Supplementary Information for

Genome-wide RNA structurome reprogramming by acute heat shock globally regulates mRNA abundance

Zhao Su[a,1], Yin Tang[b,1,2], Laura E. Ritchey[c,d], David C. Tack[a], Mengmeng Zhu[a], Philip C. Bevilacqua[c,d,e,3], and Sarah M. Assmann[a,d,3]

[3]Corresponding authors' email: sma3@psu.edu (S.M.A.); pcb5@psu.edu (P.C.B.).

This PDF file includes:

Materials and Methods

Plant material and growth conditions

Seeds of rice (*Oryza sativa* ssp. *japonica* cv. Nipponbare) were sown on wet filter paper in a petri dish and geminated for five days in a greenhouse with 16h/8h day/night photoperiod, with light intensity ~500 µmol m^{-2} s^{-1} supplied by natural daylight supplemented with 1000 W metal halide lamps (Philips Lighting Co). The temperature was 28–32 °C during the day and 25–28 °C during the night. The rice seedlings were then transferred to 6 × 6 inch nursery pots filled with water-saturated soil (Metro Mix 360 growing medium, Sun Gro Horticulture, Bellevue, WA). Nine plants were grown per pot and were watered once a week after transferring the seedlings to the pots. Shoot tissue of two-week-old plants was used for *in vivo* DMS probing. All tissue collection started at ~4 p.m. for all genome-wide experiments to minimize circadian effects.

In vivo DMS probing under two temperature conditions

All manipulations using DMS were conducted with proper safety equipment including lab coats and double gloves. All disposables were disposed of as hazardous waste. DMS treatment was applied in a chemical fume hood with strong airflow (> 200 fpm). For the 22 °C treatment, non-DMS-treated (−DMS) and DMS-treated (+DMS) samples were prepared. One g of shoot tissue was excised from the plant immediately before each treatment. For the +DMS sample, the material was immersed in 20 mL DMS reaction buffer (40 mM HEPES (pH 7.5), 100 mM KCl, and 0.5 mM $MgCl_2$) in a 50 mL conical centrifuge tube. Then 150 µl DMS (D186309, Sigma-Aldrich) was immediately added to the solution to a final concentration of 0.75% (~75 mM), followed by 10 min of gentle inversion and mixing for every 30 seconds. Next, to quench DMS in the reaction (1), dithiothreitol (DTT) at a final concentration of 0.5 M was supplied by adding 1.5 g DTT powder into the solution. After vigorous vortexing to dissolve the DTT, the quench proceeded for 2 min. The solution was decanted, and each sample was washed twice with distilled deionized water. Residual water was removed by inverting the tube onto paper towels, and the tissue was immediately frozen in liquid nitrogen. The −DMS sample was processed through the same procedure without addition of DMS. Three biological replicates were prepared for each −/+ DMS sample for a total of six samples.

For the heat shock treatment, −DMS and +DMS samples were similarly prepared. For the DMS treatment, 1 g of shoot was excised and placed into 20 mL of 42 °C pre-warmed DMS reaction buffer for 30 seconds in a 50 mL centrifuge tube for temperature equilibration of the tissue. Then 150 µl DMS was added, followed by 10 min of intermittent inversion and mixing in a 42 °C water bath to maintain the temperature. Then 1.5 g of DTT powder was added into the reaction solution for a final DTT concentration of 0.5 M to quench the DMS with the tube immersed in the 42 °C water bath for 2 min. The solution was decanted, and samples were washed twice and immediately frozen in liquid nitrogen. The −DMS 42 °C samples were processed through the same procedure, without DMS addition. Three biological replicates were prepared for each sample, for a total of six additional samples.

Structure-seq library generation

Library generation followed our previous library construction pipeline (2, 3) with some optimization. Total RNA for the 12 individual biological samples was obtained in a chemical fume hood using the NucleoSpin RNA Plant kit (Cat# 740949, Macherey-Nagel, Germany) following the manufacturer's protocol. For each sample, 300 µg total RNA comprised the starting material for two rounds of poly(A) selection using the Poly(A)Purist MAG Kit (Cat# AM1922, ThermoFisher), which provided high purity mRNA for library

construction. poly(A) purified mRNA (500 ng) was used as the input for Structure-seq library construction following the Structure-seq2 protocol (3). Reverse transcription was performed using SuperScript III First-Strand Synthesis System kit (Cat# 18080051, ThermoFisher) using the same RT primer as previously used (1): 5′CAGACGTGTGCTCTTCCGATCNNNNNN3′ which is a fusion of a random hexamer and an Illumina TruSeq Adapter. The first-strand cDNA was size-selected above 52 nt on a 8M urea 10% polyacrylamide gel to remove excess RT primer and increase the ligation efficiency in the next step. After recovering cDNA using the crush-soak method, the cDNA was dissolved in 5 μL RNase-free water. Ligation was performed using T4 DNA ligase (Cat# M0202, New England Biolabs) which ligated the 3′ end of the cDNA to a low bias single-stranded DNA linker (4) /5Phos/TGAAGAGCCTAGTCGCTGTTCANNNNNNCTGCCCATAGAG/3SpC3/ where the underlined sequence can form a hairpin structure and the random hexamer can then hybridize to any cDNA fragment (4). Reagents were added into the cDNA solution as follows: 2 μL 10x buffer, 2 μL 5M betaine, 2 μL 100 μM linker DNA, 8 μL 50% PEG8000, 1 μL T4 DNA ligase (400U/μL). The ligation was performed at 16 °C for 6 h and then 30 °C for 6 h, and the ligase was then deactivated at 65 °C for 15 min. The ligation product was size selected above 90 nt on 8M urea 10% polyacrylamide gels to remove extra single-stranded linker DNA and a 67 nt ligation byproduct, consisting of one copy of the hexamer and one copy of the linker DNA. After recovery using the crush-soak method, the purified ligation product was dissolved in 10 μL RNase-free water. PCR amplification (20 cycles) was performed using a primer specific to the single-stranded linker DNA and fused with an Illumina TruSeq Universal Adapter: 5′AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTGAACAGCGACTAGGCTCTTCA3′ (the sequence to prime single-stranded linker DNA is underlined and also needs to be trimmed from sequencing reads), and 12 different Illumina TruSeq Index Adapter reverse complementary primers:

Adapter	Sequences
Index 4	5′CAAGCAGAAGACGGCATACGAGAT**TGGTCA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC3′
Index 7	5′CAAGCAGAAGACGGCATACGAGAT**GATCTG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC3′
Index 21	5′CAAGCAGAAGACGGCATACGAGATTC**CGAAAC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC3′
Index 25	5′CAAGCAGAAGACGGCATACGAGATAT**ATCAGT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC3′
Index 2	5′CAAGCAGAAGACGGCATACGAGAT**ACATCG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC3′
Index 18	5′CAAGCAGAAGACGGCATACGAGATGT**GCGGAC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC3′
Index 1	5′CAAGCAGAAGACGGCATACGAGAT**CGTGAT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC3′
Index 19	5′CAAGCAGAAGACGGCATACGAGATCG**TTTCAC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC3′
Index 20	5′CAAGCAGAAGACGGCATACGAGATAA**GGCCAC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC3′
Index 22	5′CAAGCAGAAGACGGCATACGAGATTA**CGTACG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC3′
Index 23	5′CAAGCAGAAGACGGCATACGAGATAT**CCACTC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC3′
Index 27	5′CAAGCAGAAGACGGCATACGAGATAA**AGGAAT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC3′

The product was run on an 8M urea 10% polyacrylamide gel for DNA size separation to remove primer dimers and further eliminate byproduct contamination. DNA between 200 bp and 600 bp was collected by reference to both an Ultra Low Range DNA Ladder (Cat# SM1213, ThermoFisher) and a Low Range DNA Ladder (Cat# SM1193, ThermoFisher). Library DNA size distribution and consistency between biological replicates was assessed by Agilent 2100 Bioanalyzer (Agilent Technologies). After qPCR to quantify the library molarity, a pool of all libraries at equal molarity was made, and libraries were subjected to next-generation sequencing on an Illumina HiSeq 2500 at the Genomics Core Facility of the Penn State University to generate 150 nt single-end reads. The Strucutre-seq2 raw sequencing reads are available at the Gene Expression Omnibus (GEO) at the National Center for Biotechnology Information (NCBI) with the series entry GSE100714.

RNA-seq library preparation and sequencing

To impose the same heat shock as in the Structure-seq experiment, two week old rice plants in pots were inverted and the shoots were immersed in a water bath at 22 °C or 42 °C for 10 min treatment, and the plants were then transferred to a growth chamber set at the same temperature as in the greenhouse (30 °C) for ease of sampling during the recovery period. Three rice shoots comprised one biological replicate, and two biological replicates were obtained for each treatment and time-point, as indicated in Figure 1a. Total RNA was extracted from each sample using the NucleoSpin RNA plant kit (Cat# 740949, Macherey-nagel). After examination of the quantity and quality of these RNA samples by NanoDrop 2000 (Thermo Fisher Scientific, USA) and Bioanalyzer 2100 (Agilent Genomics, USA), total RNA samples were sent to the Genomics Core Facility at Penn State University for RNA-seq library preparation and next generation sequencing (Hiseq 2500, Illumina). Approximately 40-50 million 150 bp single-end sequencing reads were obtained for each library.

Ribosome profiling library preparation and sequencing

To test the effect of heat on ribosome footprinting, two-week-old rice plants were grown under the same conditions as described for Structure-seq probing. Ten shoots were harvested at 10 min. as described above for the RNA-seq time course experiment. Isolation of RPFs (ribosome protected fragments) and library construction were performed as described in the literature (5) with some major changes. Rice shoots were ground into powder with liquid nitrogen. For each sample, two mL of tissue powder was dissolved and homogenized in 10 mL polysome extraction buffer on ice. The buffer contains 200 mM Tris-Cl (pH 8.0), 100 mM KCl, 25 mM $MgCl_2$, 5 mM DTT, 1 mM PMSF, 100 μg/mL cycloheximide, 1% Brj-35, 1% TritonX-100, 1% Igepal CA630, 1% Tween-20, 1% polyoxyethylene 10 tridecyl ether. After centrifugation at 16 000 *g* for 10 min at 4 °C, the supernatant was collected. The supernatant was then layered on top of an 8 mL sucrose cushion (1.75 M sucrose in 200 mM Tris (pH 8.0), 100 mM KCl, 25 mM $MgCl_2$, 5 mM DTT, 100 μg/mL cycloheximide), and centrifuged at 170 000 *g* at 4 °C for 3 h. The pellet was resuspended in 400 μL RNase I digestion buffer (50 mM Tris-Cl (pH 8.0), 100 mM KCl, 20 mM $MgCl_2$, 1 mM DTT and 100 μg/mL cycloheximide). After adding 20 μL RNase I (Cat# AM2294, Thermo Fisher), RNase digestion was performed at room temperature with rotation for 2 h. TRIzol reagent (Cat# 15596026, Thermo Fisher) was used to extract the RPFs followed by fragment size selection using a NucleoSpin miRNA kit (Cat# 740971, Macherey Nagel) to collect the fragments smaller than 200 nt. A Urea-PAGE gel (10%) was then applied to size select 28-32 nt fragments. After dephosphorylation using PNK (Cat# M0201S, NEB), the RPFs were ligated to AIR adenylated RNA linker (Cat# 510201, BIOO Scientific). The ligation products were then subjected to reverse transcription using SuperScript III (Cat# 18080093, Thermo Fisher) and circularization using Circligase II (Cat# CL9021K, Illumina). Sequence libraries were ultimately obtained through PCR amplification by Q5 polymerase (Cat# M0491S, NEB). The resultant ribosome profiling libraries were sequenced at the Genomics Core Facility at Penn State University to generate single-end 100 nt reads.

Sequence mapping and treatment

FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) software was used to check the quality of the sequencing reads. To remove the adapters at both ends of the reads, cutadapt (6) was employed. Any reads shorter than 20 nt or with a quality score < 30 (-q flag of cutadapt) were discarded. Reads were then mapped to rice reference cDNA and rRNA libraries using Bowtie2 (7). Reads with more than 3 mismatches or a mismatch on the first nucleotide at the 5′ end were discarded. As we obtained a high correlation between the three biological replicates in each condition, replicates were combined for further analysis.

Determination of DMS reactivity

The method employed to derive DMS reactivity on each nucleotide was similar to that used in our previous studies (1, 2, 8, 9) with additional steps of normalization between the different temperature conditions. The steps to calculate DMS reactivity from (–) DMS and (+) DMS libraries are as follows:

Step 1. Normalization of RT stop counts. For each transcript, the RT stop counts on each nucleotide are incremented by 1 and then the natural log (*ln*) is taken, followed by normalization by the transcript's abundance and length (Equation 1 and 2).

$$P(i) = \frac{ln[P_r(i)+1]}{\{\sum_{i=0}^{l} ln[P_r(i)+1]\}/l} \quad (1)$$

$$M(i) = \frac{ln[M_r(i)+1]}{\{\sum_{i=0}^{l} ln[M_r(i)+1]\}/l} \quad (2)$$

Here, $P_r(i)$ and $M_r(i)$ are the raw 'r' numbers of RT stops mapped to nucleotide i (all four nucleotides are included) on the transcript in the plus (P) and minus (M) reagent libraries, respectively, and l is the length of the transcript. $P_r(0)$ and $M_r(0)$ are the raw numbers of 5′-runoff RT reads.

Step 2. Calculation of DMS reactivity. The raw DMS reactivity is calculated by subtracting the normalized RT stop counts between (+) DMS and (–) DMS libraries with all negative values set to 0. For each nucleotide I, the DMS reactivity is calculated as follows:

$$\theta(i) = \max[P(i) - M(i), 0] \quad (3)$$

Step 3. Normalize the raw DMS reactivity $\theta(i)$ of all the nucleotides on all the transcripts to obtain the derived DMS reactivity of each nucleotide as described below. In order to make account for the greater intrinsic reactivity of the DMS at 42 °C, the normalization process is performed differently on the two conditions.

a. 22 °C
Perform 2%/8% normalization(10) on the raw DMS reactivity $\theta(i)$ of all the nucleotides on all the transcripts to obtain the derived DMS reactivity of each nucleotide, with the normalization scale derived from the 2%/8% normalization of each transcript. Here, the normalization scale is the average of the bottom four-fifths (80%) of the top 10% of the nucleotide reactivity values on each transcript.

b. 42 °C
Perform normalization on the raw DMS reactivity $\theta(i)$ of all the nucleotides on all the transcripts using the normalization scale from the 22 °C condition of each transcript to obtain the final DMS reactivity of each nucleotide.

The normalized reactivity is capped at 7 (11).

Step 4 Normalize DMS reactivities between conditions to obtain the final reactivity. Suppose $\theta_{heat}(i)$ and $\theta_{rt}(i)$ are reactivities at 42 °C and 22 °C for nucleotide i after step 3. Final reactivities are derived as follows:

$$\theta_{final(heat)}(i) = \theta_{heat}(i) \cdot c_{heat} \quad (4)$$
$$\theta_{final(rt)}(i) = \theta_{rt}(i) \cdot c_{rt} \quad (5)$$

Here,

$$c_{heat} = \frac{(\theta_{rt} + \theta_{heat})}{2\theta_{heat}}, \qquad c_{rt} = \frac{(\theta_{rt} + \theta_{heat})}{2\theta_{rt}},$$

$$\theta_{rt} = \sum_{i \in S} \theta_{rt}(i), \qquad \theta_{heat} = \sum_{i \in S} \theta_{heat}(i)$$

S is the set of all nucleotides on all RNAs with coverage ≥ 1 at 22 °C and 42 °C.

RNA-seq library data analysis

After sequencing, adapter contamination was computationally removed from the libraries and adapter sequences were trimmed from the 3′ ends of the raw reads using cutadapt (6). Low-quality bases (Q < 30) were also trimmed from both the 5′ and 3′ ends of the reads. Next, reads from each of the four libraries were mapped independently to the rice genome (IRGSP-1.0) using STAR (12), with a GTF (Gene Transfer File) annotation file supplied as an argument. Mapping information is provided in Supplementary Table S3. Transcript abundance and differential gene expression were calculated using DEseq2 (13). TPM (transcripts per million)-based gene expression levels were generated for downstream analysis. The RNA-seq raw sequencing reads are available at the Gene Expression Omnibus (GEO) at the National Center for Biotechnology Information (NCBI) with the series entry GSE100713.

Analysis of the degradome dataset

We downloaded the supplementary file from GEO accession GSM1040649, rice degradome data under 28 °C from ZH11 WT plants, GSM1040649_ZH11.fa.gz. We mapped the fragment sequences to the rice transcriptome (Oryza_sativa.IRGSP-1.0.30.cdna.all.fa) using Bowtie2, and used a custom Python script to combine the mapping results (.sam) with the fragment counts, generating a combined count of all degradome fragments per transcript. The degradome data of each transcript were merged with our calculated average reactivity data and imported into R. We used the correlation function (cor()) to test correlation between number of normalized fragments ($\log_2$(#fragments)/transcript length) and transcript reactivity at both temperatures. We then used the quantile function to subset the data into the 5% highest and 5% lowest average transcript reactivity groups and then compared the mean number of fragments in each of these groups using two-tailed Student's t-test. To compare the shape of the distribution from each group (abundance increases or decreases) we used the Matching package (14) to run a bootstrapped KS test (boot.ks, nboots=4000) between the increased and decreased distributions at each respective time point.

Motif analysis

Sequences and reactivity values for 3'UTR regions of transcripts were extracted from the whole transcript sequence and reactivity data. We searched for all instances of the UUAG motif within the 3'UTR of transcripts with coverage over one and cataloged the reactivity change within the UUAG motif via the react_static_motif.py (SF2) module (9). The 3'UTR regions of transcripts with coverage over one were then subdivided via a sliding window analysis into windows of 50 nt by 20 nt steps and ranked by total increase and decrease of reactivity via the react_windows.py (SF2) module (9). Fasta formatted files corresponding to the top and bottom 1% of reactivity increases and decreases among these windows were used as the input to MEME suite analysis. We compared our discovered enriched motifs to the protein-binding motifs published in (15).

Ribosome profiling data analysis

To calculate ribosome association and its modulation by temperature, the adapter 5′-ACTGTAGGCACCATCAAT-3′ at the 3′ end of the reads was first removed using cutadapt. Any reads shorter than 20 nt or longer than 40 nt or with a quality score < 30 (-q flag of cutadapt) were discarded. Reads were then mapped to the rice reference genome and cDNA libraries using Bowtie2. Since we obtained a high correlation between the 2 biological replicates in each condition, replicates were combined for further analysis. Ribosome association in each condition was derived using the resultant ribosome profiling library, with the RNA-seq library at 10 min as the control library. Read depth of each nucleotide on each RNA was normalized by the total number of reads in each library and then the natural log (ln) was taken on the normalized read depth. The Ribo-seq signal of each nucleotide was calculated by subtracting the natural log of the normalized read depth of each nucleotide in the RNA-seq library from that in the ribosome profiling library. The Ribo-seq signal per transcript is the average of the value of all nucleotides in the transcript. The change in Ribo-seq signal was calculated by subtracting the average Ribo-seq signal in heat (42 °C) from that in the control condition (22 °C). The Ribo-seq raw sequencing reads are available at the Gene Expression Omnibus (GEO) at the National Center for Biotechnology Information (NCBI) with the series entry GSE102216.

Optical Melting

As is standard for analyses of optical melting, RNA was denatured at 95 °C for 90 sec in water and then allowed to refold at 4 °C for 90 sec, then room temperature for 5 min. After the 5 min, buffer was adjusted to 40 mM HEPES pH 7.5, 100 mM KCl, and 0.5 mM Mg^{2+}, and allowed to equilibrate at room temperature for 10 min. Samples were spun down at 14,000 rpm for 5 min at room temperature to remove air bubbles and particulates, then transferred to a quartz cuvette. Final sample concentrations were 1.1 μM RNA. The transitions for T2 and T3 were confirmed to be independent over a range of concentrations from 0.55 μM to 5.5 μM, supporting that transition is from the hairpin rather than duplex state.

Name	Transcript ID	Description	Sequence
T1	OS06T0105350-00	Similar to Scarecrow-like 6	GCGATGCTAGAAAAAAAAAAAAAA
T2	OS02T0662100-01	Similar to Tfm5 protein	TTTTAGATTGAAAAAAAAAAAAAA
T3	OS03T0159900-02	Hypothetical conserved gene	ATTAATTTTTAAAAAAAAAAAAAA
T4	OS02T0769100-01	Auxin responsive SAUR protein family protein	CACATTTTATAAAAAAAAAAAAAA

Thermal denaturation experiments were performed on an HP 8452 diode-array refurbished by OLIS, Inc. with a data point collected every 0.5 °C with absorbance detection from 200-600 nm. Data at 260 nm were converted to fraction folded assuming linear baselines.

mRNA decay analysis

mRNA decay rate determination was performed by following a previously described method (16) with modifications. The conditions of rice seedling growth were the same as for our other genome-wide assays. After 13 days of growth, rice seedlings were gently removed from the soil, carefully washed to remove dirt from the root tissue and transferred to tap water to recover for 1 day, similar to the method of Park et al. (2012). Cordycepin solution with a final concentration of 1 mM was prepared in tap water and equilibrated at the prevailing temperature in the greenhouse (30 °C) before treatment. Rice seedlings were then transferred to cordycepin solution, with the roots immersed, and pretreated for 30 min before the start of temperature treatment. For temperature treatment, 1 mM cordycepin solutions were prepared before use and equilibrated

in a water bath for 42°C treatment and on the lab bench for 22°C treatment. After the 30 min. pretreatment, seedlings were quickly transferred to either 42°C cordycepin solution for heat treatment or 22°C solution for room temperature (control) treatment, for 10 min. This protocol followed the identical protocol as used to obtain the Structure-seq and ribo-seq 10 min datasets (cf. Fig. 1A). The seedlings were then transferred back to the 30°C cordycepin solution and placed in a 30°C growth chamber for recovery, identical to the recovery protocol used for the RNA-seq timecourse (cf. Fig. 1A). Plant materials were sampled at the end of the cordycepin pretreatment as control sample (C0), then immediately after 10 min of the two temperature treatments (H10m and C10m), then after 50 min of "heat recovery" (HR) in the growth chamber (HR1h and C1h). Three biological replicates were prepared for each sample.

After RNA extraction using the RNA Plant kit (Cat# 740949, Macherey-nagel), cDNA was synthesized using the SuperScript III first-strand synthesis system (Cat#: 18080051, ThermoFisher). qRT-PCR analysis was performed using a Bio-Rad real-time PCR detection systems with SYBR Green Supermix (Cat. #. 1708880, Bio-Rad). Primers used for real-time PCR analysis are listed in Supplementary Table S3. qRT-PCR was performed using the following protocol: 95°C for 5 min, followed by 49 cycles of 95°C for 20 s, 53°C for 20 s, and 72°C for 30 s, and then melting curve analysis (60°C–95°C at a heating rate of 0.1°C/S). qRT-PCR was performed in triplicate for each cDNA sample. Using rice *Ubiquitin1* (*Ubi1*, Os06g0681400) as the internal control, the relative abundance of each transcript at each time point was normalized by *Ubi1* abundance within the same sample. Relative decay post temperature treatment was then normalized by comparison to the relative abundance at the 0 min time point. Relative decay was plotted as a line graph to show the trend of change in remaining mRNA abundance.

To identify candidate rice XRN targets in rice, we first used the most reliable XRN targets list in Arabidopsis, designated "Class II" by (Merret et al., 2015), then we consulted the PLAZA database to identify the best BLAST hits from *A. thaliana* to *O. sativa*. Every *O. sativa* ortholog of each *A. thaliana* XRN responsive gene was converted from MSU format to Ensemble format before use in our data analyses.

Data and Software Availability

Next-generation sequencing data for our structure-probing libraries have been deposited at the NCBI Gene Expression Omnibus and are available under the series entry GSE100714, The DMS reactivity files containing all the single nucleotide reactivity scores can be found in the supplementary files in this accession. RNA-seq data have been deposited under the series entry GSE100713. Ribo-seq data have been deposited under the entry GSE102216. Supplementary Dataset S1 contains all the analyzed data (and associated *P* values) for Structure-seq, Ribo-seq, and RNA-seq datasets. Our RNA structure analysis code has been deposited at github (https://github.com/StructureFold/StructureFold) and (https://github.com/StructureFold2/StructureFold2).

Figs. S1 to S11

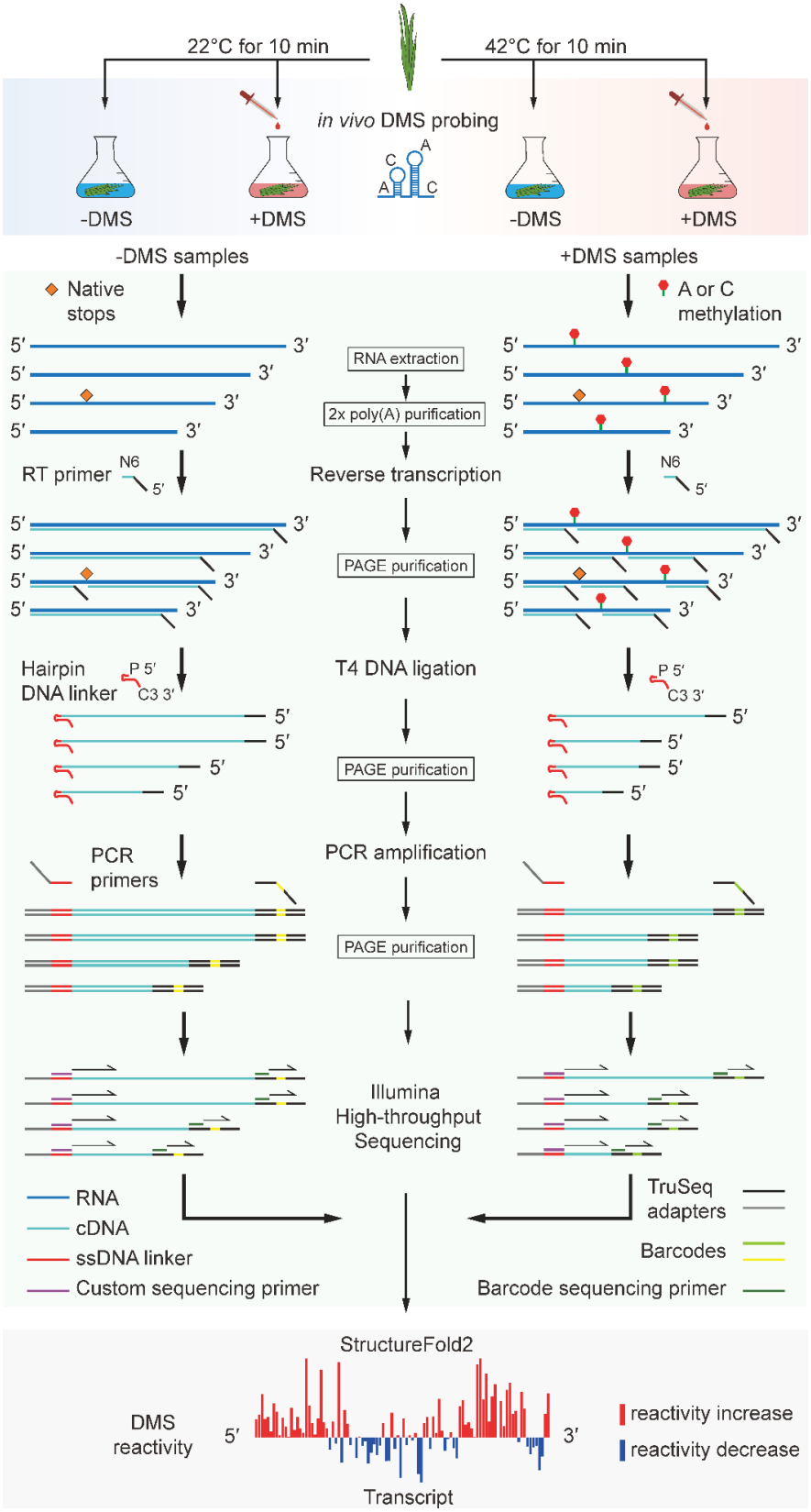


Fig. S1. Workflow of temperature treatment and rice library construction using Structure-seq2.

Two-week-old rice shoots were treated with DMS (+DMS sample) for 10 min at 22 °C or 42 °C. DMS covalently modifies single-stranded As and Cs. These modifications cause reverse transcription to stop one nucleotide before the modification; occasional native RNA modifications or strong *in vitro* RNA structure can also cause stops, which are accounted for using control (−DMS) libraries. Random hexamers (N6) with a TruSeq adaptor were employed for reverse transcription. DNA ligation was performed using T4 DNA ligase, which can ligate a hairpin DNA linker donor to the 3′ end of cDNAs. Library amplicons were then generated by PCR using Q5 high fidelity polymerase. Urea polyacrylamide gel electrophoresis (Urea-PAGE) was used for all DNA purifications. Illumina MiSeq sequencing was used for library quality determination and Illumina HiSeq sequencing was used for final data generation. DMS reactivity at nucleotide resolution was generated using the StructureFold program. Boxes indicate steps in the current Structure-seq2 protocol (3) that are improvements from our original Structure-seq method (1, 2).

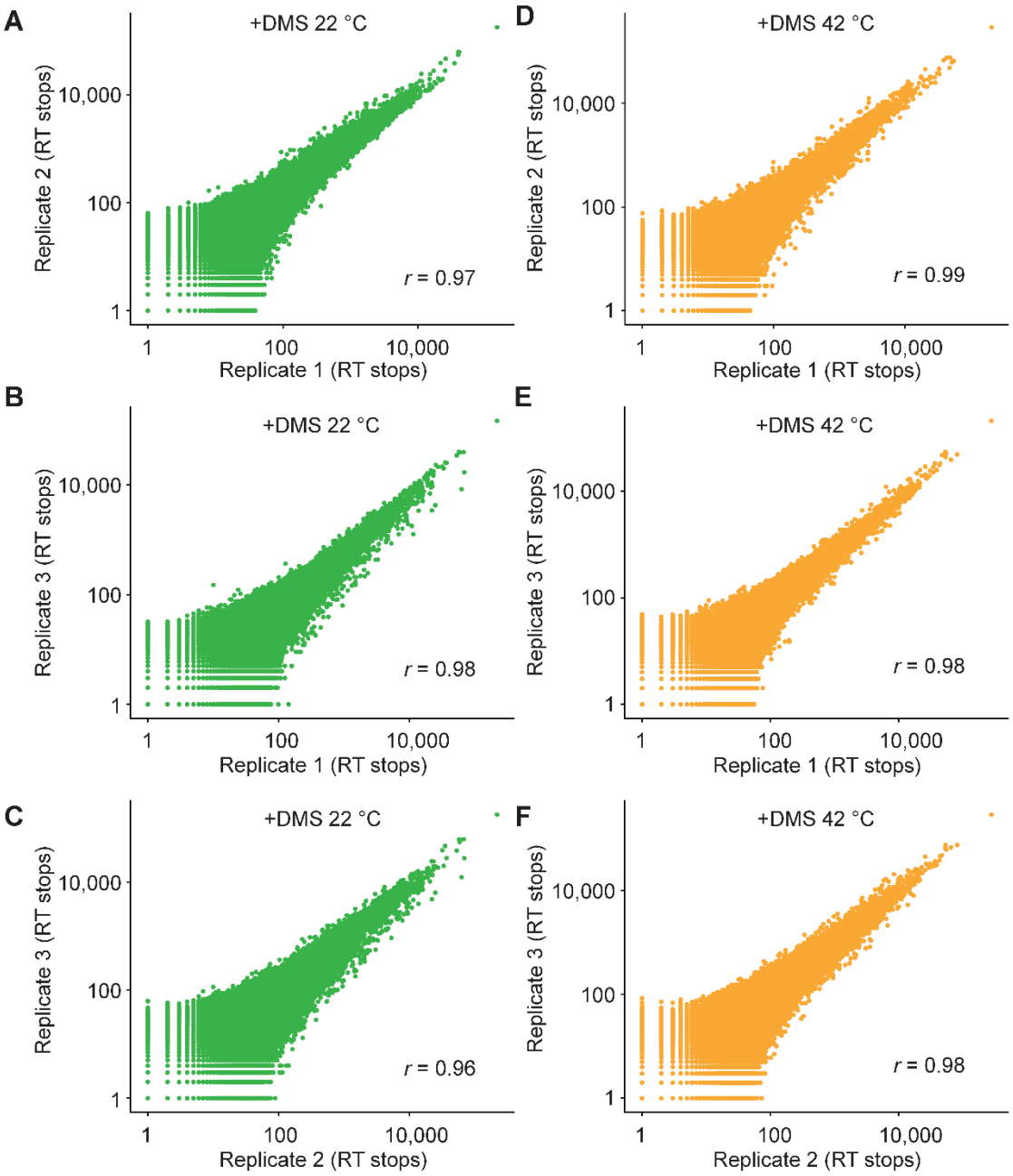
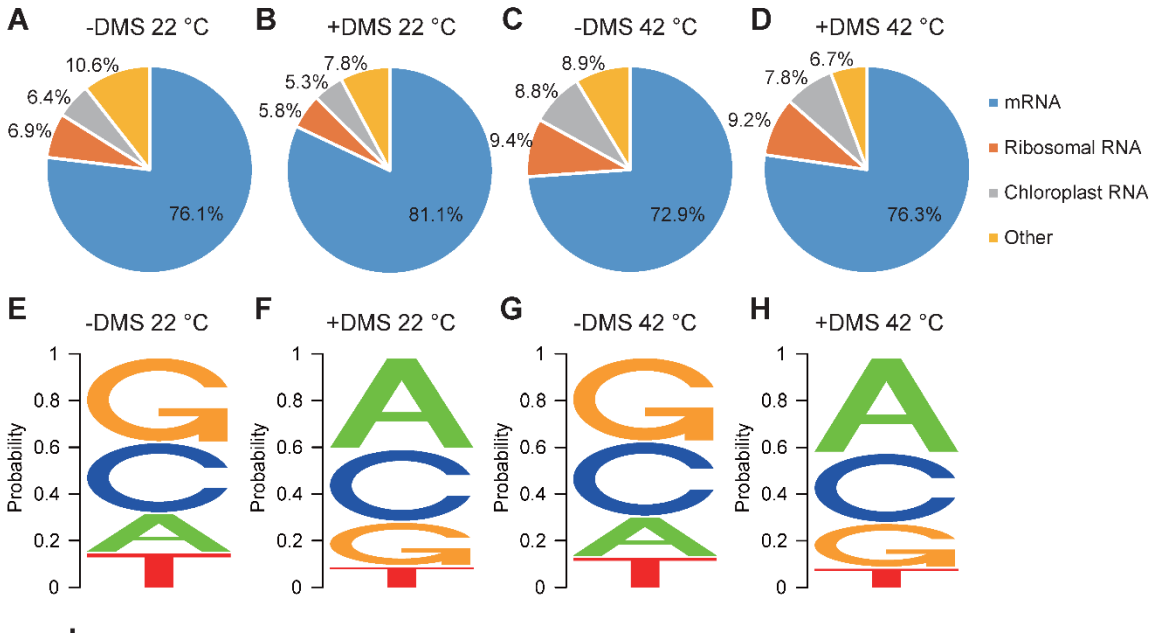


Fig. S2. High correlation of single nucleotide reverse transcription (RT) stop counts between replicates in +DMS libraries.

(A–C), Correlation between 3 biological replicates at 22 °C. (D–E), Correlation between 3 biological replicates at 42 °C. All of the biological replicates at each temperature are highly correlated.



I

	-DMS 22 °C	+DMS 22 °C	-DMS 42 °C	+DMS 42 °C
A	16.7%	40.3%	17.1%	41.9%
C	31.0%	31.8%	32.8%	30.6%
G	37.2%	18.9%	36.8%	19.1%
T	15.1%	9.1%	13.3%	8.4%

Fig. S3. The majority of Structure-seq reads are from mRNAs.

(A), −DMS library at 22 °C (136,504,440 total mapped reads). (B), +DMS library at 22 °C (152,310,815 total mapped reads). (C), −DMS library at 42 °C (125,305,132 total mapped reads). (D), +DMS library at 42 °C (141,636,436 total mapped reads). (E–H): Nucleotide modifications in the +DMS libraries are specific to As and Cs. (E) −DMS library at 22 °C. (F) +DMS library at 22 °C. (G) −DMS library at 42 °C. (H) +DMS library at 42 °C. (I) +DMS libraries show greater modification of A and C than of U and G.

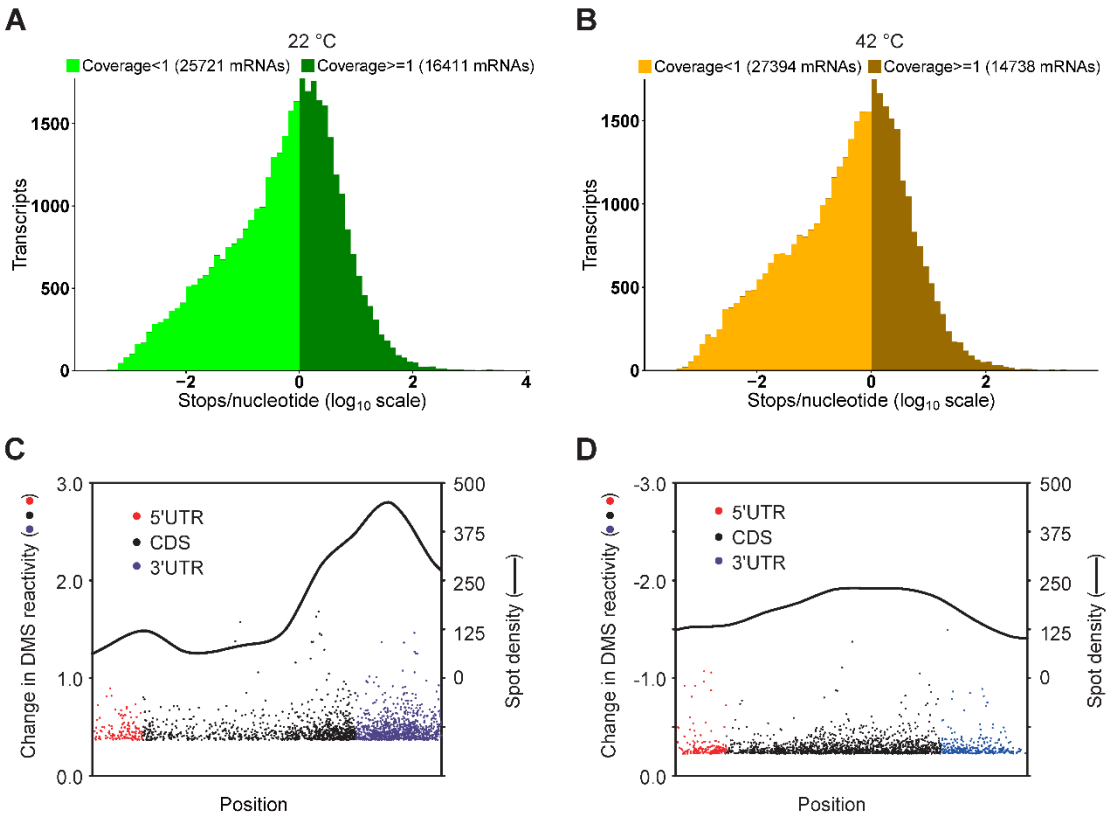


Fig. S4. Distribution of structure-probing coverage, and 3′UTRs show greatest heat-induced change in DMS reactivity (42 °C- 22 °C).

(A) Distribution of coverage of all transcripts in Structure-seq datasets at 22 °C. Structure-seq provided structural information at nucleotide resolution on 16,411 RNAs with coverage over 1 at 22 °C. (B) Distribution of coverage of all transcripts in Structure-seq datasets at 42 °C. Structure-seq provided structural information at nucleotide resolution on 14,738 RNAs with coverage over 1 at 42 °C. Lengths of regions (5′ UTR, CDS, 3′UTR) on each mRNA were normalized and aligned for plotting. Red indicates 5′UTR, black indicates CDS, and blue indicates 3′UTR (Zero value is included for clarity, as indicated). (C) Distribution of the 2,000 spots with the most elevated DMS reactivity at 42 °C as compared to 22 °C (change in DMS reactivity; left axis). A 'spot' is defined as average reactivity in a 100 nt window. The 2,000 spots were identified solely based on reactivity change, independent of location on the mRNA. Distribution shows enrichment of the hot spots in 3′UTRs. Line shows the distribution of the total number of spots (spot density; right axis) along each normalized region, for the 1,170 mRNAs harboring the 2,000 spots. (D) Distribution of the 2,000 spots with the most reduced DMS reactivity at 42 °C as compared to 22 °C. Line shows the distribution of the total number of spots along each normalized region, for the 982 mRNAs harboring the 2,000 spots.

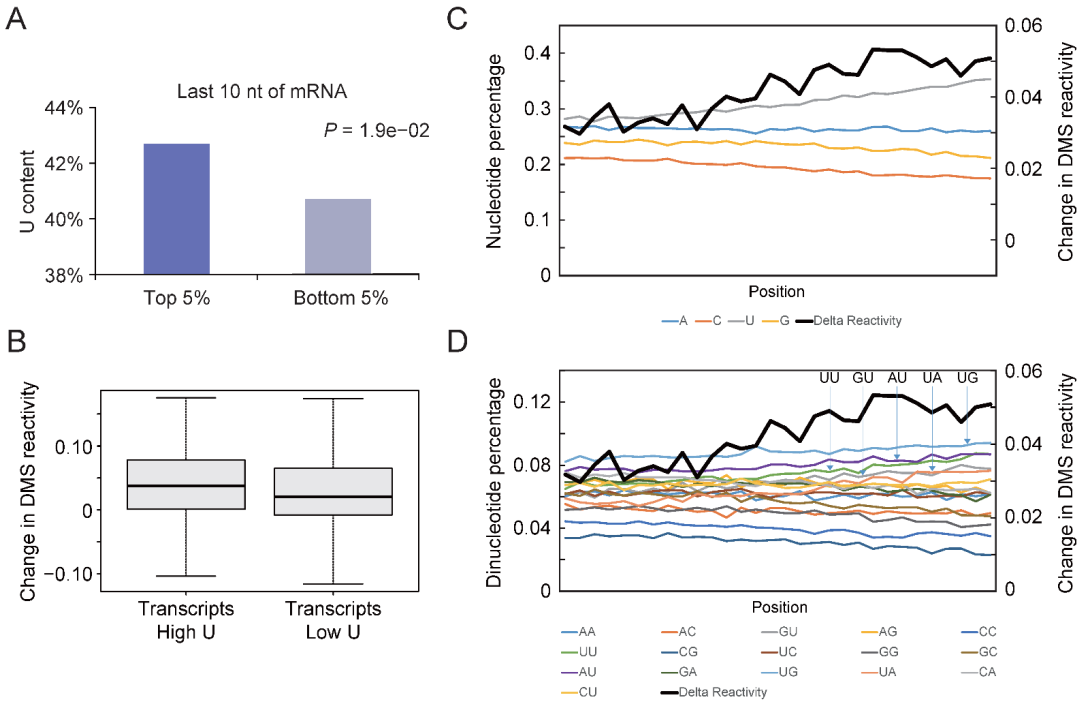


Fig. S5. Correlations between U and AU content at the 3′ ends of 3′UTRs and heat-induced DMS reactivity changes

(A) U content of the last 10 nt at the 3′ end of the 5% of mRNAs with most elevated (Top 5%) or reduced (Bottom 5%) DMS reactivity at 42 °C as compared to 22 °C. (B) Transcripts with high U content (≥8) in the last 10 nt of the 3′UTR showed significantly higher heat-induced change in average DMS reactivity of the entire 3′UTR than the ones with low U content (≤3) in the last 10 nt of 3′UTR ($P = 0.03$). (C) Single nucleotide frequency (left y-axis) and (D) dinucleotide frequency (left y-axis) and DMS reactivity change (42 °C - 22 °C; right y-axis) along the 3′UTRs. Nucleotide frequencies and DMS reactivities are binned into 40 bins (10 nt per bin). The UTR region depicted excludes the very 3′ end where DMS reactivity data do not meet the minimum coverage requirement. The five most common dinucleotides near the 3’ end are UU, GU, AU, UA, and UG (annotated), suggesting that melting of AU and GU base pairing may contribute to enhanced DMS reactivity under heat.

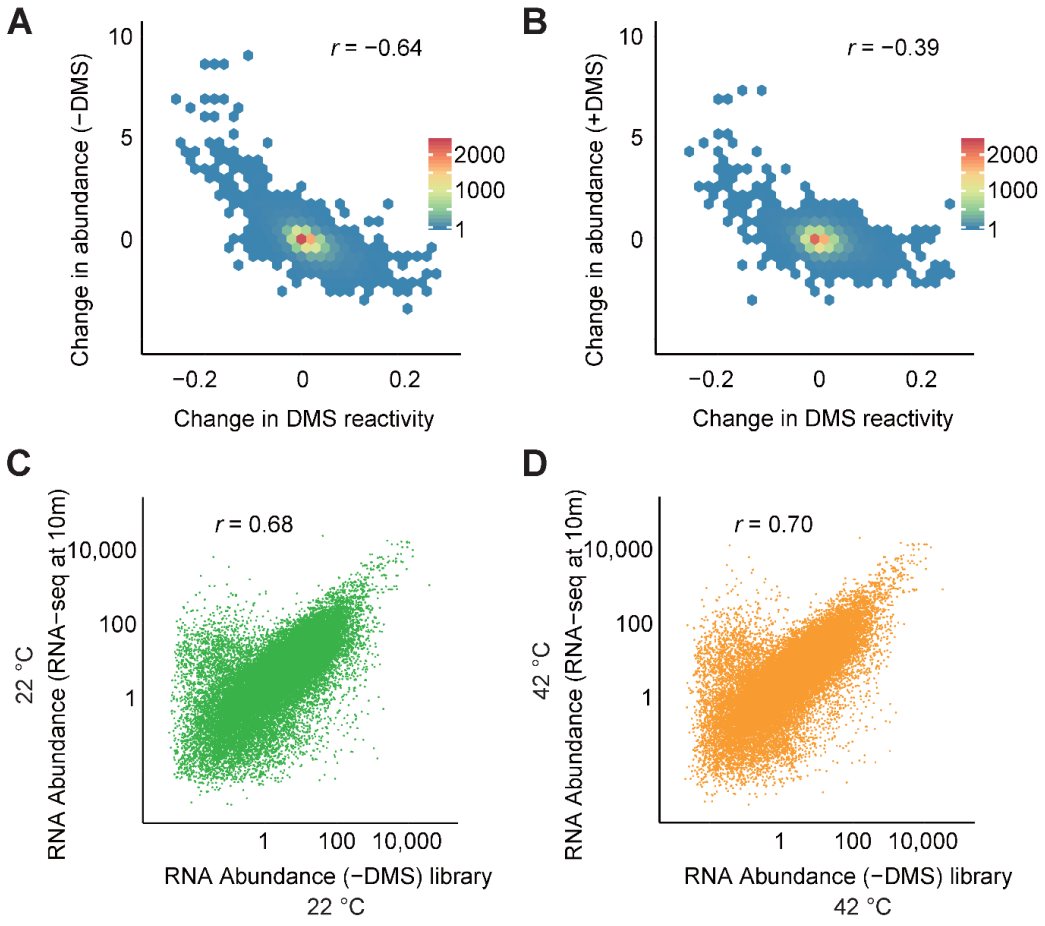


Fig. S6. Negative correlation between DMS reactivity and mRNA abundance change as measured from DMS Structure-seq libraries, and high correlation of mRNA abundance between Structure-seq and RNA-seq libraries.

(A, B) Negative correlation between change of average DMS reactivity (42 °C−22 °C) and RNA abundance change (42 °C−22 °C), measured from Structure-seq libraries as $\log_2$(TPM) at 22 °C and 42 °C for the 14,292 mRNAs with coverage above 1 in Structure-seq analysis. Colors indicate numbers of mRNAs. (A) –DMS libraries. (B) + DMS libraries. (C, D) Strong positive correlation between mRNA abundance as calculated from Structure-seq −DMS libraries and mRNA abundance as calculated from RNA-seq 10 min libraries at (C) 22 °C and (D) 42 °C.

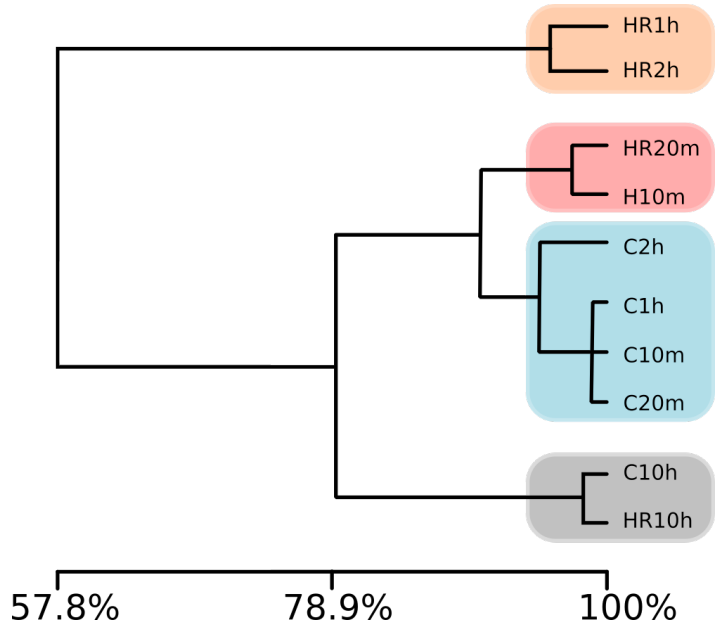


Fig. S7. Hierarchical clustering of RNA-seq datasets indicates the relationships of the samples and the recovery of the transcriptome following 10 min of 42 °C heat shock.

C=control, H=heat shock for 10 min, HR=heat recovery. Scale indicates transcriptome percent similarity between samples. The tree was generated using MEV software (http://mev.tm4.org). TPM-based RNA-seq timecourse datasets were analyzed using hierarchical clustering to show the relationship between the samples.

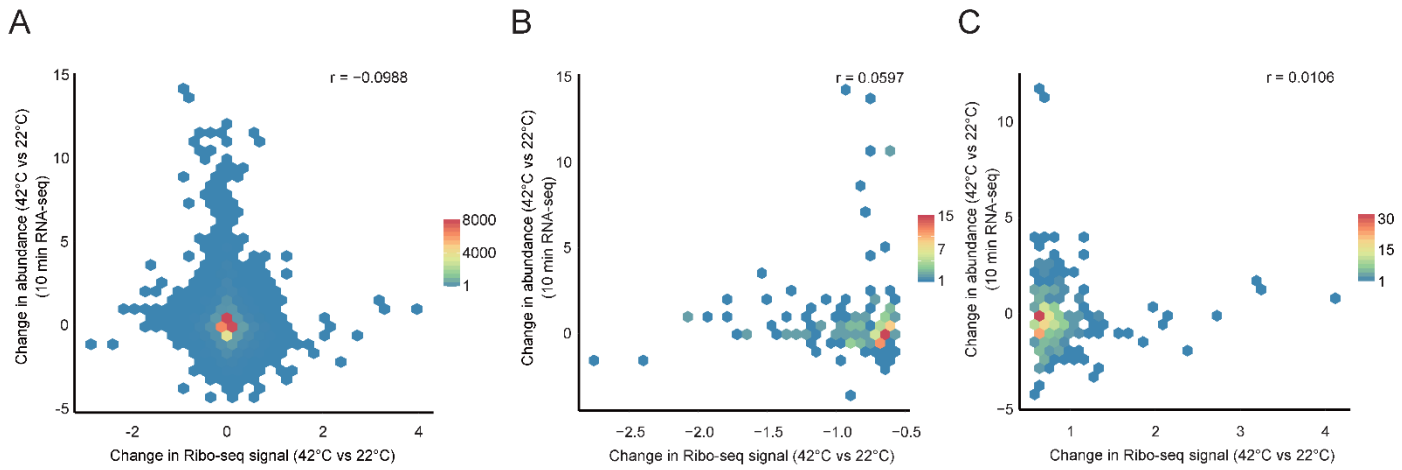


Fig. S8. No correlation detected between heat shock induced change in Ribo-seq signal (42 °C−22 °C) and mRNA abundance change between 42 °C and 22 °C at 10 min (= end of 42 °C treatment).

(A) Correlation of abundance change with Ribo-seq signal change for the whole transcripts. (B) Correlation of the transcripts with 1.5 fold decrease in Ribo-seq signal ($\log_2$(ribo-seq signal) < −0.58) (zoom-in of left-hand portion of panel A). (C) Correlation of the transcripts with 1.5 fold increase in Ribo-seq signal ($\log_2$(ribo-seq signal) > 0.58) (zoom-in of right-hand portion of panel A).

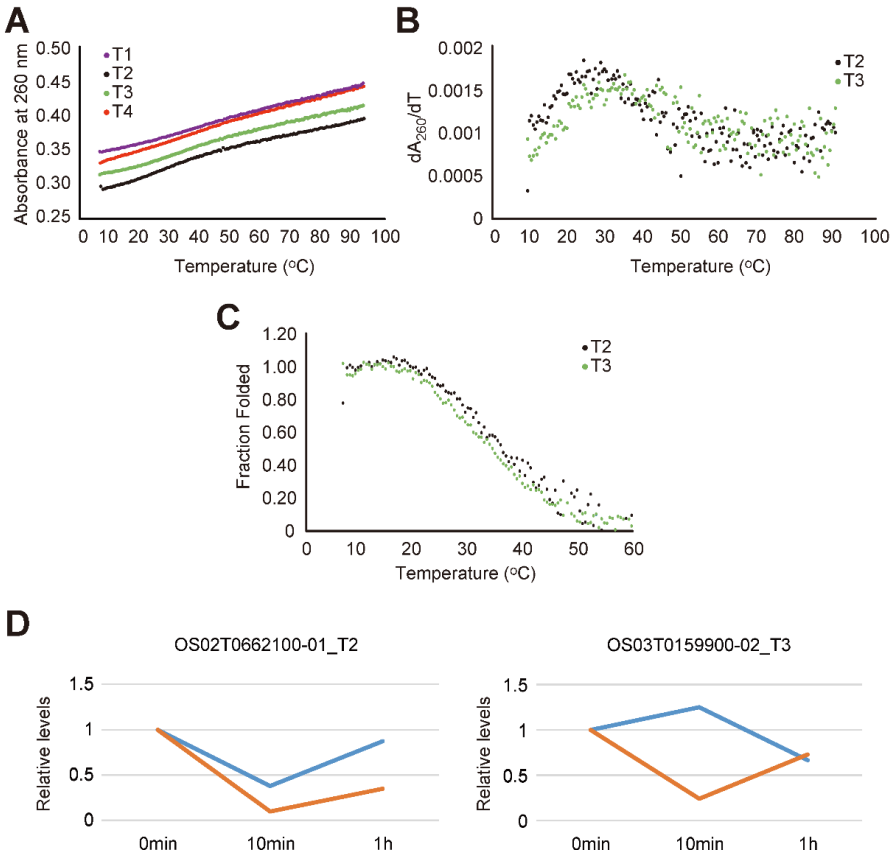


Fig. S9. 3′-end + A15 polyA tail RNA unfold in the range of heat treatment and the mRNAs of T2 and T3 decay faster under heat.

(A) Raw melts of four candidate RNAs from the top 5% that lose abundance under heat treatment. Sloping baselines are likely due to the 15 A's unstacking, given the tendency of polyA to stack (17). T1 (purple), T2 (black), T3 (green), and T4 (red). (B) Derivatives of the optical melting data from T2 and T3, which show appreciable sigmoidal characteristic in panel A. (C) Fraction folded of T2 and T3. Fraction folded is calculated from the equation Fraction Folded=$(A-A_u)/(A_f-A_u)$, where A is the absorbance at a given temperature, A_u is the absorbance of the unfolded RNA which is calculated from the linear fit of the upper baseline, and A_f is the absorbance of the folded RNA which is calculated from the linear fit of the lower baseline. Sequences were derived from the following genes: T1 (OS06T0105350-00) *Similar to Scarecrow-like 6*; T2 (OS02T0662100-01) *Similar to Tfm5 protein*; T3 (OS03T0159900-02) *Hypothetical conserved gene*; T4 (OS02T0769100-01) *Auxin responsive SAUR protein family protein*. See Materials and Methods for specific sequences and methodological details. (D) RNA decay rate analysis of T2 and T3 under two temperature conditions (42 °C (orange) vs 22 °C (blue)) in the presence of cordycepin shows accelerated decay at 42 °C.

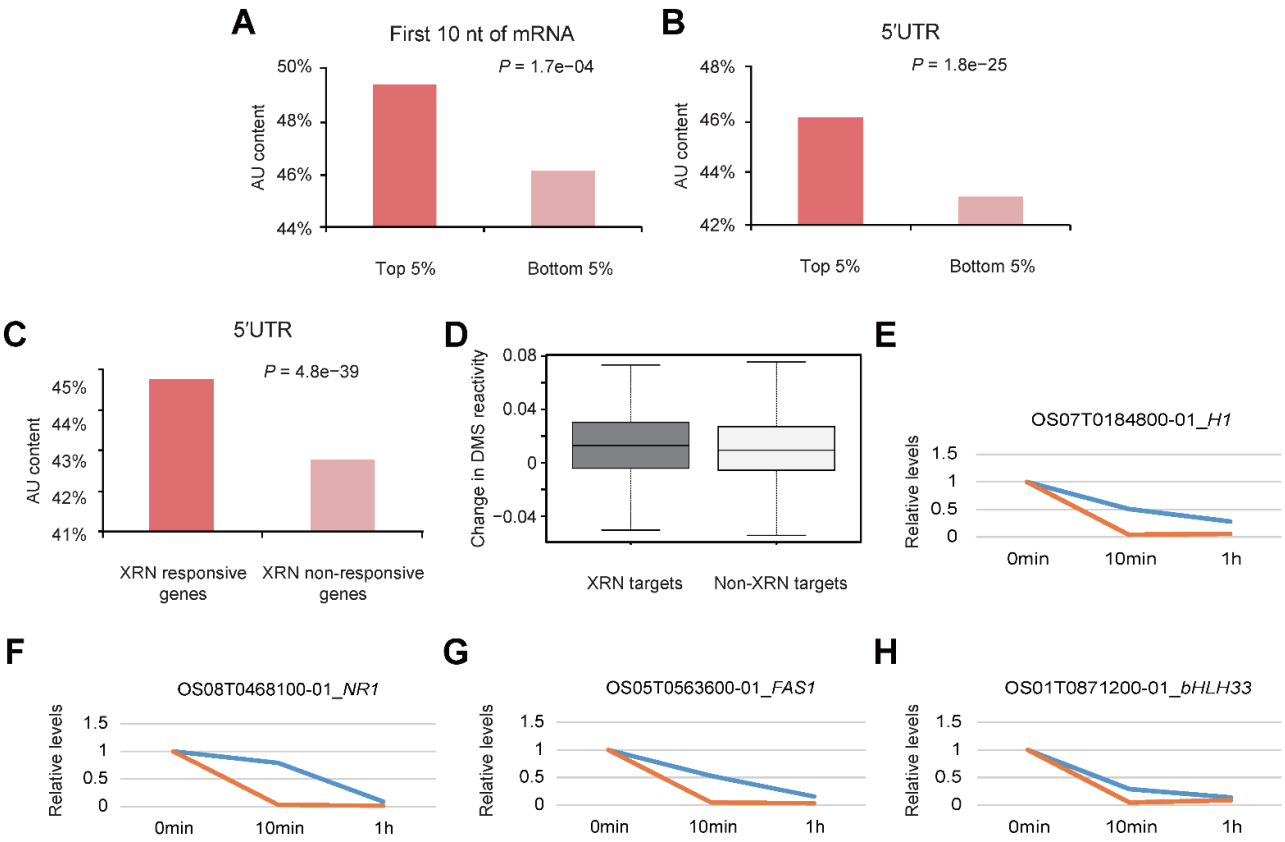


Fig. S10. AU content and U content at the 5′ end are significantly different between top 5% and bottom 5% of mRNAs; XRN targets show significantly higher 5′UTR AU content and DMS reactivity change (42°C−22°C) than non-XRN targets and decay rapidly under heat.

(A) AU content of the first 10 nt at the 5′ end of the 5% mRNAs with most elevated (Top 5%) or reduced (Bottom 5%) DMS reactivity at 42 °C as compared to 22 °C. (B) AU content of the 5′UTRs of the 5% of mRNAs with most elevated (Top 5%) and reduced (Bottom 5%) DMS reactivity at 42 °C as compared to 22 °C. (C) Higher AU content of the 5′UTRs of rice orthologs (derived from the MSU Rice Genome Annotation Project; http://rice.plantbiology.msu.edu/index.shtml) of mRNAs subject to heat-induced XRN4-mediated decay vs. XRN4 non-responsive mRNAs from the published datasets (18). *P* values are from Chi-squared tests. (D) Distribution of change in DMS reactivity of rice orthologs of XRN targets identified from (Merret et al., 2015) at 42°C as compared to 22°C. The average change in DMS reactivity (42°C compared to 22°C) of rice orthologs of XRN target mRNAs is significantly higher than that of mRNAs which are not XRN target orthologs.($p = 0.02$, two sample t-test). (E-H) mRNA decay rate analysis of XRN target transcripts under two temperature conditions (42 °C (orange) vs 22 °C (blue)) in the presence of cordycepin shows accelerated decay at 42 °C.

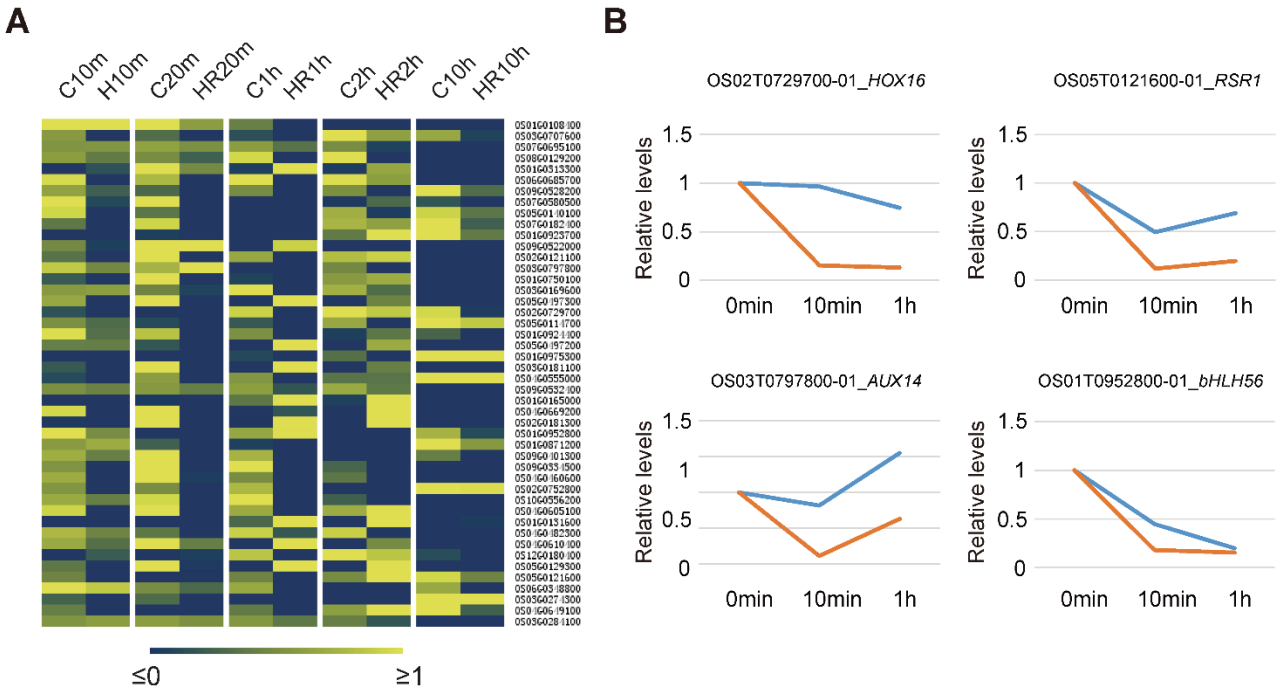


Fig. S11. mRNAs of transcription factors with increased DMS reactivity present in the top 5% group show decreased abundance post-heat shock, as compared to the control, and show accelerated heat-induced decay.

(A) mRNAs of transcription factors present in the top 5% of transcripts with increased DMS reactivity after heat shock show obvious heat shock-induced decreases in abundance over the time-course, especially at 10 and 20 min (H10 min and HR20 min), as compared to their abundance in the control (C10 min and C 20 min). Each expression value (Log_2(TPM)) was normalized by the average value of each row (i.e. the average expression value of that mRNA). In the heat map, blue represents low relative expression values ($(Log_2TPM)_{actual} - (Log_2TPM)_{average} \leq 0$) and yellow represents high relative expression ($(Log_2TPM)_{actual} - (Log_2TPM)_{average} \geq 1$). "HR" denotes recovery after heat shock. (B) mRNA decay analysis of transcription factors that showed increased DMS reactivity under heat. In the presence of cordycepin, these transcription factors show an accelerated decrease in mRNA abundance at 10 min after 42 °C treatment (orange) as compared to 22 °C (blue).

Tables S1 to S6

Table S1. Mapping statistics of Structure-seq libraries, generated using the Structure-seq2 protocol

Condition	Library replicate	All reads[a,b]	Genome mapped reads (GMR)	% of All	cDNA mapped reads (CMR)	% of GMR
−DMS 22 °C	1	58,686,867	51,393,158	87.6%	39,320,200	76.5%
	2	54,050,775	46,705,657	86.4%	35,187,054	75.3%
	3	43,985,174	38,405,625	87.3%	29,206,155	76.1%
	total	156,722,816	136,504,440	87.1%	103,713,409	76.0%
+DMS 22 °C	1	67,866,993	60,199,941	88.7%	49,203,569	81.7%
	2	44,979,760	39,058,142	86.8%	31,845,670	81.5%
	3	59,668,373	53,052,732	88.9%	42,483,881	80.1%
	total	172,515,126	152,310,815	88.3%	123,533,120	81.1%
−DMS 42 °C	1	49,519,857	43,477,224	87.8%	30,058,553	69.1%
	2	46,376,890	41,483,875	89.5%	29,681,209	71.6%
	3	45,514,706	40,344,033	88.6%	31,541,691	78.2%
	total	141,411,453	125,305,132	88.6%	91,281,453	72.8%
+DMS 42 °C	1	59,495,970	52,715,770	88.6%	40,814,235	77.4%
	2	43,235,010	38,268,283	88.5%	28,967,456	75.7%
	3	57,443,353	50,652,383	88.2%	38,257,273	75.5%
	total	160,174,333	141,636,436	88.4%	108,038,964	76.3%

[a]High quality (Q > 30) with length over 20 nt which is the minimum required for mapping in this study.
[b]Total of 630 million high quality reads in all libraries combined.

Table S2. *P* values for the comparisons of the average DMS reactivity and the average single-strandedness between 5′UTR, CDS, 3′UTR in Fig. 2 panels E to J.

	22 °C (Panel E)			42 °C (Panel F)			42 °C−22 °C (Panel G)		
	5′UTR vs 3′UTR	5′UTR vs CDS	3′UTR vs CDS	5′UTR vs 3′UTR	5′UTR vs CDS	3′UTR vs CDS	5′UTR vs 3′UTR	5′UTR vs CDS	3′UTR vs CDS
Mean average reactivity	$P = 9.9\times10^{-19}$	$P < 10^{-100}$	$P < 10^{-100}$	$P = 3.1\times10^{-65}$	$P < 10^{-100}$	$P < 10^{-100}$	$P = 7.7\times10^{-57}$	$P = 5.1\times10^{-18}$	$P < 10^{-100}$

	22 °C (Panel H)			42 °C (Panel I)			42 °C−22 °C (Panel J)		
	5′UTR vs 3′UTR	5′UTR vs CDS	3′UTR vs CDS	5′UTR vs 3′UTR	5′UTR vs CDS	3′UTR vs CDS	5′UTR vs 3′UTR	5′UTR vs CDS	3′UTR vs CDS
Mean single strandedness	$P = 4.8\times10^{-26}$	$P = 7.7\times10^{-12}$	$P = 4.5\times10^{-15}$	$P = 0.98$	$P = 4.6\times10^{-7}$	$P = 6.7\times10^{-19}$	$P = 1.9\times10^{-27}$	$P = 0.13$	$P = 1.8\times10^{-94}$

Table S3. Mapping statistics of Ribo-seq libraries

Sample name	Temp.	Rep.	All reads[a]	Genome mapped reads[b] (% of All)	Nuclear-encoded rRNA mapped reads[b] (% of All)	Chloroplast-encoded rRNA mapped reads[b] (% of All)	cDNA mapped reads (% of All)
C10m[c]	22 °C	1	144,214,556	133,066,498 (92.3%)	64,790,999 (44.9%)	31,472,873 (21.8%)	55,216,455 (38.3%)
		2	133,960,640	123,943,648 (92.5%)	60,051,024 (44.8%)	31,062,270 (23.2%)	43,932,282 (32.8%)
H10m[c]	42 °C	1	134,512,418	124,759,176 (92.7%)	52,210,311 (38.8%)	17,063,054 (12.7%)	29,721,738 (22.1%)
		2	133,690,235	124,781,078 (93.3%)	61,442,088 (46.0%)	20,380,018 (15.2%)	37,142,708 (27.8%)

[a]High quality (Q > 30) and adapter trimmed reads.
[b]Some reads map to both nuclear and chloroplast genomes.
[c]In the sample name, "C" indicates control, "H" indicates 42 °C x 10 min heat treatment.

Table S4. Mapping statistics of time course RNA-seq libraries

Treatment Temp. & duration	Recovery time	Sample Name[a]	Rep.	Total reads[b]	Genome mapped reads (GMR)	% of total	Multi-mapped reads	% of GMR
22 °C for 10 min	0 min	C10m	1	33,869,785	32,978,562	97.4%	1,776,249	5.2%
			2	33,531,678	32,460,700	96.8%	1,892,381	5.6%
	10 min	C20m	1	31,261,223	30,760,972	98.4%	1,418,674	4.5%
			2	43,170,666	42,285,712	98.0%	2,246,220	5.2%
	50 min	C1h	1	35,228,054	34,510,009	98.0%	1,387,559	3.9%
			2	36,928,419	35,989,953	97.5%	1,791,360	4.9%
	1h 50 min	C2h	1	29,090,627	28,391,336	97.6%	1,196,994	4.1%
			2	32,243,175	31,408,171	97.4%	1,386,573	4.3%
	9h 50 min	C10h	1	41,378,526	40,347,435	97.5%	2,057,179	5.0%
			2	33,635,765	32,669,302	97.1%	1,810,499	5.4%
42 °C for 10 min	0 min	H10m	1	45,314,781	44,186,535	97.5%	2,320,600	5.1%
			2	33,168,426	32,416,819	97.7%	1,534,421	4.6%
	10 min	HR20m	1	33,440,984	32,159,843	96.2%	3,868,691	11.6%
			2	33,604,361	32,695,384	97.3%	1,596,237	4.8%
	50 min	HR1h	1	38,936,971	37,407,796	96.1%	3,380,004	8.7%
			2	29,188,625	28,375,413	97.2%	1,722,330	5.9%
	1h 50 min	HR2h	1	32,705,369	31,808,867	97.3%	1,580,836	4.8%
			2	41,806,142	40,970,086	98.0%	2,273,954	5.4%
	9h 50 min	HR10h	1	33,685,237	32,834,538	97.5%	3,069,028	9.1%
			2	33,292,915	31,760,585	95.4%	1,896,497	5.7%

[a]In the sample name, "C" indicates control, "H" indicates 42 °C x 10 min heat treatment and "HR" means recovery after the 42 °C x 10 min heat treatment. Timepoints are identical to those shown in Figure 1a.
[b]Total of 707 million reads in all libraries.

Table S5. Heat shock transcription factors (HSFs) with coverage in Structure-seq datasets show diverse changes in average DMS reactivity at 42°C as compared to 22°C.

Transcript ID	Description	Heat(42°C)	RT(22°C)	Difference
OS01T0733200-01	Similar to Heat shock transcription factor 29	0.17	0.18	0.01
OS01T0749300-01	Heat shock transcription factor	0.19	0.14	-0.05
OS01T0749300-02	Heat shock transcription factor	0.18	0.14	-0.04
OS02T0527300-01	Similar to Heat shock transcription factor 31	0.23	0.21	-0.02
OS03T0161900-01	Similar to Heat stress transcription factor A-2d	0.24	0.06	-0.18
OS03T0795900-01	Similar to Heat shock transcription factor 31	0.23	0.18	-0.05
OS03T0854500-01	Similar to Heat shock transcription factor 31	0.25	0.27	0.02
OS03T0854500-02	Similar to Heat shock transcription factor 31	0.26	0.27	0.01

Table S6. Primer sets used in qRT-PCR in RNA decay analysis

Transcript ID		Primer sequences
OS02T0729700-01_*HOX16*	5'	TTTAGCTTCAGAGCCCTCCA
	3'	CGGAGAAGAGGCACAATTCG
OS03T0797800-01_*AUX14*	5'	TCAGTTGGATCGGAGTTGGT
	3'	CTGCAGAACTAGCAGTTGCC
OS05T0121600-01_*AP2*	5'	TTCAACCTGAGCGACTACGA
	3'	CACGAACTCCTCCTTGGACA
OS01T0952800-01_*bHLH56*	5'	CGGCTACCTGCATCAATGAC
	3'	TGTTCCCAAATCCGGAGGAA
OS08T0468100-01_*NR1*	5'	TGTACCAGGTCATCCAGTCG
	3'	TCGATGACGTACCACACCTT
OS05T0563600-01_*FAS1*	5'	GCTGATGGACAGCTACAACG
	3'	CATCAGCTGGATCTGGTCCT
OS07T0184800-01_*H1*	5'	CTGAGGAGGTTGTCCCTGAG
	3'	GCCTTCTTCTTCTCCTTCGC
OS01T0871200-01_*bHLH33*	5'	ACAGTGGCTTGGGAAGAAGT
	3'	CCCTGTTGGTCAGGATCAGT
OS02T0662100-01_T2	5'	GCAGACCATCACTTACGTAGC
	3'	ACGGTGAAGAAGAGGACCAG
OS03T0159900-02_T3	5'	TGCTGCAATGGAGGGAAGAT
	3'	CGCAGACAAAGGGCTATCAC

Supplementary Dataset S1. All the analyzed data (and associated *P* values) for Structure-seq, Ribo-seq, and RNA-seq datasets.

Supporting File for

Genome-wide RNA structurome reprogramming by acute heat shock globally regulates mRNA abundance

Zhao Su[a,1], Yin Tang[b,1,2], Laura E. Ritchey[c,d], David C. Tack[a], Mengmeng Zhu[a], Philip C. Bevilacqua[c,d,e,3], and Sarah M. Assmann[a,d,3]

Affiliations:

[a]Department of Biology, Pennsylvania State University, University Park, PA 16802, USA.

[b]Bioinformatics and Genomics Graduate Program, Pennsylvania State University, University Park, PA 16802, USA.

[c]Department of Chemistry, Pennsylvania State University, University Park, PA 16802, USA.

[d]Center for RNA Molecular Biology, Pennsylvania State University, University Park, PA 16802, USA.

[e]Department of Biochemistry & Molecular Biology, Pennsylvania State University, University Park, PA 16802, USA.

[1]These authors contributed equally to this work.

[2]Current address: Department of Genetics, Yale University School of Medicine, New Haven, CT, 06520, USA.

[3]Corresponding authors. Email: sma3@psu.edu (S.M.A.); pcb5@psu.edu (P.C.B.).

Evaluating the *Oryza sativa* RNA structurome for the presence of prokaryotic-type RNA thermometers

RNA secondary structures are known to modulate translation initiation in prokaryotes; for example, strong mRNA structure can impede ribosome binding to the Shine-Dalgarno (SD) sequence (AGGA) (19). RNA thermometers (RNATs) in prokaryotes function by temperature-dependent changes in secondary structure that alter accessibility of the SD sequence to the ribosome, thereby controlling translation initiation in a temperature-dependent manner (20, 21). The repression of heat shock gene expression (ROSE) element and fourU element are two common types of RNA thermometers found in prokaryotes. These two types of RNATs operate in similar ways: the SD sequence is harbored in a hairpin structure at low temperature and the local hairpin melts at high temperature to expose the SD sequence, allowing ribosome binding. Another type of RNAT, found in *Synechocystis sp. PCC6803* (22), is similar to the fourU element but has UCCU, rather than four U's, base-pairing with the SD sequence. Two other RNATs are associated with two specific genes in prokaryotes: the *prfA* RNAT found in the 5′UTR of the *prfA* gene in *Listeria monocytogenes* (23) and the *cssA* RNAT found in the 5′UTR of the *cssA* gene in *Neisseria meningitides* (24). These thermometers are characterized by a strong hairpin located upstream and nearby the start codon, and have SD sequences within the hairpin that differ from the standard AGGA sequence.

Other types of RNATs in prokaryotes also employ similar mechanisms for controlling translation initiation. Narbenhaus and colleagues (25) identified multiple candidate RNATs in *Yersinia pseudotuberculosis* from genome-wide *in vitro* RNA structure data by identifying transcripts with a decreased average PARS score (less RNA structure) at the SD region (located 10 nt ± 4 nt upstream of the start codon) under elevated temperature (25). A subset of these RNATs were validated by observation of significant protein abundance increase under elevated temperatures in transient reporter assays conducted in *E. coli*.

Our study provides the first *in vivo* genome-wide datasets on temperature regulation of a eukaryotic RNA structurome, affording an opportunity to investigate the possible presence of prokaryotic or other types of RNA-based thermometers. Our RNA-seq and Ribo-seq data also allow direct assessment in the organism of interest of possible correlations between temperature-regulated RNA structure and transcript abundance or translation. However, as described below, we found no evidence for prokaryotic-type RNA thermometers in our datasets.

1. RNA thermometers search based on SD sequence

a. ROSE element

The repression of heat shock gene expression (ROSE) element is an RNA element that regulates translation and is found in the 5′UTRs of some bacterial heat shock genes (26). This element consists of a conserved SD sequence that base pairs with a UYGCU region, where Y represents a pYrimidine (C or U). Fig. 1 shows the RNA structure model of the ROSE element (20, 21). We performed a sequence search of the *Oryza sativa* reference transcriptome for ROSE elements present in the region 50 nt upstream of the start codon in mRNAs that contain a SD sequence located 10 nt ± 4 nt upstream of the start codon. We identified 1,621 candidates with a SD sequence within this region. Among these, five contained a ROSE element based on sequence identity. Of these, four had sufficient coverage in our RNA structuromes. Structures for these four mRNAs were predicted using RNAstructure (27) with and without DMS reactivities as restraints. The fifth mRNA, which does not meet our coverage requirement, was predicted *in silico* only. However, none of the candidates

are predicted to form an RNA secondary structure similar to that of the ROSE element (Fig. 2) at 22 °C; moreover none of these are a heat shock gene. Temperature change has little effect on the predicted RNA structures. None of these candidates exhibit a significant elevation in RNA abundance between 22 °C and 42 °C at any time point (Table 1).

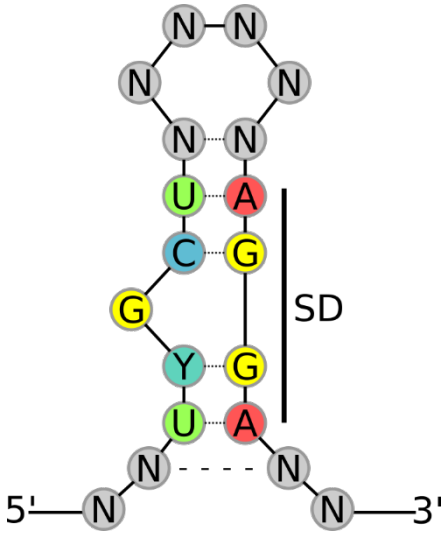
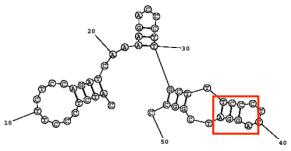
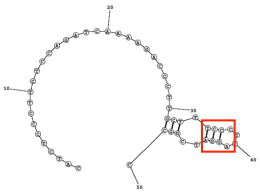


Fig. 1. RNA structure model of the ROSE element, modified from (20, 21). Colors represent different nucleotides (Red: A; Green: U; Blue: C; Yellow: G; Teal: Y = G or C, Gray: N = Any).

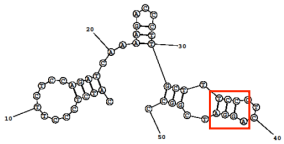
in silico **OS03T0784700-02** ***in vivo***

22 °C



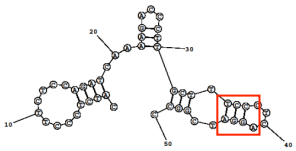


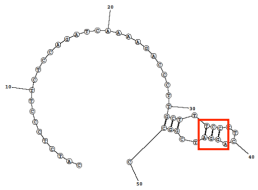
42 °C



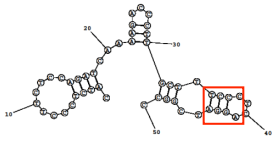
in silico **OS03T0784700-01** ***in vivo***

22 °C

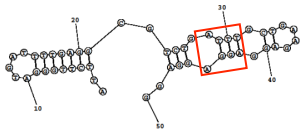




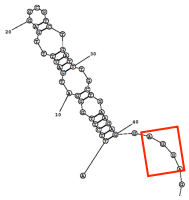
42 °C

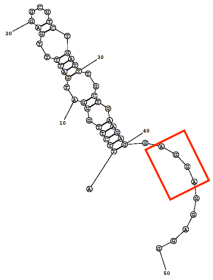


in silico **OS03T0413000-01** ***in vivo***

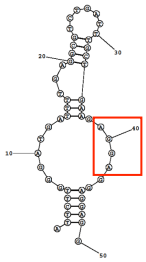


22 °C





42 °C



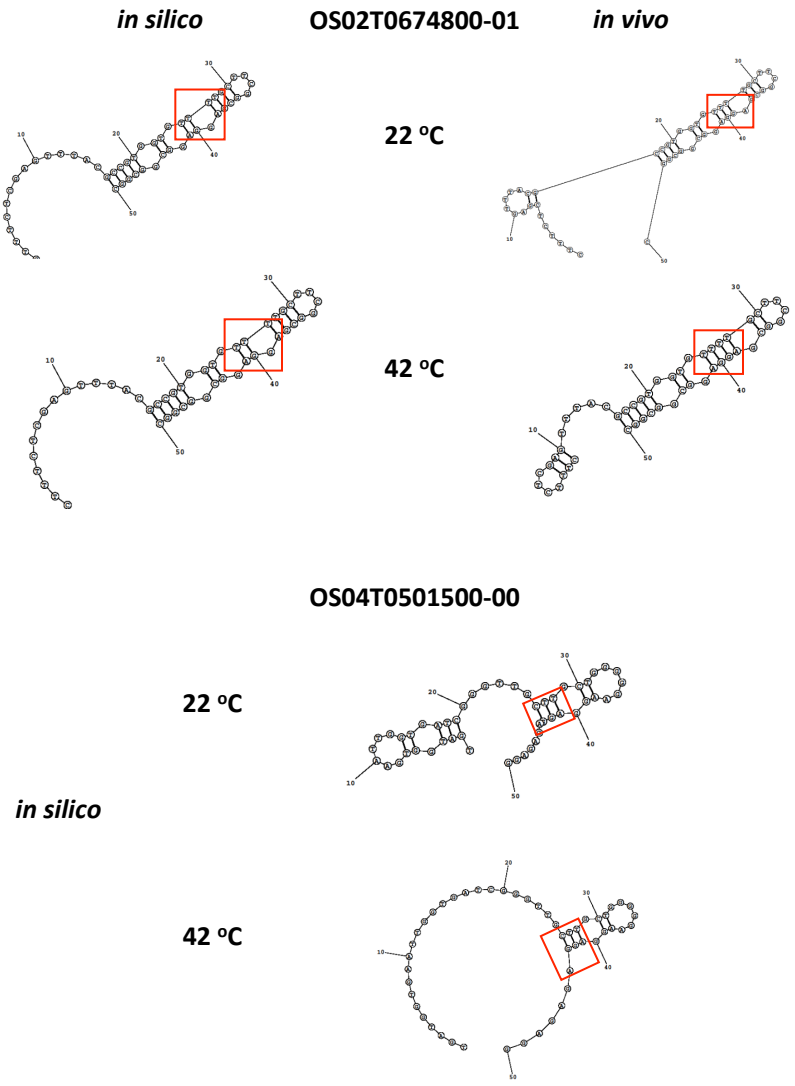


Fig. 2. Predicted RNA structures at 22 °C and 42 °C *in silico* and *in vivo* (with DMS reactivities as restraints) of ROSE element candidates in *Oryza sativa*. The red squares mark the SD sequence region. Structures were predicted using RNAstructure (27).

Table 1. Change of DMS reactivity, mRNA abundance, and Ribo-seq signal of the identified ROSE element candidates in *Oryza sativa*. Reactivity difference is the difference in average DMS reactivity between 22 °C and 42 °C (from Structure-seq data); RNA abundance fold change is the fold change of mRNA abundance between 22 °C and 42 °C at each time point (from time-series RNA-seq data); Ribo-seq difference is the difference in average Ribo-seq signal between 22 °C and 42 °C (from Ribo-seq data). SD stands for the Shine-Dalgarno sequence (AGGA) and the table shows the average reactivity of the four nucleotide. "Whole" stands for the whole transcript and the table shows the average reactivity of the whole transcript. NA indicates data not available in the dataset. Asterisks mark statistically significant changes of abundance (t-test, p value < 0.05)

		Reactivity difference		**RNA abundance fold change**					**Ribo-seq difference**
ID	Annotation	SD	Whole	10m	20m	1h	2h	10h	10m
OS03T0784700-02	Ferredoxin--NADP reductase, root isozyme, chloroplast precursor (EC 1.18.1.2)	-0.0035	-0.0011	0.92	0.91	0.41*	0.59	0.65	0.312
OS03T0784700-01	Ferredoxin--NADP reductase, root isozyme, chloroplast precursor (EC 1.18.1.2)	-0.0045	-0.0020	0.92	0.91	0.41*	0.59	0.65	0.283
OS03T0413000-01	Similar to Potential histone-like transcription factor	-0.5956	-0.0147	0.90	0.70	0.51	0.94	1.02	-0.004
OS02T0674800-01	Similar to °CL1 homeobox protein	0.2840	0.0121	1.03	0.92	0.56	0.71	0.70	-0.060
OS04T0501500-00	BTB domain containing protein	NA	NA	0.73	0.76	0.47	0.61	0.71	-0.176

FourU element

FourU thermometers are a type of RNA thermometer found in *Salmonella* (28), *E.coli* (29) and *V. cholerae* (30). This element consists of a conserved SD sequence that base pairs with a UUUU region. Fig. 3 shows the RNA structure model of the fourU element (20, 21). We performed a sequence search for fourU elements in the region 50 nt upstream of the start codon on all the mRNAs with a SD sequence present 10 nt ± 4 nt upstream of the start codon, and identified 11 fourU candidates with sequences which match that of the fourU element. Of these, five had sufficient coverage in our RNA structuromes for structure prediction. However, only one of these candidates (OS09T0572000-01) forms a predicted RNA secondary structure similar to that of the fourU element (Fig. 4) at 22 °C. While the SD sequence part of OS09T0572000-01 melts *in silico* at 42 °C, the RNA abundance of OS09T0572000-01 is only 0.07 (TPM), which is too low for RNA structure probing *in vivo*. Temperature change also has little effect on the remaining 10 RNA structures predicted either with or without DMS reactivities as restraints. One of these candidates (OS05T0542500-02) exhibits a dramatic change in RNA abundance between 22 °C and 42 °C at 1 hr, 2 hrs and 10 hrs time points (Table 2). However, the predicted RNA secondary structure of OS05T0542500-02 is not similar to that of the fourU element.

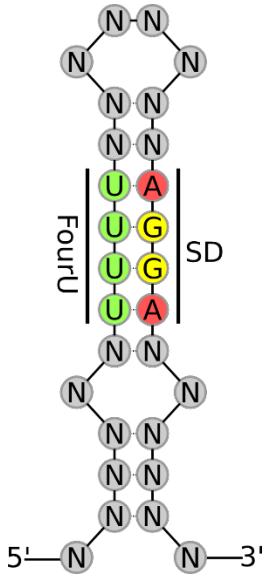
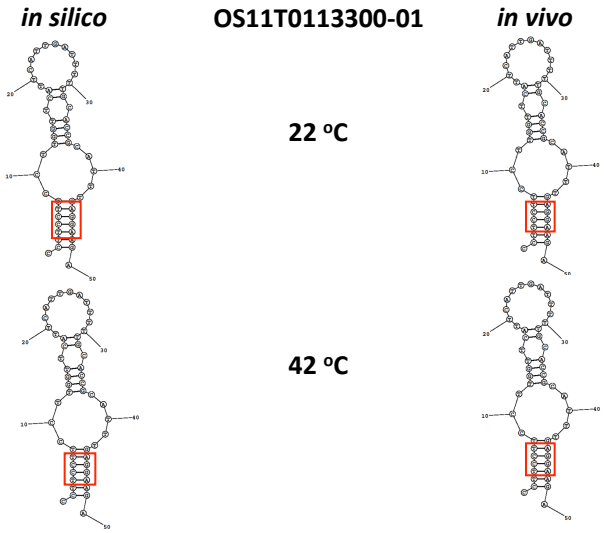
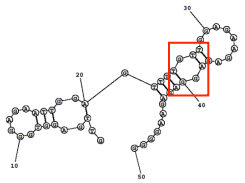


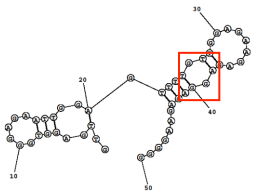
Fig. 3. RNA structure model of the fourU element, modified from (20, 21). Colors represent different nucleotides (Red: A; Green: U; Blue: C; Yellow: G, Gray: N = Any).

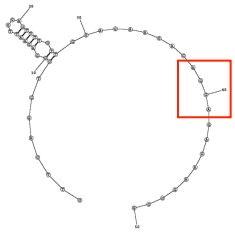


in silico **OS08T0506700-01** ***in vivo***

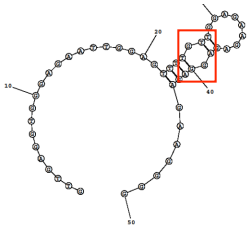


22 °C

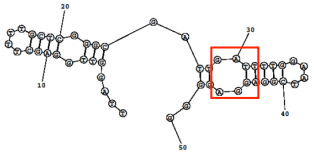




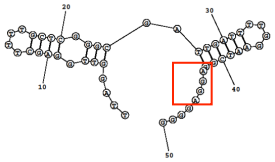
42 °C

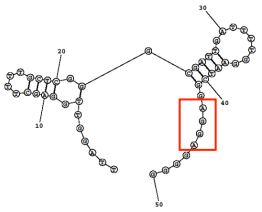


in silico **OS05T0545000-01** ***in vivo***

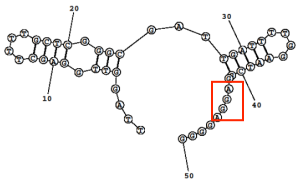


22 °C

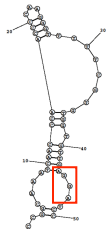




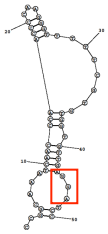
42 °C

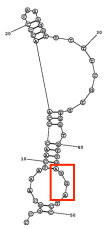


in silico **OS02T0813500-01** ***in vivo***

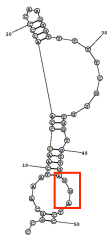


22 °C

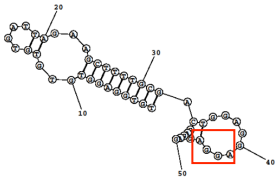




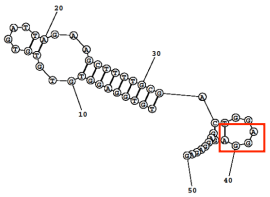
42 °C

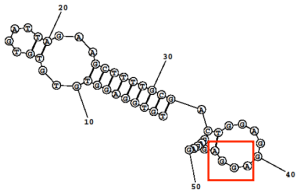


in silico **OS10T0522400-01** *in vivo*

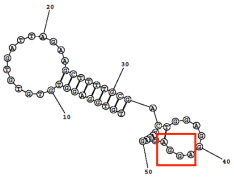


22 ºC



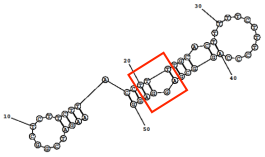


42 ºC



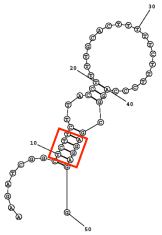
OS01T0889400-02

22 ºC



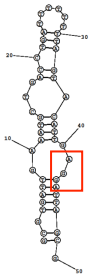
in silico

42 ºC



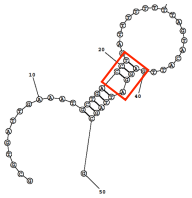
OS05T0227800-01

22 ºC



in silico

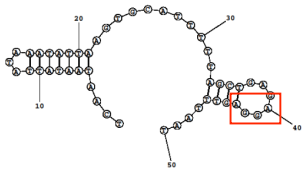
42 ºC



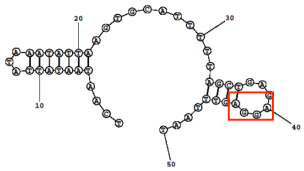
OS05T0492000-00

in silico

22 °C



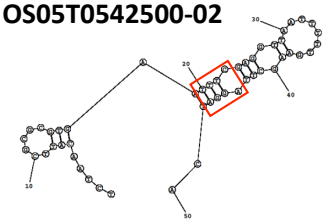
42 °C



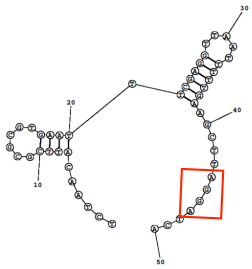
OS05T0542500-02

in silico

22 °C



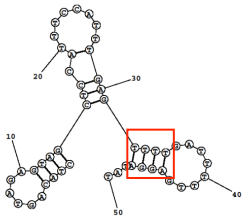
42 °C



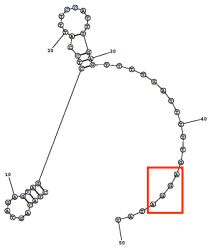
OS09T0572000-01

in silico

22 °C



42 °C



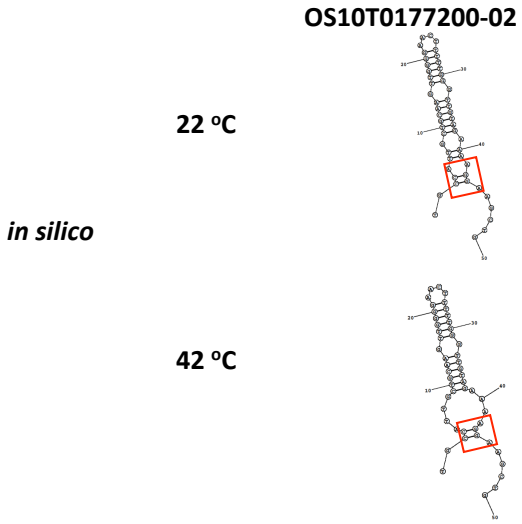


Fig. 4. Predicted RNA structures at 22 °C and 42 °C *in vivo* and *in silico* (with and without DMS reactivity as restraints) of fourU element candidates in *Oryza sativa*. The red squares mark the SD sequence region.

Table 2. Change of DMS reactivity, mRNA abundance, and Ribo-seq signal of the identified candidates in *Oryza sativa* with fourU elements. Reactivity difference is the difference in average DMS reactivity between 22 °C and 42 °C (from Structure-seq data); RNA abundance fold change is the fold change of mRNA abundance between 22 °C and 42 °C at each time point (from time-series RNA-seq data); Ribo-seq difference is the difference in average Ribo-seq signal between 22 °C and 42 °C (from Ribo-seq data). SD stands for the Shine-Dalgarno sequence (AGGA) and the table shows the average reactivity of the four nucleotides. "Whole" stands for the whole transcript and the table shows the average reactivity of the whole transcript. Inf indicates infinite value (division by 0). Asterisks mark statistically significant changes of abundance (t-test, p value < 0.05).

		Reactivity difference		RNA abundance fold change					Ribo-seq difference
ID	Annotation	SD	Whole	10m	20m	1h	2h	10h	10m
OS11T0113300-01	Conserved hypothetical protein	-0.023	0.0387	1.07	0.56	0.51	1.18	0.83	0.018
OS08T0506700-01	Helix-loop-helix DNA-binding domain containing protein	0.1319	0.0526	0.81	0.54*	0.60	1.51*	0.76	-0.011
OS05T0545000-01	Similar to Phosphatidylinositol transfer-like protein IV	0.4420	-0.0008	1.11	0.80	0.53	0.80	0.69	0.001
OS02T0813500-01	Glutathione reductase, cytosolic (EC 1.8.1.7) (GR) (GRase)	0.1580	0.0094	1.09	0.82	0.66	1.53*	0.76	0.094
OS10T0522400-01	Cyclin-like F-box domain containing protein	0.8569	0.1035	0.85	0.53	1.42*	1.33	1.09	0.272
OS01T0889400-02	Similar to LOB domain protein 6	NA	NA	1.22	1.00	1.12	2.21	0.90	-0.201
OS05T0227800-01	Homeodomain-like containing protein	NA	NA	1.41	1.03	1.16	1.79	0.88	-0.117
OS05T0492000-00	Similar to BZIP transcription factor	NA	NA	0.88	Inf	Inf	0.36	0.21	0
OS05T0542500-02	LEA-like protein, HSP12	NA	NA	1.73	3.08	14.6*	124*	2.37	0.065

OS09T0572000-01	Pathogenesis-related transcriptional factor and ERF domain containing protein	NA	NA	1.06	Inf	Inf	2.40	0.12	NA
OS10T0177200-02	EF-HAND 2 domain containing protein	NA	NA	1.29	0.96	2.01*	1.89*	0.67	-0.087

c. UCCU element

UCCU thermometers are a type of RNA thermometer found in *Synechocystis sp. PCC6803* (22). Fig. 5 shows the RNA structure model of the UCCU element (20, 21). We performed a sequence search for this type of RNAT in the region 50 nt upstream of the start codon. Among these, five contained a UCCU element based on sequence identity. Of these, four had sufficient coverage in our RNA structuromes. Three of these candidates form a predicted RNA secondary structure similar to the UCCU element both *in silico* and *in vivo* at 22 °C (Fig. 6). However, none of these structures melts out at the SD region at 42 °C either *in silico* or *in vivo* (Fig. 6). Moreover, unlike the *Synechocystis* UCCU thermometer, none of these candidates is a heat shock mRNA. OS06T0114000-02 has significant elevation of mRNA abundance at 42 °C as compared to 22 °C at 20 min, 1 hr and 2 hrs time points, and OS12T0167900-01 has significant elevation of mRNA abundance at 42 °C as compared to 22 °C at 20 min, however, neither of the candidates shows marked change in Ribo-seq signal between 22 ^{o}C and 42 ^{o}C (Table 3).

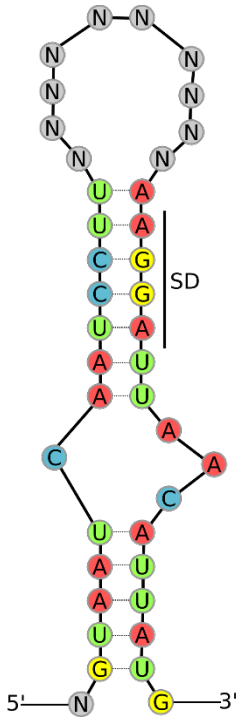
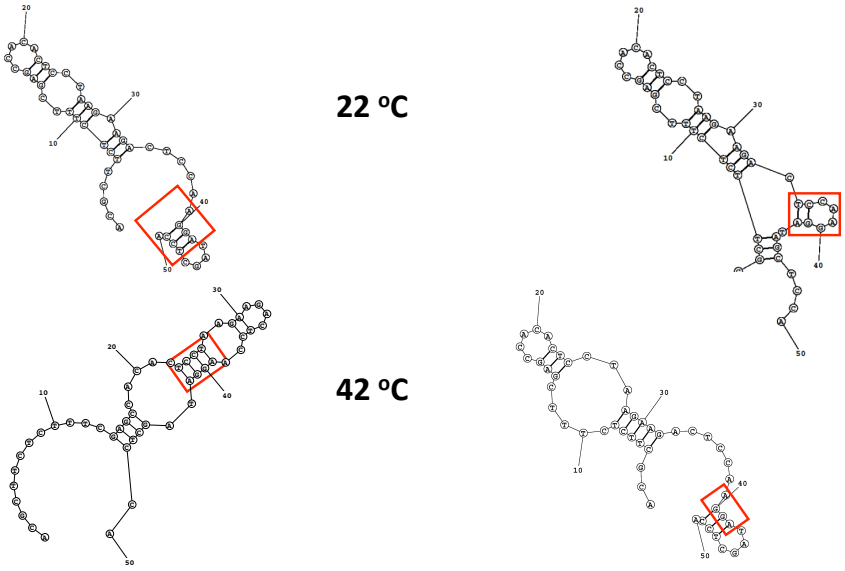
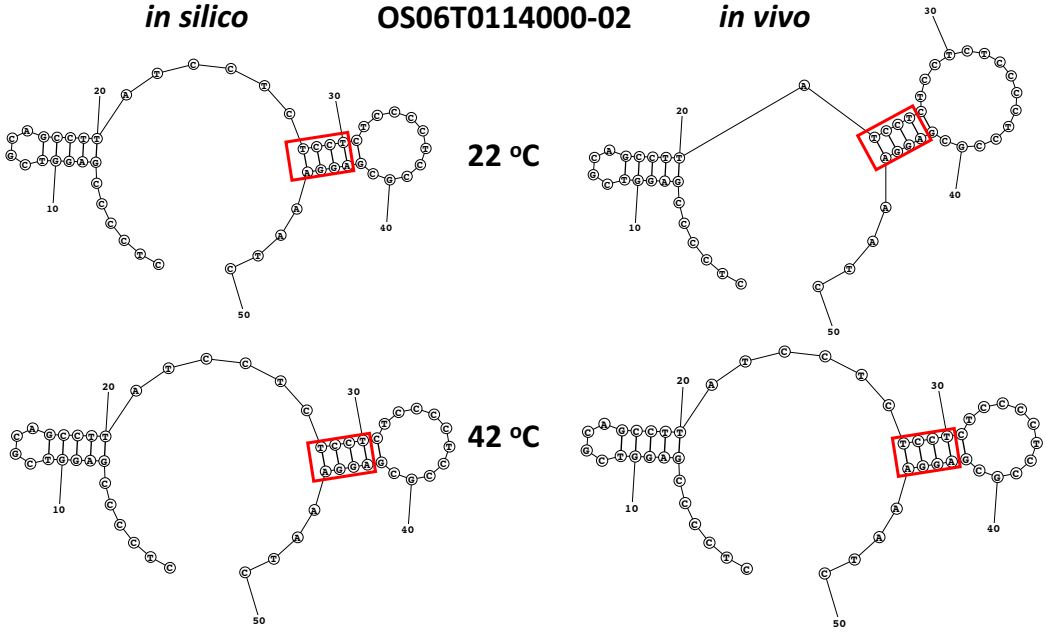


Fig. 5. RNA structure model of the UCCU element, modified from (20, 21). Colors represent different types nucleotides (Red: A; Green: U; Blue: C; Yellow: G; Gray: N = Any).

in silico **OS06T0148200-01** ***in vivo***



in silico **OS06T0114000-02** ***in vivo***



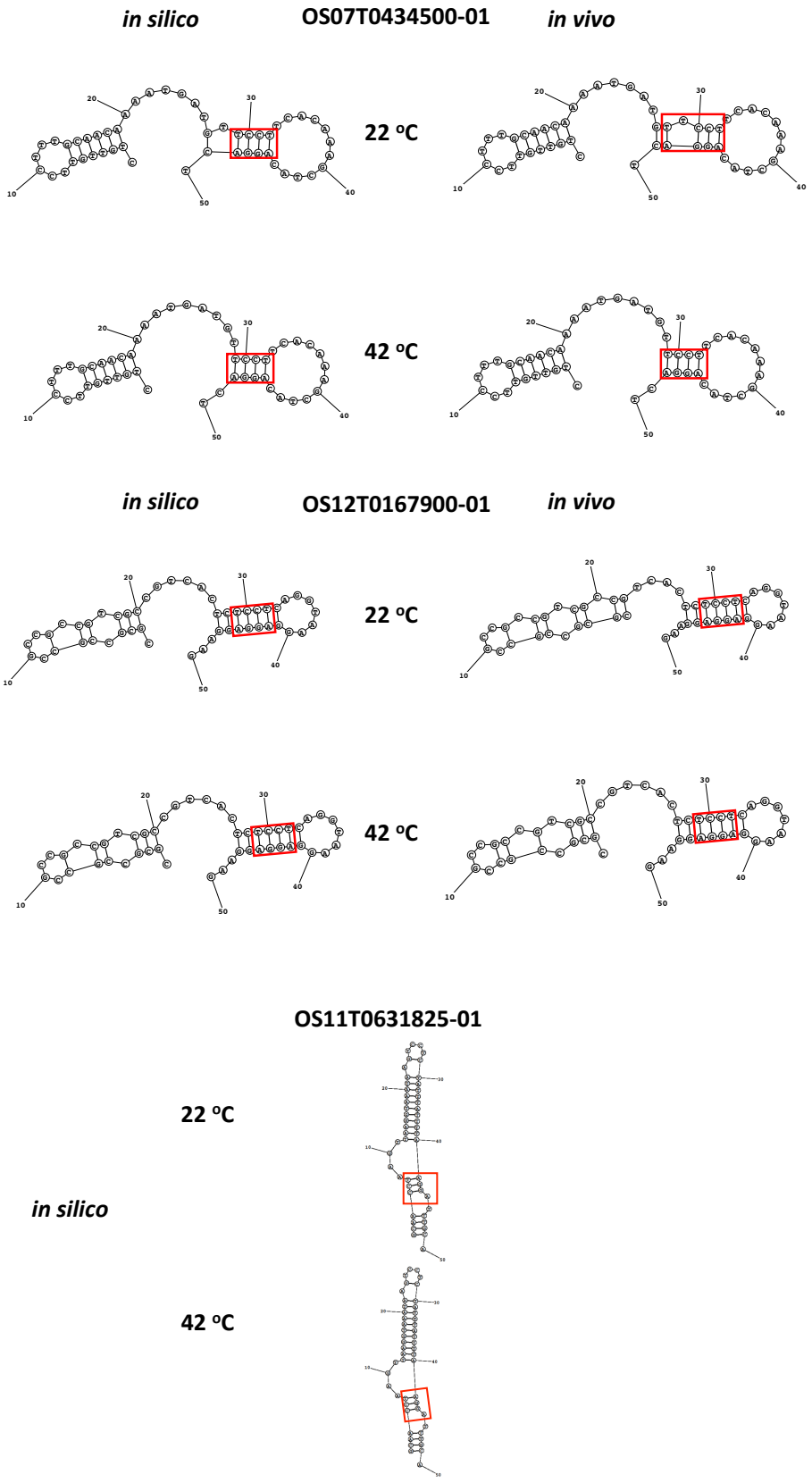


Fig. 6. Predicted RNA structures at 22 °C and 42 °C *in silico* and *in vivo* (with DMS reactivities as restraints) of UCCU element candidates in *Oryza sativa*. The red squares mark the SD sequence region.

Table 3. Change of DMS reactivity, mRNA abundance, and Ribo-seq signal of identified UCCU element candidates in *Oryza sativa*. Reactivity difference is the difference in average DMS reactivity between 22 °C and 42 °C (from Structure-seq data); RNA abundance fold change is the fold change of mRNA abundance between 22 °C and 42 °C at each time point (from time-series RNA-seq data); Ribo-seq difference is the difference in average Ribo-seq signal between 22 °C and 42 °C (from Ribo-seq data). SD stands for the Shine-Dalgarno sequence (AGGA) and the table shows the average reactivity of the four nucleotides. "Whole" stands for the whole transcript and the table shows the average reactivity of the whole transcript. Asterisks mark statistically significant changes of abundance (t-test, p value < 0.05).

		Reactivity difference		**RNA abundance fold change**					**Ribo-seq difference**
ID	Annotation	SD	Whole	10m	20m	1h	2h	10h	10m
OS06T0148200-01	Lipase, GDSL domain containing protein	0.9793	0.0791	0.56	0.62	0.18*	0.18*	0.90	-0.031
OS06T0114000-02	Similar to 60 kDa chaperonin (Protein Cpn60) (groEL protein) (63 kDa stress protein)	-0.2769	-0.0594	1.48	2.29*	4.18*	3.47*	1.62	-0.074
OS07T0434500-01	Similar to chromatin remodeling complex subunit	0.5485	0.0056	1.10	0.71	0.78	0.70	0.74	0
OS12T0167900-01	60S ribosomal protein L3	-0.4771	0.0161	0.92	1.42*	0.66	0.95	1.27	0.015
OS11T0631825-01	Conserved hypothetical protein	NA	NA	0.85	0.69	0.57	0.60	1.05	0

d. Other types of RNATs in bacteria

Fig. 7 shows RNA structure models of the *prfA* 5′UTR RNAT of *Listeria monocytogenes* (23) and the *cssA* 5′UTR RNAT of *Neisseria meningitides* (24). We did not find any exact matches to these sequences in the 5′UTRs of any *Oryza sativa* mRNAs.

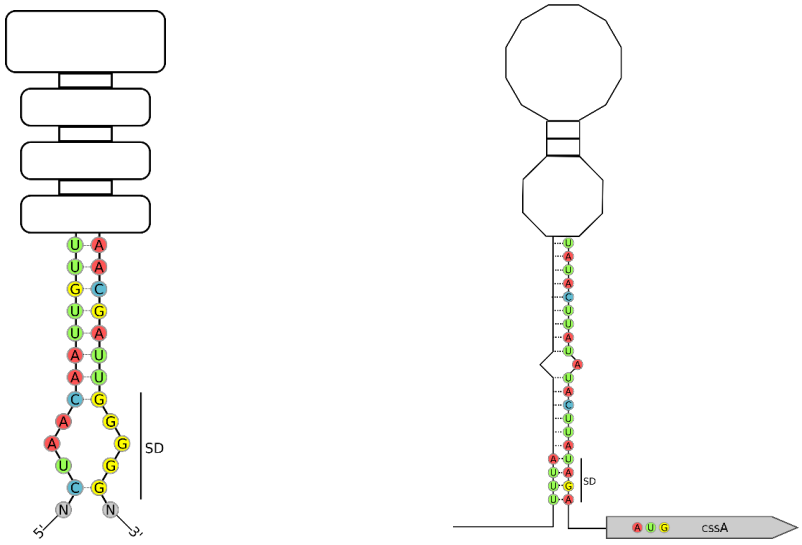


Fig. 7. RNA structure models of *prfA* (left) and *cssA* (right) RNATs, modified from (21). The elongated nucleotide hairpin with internal loops and bulges of the *prfA* RNAT is drawn schematically. The symbols at the tops of the structures represent non-obligatory parts of the RNAT. Colors represent different types of nucleotides (Red: A; Green: U; Blue: C; Yellow: G; Gray: N = Any).

2. RNA thermometer search in rice chloroplast transcriptome

Since chloroplasts are of prokaryotic origin, we performed a search for prokaryotic types of RNA thermometers in the chloroplast transcriptome of rice. We did not find any sequence matches to the ROSE element or UCCU element types of RNA thermometers within the region 50 nt upstream of the start codon of chloroplast mRNAs. We identified only one candidate that matches the fourU element sequence, located in the region 50 nt upstream of the start codon of the *atpH* (ATP synthase subunit c) transcript. However, the SD sequence (marked by a red square) is not open at 42 °C (Fig. 8) in either the *in silico* or the *in vivo* structures of this region, indicating that this candidate is not likely to be an RNA thermometer.

Table 4. Coverage and change of DMS reactivity of the identified candidate of the four U element in the *Oryza sativa* chloroplast transcriptome.

ID (Symbol)	Coverage (22 °C)	Coverage (42 °C)	Reactivity difference (whole) [per nt]	Reactivity difference (SD) [per nt]
atpH	1.17	1.89	0.015	0.39

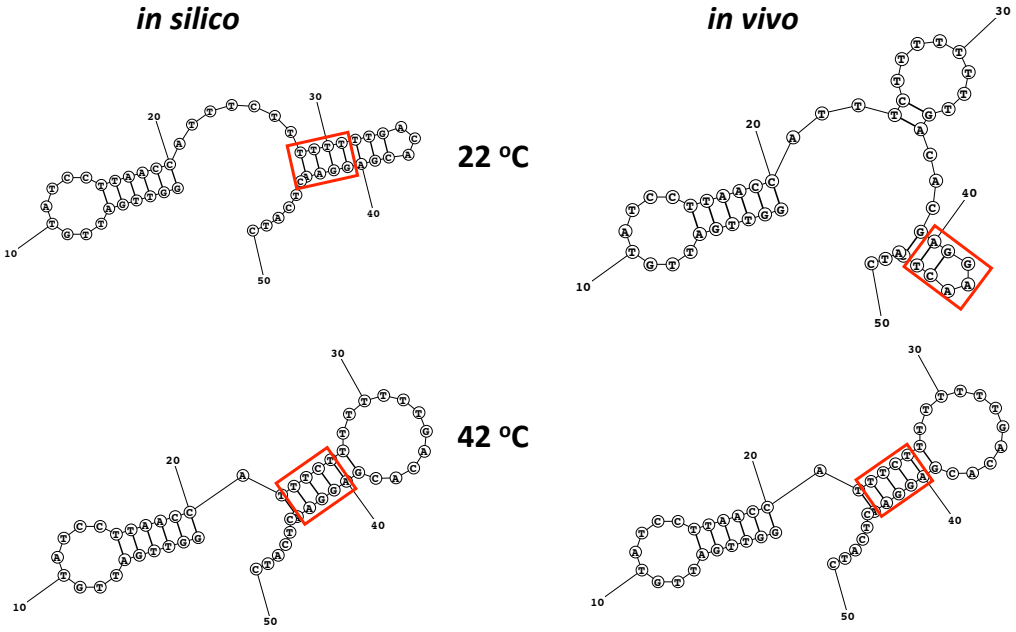


Fig. 8. Predicted *in silico* and *in vivo* RNA secondary structures of the 50 nt upstream of start codon of *atpH* at 22 °C and 42 °C. The red squares mark the SD sequence. Structures were predicted using RNAstructure (27).

3. RNA thermometers in eukaryotes

A cis-regulatory element thermometer was proposed for the HSP90 mRNA of the eukaryote, *Drosophila melanogaster* (31). As for most eukaryotic transcripts, the HSP90 transcript does not contain a SD sequence, but has a ~3-4 fold increase in protein abundance under heat shock compared to a normal growth temperature. In *D. melanogaster* the 5′UTR of HSP90 had greater stability (significantly lower free energy per nucleotide) than other HSP mRNAs. In contrast, we identified the ortholog of the HSP90 mRNA in rice (OS06G0716700) by sequence alignment and found that the free energy per nucleotide of the 5′UTR of the rice HSP90 mRNA does not differ significantly as compared to other mRNAs that code for HSPs, based on predicted RNA structures *in silico* or with DMS reactivities as restraints at 22 °C and 42 °C (Fig. 9A,B).

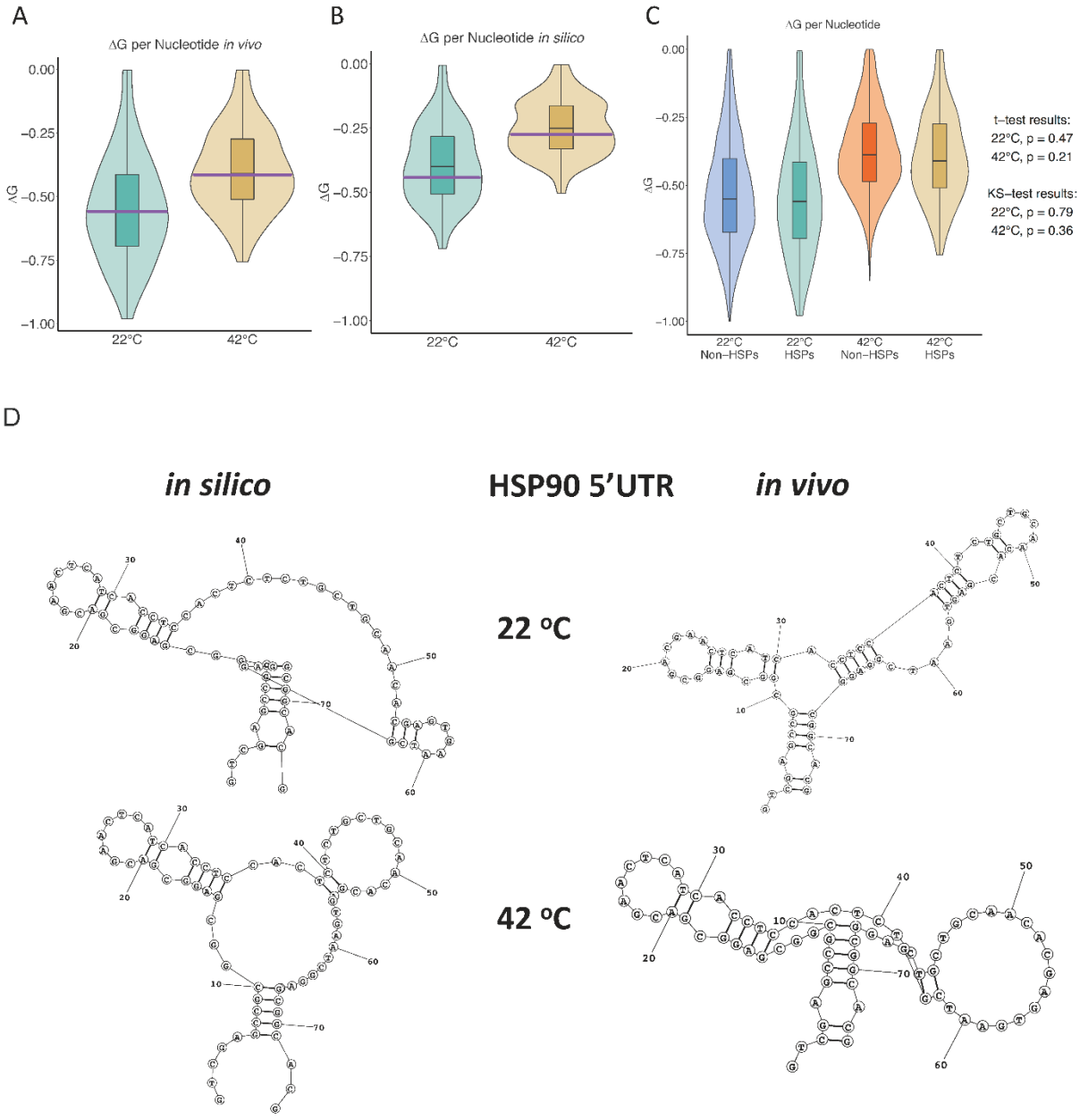


Fig. 9. **Distribution of free energy per nucleotide within the entire 5′UTR of HSP mRNAs and other mRNAs in our Structure-seq dataset.** (A) Distribution of the free energy per nucleotide within the 5′UTRs

of all *Oryza sativa* HSP mRNAs with sufficient coverage from Structure-seq (n = 93), based on RNA structure prediction using DMS reactivities as restraints (B) Distribution of the free energy per nucleotide within the 5′UTRs of all HSP mRNAs (n = 168), based on structures predicted *in silico*. (C) Comparison of the distribution of the free energy per nucleotide of the 5′UTRs of all HSP mRNAs (n = 93) and all other mRNAs (n = 9,875) with 5′UTR annotation and with sufficient coverage from Structure-seq. In panels A and B, the data for HSP90 mRNA are marked with a purple horizontal line. In the violin plots of panels A-C, green indicates the distribution of free energy per nt of the 5′UTR of all HSP mRNAs at 22 °C; dark yellow indicates the distribution of free energy per nt of the 5′UTR of all HSP mRNAs at 42 °C; blue indicates the distribution of free energy per nt of the 5′UTR of all mRNAs other than HSPs with 5′UTR annotation and with sufficient coverage from Structure-seq (n = 9,875); red indicates the distribution of free energy per nt of the 5′UTR of all mRNAs other than HSPs at 42 °C with 5′UTR annotation and with sufficient coverage from Structure-seq (n = 9,875). (D) The predicted RNA structure of the 5′UTR of rice HSP90 *in silico* or with DMS reactivities as restraints at 22 °C and 42 °C.

The authors (31) also proposed that unlike the HSP70 and HSP22 mRNA which have minimal 5′UTR RNA secondary structure in *D. melanogaster*, the *Drosophila* HSP90 mRNA may adopt a similar mechanism as prokaryotic RNATs, consisting of thermal melting of a stem-containing region near start codon, although no direct evidence was provided. Fig. 9D shows the predicted RNA structure of the 5′UTR of rice HSP90 *in silico* or with DMS reactivities as restraints at 22 °C and 42 °C. We did not observe obvious thermal melting of the RNA secondary structure near the start codon predicted either *in silico* or with DMS reactivities as restraints at 42 °C. In fact, in rice, there is no significant difference in free energy per nucleotide in the 5′UTRs of mRNAs that code for HSPs versus all other mRNAs with sufficient coverage (Fig. 9C). Together, these results provide no evidence that HSP mRNAs in rice function as thermosensors in a similar way to that proposed for the HSP90 cis-regulatory element in *D. melanogaster*.

4. Kozak sequence

The Kozak consensus sequence is a sequence in eukaryotic mRNAs that plays an important role in translation initiation. We hypothesized that RNA thermometers in plants may function by temperature-dependent changes in secondary structure that alter accessibility of the Kozak sequence to the ribosome, thus regulating translation. The Kozak sequence in plants is AACA(AUG) as suggested in (32). We identified 158 sequence matches to the Kozak sequence within the set of 14,292 mRNAs with sufficient Structure-seq coverage. We checked the correlation between the average DMS reactivity change on the Kozak sequence between 22 °C and 42 °C of the identified 158 Kozak sequence-containing transcripts and their mRNA abundance fold change ($\log_2$). However, the DMS reactivity change of these mRNAs is not correlated with their abundance fold change ($\log_2$) at any time point (Fig. 10A-E). In addition, we have also investigated the correlation between the average DMS reactivity change on the Kozak sequence between 22 °C and 42 °C of these 158 Kozak sequence-containing transcripts and their Ribo-seq signal change (Fig. 10F), but observed no correlation between average DMS reactivity change and Ribo-seq signal change between 22 °C and 42 °C. These results indicate that the Kozak sequence may not be a target for RNA structure-based regulation of gene expression in *Oryza sativa*.

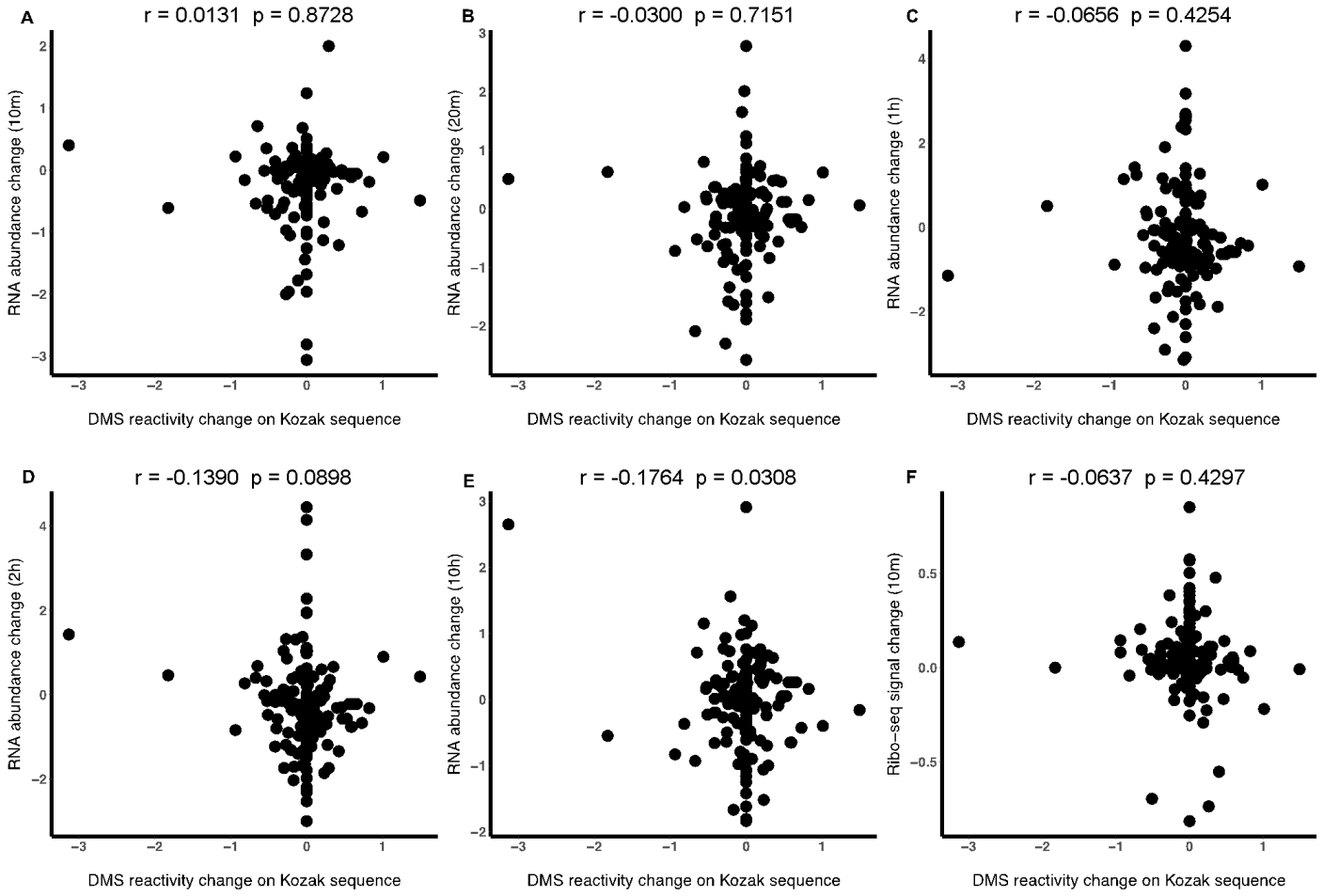


Fig. 10. Lack of correlation between change of DMS reactivity on Kozak sequences and mRNA abundance changes (log_2) at 22 °C and 42 °C at different time points (A-E), and Ribo-seq signal change at 22 °C and 42 °C (F). (A) 10 min (B) 20 min (C) 1 hr (D) 2 hrs (E) 10 hrs (F) Ribo-seq 10min.

5. RNA thermometer search in 5′UTRs within the 50 nt upstream of the start codon in rice

We performed a sequence motif search with the idea that rice might employ a temperature-regulated sequence motif near the start codon that is different from known RNAT translation-related motifs. Motif search was performed using MEME (33) on the 50 nt upstream of the start codon of the "top group" (Fig. 11a) and "bottom group" (Fig. 11b) of mRNAs. Here, the top group is the 5% of mRNAs with the most *elevated* average DMS reactivity at 42 °C as compared to 22 °C and the bottom group is the of 5% mRNAs with the most *reduced* average reactivity at 42 °C as compared to 22 °C. Sequence motif search was also performed on all mRNAs (n = 4,308) with *elevated* Ribo-seq signal at 42 °C and with 5′UTR length ≥ 50 nt (Fig. 11C), and all mRNAs (n = 8,739) with sufficient coverage and with 5′UTR length ≥ 50 nt in our RNA structuromes (Fig. 11D). We observed a similar AG rich motif as the most overrepresented sequence motif among the four groups. In addition, we did not observe any strongly overrepresented motif within each group: motifs within each group are quite different from each other. These results suggest that *Oryza sativa* may not employ conserved sequence motifs as analogous to the SD sequence in RNATs of bacteria.

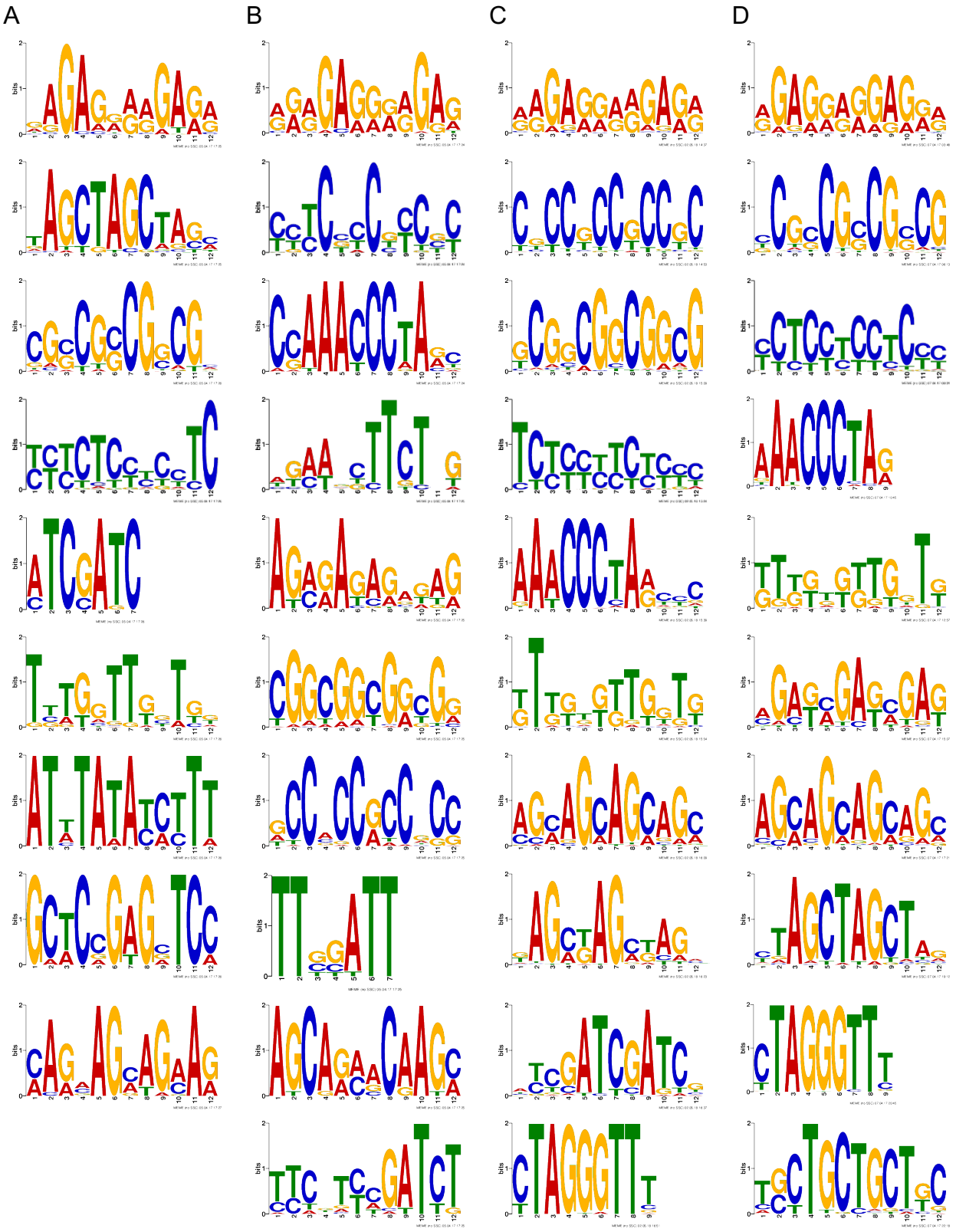


Fig. 11. Overrepresented sequence motifs in different mRNA classes. Overrepresented sequence motifs in the 50 nucleotides upstream of the start codon within (A) top group (B) bottom group (C) all mRNAs with elevated Ribo-seq signal at 42 °C based on Ribo-seq data and with 5′UTR length ≥ 50 nt (D) all mRNAs with

sufficient coverage from Structure-seq and with 5′UTR length ≥ 50 nt. Here, motifs are ranked according to the significance of overrepresentation.

Conclusions

Based on the above results, we found no evidence in rice for RNA thermometers of the prokaryotic type. We also failed to find evidence of any HSP mRNA functioning as a thermosensor in the manner proposed for the HSP90 cis-regulatory element thermometer in *Drosophila melanogaster* (31), nor did we find any Kozak sequence acting like the SD sequence of prokaryotic RNA thermometers. In addition, we did not find any clear evidence for any conserved mRNA sequence motif that functions as a RNA thermometer. In summary, we did not find evidence in rice for discrete RNA-based thermometers.

SI Reference Citations

1. Ding Y, Kwok CK, Tang Y, Bevilacqua PC, & Assmann SM (2015) Genome-wide profiling of *in vivo* RNA structure at single-nucleotide resolution using structure-seq. *Nat. Protoc.* 10(7):1050-1066.
2. Ding Y*, et al.* (2014) *In vivo* genome-wide profiling of RNA secondary structure reveals novel regulatory features. *Nature* 505(7485):696-700.
3. Ritchey LE*, et al.* (2017) Structure-seq2: sensitive and accurate genome-wide profiling of RNA structure *in vivo*. *Nucleic Acids Res.* 45(14):e135.
4. Kwok CK, Ding Y, Sherlock ME, Assmann SM, & Bevilacqua PC (2013) A hybridization-based approach for quantitative and low-bias single-stranded DNA ligation. *Anal. Biochem.* 435(2):181-186.
5. Juntawong P, Girke T, Bazin J, & Bailey-Serres J (2014) Translational dynamics revealed by genome-wide profiling of ribosome footprints in *Arabidopsis*. *Proc. Natl Acad. Sci. USA* 111(1):E203-212.
6. Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17(1):10-12.
7. Langmead B & Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9(4):357-359.
8. Tang Y*, et al.* (2015) StructureFold: genome-wide RNA secondary structure mapping and reconstruction *in vivo*. *Bioinformatics* 31(16):2668-2675.
9. Tack DC, Tang Y, Ritchey LE, Assmann SM, & Bevilacqua PC (2018) StructureFold2: Bringing chemical probing data into the computational fold of RNA structural analysis. *Methods*.
10. Low JT & Weeks KM (2010) SHAPE-directed RNA secondary structure prediction. *Methods* 52(2):150-158.
11. Kertesz M*, et al.* (2010) Genome-wide measurement of RNA secondary structure in yeast. *Nature* 467(7311):103-107.
12. Dobin A*, et al.* (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29(1):15-21.
13. 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.
14. Sekhon JS (2011) Multivariate and propensity score matching software with automated balance optimization: the Matching package for R. *J. Stat. Softw.* 42:1-52.

15. Gosai SJ, *et al.* (2015) Global analysis of the RNA-protein interaction and RNA secondary structure landscapes of the *Arabidopsis* nucleus. *Mol. Cell* 57(2):376-388.
16. Park SH, *et al.* (2012) Posttranscriptional control of photosynthetic mRNA decay under stress conditions requires 3' and 5' untranslated regions and correlates with differential polysome association in rice. *Plant Physiol.* 159(3):1111-1124.
17. Leng M & Felsenfeld G (1966) A study of polyadenylic acid at neutral pH. *J. Mol. Biol.* 15(2):455-466.
18. Merret R, *et al.* (2015) Heat-induced ribosome pausing triggers mRNA co-translational decay in *Arabidopsis thaliana*. *Nucleic Acids Res.* 43(8):4121-4132.
19. Laursen BS, Sorensen HP, Mortensen KK, & Sperling-Petersen HU (2005) Initiation of protein synthesis in bacteria. *Microbiol. Mol. Biol. Rev.* 69(1):101-123.
20. Narberhaus F (2010) Translational control of bacterial heat shock and virulence genes by temperature-sensing mRNAs. *RNA Biol.* 7(1):84-89.
21. Krajewski SS & Narberhaus F (2014) Temperature-driven differential gene expression by RNA thermosensors. *Biochim. Biophys. Acta* 1839(10):978-988.
22. Waldminghaus T, Gaubig LC, & Narberhaus F (2007) Genome-wide bioinformatic prediction and experimental evaluation of potential RNA thermometers. *Mol. Genet. Genomics* 278(5):555-564.
23. Johansson J, *et al.* (2002) An RNA thermosensor controls expression of virulence genes in *Listeria monocytogenes*. *Cell* 110(5):551-561.
24. Loh E, *et al.* (2013) Temperature triggers immune evasion by *Neisseria meningitidis*. *Nature* 502(7470):237-240.
25. Righetti F, *et al.* (2016) Temperature-responsive in vitro RNA structurome of *Yersinia pseudotuberculosis*. *Proc. Natl Acad. Sci. USA* 113(26):7237-7242.
26. Nocker A, *et al.* (2001) A mRNA-based thermosensor controls expression of rhizobial heat shock genes. *Nucleic Acids Res.* 29(23):4800-4807.
27. Reuter JS & Mathews DH (2010) RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics* 11:129.
28. Waldminghaus T, Heidrich N, Brantl S, & Narberhaus F (2007) FourU: a novel type of RNA thermometer in *Salmonella*. *Mol. Microbiol.* 65(2):413-424.
29. Klinkert B, *et al.* (2012) Thermogenetic tools to monitor temperature-dependent gene expression in bacteria. *J. Biotechnol.* 160(1-2):55-63.
30. Weber GG, Kortmann J, Narberhaus F, & Klose KE (2014) RNA thermometer controls temperature-dependent virulence factor expression in *Vibrio cholerae*. *Proc. Natl Acad. Sci. USA* 111(39):14241-14246.
31. Ahmed R & Duncan RF (2004) Translational regulation of Hsp90 mRNA. AUG-proximal 5'-untranslated region elements essential for preferential heat shock translation. *J. Biol. Chem.* 279(48):49919-49930.
32. Lutcke HA, *et al.* (1987) Selection of AUG initiation codons differs in plants and animals. *EMBO J.* 6(1):43-48.
33. Bailey TL, *et al.* (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* 37(Web Server issue):W202-208.