

RESEARCH

Open Access



Circularization of 23S rRNA but not 16S rRNA within archaeal ribosomes

Ling-Dong Shi^{1*}, Petar I. Penev^{1,8}, Amos J. Nissley², Dipti D. Nayak^{3,4}, Rohan Sachdeva¹, Jamie H. D. Cate^{1,2,3,5} and Jillian F. Banfield^{1,6,7*}

*Correspondence:

Ling-Dong Shi
lingdong_shi@berkeley.edu
Jillian F. Banfield
jbanfield@berkeley.edu

¹Innovative Genomics Institute,
University of California, Berkeley,
CA, USA

²Department of Chemistry,
University of California, Berkeley,
CA, USA

³Department of Molecular and Cell
Biology, University of California,
Berkeley, CA, USA

⁴Department of Plant and Microbial
Biology, University of California,
Berkeley, CA, USA

⁵Molecular Biophysics and
Integrated Bioimaging Division,
Lawrence Berkeley National
Laboratory, Berkeley, CA, USA

⁶Earth and Planetary Science,
University of California, Berkeley,
CA, USA

⁷Environmental Science, Policy
and Management, University of
California, Berkeley, CA, USA

⁸Present address: Eligo Bioscience,
111 Avenue de France, Paris, France

Abstract

Background Processing of archaeal 16S and 23S rRNAs is believed to involve excision of individual rRNAs from polycistronic precursors, circularization of excised rRNAs, and re-linearization before the incorporation into ribosomes. However, all the knowledge is derived from several isolated species, leaving open the possibility that different processes may occur in other archaeal groups.

Results Here, we investigate rRNAs from diverse and mostly uncultivated archaea. Sequencing of total cellular RNA from eight phylum-level lineages indicates that archaeal circular 23S rRNA transcript abundances vastly exceed those of linear counterparts, and linear versions are often undetectable. As the majority of rRNAs derive from mature ribosomes, the data suggest that ribosomes contain circular 23S rRNAs. Thus, we directly sequence RNA extracted from isolated ribosomes of a model archaeon, *Methanosarcina acetivorans*, and confirm that the 23S rRNAs in the ribosomes are circular. Structural modeling places the 5' and 3' ends of the linear precursors of archaeal 23S rRNAs in close proximity to form a GNRA tetraloop (in which N is A, C, G, or U and R is A or G), consistent with their existence as circular molecules. We also confirm the existence of circular 16S rRNA intermediates in transcriptomes of most archaea, yet a circular form is not evident in some distinct archaeal groups, suggesting that certain archaea do not circularize 16S rRNA during processing.

Conclusions Our findings uncover unexpected variations in the processing required to generate mature rRNAs and the conformation of functional molecules in archaeal ribosomes.

Keywords Archaea, rRNA processing, Circular 23S rRNA, Metatranscriptomics, *Methanosarcina*

Background

Translation is less well studied in Archaea than in the other Domains of life, yet this process is of great interest because archaeal systems are more similar to those of Eukaryotes than those of Bacteria [1–3]. Once transcribed, ribosomal RNAs (rRNAs) must be processed to maturation before assembly into ribosomes [4–6]. Several studies of pure cultures demonstrate that processing of archaeal 16S and 23S rRNAs, but not 5S rRNAs,



© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

often involves excision of individual rRNAs from polycistronic rRNA precursors and circularization of the excised rRNAs [7–12]. It is widely believed that re-linearization of circular 16S and 23S rRNA intermediates occurs prior to their incorporation into the ribosome [7–12]. However, all the knowledge is derived from several isolated species, leaving open the possibility that different processes occur in other archaeal groups.

Since the advent of genome-resolved metagenomics [13], culture-independent *in silico* analyses have greatly increased genomic sampling of Archaea and brought to light numerous lineages that were not known based on laboratory cultures [14–17]. More recently, environmental microbiomes have been studied by directly sequencing extracted RNA, an approach referred to as metatranscriptomics [18]. The messenger RNA (mRNA) sequences are used to document *in situ* microbial activity, typically after depletion of the rRNAs that comprise the vast majority of transcriptomes and hinder the resolution of the mRNA pool [19]. However, the rRNAs can be used to investigate the steps by which rRNAs are processed in diverse, usually uncultivated organisms and to predict the form of mature molecules within ribosomes.

Here we performed short- and long-read deep DNA and RNA sequencing to access the genomes and transcriptomes of diverse and mostly uncultivated archaea in complex wetland soil samples. We mapped the rRNA transcripts to genomes and used read alignment discrepancies to document possible circularization of 16S and 23S rRNAs. Results of total cellular RNA sequencing indicated that 23S rRNAs from eight phylum-level archaeal lineages are mostly in circular forms, suggesting that ribosome-associated 23S rRNAs are circularized. We augmented these analyses with direct sequencing of rRNAs from ribosomes isolated from a model archaeon, *Methanosarcina acetivorans*, and concluded that the investigated archaea have circular 23S rRNAs within their ribosomes. We also uncovered evidence suggesting that the steps in processing 16S rRNAs vary across archaeal lineages. These findings expand our understanding of rRNA processing and the structure of ribosomes in the Archaeal Domain of life.

Results

DNA and RNA were extracted from samples of wetland soil collected from a site in Lake County, California, USA. Genome-resolved metagenomic data have revealed abundant, diverse, and mostly uncultivated archaea in this soil ecosystem, including methane-oxidizing archaea, methanogenic archaea, Asgard archaea, and their associated extra-chromosomal elements [20–23]. Here, we performed metatranscriptomic sequencing to determine the structures of archaeal rRNAs and provided insights into archaeal rRNA processing.

A circularized *Methanoperedens* genome was reconstructed using long-read PacBio data and completion was confirmed using short Illumina reads. Based on the genome sequence, we predicted the locations of the rRNA genes. DNA reads spanned the transition from intergenic regions into the rRNA genes (Additional file 1: Fig. S1a). We then mapped Nanopore and Illumina transcript reads to the *Methanoperedens* genome and observed that reads placed at the starts and ends of 16S and 23S rRNA genes only partially agreed with the reference sequence. The discrepant portions of the transcript reads from the region preceding the genes exactly matched the ends of the genes and vice versa, indicating that these reads derived from circularized rRNAs (Additional file 1: Figs. S1b–1c). As this method does not require pure cultures and can document the

various rRNA forms present (Fig. 1), we used it to investigate whether rRNA processing that involves a circularization step occurs in other archaea.

We analyzed all 16S rRNA genes from the wetland soil dataset and assigned taxonomic classifications to each at the highest resolution possible. Long-read PacBio HiFi sequences from several samples were used to capture entire rRNA gene sequences (Additional file 1: Fig. S2). In total, 11,710 16S rRNA sequences ≥ 500 bp were assembled. The 4,580 full-length and potentially complete 16S rRNA genes were clustered into 2,576 species-level groups at 99% identity (Additional file 1: Fig. S3) [24]. Of the 1,125 16S rRNA genes with sequence coverage in start and end regions needed to evaluate circularization, 963 were from Bacteria. As expected based on prior studies of bacterial rRNA processing [4], no evidence supported the presence of circularized bacterial 16S rRNA transcripts. The remaining 162 genes were from Archaea. Of these, 78 showed evidence that a fraction of the 16S rRNA transcripts were in circularized forms. Archaea with circular 16S rRNAs are phylogenetically distributed in eight phyla (Fig. 2), including lineages that currently do not have representative isolates (e.g., Bathyarchaeia), thus expanding the diversity of microorganisms that are capable of circularizing 16S rRNAs. Circularization was also observed for 23S rRNAs in all investigated archaea, but consistent with previous reports, not in bacteria [4].

Archaeal 16S rRNA transcripts that circularize exhibit an extension of 105.6 ± 22.2 bp ($n = 78$) relative to the predicted 16S rRNA genes (Additional file 1: Fig. S4). It is intriguing that archaeal 16S rRNAs from 18 organisms showed no evidence for circularization, given that the current view is that all archaeal 16S rRNA genes go through a circular stage during processing [6]. Taking a Bathyarchaeia (formerly Bathyarchaeota) genome as an example, although abundant mapped transcripts extended gene ends for up to 246 bp, no transcripts supported the existence of circular 16S rRNA (Additional file 1: Fig. S5). Absence of circular 16S rRNAs was also observed for Thermoplasmata and Nitrososphaeria (formerly Thaumarchaeota), yet circularized 23S rRNAs in these archaea were detected (Fig. 2). Notably, all archaea in the three phyla without circular

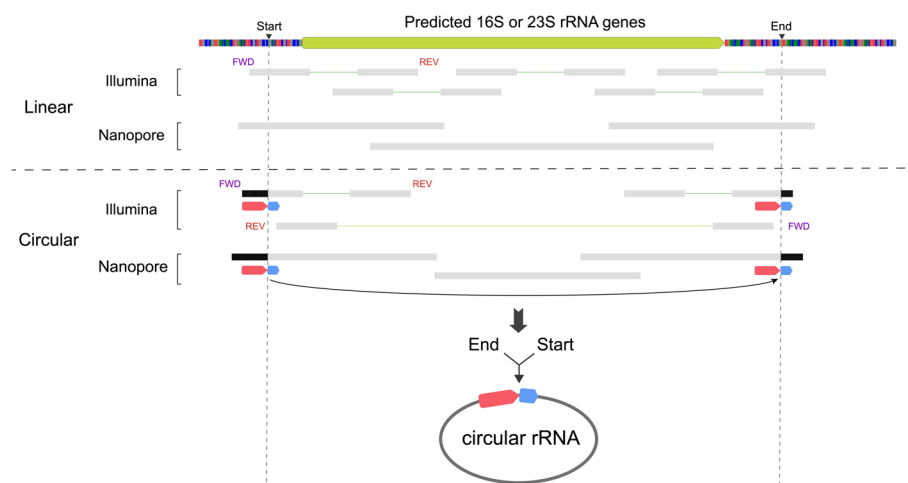


Fig. 1 Diagram of transcript mapping that predicts rRNA forms. Black triangles on the top genome reference indicate inferred start and end sites of circular rRNAs. Gray bars indicate reads that agree with the reference sequence. Black regions of reads are segments that do not match the reference sequence; red and blue arrows indicate where these sequences match the reference at the other end of the gene. Reads are linked to their paired reads by lines. When the forward (FWD) read is to the left of the reverse (REV) read, as expected, the reads are linked by a green line. When the REV read is to the left of the FWD read, not as expected, the reads are linked by a yellow line

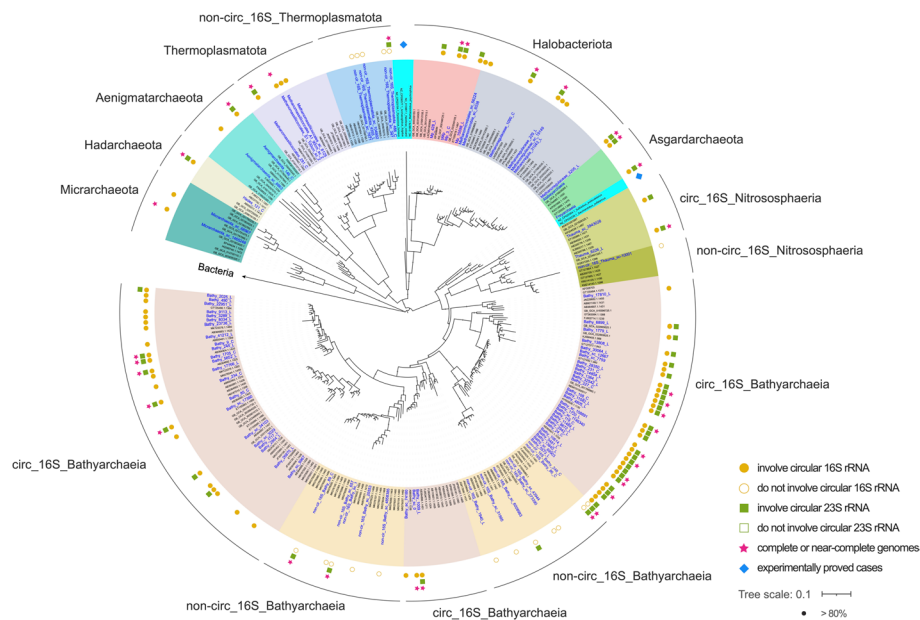


Fig. 2 Phylogeny of archaeal 16S rRNA genes with confident determination in rRNA forms. Sequences highlighted in blue are identified in this study. The tree is rerooted using Bacteria (*Escherichia coli* and *Bacillus subtilis*) as the outgroup. Filled/open circles and rectangles beside genomes indicate observed circular/non-circular 16S rRNAs and 23S rRNAs, respectively. Stars represent potentially complete or near-complete genomes, the latter of which include most single-copy genes and < 20 scaffolds. Taxonomic classifications are present in Additional file 2: Table S1. Diamonds show cases that are experimentally proven to circularize 16S and 23S rRNAs. Support values were calculated based on 1000 replicates and labeled as greater than 80

16S rRNAs are in phylogenetic clades that are distinct from related organisms with circular 16S rRNAs (Fig. 2).

In model archaea, the first step of processing for both 16S and 23S rRNAs involves an RNA splicing endonuclease (EndA) that recognizes bulge-helix-bulge (BHB) motifs and cleaves transcribed rRNA into precursors [6, 25]. Excised precursor rRNAs are then ligated by RNA splicing ligase (RtcB), generating circular intermediates [11, 26]. We thus searched for EndA and RtcB in our circularized and near-complete archaeal genomes (Fig. 2), and identified them in all cases, regardless of whether or not the data supported a circularization step during 16S rRNA processing (Additional file 1: Figs. S6 and S7). EndA proteins are classified into four different forms, but all are reported to recognize and cleave BHB motifs [27–29]. Each type includes sequences from archaeal species that have been experimentally demonstrated to circularize rRNAs (Additional file 1: Fig. S6). Therefore, we speculate that the four types of EndA proteins in our archaeal genomes likely have the function of splicing, enabling subsequent rRNA circularization. This is not unexpected, given the observation of circular 23S rRNAs in all the archaea studied here.

We next sought EndA recognition sites, the BHB motifs, in 16S and 23S rRNAs predicted secondary structures. Consistent with the inferred RNA forms from transcript mapping, sequences with circularized 16S and/or 23S rRNAs have canonical BHB motifs in the region where circularization occurs (Additional file 1: Figs. S1 and S8). In contrast, sequences with non-circular 16S rRNAs do not have BHB motifs (Additional file 1: Figs. S5 and S9). Without a BHB motif, EndA cannot recognize and excise 16S rRNA from precursors, and as a consequence, no circular 16S rRNA transcripts would be generated.

We found no correspondence between the genomic organization of the rRNA genes and whether or not their 16S rRNA genes go through a circularization step. Of 78 archaea that can generate circular 16S rRNA transcripts, 41 encode rRNA genes in operons and 35 encode separate rRNA genes. Of 18 archaea that are unable to circularize 16S rRNA transcripts, 3 have rRNA genes arranged in operons and 12 have them separated. In the other cases both for archaea that do and do not circularize their 16S rRNA transcripts, genome fragmentation precluded determination of the coding pattern.

Notably, some archaeal phyla (e.g., Bathyarchaeia) include organisms that do and do not circularize their 16S rRNA transcripts during processing (Fig. 2). We wondered if this difference may relate to growth rates, as rRNA transcription is closely associated with ribosome synthesis and cell growth [30]. The GC skew profiles of some of the circular Bathyarchaeia genomes resemble those associated with bidirectional replication from a single origin in Bacteria (Additional file 1: Fig. S10), although not all of the trends are well defined. It is likely that some of these Bathyarchaeia undergo bidirectional replication. Thus, we used iRep to evaluate coverage differences from the origin to the terminus of replication as an indication of replication rate in the two groups of Bathyarchaeia [31]. Bathyarchaeia that apparently circularize their 16S rRNA transcripts had significantly higher replication rates than their counterparts (average iRep values: 1.67 vs. 1.30, $p = 0.000015$; Additional file 1: Fig. S11). Circular RNAs are typically harder to degrade than linear ones [10], and the greater stability of circular rRNAs may promote ribosome synthesis and thus facilitate faster microbial growth.

Where there is evidence for an intermediate circularized form of 16S rRNA, Nanopore long-read sequencing without rRNA pre-depletion showed that the amounts of circular and linear 16S rRNA transcripts were of similar magnitudes (Fig. 3). Thus, it was surprising to note up to thousands of 23S rRNA transcripts in the circular form, with very little evidence for linear counterparts. For example, ~6,000 Nanopore transcripts showed circularization of 23S rRNA from a near-complete *Methanoperedens* genome (bMp) [32] but no transcripts indicated linear versions (Fig. 3). Given that the majority of rRNAs derive from mature ribosomes [30, 33, 34], this finding suggests that the 23S rRNAs in circular forms in the studied archaea are within ribosomes.

To test the hypothesis that some archaeal ribosomes contain a circularized form of the 23S rRNA, we isolated large and small ribosomal subunits from the model archaeon, *Methanosarcina acetivorans* [35], and directly sequenced the extracted 23S and 16S rRNAs. As expected based on our quantitative analysis of transcript mapping (Fig. 3),

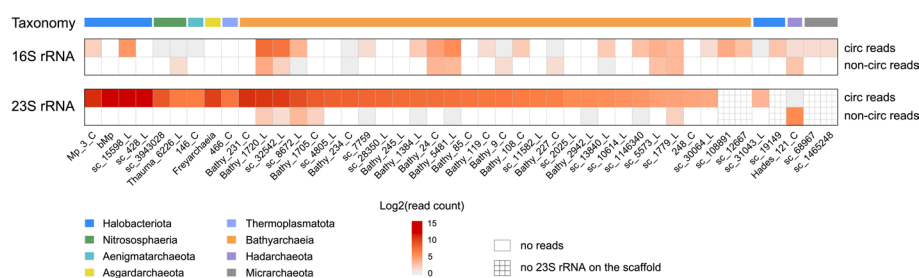


Fig. 3 Quantitation of Nanopore transcripts indicative of different rRNA forms. Nanopore transcript reads were mapped to rRNA genes with maximum mismatches of 1%. The lines “circ reads” indicate transcripts that support the circular form of rRNAs, and “non-circ reads” denote transcripts that support the linear form (see Fig. 1). Read counts were converted logarithmically with base 2 for visualization purposes. The bottom x-axis represents genome names and the top, colored bars indicate taxonomic affiliations. Detailed classifications are present in Additional file 2: Table S1

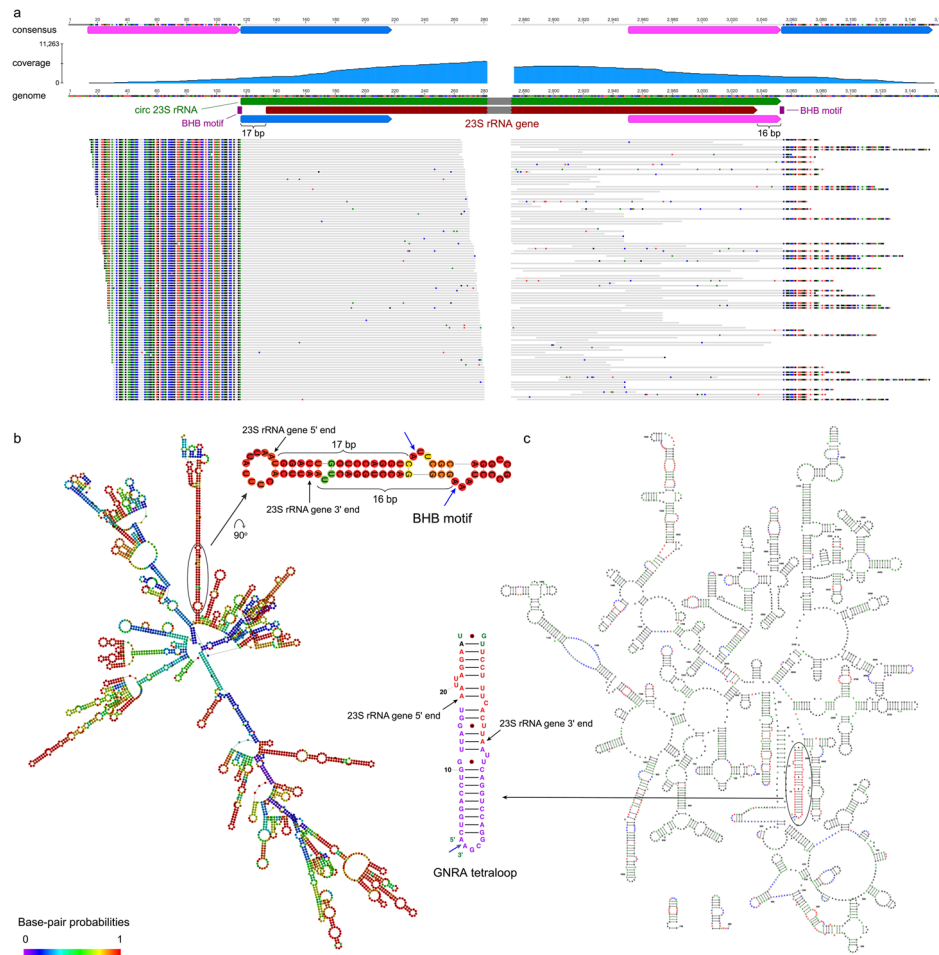


Fig. 4 Circular 23S rRNA within the *Methanosarcina acetivorans* ribosome. **a** *M. acetivorans* genome mapped by 250-bp transcript reads. The dark red arrow indicates the 23S rRNA gene predicted by Rfam and the green arrow indicates its circular transcript form inferred from transcript read mapping. Purple boxes show the bulge regions in the BHB splicing motif. Pink and blue arrows demonstrate identical sequences in the genome reference and in the consensus sequence of mapped reads. Gray bars are mapped reads, in which colored dots are mismatched nucleotides to the genome. Details can be found in Additional file 1: Fig. S13. **b** Predicted secondary structure of the *M. acetivorans* 23S rRNA gene region by the Vienna RNAfold web server. Colors indicate base-pair probabilities. The ellipse region is magnified, showing the locations of the BHB motif and the 5' and 3' ends of the predicted 23S rRNA gene. Blue arrows indicate the cleavage sites on the BHB motif. **c** 2D structure of the circular 23S rRNA in the *M. acetivorans* ribosome predicted by RNAcentral and curated manually. The sequence is capped with the GNRA tetraloop located outside the 5' and 3' ends of the predicted 23S rRNA gene. The blue arrow indicates the site where circularization occurs. The purple nucleotides show the extra 33 bp of the circular 23S rRNA relative to the predicted gene as depicted in (a)

16S rRNA in the *M. acetivorans* ribosomes is linear (Additional file 1: Fig. S12). However, all mapped transcript reads at the start and ends of the 23S rRNA genes (> 3,000) showed discrepant read parts relative to the reference, indicative of circularization, whereas no reads extended into the flanking intergenic regions consistent with linearization (Fig. 4a). The results clearly indicate that the 23S rRNAs within the ribosomes are circular (Additional file 1: Fig. S13). Notably, discrepant regions start at the BHB splicing motif. The *M. acetivorans* circular 23S rRNA is 33 bp longer than the initially predicted gene and is terminated by a GNRA tetraloop that may cap the RNA sequence in circularized form (Fig. 4b and c).

We considered possible permutation events and further re-linearization of circular 23S rRNAs that would require an excision of some parts. It is reported that the 23S rRNA from *Pyrococcus furiosus* is circularly permuted and re-linearized by excising Helix 98 [36]. However, Helix 98 is absent in the 23S rRNA of *M. acetivorans* (Fig. 4c). There is no evidence of an excised region in the alignment between the rRNA transcripts and the rRNA genes. The transcript read mapping showed globally even coverage along the full-length 23S rRNA, except for a region located at positions 1,837-2,007 of the gene (Additional file 1: Fig. S14). It is unlikely that this region has been excised from some transcripts as this region spans domain IV, which crucially contributes to ribosome subunit-subunit interactions and the proper assembly and function of the peptidyl transferase center [37, 38]. Additionally, domain IV has been observed in the cryo-EM structure of the *M. acetivorans* ribosome [39], confirming its presence at the predicted location. The low coverage here might be due to strong secondary structure or RNA modifications that could interfere with reverse transcription during sequencing library preparation, as observed in other archaeal rRNAs [36]. Overall, we conclude that permutation and re-linearization events did not occur after the circularization of 23S rRNAs in the *M. acetivorans* ribosomes.

To examine the potential consequences of incorporating circular 23S rRNAs into ribosomes, we analyzed the cryo-EM structure of a 70S ribosome from a model archaeon, *Thermococcus kodakarensis* (PDB: 6SKF) [40]. The structure shows that the 5' and 3' ends of the 23S rRNA form a helix and the ribosome provides sufficient space for an extended stem-loop, supporting the possibility that the 23S rRNA can be circularized (Additional file 1: Fig. S15a). Additionally, we analyzed the cryo-EM structure of the *M. acetivorans* large ribosomal subunit [39]. In this structure, the helix containing the 5' and 3' ends is not resolved, likely due to dynamics on the surface of the ribosome. We applied a low-pass filter to the associated cryo-EM map (EMD-70864) and observed low-resolution tubular density extending outward from the surface of the ribosome (Additional file 1: Fig. S15b). Although this region lacks sufficient resolution to confirm circularization based solely on cryo-EM density, the observed features are consistent with an AlphaFold 3 prediction of the extended circular stem-loop [41] (Additional file 1: Fig. S15b). In contrast, the start and end of the 16S rRNA are spatially separated in the small ribosomal subunit (Additional file 1: Fig. S15c). As circularization of 16S rRNA would disrupt the ribosome structure, it is unsurprising that we found no evidence for the presence of circularized 16S rRNA within mature ribosomes.

In combination, the extremely low incidence of linear forms of 23S rRNAs and the findings for *M. acetivorans* suggest that many archaea incorporate circular 23S rRNAs into their ribosomes. We considered other explanations for the lack of detectable (or appreciable) linear 23S rRNAs in cases without experimental evidence. One possibility may be the dominance of 23S rRNA pools by circular molecules not yet incorporated into ribosomes. This might be possible if the cells were growing extremely rapidly (but rapid growth would require ribosomes containing mature 23S rRNA). Illumina and Nanopore sequencing with and without rRNA pre-depletion showed very low average transcript coverages in genomic regions that flank rRNA genes (Additional file 1: Fig. S16), indicating that most archaea were not growing. An exception was the Hadarchaeota, whose genes were highly expressed in both Illumina and Nanopore RNA sequencing data (Additional file 1: Fig. S16). This may explain why transcripts indicative

of linear, intermediate 23S rRNAs were more common than their circular counterparts (Fig. 3). It would be surprising if microorganisms transcribe and accumulate, but do not use rRNA, as ribosome assembly is more rapid than rRNA transcription [30]. In conclusion, our findings indicate that circular 23S rRNAs exist within many archaeal ribosomes.

Discussion

Despite the significance of Archaea in the evolution of life, archaeal translation is largely understudied, partially due to the scarcity of laboratory cultures. Here we leveraged culture-independent RNA sequencing to investigate transcriptomes of diverse, mostly uncultivated archaea and revealed unexpected variations in rRNA processing and the conformation of functional molecules in ribosomes.

Our transcriptomic data demonstrated that some lineages of archaea appear not to generate a circular intermediate of 16S rRNA, establishing that the 16S rRNA processing pathway is not universal across the archaeal domain (Fig. 2 and Additional file 1: Fig. S5). This finding raises the question of how archaeal 16S rRNA genes are processed if they cannot go through a circularization step. If the rRNA genes are arranged in a single operon and co-transcribed into 16S–23S–5S rRNA precursors, the 23S rRNAs can be excised at their BHB motifs by EndA, generating linear 16S rRNAs that may be trimmed prior to incorporation into ribosomes. However, if the rRNA genes are encoded at separate locations in genomes [42], the 16S rRNA gene is transcribed by an independent promoter, so EndA cleavage and circularization may not be required. Although trimming is necessary to form mature molecules, it remains unclear which enzymes perform this function. Perhaps these archaea exploit RNases, analogous to bacterial and eukaryotic processing pathways [5, 6, 43].

We observed that, in most archaea, the abundances of circular 23S rRNA transcripts vastly exceeded those of linear counterparts. This contrasts with results of a prior study of a *Methanobrevibacter psychrophilus* species that reported that only 12% of the total 23S rRNA was circular [11]. In part, this may be because the *M. psychrophilus* was either actively growing or had been actively growing recently, so much of the transcribed rRNA was likely intermediates. In contrast, most of the archaea investigated here were not active when sampled (Additional file 1: Fig. S16); thus, ongoing production of linear 23S rRNA intermediates is not expected, and only 23S rRNAs in ribosomes were sampled. However, we acknowledge that degradation of the ends of linear 23S rRNA intermediates may have prevented detection using our methods, so the fraction of linear intermediates may be underestimated.

It is interesting that some (and maybe most) archaea adopt a circularization step during rRNA processing, which does not occur in bacteria or eukaryotes. In the current study, we observed evidence for the existence of circular 16S and 23S rRNA transcripts of Asgard archaea (Fig. 2), which are argued to share a common ancestor with eukaryotes [44–46]. This suggests that eukaryotic RNA processing mechanisms either evolved after divergence from Asgard archaea or were acquired from Bacteria.

We also showed for one archaeal isolate, and inferred for many uncultivated archaea, that the ribosomes incorporate a circularized form of the 23S rRNA. This is a revision of our understanding of ribosome biology and raises the intriguing question of how the inclusion of circular molecules may be advantageous. One conceivable benefit is the

greater stability and durability of ribosomes, as increased interactions between the 5' and 3' ends of 23S rRNA can improve the thermostability of the large subunit [47]. It may also confer resistance to exonucleases and thus reduce ribosome turnover for lower metabolic costs of organisms [48].

It would be interesting to verify the prediction of circularized 23S rRNAs within mature ribosomes of the uncultivated archaea studied here. To do this would require isolation of ribosomes from low-abundance archaea from extremely microbially complex wetland soil samples. We attempted to do this using several strategies (see Methods), but the efforts were unsuccessful.

Overall, our findings of circular 23S rRNA within archaeal ribosomes contrasts with the long-standing view that, across all the domains of life, only linear rRNAs can be assembled into ribosomes [49]. How archaea assemble circular 23S rRNAs into ribosomes remains unclear.

One possibility could be that ribosomal proteins bind consecutively to their corresponding sites on circularized 23S rRNAs directly, with the assembly proceeding in the 5'–3' direction [49–52]. However, circular RNAs impose stronger topological constraints relative to linear counterparts that limit their conformational flexibility [53], which may impede ribosome biogenesis. The other possibility is that linear precursors of circular 23S rRNAs interact with ribosomal proteins and assemble into the large ribosomal subunit, and their free ends are ligated on mature ribosomes. Future imaging experiments based on pure cultures may address these possibilities.

Conclusions

In this study, we discovered unexpected variations in the processing of archaeal 16S and 23S rRNAs and the conformation of mature 23S rRNAs in archaeal ribosomes. We provided evidence that, in at least eight phylum-level archaeal lineages, the 23S rRNAs within mature ribosomes are circular, suggestive of no involvement of re-linearization steps in 23S rRNA processing. We also demonstrated that circular 16S rRNA intermediates are not evident in some archaeal groups, suggesting that 16S rRNA processing in certain archaea does not involve a circularization step.

Methods

Nucleic acid extraction and sequencing

Samples were collected at various depths from wetland soil in Lake County, California, USA. Approximately 5.0 g of soil samples at 75 cm and 140 cm depths were used for DNA extraction using Qiagen DNeasy PowerMax Soil Kit. This DNA was sequenced using PacBio HiFi long reads to generate new metagenomic datasets for recovery of high-quality rRNA genes and for comparison with previous short-read datasets. Other samples from the same site at depths from 40 cm to 115 cm were used for RNA extraction by Qiagen RNeasy PowerSoil Total RNA Kit. A subset of RNA samples from 50 cm, 90 cm, 100 cm, and 115 cm were sequenced using Nanopore long reads [32], and others from 40 cm, 60 cm, 80 cm, and 100 cm were sequenced using Illumina PE150 short reads [23]. Nanopore RNA sequencing did not pre-deplete rRNAs. In brief, all the RNA were polyadenylated using poly(A) polymerase (NEB, M0276L) and purified by GeneJET RNA Cleanup and Concentration Micro Kit (ThermoFisher, K0841). The poly(A)-tailed RNA was reverse transcribed and amplified to generate full-length cDNA according to

the Oxford Nanopore Technologies PCR cDNA Synthesis (PCS109) protocol. The cDNA amplicons were barcoded (EXP-NBD114), pooled and sequenced using FLO-PRO114M flow cells on PromethION.

Processing of metagenomic and metatranscriptomic sequencing data

All raw sequencing reads were processed to filter out low-quality data using BBDuk (<https://jgi.doe.gov/data-and-tools/software-tools/bbtools/>). PacBio HiFi reads were assembled using hifiasm-meta (v0.13-r308) [54]. Scaffolds ≥ 1 kb were binned by a combination of CONCOCT (v1.1.0) [55], MaxBin2 (v2.2.7) [56], MetaBAT2 (v2.15) [57], and VAMB (v3.0.2) [58]. Self-circular PacBio genomes and near-complete genomes that include most single-copy genes and < 20 scaffolds were classified using GTDB-Tk (v2.4.0) based on the GTDB taxonomy (r220) [59, 60]. Open reading frames (ORFs) and ribosomal RNA genes in assembled scaffolds were predicted using Prodigal (v2.6.3) [61] and Rfam 14 [62], respectively.

Identification of genomes that can circularize rRNAs

A total of 11,710 16S rRNA genes were identified from the wetland soil site (Additional file 1: Fig. S3). We removed genes that were shorter than 1,400 bp, likely fragmental genes, and that were longer than 2,000 bp, likely false positives, which resulted in 4,580 putative complete 16S rRNA genes. These genes were clustered at $\geq 99\%$ identity and $\geq 90\%$ coverage, yielding 2,576 species-level 16S rRNA genes. As Nanopore RNA sequencing did not involve the pre-depletion of rRNAs, we mapped processed Nanopore transcript reads to the species-level 16S rRNA genes using Minimap2 (v2.28-r1209) with default parameters in the “map-ont” mode [63]. Resulting SAM files were converted to BAM formats using SAMtools (v1.17) [64], followed by calculations of transcript coverages and covered fractions using CoverM (v0.6.1) [65]. We further discarded genes that have coverages < 5 and/or covered fractions $< 50\%$, leading to a final collection of 1,125 species-level 16S rRNA genes. We retrieved the corresponding genomes of these genes and mapped both Nanopore and Illumina transcript reads to them. Mapping files were visualized in Geneious Prime 2024.0.4 (<https://www.geneious.com>) to figure out the forms of rRNAs, as depicted in Fig. 1. We also preliminarily classified the 16S rRNA genes using the SILVA ACT web service [66].

Gene annotation and phylogeny construction

Gene analyses were conducted on self-circular PacBio genomes and near-complete genomes that include most single-copy genes and < 20 scaffolds. RNA splicing endonuclease (EndA) was identified with HMMER (v3.3) using profile hidden Markov Models (HMM) PF01974 and PF02778 [67]. RNA ligase (RtcB) was identified according to HMM PF01139.

Proteins of interest were blasted against the Genbank nr database to recruit homologues [68]. Query sequences and references were aligned using MAFFT (v7.453) [69]. Alignments were trimmed by trimAl (v1.4.rev15) and then used for phylogeny construction by IQ-TREE (v1.6.12) with automatically-selected best-fit models [70, 71]. Trees were visualized and decorated on the iTOL server [72].

Predicted 16S rRNA genes were blasted against the SILVA database (v138) [73] and the GTDB SSU database (v214) [60] for the recruitment of homologous sequences. Two

Escherichia coli 16S rRNA genes (A14565.1 and AB045730.1) and three *Bacillus subtilis* genes (AB016721.1, AB042061.1 and AB055007.1) were used as the outgroup sequences. Further alignment, trimming, and phylogeny construction were the same as described above.

Measurement of in situ replication rates and theoretical minimal doubling time

Metagenomic sequencing reads from samples at depths of 60–175 cm were mapped to self-circular and near-complete Bathyarchaeia genomes using BBMap with a minimal read identity of 97%. Generated SAM files were sorted by SAMtools (v1.17) and input into iRep (v1.10) to calculate in situ genome replication rates and profile GC skew [31]. Theoretical minimal doubling times of these genomes were estimated using the gRodon package with the “metagenome_v2” mode [74].

Transcript quantification of rRNAs and flanking genomes

To calculate the transcript abundances of circular and linear rRNA forms, Nanopore transcript reads were mapped to rRNA genes with maximum mismatches of 1%. Reads extending beyond rRNA gene ends were analyzed, according to Fig. 1, to quantify circular versus linear forms. Read counts were converted logarithmically with base 2 for visualization purposes.

To calculate the transcript abundances of genomes flanking rRNA genes, rRNA gene sequences were masked with ‘N’, and the modified genomes were mapped with Nanopore and Illumina transcript reads using maximum mismatches of 3%. Average transcript coverages were calculated by dividing total mapped transcript bases by flanking genome length. The resulting values were converted logarithmically with base 2. To avoid biases caused by too short reference length, only genomes and/or scaffolds ≥ 50 kb were included for analysis.

Modeling of RNA 2D structures

Secondary structures of rRNA sequences were predicted on the RNAfold web server [75]. To locate the possible BHB splicing motifs, an extension of sequences flanking the rRNA genes (e.g., 200 bp) was included. rRNA 2D structural diagrams with consistent, reproducible, and recognizable layouts were modeled with RNAcentral [76] and manually curated using XRNA-React [77].

Extraction and sequencing of RNA from *Methanosarcina acetivorans* ribosomes

M. acetivorans ribosomes were isolated from the pure cultures as described previously [35]. RNA was extracted from the ribosomes using Qiagen RNeasy Mini Kit (Cat. No. 74104). Libraries were prepared and sequenced using the Illumina MiSeq 250PE platform in the QB3 Genomics, UC Berkeley. Metatranscriptomic reads were first processed to filter out low-quality data using BBDuk and then mapped to the *M. acetivorans* 16S and 23S rRNA genes, which are predicted by Rfam 14, using BBMap with a minimal read identity of 97%. Matched reads were re-mapped to the *M. acetivorans* genome (GenBank: AE010299.1) with a loose identity of $\geq 70\%$ in Geneious Prime 2024.0.4 (<http://www.geneious.com>) to visualize transcript reads that support circular or linear rRNA forms (Fig. 1). Secondary structures of rRNAs were predicted as described above.

Attempt for ribosome isolation from complex soil samples

Approximately 5.0 g of soil samples were resuspended in 25 mL ribosome buffer A (20 mM Tris-HCl pH 7.5, 100 mM NH₄Cl, and 10 mM MgCl₂) and sonicated in a Q500 sonicator (Qsonica) for 20 min. Another portion of samples were resuspended in 12 mL PowerBead solution and 0.5 mL Solution SR1 with PowerMax beads from the Qiagen RNeasy PowerSoil Total RNA Kit and vortexed at maximum speed for 30 min. The solutions were clarified by centrifugation and loaded onto sucrose cushions containing 24 mL buffer B (20 mM Tris-HCl pH 7.5, 500 mM NH₄Cl, and 10 mM MgCl₂) with 0.5 M sucrose and 17 mL buffer C (20 mM Tris-HCl pH 7.5, 60 mM NH₄Cl, and 6 mM MgCl₂) with 0.7 M sucrose. Ribosomes were pelleted at 57,000 × *g* for 16 h at 4 °C in a Ti-45 rotor (Beckman-Coulter). Pellets were resuspended in ribosome dissociation buffer (20 mM Tris-HCl pH 7.5, 60 mM NH₄Cl, and 1 mM MgCl₂) and were loaded onto a 25–40% (w/v) sucrose gradient in ribosome dissociation buffer. The gradients were centrifuged in a SW-32 rotor (Beckman-Coulter) at 97,000 × *g* for 16 h at 4 °C. Gradients were analyzed on a Biocomp piston fractionator.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13059-025-03903-0>.

Additional file 1: Figures S1–S16. Supplementary information of the paper

Additional file 2: Table S1. Taxonomic classifications of genomes used in the paper

Acknowledgements

We would like to thank Shufei Lei for the support in ggKbase usage and Leylen Miloslavich for the help with lab stuff.

Peer review information

Veronique van den Berghe and Andrew Cosgrove were the primary editors of this article and managed its editorial process and peer review in collaboration with the rest of the editorial team. The peer-review history is available in the online version of this article.

Authors' information

Bluesky handles: @lingdong-shi.bsky.social (Ling-Dong Shi); @banfieldlab.bsky.social (Jillian F. Banfield).

Authors' contributions

The study was designed by L-D.S. and J.F.B. L-D.S. took soil samples, extracted RNA, and processed metatranscriptomic sequencing data. L-D.S. analyzed rRNA forms, gene functions and phylogenies, transcript abundances, and genome replication patterns and rates. P.I.P. and L-D.S. performed 2D structural analysis on rRNAs. A.J.N. performed 3D structural analysis on *T. kodakarensis* and *M. acetivorans* ribosomes. A.J.N. isolated ribosomes from *M. acetivorans* cell pellets with oversight by D.D.N. and J.H.D.C. L-D.S. extracted RNA from *M. acetivorans* ribosomes and analyzed transcriptomic sequencing data. R.S. processed metagenomic sequencing data and provided bioinformatic support. L-D.S. and J.F.B. wrote the paper with input from all the authors. All authors read and approved the final manuscript.

Funding

This work was supported by the Chan Zuckerberg Climate Fund and by the funding from the U.S. Department of Energy (DOE) Chemical Sciences, Geosciences, and Biosciences Division under Contract DE-AC02-05CH11231. Ribosome isolation experiments were supported by the NSF Center for Genetically Encoded Materials (C-GEM; CHE 2002182).

Data availability

Nanopore transcript data can be accessed in the NCBI database under BioProject accession: PRJNA1119519 (BioSample: SAMN41664889–SAMN41664893) [32]. A subset of Illumina transcript data has been deposited in the NCBI database under BioProject accession: PRJNA1050611 (BioSample: SAMN41840554–SAMN41840557) [23]. Other Illumina metagenomic and metatranscriptomic raw sequences, PacBio long reads data, and metagenome-assembled genomes have been deposited under the BioProject accession of PRJNA1247515 [78]. The genomes can also be accessed at: https://ggkbase.berkeley.edu/circ_rna_paper_related. No custom code was used beyond the methods and packages listed in the Methods.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

J.F.B. is a co-founder of Metagenomi. The remaining authors declare no competing interests.

Received: 24 July 2025 / Accepted: 9 December 2025

Published online: 27 January 2026

References

1. Dennis PP. Ancient ciphers: translation in archaea. *Cell*. 1997;89:1007–10. [https://doi.org/10.1016/S0092-8674\(00\)80288-3](https://doi.org/10.1016/S0092-8674(00)80288-3).
2. Becker T, Franckenberg S, Wickles S, Shoemaker CJ, Anger AM, Armache J-P, et al. Structural basis of highly conserved ribosome recycling in eukaryotes and archaea. *Nature*. 2012;482:501–6. <https://doi.org/10.1038/nature10829>.
3. Penev PI, Fakhretaha-Aval S, Patel VJ, Cannone JJ, Gutell RR, Petrov AS, et al. Supersized ribosomal RNA expansion segments in Asgard Archaea. *Genome Biol Evol*. 2020;12:1694–710. <https://doi.org/10.1093/gbe/evaa170>.
4. Giuliano MG, Engl C. In: Kotta-Loizou I, editor. The lifecycle of ribosomal RNA in bacteria. *RNA Damage Repair*. Cham (CH). 2021.
5. Tomecki R, Sikorski PJ, Zakrzewska-Placzek M. Comparison of preribosomal RNA processing pathways in yeast, plant and human cells - focus on coordinated action of endo- and exoribonucleases. *FEBS Lett*. 2017;591:1801–50. <https://doi.org/10.1002/1873-3468.12682>.
6. Clouet-d'Orval B, Batista M, Bouvier M, Quentin Y, Fichant G, Marchfelder A, et al. Insights into RNA-processing pathways and associated RNA-degrading enzymes in archaea. *Fems Microbiol Rev*. 2018;42:579–613. <https://doi.org/10.1093/femsre/fuy016>.
7. Durovic P, Dennis PP. Separate pathways for excision and processing of 16S and 23S rRNA from the primary rRNA operon transcript from the hyperthermophilic archaeobacterium *Sulfolobus acidocaldarius*: similarities to eukaryotic rRNA processing. *Mol Microbiol*. 1994;13:229–42. <https://doi.org/10.1111/j.1365-2958.1994.tb00418.x>.
8. Dennis PP, Ziesche S, Mylvaganam S. Transcription analysis of two disparate rRNA operons in the halophilic archaeon. *J Bacteriol*. 1998;180:4804–13. <https://doi.org/10.1128/Jb.180.18.4804-4813.1998>.
9. Tang TH, Rozhdestvensky TS, d'Orval BC, Bortolin ML, Huber H, Charpentier B, et al. RNomics in archaea reveals a further link between splicing of archaeal introns and rRNA processing. *Nucleic Acids Res*. 2002;30:921–30. <https://doi.org/10.1093/Nar/30.4.921>.
10. Danan M, Schwartz S, Edelheit S, Sorek R. Transcriptome-wide discovery of circular RNAs in archaea. *Nucleic Acids Res*. 2012;40:3131–42. <https://doi.org/10.1093/nar/gkr1009>.
11. Qi L, Li J, Jia J, Yue L, Dong X. Comprehensive analysis of the pre-ribosomal RNA maturation pathway in a methanococcus exposes the conserved circularization and linearization mode in archaea. *RNA Biol*. 2020;17:1427–41. <https://doi.org/10.1080/15476286.2020.1771946>.
12. Grunberger F, Juttner M, Knuppel R, Ferreira-Cerca S, Grohmann D. Nanopore-based RNA sequencing deciphers the formation, processing, and modification steps of rRNA intermediates in archaea. *RNA*. 2023;29:1255–73. <https://doi.org/10.1261/rna.079636.123>.
13. Tyson GW, Chapman J, Hugenholtz P, Allen EE, Ram RJ, Richardson PM, et al. Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature*. 2004;428:37–43. <https://doi.org/10.1038/nature02340>.
14. Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson U, Cheng J-F, et al. Insights into the phylogeny and coding potential of microbial dark matter. *Nature*. 2013;499:431–7. <https://doi.org/10.1038/nature12352>.
15. Castelle CJ, Wrighton KC, Thomas BC, Hug LA, Brown CT, Wilkins MJ, et al. Genomic expansion of domain Archaea highlights roles for organisms from new phyla in anaerobic carbon cycling. *Curr Biol*. 2015;25:690–701. <https://doi.org/10.1016/j.cub.2015.01.014>.
16. Parks DH, Rinke C, Chuvochina M, Chaumeil PA, Woodcroft BJ, Evans PN, et al. Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nat Microbiol*. 2017;2:1533–42. <https://doi.org/10.1038/s41564-017-0012-7>.
17. Ma B, Lu C, Wang Y, Yu J, Zhao K, Xue R, et al. A genomic catalogue of soil microbiomes boosts mining of biodiversity and genetic resources. *Nat Commun*. 2023;14:7318. <https://doi.org/10.1038/s41467-023-43000-z>.
18. Frias-Lopez J, Shi Y, Tyson GW, Coleman ML, Schuster SC, Chisholm SW, et al. Microbial community gene expression in ocean surface waters. *Proc Natl Acad Sci U S A*. 2008;105:3805–10. <https://doi.org/10.1073/pnas.0708897105>.
19. Deng Z-L, Münch PC, Mreches R, McHardy AC. Rapid and accurate identification of ribosomal RNA sequences via deep learning. *Nucleic Acids Res*. 2022;50:e60–e60. <https://doi.org/10.1093/nar/gkac112>.
20. Al-Shayeb B, Schoelmerich MC, West-Roberts J, Valentin-Alvarado LE, Sachdeva R, Mullen S, et al. Borgs are giant genetic elements with potential to expand metabolic capacity. *Nature*. 2022;610:731–6. <https://doi.org/10.1038/s41586-022-05256-1>.
21. Schoelmerich MC, Oubouter HT, Sachdeva R, Penev P, Amano Y, West-Roberts J, et al. A widespread group of large plasmids in methanotrophic methanoperedens archaea. *Nat Commun*. 2022;13. <https://doi.org/10.1038/s41467-022-34588-9>.
22. Shi L-D, West-Roberts J, Schoelmerich MC, Penev PI, Chen L, Amano Y, et al. Methanotrophic methanoperedens archaea host diverse and interacting extrachromosomal elements. *Nat Microbiol*. 2024;9:2422–33. <https://doi.org/10.1038/s41564-024-01740-8>.
23. Valentin-Alvarado LE, Appler KE, De Anda V, Schoelmerich MC, West-Roberts J, Kivenson V, et al. Asgard archaea modulate potential methanogenesis substrates in wetland soil. *Nat Commun*. 2024;15:6384. <https://doi.org/10.1038/s41467-024-49872-z>.
24. Johnson JS, Spakowicz DJ, Hong BY, Petersen LM, Demkowicz P, Chen L, et al. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat Commun*. 2019;10:5029. <https://doi.org/10.1038/s41467-019-13036-1>.
25. Hirata A. Recent insights into the structure, function, and evolution of the RNA-splicing endonucleases. *Front Genet*. 2019;10:103. <https://doi.org/10.3389/Fgene.2019.00103>.

26. Englert M, Sheppard K, Aslanian A, Yates JR, Söll D. Archaeal 3'-phosphate RNA splicing ligase characterization identifies the missing component in tRNA maturation. *Proc Natl Acad Sci U S A*. 2011;108:1290–5. <https://doi.org/10.1073/pnas.1018307108>.
27. Li H, Trotta CR, Abelson J. Crystal structure and evolution of a transfer RNA splicing enzyme. *Science*. 1998;280:279–84. <https://doi.org/10.1126/science.280.5361.279>.
28. Tocchini-Valentini GD, Fruscoloni P, Tocchini-Valentini GP. Structure, function, and evolution of the tRNA endonucleases of archaea: an example of subfunctionalization. *Proc Natl Acad Sci*. 2005;102:8933–8. <https://doi.org/10.1073/pnas.0502350102>.
29. Fujishima K, Sugahara J, Miller CS, Baker BJ, Di Giulio M, Takesue K, et al. A novel three-unit tRNA splicing endonuclease found in ultrasmall Archaea possesses broad substrate specificity. *Nucleic Acids Res*. 2011;39:695–704. <https://doi.org/10.1093/nar/gkr692>.
30. Gourse RL, Gaal T, Bartlett MS, Appleman JA, Ross W. rRNA transcription and growth rate-dependent regulation of ribosome synthesis in *Escherichia coli*. *Annu Rev Microbiol*. 1996;50:645–77. <https://doi.org/10.1146/annurev.micro.50.1.645>.
31. Brown CT, Olm MR, Thomas BC, Banfield JF. Measurement of bacterial replication rates in microbial communities. *Nat Biotechnol*. 2016;34:1256–63. <https://doi.org/10.1038/nbt.3704>.
32. Schoelmerich MC, Ly L, West-Roberts J, Shi L-D, Shen C, Malvankar NS, et al. Borg extrachromosomal elements of methane-oxidizing archaea have conserved and expressed genetic repertoires. *Nat Commun*. 2024;15:5414. <https://doi.org/10.1038/s41467-024-49548-8>.
33. Zhu M, Mu H, Jia M, Deng L, Dai X. Control of ribosome synthesis in bacteria: the important role of rRNA chain elongation rate. *Sci China Life Sci*. 2021;64:795–802. <https://doi.org/10.1007/s11427-020-1742-4>.
34. Rio DC, Ares M, Hannon GJ, Nilsen TW. Removal of ribosomal subunits (and rRNA) from cytoplasmic extracts before solubilization with SDS and deproteinization. *Cold Spring Harb Protoc*. 2010;2010:prot5442. <https://doi.org/10.1101/pdb.prot5442>.
35. Nissley AJ, Shulgina Y, Kivimäe RW, Downing BE, Penev PI, Banfield JF, et al. Structure of an archaeal ribosome reveals a divergent active site and hibernation factor. *Nat Microbiol*. 2025;10:1940–53. <https://doi.org/10.1038/s41564-025-02065-w>.
36. Birkedal U, Beckert B, Wilson DN, Nielsen H. The 23S ribosomal RNA from *Pyrococcus furiosus* is circularly permuted. *Front Microbiol*. 2020;11:582022. <https://doi.org/10.3389/fmicb.2020.582022>.
37. Leviev L, Levieva S, Garrett RA. Role for the highly conserved region of domain IV of 23S-1 lke rRNA in subunit-subunit interactions at the peptidyl transferase centre. *Nucleic Acids Res*. 1995;23:1512–7. <https://doi.org/10.1093/nar/23.9.1512>.
38. Egebjerg J, Christiansen J, Garrett RA. Attachment sites of primary binding proteins L1, L2 and L23 on 23S ribosomal RNA of *Escherichia coli*. *J Mol Biol*. 1991;222:251–64. [https://doi.org/10.1016/0022-2836\(91\)90210-W](https://doi.org/10.1016/0022-2836(91)90210-W).
39. Fordjour GNR, Ghosh A, Ferry JG, Armache J-P, Bevilacqua PC, Murakami KS. Cryo-EM study and in vivo chemical mapping of the Methanosarcina acetivorans ribosome and its dimerization via a repurposed enzyme and translation factor. *J Biol Chem*. 2025;110686. <https://doi.org/10.1016/j.jbc.2025.110686>.
40. Sas-Chen A, Thomas JM, Matzov D, Taoka M, Nance KD, Nir R, et al. Dynamic RNA acetylation revealed by quantitative cross-evolutionary mapping. *Nature*. 2020;583:638–43. <https://doi.org/10.1038/s41586-020-2418-2>.
41. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, et al. Accurate structure prediction of biomolecular interactions with alphafold 3. *Nature*. 2024;630:493–500. <https://doi.org/10.1038/s41586-024-07487-w>.
42. Brewer TE, Albertsen M, Edwards A, Kirkegaard RH, Rocha EPC, Fierer N. Unlinked rRNA genes are widespread among bacteria and archaea. *ISME J*. 2020;14:597–608. <https://doi.org/10.1038/s41396-019-0552-3>.
43. Bechthofer DH, Deutscher MP. Bacterial ribonucleases and their roles in RNA metabolism. *Crit Rev Biochem Mol Biol*. 2019;54:242–300. <https://doi.org/10.1080/10409238.2019.1651816>.
44. Zaremba-Niedzwiedzka K, Caceres EF, Saw JH, Bäckström D, Juzokaite L, Vancaester E, et al. Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature*. 2017;541:353–8. <https://doi.org/10.1038/nature21031>.
45. Liu Y, Makarova KS, Huang WC, Wolf YI, Nikolskaya AN, Zhang XX, et al. Expanded diversity of Asgard archaea and their relationships with eukaryotes. *Nature*. 2021;593:553. <https://doi.org/10.1038/s41586-021-03494-3>.
46. Eme L, Tamarit D, Caceres EF, Stairs CW, De Anda V, Schön ME, et al. Inference and reconstruction of the Heimdallarchaeal ancestry of eukaryotes. *Nature*. 2023;618:992–9. <https://doi.org/10.1038/s41586-023-06186-2>.
47. Nissley AJ, Kamal TS, Cate JHD. Interactions between terminal ribosomal RNA helices stabilize the *E. coli* large ribosomal subunit. *RNA*. 2023;29:1500–8. <https://doi.org/10.1261/rna.079690.123>.
48. Warner JR. The economics of ribosome biosynthesis in yeast. *Trends Biochem Sci*. 1999;24:437–40. [https://doi.org/10.1016/S0968-0004\(99\)01460-7](https://doi.org/10.1016/S0968-0004(99)01460-7).
49. Berg JM, Tymoczko JL, Gatto GJ, Stryer L. *Biochemistry*. 8th ed. New York, NY: W.H. Freeman & Company; 2015.
50. Londei P, Ferreira-Cerca S. Ribosome biogenesis in archaea. *Front Microbiol*. 2021;12:686977. <https://doi.org/10.3389/fmicb.2021.686977>.
51. Sheng K, Li N, Rabuck-Gibbons JN, Dong X, Lyumkis D, Williamson JR. Assembly landscape for the bacterial large ribosomal subunit. *Nat Commun*. 2023;14:5220. <https://doi.org/10.1038/s41467-023-40859-w>.
52. Qin B, Lauer SM, Balke A, Vieira-Vieira CH, Bürger J, Mielke T, et al. Cryo-EM captures early ribosome assembly in action. *Nat Commun*. 2023;14:898. <https://doi.org/10.1038/s41467-023-36607-9>.
53. Liu C-X, Yang L, Chen L-L. Dynamic conformation: marching toward circular RNA function and application. *Mol Cell*. 2024;84:3596–609. <https://doi.org/10.1016/j.molcel.2024.08.020>.
54. Feng X, Cheng H, Portik D, Li H. Metagenome assembly of high-fidelity long reads with hifiasm-meta. *Nat Methods*. 2022;19:671–4. <https://doi.org/10.1038/s41592-022-01478-3>.
55. Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, et al. Binning metagenomic contigs by coverage and composition. *Nat Methods*. 2014;11:1144–6. <https://doi.org/10.1038/Nmeth.3103>.
56. Wu YW, Simmons BA, Singer SW. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics*. 2016;32:605–7. <https://doi.org/10.1093/bioinformatics/btv638>.
57. Kang DWD, Li F, Kirton E, Thomas A, Egan R, An H, et al. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ*. 2019;7:e7359. <https://doi.org/10.7717/PeerJ.7359>.
58. Nissen JN, Johansen J, Allesoe RL, Sonderby CK, Armenteros JJA, Gronbech CH, et al. Improved metagenome binning and assembly using deep variational autoencoders. *Nat Biotechnol*. 2021;39:555–60. <https://doi.org/10.1038/s41587-020-00777-4>.

59. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk v2: memory friendly classification with the genome taxonomy database. *Bioinformatics*. 2022;38:5315–6. <https://doi.org/10.1093/bioinformatics/btac672>.
60. Parks DH, Chuvochina M, Chaumeil PA, Rinke C, Mussig AJ, Hugenholtz P. A complete domain-to-species taxonomy for bacteria and archaea. *Nat Biotechnol*. 2020;38:1079–86. <https://doi.org/10.1038/s41587-020-0501-8>.
61. Hyatt D, Chen GL, LoCascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*. 2010. <https://doi.org/10.1186/1471-2105-11-119>.
62. Kalvari I, Nawrocki EP, Ontiveros-Palacios N, Argasinska J, Lamkiewicz K, Marz M, et al. Rfam 14: expanded coverage of metagenomic, viral and MicroRNA families. *Nucleic Acids Res*. 2021;49:D192–200. <https://doi.org/10.1093/nar/gkaa1047>.
63. Li H, Birol I, editor. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*. 2018;34:3094–100. <https://doi.org/10.1093/bioinformatics/bty191>.
64. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and samtools. *Bioinformatics*. 2009;25:2078–9. <https://doi.org/10.1093/bioinformatics/btp352>.
65. Aroney STN, Newell RJP, Nissen JN, Camargo AP, Tyson GW, Woodcroft BJ, et al. Coverm: read alignment statistics for metagenomics. *Bioinformatics*. 2025;41:btaf147. <https://doi.org/10.1093/bioinformatics/btaf147>.
66. Pruesse E, Peplies J, Glöckner FO. Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics*. 2012;28:1823–9. <https://doi.org/10.1093/bioinformatics/bts252>.
67. Eddy SR. Accelerated profile HMM searches. *PLoS Comput Biol*. 2011;7:e1002195. <https://doi.org/10.1371/journal.pcbi.1002195>.
68. Boratyn GM, Camacho C, Cooper PS, Coulouris G, Fong A, Ma N, et al. BLAST: a more efficient report with usability improvements. *Nucleic Acids Res*. 2013;41:W29–33. <https://doi.org/10.1093/nar/gkt282>.
69. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast fourier transform. *Nucleic Acids Res*. 2002;30:3059–66. <https://doi.org/10.1093/Nar/Gkf436>.
70. Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. TrimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*. 2009;25:1972–3. <https://doi.org/10.1093/bioinformatics/btp348>.
71. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol*. 2015;32:268–74. <https://doi.org/10.1093/molbev/msu300>.
72. Letunic I, Bork P. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res*. 2021;49:W293–6. <https://doi.org/10.1093/nar/gkab301>.
73. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res*. 2013;41:D590–6. <https://doi.org/10.1093/nar/gks1219>.
74. Weissman JL, Hou S, Fuhrman JA. Estimating maximal microbial growth rates from cultures, metagenomes, and single cells via codon usage patterns. *Proc Natl Acad Sci U S A*. 2021. <https://doi.org/10.1073/pnas.2016810118>.
75. Gruber AR, Lorenz R, Bernhart SH, Neubock R, Hofacker IL. The Vienna RNA websuite. *Nucleic Acids Res*. 2008;36:W70–4. <https://doi.org/10.1093/nar/gkn188>.
76. RNAcentral Consortium, Sweeney BA, Petrov AI, Ribas CE, Finn RD, Bateman A, et al. RNAcentral 2021: secondary structure integration, improved sequence search and new member databases. *Nucleic Acids Res*. 2021;49:D212–20. <https://doi.org/10.1093/nar/gkaa921>.
77. McCann H, Meade CD, Williams LD, Petrov AS, Johnson PZ, Simon AE, et al. R2dt: a comprehensive platform for visualizing RNA secondary structure. *Nucleic Acids Res*. 2025;53:gkaf032. <https://doi.org/10.1093/nar/gkaf032>.
78. Shi L-D, Banfield JF. Circular 23S rRNA within archaeal ribosomes. PRJNA1247515. Sequence Read Archive. 2025. <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1247515>.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.