NAD+-capped RNAs are widespread in the Arabidopsis transcriptome and can probably be translated

Wang et al.

SUPPLEMENTAL INFORMATION

Supplemental Figures and Legends

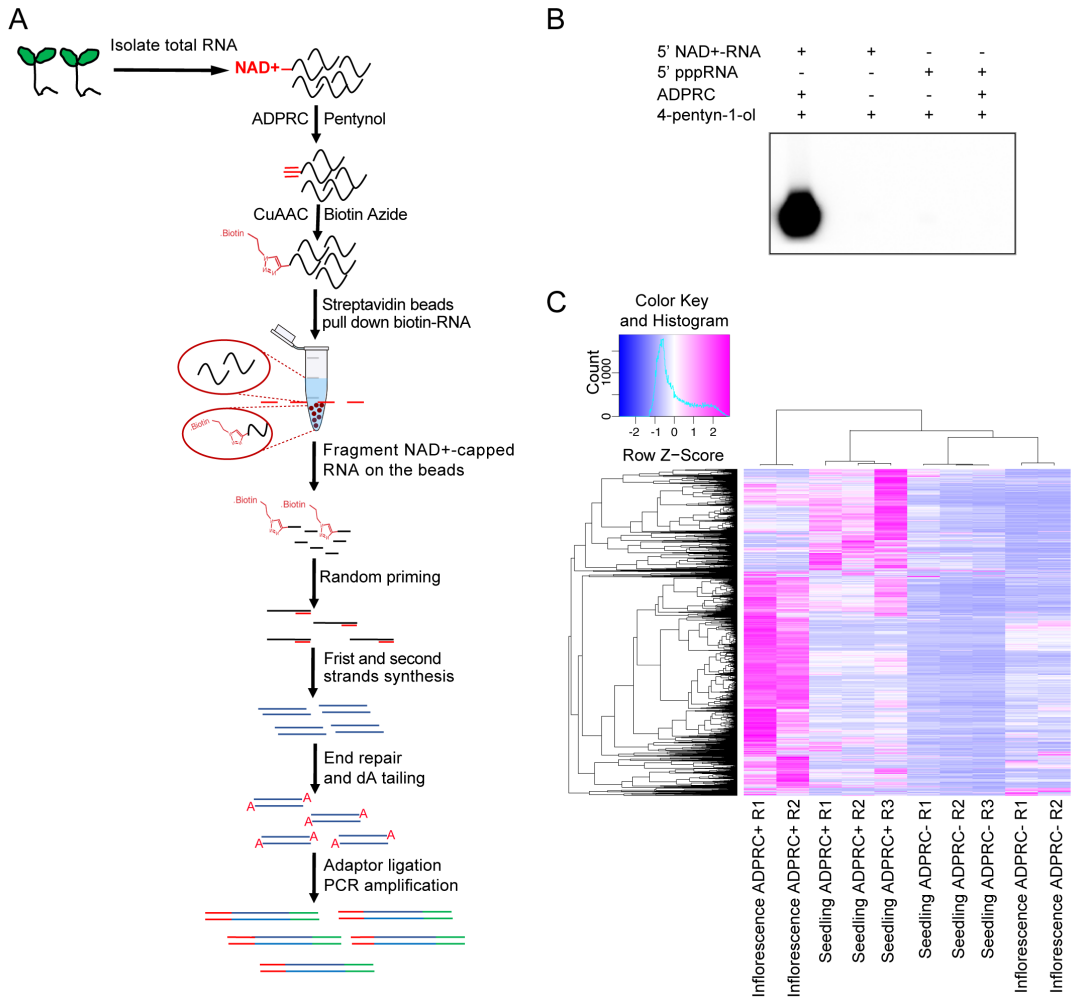

Figure S1. NAD captureSeq workflow and quality control. (A) Diagram of the workflow of NAD captureSeq. (B) A gel blot assay to demonstrate the selectivity of NAD capture. In vitro transcribed 5' pppRNA and 5' NAD+-RNA were treated or not with ADPRC, and then subjected to the click chemistry with biotin azide. The RNAs were resolved in a 2% agarose gel, transferred to a nylon membrane, and probed by Streptavidin-HRP to detect biotinylated RNA. (C) Clustering analysis of NAD

www.pnas.org/cgi/doi/10.1073/pnas.1903682116

captureSeq samples. The 10000 top-varying transcripts in seedling and inflorescence libraries were used for the analysis.

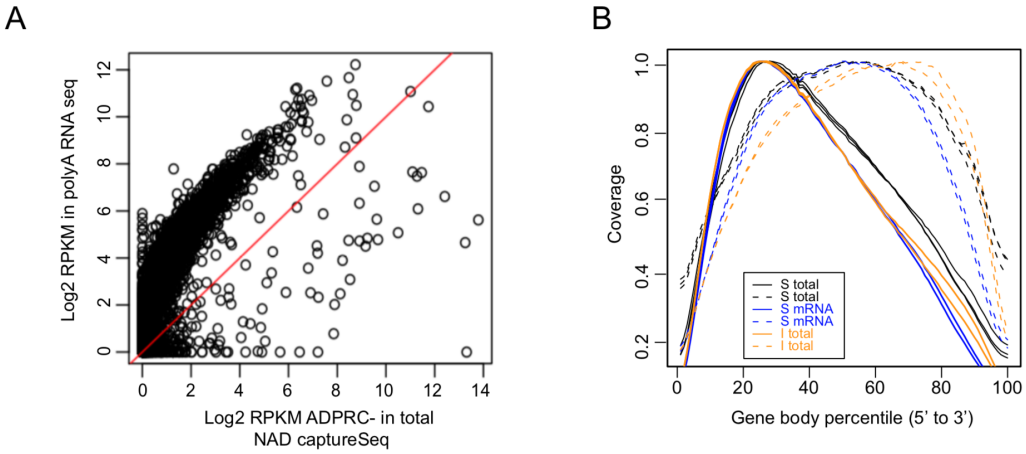

Figure S2. Features of NAD captureSeq read output. (A) The abundance of RNAs in ADPRC- samples was largely proportional to that in polyA RNA-seq libraries. (B) NAD captureSeq read distribution along gene bodies. The solid and dotted lines represent the ADPRC+ and ADPRC- libraries, respectively, in various samples as indicated. S, seedling; I, inflorescence. “total”, total RNAs. “mRNA”, polyA RNAs. Two to three lines of the same color represent biological replicates.

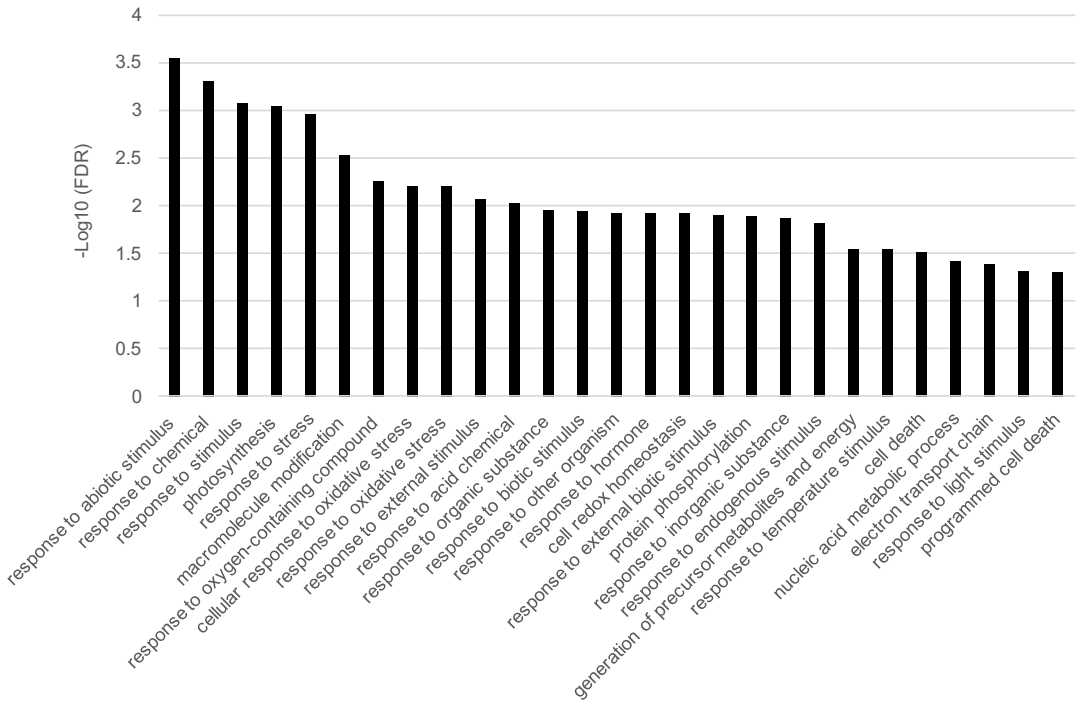

Figure S3. GO enrichment analysis of 1000 genes with the highest ADPRC+ /ADPRC- ratio from seedling NAD captureSeq.

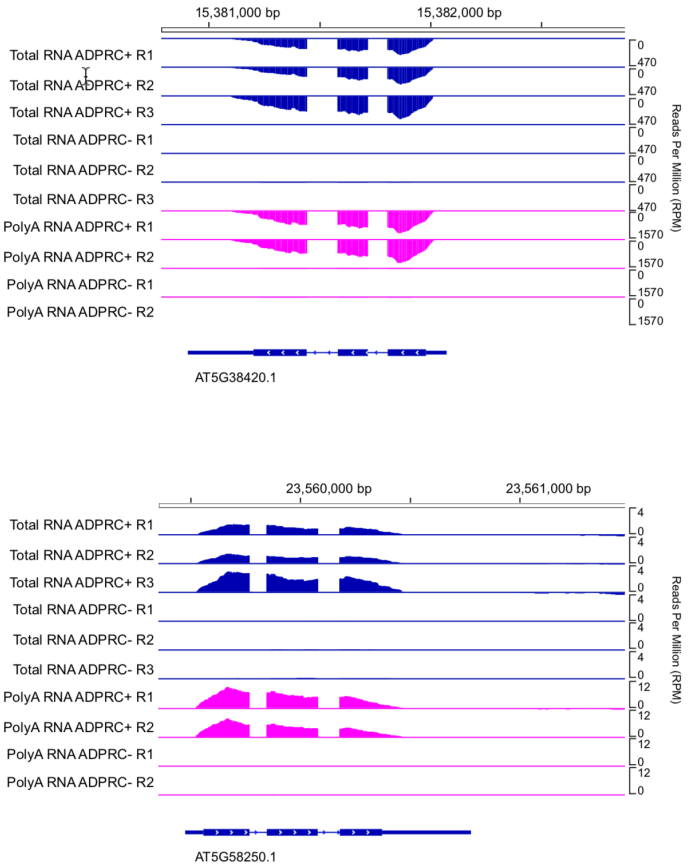

Figure S4. Genome browser views of reads from seedling total RNA and polyA RNA NAD captureSeq for two genes. The gene models are at the bottom, with boxes and thin lines representing exons and introns, respectively. Three and two biological replicates (R) were performed.

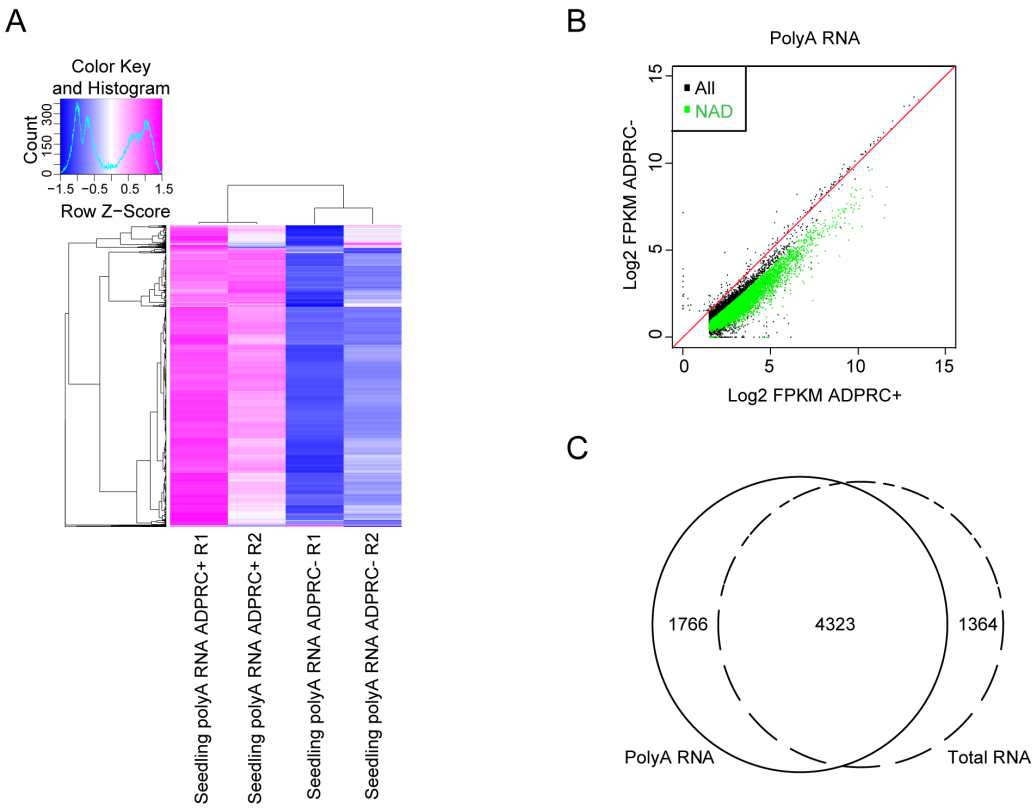

Figure S5. NAD captureSeq with polyA RNAs from seedlings. (A) Clustering analysis of NAD captureSeq with polyA RNAs from seedlings. Two biological replicates were performed. The 10000 top-varying transcripts were used in the analysis. (B) Comparison of read counts in ADPRC-treated (ADPRC+) and mock-treated (ADPRC-) samples of polyA RNA NAD captureSeq. Green dots indicate NAD-RNAs; black dots indicate non-NAD-RNAs. **"**All", all genes with detectable transcripts. "NAD", genes producing NAD-RNAs. (C) Venn diagram showing the degree of overlap among genes producing NAD-RNAs in total RNA and polyA RNA NAD caputreSeq.

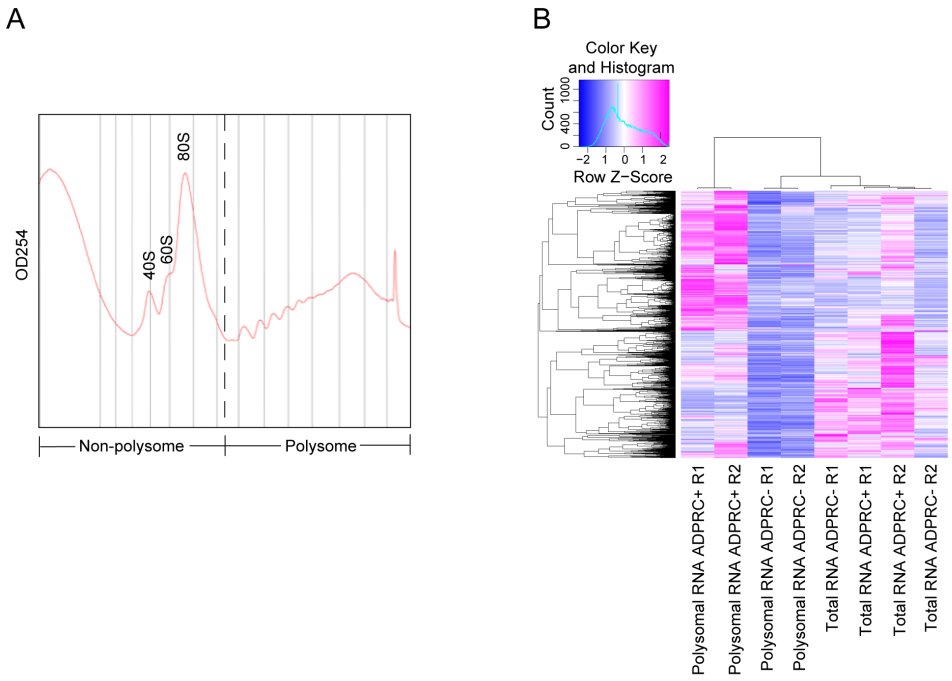

Figure S6. NAD captureSeq with total RNAs and polysomal RNAs. (A) A representative profile of ribosomes along a 15-60% sucrose gradient. The fractions to the left of the dotted line contain ribosomal subunits and monosomes, while the fractions to the right contain polysomes. (B) Clustering analysis of NAD captureSeq with polysomal RNAs and total RNAs isolated from the same plant samples. The 10000 top-varying transcripts were used in the analysis.

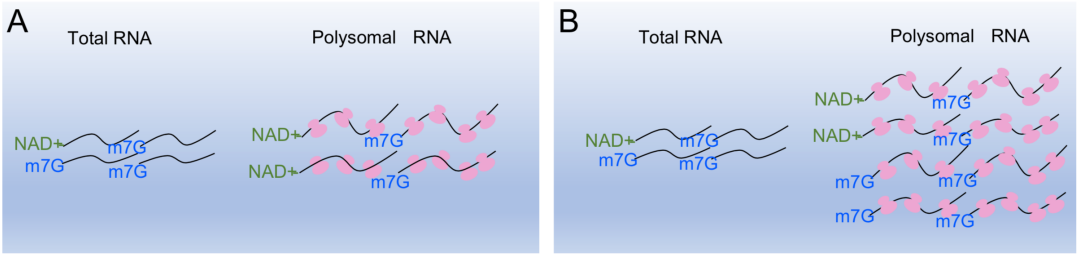

Figure S7. Diagrams depicting two scenarios with respect to NAD-RNAs on polysomes. Black curves represent multiple transcripts derived from one gene. (A) NAD+-capped transcripts are preferentially polysome-associated relative to m7G-capped transcripts. The proportion of NAD-RNAs is higher in polysomal RNAs than in total RNAs. (B) The proportion of NAD-RNAs is similar in total RNAs and polysomal RNAs, but the level of NAD-RNAs maybe higher in the polysomal fraction as the level of total transcripts (both NAD+-capped and m7G-capped) is higher in the polysomal fraction.

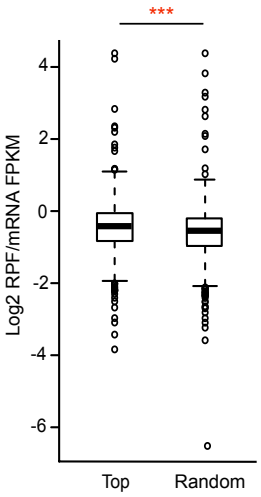

Figure S8. Genes with higher levels of NAD-RNAs have higher translation efficiency. Published ribo-seq data (GSE50597, no stress group)(1) from plants at a similar developmental stage as those in our NAD captureSeq were used to derive the levels of ribosome protected fragments (RPF). RNA-seq from the same published study was used to derive the mRNA FPKM. The RPF/mRNA FPKM ratio represents translation efficiency. The translation efficiencies of 1000 genes with the highest ADPRC+/ADPRC- ratio (labeled as "Top") and 1000 randomly selected genes (labelled as "Random") were compared. (***) $p < 0.001$ (two-tailed Wilcoxon test).

References

1. P. Juntawong, T. Girke, J. Bazin, J. Bailey-Serres, Translational dynamics revealed by genome-wide profiling of ribosome footprints in Arabidopsis. *Proc Natl Acad Sci U S A* **111**, E203-212 (2014).

Supplemental Table

Table S1. Oligonucleotides used in this study

Name	**Sequence**
IVT template	CAG**TAATACGACTCACTAT**TAGGCCTCTCGCTCTGCTGGGTGTGCG CTTGCTTGGCTTGCTT GTGGTCTGCGGTTCGTTCCCGCTTTGGT
IVT template PCR	Forward: CAGTAATACGACTCACTATTAGGCC Reverse: CCCCCCCCCCCCCCCCCACCAAAGCGGGAACGAAC
AT5G38420 RT-PCR	Forward: AGTCATCCGCTTCTTTCCCG Reverse: ACGGTACACAAATCCGTGCT
AT5G58250 RT-PCR	Forward: ACCTCCATCGAGCAACAGTC Reverse: ATTGCTCCTGGAAGTGCTCC
AT5G42530 RT-PCR	Forward: GCACAAGAGGAAGCCCAGA Reverse: CCAGGTTCGAGTGTGCCTTT
AT2G03340 RT-PCR	Forward: CTCTGATCCGGAGCCAAACC Reverse: GTAGCCGGTCTGTGTGTTGT
AT4G38840 RT-PCR	Forward: AGCAGATTCTCCGACAAGCC Reverse: CAAGGGATTGTGAGGCCACC
ATMG00110 RT-PCR	Forward: CGTCGTAACGCCCTTAATGC Reverse: ATTCTCATACCGGTGGCAGC
ATMG01170 RT-PCR	Forward: ACTAACGACGTTCTGCCAGG Reverse: CAGTTTGCTCCCGAGGAGTT
ATMG01130 RT-PCR	Forward: ACGAAGCCTCCTCCTCAGAT Reverse: CTATTCCTCGGTAAGCGGGC
ATMG00160 RT-PCR	Forward: TCGCGCTTTATGGCATTTCC Reverse: AGTGAGTGACTGCTCATCGG
ATMG00516 RT-PCR	Forward: TGTTCTTTCCAGGAGGTTGGC Reverse: CTTTCCGGCCAAGTCCCATT
ATMG00670 RT-PCR	Forward: ACGTGCGTTTCATTCTGTGC Reverse: AGCCTGACCACCCAGTAGAT
ATCG01110 RT-PCR	Forward: ACCGTATGTAACGCGATGGG Reverse: ATTCTCATACCGGTGGCAGC

ATCG00430 RT-PCR	Forward: TCCTAGACAGGCGGACCTTA Reverse: TGGTTTAGGTGGACAACCCG
5.8S rRNA RT-PCR	Forward: CAACGGATATCTCGGCTCTC Reverse: AACTTGCGTTCAAAGACTCG
18S rRNA RT-PCR	Forward: AATGAACGAGACCTCAGCCT Reverse: CCCAGAACATCTAAGGGCAT
SNOR77 RT-PCR	Forward: GACATGAGAGGTGTTTAACT Reverse: GTGCTTAAGACTACTAGACCA
SNO111 RT-PCR	Forward: GTGCTTAAGACTACTAGACCA Reverse: GTGCTTAAGACTACTAGACCA
RBCS1B/*2B*/ *3B* probe	Forward: CGGCTTGAAGTCATCCGCT Reverse: TCCGTGTTTTGTGTTTATTACCTC