Supplemental Methods

Mass spectrometry sample preparation

Eluted 1x LiDS protein fraction was heated to 80°C and separated on a 4%-12% NuPage NOVEX gradient gel in MOPS buffer (Thermo). In-gel digestion was performed essentially as previously described (Rappsilber et al. 2007). Peptides were desalted on stage tips and stored at room temperature until analysis.

Mass spectrometry measurement

Peptides were separated on a C18-reversed phase column (15-20 cm, 75 μm inner diameter, packed in-house with ReproSil-Pur C18-AQ 1.9 μm resin (Dr. Maisch)) by nanoflow liquid chromatography on an EASY-nLC series uHPLC system (Thermo) coupled to a Q Exactive series mass spectrometer (Thermo). For elution, a gradient of 2% to 60% acetonitrile in 0.5% formic acid at a flow of 200 nl/min was applied within an overall 105 min (Q Exactive Plus) or 240 min (Q Exactive) instrument method. The instruments were operated in a top10 data-dependent acquisition with a full scan (QE: 70k, 300-1750 m/z; QEP: 70k, 300-1650 m/z) and up to 10 HCD fragmentations.

MS analysis

Raw files were processed with MaxQuant v.1.5.0.0 (Cox and Mann 2008) and searched with the incorporated Andromeda search engine against the supplied Uniprot danio rerio fasta file (39.559 entries). Carbamidomethylation was set as fixed modification while methionine oxidation and protein N-acetylation were considered as variable modifications. The search was performed with an initial mass tolerance of 7 ppm mass accuracy for the precursor and 20 ppm for the MS/MS spectra in the HCD fragmentation mode. Standard settings were applied except match between runs and the LFQ quantitation was activated. Search results were filtered with MaxQuant at false discovery rate of 0.01 on protein and peptide level.

Data Analysis

Prior to statistical analysis, known contaminants, protein groups identified by site and reverse hits were removed. The dataset was filtered for at least 1 unique peptide. To calculate enrichment, LFQ values were $\log_2$ transformed. Measured LFQ values in a least two replicates were required. Missing LFQ values were imputed with a normal distribution shifted by -4 and sharpened with a standard deviation factor of 0.3. For the volcano plots, s_0 and c value were adapted to reach FDR=0.01.

Gene ontology and protein domain analyses of the zebrafish MZT mRNA interactome

Protein IDs were converted into respective Ensembl gene IDs. If the protein group had two or more associated genes, one representative Ensembl gene ID was chosen for further analysis to circumvent data overestimation. GO terms were downloaded from Zv10 (v82) Ensembl annotation using Biomart. All protein-coding genes with available annotations served as a gene universe for GO term enrichment analyses. PFAM domains were downloaded from Ensembl using Biomart and compared to all protein-coding genes in zebrafish genome (Zv10, v82). P values were derived using Fisher exact two-sided test and corrected using Benjamini-Hochberg method. For comparison analyses with mammalian mRNA interactomes, orthologs and their respective GO terms were retrieved from Ensembl (h38, m38) using Biomart.

Splicing score per intron calculation

In order to derive accurate estimate of splicing values per intron from the zebrafish mRNA-seq dataset (Pauli et al. 2012), we used the embryonic transcriptome annotation from the same study. Splicing values per intron was calculated as previously described (Herzel and Neugebauer 2015). We required support of at least 10 junction reads (exon-intron/intron-exon/exon-exon) and allowed a maximal difference between 5' and 3' splicing value of 0.2, which resulted in the analysis of 116,385 introns in total.

iCLIP

2x 200 zebrafish embryos at 32-64 (1.75-2 hpf, preZGA) and 512-1K cell (2.75-3 hpf, ZGA) stages were irradiated twice with 0.8 J/cm^2 (Stratalinker 2400, Stratagene). Irradiated embryos were collected in 1.5 ml

tubes and snap frozen in liquid N_2. Frozen embryos were thawed and immediately homogenized on ice in 400 µl iCLIP lysis buffer (50 mM Tris HCl pH 7.4, 100 mM NaCl, 1% (v/v) NP40, 0.1% (w/v) SDS, 0.5% (w/v) Sodium deoxycholate, 7.5 µl RNase OUT (Invitrogen), complete Mini EDTA-free protease inhibitors (Roche)) by 10 passages through 1 ml syringe and 22G $1^{1/2}$ needles. Lysates were centrifuged (10 min, 14,000 rpm, 4°C), two supernatants from the same developmental stage combined and adjusted with iCLIP lysis buffer to the total volume of 1 ml. Pooled lysates were incubated 30 min with 25 µl Protein G Dynabeads (1:5 dilution; Invitrogen) on a rotating wheel at 4°C. Cleaned lysates were incubated 10 min with 5 µl TURBO DNase (Ambion) and 10 µl RNase I dilution (Ambion) at 37°C, 800 rpm and chilled on ice for at least 5 min. 10 µg of mouse monoclonal α-Hnrnpa1 antibody (Douglas Black Lab) or 10 µg of mouse non-immune IgG antibody (negative control; Sigma-Aldrich) were added and incubated over night at 4°C on a rotating wheel. 120 µl Protein G Dynabeads (1:5 stock dilution) were added to the antibody-lysate mix and incubated 1h at 4°C. Beads were washed twice with High Salt buffer (50 mM Tris-HCl pH 7.4, 1 M NaCl, 1 mM EDTA, 0.1 % (w/v) SDS, 0.5% (w/v) Sodium deoxycholate, 1% (v/v) NP40) and twice with PNK buffer (20 mM Tris-HCl pH 7.4, 10 mM $MgCl_2$, 0.2% Tween-20). The rest of the iCLIP experiment was performed as previously described (Konig et al. 2010). For final library amplification, 9 µl linearized cDNA were mixed with 10 µl 2x Phusion Flash High-Fidelity PCR master mix (Thermo), 0.5 µl nuclease-free water, and 0.5 µl P3/P5 primer mix (10 µM each) and amplified with 24 PCR cycles. Quality of the library was determined by resolving an aliquot of the PCR products on 6% TBE gel (Life Technologies), and staining with GelStar (Lonza) for 5 min. Three different size selected fractions of PCR amplified cDNA libraries were mixed in the same ratio, quantified by RT-qPCR and sequenced on the Illumina 2000 platform to obtain 75-nucleotide single-end reads containing a nine-nucleotide barcode. All oligonucleotides used in described iCLIP procedure are provided in Supplemental Table 6.

iCLIP data analysis

Sequencing reads obtained from individual replicates were sorted using the sequence information embedded within the experiment-specific barcodes introduced by RT primers. After removal of duplicate reads and

barcode sequence trimming, remaining reads were mapped to the reference genome assembly Zv9 with Bowtie v0.12.4 allowing two mismatches. Uniquely mapped reads were kept to determine statistically significant cross-link sites. Crosslink position was defined as the first nucleotide upstream of the read 5'-end. For identification of significant cross-link sites, iCLIP tag positions were randomized 100 times. For each experiment and all randomized samples, iCLIP tags were summed up in ± 15 nucleotide window around each cross-link site and probabilities of observation of the same or higher iCLIP tags was scored (Konig et al. 2010). iCLIP tags with FDR<0.01 were considered as significant. Since high degree of correlation was observed between replicates, data derived from all replicates were summed up and reanalyzed as described above. Protein-coding genes with more than 5 significant iCLIP tags were considered for further analysis. To perform direct comparison of iCLIP data between two developmental stages (**Fig. 4B**), the absolute number of iCLIP tags per mRNA was corrected to account for differences in both sequencing depth and the resulting number of significant iCLIP tags between developmental stage (**Supplemental Table S5**). Ensembl Zv9 (v72) annotation was used for initial iCLIP tag assignment to genes and their regions. Due to the observation of high number of iCLIP tags in intergenic regions, owing to the poor annotation of zebrafish genome, we reassigned significant iCLIP tags (FDR<0.01) to the custom gene annotation, created through the 3'UTR extensions based on the recent study (Ulitsky et al. 2012). Sequencing reads that mapped to multiple places were analyzed as following: reads whose length was within 17-30 nt range were allowed to have 20 hits at most, and 50 hits for reads whose length was 31 nt or higher. Multi mapped reads were weighted; each read contributed to the positions it was mapped to in a relative proportion, which was based on the number of hits the read had and its length. iCLIP enrichment (**see Supplemental Fig. S2E**) for preZGA stage was calculated as iCLIP tag density (number of iCLIP tags divided by the length of the transcript) normalized to the expression level of the transcript at 2 hpf using processed data from (Mishima and Tomari 2016).

Motif analysis

Motif analysis was performed using DREME. We selected iCLIP tags with FDR<0.01 for preZGA and ZGA samples. Sequences that lied within 25 nt upstream (-30 to -5) and 25 nt downstream (+5 to +30) of each

iCLIP tag were extracted strand-specifically to form the input. Nine nucleotides around the tag (including the tag itself) were skipped to avoid T-bias. To facilitate the motif finding process, we restricted the core motif to the hexanucleotide sequence. Shuffled sequences generated from input served as a background.

Plasmids

Hnrnpa1 shRNA-expressing plasmid (pSUPER-A1; 5'-AGCAAGAGATGGCTAGTGC-3') was a generous gift from Javier F. Caceres. Empty pSUPER plasmid was made from pSUPER-A1 construct by site-directed mutagenesis kit (Thermo) according to the manufacturer's instructions. GFP shRNA construct was created through annealing of two DNA oligos and their subsequent cloning into an empty pSUPER vector. Pri-miR430 construct was generated by PCR amplification of zebrafish miR430 polycistron from oligo(dT) reverse transcribed total RNA from late ZGA embryos (sphere stage; 6 hpf). Obtained amplicon was gel extracted and cloned into pcDNA3.0 plasmid between BamHI and XhoI restriction sites. 3'UTR of *npc1* gene was amplified from poly(A)+ selected cDNA. eGFP sequence was inserted into pCS2+ vector between BamHI and ClaI sites, whereas amplified 3'UTR was subsequently inserted between ClaI and XbaI restriction sites. AGGGA>TCCCT mutation was created using Site-Directed Mutagenesis Kit (Thermo Fisher) according to the manufacturer's instructions. RLuc containing plasmid was a generous gift from Susan Baserga Lab. RLuc sequence was HA tagged and inserted into pCS2+ plasmid. DNA oligos used for molecular cloning are listed in the Supplemental Table 6.

Northern blot

7.5 µg total RNA extracted from HeLa cells were loaded on the 10% TBE-Urea gel (Life Technologies) and ran 70 min at 180 V in 1x TBE buffer. RNA was subsequently electro-blotted onto the BrightStar-Plus, positively charged nylon membrane (Thermo) 1 hour at 350 mA in 0.5x TBE buffer. RNA was crosslinked to the membrane using UV light (254 nm). The membrane was prehybridized in 20 mL UltraHyb-Oligo buffer (Thermo) for 1 hour at 42°C. 2 µl 5'-biotinylated miR-430b DNA probe (100uM) was hybridized in the same buffer over night at 34°C. The membrane was washed twice 30 min with wash buffer (2x SSC, 0.5% SDS) at

room temperature. RNA detection was performed using Chemiluminescent Nucleic Acid Detection Module Kit (Thermo) according to the manufacturer's instructions.

Western blot

Total protein extracts from either dechorionated and deyolked embryos or HeLa cells were loaded onto the 4-12% polyacrylamide gels (Life Technologies) and transferred to the nitrocellulose membrane (Biorad). Blots were blocked with 5% non-fat milk in PBS-T (Tween, 0.1%) and probed for Hnrnpa1, RLuc-HA, splicing factors, eGFP and β-Tubulin, using mouse monocolonal α-Hnrnpa1 antibody (1:2,500; Douglas Black Lab) α-HA tag antibody (1:2,500; Millipore), 16H3 (undiluted hybridoma supernatant), goat α-eGFP (1:2,000; MPI-CBG) and rabbit α-β-Tubulin (1:1,500; Sigma-Aldrich) antibodies. Quantitative measurements of band intensities were performed using ImageJ software.

Definition of ZGA-independent (ZI) and ZGA-dependent (ZD) decaying genes

Decaying genes were defined based on their mRNA expression levels (RPKM) during MZT using the processed RNA-seq dataset (Mishima and Tomari 2016). Protein-coding genes with reliable ORF and 3'UTR annotations (>100 and 50-10,000 nt range for ORFs and 3'UTRs, respectively) were used in the analysis. Of those, only genes whose mRNA expression levels were >1 RPKM at 6 hpf were considered for further analysis and comparisons with Hnrnpa1 iCLIP data. Genes whose mRNA expression levels significantly dropped for more than 2 fold between 2 and 6 hpf wt were identified as destabilized (decay) genes. All decay genes were subsequently subdivided into two groups based on their expression response to α-amanitin treatment; genes whose mRNA levels significantly decreased more than 2 fold in α-amanitin injected 6 hpf embryos compared to preZGA (2 hpf) were considered as degraded independently of zygotic transcription and called ZGA-independent (ZI) decay genes (n=1337). Genes whose degradation was dependent on zygotic transcription (significant increase of more than 2 fold in α-amanitin injected embryos at 6 hpf compared to 6 hpf wt and simultaneously did not change expression levels between α-amanitin injected

embryos at 6 hpf compared to 2 hpf wt embryos were identified as ZGA-dependent (ZD) decay genes (n=1138).

Validation of Hnrnpa1 dissociation from maternal mRNAs and Hnrnpa1 interaction with *pri-mir-430* transcripts

200 and 250 zebrafish embryos at 32 cell and 1K cell stage or 32 cell and sphere stage were irradiated in E3 medium 2x with 0.8 J/cm^2 for validation of Hnrnpa1 dissociation from maternal mRNAs and Hnrnpa1 interaction with *pri-mir-430* transcripts, respectively. Irradiated embryos were homogenized on ice in 500 μl iCLIP lysis buffer. The lysates were subsequently centrifuged (10 min, 4°C, 14,000 rpm) and supernatants adjusted to the total volume of 1 mL with iCLIP lysis buffer. Lysates were incubated 10 min with TURBO DNase at 37°C and 800 rpm and chilled on ice for at least 5 min. 25 μl of each lysate was saved as input. 10 or 6 μg α-Hnrnpa1 antibody were added and incubated over night at 4°C. 42 μl of the stock or 100 μl Protein G-Dynabeads (1:5 dilution) were added to the antibody-lysate mix and incubated 1 hour at 4°C. Beads were washed 2x with High Salt buffer (3 min each), followed by 2 washing steps with PNK buffer (3 min each). 300 μl PK buffer were added together with 10 μl Proteinase K (10 mg/mL, Roche) to the beads. Thawed inputs were supplemented with 275 μl PK buffer and 10 μl proteinase K. Samples were incubated 30 min at 56°C at 1,000 rpm. After incubation, each sample volume was recovered from the beads and RNA extracted using Phenol-Chloroform extraction procedure. RNA was treated with TURBO DNase and re-extracted using the same procedure and ethanol precipitation. Extracted RNA was converted into cDNA using hexanucleotide mix (Roche). qPCR and PCR were performed using DNA oligos listed in the Supplemental Table 6.

Validation of Hnrnpa1 binding to AGGGA motif

80 embryos were injected with 150 ng wild type or mutated eGFP-*npc1*-3'UTR mRNA reporters at 1 cell stage. Two hours after injections, embryos were irradiated twice with 0.8 J/cm^2 and lysed in 500 μl iCLIP lysis buffer. Lysates were centrifuged at 14,000 rpm at 4°C, supernatants saved and adjusted to 1 ml with

iCLIP lysis buffer. 0.5% of the lysate was saved and frozen as input material. After 5 min DNase treatment, the lysates were further cleared with 8 µl Protein G-Dynabeads and subsequently incubated with 10 µg α-Hnrnpa1 antibody for 3 hours at 4°C. 42 µl Protein G-Dynabeads were added to the antibody-lysate mix and incubated 1 hour at 4°C. The rest of the protocol was performed as described in the section above. qPCR was performed with eGFP-specific oligo pair listed in the Supplemental Table 6.

ePAT

500 ng total RNA extracted from ten 64 cell stage embryos injected with 100 ng wild type or mutated eGFP mRNA reporters was used for ePAT. ePAT assay was performed as previously described with minor modifications (Janicke et al. 2012). To avoid amplification of poly(A) tails from endogenous transcripts, first PCR was performed using eGFP-specific and ePAT anchor primer with 18 cycles. To obtain shorter PCR products, 1:5 dilution of the first PCR was used as a template for the second PCR (20 cycles) where *npc1* UTR-specific forward primer was used together with the ePAT anchor primer. PCR products were loaded on 1.5% agarose gel and ran 65 min at 110 V. DNA oligos are listed in the Supplemental Table 6.

RT-PCR

1 µg total RNA extracted from HeLa cells subjected to different treatments was converted into cDNA using hexanucleotide mix (Roche). Subsequently, PCR was performed for *ndufa2* (NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 2) and *pri-mir-430* using the DNA oligonucleotides listed in Supplemental Table 6.

Supplemental references:

Cox J, Mann M. 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature biotechnology* **26**: 1367-1372.

Herzel L, Neugebauer KM. 2015. Quantification of co-transcriptional splicing from RNA-Seq data. *Methods* **85**: 36-43.

Janicke A, Vancuylenberg J, Boag PR, Traven A, Beilharz TH. 2012. ePAT: a simple method to tag adenylated RNA to measure poly(A)-tail length and other 3' RACE applications. *RNA* **18**: 1289-1295.

Konig J, Zarnack K, Rot G, Curk T, Kayikci M, Zupan B, Turner DJ, Luscombe NM, Ule J. 2010. iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nature structural & molecular biology* **17**: 909-915.

Mishima Y, Tomari Y. 2016. Codon Usage and 3' UTR Length Determine Maternal mRNA Stability in Zebrafish. *Mol Cell* **61**: 874-885.

Pauli A, Valen E, Lin MF, Garber M, Vastenhouw NL, Levin JZ, Fan L, Sandelin A, Rinn JL, Regev A et al. 2012. Systematic identification of long noncoding RNAs expressed during zebrafish embryogenesis. *Genome research* **22**: 577-591.

Rappsilber J, Mann M, Ishihama Y. 2007. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nature protocols* **2**: 1896-1906.

Ulitsky I, Shkumatava A, Jan CH, Subtelny AO, Koppstein D, Bell GW, Sive H, Bartel DP. 2012. Extensive alternative polyadenylation during zebrafish development. *Genome research* **22**: 2054-2066.