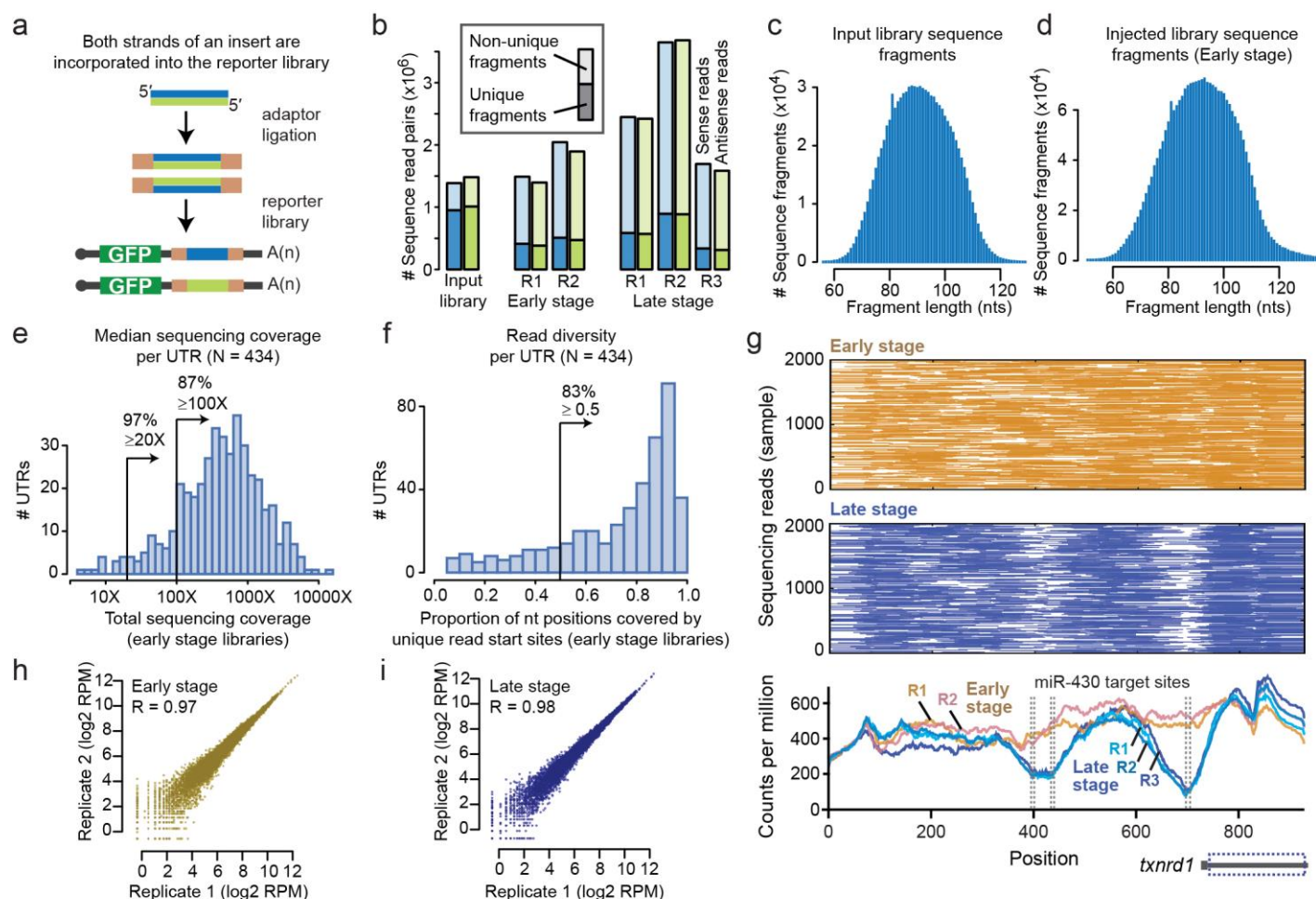
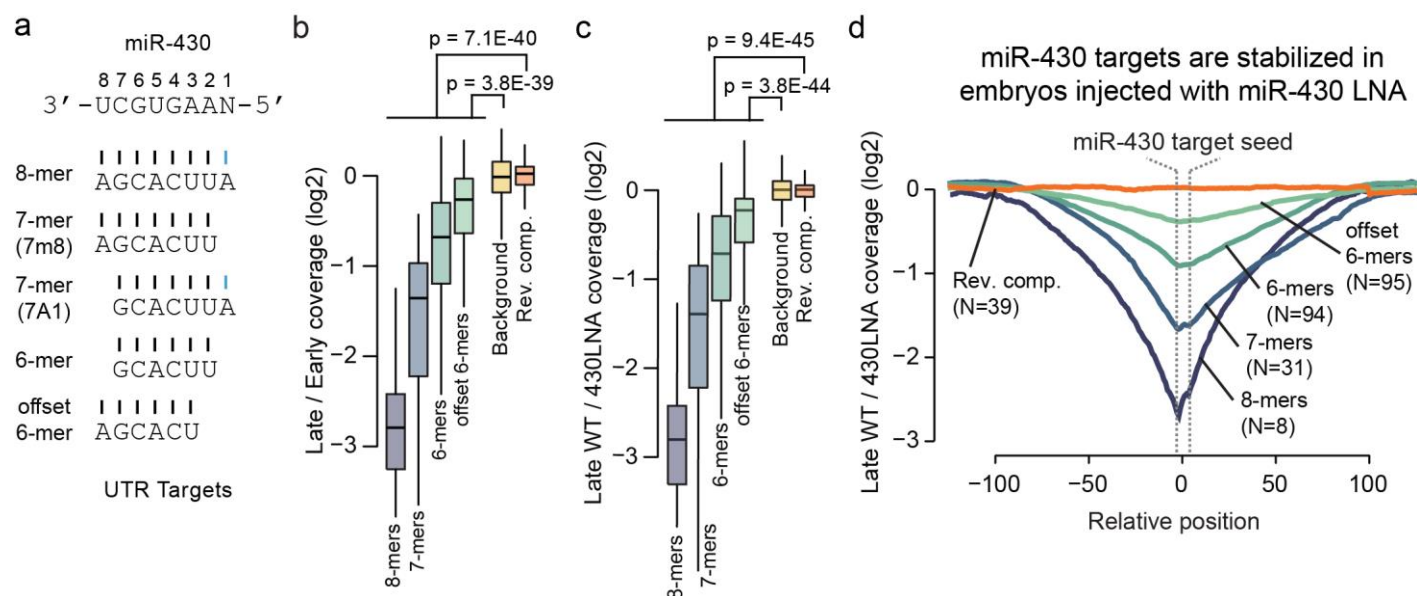




Supplementary Figure 1

RESA libraries are complex and reproducible.

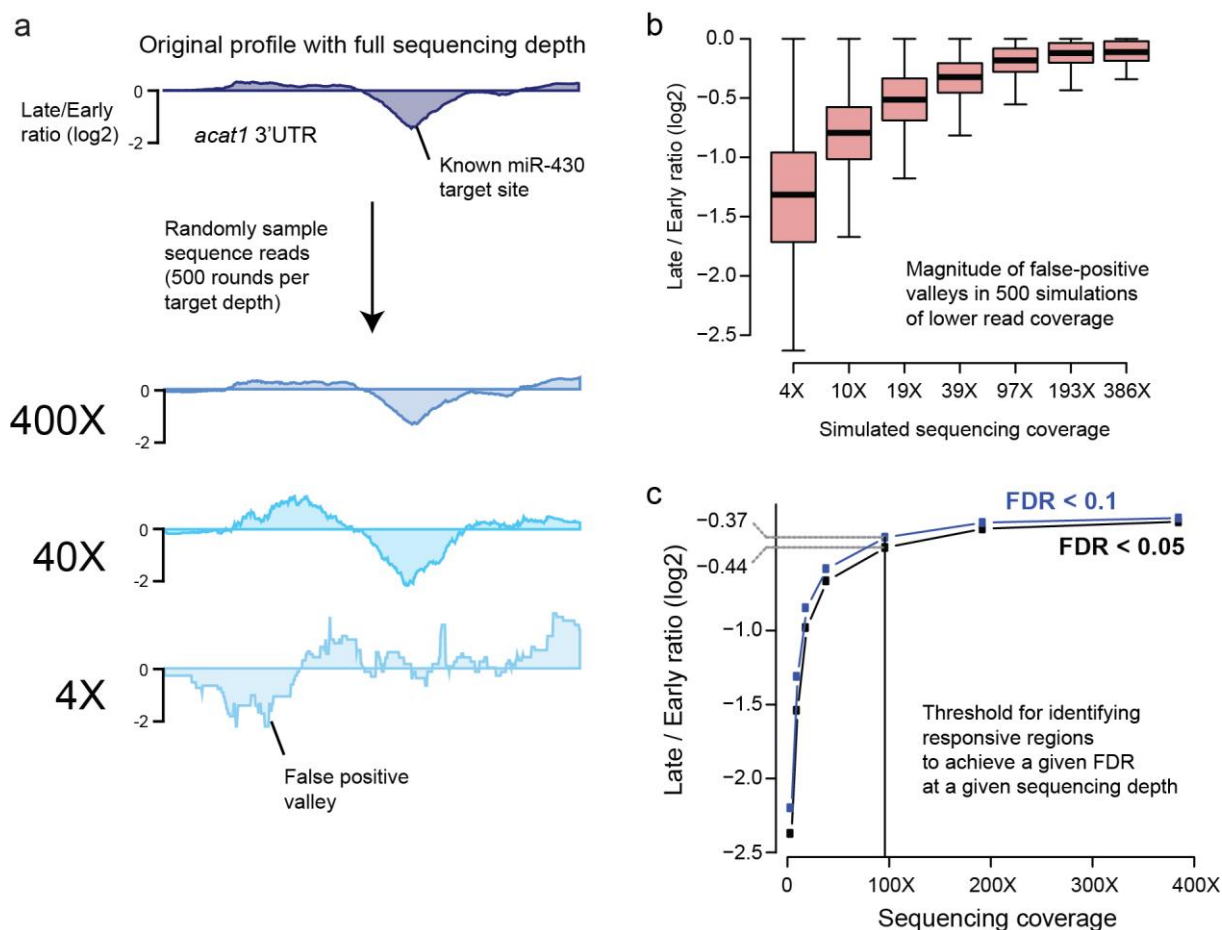
(a) Adaptor ligation to DNA fragments results in incorporation of both strands into the mRNA reporter library. Reverse complement variants can serve as internal controls in RESA experiments. **(b)** Proportion of unique inserts in each sample (defined as having the same aligned start and end position), segregated by read orientation. R1 = biological replicate 1, etc. **(c)** Distribution of sequence fragment length in library input sample. **(d)** Distribution of sequence fragment length recovered after injection into zebrafish. **(e)** Histogram of median sequencing coverage for successfully amplified 3'UTRs (N = 434). 97% of assayed 3'UTRs have coverage $\geq 20X$ and 87% have coverage $\geq 100X$. **(f)** Histogram of nucleotides proportion overlapped by unique read start sites across the UTRs (N = 434). For 83% of 3'UTRs, greater than 50% of positions are unique read start sites. **(g)** Randomly selected 2000 reads derived from early stage (yellow) and late stage (blue) samples overlapping miR-430 target locus *txnrd1*. Below, positional coverage for each condition separated by replicate. miR-430 seed sites are marked with dashed lines. **(h)** Sequencing read coverage over non-overlapping 50-nt windows across the 3'UTRs comparing early stage replicates and **(i)** late stage replicates. Pearson's correlation between early and late stage replicates is $R = 0.97$ and $R = 0.98$, respectively.



Supplementary Figure 2

miR-430-dependent regulation demonstrates the efficacy of the RESA platform.

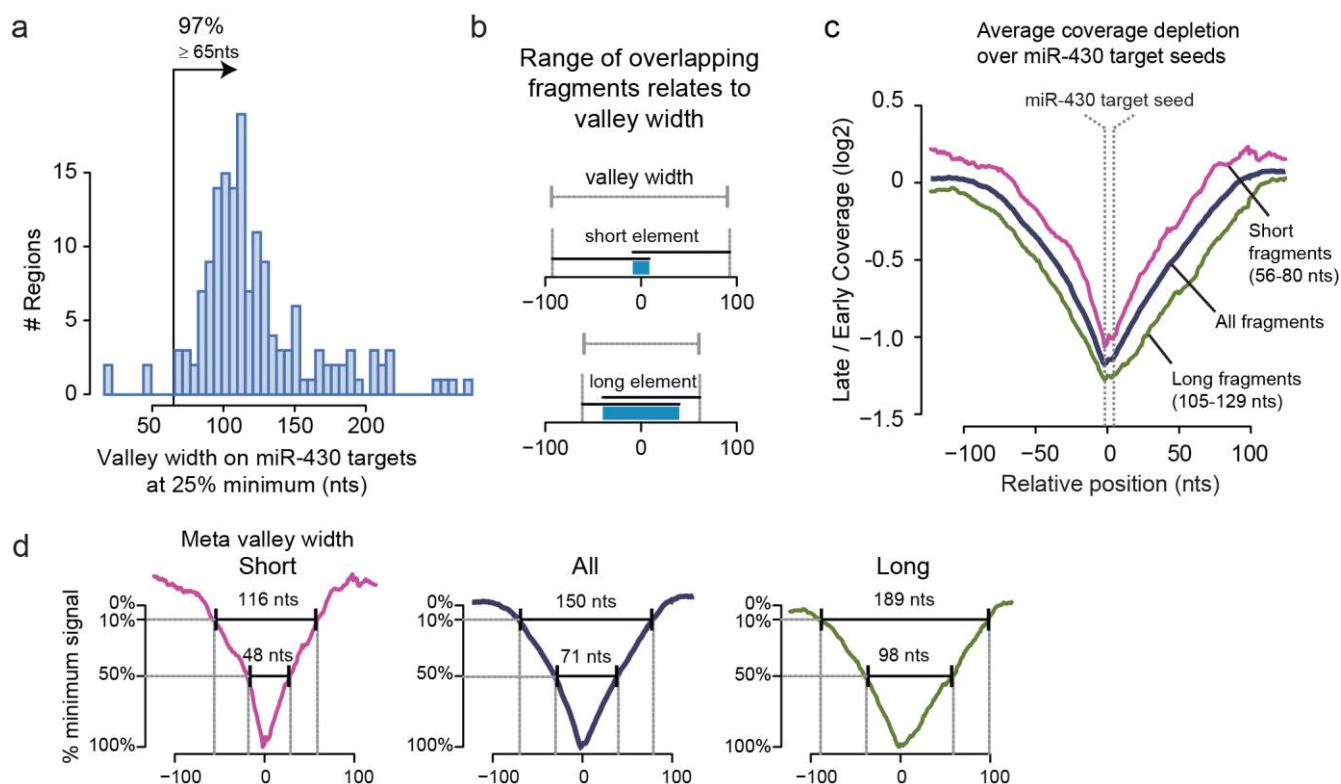
(a) miR-430 target seed definitions where Watson-Crick pairing occurs between positions 2-8 of the mature miR-430 sequence and complementary target sequence. An A in the target site opposite position 1 is also favorable. **(b)** Boxplot showing RESA coverage fold change late vs early stage centered on miR-430 target 8-mers (AGCACUUA), 7-mers (AGCACUU or GCACUUA), 6-mers (GCACUU), offset 6-mers (AGCACU), sequences encoding the non target seed CCACUU, and reverse complement coverage over the 7-mers and 8-mers. Depletion at late stage for the miRNA targets is significant (comparing all targets to non targets, $p = 3.8E-39$, Kruskal-Wallis Test; comparing sense 7-mers and 8-mers to their reverse complement, $p = 7.1E-40$, Wilcoxon Signed-Rank Test). **(c)** Similar to (b), comparing RESA coverage fold change wild type late stage vs 430LNA treated, which lacks miR-430 activity. Depletion at wild type late stage for the miRNA targets is significant (comparing all targets to non targets, $p = 3.8E-44$, Kruskal-Wallis Test; comparing sense 7-mers and 8-mers to their reverse complement, $p = 9.4E-44$, Wilcoxon Signed-Rank Test). **(d)** Meta profiles of coverage ratios (wild type late stage vs 430LNA) centered on all the miR-430 target seed in the study.



Supplementary Figure 3

FDR for regulated-region identification.

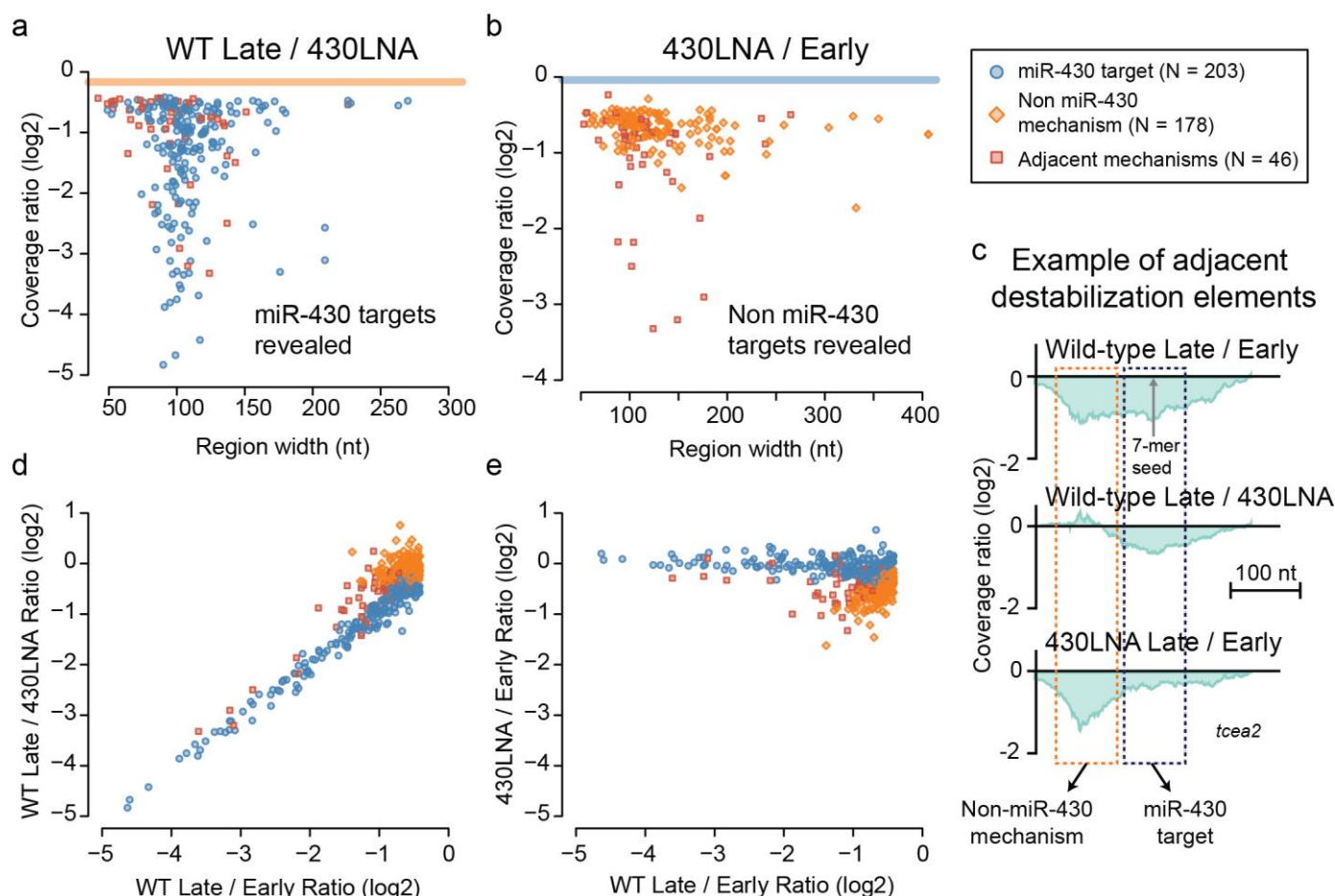
(a) Full depth sequencing profile for the *acat1* locus with miR-430 target seed. Below profiles generated using randomly sampled reads from early and late stage embryos at varying depth (500 rounds per target depth). False-positive valley is highlighted in 4X coverage condition. **(b)** Distribution of false-positive valley magnitude in 500 simulations of varied read coverage for the *acat1* locus. **(c)** Determination of a magnitude threshold for identifying regulated regions to reduce false discovery. For each simulation in (b), the *n*th percentile of the distribution of false positive magnitudes is plotted, for *n* = 0.05 and 0.1; this implies that 95% or 90% of false positive valleys have shallower magnitude (less negative) than this value, respectively. Thus, at 100X coverage, a threshold of -1.33-fold (-0.41 on a log2 scale) falls between a false-discovery rate (FDR) of 0.05 and 0.1. This parameter becomes the maximum (least negative) log2 coverage ratio allowed to confidently define a regulated region.



Supplementary Figure 4

Regulatory-element width measurement with the RESA platform.

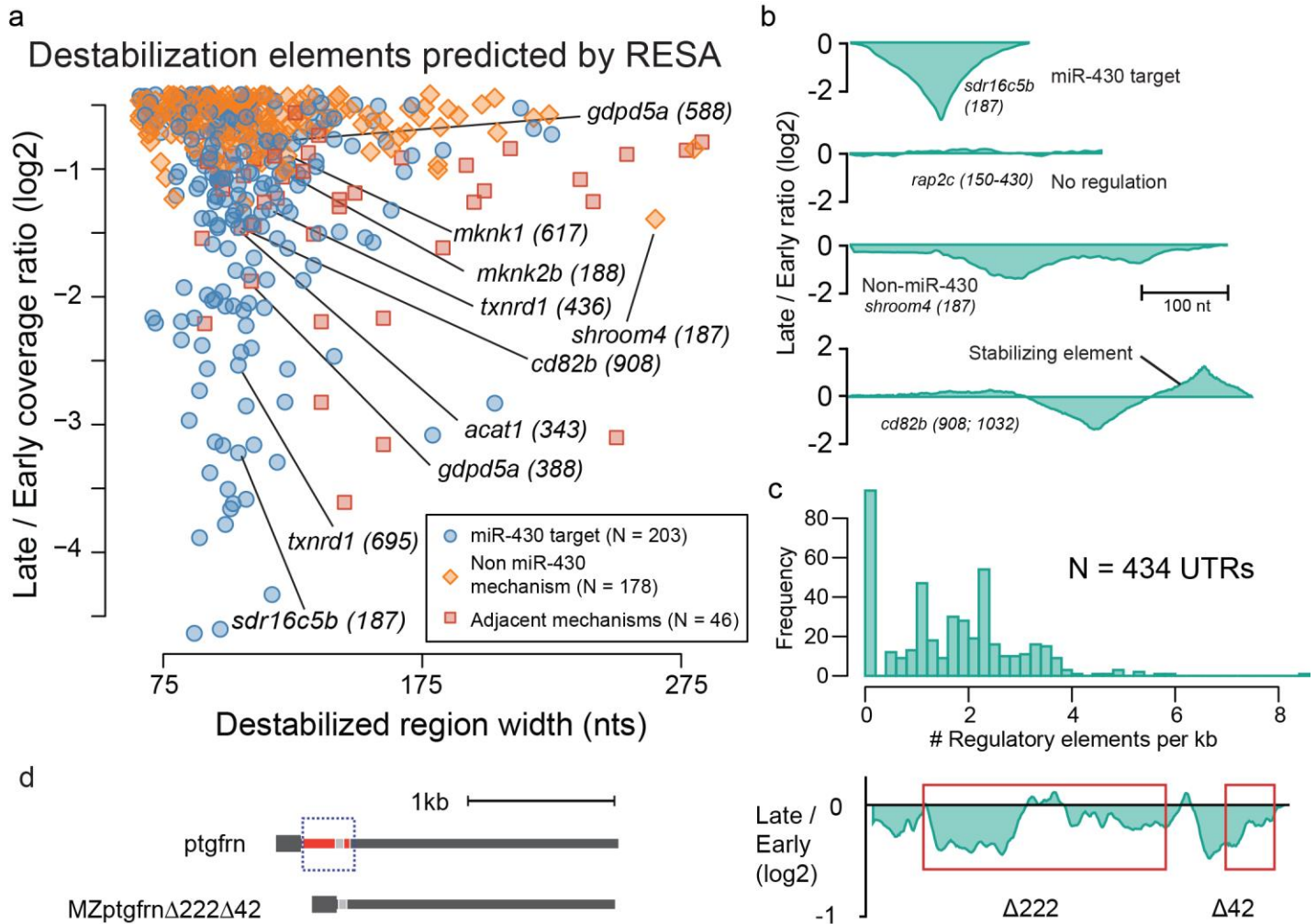
(a) Histogram of RESA destabilized region (valley) for miR-430 targets seeds. 97% of mir-430 target seeds overlap valleys that are ≥ 65 nt wide where width is measured at 25% of valley minimum magnitude. **(b)** Model of the relationship between regulatory element size and RESA valley widths. For reporter fragments (black bars) of a constant length distribution, longer regulatory elements (blue) will be fully overlapped by fewer fragments than elements with short sequence requirements, resulting in correspondingly narrower valleys. **(c)** Average coverage depletion over miR-430 target seeds (GCACUU) comparing short reporter inserts (56-80nt, pink), long fragments (105-129nt, green), and all fragments (56-129, blue). **(d)** The three profiles in (c) separated, showing nucleotide width as measured at 50% or 10% of the valley minimum (maximum depletion point).



Supplementary Figure 5

RESA approach to distinguish mRNA-regulatory mechanisms.

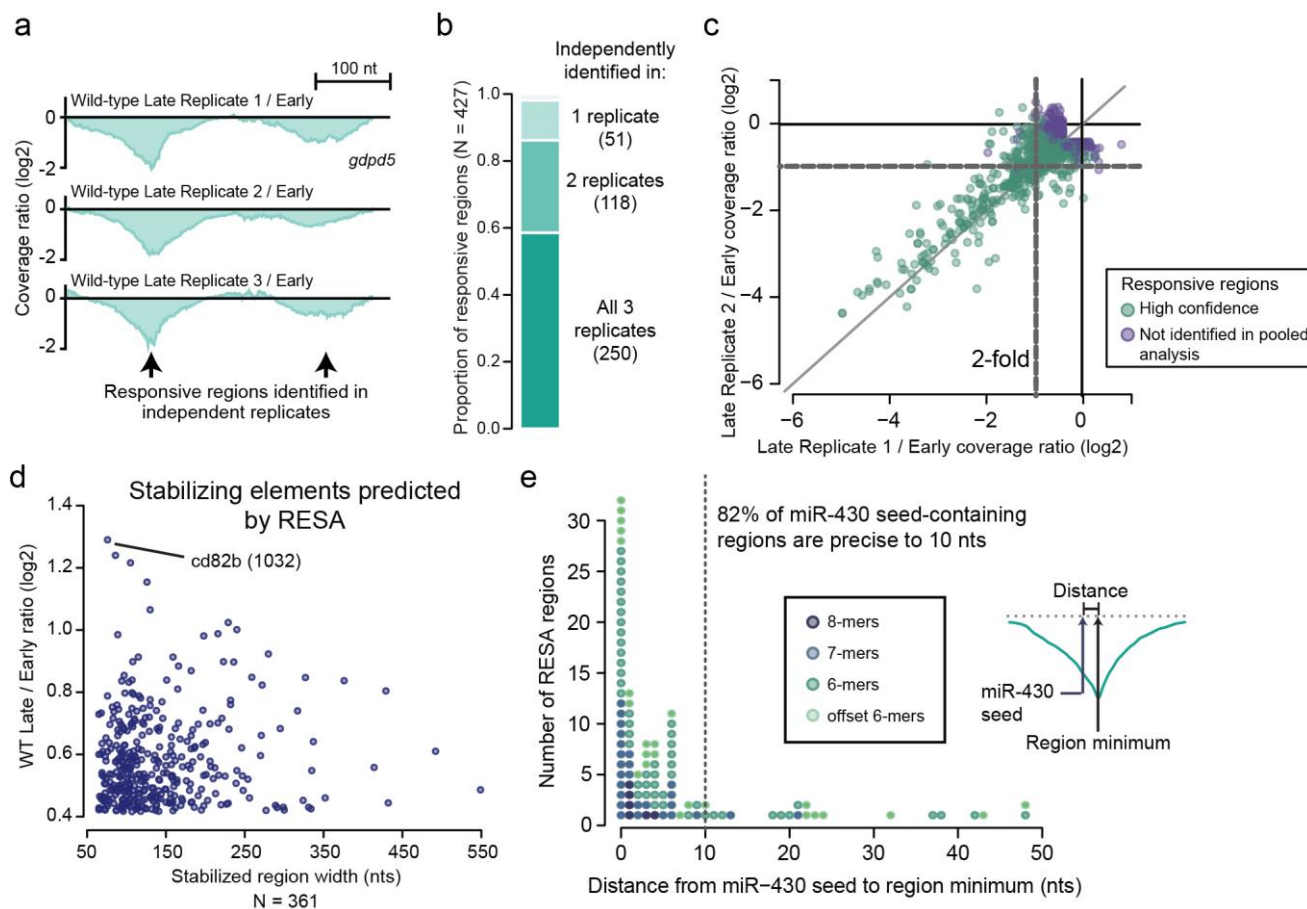
(a-c) Plots showing the dissection of mechanisms regulating mRNA stability. Destabilized regions that depend on miR-430 activity are revealed by comparing wild type late stage RESA coverage to 430LNA coverage, in which miR-430 is inhibited **(a)**, while regions that do not depend on miR-430 activity are revealed by comparing 430LNA coverage to early stage **(b)**. In each comparison, the horizontal line indicates the median coverage ratio of the regions that are not regulated. For a subset (red squares), the region is composed of adjacent miR-430-dependent and miR-430-independent regulatory elements, as they show regulation in both comparisons. **(c)** An example of a region (*tcea2*) with adjacent miR-430 dependent and independent mechanisms that are revealed using the 430LNA condition. **(d)** Plot comparing destabilization magnitudes at late stage when compared against early stage (x axis) or 430LNA (y axis). miR-430 dependent regions (blue) are correlated in these conditions. **(e)** Plot comparing destabilization magnitudes when comparing wild-type (x axis) or 430LNA (y axis) late stage to early stage. Regions that do not depend on miR-430 (orange) are correlated in these conditions.



Supplementary Figure 6

RESA identifies differentially regulated mRNA regions.

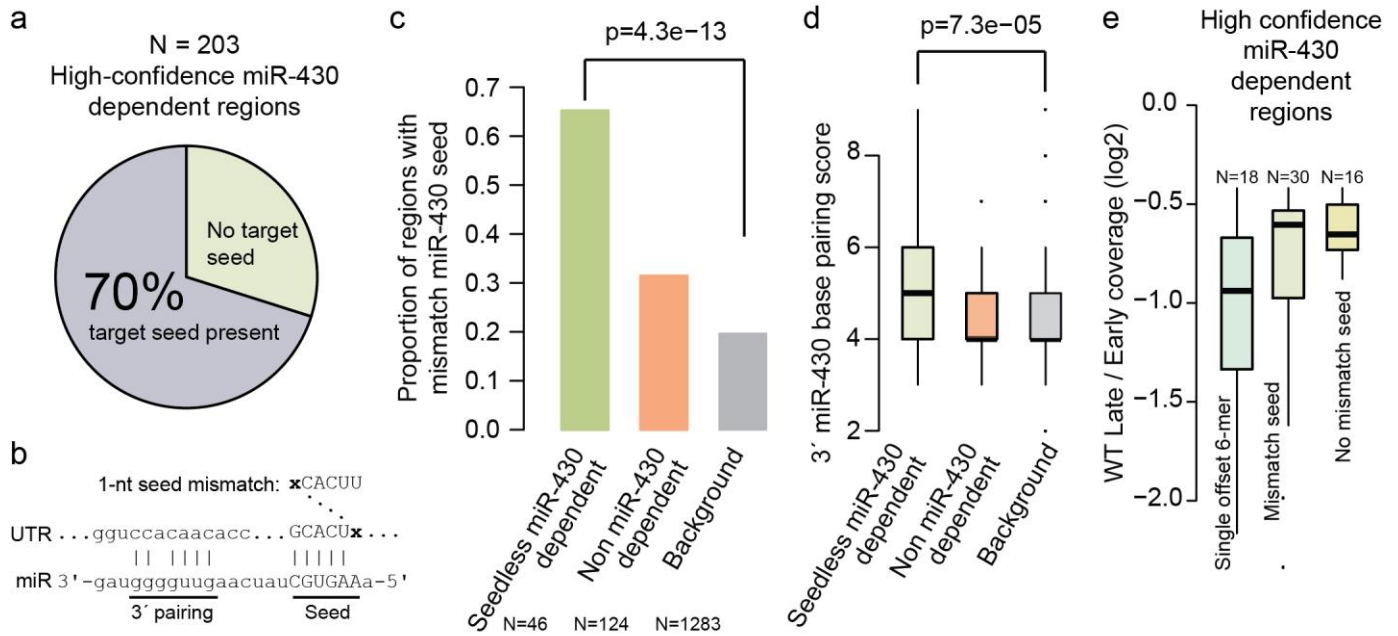
(a) High confidence (see Online Methods) miR-430 dependent (blue) regions, miR-430 independent regions (orange), and regions with adjacent mechanisms (red). Source gene name and relative position within the respective 3'UTR is marked for referenced regions (see Supplementary Data 3). **(b)** Log2 coverage profiles showing examples of a miR-430 dependent destabilized region (top), a non-regulated region (second), a non-miR-430 dependent destabilized region (third), and adjacent destabilized and stabilized regions (fourth). Source gene and relative position in the UTR amplicon are shown below each. **(c)** Histogram of regulatory element density in 3'UTRs. **(d)** (left) Diagram of the predicted destabilization element in the *ptgfrn* 3'UTR (dashed box) and the CRISPR-mediated deletion (red) in *MZptgfrn*Δ222Δ42. (right) RESA coverage profile over the destabilized region.



Supplementary Figure 7

RESA identifies regulated regions with high precision.

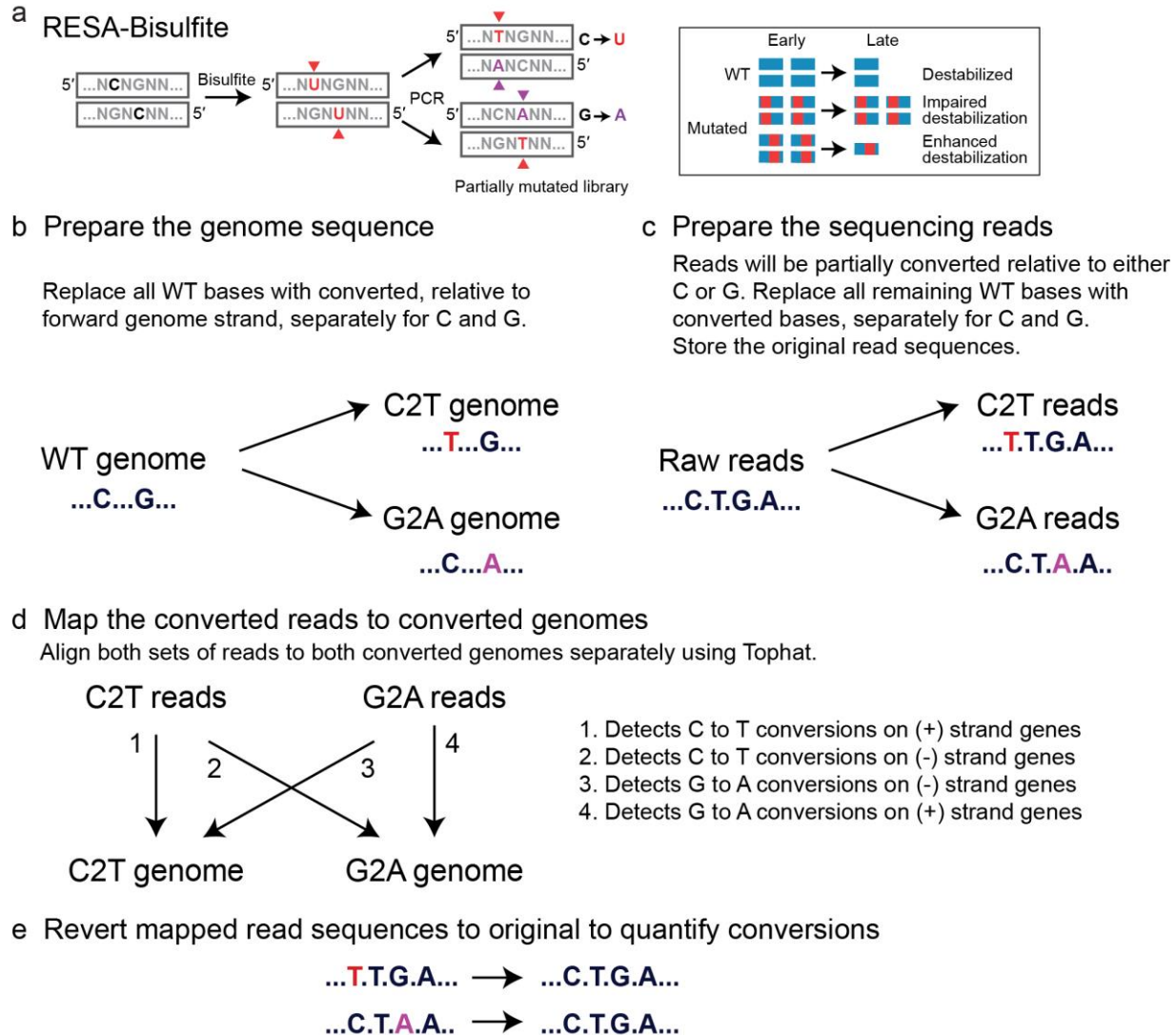
(a) Coverage ratio plots over the *gdpd5a* UTR separated by three replicate late stage samples when compared against the pooled early stage each. Responsive regions are marked with arrow. **(b)** RESA destabilized regions from the pooled analysis are also identified when single replicate late-stage samples are used individually. 86% of the regions are found individually by at least 2 separate replicates. **(c)** Plot comparing log2 coverage ratio using 2 different single replicate late-stage samples, showing high agreement for the RESA destabilized regions from the pooled analysis (green). Additional destabilized regions are predicted in one but not the other replicate (purple); however, these are only weakly predicted. **(d)** High confidence (see Online Methods) stabilized regions. Wild-type sequence coverage enrichment over time indicates stability-promoting regions, since stable regions accumulate over time with respect to the overall mRNA population. **(e)** Dot plot of miR-430 dependent destabilized regions containing miR-430 target seeds (within 50 nts of the region minimum, for the late versus early stage comparison ratio). 82% of miR-430 seed-containing regions are located within 10 nts of the region minimum, demonstrating the high predictive power of RESA valley shape. Only regions in the interior of amplicons (i.e., not truncated by edge effects) were considered.



Supplementary Figure 8

Compensatory pairing and mismatch seeds are detected in miR-430-dependent regions.

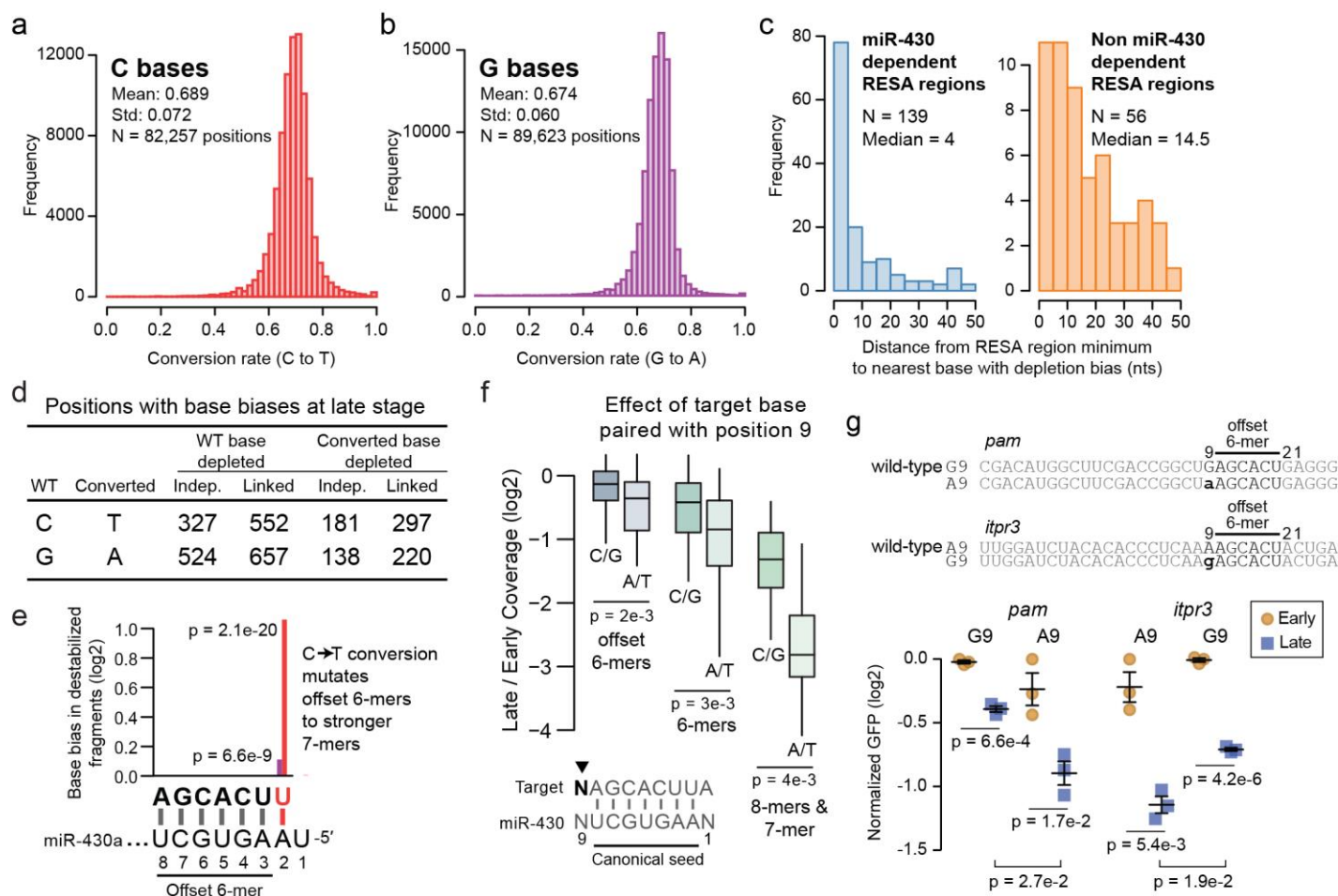
(a) Proportion of RESA predicted miR-430 dependent destabilized regions that contains a canonical miR-430 target seed match is 70%. Of the remaining regions (non target seed), the majority (30 out of 61) contains a 1-nt mismatch seed, as illustrated below **(b)** Scheme illustrating 1-nt mismatch and compensatory pairing for miR-430 target region. **(c)** The “non target seed” regions are significantly enriched for containing mismatch seeds, as well as base-pairing complementarity to the 3' end of mature miR-430 **(d)**, compared to the non-miR-430 dependent RESA destabilized regions or a set of background, unregulated regions (see Online Methods). **(e)** Target seed match has a significant effect ($P = 3.2e-9$, Kruskal-Wallis Test) on destabilization strength, with non-seed match containing regions weaker than offset 6-mer containing regions.



Supplementary Figure 9

RESA-bisulfite mapping strategy.

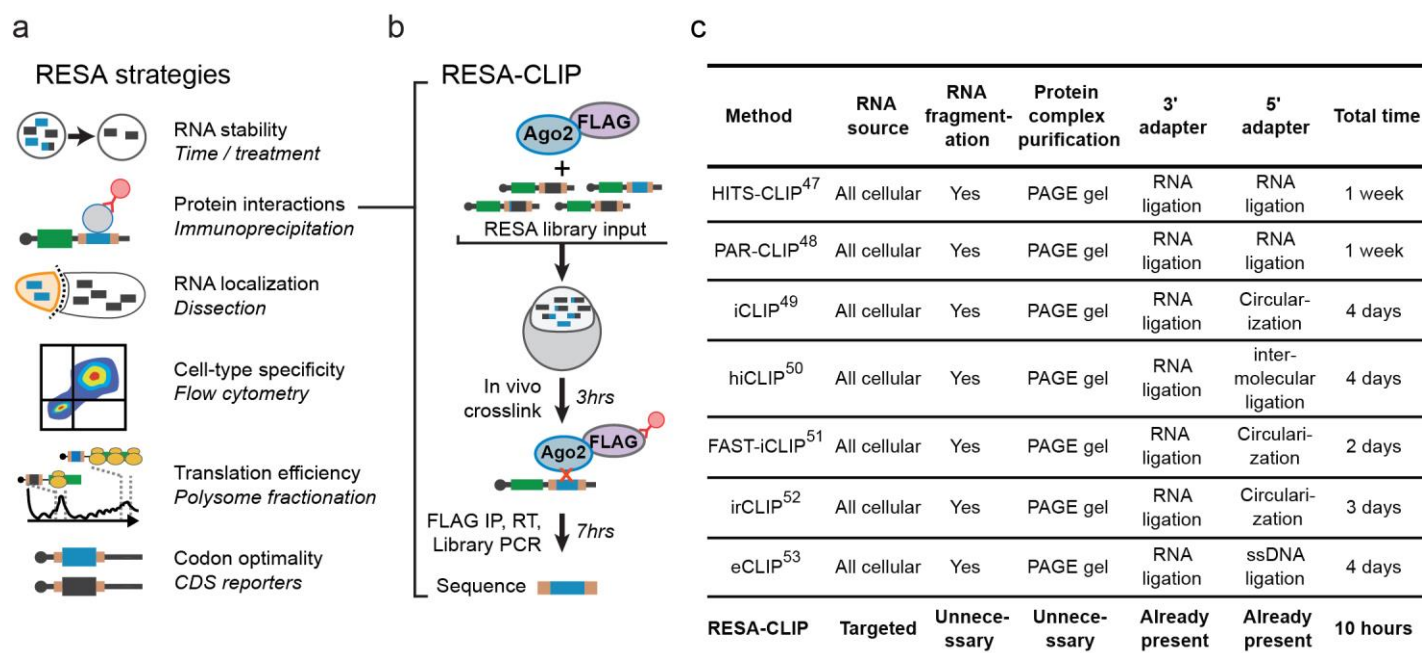
(a) RESA-Bisulfite used bisulfite treatment to induce C to U or G to A mutations within the reporter library pool. Partially mutated reporters are injected in vivo to determine the effect of mutations (red, purple) on the destabilization of wild type (WT, blue) regions. **(b)** Mapping partially converted reads to the genome first requires building a reference genome where all C bases are converted to T (top) and one where all G bases are converted to A (bottom). Reads are expected to encode either C to T conversions or G to A conversions (relative to the sense orientation of the read pair), but not both. **(c)** Because fragments are only partially converted, all reads are subjected to in silico conversion to completion, by changing all C bases to T (top) or separately all G bases to A (bottom). This is to ensure that they align to the converted genomes from (b). These two sets of reads will be treated separately moving forward. Base identity is with respect to the orientation of the fragment, which is defined by read 1 of the read pair; thus, conversions on read 2 are reverse complements due to the design of Illumina paired-end reads (i.e., if read 1 is subjected to C to T conversion, read 2 is subjected to G to A). For a given read pair, it is unknown a priori whether the fragment originally encodes C to T or G to A conversions, so both in silico conversions are always applied, with the expectation that only one will yield an alignable read pair. **(d)** Mapping is done four ways, with each set of converted reads aligned to each of the two converted genomes; only one of the four pairings should produce a successful alignment per read pair. If a gene is in the sense orientation with respect to the reference genome sequence, then successful alignment will occur sense to sense (1 or 4). If a gene is in the antisense orientation with respect to the genome, successful alignment will occur sense to antisense (2 or 3). **(e)** Once read pairs are mapped, the reads are reverted to their original, partially converted sequences. This is easily implemented by encoding the original sequence as part of the read ID during the in silico conversion phase.



Supplementary Figure 10

RESA-bisulfite reveals nucleotide positions associated with differential regulation.

(a, b) Comparison of the reference base identities of each nucleotide position to the bases in the aligned reads induces a conversion rate, which follows an approximately normal distribution centered at 0.67-0.69. Only positions with at least 10 overlapping reads were used to calculate this distribution. **(c)** Histograms of the positions of bases where significant depletion of the wild-type allele was observed, with respect to RESA-predicted destabilized regions. Distance is measured from the destabilized region minimum to the nearest depleted wild-type base, up to 50 nts. **(d)** Summary of positions with nucleotide biases (depletion of wild-type or converted base) at late stage. Positions with biases observed as linked dibases or tribases are distinguished from positions with independent effects. **(e)** Pooled base conversion adjacent to miR-430 offset 6-mer target sites enhances destabilization. Y-axis shows the magnitude of change in the proportion of converted bases observed (U, red; A, purple) from early to late stage. X-axis is oriented sense to the target. P value from a G test of independence is shown. **(f)** miR-430 target seeds with A or U at position 9 (outside of the canonical microRNA seed region) are significantly more destabilized than those with C or G, across target types. 8-mers and 7-mers (pairing positions 2-8) are pooled. **(g)** GFP reporter analysis comparing the stability of endogenous loci containing offset 6-mer miR-430 sites with either G (*pam*) or A (*itpr3*) at position 9, versus mutant variants where the G is mutated to A (*pam*) or A to G (*itpr3*). GFP mRNA level is normalized to control dsRed mRNA and assayed by qRT-PCR. In both cases reporters with A in position 9 show significantly stronger destabilization: substituting the endogenous G with an A in the *pam* sequence increased destabilization 35%, while conversely substituting the endogenous A with a G in *itpr3* attenuated destabilization 29% (Welch's t-test; bars show mean $\pm$ SEM).



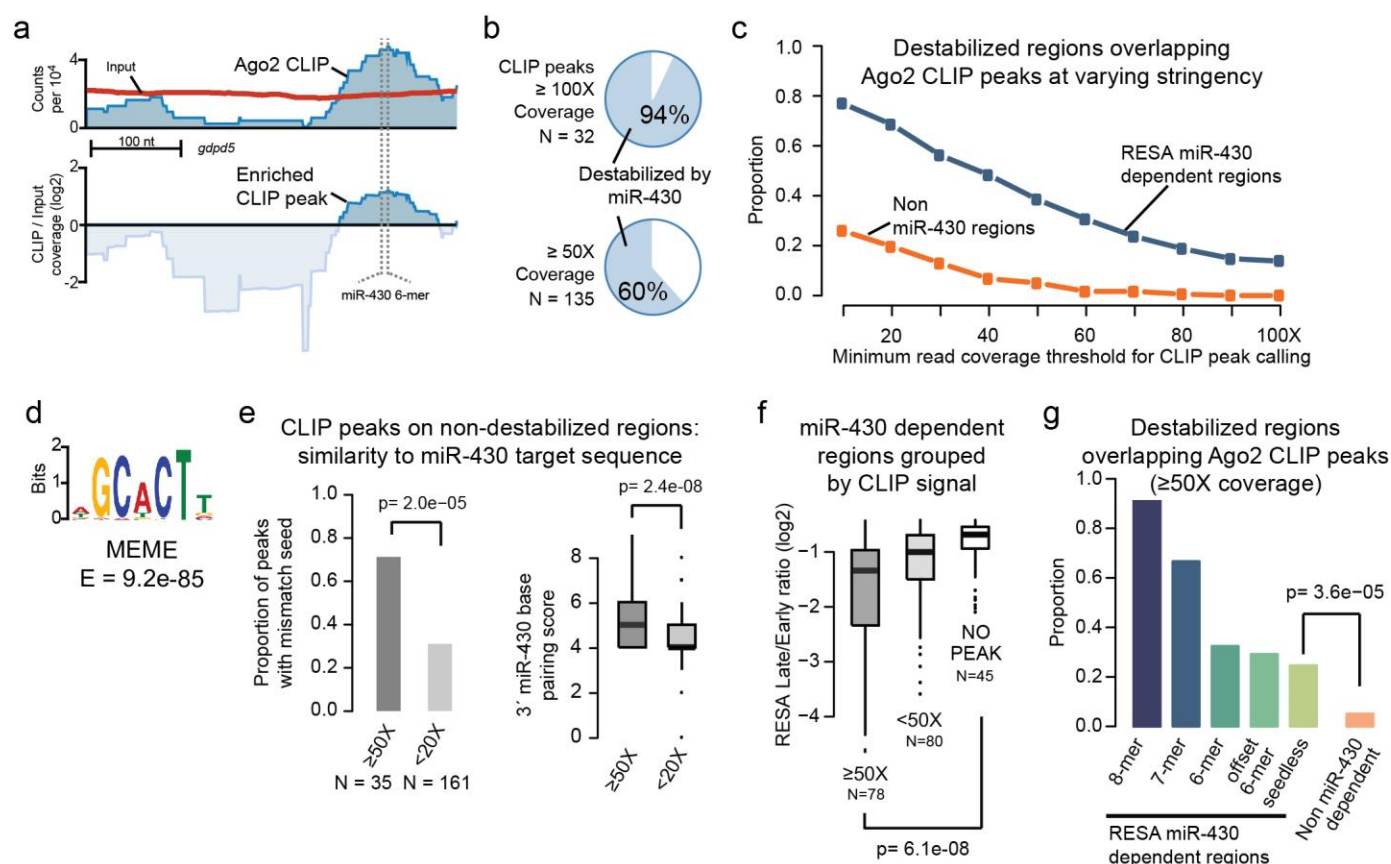
Supplementary Figure 11

RESA applications and CLIP method comparison.

(a) Diagram of proposed examples of RESA applications in diverse biological contexts. **(b)** Overview of RESA-CLIP strategy. Tagged Ago2 is co-injected with the RESA library, injected into zebrafish embryos, crosslinked, immunoprecipitated, and sequenced to identify Ago2-bound regions. **(c)** Comparison of RESA-CLIP to other methods to interrogate protein-RNA interactions.

Supplementary References

[47] Licatalosi, D. D. et al. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* 456, 464–470 (2008). [48] Hafner, M. et al. Transcriptome wide identification of RNA binding protein and microRNA target sites by PAR-CLIP. *Cell* 141, 129–141 (2010). [49] Huppertz, I. et al. iCLIP: Protein-RNA interactions at nucleotide resolution. *Methods* 65, 274–287 (2014). [50] Sugimoto, Y. et al. hiCLIP reveals the in vivo atlas of mRNA secondary structures recognized by Staufen 1. *Nature* 519, 491–494 (2015). [51] Flynn, R. A. et al. Dissecting noncoding and pathogen RNA – protein interactomes. *RNA* 21, 135–143 (2014). [52] Zarnegar, B. J. et al. irCLIP platform for efficient characterization of protein – RNA interactions. *Nat. Methods* 13, (2016). [53] Van Nostrand, E. L. et al. Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat. Methods* 1–9 (2016).



Supplementary Figure 12

RESA–CLIP shows specific Ago2 association with miR-430-dependent destabilized regions.

(a) (Top) Plot of reads-per-million normalized coverage over the *gdpd5a* 3'UTR showing RESA-CLIP levels (shaded blue) compared to input RNA levels (orange line). (Bottom) Corresponding plot of log2 ratio CLIP over input. Only the region overlapping the miR-430 target seed has enrichment of CLIP signal. **(b)** Proportion of CLIP peaks overlapping RESA-predicted miR-430 dependent regions, at minimum CLIP read threshold 100X (top) or 50X (bottom). **(c)** Proportion of RESA destabilized regions classified as miR-430 dependent (blue) and non miR-430 dependent (orange) overlapped by CLIP peaks, as minimum read threshold is varied. A higher proportion of miR-430 dependent regions are overlapped throughout the threshold range. **(d)** Motif analysis in Ago2 CLIP peaks shows significant enrichment of the miR-430 target seed sequence. **(e)** CLIP peaks that do not overlap RESA destabilized regions are enriched in partial sequence matches to mature miR-430, correlated with the strength of the CLIP signal ($P = 2.0 \times 10^{-5}$, Chi-squared test comparing presence/absence of mismatch seeds; $P = 2.4 \times 10^{-8}$, Wilcoxon rank sum test comparing miR-430 base-pairing score, see Online Methods). This suggests that RESA-CLIP is capturing RISC interactions as it samples for matches to the miR-430 sequence. **(f)** Boxplot showing RESA late versus early stage coverage ratio of miR-430 dependent destabilized regions, grouped according to CLIP sequencing read coverage. Regions with greater CLIP coverage have significantly stronger destabilization ($P = 6.1 \times 10^{-8}$, Wilcoxon rank sum test). **(g)** The proportion of miR-430 dependent RESA destabilized regions overlapped by CLIP peaks ($\geq 50X$ coverage) is correlated with the strength of the miR-430–target seed interaction. However, even miR-430 dependent destabilized regions lacking a target seed match are significantly enriched for CLIP peaks compared to non-miR-430 dependent destabilized regions ($P = 3.6 \times 10^{-5}$, Chi-squared test).

Supplementary Table 1. Comparison of approaches to discover RNA regulatory elements.

Strategy	Implementation	Advantages	Disadvantages	References
Full-length biological sequences	PCR or endogenous half-life measurements	Potential to capture regulatory RNA activity which require long-range interactions	Limited insight on the minimal regulatory element; multiple regulatory elements may be present per sequence whose effects can't be separated	Geisberg, 2014, Wolter, 2015 (3'LIFE)
Random oligonucleotide sequences	Custom synthesis	Unbiased sampling of potential regulatory sequences	Number of possible sequences grows exponentially for longer oligos, making it difficult/inefficient to assay longer regulatory elements	Rosenberg, 2015, Wissink, 2016
Targeted short biological sequences	Custom synthesis	Measuring activity of biological sequence fragments	Labor intensive curation, biased selection of sequences to assay (e.g., using sequence conservation level)	Oikonomou, 2014, Zhao 2014 (fast-UTR), Weingarten-Gabbay, 2016
Randomly fragmented biological sequences	PCR followed by sonication, or pre-built high-throughput sequencing libraries can also be used	Coverage by many partially overlapping random fragments yields high nucleotide resolution of regulatory element position	High nucleotide resolution depends on sufficient depth of sequencing (100X)	Current study (RESA)

Supplementary Table 2. Summary of sequencing results.

Condition	Replicate	Experiment	Total read pairs	Aligned read pairs		
				WT	C to T converted	G to A converted
Library Input	1	RESA	3530868	2928926	NA	NA
Early	1	RESA	3494047	2946878	NA	NA
Early	2	RESA	4766081	4020944	NA	NA
Late	1	RESA	5895953	4975343	NA	NA
Late	2	RESA	8845108	7477684	NA	NA
Late	3	RESA	3958477	3346324	NA	NA
430LNA	1	RESA	9489219	8013334	NA	NA
430LNA	2	RESA	8640935	7325241	NA	NA
430LNA	3	RESA	11840412	10020842	NA	NA
Early	1	RESA-Bisulfite	13487785	36351	2295038	7906741
Early	2	RESA-Bisulfite	11984468	31996	2195578	6866070
Early	3	RESA-Bisulfite	14991744	42034	2733234	8616533
Late	1	RESA-Bisulfite	4166676	11298	724906	2426895
Late	2	RESA-Bisulfite	7302088	19730	1243021	4278012
Late	3	RESA-Bisulfite	7412900	19946	1250178	4355579
Ago2	1	RESA-CLIP	61467344	47754369	NA	NA
CLIP Input	1	CLIP-Input	64900157	52860980	NA	NA

Each entry lists an independent biological sample. Total read pairs are reported, where each pair equals two sequencing reads (read 1 and read 2). Alignment count is with respect to concordantly aligned read pairs. For bisulfite-converted reads, alignment totals are reported separately for the two different converted genomes.

Supplementary Table 3. Summary of miR-430 target sequences across the UTR library.

	Within RESA destabilized region			Poly-morphic seed	Low Seq coverage	No destabilization detected
	miR-430 dependent	Non miR-430 dependent	Low confidence region			
8-mers	13	0	1	0	0	0 (0%)
7-mers	38	2	1	0	2	1 (2%)
6-mers	93	9	22	2	8	25 (16%)
Offset 6-mers	40	5	25	7	15	49 (35%)

miR-430 target seed sequences (8-mer, 7-mer, 6-mer and offset 6-mer matches) are enumerated across the 3'UTRs using the Zv9 reference genome sequence. For each category, a seed may fall within a RESA destabilized region (assigned to the miR-430 dependent, non miR-430 dependent, or low-confidence destabilized categories). Otherwise, a seed sequence may have nucleotide polymorphisms compared to the reference genome (affecting target complementarity) or low sequencing read coverage (<100X at early stage).