-junction coordinates from Ensembl r78 (Aken et al, 2017).

UMI

To filter unique molecules, reads were considered unique when they had a unique UMI, as well as unique start and end positions.

Binding profiles

Using Ensembl 78, per chromosome and per transcript annotation binding profiles were created by summing the first position (5' end) of reads.

RBP binding metagene profiles

To summarize RBP binding within protein-coding transcripts, all UTRs and CDSs iCLIP profiles were split into 50 bins (bin length is different for different transcript lengths). Transcripts with less than 50 reads total or less than 10 reads within a UTR or CDS were discarded. Bins were then summed from the most 5' to the 3' end for all transcripts (first bin of first transcript with first bin of second transcript, etc.) to obtain metagene profiles. Hierarchical clustering with complete-linkage method using an Euclidean distance matrix was applied to group similar binding profiles together. For clustering, biological replicates were merged together by summing replicates.

RBP motif and overlap heatmap

To define binding preference of RBP, bound positions were first defined within transcripts for nucleotide with at least 2 unique reads. These bound positions extended 5 nucleotides upstream and 10 nucleotides downstream defined bound windows. In each window, 6-mers were counted. Transcriptome-wide 6-mer frequencies were then normalized per RBP. To take into account the iCLIP experimental background, these frequencies were normalized by the average 6-mer frequencies measured for the 6 controls. For the khsrp iCLIP experiment using an endogenous antibody, a minimum of 10 unique reads was required to define bound windows to take into account the higher sequencing depth of these samples, while the cutoff for the controls remained at 2.

To build the logo representation, the top 10 most frequent normalized 6-mer were aligned with MAFFT (Katoh et al, 2013) using the parameters *--reorder --lop -10 --lexp -10 --localpair --genafpair --maxiterate 1000*. The multiple sequence alignment positions with more than 50% gaps were trimmed on both ends. Nucleotide frequencies were calculated using the final trimmed alignment, and represented as a logo.

To cluster RBPs on their motif, the overlap of the top 20 normalized 6-mer frequencies with iCLIP controls were used to perform hierarchical clustering similarly to metagene profiles. Biological replicates were merged together by summing counts of all replicates.

Motif-centered metaplots

To simultaneously analyze binding and regulatory activity, sequences matching RBP motifs were searched within the transcriptome. Windows were defined 100 nt upstream and downstream of the motif occurrence. Windows with minimum

RESA coverage below 0.05 CPM in one of the 2 conditions used (e.g. 6 hpf and 2 hpf) were discarded. Finally, iCLIP (5'end) and RESA (coverage ratio) windows were averaged.

To define RBP binding context, all 3-mers were counted within 20 nt windows upstream and downstream of the motif (not including it). The 10% most regulated (highest RESA) motif occurrences were used to define the 3-mer frequencies of favorable context, while the 10% least regulated defined unfavorable context. Strong sites had a favorable context and no unfavorable context, while weak or control sites had neither. For each set of sites, the corresponding iCLIP signal was averaged before plotting.

Modeling the effect of sequence on mRNA stability

Computational model training

A RESA profile integrates the stability values across multiple sequence reporters. RESA stability values are calculated based on the weighted average of all inserts intersecting at center of the window. To analyze the RESA profiles, a sliding window approach was applied: a window of 100 nucleotides (nt) was used with a sliding step-size of 10 nt to scan the 3'-UTR profile of each reporter. For each window, the frequency of all *k*-mers was computed (1-8 nt). The RESA stability value of the center of each window (the 50^{th} nt) was computed. This value represents the effect of that single position along with its sequence context (preceding and following 50 nt within that window). This approach maintained the resolution of RESA while integrating the effects of the flanking sequences. In addition, it also reduced training data redundancy as opposed to considering every single position (sliding window with step-size of 1nt or summing reads over entire window), both of these latter approaches introduce large redundancy in the training data and eventually bias model training and lead to overfitting.

After scanning all the profiles, this window-based data was used to train a Random Forest (RF) model with 500 trees (R 3.4.2, package of "*randomForest*" 4.6-12). The objective of the training process was to build a set of decision trees that are able to capture the association between *k*-mer frequencies and their correlation with RESA stability values. The trained RF model was used to predict the stability of new transcripts in a similar window scanning procedure: for each new transcript, a scanning window of 100 nt and slide size 10 nt was applied, where the stability of each window was predicted by the trained RF model. Next, this prediction process was repeated for each sliding window, providing a prediction value for each 10nt in the 3'-UTR. Finally, predicted values were connected using curve fitting approach which constituted the predicted RESA profile for this transcript.

Feature selection

To make the learning process more efficient, a pre-processing, unbiased, and fast *k*-mer filtering step was required i) to reduce the number of features and focus on those with potential effects on RESA fold-change, ii) to increase the training speed of the RF model, iii) to reduce potential bias towards irrelevant, uncorrelated features with respect to RESA stability score, and iv) to decrease the model chance of overfitting by only considering correlated features to RESA fold-change. First, to filter rare (never occurring) *k*-mers, all *k*-mers with non-zero occurrence frequency were retained (31,548 *k*-mers out of 87,380 *k*-mers). Second, to filter irrelevant and not correlated *k*-mers, *k*-mers with absolute correlation (Spearman's correlation coefficient) greater than or equal to 0.1 (with respect to RESA stability values) were selected as potentially correlated (387 *k*-mers). Finally, the importance of the retained features was assessed using feature importance function in trained random forest model (R package of "randomForest" 4.6-12), which uses out-of-bag error to assess the importance of each feature. Feature selection through permutation (out-of-bag) suffers from overestimating correlated features. To reduce this bias, the feature importance process was computed applying cross-validation. The final importance of each feature was the average importance from the cross-validation process. For each cross-validation step, the RF model was used to calculate Mean-Decrease-Accuracy (measured by Mean Square Error "MSE") for measuring feature importance. This method is more reliable than the default permutation importance through Mean-Decrease-in-impurity. Only features

whose importance value was greater than the average importance values for all the features were retained. This last feature importance step highlighted that 57 motifs had the most significant contribution out of 387 *k*-mers that were used to train the random forest model. For these 57 motifs, it is likely several of them are related and interdependent.

5-fold cross-validation

To assess how the trained model would generalize to an independent dataset, a 5-fold cross-validation was used to evaluate the predictive performance of the model in practice. All the transcripts were randomly partitioned into 5 equally sized subsamples. Of these 5 subsamples, a single subsample was retained as the validation set of transcripts for testing the RF model, while the remaining 4 subsamples were used as training data for building the RF model. This process of training and validation was then repeated 5 times, with each of the 5 subsamples used exactly once as the validation data.

Model performance assessment through cross-validation was performed for each library (RESA targeted and poly(A) selected), separately. For each of these two libraries, a separate RF model was trained and validated for each different treatment conditions (WT, LNA-430, and α-amanitin), individually. The score per transcript was calculated as the average of all the sliding windows predicted scores across the 3'-UTR of that transcript. Model assessment was performed using the "cor.test" function in R to calculate the Pearson correlation between the predicted fold-change and the RESA measured fold-change. Asymptotic *P* was computed using the same correlation function for each of the computed correlation values.

Selected motifs analysis

For each selected motif, the differential stability between windows that did and did not contain that motif were compared. The mean RESA values between these two groups of windows (motif present vs. absent) was used as a proxy to reflect the average differential effect of each individual motif on stability. All the windows were grouped into three categories according to their RESA stability profile; the top 25% (stable), middle 50%, and bottom 25% (unstable). For each selected motif, the motif mean frequency of each group (top 25% stable and bottom 25% unstable) was compared against the middle group, which highlighted the preferential enrichment of each motif in the stabilizing/destabilizing groups compared to the middle control group.

Model performance on endogenous transcripts

To evaluate the random forest model performance in predicting stability of endogenous mRNAs, the model trained on average stability profiles from the RESA targeted library was used. A similar sliding window approach was applied to the 3'-UTR of endogenous genes. The predicted stability for each mRNA was calculated as the average of all the sliding windows predicted stability scores across the 3'-UTR. The predicted fold-change of each endogenous mRNA was then compared against the fold-changes measured by mRNA-seq.

Model training on trip10 locus

To avoid any bias in model training, the RF model was trained on all the RESA targeted transcripts after excluding the *trip10* transcript. After training, the model performance was assessed by comparing the experimental profile of the *trip10* transcript according to RESA against the predicted profile according to the trained RF model.

Model validation on trip10 locus

Reporter and library preparation

WT: `NNNNAAATAATATAATTTATTGAGTAAATaagcgcttGTATATTAAATAAACATGTATGTAAGANNNN`

WT_F: ccctacacgacgctcttccgatctNNNNAAATAATATAATTTATTGAGTAAATaagcgcttGTATA

WT_R: gAGTTCAGacgtgtgctcttccgatctNNNNTCTTACATACATGTTTATTTAATATACaagcgcttATTTACTCAAT

Mut1: NNNNAAATAATATAATTTATTGAGTAAATaaAAgGttGTATATTAAATAAACATGTATGTAAGANNNN

Mut1_F: ccctacacgacgctcttccgatctNNNNAAATAATATAATTTATTGAGTAAATaaAAgGttGTATA

Mut1_R: gAGTTCAGacgtgtgctcttccgatctNNNNTCTTACATACATGTTTATTTAATATACaaCcTTttATTTACTCAAT

Mut2: NNNNAcATAAcATcActcATcGAGTAAATaagcgcttGTATAccAcATAcACATGcATGTAAGANNNN

Mut2_F: ccctacacgacgctcttccgatctNNNNAcATAAcATcActcATcGAGTAAATaagcgcttGTATA

Mut2_R: gAGTTCAGacgtgtgctcttccgatctNNNNTCTTACATgCATGTgTATgTggTATACaagcgcttATTTACTCgAT

Mut3: NNNNAcATAAcATcActcATcGAGTAAATaaAAgGttGTATAccAcATAcACATGcATGTAAGANNNN

Mut3_F: ccctacacgacgctcttccgatctNNNNAcATAAcATcActcATcGAGTAAATaaAAgGttGTATA

Mut3_R: gAGTTCAGacgtgtgctcttccgatctNNNNTCTTACATgCATGTgTATgTggTATACaaCcTTttATTTACTCgAT

Mut4: NNNNAAgcAAgcTAgcTgcTTGAGgcAAgcaAAgGttGgcTATgcAAgcAACATGgcTGgcAGANNNN

Mut4_F: ccctacacgacgctcttccgatctNNNNAAgcAAgcTAgcTgcTTGAGgcAAgcaAAgGttGgcTA

Mut4_R: gAGTTCAGacgtgtgctcttccgatctNNNNTCTgcCAgcCATGTTgcTTgcATAgcCaaCcTTtgcTTgcCTCAAg

sticky ends_F 10353: CCTGACTCGACTCGAGCCCTACACGACGCTCTTC

sticky ends_R 10354: TCACTATAGTTCTAGAGAGTTCAGACGTGTGCTC

Hi-Fidelity PCR (using Phusion NEB-F-530S) was run with Forward and Reverse oligos. The PCR product was run on gel, then extracted and purified using Qiagen Gel Extraction Kit. A second PCR was run to add the "sticky ends" to the previous PCR products. The final product was cloned into pCS2-GFP vector (cut with XhoI-XbaI) using In-Fusion® HD Cloning Plus-Clonetech- 639642). Each construct was sequenced to confirm cloning of the expected sequence. Each of the final *trip10* plasmids was linearized with NotI and used as a template for *in vitro* transcription using (SP6 mMessage Machine-Ambion AM1340).

To test the effect of each *trip10* sequences on gene expression, a mix of 100 pg *trip10* and 70 pg dsRed mRNA as a control was injected for each *trip10* sequence into one-cell stage embryos as described previously. The injected embryos were incubated for 24h at 28 °C, and the GFP expression was compared to the control dsRed expression using fluorescence microscopy.

To measure the effect of each *trip10* sequence on mRNA stability, a second set of embryos was injected at the one-cell stage with an mRNA cocktail containing 5 pg of each of the GFP *trip10* sequences. Then, 30 embryos were collected at 64-cell and shield stages (in triplicate). Frozen embryos were lysed with 1 mL of TRIzol and total RNA was extracted following the manufacturer's protocol. Total RNA was dissolved in 11 µL of water and reverse transcribed using SuperScript III with 100 nM of reverse transcription primer (5'- CTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNN-barcode-CATGTCTGGATCTACGTAATACG-3') in 20 µL reactions following the manufacturer's protocol. 64-cell and shield samples were reverse transcribed with primers containing the 5'-TAGCTT-3' and 5'-ACTTGA-3' homemade barcodes, respectively. After reverse transcription, 64-cell and shield samples were pooled per-replicate and cDNA was purified using 108 µL of AMPure XP beads (Beckman Coulter Inc.), following the manufacturer's protocol, and then dissolved in 20 µL of water. 5'- and 3'-Illumina adapters were added by PCR with the following primers (forward 5'-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCC-3' and reverse: 5'- CAAGCAGAAGACGGCATACGAGAT-illumina_barcode-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-3'), and amplified using the lowest amount of PCR circles necessary to obtain enough material for sequencing. Replicate 1, 2, and 3 were amplified using the 5'-CCACTC-3', 5'-ATCAGT-3', and 5'-AGGAAT-3' Illumina barcodes, respectively. PCR products were precipitated using ethanol and purified on an 8% non-denaturing polyacrylamide gel. Libraries were sequenced at the Yale Center for Genome Analysis using HiSeq 2500 platform. Raw reads are publicly accessible in the Sequence Read Archive (see Data access section).

Demultiplexing samples

The sequenced pairs had the following structure: R1-4N-insert-4N-plasmid-barcode(6-mer)-10N-R2. The second read R2 was used to demultiplex the samples while R1 was used to measure the WT and mutant sequences. R2 was composed as `NNNNNNNNNN`-barcode(6-mer)-plasmid where (i) the 10N are random nucleotides composing the Unique Molecular Identifier (UMI), (ii) the 6-mer barcode is the in-house barcode used to demultiplex samples. To demultiplex samples barcoded with in-house barcode, reads were first clipped of 10N on their 5' end composing the UMI and added to the read name for downstream analysis. Reads with UMI containing unidentified nucleotide(s) were discarded. Reads were then demultiplexed, allowing a single mismatch with the expected barcodes.

Mapping

The first read of demultiplexed read-pair was composed as NNNN-insert-NNNN-plasmid where the 4N are random nucleotides. First reads were (i) clipped of 4N on their 5' end, (ii) trimmed on their 3' end of the sequence `AGATCGGAAGAGC ACACGTCTGAACTCTCTAGA` by alignment, requiring 100% match of the first five base pairs and a minimum global alignment score of 60 (Matches: 5, Mismatches: −4, Gap opening: −7, Gap extension: −7, Cost-free ends gaps), (iii) clipped of 4N on their 3' end, then (iv) mapped to the *trip10* WT and mutant sequences using Bowtie 2 2.3.4.1 (Langmead et al, 2012) with the following non-default parameter: *--score-min L,-0.6,-0.1*. Because WT and mutant sequences were similar, this parameter limited the number of allowed mismatches to get unique mapping of first reads on WT and mutant sequences.

UMI counts

To filter unique molecules, reads were considered unique when they had a unique UMI. Reads with unique UMI were counted for WT and mutant sequences. Counts were then normalized to CPM (Count Per Million) using total counts of all WT and mutant sequences per sample.

References

Aken BL, Achuthan P, Akanni W, Amode MR, Bernsdorff F, Bhai J, Billis K, Carvalho-Silva D, Cummins C et al. (2017). Ensembl 2017. *Nucleic Acids Res*, *45*(D1), D635–D642. doi:10.1093/nar/gkw1104

Bazzini AA, Lee MT, & Giraldez AJ (2012). Ribosome profiling shows that miR-430 reduces translation before causing mRNA decay in zebrafish. *Science*, *336*(6078), 233–7. doi:10.1126/science.1215704

Castello A, Fischer B, Eichelbaum K, Horos R, Beckmann BM, Strein C, Davey NE, Humphreys DT, Preiss T et al. (2012). Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell*, *149*(6), 1393–406. doi:10.1016/j.cell.2012.04.031

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*(1), 15–21. doi:10.1093/bioinformatics/bts635

Elemento O, Slonim N, & Tavazoie S (2007). A universal framework for regulatory element discovery across all genomes and data types. *Mol Cell*, *28*(2), 337–50. doi:10.1016/j.molcel.2007.09.027

Huppertz I, Attig J, D'Ambrogio A, Easton LE, Sibley CR, Sugimoto Y, Tajnik M, König J, & Ule J (2014). ICLIP: protein-RNA interactions at nucleotide resolution. *Methods*, *65*(3), 274–87. doi:10.1016/j.ymeth.2013.10.011

Jiang H, Lei R, Ding SW, & Zhu S (2014). Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinformatics*, *15*, 182. doi:10.1186/1471-2105-15-182

Katoh K & Standley DM (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*, *30*(4), 772–80. doi:10.1093/molbev/mst010

Langmead B & Salzberg SL (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods*, *9*(4), 357–9. doi:10.1038/nmeth.1923

Love MI, Huber W, & Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*, *15*(12), 550. doi:10.1186/s13059-014-0550-8

Moreno-Mateos MA, Vejnar CE, Beaudoin JD, Fernandez JP, Mis EK, Khokha MK, & Giraldez AJ (2015). CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. *Nat Methods*, *12*(10), 982–8. doi:10.1038/nmeth.3543

Oikonomou P, Goodarzi H, & Tavazoie S (2014). Systematic identification of regulatory elements in conserved 3' UTRs of human transcripts. *Cell Rep*, *7*(1), 281–92. doi:10.1016/j.celrep.2014.03.001

Thisse C & Thisse B (2008). High-resolution in situ hybridization to whole-mount zebrafish embryos. *Nat Protoc*, *3*(1), 59–69. doi:10.1038/nprot.2007.514

Yartseva V, Takacs CM, Vejnar CE, Lee MT, & Giraldez AJ (2017). RESA identifies mRNA-regulatory sequences at high resolution. *Nat Methods*, *14*(2), 201–207. doi:10.1038/nmeth.4121