nature communications

Article

https://doi.org/10.1038/s41467-026-71487-9

High-resolution mapping reveals features of bacterial NAD-capped RNAs and stress-responsive transcription initiation

Received: 13 June 2025

Accepted: 20 March 2026

Published online: 07 April 2026

Check for updates

A list of authors and their affiliations appears at the end of the paper

The NAD^+ cap has been discovered in RNAs across prokaryotes and eukaryotes, suggesting a possible role of NAD capping in gene regulation. Current NAD-capped RNA (NAD-RNA) profiling methods lack precision in 5′-end mapping or bias against small NAD-RNAs. Here, we introduce precision NAD-RNA sequencing (pNAD-seq), which combines a two-step enrichment strategy with high-throughput sequencing to achieve single-nucleotide resolution of 5′-ends and unprecedented sensitivity for identifying small NAD-RNAs. We further develop NAD-linkSeq to determine full-length NAD-RNA sequences. Applying these methods to *E. coli*, we uncover a vast repertoire of NAD-RNAs, including tRNAs, rRNAs, intragenic transcripts, and antisense RNAs, many of which are significantly shorter than regular mRNAs, implying specialized biogenesis. High-resolution mapping reveals conserved promoter architectures driving NAD-RNA production and condition-dependent initiation dynamics: under nitrogen limitation, some RNAs, alternative promoter usage, and coordinated expression shifts, correlating with metabolic stress responses. This report presents findings that offer a comprehensive view of NAD-RNAs in *E. coli* and introduces reliable methods for genome-wide profiling of NAD-RNAs across different organisms, which will facilitate the functional characterization of NAD-capping.

Eukaryotic messenger RNAs (mRNAs) typically feature a 7-methylguanylate (m^7G) cap at their 5′-end, a critical modification that protects them from 5′−3′ exonucleolytic degradation and promotes essential processes such as mRNA processing, nuclear export, and translation initiation via interactions with cap-binding proteins[1–3]. In contrast, prokaryotic RNAs were long thought to possess only a 5′-triphosphate terminus derived from the initiating nucleotide of transcription[4]. However, recent studies have revealed the presence of nicotinamide adenine dinucleotide at the 5′-end of certain bacterial transcripts[5,6]. Intriguingly, NAD-capped RNAs (NAD-RNAs) are not restricted to prokaryotes; they have also been identified in diverse eukaryotes, including yeast[7], mammals[8,9], and plants[10,11], suggesting a broader biological significance of this modification.

The NAD captureSeq technique[5] was the first method developed for transcriptome-wide profiling of NAD-RNAs. In this approach, the nicotinamide moiety of NAD is enzymatically modified with an alkyne group using ADP-ribosyl cyclase (ADPRC). The alkyne-labeled NAD-RNAs are then biotinylated via copper-catalyzed azide-alkyne cycloaddition (CuAAC), enabling their selective enrichment with streptavidin-coated beads. The purified RNAs are subsequently converted into cDNA libraries for high-throughput sequencing. This method has successfully identified NAD-RNAs in diverse organisms, including bacteria[5], yeast[7], mammalian cells[8], and the plants Arabidopsis[10], and rice[12].

Alternative methods for transcriptome-wide identification of NAD-RNAs include NAD tagSeq[11] and NAD tagSeq II[13]. Both techniques use ADP-ribosyl cyclase (ADPRC) to label NAD caps, but differ in their

e-mail: hl.zhang@zwu.edu.cn; zwcai@eitech.edu.cn; yxia@hkbu.edu.hk

downstream chemistry: NAD tagSeq employs copper-catalyzed azide-alkyne cycloaddition (CuAAC) to attach an RNA tag for Oxford Nanopore sequencing, while NAD tagSeq II replaces CuAAC with copper-free strain-promoted azide-alkyne cycloaddition (SPAAC) to minimize copper-induced RNA fragmentation. Unlike enrichment-based approaches, these methods enable direct quantification of both NAD-RNAs and untagged transcripts in the same sequencing run. The SPAAC chemistry has also been adapted into SPAAC-NAD-seq for profiling NAD-RNAs in Arabidopsis and a similar method, NADcapPro[14], has been applied in yeast.

Current methods for NAD-RNA detection face significant challenges in precisely mapping the 5′-ends of NAD-capped transcripts. This limitation impedes both mechanistic studies of NAD capping and functional characterization of NAD-RNAs. In the original NAD captureSeq protocol[5], first-strand cDNA synthesis occurs while NAD-RNAs remain immobilized on streptavidin resin, where the bulky biotin moiety at the 5′-end can interfere with reverse transcriptase processivity, preventing full-length cDNA synthesis. Modified versions of NAD captureSeq[7,8,10] and ONE-Seq[15] employ standard RNA-seq library preparation strategies (random priming or template switching) to improve transcript coverage, but these approaches compromise accurate 5′-end mapping[16,17]. Similarly, while NAD tagSeq[11] and NAD tagSeq II[13] enable direct sequencing of NAD-RNAs without enrichment, the non-canonical nucleotides introduced during click chemistry prevent precise identification of the native 5′-terminus.

Current NAD-RNA detection methods also struggle to effectively identify short NAD-capped transcripts. Standard cDNA library preparation protocols, optimized for typical mRNAs, and the original NAD captureSeq protocol[5] that omits RNA fragmentation, both introduce biases against shorter RNAs through size selection, adapter ligation, and PCR amplification steps. This results in inefficient capture and underrepresentation of small NAD-RNAs in sequencing libraries. While NAD tagSeq[11] and NAD tagSeq II[13] offer direct sequencing capabilities, they are similarly limited by Nanopore technology's inherent bias against small transcripts. These technical constraints create a significant gap in our ability to comprehensively profile the NAD-capped transcriptome, as shorter but potentially important NAD-RNA species are frequently lost or underrepresented in current analyses.

NAD-RNAs are primarily formed when RNA polymerase initiates transcription using NAD^+ instead of ATP as the starting nucleotide. This co-transcriptional capping occurs at specific transcription start sites (TSS) and is influenced by promoter sequences[18,19]. However, evidence suggests that NAD^+ may also be incorporated post-transcriptionally in some instances[8,14]. Using NAD tagSeq, Zhang et al. demonstrated that while NAD-capped RNAs represent only ~0.5% of non-rRNA transcripts in *E. coli* overall, certain genes produce predominantly NAD-RNAs under specific growth conditions[13]. This finding indicates that NAD-capping may be a regulated process with potential biological significance.

To overcome existing limitations in NAD-RNA profiling, we developed precision NAD-RNA sequencing (pNAD-seq) and its modified version, NAD-linkSeq, which combine SPAAC-mediated biotinylation for specific NAD-RNA enrichment, NudC processing to convert NAD-capped RNAs to 5′-monophosphate RNAs for selective amplification, and a linker-based cDNA library construction strategy for high-resolution sequencing. When applied to *E. coli*, pNAD-seq demonstrated superior sensitivity, accuracy, and 5′-end resolution compared to existing methods, enabling full-length NAD-RNA characterization. Our analysis revealed an unanticipated diversity of NAD-capped transcripts, including small intragenic and antisense RNAs. Our findings indicate that NAD-RNAs represent a diverse class of RNAs that may play significant regulatory roles. These findings not only expand the known repertoire of NAD-RNAs but also establish a robust framework for investigating the biological significance of NAD-capping.

Results

High-efficiency biotinylation of NAD-RNA through SPAAC tagging

Our previous work established SPAAC-based tagging for NAD-RNA labeling in methods such as NAD tagSeq II[13] and SPAAC-NAD-seq[20]. Building on this, we developed precision NAD-RNA sequencing (pNAD-seq), which relies on ADPRC-mediated transglycosylation of the 5′-NAD moiety in the presence of azide to generate clickable azide-NAD-RNA. This product is then conjugated to biotin via SPAAC reaction with biotin-DBCO (Fig. 1a). To optimize the SPAAC tagging efficiency in pNAD-seq, we first tested the system using an in vitro-transcribed NAD-capped 24mer RNA. Following ADPRC treatment and biotin-DBCO conjugation, we observed a slow-running band in urea-PAGE, absent in the no-ADPRC control (Fig. 1b), confirming successful biotinylation. Quantification of the biotinylated and unbiotinylated bands revealed a tagging efficiency of ~94%. Next, we determined the optimal ADPRC concentration by titrating the enzyme and monitoring substrate conversion. After incubating NAD-RNA with varying ADPRC concentrations and subsequent biotin-DBCO labeling, urea-PAGE separation allowed quantification of biotinylated vs. unbiotinylated RNA. The resulting titration curve (Fig. 1c) indicated that 42 nM ADPRC was sufficient to achieve 96% conversion efficiency. Our tests showed that the amount of ADPRC enzyme used in previous studies can be significantly reduced while still working effectively. We used the original, higher concentration[5] in this study to guarantee reliable results, but this finding means future studies can achieve the same outcome at a much lower cost. We further compared SPAAC tagging with the one-step biotinylation approach, where NAD-RNA is directly labeled using biotin-PEG-OH via ADPRC[15]. Under identical conditions, SPAAC tagging achieved 92% efficiency, whereas the one-step method reached only 50% (Fig. 1d), demonstrating the superior efficiency of SPAAC-based biotinylation.

Specificity of SPAAC tagging for NAD-RNA

Having validated the high efficiency of SPAAC tagging, we assessed its specificity using RNAs with diverse 5′-ends, including NAD-capped 25mer RNA (NAD+24nt), uncapped A-RNA, and RNAs with various biologically relevant 5′-modifications (pppA[4], ppRNA[21], pRNA[22], Ap3A-RNA[23], Ap4A-RNA[24], 5′-OH RNA[25], and dpCoA-RNA[26]). While NAD-RNA showed near-complete biotinylation (98%), none of the other tested RNAs were labeled, confirming high specificity (Supplementary Fig. 1a–c). NADH-RNA, though recognized by ADPRC, exhibited significantly lower biotinylation efficiency (~20%) compared to NAD-RNA (Supplementary Fig. 1d), suggesting that ADPRC-mediated tagging selectively enriches NAD-RNAs with minimal cross-reactivity toward NADH-RNA.

NudC efficiently hydrolyzes the pyrophosphate bond of biotinylated NAD-RNA

The *E. coli* Nudix hydrolase NudC[5], a known NAD-RNA decapping enzyme, cleaves the pyrophosphate bond of NAD-RNA. Using purified recombinant NudC (Supplementary Fig. 2), we confirmed that NudC also hydrolyzes biotinylated NAD-RNA generated via SPAAC tagging, producing a biotin-modified riboside moiety and 5′-monophosphorylated RNA (pA-RNA) (Fig. 1e). Urea-PAGE analysis revealed faster migration of the NudC-treated product (Fig. 1f), and mass spectrometry identified the expected biotin-containing fragment ($m/z = 1063.421$) (Supplementary Fig. 3), confirming pyrophosphate bond cleavage. Time-course experiments showed complete decapping of biotinylated NAD-RNA within 4 min (Fig. 1g). We further examined how 5′-secondary structures affect decapping efficiency[27] by testing linear, extended, blunt, and recessed ends. While linear and extended structures required lower NudC concentrations, complete decapping of blunt or recessed ends needed ≥500 nM enzyme (Fig. 1h), demonstrating that 5′-end accessibility influences NudC activity. These results

https://doi.org/10.1038/s41467-026-71487-9

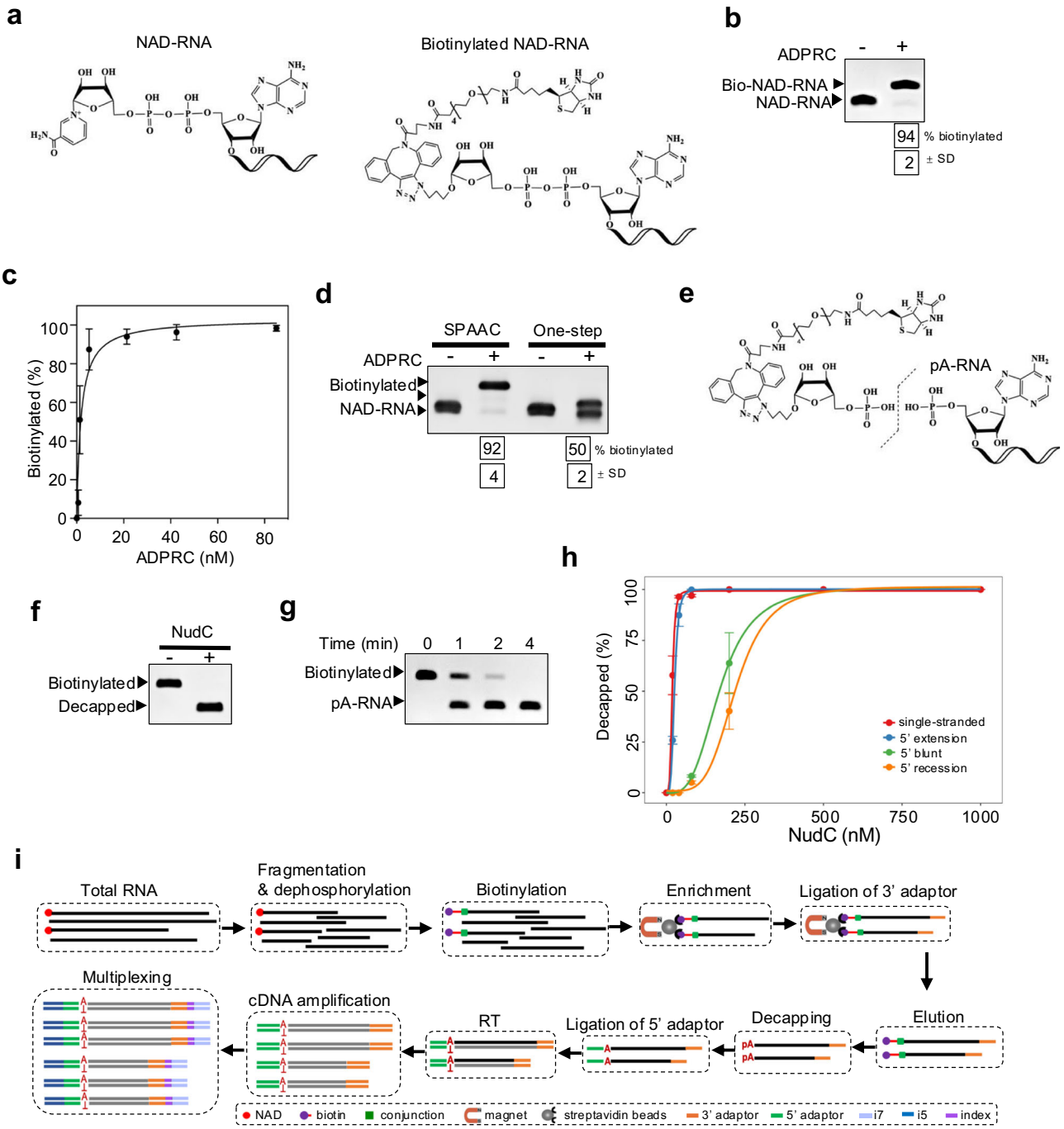


Fig. 1 | The pNAD-seq method for high-efficiency NAD-RNA capture and sequencing. a Schematic of NAD-RNA biotinylation via SPAAC tagging. The 5′-NAD cap is modified by ADPRC with azido-propanol, followed by biotin-DBCO conjugation. Expected products after NudC decapping are shown. **b** Biotinylation efficiency of NAD-24mer RNA. NAD-RNA was treated with or without ADPRC (in the presence of azido-propanol), then reacted with biotin-DBCO. Products were resolved by urea-PAGE and stained. Biotinylation efficiency was calculated as the fraction of biotinylated RNA relative to total RNA. Data are presented as mean ± SD of three independent experiments ($n = 3$). **c** ADPRC titration for optimal biotinylation. NAD-RNA was incubated with varying ADPRC concentrations, followed by biotin-DBCO labeling. The biotinylated fraction was quantified as in (**b**). Data are presented as mean ± SD of three independent experiments ($n = 3$). **d** Comparison of SPAAC tagging vs. one-step biotinylation. Data are mean ± SD of three independent experiments ($n = 3$). **e** Expected NudC decapping products. NudC cleaves the biotinylated phosphoribosyl moiety, leaving 5′-monophosphate RNA (5′-pA-RNA). **f** NudC decapping assay. Biotinylated NAD-24mer RNA was incubated ± NudC and analyzed by urea-PAGE. Shown is a representative gel image from at least three independent assays. **g** Time-course decapping by NudC (230 nM). Reaction products were resolved by urea-PAGE and quantified over time. Data reflect three independent experiments. **h** Impact of 5′ structures on NudC-mediated decapping. Decapping efficiency was measured across NudC concentrations. Decapping (%) was quantified across NudC concentrations for RNAs with single-stranded, 5′ extension, 5′ blunt, and 5′ recessed ends. Points show mean ± SD from three independent experiments ($n = 3$). **i** pNAD-seq workflow. Total RNA is fragmented, dephosphorylated (to block non-NAD-RNA ligation), and subjected to SPAAC tagging (in parallel with no-ADPRC control). Biotinylated NAD-RNAs are enriched, ligated with 3′- and 5′-adapters post-decapping, and sequenced. Structures of the reaction products are illustrated where relevant. Source data are provided as a Source data file.

establish that NudC efficiently processes biotinylated NAD-RNA for downstream applications while being sensitive to RNA structural context.

Precision NAD-RNA sequencing (pNAD-seq) method

We developed pNAD-seq as a two-step enrichment strategy for high-specificity NAD-RNA identification (Fig. 1i). The method proceeds as follows: (1) Total RNA is fragmented to ~200 nt and treated with phosphatase to remove 5′ phosphates (preventing non-NAD-RNA ligation in later steps) while preserving 3′-OH groups for adapter ligation; (2) SPAAC tagging biotinylates NAD-capped RNAs (with ADPRC− control); (3) Streptavidin bead enrichment isolates NAD-RNAs; (4) A 3′-RNA adapter is ligated to bead-bound RNAs; (5) After elution, NudC decapping generates 5′-monophosphate ends (5′-pA-RNA); (6) A 5′-RNA adapter is ligated to the decapped ends; (7) cDNA synthesis and PCR amplification are performed using adapter-specific primers (forward primer complementary to 5′-adapter and reverse primer complementary to 3′-adapter); (8) Final libraries are indexed and sequenced. NAD-RNAs are identified by: precise 5′-end adenosine mapping and ADPRC+ enrichment over control. Compared to NAD captureSeq, pNAD-seq provides nucleotide-resolution 5′-end mapping and enhanced specificity through both streptavidin enrichment and selective PCR amplification of properly adapter-ligated, decapped NAD-RNAs.

Profiling the NAD-capped transcriptome in *E. coli* by pNAD-seq

We validated pNAD-seq using stationary-phase *E. coli* total RNA spiked with a synthetic NAD-RNA standard (Supplementary Fig. 4). Three technical replicates demonstrated excellent reproducibility (Supplementary Fig. 5), with 55.8-fold enrichment of standard in ADPRC+ versus ADPRC− samples. The intentional depletion of non-NAD-capped RNAs during the ADPRC+ treatment was highly effective, as evidenced by the dramatically higher levels of rRNA and tRNA retained in the ADPRC− controls (Supplementary Fig. 6). Due to this significant and intentional difference in library composition, standard differential expression tools like DESeq2 were unsuitable for our enrichment analysis. Instead, we identified high-confidence NAD-RNAs by applying stringent thresholds (>20-fold enrichment in Reads Per Million (RPM), $p < 0.05$, and a count >150), yielding a final set of 820 NAD-RNAs (Fig. 2a). The full pNAD-seq results for all detected transcripts are provided in Supplementary Data 1, and the curated list of the 820 high-confidence NAD-RNAs (with counts, enrichment values, and statistics) is provided in Supplementary Data 2.

The (+1) A-read ratio (proportion of reads starting with NAD-derived adenosine) served as a key specificity metric. While background signals from incomplete dephosphorylation or RNA degradation showed the expected ~25% probability of A-starts (matching random nucleotide distribution), ADPRC+ samples exhibited strong 5′-A preference (71–100% A-read ratios) (Fig. 2b). High-confidence NAD-RNAs showed exceptional specificity with >94% (+1) A-read ratios, demonstrating pNAD-seq's ability to distinguish true NAD-capped transcripts from background.

pNAD-seq specifically captures 5′-terminal fragments of NAD-capped RNAs through selective streptavidin enrichment following fragmentation, enabling precise 5′-end mapping as demonstrated by the synthetic NAD-RNA standard, *gcvB*, *ssrA*, and *raiA* (Fig. 2c–f). Approximately 50% of the NAD-RNAs originated from canonical transcription start sites (TSSs), consistent with NAD^+ serving as an initiating nucleotide, while the remaining transcripts revealed unexpected NAD-capping events: one-third of protein-coding NAD-RNAs (e.g., *prmB*, Fig. 2g) mapped to intragenic regions, suggesting internal alternative promoters; 118 antisense RNAs (e.g., *glpF*, Fig. 2h) implied NAD-dependent initiation on the opposite strand; and structured ncRNAs, including tRNA glyW (Fig. 2i) and 16S rRNA rrsA fragments (Fig. 2j), expanded the diversity of NAD-capped transcripts beyond mRNA and non-coding sRNA. In contrast, 5S rRNA, previously reported to lack an NAD cap[5], showed neither enrichment nor a +1 A profile (Fig. 2k), serving as a negative control.

pNAD-seq significantly outperformed previous methods, identifying 42 of the 47 *E. coli* NAD-RNAs reported by NAD captureSeq and all 279 previously detected by NAD tagSeq II (Fig. 2l), demonstrating superior sensitivity and expanding the known catalog of NAD-capped transcripts. Beyond increased detection, the method's precision revealed unexpected complexity in NAD capping—including widespread intragenic transcription and antisense NAD-RNAs—suggesting regulatory mechanisms more diverse than previously recognized.

To validate the NAD-RNAs identified by pNAD-seq, we selected 14 candidates for confirmation: six previously reported NAD-RNAs and eight candidates uniquely detected by our method (Fig. 2m and Supplementary Fig. 7). Using biotinylated oligonucleotides, we affinity-purified these transcripts from *E. coli* total RNA and analyzed their 5′-caps by nuclease P1 digestion followed by LC-MS. All 14 RNAs showed a clear NAD^+ peak ($m/z = 664.1$), confirming their NAD-capped status. The eight pNAD-seq specific targets included both non-coding RNAs (glmZ sRNA, tRNAs glyW/metZ/glyU, and tmRNA ssrA) and protein-coding transcripts (spy, glpF, and phoH) (Fig. 2k). In contrast, no NAD^+ peak was detected in 5S rRNA, which was also negative by pNAD-seq. Notably, we detected no NADH ($m/z = 666.1$) in these samples (Supplementary Fig. 8), suggesting either minimal abundance of NADH-capped RNAs or technical limitations in their detection.

Classification of NAD-RNAs

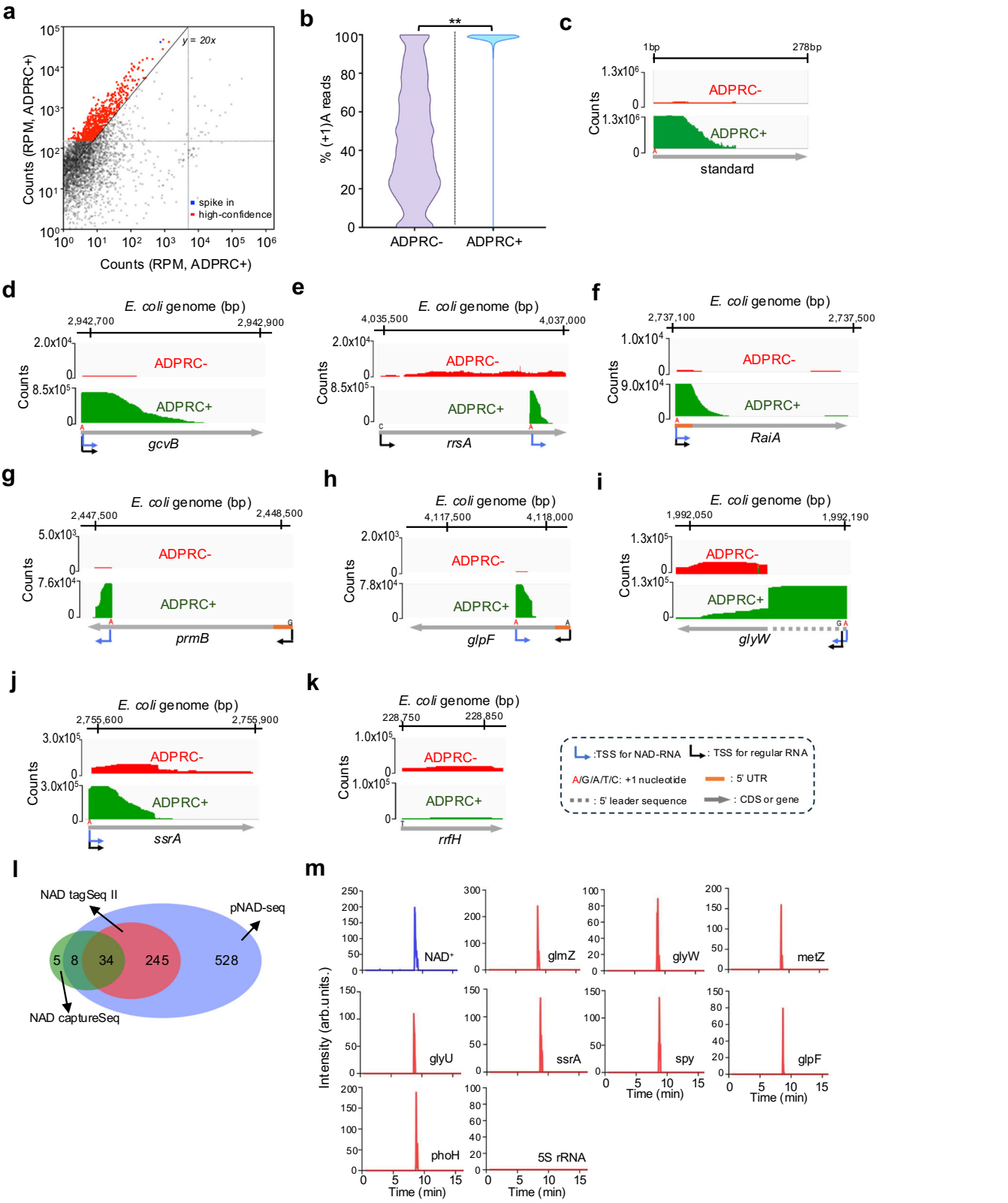
Among the 820 NAD-RNAs identified, 35 originated from non-coding genes, including the 16S rRNA gene (Fig. 2j), 22 sRNAs (Fig. 3a), 11 tRNAs, and one tmRNA. Gene-body coverage of the NAD-sRNA reads showed that they closely matched their annotated sequences (Fig. 3b). Notably, 9 of the 11 NAD-tRNAs were transcribed from alternative TSSs, suggesting regulatory variants (Fig. 3c). The remaining 785 NAD-RNAs derived from protein-coding genes were classified into three subgroups based on TSS position and strand orientation (Fig. 3d, Supplementary Data 3): 5′-Region ($n = 401$): canonical 5′-UTR sense transcripts; Internal ($n = 268$): intragenic sense transcripts; Antisense ($n = 116$): antisense-strand transcripts, implicating potential regulatory roles. Gene Ontology (GO) enrichment analysis of NAD-RNA-producing genes revealed associations with metabolic regulation, particularly nucleobase-containing compound metabolism, and negative regulation of macromolecular processes (Fig. 3e).

Identification of promoter elements associated with NAD capping

To elucidate the mechanism of NAD capping, we analyzed the promoter sequences of NAD-RNA genes using the precise transcription start sites mapped by pNAD-seq. We focused on the top 199 NAD-RNA genes (representing >76% of total NAD-RNA reads), grouping their promoters as 5′-region, intragenic, or antisense (Supplementary Data 4). Sequence alignment revealed significant enrichment of a −10 TATA box and a −35 TT element in both promoter classes (Fig. 3f). Furthermore, we identified a conserved RDAY motif (positions −2 to +2) in a subset of 75 promoters (Fig. 3g). The RDAY motif includes a +1 A, which is necessary for NAD incorporation; however, +1 A alone is not sufficient; the corresponding sequences are listed in Supplementary Data 5. Promoters lacking RDAY also show a strong preference for +1 A transcription start sites (Fig. 3h), and only a subset of +1 A promoters yield NAD-RNAs. Accordingly, RDAY is enriched among promoters that produce NAD-capped transcripts but is not exclusive to them; it should be viewed as a predictive feature rather than a unique signature of NAD capping.

To test whether the RDAY motif affects transcriptional output of NAD-RNA-producing genes, we compared total RNA levels between promoters with and without the motif and observed no significant

https://doi.org/10.1038/s41467-026-71487-9



difference (Fig. 3i). We then assessed NAD incorporation efficiency by comparing the ADPRC+/ADPRC− read count ratio for transcripts arising from RDAY versus non-RDAY promoters. Because ADPRC− broadly reflects total transcript abundance while ADPRC+ corresponds to NAD-RNA levels, this ratio serves as a proxy for NAD-capping efficiency. A statistically significant difference was observed between the two groups (Fig. 3i), suggesting a potential association between the RDAY motif and higher NAD capping efficiency, which appears to be independent of transcriptional activity. Together, these results define distinct promoter features linked to NAD capping: a core architecture, comprising a −10 TATA box, an extended −10 element, and a +1 A start site, accounts for most NAD-RNAs, whereas the RDAY motif marks a

Fig. 2 | Identification of NAD-RNAs in *E. coli* by pNAD-seq. a Scatter plot comparing RNA abundance between ADPRC-treated and control samples. Red and gray dots denote RNAs with average counts >150 and ≤150 in Reads Per Million (RPM), respectively, across three replicates; the blue square indicates the synthetic NAD-RNA standard. The diagonal line (y = 20x) marks the 20-fold enrichment threshold. High-confidence NAD-RNAs were defined as those with mean abundance >150 RPM across replicates, >20-fold enrichment, and a paired two-tailed Student's *t*-test $p < 0.05$; these are shown in red, and all other RNAs are in gray. Per-RNA *p* values are reported in Supplementary Data 1. **b** Violin plot depicts the distribution of ratios of (+1) A reads for each gene in ADPRC-treated (ADPRC +) samples and control (ADPRC−) samples. The y-axis represents the (+1) A ratio, calculated as the percentage of reads with an A as the first nucleotide to the total reads for each gene. The x-axis indicates the treatment conditions. Each dot represents the average of three replicates. Statistical significance was determined by a paired Student's *t*-test (**$p < 0.01$; two-tailed). **c**–**k** Genomic mapping of NAD-RNA sequencing reads. IGV tracks display read alignments from ADPRC-treated (ADPRC +, green) and untreated control (ADPRC−, red) libraries across representative loci: spike-in NAD-RNA control (**c**), small RNA gene *gcvB* (**d**), tmRNA gene *ssrA* (**e**), protein-coding genes *raiA* (**f**), *prmB*(g), and *glpF* (**h**), tRNA gene *glyW* (**i**), 16S ribosomal RNA gene *rrsA* (**j**), and 5S rRNA gene *rrfH* (**k**). Tracks display genomic coordinates along the top and raw read counts on the y axis. Gene models depict coding sequences/genes as gray arrows, 5′ UTRs in orange, and 5′ leader sequences as gray dashed lines. The +1 nucleotide/TSS is indicated: NAD-RNA TSSs in blue and EcoCyc[37]–annotated TSSs for regular RNA in black; where noted, the identity of the +1 base (A/G/T/C) is labeled. **l** The Venn diagram depicts the overlap of the NAD-RNA species identified by the previous methodologies with those identified by the pNAD-seq method. Each circle within the diagram corresponds to the NAD-RNA species identified by a specific method, and the intersecting areas indicate the species that are commonly identified across the methods. The pNAD-seq method detected nearly all previously known NAD-RNA species and uncovered an additional 528 species, showcasing its enhanced sensitivity. **m** Validation of NAD-capping on pNAD-seq-identified RNAs by LC-MS. LC-MS analysis confirms the presence of the NAD cap on a subset of RNAs identified by pNAD-seq. The histogram shows the detection of NAD on eight selected RNAs from the 528 specifically enriched targets, with 5S rRNA serving as a control. RNA was purified from *E. coli* total RNA using a biotinylated DNA probe. The purified RNAs were digested with nuclease P1 and subjected to LC-MS analysis. The histogram displays a specific fragment ion (*m*/*z* 428.001), which originates from a precursor ion with a mass-to-charge ratio (*m*/*z*) of 624.001, indicative of NAD presence. (arb. units) denotes arbitrary units.

smaller subset. These conserved sequences could contribute to the recruitment or orientation of transcription machinery to facilitate co-transcriptional NAD incorporation.

NAD-linkSeq demonstrates that most of NAD-RNAs are shorter than their regular mRNA counterparts

The pNAD-seq results indicated that NAD-RNAs are predominantly small (Supplementary Fig. 9). However, because pNAD-seq identifies only the 5′ regions of NAD-RNAs, it cannot resolve full-length sequences. Our previous method, NAD tagSeq II[13], enables direct analysis of full-length NAD-RNAs by Oxford Nanopore direct RNA sequencing, but the SPR bead-based purification step removes short RNAs, potentially biasing size representation. To overcome these limitations, we developed a modified pNAD-seq workflow, termed NAD-linkSeq, for full-length NAD-RNA sequencing (Fig. 4a).

In NAD-linkSeq, total RNA is spiked with four NAD-RNA standards (see Methods) and left unfragmented. As in pNAD-seq, NAD-RNAs are biotinylated via SPAAC tagging and enriched. Following ligation of a 3′ adapter, full-length RNAs undergo NudC-mediated decapping and are subsequently ligated to an extended 5′ adapter. The resulting RNAs are then reverse transcribed to generate full-length cDNAs, which are directly sequenced. Because Nanopore DNA (cDNA) sequencing more efficiently captures shorter fragments than direct RNA sequencing, cDNAs are analyzed using Oxford Nanopore direct cDNA sequencing. Following adapter trimming, clean cDNA reads with A as the first nucleotide were mapped to the *E. coli* reference genome.

This analysis showed that NAD-RNAs originating from non-coding genes generally matched the annotated full transcript lengths (Fig. 3b), whereas NAD-RNAs derived from protein-coding genes typically ranged from 50 to 450 nt, with an average length of ~200 nt; the four internal standards were recovered at their expected sizes (Fig. 4b, Supplementary Data 6). These findings are consistent with previous reports that, even in the absence of RNA fragmentation, specific 5′ terminal fragments, rather than full-length mRNAs, are preferentially enriched in *E. coli* NAD-RNAs[5]. Comparing gene-body coverage for NAD-RNAs (NAD-linkSeq) with their mRNA counterparts (RNA-seq) further revealed that NAD-RNAs predominantly occupy the 5′ regions of protein-coding loci (Fig. 4c), as exemplified by read mapping to *trmJ* (Fig. 4d).

To validate the observed size differences between NAD-RNAs and their cognate mRNAs, total RNA was subjected to SPAAC-mediated biotinylation and streptavidin pull-down to enrich NAD-RNAs. The enriched fraction and the input RNA were then analyzed by Northern blot using a trmJ-specific probe. The trmJ NAD-RNA size determined by Northern blot (Fig. 4e) matched the NAD-linkSeq result. Notably, gene-coverage analysis identified 40 cases in which NAD-RNAs correspond to full-length mRNAs (Fig. 4g). Among these, NAD-acpP was previously reported by NAD tagSeq II[13]; consistent with that, Northern analysis confirmed that NAD-acpP and acpP mRNA are similar in size (Fig. 4h). A complete list of these NAD-RNAs is provided in Supplementary Data 7.

Dynamic remodeling of the NAD-RNA landscape in response to metabolic stress

We investigated the dynamic remodeling of the NAD-RNA landscape in response to nitrogen deficiency by comparing *E. coli* cells grown in nitrogen-reduced M9 medium versus standard M9 conditions. Using pNAD-seq with spike-in NAD-RNA standards for normalization, we identified 811 NAD-RNAs, including 178 that showed significant expression changes (160 upregulated and 18 downregulated, $p < 0.05$) under nitrogen limitation (Fig. 5a; dataset in Supplementary Data 8). To validate these findings, we developed a NAD-PEGylation method using ADPRC and SPAAC to selectively label NAD-RNAs with a 10 kDa PEG tag, enabling their separation from regular transcripts by gel electrophoresis (Fig. 5b). Northern blot analysis of PEGylated RNAs confirmed the pNAD-seq results, showing increased NAD-gcvB and decreased NAD-gadY levels under nitrogen stress (Fig. 5c, d; see Supplementary Fig. 10 for lower-contrast versions of the blots). We then asked whether the observed shifts in NAD-RNA levels could be attributed primarily to transcriptional changes. To address this, we performed RNA-seq to assess the differential expression of total RNA between the two growth conditions (Supplementary Data 9). Comparison of $\log_2$ fold changes for NAD-RNA and total RNA under the two conditions revealed a positive overall correlation ($R^2 = 0.35$; Fig. 5e), suggesting that changes in NAD-RNA abundance often parallel changes in transcript levels. However, a notable subset of genes departed from this trend, displaying marked changes in NAD-RNA with little shift in total RNA, or vice versa. This pattern implies that, alongside transcriptional co-variation, NAD-capping efficiency may be subject to additional gene-specific modulation.

Conventional RNA-seq analysis indicated a metabolic shift under nitrogen stress, with upregulation of nitrogen scavenging genes and downregulation of core energy metabolism genes (Supplementary Fig. 11). In contrast, Gene Ontology (GO) enrichment analysis of genes with upregulated NAD-RNAs showed significant enrichment for transcriptional regulatory processes, including RNA biosynthetic process and regulation of DNA-templated transcription (Supplementary

https://doi.org/10.1038/s41467-026-71487-9

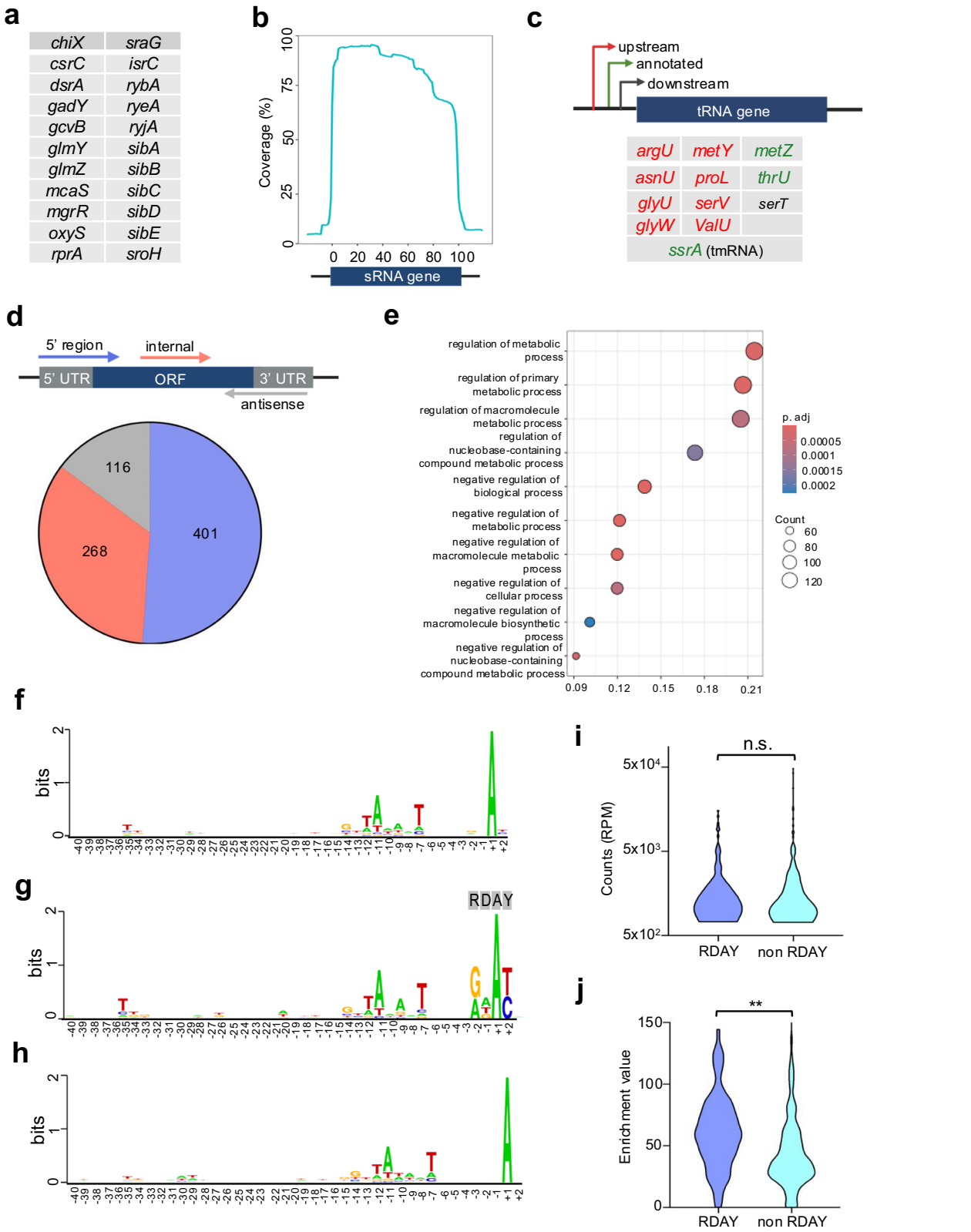


Fig. 12). The distinct functional categories associated with total mRNA changes versus NAD-RNA changes suggest these RNA populations may participate in different aspects of the nitrogen stress response.

Beyond expression-level changes, we identified significant condition-dependent alterations in transcriptional initiation patterns. Under nutrient stress conditions, several genes exhibited striking shifts in their transcription start site (TSS) selection. Notably, *mqsR* and *sibD* transitioned from utilizing their annotated TSSs in nutrient-rich LB medium to preferentially initiating at downstream TSSs in minimal media. This redistribution was particularly pronounced for *sibD*, which showed 46% downstream TSS usage in standard M9 medium, increasing to near-exclusive downstream initiation under nitrogen-limited conditions (Fig. 5f–i). A parallel phenomenon occurred with promoter selection, as demonstrated by *ftnB*'s switch from its

Fig. 3 | Characterization of *E. coli* NAD-RNAs identified by pNAD-seq. a Small regulatory RNAs with NAD caps include 22 sRNAs sharing an A-start TSS (bent arrows, EcoCyc-annotated). **b** Gene-body coverage profile of NAD-sRNA reads. Coverage is plotted across the gene body from 5′ to 3′ (percentile). The green bent arrow marks the transcription start site (TSS). Gene structure schematic was created with BioRender.com. **c** NAD-capped tRNAs/tmRNAs are classified by TSS position relative to reference: annotated, upstream, or downstream. Categories indicated by color-coded arrows. Gene structure schematic was created with BioRender.com. **d** Protein-coding gene derived NAD-RNAs fall into three groups: 5′ region (canonical 5′-UTR sense transcripts; internal (intragenic sense transcripts); and antisense (antisense strand transcripts). Created in BioRender. Zhang, B. (2026) https://BioRender.com/ngvvmvc. **e** Gene Ontology enrichment for NAD-RNA genes highlighting key biological processes. Enrichment was tested using a one-sided hypergeometric (Fisher's exact) over-representation test with Benjamini–Hochberg FDR correction; terms with FDR < 0.05 were considered significant. Sequence logos depicting conserved promoter motifs for NAD-RNAs. Logos show nucleotide conservation for all the 199 promoters (**f**), 75 promoters containing the RDAY motif (**g**), and 124 promoters lacking the RDAY motif (**h**). Nucleotide abbreviations: R (A/G), D (A/G/T), Y (C/T). Sequence logos were generated using WebLogo[38]. **i** Comparison of NAD-capped transcript levels (RPM) between promoters with and without the RDAY motif. n.s. ($p = 0.9693$; not significant); significance assessed by a two-sided Mann–Whitney U test. **j** The RDAY promoter motif is associated with higher NAD^+ incorporation efficiency. Violin plots show fold enrichment (ratio of ADPRC+/ADPRC− read counts) for NAD-capped transcripts from promoters containing the RDAY motif versus non-RDAY promoters. Statistical significance was determined by a two-sided Mann–Whitney U test; $^{**}p < 0.0001$.

annotated LB-medium promoter to a downstream promoter in minimal media conditions (Fig. 5j).

Discussion

Our development of pNAD-seq, which integrates optimized SPAAC tagging, NudC-mediated decapping, and adapter-ligation strategies, represents a significant advance in NAD-RNA detection and characterization that overcomes some key limitations of previous methods. The technique provides precise 5′-end mapping, enhanced capture of short transcripts, and exceptional sensitivity capable of detecting capping frequencies below 0.1%. This high sensitivity enabled the discovery of NAD-capped RNAs, including classes such as 16S rRNA fragments and tRNA that were not previously known to undergo NAD capping. Notably, we identified 11 NAD-tRNAs originating from both monocistronic and polycistronic clusters, with most utilizing alternative promoters containing functional −10 TATA elements. Gene ontology analysis revealed an enrichment of NAD-RNAs in metabolic and transcriptional regulatory pathways. Collectively, these findings, made possible by the superior performance of pNAD-seq, expand our understanding of NAD-RNA diversity.

While 5′-end secondary structures in decapped NAD-RNAs could theoretically hinder ligation efficiency, our data show this does not meaningfully impact pNAD-seq performance. The method successfully detected 42 of 47 known *E. coli* NAD-RNAs from NAD captureSeq[5] and all 279 NAD-RNAs reported by NAD tagSeq II[13], demonstrating superior sensitivity and coverage. This expanded detection reveals NAD-capping as a more widespread phenomenon than previously recognized.

Our analysis of transcription start site (TSS) reveals that NAD-capped RNA synthesis is associated with conserved promoter features consistent with canonical bacterial initiation mechanisms. While previous in vitro studies proposed a σ70-dependent HRRAS motif (positions −3 to +2) as a universal promoter element[18], we identified three predominant signatures in NAD-RNA gene promoters: a conserved −10 TATA box often accompanied by an extended −10 GT element, a less enriched −35 TT element, and a universal adenine at the +1 position. Furthermore, we discovered an enriched RDAY motif (positions −2 to +2) in a subset of 75 promoters of the genes producing the most abundant NAD-RNAs. The presence of these promoter-specific signatures, including for NAD-RNAs initiated from intragenic regions, strongly supports the previous findings that the NAD^+ cap is formed during transcription initiation.

Our observation that protein-coding gene-derived NAD-RNAs are much shorter than their full-length mRNA counterparts extends the earlier finding by Höfer et al.[5], who reported this phenomenon for specific transcripts. Using the NAD-linkSeq method, our study provides a transcriptome-wide analysis of NAD-RNA sizes, demonstrating that this shortened length is a prevalent characteristic of the NAD-capped RNA landscape from protein-coding genes in *E. coli*. This observation argues against a protein-coding function for NAD-RNAs.

NAD-linkSeq shows that most *E. coli* NAD-RNAs are shorter than their mRNA counterparts, whereas the results from NAD tagSeq II analysis, which uses Nanopore direct RNA sequencing, suggested that NAD-RNAs are similar in size to mRNAs[13]. This discrepancy reflects platform-specific biases: direct RNA library preparation includes bead-based size selection that enriches longer RNAs, and pore loading during direct RNA sequencing favors long molecules while capturing short RNAs inefficiently. These features skewed the NAD tagSeq II profile toward longer species. In contrast, NAD-linkSeq sequences cDNAs generated from adapter-ligated NAD-RNAs, and Nanopore cDNA sequencing efficiently recovers short fragments. By sequencing cDNA, NAD-linkSeq overcomes this bias and sensitively detects short NAD-RNAs, yielding a more comprehensive and accurate view of the NAD-capped transcriptome. Its orthogonal design complements existing methods for cross-validating and refining maps of NAD-capped RNAs.

Our discovery of condition-dependent NAD-RNA isoforms and promoter switching under nitrogen limitation aligns with emerging evidence that NAD capping is dynamically regulated by cellular metabolic state. While prior work showed that NAD-RNA levels are controlled by substrate competition (ATP/NAD^+) and NudC-mediated decay[28], our high-resolution mapping reveals an additional layer of regulation: transcriptional reprogramming via alternative TSS usage. The prevalence of short, non-coding NAD-RNAs (e.g., intragenic RNAs) suggests their potential roles in stress adaptation, possibly through modulation of RNA stability. Future studies should explore whether the observed promoter architectures and small NAD-RNAs coordinate with ppGpp signaling to fine-tune NAD-RNA turnover during metabolic stress.

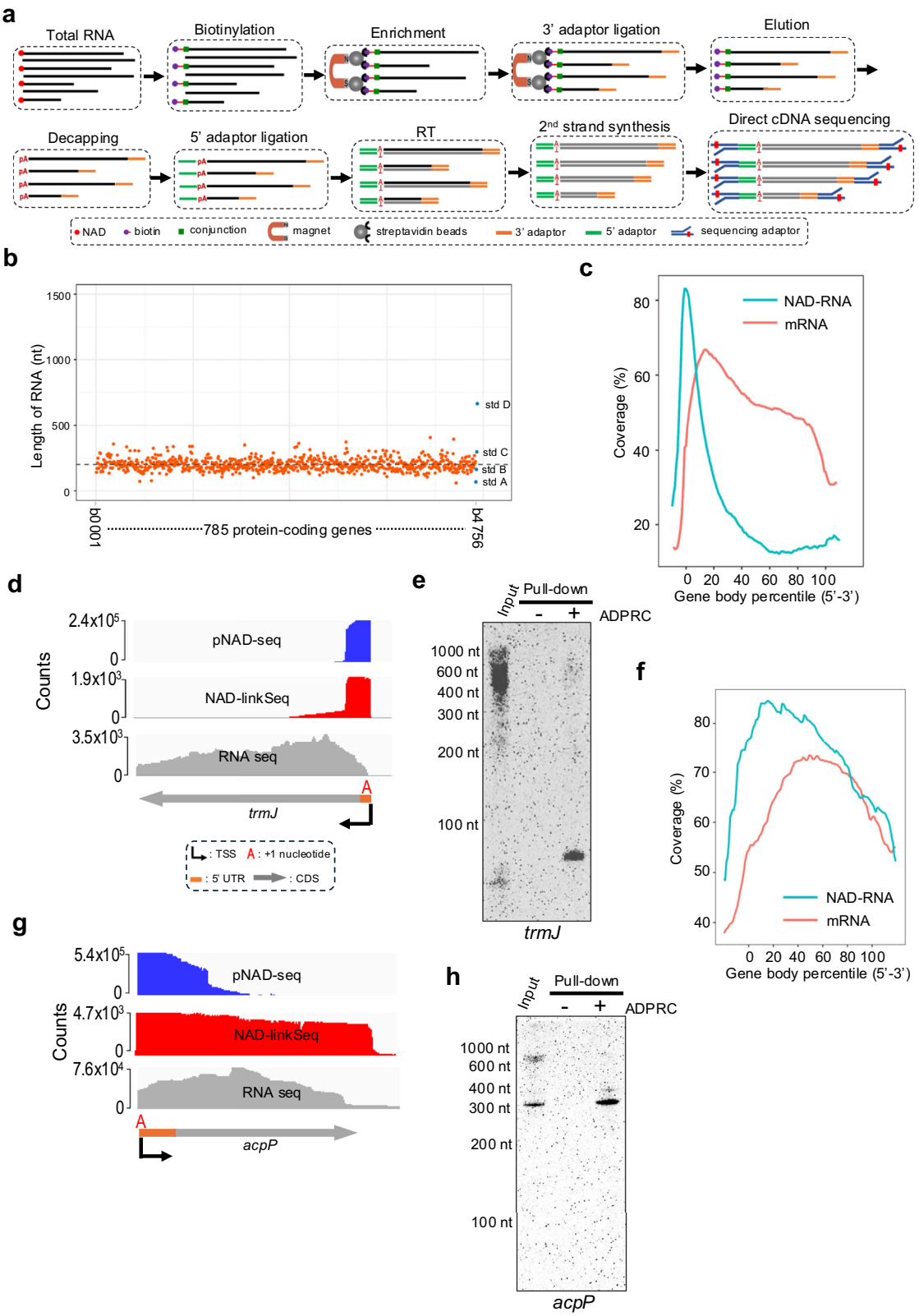
The detection of NAD-capped tRNAs and 16S rRNA fragment, some of which rank among the most abundant NAD-RNAs yet were undetected by previous methods, demonstrates the superior sensitivity of pNAD-seq. Although NAD-RNAs have been documented in other bacteria such as *Bacillus subtilis*[19] and *Staphylococcus aureus*[6], our findings indicate that a substantial portion of the NAD-modified transcriptome remains uncharacterized across bacterial species. These methods can also be used for profiling NAD-RNAs in prokaryotic cells. However, the methods rely on labeling NAD-RNAs with a reaction catalyzed by ADPRC, which has also been found to label m^7G-capped mRNAs, albeit with lower efficiency[13,20]. Therefore, when profiling eukaryotic NAD-RNAs using these methods, m^7G-capped mRNAs must be removed prior to the ADPRC reaction. By generating an accurate, comprehensive, high-resolution map of NAD-RNAs, pNAD-seq and NAD-linkSeq provide valuable tools will facilitate future functional investigations of NAD capping in both prokaryotic and eukaryotic organisms.

Methods

E. coli strain, culture conditions and RNA extraction

E. coli K-12 strain MG1655 was grown aerobically at 37 °C in Luria-Bertani (LB) broth. Overnight cultures were diluted 1:100 into fresh medium and grown until stationary phase ($OD_{600} = 2.4$). For nitrogen

https://doi.org/10.1038/s41467-026-71487-9



reduction experiments, cells were initially grown in M9 minimal medium at 37 °C to mid-log phase (OD_{600} = 0.5), harvested by centrifugation (2500 × *g*), and resuspended in either: standard M9 medium (18.7 mM NH_4Cl), or nitrogen-reduced M9 medium (5.6 mM NH_4Cl). Resuspended cultures were incubated for 3.5 h in their respective media before harvesting. Total RNA was extracted using the hot phenol method[13].

In vitro transcription

A double-stranded DNA template was generated by annealing oligonucleotides φ2.5-24mer_F and φ2.5-24mer_R (Supplementary Data 10). Transcription reactions contained 40 mM Tris-Cl (pH 8.1), 1 mM spermidine, 22 mM $MgCl_2$, 0.01% Triton X-100, 5 mM DTT, 5% DMSO, 1 mM each of GTP, CTP and UTP, plus 4 mM of either NAD (for NAD-RNA), AMP (for 5′-pA-RNA), ADP (for 5′-ppA-RNA), ATP (for 5′ pppA-RNA),

Fig. 4 | Development and application of NAD-linkSeq for systematic analysis of NAD-RNA lengths. a Schematic of NAD-linkSeq workflow. Total RNA, spiked with NAD-RNA standards of varying lengths, was subjected to SPAAC tagging for biotinylation of NAD-RNAs, followed by enrichment and 3′ adapter ligation. NudC decapping generated 5′-pA-RNAs, enabling ligation with an extended 5′-adapter. Full-length cDNAs were synthesized by reverse transcription for Nanopore direct cDNA sequencing. Structures of the reaction products are illustrated where relevant. **b** NAD-RNA transcript length distribution. Mean lengths for 785 protein-coding genes are plotted by EcoCyc locus ID on the x-axis (b0001 *thrL* to b4756 *yqiD*); the y-axis shows length in nucleotides. NAD-linkSeq-identified NAD-RNAs and four synthetic spike-in standards are included. **c** Coverage profiles comparing full-length NAD-RNA reads and mRNA reads across NAD-RNA genes classified in the 5′-region group. Coverage is plotted along the gene body from 5′ to 3′ (percentile; mRNA in red, NAD-RNA in teal). **d** Read coverage at the *trmJ* gene locus. IGV tracks compare pNAD-seq (blue) and NAD-linkSeq (red) NAD-RNA reads with RNA-seq mRNA reads (gray). The y axis shows raw read counts. The gene model shows the *trmJ* coding sequence (gray arrow) and 5′ UTR (orange). The transcription start site (+1) is marked by a bent black arrow, with the identity of the +1 base (A) indicated. **e** Northern blot validation of trmJ NAD-RNA size. NAD-RNAs were biotinylated via ADPRC/SPAAC (biotin-DBCO), enriched with streptavidin beads (with a no-ADPRC control), and resolved by urea-PAGE. Transcripts were detected using digoxigenin-labeled probes and anti-digoxigenin antibodies, comparing NAD-RNA and mRNA sizes. Shown is a representative gel from two independent experiments. **f** Coverage profiles of full-length NAD-RNA and mRNA reads across genes where the NAD-RNAs span the entire mRNA. Coverage is plotted along the gene body from 5′ to 3′ (percentile; mRNA in red, NAD-RNA in teal). **g** Read coverage at the *acpP* gene locus. IGV tracks compare pNAD-seq (blue) and NAD-linkSeq (red) NAD-RNA reads with RNA-seq mRNA reads (gray). Gene models are depicted as in (**b**). **h** Northern blot validation of acpP NAD-RNA size. NAD-RNA pulldown and Northern blot analysis were performed as in (**e**). Shown is a representative gel from two independent experiments. Source data are provided as a Source data file.

Ap3A (for 5′-ApppA-RNA), Ap4A (for 5′-Ap4A-RNA), dpCoA (for 5′-dpCoA-RNA), or NADH (for 5′-NADH-RNA), along with 0.1 μM DNA template and 5.2 U/μL T7 RNA polymerase. To generate longer NAD-RNA standards, we performed in vitro transcription reactions using 2 μg of double-stranded DNA template (Supplementary Table 1) for two NAD-RNA spike-in controls (Standard 1 and Standard 2) and a 90-nt NAD-RNA for PEGylation analysis (Standard 3). All transcription reactions were conducted under identical buffer conditions in the presence of 4 mM NAD. All reactions were incubated at 37 °C for 4 h, treated with 50 U of DNase I (Takara, 2270B) at 37 °C for 20 min, and products were resolved on 8-10% polyacrylamide/8 M urea gels (1× TBE, 40 mA). RNA bands were excised and eluted in 0.3 M NaOAc (pH 5.5), followed by ethanol precipitation. To generate 5′-OH-A-RNA, 5′-pA-RNA was treated with 0.1 U/μL calf intestinal alkaline phosphatase (Invitrogen, 18009019) in 1× CutSmart buffer for 20 min, then purified by acid phenol/chloroform extraction and ethanol precipitation, with concentrations measured using a Qubit fluorometer (Thermo Scientific).

SPAAC biotinylation

Each 25-mer RNA oligonucleotide (30 pmol) with a distinct 5′ end was incubated in a 50 μL reaction containing 50 mM Na-HEPES (pH 7.0), 5 mM $MgCl_2$, 4.25 μM ADPRC (Sigma, A9106), and 5 μL of 3-azido-1-propanol (Sigma, 776130) at 37 °C for 30 min. Enzyme-free reactions were included as negative controls. Following incubation, all RNA samples were purified by acid phenol/chloroform extraction and ethanol precipitation. The purified RNAs were then resuspended in 10 μL of PBS (pH 7.4) containing 1 mM biotin-PEG4-DBCO (BroadPharm, BP-22295) and incubated at 37 °C for 1 h to facilitate biotinylation. Reactions were halted by adding 10 μL of RNA loading dye (Thermo Scientific, R0641). The samples were resolved on a 12% TBE-urea polyacrylamide gel (containing 8 M urea), and the biotinylated RNA bands were excised and eluted in 0.3 M NaOAc (pH 5.5), followed by ethanol precipitation.

One-step biotinylation

NAD-24mer RNA (30 pmol) was biotinylated by incubating with 100 mM HEEB (Amatek, B-1328) and 4.25 μM ADPRC (Sigma) in 100 μL of 50 mM Na-HEPES (pH 7.0), 5 mM $MgCl_2$ at 37 °C for 30 min, with parallel enzyme-free reactions serving as negative controls. Reactions were terminated and purified by acid phenol/chloroform extraction, followed by ethanol precipitation.

Recombinant NudC purification

The *E. coli* nudC gene was amplified by PCR from K12 MG1655 genomic DNA using primers NudC_BamHI and NudC_SalI (Supplementary Data 10). The resulting PCR product was digested with BamHI and SalI and cloned into the pET28a vector (Novagen, 69864). The constructed plasmid, pET28a-NudC, was transformed into *E. coli* Rosetta (DE3) cells (Novagen, 70954) for protein expression. Expression of NudC was induced by adding 0.3 mM IPTG when the culture reached an OD_{600} of ~0.6, followed by incubation at 16 °C for 12 h. Cells were harvested by centrifugation and lysed via sonication in Lysis Buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 10 mM imidazole). The lysate was clarified by centrifugation at 12,500 × *g* for 20 min at 4 °C. The supernatant was applied to Ni-NTA resin (Thermo Scientific, 88222) and washed with Wash Buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 50 mM imidazole). The His-tagged NudC protein was eluted with Elution Buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 250 mM imidazole). The eluted protein was buffer-exchanged into 50 mM Tris-HCl (pH 8.0) using a 10-kDa centrifugal filter (Millipore, UFC501008) and further purified by FPLC (Bio-Rad, 7880003) on a MonoQ column (Cytiva, 29275878) with a linear 0 to 1 M NaCl gradient. Fractions containing pure NudC, as confirmed by SDS-PAGE, were pooled, concentrated, and buffer-exchanged into Storage Buffer (25 mM Tris-HCl, pH 8.0, 50 mM NaCl, 1 mM $MgCl_2$, 1 mM DTT, 50% glycerol) using a 10-kDa centrifugal filter. The final protein concentration was determined using a Qubit fluorometer (Thermo Scientific, Q33216).

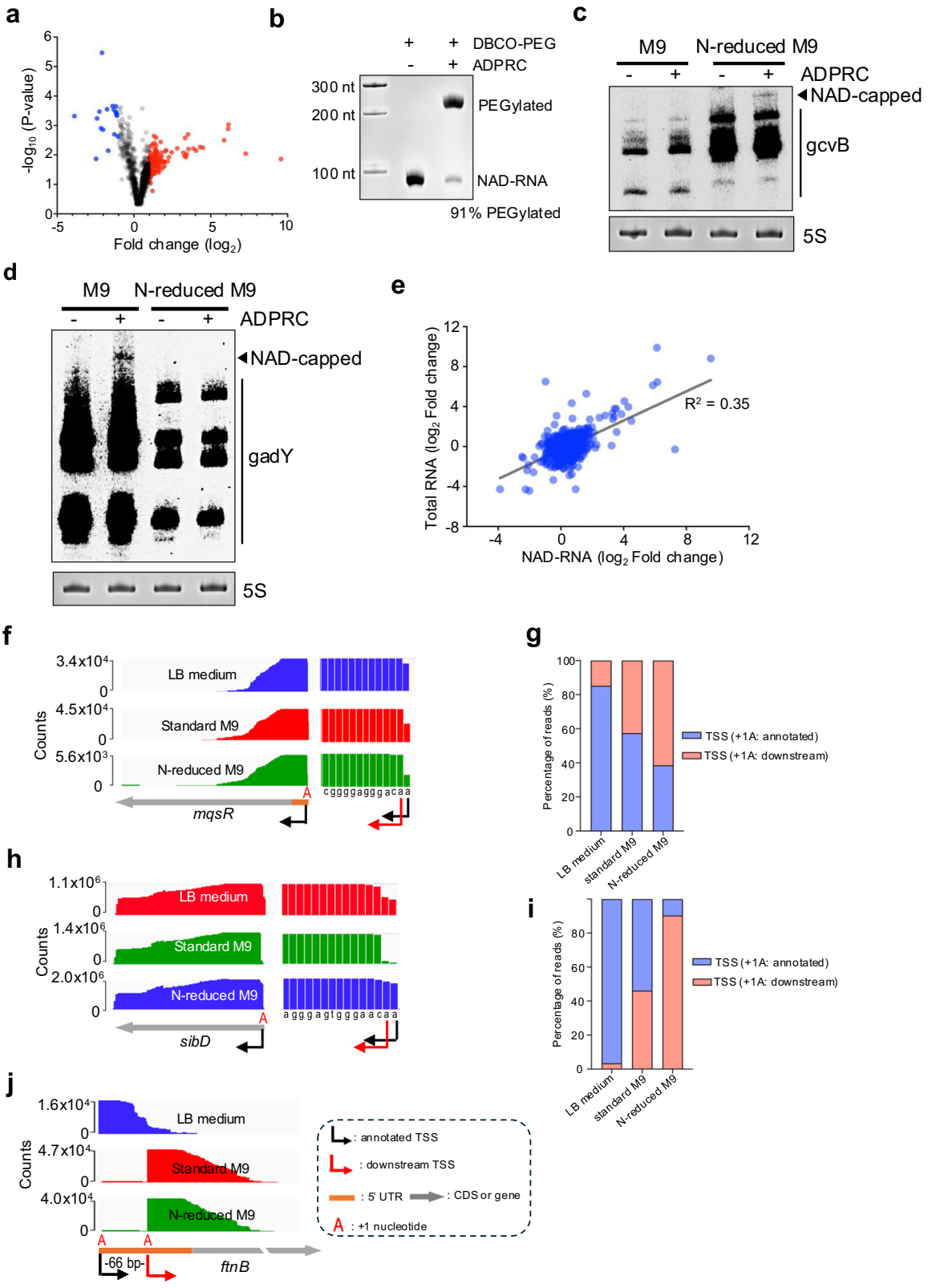
NudC decapping assay

To assess the impact of 5′-end secondary structure on decapping, biotinylated NAD-24mer RNA was annealed to complementary oligonucleotides (Supplementary Data 10) to generate substrates with extended, blunt, or recessed ends. For the decapping assay, 30 pmol of either the linear RNA substrate or the structured substrates were incubated in a 10 μL reaction buffer containing 10 mM Tris-HCl (pH 8.0), 100 mM KCl, 2 mM $MgCl_2$, 1 mM $MnCl_2$, and 2 mM DTT, along with specified concentrations of NudC (typically 0.58 μM, unless otherwise indicated for titration) at 37 °C for 10 min. Reactions were terminated by adding 10 μL of 2× RNA loading dye (Thermo Scientific; R0641), followed by denaturation, separation on a 12% denaturing urea-polyacrylamide gel, and visualization with RedSafe (INTRON, 21141) staining.

pNAD-seq library preparation

A 600 μg aliquot of RNA was spiked with either 5 ng of standard 1 RNA (for normal conditions) or 6 ng of standard 1 plus 3 ng of standard 2 RNA (for nitrogen reduction experiments). RNA fragmentation was performed in 600 μL of 10 mM Tris-HCl (pH 7.4) and 10 mM $ZnCl_2$ at 75 °C for 2.5 min, immediately quenched with 20 μL of 0.5 M EDTA (pH 8.0), and purified by acid phenol:chloroform:isoamyl alcohol (PCI, Invitrogen) extraction followed by ethanol precipitation. Dephosphorylation was carried out with 25 U of calf intestinal alkaline phosphatase (NEB, M0290) in 1× CutSmart Buffer (NEB, B6004) at 37 °C for 20 min, followed by purification as above. For NAD-RNA capture, the sample was divided into two equal halves. One half was

https://doi.org/10.1038/s41467-026-71487-9



incubated in a 150 µL reaction containing 50 mM Na-HEPES (pH 7.0), 5 mM $MgCl_2$, 8.5 µM ADPRC, and 15 µL of (Sigma, 776130) at 37 °C for 30 min. The other half underwent the same treatment but without ADPRC, serving as a negative control. Both RNA samples were then purified by phenol/chloroform extraction and ethanol precipitation. The resulting RNA pellets were resuspended in 30 µL of 1 mM biotin-PEG4-DBCO (BroadPharm, BP-22295) in 1× PBS (pH 7.4) and incubated at 37 °C for 1 h. Streptavidin capture was performed by incubating the RNA in 200 µL of binding buffer (50 mM Tris-Cl, pH 7.5, 500 mM NaCl, 5 mM EDTA) with 600 µg of MyOne T1 beads (Thermo Scientific, 65601) at 37 °C for 20 min. Beads were subsequently washed twice with 200 µL of nuclease-free water and once with 200 µL of T4 RNA ligation buffer (50 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 1 mM DTT). 3′-adapter ligation was performed in a 30 µL reaction containing T4 RNA ligation

Fig. 5 | Nitrogen limitation alters NAD-RNA expression, capping, and transcriptional initiation in *E. coli*. **a** Volcano plot displays 811 detected NAD-RNAs (upregulated in red, downregulated in blue; $p < 0.05$, Log2FC > |1|) comparing nitrogen-reduced M9 versus standard M9 conditions. Differential testing used a negative-binomial generalized linear model (Wald test, two-sided). *P*-values were adjusted for multiple comparisons using the Benjamini–Hochberg false discovery rate procedure. **b** NAD PEGylation-mediated separation of NAD-RNA from uncapped RNA. 90-nt NAD-RNA was incubated with azido-propanol in the presence or absence of ADPRC, followed by 10-kDa PEG-DBCO conjugation. The resulting mass shift enables resolution from uncapped RNA on 8% denaturing PAGE. RedSafe staining was used to quantify relative abundance. A representative gel from two independent assays is shown. **c**, **d** Northern blot analysis comparing NAD-capped (PEGylated) and regular transcripts for *gcvB* (upregulated) and *gadY* (downregulated) between growth conditions, using 5S rRNA as loading control. Shown is a representative blot from two independent experiments. **e** Scatter plot showing the $\log_2$ fold change in NAD-RNA abundance (x-axis) versus the corresponding $\log_2$ fold change in total RNA levels (y-axis) for genes detected as NAD-capped. **f**, **h** IGV tracks (left) showing 5′ end read coverage at representative loci for cultures grown in LB, standard M9, and N-limited M9. Insets (right of each track) zoom in on the promoter region; the annotated TSS and an alternative downstream TSS are indicated. Read counts are shown on the y axis. **g**, **i** Quantification of TSS usage corresponding to the IGV tracks: stacked bars show the fraction of reads initiating at the annotated (+1 A: annotated) versus the downstream (+1 A: downstream) TSS. **j** IGV snapshot of the *ftnB* locus across three conditions (LB, standard M9, and N-reduced M9) showing a condition-dependent switch in promoter usage from a proximal to a distal TSS (-66 bp apart). Read counts are on the y axis. The gene body (gray) and 5′ UTR (orange) are shown; bent arrows denote TSSs (black, annotated; red, downstream), with +1 A positions and strand orientation indicated. Source data are provided as a Source data file.

buffer, 5 µM pre-adenylated adapter (Supplementary Data 10), 20% PEG 8000, 30 U of murine RNase inhibitor (NEB, M0314S), and 30 U of T4 Rnl2tr K227Q (NEB, M0351L) at 25 °C for 4 h. Ligation products were then washed six times with 200 µL of 50 mM Tris-Cl (pH 7.5), 8 M urea, 0.5% SDS, and 10 mM EDTA. Beads-bound RNA was eluted by incubation in 80 µL of 50 mM Tris-Cl (pH 7.5), 95% formamide, 10 mM EDTA, and 2 mM biotin at 65 °C for 2.5 min. The supernatant was collected and combined with 320 µL of 0.3 M sodium acetate (pH 5.5) and 50 µg of linear acrylamide (Thermo Scientific, AM9520), and the RNA was ethanol-precipitated. To remove the biotin-labeled NAD cap, the RNA was incubated in a 40 µL reaction containing 10 mM Tris-HCl (pH 8.0), 100 mM KCl, 2 mM $MgCl_2$, 1 mM $MnCl_2$, 2 mM DTT, and 0.58 µM NudC at 37 °C for 20 min. The decapped RNA was purified using a RNA Clean & Concentrator kit (Zymo, R1016) and eluted in 11 µL of nuclease-free water. 5′-adapter ligation was performed in a 30 µL reaction containing 1× T4 RNA ligase reaction buffer (NEB, B0216L), the purified RNA, 3.3 µM 5′-adapter (Supplementary Data 10), 15% PEG 8000, 1 mM ATP, 30 U of murine RNase inhibitor, and 30 U of T4 RNA ligase 1 (NEB, M0204S) at 25 °C for 3 h. The ligation product was combined with 10 pmol of RT primer (Supplementary Data 10), heated to 70 °C for 5 min, and immediately chilled on ice. Reverse transcription was performed in a 40 µL reaction containing the primed RNA, 1× Maxima RT Buffer, 1 mM dNTPs, 40 U of murine RNase inhibitor (NEB; M0314L), and 200 U of Maxima Reverse Transcriptase (Thermo Scientific, EP0742) at 50 °C for 30 min. The resulting cDNA was amplified by PCR for 12–13 cycles using Q5 High-Fidelity DNA Polymerase (NEB; M0491L) and primers listed in Supplementary Data 10, followed by a 5-cycle indexing PCR. The final library was purified with AMPure XP beads (Beckman, A63880) at a 1.2:1 bead-to-sample ratio, and its quality was assessed on an Agilent 2100 Bioanalyzer. Sequencing was performed on an Illumina HiSeq 2000 platform (Novogene) for 150 bp paired-end reads.

qPCR analysis

Relative transcript levels of selected tRNAs and rRNAs were measured by qPCR in pNAD-seq libraries prepared with or without ADPRC treatment. Reactions were performed using SupRealQ Ultra Hunter SYBR qPCR Master Mix (Vazyme, Q713-02) on a StepOne Plus RT-PCR system (Applied Biosystems). Transcript levels were normalized to a synthetic NAD-RNA spike-in control for each RNA species. Primer sequences are listed in Supplementary Data 10.

RNA-Seq and data analysis

For RNA sequencing, rRNA depletion was performed using the Ribo-Zero Plus rRNA Depletion Kit for bacteria (Illumina, ILL20037135EA). Library preparation was then carried out using the Next Ultra Directional RNA Library Prep Kit for Illumina (NEB, E7420). The prepared samples were sequenced on an Illumina HiSeq 2000 platform, generating 150 bp paired-end reads (Novogene). The raw reads were trimmed using fastp v0.22.0[29]. The cleaned reads were then aligned to the *E. coli* str. K-12 substr. MG1655 genome (ASM584v2) using HISAT2[30]. Expression levels of genes were calculated using featureCounts v2.0.1[31]. Finally, differential gene expression analysis was conducted on the count data using DESeq2[32]. A Benjamini–Hochberg-adjusted *p*-value of less than 0.05 was considered statistically significant. Data were visualized in GraphPad Prism 10.

NAD-linkSeq library preparation

A 400 µg aliquot of total RNA from stationary phase *E. coli* was spiked with four NAD-RNA standards: 150 pg of 50 nt, 420 pg of 150 nt, 800 pg of 277 nt, and 200 pg of 700 nt transcripts (Supplementary Table 2). The mixture was incubated for 30 min at 37 °C with 8.5 µM ADPRC and 15 µL 3 azido 1 propanol in a 150 µL reaction containing 50 mM Na HEPES (pH 7.0) and 5 mM $MgCl_2$. RNA was purified by acid phenol–chloroform extraction followed by ethanol precipitation. For biotinylation, the RNA was resuspended in 40 µL PBS (pH 7.4) with 1 mM biotin-PEG4-DBCO (BroadPharm, BP-22295) and incubated at 37 °C for 1 h, then ethanol-precipitated to remove excess reagent. Subsequent steps, from streptavidin bead capture through reverse transcription, followed the pNAD-seq protocol, with the modification of using a longer 5′ adapter and the linkSeq RT primer (Supplementary Data 10). After reverse transcription, RNA templates were digested with 1 µL RNase A/T1 mix (Thermo Scientific, EN0551) at 37 °C for 10 min, and cDNA was purified using AMPure XP beads. The second cDNA strand was synthesized with Q5 DNA polymerase after annealing 1 pmol linkSeq FP (Supplementary Data 10). Finally, the resulting double-stranded cDNA was purified using AMPure XP beads (Beckman, A63880).

Nanopore sequencing and data analysis

DNA library was prepared using Ligation sequencing DNA V14 kit (ONT, SQK-LSK114) and loaded on a flowcell (ONT, R10.1 version) and sequencing was run on a GridIon sequencer (ONT). The adapter sequences, 5′-CAGAGTTCTACAGTCCGACGATC-3′ present at 5′-end and 5′-GATCGGAAGAGCACACGTCTG-3′ present at 3′-end of each read, were trimmed using Cutadapt[33]. Sequencing reads were aligned to the reference using the minimap2[34]. Length distribution was analyzed using Samtools[35]. Candidate sequences were visualized in the Integrated Genome Viewer[36].

RNA pulldown and LC-MS/MS

RNA pulldown was performed by hybridizing 100 pmol biotin-DNA probe (Supplementary Data 10) with 1 mg total RNA in 400 µL solution (50 mM Tris pH 7.5, 500 mM LiCl, and 5 mM EDTA) at 80 °C for 5 min and slowly cooled to 25 °C, followed by capture with 500 µg MyOne Streptavidin T1 beads (Thermo Scientific, 65601) at 37 °C for 20 min. After washing with 50 mM Tris (pH 7.5), 500 mM LiCl, 5 mM EDTA, and 0.05% Tween, elution was performed by incubating the beads-bound RNA in 80 µL elution buffer (10 mM Tris-Cl pH 7.5, 95% formamide, 5 mM EDTA, and 2 mM biotin) at 65 °C for 2.5 min, then ethanol-

precipitated after mixing the supernatant with 320 µL of 0.3 M sodium acetate (pH 5.5) containing 30 µg linear acrylamide. The RNA was digested with 50 U of nuclease P1 (NEB, M0660S) in 1× nuclease P1 buffer at 37 °C for 1 h and analyzed by LC-MS/MS using a TSQ Altis mass spectrometer (negative ion mode) with an ACQUITY UPLC HSS T3 column (2.1 × 100 mm, 1.8 µm) at 0.1 mL/min flow rate. Separation was performed at a flow rate of 0.1 mL/min with mobile phase A (0.3% formic acid) and B (80% acetonitrile) using the following gradient: 18–21% B (0–2 min), 21–40% B (2–4 min), 40–60% B (4–14 min), and 100% B (14–18 min), while monitoring *m/z* 428.001. The column was re-equilibrated by returning to 18% B for 2 min between injections.

NAD PEGylation

Standard 90-nt NAD-RNA (200 ng) or *E. coli* total RNA (200 µg) was PEGylated by incubation with 8.5 µM ADPRC and 15 µL of 3-azido-1-propanol (Sigma, 776130) in 150 µL reaction containing 50 mM Na-HEPES, pH 7.0, 5 mM $MgCl_2$ at 37 °C for 30 min, followed by phenol/chloroform extraction and ethanol precipitation. The RNA was then reacted with 1 mM DBCO-mPEG (10 kDa) (BroadPharm, BP-22462) in 40 µL 1xPBS (pH 7.4) at 37 °C for 1 h, ethanol-precipitated, and analyzed by urea-PAGE or Northern blot.

Digoxigenin-labeled RNA probes

DNA template was generated by PCR amplification with primers (Supplementary Data 10) using *E. coli* K-12 genomic DNA as template. Transcription mixture containing 2 µg DNA template in 40 µL transcription solution containing 40 mM Tris-Cl, pH 8.1, 1 mM spermidine, 22 mM $MgCl_2$, 0.01% Triton-X-100, 5 mM DTT, 5% DMSO, and 2.5 mM of each of NTP and 0.1 mM digoxigenin-11-UTP (Roche, 11209256910), along with 208 U T7 RNA polymerase was incubated at 37 °C for 3 h. To the reaction 10 U DNase I (Takara, 2270B) was added to digest DNA template by incubation at 37 °C for 30 min. Transcription products were resolved on 8% polyacrylamide/8 M urea gel. RNA probes were extracted from the gel for immediate use in Northern blot.

Northern blot

20 µg aliquot of total RNA from the PEGylation step was denatured by incubation at 95 °C for 2 min in 1× RNA gel loading dye. The RNA was then separated on a 6.5% polyacrylamide/8 M urea gel in 1× TBE buffer at a constant current of 40 mA. Following electrophoresis, the RNA was transferred onto a Hybond-N+ membrane (Cytiva, RPN303B) by electroblotting in 0.5× TBE buffer at 250 mA for 30 min. The membrane was cross-linked using UV light at 120 mJ/cm². For hybridization, the membrane was first prehybridized in DIG Easy Hyb solution (Roche, 11603558001) for 1 h. A digoxigenin-labeled RNA probe was then added, and hybridization was carried out overnight at 62 °C. The membrane was washed twice for 15 min each at 62 °C, first with 2× SSC containing 0.1% SDS, followed by 0.5× SSC containing 0.1% SDS. The membrane was subsequently washed for 5 min in maleic acid wash buffer (0.1 M maleic acid, 0.15 M NaCl, 0.3% Tween-20, pH 7.5) and blocked with blocking solution (Roche, 11585762001) for 30 min. It was then incubated with anti-digoxigenin-AP antibody (Roche, 11093274910) diluted to 75 mU/mL in blocking solution for 45 min. After incubation, the membrane was washed three times for 15 min each in maleic acid wash buffer and equilibrated in detection buffer (0.1 M Tris, 0.1 M NaCl, pH 9.5). Signals were detected using the chemiluminescent substrate CSPD (Thermo Scientific, T2142) and visualized on a ChemiDoc XRS+ system (Bio-Rad). Uncropped and unprocessed scans of the most important blots are provided in the Source Data file.

Statistics and reproducibility

Statistical analyses and data visualization were performed using GraphPad Prism 10. Significance thresholds were set at as a *p*-value of <0.05 as [*] or <0.01 as [**] as indicated in the respective figure legends. All experiments were performed at least two times.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The pNAD-seq, RNA-seq, and NAD-linkSeq data generated in this study have been deposited in NCBI's Sequence Read Archive (SRA) under accession code PRJNA1125588. Source data are provided with this paper.

References

1. Wei, C. M., Gershowitz, A. & Moss, B. Methylated nucleotides block 5′ terminus of HeLa cell messenger RNA. *Cell* **4**, 379–386 (1975).
2. Topisirovic, I., Svitkin, Y. V., Sonenberg, N. & Shatkin, A. J. Cap and cap-binding proteins in the control of gene expression. *Wiley Interdiscip. Rev.: RNA* **2**, 277–298 (2011).
3. Grudzien-Nogalska, E. & Kiledjian, M. New insights into decapping enzymes and selective mRNA decay. *WIREs RNA* **8**, e1379 (2017).
4. McClure, W. R. Mechanism and control of transcription initiation in prokaryotes. *Annu. Rev. Biochem.* **54**, 171–204 (1985).
5. Cahová, H., Winz, M.-L., Höfer, K., Nübel, G. & Jäschke, A. NAD captureSeq indicates NAD as a bacterial cap for a subset of regulatory RNAs. *Nature* **519**, 374–377 (2015).
6. Morales-Filloy, H. G. et al. The 5′ NAD Cap of RNAIII modulates toxin production in Staphylococcus aureus isolates. *J. Bacteriol.* https://doi.org/10.1128/jb.00591-00519 (2020).
7. Walters, R. W. et al. Identification of NAD^+ capped mRNAs in Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci. USA* **114**, 480–485 (2017).
8. Jiao, X. et al. 5′ end nicotinamide adenine dinucleotide cap in human cells promotes RNA decay through DXO-mediated deNADding. *Cell* **168**, 1015–1027.e1010 (2017).
9. Grudzien-Nogalska, E. et al. Structural and mechanistic basis of mammalian Nudt12 RNA deNADding. *Nat. Chem. Biol.* **15**, 575–582 (2019).
10. Wang, Y. et al. NAD^+-capped RNAs are widespread in the Arabidopsis transcriptome and can probably be translated. *Proc. Natl. Acad. Sci. USA* **116**, 12094–12102 (2019).
11. Zhang, H. et al. NAD tagSeq reveals that NAD^+-capped RNAs are mostly produced from a large number of protein-coding genes in Arabidopsis. *Proc. Natl. Acad. Sci. USA* **116**, 12072–12077 (2019).
12. Dong, H. et al. NAD^+-capped RNAs are widespread in rice (Oryza sativa) and spatiotemporally modulated during development. *Sci. China Life Sci.* **65**, 2121–2124 (2022).
13. Zhang, H. et al. Use of NAD tagSeq II to identify growth phase-dependent alterations in *E. coli* RNA NAD^+ capping. *Proc. Natl. Acad. Sci. USA* **118**, e2026183118 (2021).
14. Sharma, S., Yang, J., Favate, J., Shah, P. & Kiledjian, M. NADcapPro and circNC: methods for accurate profiling of NAD and non-canonical RNA caps in eukaryotes. *Commun. Biol.* **6**, 406 (2023).
15. Niu, K. et al. ONE-seq: epitranscriptome and gene-specific profiling of NAD-capped RNA. *Nucleic Acids Res.* **51**, e12–e12 (2023).
16. Hansen, K. D., Brenner, S. E. & Dudoit, S. Biases in Illumina transcriptome sequencing caused by random hexamer priming. *Nucleic Acids Res.* **38**, e131 (2010).
17. Ramsköld, D. et al. Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat. Biotechnol.* **30**, 777–782 (2012).
18. Vvedenskaya, I. O. et al. CapZyme-seq comprehensively defines promoter-sequence determinants for RNA 5′ capping with NAD^+. *Mol. Cell* **70**, 553–564.e559 (2018).

19. Frindert, J. et al. Identification, biosynthesis, and decapping of NAD-capped RNAs in B. subtilis. *Cell Rep.* **24**, 1890–1901.e1898 (2018).
20. Hu, H. et al. SPAAC-NAD-seq, a sensitive and accurate method to profile NAD^+-capped transcripts. *Proc. Natl. Acad. Sci. USA* **118**, e2025595118 (2021).
21. Luciano, D. J., Vasilyev, N., Richards, J., Serganov, A. & Belasco, J. G. A novel RNA phosphorylation state enables 5′ end-dependent degradation in Escherichia coli. *Mol. Cell* **67**, 44–54.e46 (2017).
22. Deana, A., Celesnik, H. & Belasco, J. G. The bacterial enzyme RppH triggers messenger RNA degradation by 5′ pyrophosphate removal. *Nature* **451**, 355–358 (2008).
23. Hudeček, O. et al. Dinucleoside polyphosphates act as 5′-RNA caps in bacteria. *Nat. Commun.* **11**, 1052 (2020).
24. Luciano, D. J., Levenson-Palmer, R. & Belasco, J. G. Stresses that raise Np_4A levels induce protective nucleoside tetraphosphate capping of bacterial RNA. *Mol. Cell* **75**, 957–966. e958 (2019).
25. Peach, S. E., York, K. & Hesselberth, J. R. Global analysis of RNA cleavage by 5′-hydroxyl RNA sequencing. *Nucleic Acids Res.* **43**, e108–e108 (2015).
26. Kowtoniuk, W. E., Shen, Y., Heemstra, J. M., Agarwal, I. & Liu, D. R. A chemical screen for biological small molecule–RNA conjugates reveals CoA-linked RNA. *Proc. Natl. Acad. Sci. USA* **106**, 7768–7773 (2009).
27. Höfer, K. et al. Structure and function of the bacterial decapping enzyme NudC. *Nat. Chem. Biol.* **12**, 730–734 (2016).
28. Wiedermannová, J., Babu, R. & Yuzenkova, Y. Stochastic nature and physiological implications of 5′-NAD RNA cap in bacteria. *Nucleic Acids Res.* **52**, 11838–11852 (2024).
29. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).
30. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015).
31. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
32. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 1–21 (2014).
33. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10–12 (2011).
34. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
35. Li, H. et al. The sequence alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
36. Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinforma.* **14**, 178–192 (2012).
37. Keseler, I. M. et al. EcoCyc: a comprehensive view of Escherichia coli biology. *Nucleic Acids Res.* **37**, D464–D470 (2009).
38. Crooks, G. E., Hon, G., Chandonia, J.-M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).

Acknowledgements

This work was supported by Zhejiang Wanli University's startup funds (to H.Z. and Y.X.) and by the Research Grants Council of Hong Kong: GRF 12102220, 12101721, and 12102022 (to Y.X.); GRF 12103425 (to H.Z. and Y.X.); AoE AoE/M-402/25-N (to Y.X.); CRF CRS_HKBU201/22 (to Y.X.); and grants C2009-19GF and C2003-22WF (to Z.C. and Y.X.). We thank Ms. Yichen Zhou for assistance with sequencing data analysis.

Author contributions

H.Z. and Y.X. conceived and designed the study. H.Z., Y.Q., F.Z., P.W., W.L., X.S., B.Z., C.Z., Y.W., and L.H. performed the laboratory experiments. Q.L. conducted RNA-seq analyses, and Y.Q. and P.W. performed NAD-linkSeq analyses. H.Z. interpreted the sequencing data. X.Y. and Z.C. supervised the study. Y.X., H.Z., and Z.C. wrote the manuscript. All authors reviewed, edited, and approved the final version.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-026-71487-9.

Correspondence and requests for materials should be addressed to Hailei Zhang, Zongwei Cai or Yiji Xia.

Peer review information *Nature Communications* thanks Gisela Storz, Marie-Luise Winz and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Hailei Zhang[1,2] ✉, Yinyao Qi[2], Qiongfang Li[2], Feng Zhang[3], Pengxi Wang[2], Wuzhen Liu[2], Xiaojian Shao[3], Bingxu Zhang[2], Chen Zhang[2], Yiping Wang[4], Linfeng Huang[5], Zongwei Cai[3,6] ✉ & Yiji Xia[1,2,7] ✉

[1]College of Biological and Environmental Sciences, Zhejiang Wanli University, Ningbo, China. [2]Department of Biology, Hong Kong Baptist University, Hong Kong SAR, China. [3]Department of Chemistry, Hong Kong Baptist University, Hong Kong SAR, China. [4]College of Plant Protection, Yangzhou University,

Yangzhou, China. [5]Wang-Cai Biochemistry Lab, Division of Natural and Applied Sciences, Duke Kunshan University, Kunshan, China. [6]Eastern Institute of Technology, Ningbo, China. [7]State Key Laboratory of Agrobiotechnology, The Chinese University of Hong Kong, Hong Kong SAR, China.
✉e-mail: hl.zhang@zwu.edu.cn; zwcai@eitech.edu.cn; yxia@hkbu.edu.hk