



OPEN ACCESS

EDITED BY

Arthur Charles-Orszag,
University of California, San Francisco,
United States

REVIEWED BY

Fabian Blombach,
University College London, United Kingdom
Megha Patro,
European Molecular Biology Laboratory
(EMBL) Heidelberg, Germany

*CORRESPONDENCE

Ricardo Z. N. Vêncio
✉ rvencio@usp.br
Tie Koide
✉ tkoid@fmrp.usp.br

†PRESENT ADDRESS

Vinicius H. Franceschini-Santos,
Division of Molecular Genetics and Oncode
Institute, Netherlands Cancer Institute,
Amsterdam, Netherlands

RECEIVED 04 June 2025

ACCEPTED 26 September 2025

PUBLISHED 14 October 2025

CITATION

Picinato BA, Franceschini-Santos VH,
Zaramela LS, Vêncio RZN and Koide T (2025)
Archaea express circular isoforms of
IS200/IS605-associated ω RNAs.
Front. Microbiol. 16:1641342.
doi: 10.3389/fmicb.2025.1641342

COPYRIGHT

© 2025 Picinato, Franceschini-Santos,
Zaramela, Vêncio and Koide. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

Archaea express circular isoforms of IS200/IS605-associated ω RNAs

Beatriz A. Picinato¹, Vinicius H. Franceschini-Santos^{1†},
Livia S. Zaramela¹, Ricardo Z. N. Vêncio^{2*} and Tie Koide^{1*}

¹Departamento de Bioquímica e Imunologia, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, Brazil, ²Departamento de Computação e Matemática, Faculdade de Filosofia, Ciências e Letras, Universidade de São Paulo, Ribeirão Preto, Brazil

Circular RNAs (circRNAs) are RNA molecules with 5' and 3' ends covalently ligated. Their functions range from acting as genetic regulators to producing proteins, and they are often expressed in a tissue and condition-specific manner. Next-generation sequencing with prior RNA treatment with the RNase R exonuclease (circRNA-Seq) has been used to identify circRNAs in many organisms, especially in model eukaryotes. However, we know little about circRNAs in prokaryotes: they have not been consistently reported in bacteria and, to date, only a few circRNA-Seq studies have been done in archaea. We have developed a prokaryotic-specific computational pipeline, MonArch, that explores RNA-Seq reads for circRNA signatures. We annotated circRNAs in newly generated *Halobacterium salinarum* circRNA-Seq data and reanalyzed over 20 archaeal public RNA-Seq datasets with this tool. *H. salinarum* has 49 high-confidence circRNAs, with some validated by RT-PCR. We detected known circular ribosomal RNA and transfer RNA processing intermediates and novel circRNAs associated with ω RNAs (obligate mobile element-guided activity - OMEGA) and IS200/IS605 transposases. The ω RNAs circular isoforms have a growth-dependent expression pattern, distinct from the total ω RNAs levels. This is one of the few examples of prokaryotic circRNAs with a conditional expression pattern. In all the other public archaea circRNA-Seq data (*Haloferax volcanii*, *Saccharolobus solfataricus*, *Sulfolobus acidocaldarius*, and *Pyrococcus abyssi*), we found circRNAs associated with the same classes of transcripts as for *H. salinarum*, including circRNAs in IS200/IS605 transposases in the two *Sulfolobales* species. We broadened our search for circRNAs in representatives of major archaeal groups, and found that circRNAs associated with the rRNA and tRNA are widespread, indicating conserved processing of these transcripts. Circular ω RNAs are also present in other haloarchaeal species. Together, our results show that circRNAs appear to be conserved and abundant among archaea, maybe more than previously appreciated. The circular ω RNAs are present in different distant archaeal species, and are a new piece in the IS200/IS605 system.

KEYWORDS

circRNA, archaea, RNA-Seq, IS200/IS605, ω RNA, rRNA, tRNA

1 Introduction

Circular RNAs (circRNAs) are RNA molecules with their 5' and 3' ends covalently ligated. They are involved in many biological processes in eukaryotes (reviewed in [Liu and Chen, 2022](#)) and are long known to be involved in ribosomal (rRNA) and transfer RNA (tRNA) processing in archaea ([Kjems and Garrett, 1988](#); [Lykke-Andersen and Garrett, 1994](#)). They were first discovered in viruses and viroids with circular RNA genomes in

the 1980s (Sanger et al., 1976; Hsu and Coca-Prados, 1979; Kos et al., 1986). Soon after, “scrambled exons” or “mis-splicing subproducts” were found in human and mouse genes (Nigro et al., 1991; Cocquerelle et al., 1992), but they were dismissed as splicing by-products and non-functional molecules. These exons were, in fact, circRNAs (Cocquerelle et al., 1993), but even with more circRNAs being discovered in the following years, they were not given much attention (Patop et al., 2019).

It was only with the popularization of next-generation sequencing (NGS) and the development of specific bioinformatic tools that circRNAs were described as functional and abundant molecules. They are expressed in a condition and tissue-specific manner (Salzman et al., 2012; Jeck et al., 2013; Memczak et al., 2013), are conserved among different species (Westholm et al., 2014; Rybak-Wolf et al., 2015), and can act as transcriptional (Ashwal-Fluss et al., 2014; Gao et al., 2016; Zhang et al., 2016) and post-transcriptional regulators (Hansen et al., 2013; Memczak et al., 2013; Du et al., 2017). CircRNAs can generate new transcripts (Soma et al., 2007; Birkedal et al., 2020) and proteins (Legnini et al., 2017; Yang et al., 2017); their resistance to exonucleases can make them more stable molecules (Lasda and Parker, 2014). In animals, circRNAs can be associated with aging (reviewed in Cai et al., 2019), autism spectrum disorder (Chen et al., 2020), Alzheimer’s disease (reviewed in Zhang et al., 2020), cancer (reviewed in Bach et al., 2019), and viral infections (reviewed in Nahand et al., 2020). The literature on the topic is recent and growing every year.

RNA sequencing specific to find circRNAs (circRNA-Seq) uses RNA treated with RNase R exonuclease to enrich for circRNAs and avoid false positives (Jeck and Sharpless, 2014; Dodbele et al., 2021). This technique has been used extensively in model eukaryotes. However, the explosion in knowledge about circRNAs led by NGS in eukaryotes had no parallel in prokaryotes. Few circRNAs are known in bacteria, with scarce high-throughput initiatives to systematically map them in this domain of life (He et al., 2023). In archaea, only three organisms have circRNA-Seq data published: *Saccharolobus solfataricus* (Danan et al., 2012), *Pyrococcus abyssi* (Becker et al., 2017), and *Haloferax volcanii* (Schwarz et al., 2020). Given the evolutionary implications of the relationship between archaea and eukaryotes (Eme et al., 2018; Williams et al., 2020), it would be important to increase our sampling of diverse archaea to systematically search for circRNAs.

Halobacterium salinarum is an archaeon with an established transcriptional regulation network (Brooks et al., 2014) and post-transcriptional regulation information (Lorenzetti et al., 2023). Even with many types of RNAs identified in this organism (Zaramela et al., 2014; Gomes-Filho et al., 2015; Ten-Caten et al., 2018; de Almeida et al., 2019), circRNAs are still a missing part of its regulatory information and network. Insertion sequences (IS) are prokaryotic mobile genetic elements that are post-transcriptionally regulated in *H. salinarum* (Lorenzetti et al., 2023). With 80 full and 33 partial ISs (Siguier et al., 2006; Kichenaradja et al., 2010), they are believed to contribute to *H. salinarum* genome plasticity and instability (DasSarma, 1993; Dulmage et al., 2018). *H. salinarum* ISs from the IS200/IS605 family harbor sense overlapping transcripts (sotRNAs) in their *tnpB* transposase genes (Gomes-Filho et al., 2015) that were generalized for several other *Halobacteria* in the

RFAM database (RFAM families RF02656 and RF02657). Later, it was observed that many other bacteria and archaea had similar transcripts, and this system was implicated as ancestral to the CRISPR/Cas defense mechanism (Kapitonov et al., 2015; Shmakov et al., 2017; Altae-Tran et al., 2021).

The IS200/IS605 transposases (IscB and TnpB) are the ancestral proteins of Cas9 and Cas12 (Kapitonov et al., 2015; Shmakov et al., 2017; Altae-Tran et al., 2021) and were shown to act as RNA-guided endonucleases (Altae-Tran et al., 2021; Karvelis et al., 2021). Both TnpB and IscB have small RNAs (ω RNAs - obligate mobile element-guided activity, OMEGA RNAs) associated with their 3' or 5' ends, respectively, that interact with the transposase and guide its activity (Altae-Tran et al., 2021; Karvelis et al., 2021). This system has been tested as a new compact gene editing tool in mammals (Li et al., 2024; Xiang et al., 2024). *H. salinarum* sotRNAs have the same relative position to the *tnpB* as the ω RNAs and also have a conserved structure with ω RNA characteristic features. As such, we will refer to *H. salinarum* sotRNAs as ω RNAs from now on.

In this work, we generated a novel circRNA-Seq dataset for *H. salinarum* and found several circRNAs using a custom-made computational pipeline. The bioinformatics approach is suitable for generic RNA-seq data and could retrieve circRNAs in at least 20 different archaea from public databases. By analyzing our novel and public circRNA-Seq datasets, we could retrieve known circRNAs, as well as annotate new circRNAs associated with the IS200/IS605 family and their ncRNAs, the ω RNAs.

2 Materials and methods

2.1 *Halobacterium salinarum* growth conditions, RNA extraction, and sequencing

Halobacterium salinarum NRC-1 cells were grown in complex media (CM) (NaCl 250g/L (Sigma-Aldrich S9888), MgSO₄·7H₂O 20 g/L (Sigma-Aldrich M1880), KCl 2 g/L (Sigma P9541), sodium citrate 3 g/L (Sigma-Aldrich C7254), peptone 10g/L (Oxoid LP0037) (Dyall-Smith, 2009) until OD₆₀₀ ~ 0.5. We extracted small RNAs (<200 nt) using the mirVana miRNA Isolation kit (Ambion). RNA was treated with RNase R as described in (Danan et al., 2012). 120U of RNase R (Biosearch Technologies) was added to 20 μ g of RNA for 45 min at 37 °C. The samples were cleaned with the RNeasy MinElute Cleanup kit (QIAGEN) using the Supplementary Protocol “Purification of miRNA from animal cells using the RNeasy® Plus Mini Kit and RNeasy MinElute® Cleanup Kit” protocol 1, where after RLT buffer, 1.5× volumes of 100% ethanol are added to the sample, applied to the column and eluted in 30 μ L of DEPC water. The RNase R treatment was repeated twice, using 3U of RNase R for a μ g of purified RNA for subsequent treatments.

We prepared the sequencing library using the TruSeq mRNA Stranded RNA Sample Preparation kit (Illumina), and they were sequenced using the MiSeq Reagent v2 50 cycles kit (Illumina) on the Illumina MiSeq (Illumina). All reactions and preparations were made with two biological replicates.

2.2 RNA-Seq data used in reanalysis

We reanalyzed all available archaeal RNase R-treated RNA-Seq available for the identification of circRNAs (as of June 2024). We reanalyzed RNA-Seq data from *Haloferax volcanii* (PRJEB40302) (Schwarz et al., 2020), *Pyrococcus abyssi* (personal communication) (Becker et al., 2017), *Saccharolobus solfataricus* (personal communication) (Danan et al., 2012), and *Sulfolobus acidocaldarius* (PRJNA388657) (Orell et al., 2018).

To make circRNA expression profiles relative to the linear counterpart, we searched for *H. salinarum* circRNAs in RNA-Seq data from different growth stages (López García de Lomana et al., 2020).

We searched for circRNAs in regular RNA-Seq data of representatives of major archaeal groups.

The accession numbers for all data used in this study are in Table S1.

2.3 Computational pipeline to identify circRNAs

We developed MonArch, a computational pipeline to identify circRNAs in RNA-Seq data. It is based on the premise that reads from the circularization junction align regularly in a chiasmic manner in the genome (Figure 1A). The pipeline can be divided into two main parts: (1) identification of individual circularization junctions in the reads and (2) grouping of similar junctions into circRNA ensembles (Figure 1A).

MonArch uses the reference genome and the RNA-Seq reads in FASTA format as input; if there are replicates, they are merged into one file. In the first step, the reads are aligned to the reference genome with BLASTn (Altschul et al., 1990), as in other prokaryotic circRNA studies (Danan et al., 2012; Becker et al., 2017). MonArch uses the blastn-short routine with ungapped alignments. Then, a custom Python script searches for a pair of alignments from the same read that could represent a circularization junction to annotate it. The alignments must be uniquely aligned in the genome, do not have mismatches, be no further than 3,500 bases from each other (maximum size we allowed for a circRNA), and together cover at least 90% of the read (Figure S1A). Moreover, the best of the two BLASTn alignments must cover at least half of the read, and the other one should be at least eight bases long (Figure S1A). We allow alignments to have at most a 3nt overlap or gap between them (Figure S1B), as other prokaryotic studies have done before (Danan et al., 2012; Becker et al., 2017). An overlap occurs when a base in the circularization junction can be aligned to the reference genome by either alignment of the pair, while a gap occurs when a base in the circularization junction does not align with the reference genome. The coordinates of the circularization junction are adjusted accordingly.

The pipeline then groups close circularization junctions into a circRNA ensemble (Figure 1A). Circularization junctions with their start coordinates distant at most 3nt and end coordinates distant at most 3nt are grouped into one entity. The final coordinates of the circRNA ensemble are the minimum among the start coordinates and the maximum among the end coordinates in the forward

strand; the reverse is done for the reverse strand. The coordinates of the ensemble are not necessarily the same as the coordinates of the junctions that it is made of. We use “circRNA” in the rest of this manuscript to refer to the circRNA ensemble.

Some of these parameters can be altered by the user, but for the analyses of this article, we used the default parameters described here for all organisms, with a few exceptions. For all organisms except *H. salinarum*, we used the option “-dont_invert_strand” because of the nature of the sequencing kits used for each dataset. For *S. solfataricus*, because the sequencing protocol used was not strand-specific, we considered all reads to be on the forward strand after aligning for simplicity before grouping them into circRNA ensembles.

The MonArch pipeline can be found at <https://github.com/bpicinato/MonArch>.

2.4 circRNA identification and annotation in RNA-Seq data

To identify circRNAs in RNA-Seq data, we first trimmed the sequencing adapters (Table S1) from the reads and discarded reads smaller than 20nt using Trimmomatic v0.39 (Bolger et al., 2014). We did not perform this step only in *P. abyssi* data, which was sequenced in an Ion Torrent platform (Thermo Fischer Scientific). Then, if the data was from paired-end sequencing, we grouped the FASTQ R1 and R2 files in one file. Finally, we converted the FASTQ files into FASTA files and used them as input for the MonArch pipeline for circRNA identification. We considered significant circRNAs with 10 or more reads supporting the circularization junction.

Annotation of circRNAs was performed with bedtools intersect (Quinlan and Hall, 2010) using the ensembles output BED file containing the circRNAs and a GFF annotation file. Data visualization and integration were done in Integrative Genomics Viewer (Robinson et al., 2011).

To annotate *H. salinarum* NRC-1 circRNAs, we used the gene annotation from (Pfeiffer et al., 2019), the sotRNA/ωRNA annotation from (Gomes-Filho et al., 2015), and the insertion sequence annotation from ISfinder/ISbrowser (Siguier et al., 2006; Kichenaradja et al., 2010).

For *H. volcanii* DS2, we used gene annotation from NCBI (assembly ASM2568v1) and the insertion sequence annotation from HaloLex (Pfeiffer et al., 2008).

For *P. abyssi* circRNAs, we used the gene annotation from NCBI (assembly ASM19593v2), the insertion sequence annotation from ISfinder/ISbrowser (Siguier et al., 2006; Kichenaradja et al., 2010), and the C/D box RNA annotation from (Omer et al., 2000) and (Toffano-Nioche et al., 2013) and the LoweLab (<https://lowelab.ucsc.edu/>).

For *S. solfataricus* P2, we used the gene annotation from NCBI (assembly ASM1228v1), the insertion sequence annotation from ISfinder/ISbrowser (Siguier et al., 2006; Kichenaradja et al., 2010), and the annotation for non-coding RNAs and C/D box RNAs from (Tang et al., 2005) and (Zago et al., 2005).

For *S. acidocaldarius* DSM639, we used the gene annotation from NCBI (assembly ASM700v1), the insertion sequence

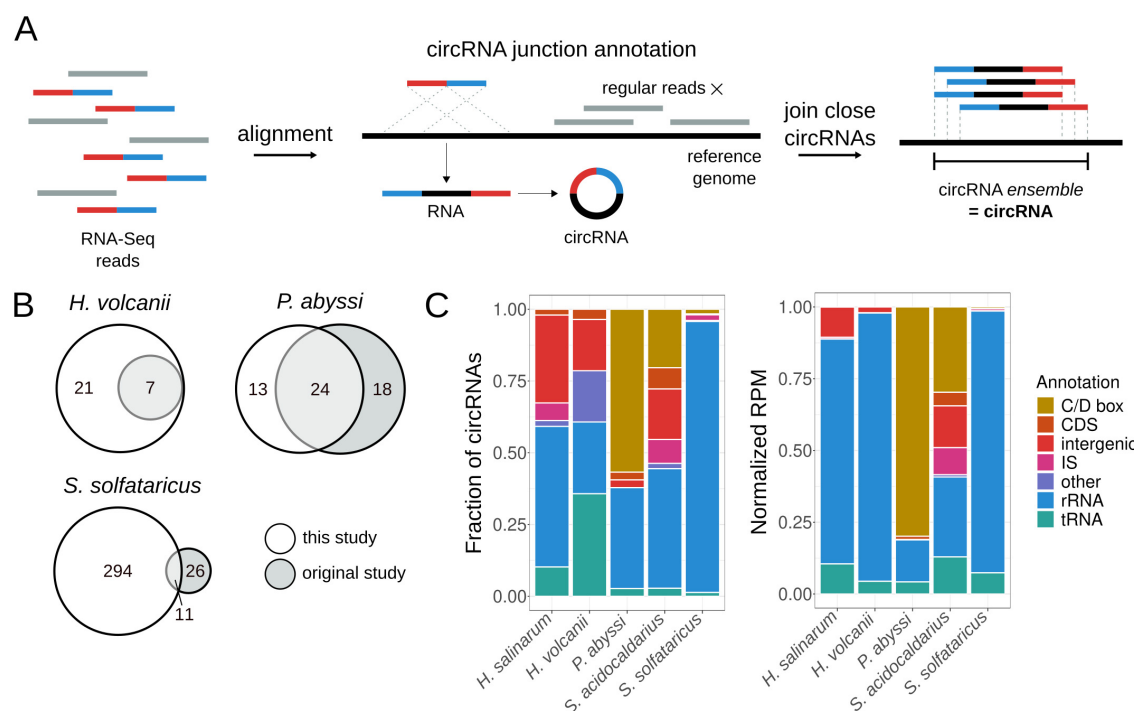


FIGURE 1

MonArch discovers circRNAs in archaeal circRNA-Seq data. **(A)** Schematic of the MonArch pipeline. First, it aligns the input RNA-Seq reads in the provided reference genome using BLASTn and searches for reads with a chiasmic alignment that came from a circularization junction. MonArch then groups close annotated circRNA junctions into one entity, a circRNA ensemble, that we call interchangeably as a circRNA. **(B)** Venn diagrams showing how many circRNAs were identified in this study and the original study. **(C)** Categorization of circRNAs annotated by MonArch in all archaeal circRNA-Seq data. Left: fraction of circRNAs in each annotation category. Right: Reads per Million (RPM) counts present in each annotation category, normalized to the circRNA total of each particular study to allow direct comparison. Annotation of circRNAs as in Table S3.

annotation from ISfinder/ISbrowser (Siguier et al., 2006; Kichenaradja et al., 2010), the small RNAs annotation from (Orell et al., 2018) and C/D box RNAs from (Omer et al., 2000) and the LoweLab (<https://lowelab.ucsc.edu/>).

For the other organisms for which RNase R-treated RNA-Seq data do not exist, we used only the gene annotation from NCBI (accession number of assemblies used in Table S1).

2.5 Expression profiles of *H. salinarum* circRNAs

We searched for *H. salinarum* circRNAs associated with its ω RNAs and IS200/IS605 in RNA-Seq data of total RNA extracted from different times of a growth curve (López García de Lomana et al., 2020). We considered for analysis circRNAs with at least (i) ten reads supporting the circularization junction in the growth-curve dataset and (ii) one read in the RNase R-treated data. Many ω RNAs had more than one circRNA associated with them; we chose the one with the most read counts across all time points and followed its expression along the growth curve. The coordinates of the circRNAs we analyzed for each ω RNA are in Table S4.

The counts of circRNAs are the counts of how many reads are aligned in their circularization junction. We also counted how many reads aligned in each ω RNA with the featureCounts function from the Rsubread R package (Liao et al., 2019); these are

considered the “total reads” (since they are a combination of the reads from linear and circRNA, but not from the circularization junction). The read counts were normalized by the total number of reads sequenced in each time point and corresponding replicate, and multiplied by a million (reads per million, RPM). We calculated the mean and standard error of the three biological replicates for each time point and transcript type (circular or total).

2.6 RT-PCR for circRNA experimental validation

Total RNA was extracted from *H. salinarum* NRC-1 grown in the same conditions described in section 2.1 with acid phenol-chloroform (Ambion) and precipitated with ethanol. Small RNAs were extracted with the mirVana miRNA Isolation kit (Ambion). RNase R treatment was performed once with 6U of enzyme (Biosearch Technologies) per μ g of RNA. The reaction was cleaned with the RNeasy MinElute Cleanup Kit (QIAGEN). Reverse transcription was done with SuperScript III Reverse Transcriptase (Invitrogen) with random primers and the PCR with GoTaq 2x Master Mix (Promega) using annealing temperatures between 50 °C and 66 °C. We used divergent primers to amplify the circularization junctions (Dodbele et al., 2021). The sequences of the primers used in this study are in Table S2.

2.7 In silico tools used for RNA structure analysis

The structures of the rRNAs were predicted using R2DT (Sweeney et al., 2021), visualized with RNACanvas (Johnson and Simon, 2023), and compared with 16S and 23S structures available at Ribovision (Bernier et al., 2014) to determine helix and motif numbering.

We determined the conserved structures of *H. salinarum* ω RNAs using LocARNA with default parameters (Will et al., 2012; Raden et al., 2018). Pseudo-knot structures were predicted with IPknot (v2.2.1) (Sato et al., 2011). The individual structures of the ω RNAs were predicted with Vienna RNAfold (Gruber et al., 2008) and visualized in RNACanvas (Johnson and Simon, 2023). We removed the base pairing of the predicted guide sequence from the final structures.

2.8 Data and code availability

H. salinarum raw RNA-Seq data is available at NCBI's Sequence Read Archive (SRA) under the BioProject accession number PRJNA1268524. All the accession numbers for data reanalyzed in this study are in Table S1. The MonArch pipeline is available at <https://github.com/bpicinato/MonArch>.

3 Results

3.1 The MonArch pipeline discovers circRNAs in archaea

We developed a computational pipeline, MonArch (<https://github.com/bpicinato/MonArch>), to identify circRNAs in RNA-Seq data (Figure 1A and Figure S1 for details). MonArch searches for RNA-Seq reads that contain the circularization junction sequence. Reads from these junctions do not align normally in the genome but in a chiasmic manner (Figure 1A), a characteristic that allows one to distinguish a circRNA from its linear cognate. MonArch then groups close circRNAs into one entity, as previous works in prokaryotes have done (Figure 1A) (Danan et al., 2012; Becker et al., 2017; He et al., 2023).

We used MonArch to annotate the circRNAs in *H. salinarum* circRNA-Seq data we generated, and reanalyzed all published archaeal circRNA-Seq data to understand circRNA prevalence and distribution among archaea. We reanalyzed data for *H. volcanii* (Schwarz et al., 2020), *P. abyssi* (Becker et al., 2017), *S. acidocaldarius* (Orell et al., 2018), and *S. solfataricus* (Danan et al., 2012). For all data analyzed, we considered significant a circRNA with 10 reads or more supporting its circularization junction.

In our reanalysis, MonArch could reasonably recover the results of the original studies, with different success rates for each organism (Figure 1B). It is important to note that different tools or approaches annotate different circRNAs (Hansen et al., 2016; Chen et al., 2021) and that the criteria for the significance of a circRNA vary among all studies. We recovered all the circRNAs originally annotated in *H. volcanii* and annotated 21 new ones associated with tRNAs, the rRNA operon, and the signal recognition particle. For *P.*

abyssi, we found the majority (57%) of the circRNAs annotated in the original study and annotated 13 new circRNAs associated with the rRNA operon. Among the circRNAs found only in the original study, 16 (89%) were present in the non-significant circRNAs of our reanalysis. For *S. solfataricus*, we recovered 11 out of the 37 (30%) previously annotated circRNAs and found 294 new circRNAs associated with rRNA, IS, and C/D box small RNAs. 19 (73%) circRNAs found only in the original study were supported by fewer than 10 reads in our analysis. The *S. acidocaldarius* original study did not use the data for circRNA annotation.

For all organisms, MonArch recovered known archaeal circRNAs, such as rRNA and tRNA circular processing intermediates, as well as new circRNAs associated with IS, the rRNA operon, and tRNAs (Figure 1C). The following sections will detail the circRNAs found in *H. salinarum* and other archaea.

3.2 MonArch recovers known circRNAs in *Halobacterium salinarum*

We identified 49 high-confidence circRNAs in *H. salinarum* NRC-1 (Table S3). They were associated with three main classes of transcripts: rRNA, tRNA, and IS (Figure 1C). Most of them were associated with the rRNA, followed by intergenic circRNAs, tRNAs, IS, and one circRNA in a hypothetical protein-coding gene (Figure 1C, Table S3). Among them, we could find known circRNAs associated with the 16S rRNA bulge-helix-bulge (BHB) motif, the 5S rRNA, and two circular tRNA introns.

The circular processing intermediate associated with the 16S rRNA BHB motif in *H. salinarum* is supported by 2357 reads (circRNA_0143, Table S3, Figure 2A). The pre-23S circular rRNA was supported by 5 reads, probably due to the enrichment for small RNA step performed before RNA sequencing in the *H. salinarum* (Figure S2).

We also identified 7 different circRNAs associated with the 5S rRNA (Table S3, Figures 2A and Figure S3). The one with the most reads (circRNA_0338) encompasses helices II-V, while the second, in the number of reads, encompasses the whole 5S (circRNA_0325) (Figure S3). Interestingly, we found a lot of variability among the start and end coordinates of these circRNAs, especially circRNA_0331 (Figure S3). This might suggest that circularization is a part of the degradation process of the 5S rRNA, as proposed for *S. solfataricus* (Danan et al., 2012).

H. salinarum has two tRNA introns (in tRNA-Trp VNG_RS03925 and tRNA-Met VNG_RS06350); both are circularized at a BHB motif (Figures 3A,B) (circRNA_0059 and circRNA_0074, respectively, in Table S3). tRNA-Trp also has a circRNA with the exact same coordinates as the circular intron but in the opposite strand (circRNA_0060, Table S3). As in *H. volcanii* (Clouet d'Orval et al., 2001) and *P. abyssi* (Omer et al., 2000), *H. salinarum* tRNA-Trp intron is also a C/D box RNA (Weisel et al., 2010).

Using RT-PCR, we validated the circular pre-16S and pre-23S, the circRNAs in the 5S rRNA (Figure 2C, Figure S4), and the tRNA-Trp intron (Figure 3C, Figure S7).

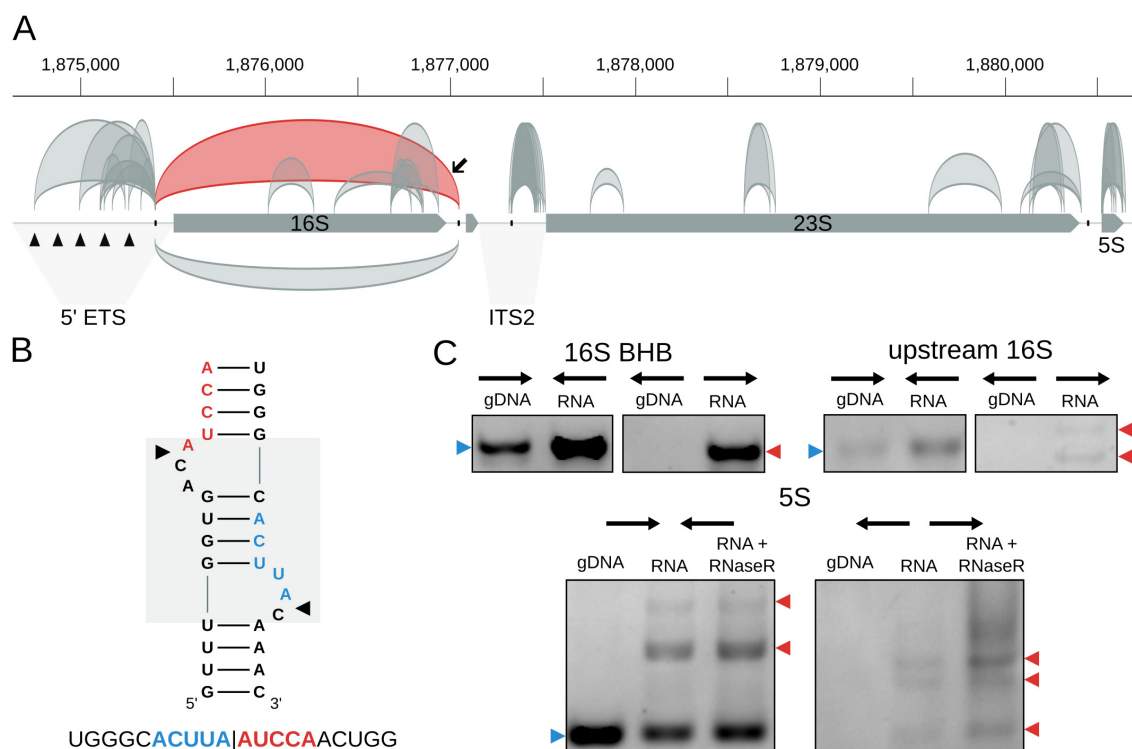


FIGURE 2

circRNAs in the *Halobacterium salinarum* rRNA operon. **(A)** Scheme of circRNAs in the rRNA operon. Gray arrows represent the genes (16S, tRNA-Ala, 23S, and 5S), while arcs represent the annotated circRNAs with 10 reads or more supporting the circularization junction. On top, circRNAs are in the forward strand; on the bottom, they are in the reverse strand. The circRNA associated with the 16S bulge-helix-bulge (BHB) structural motif is highlighted in red, indicated with a black arrow. Black triangles in the 5' external transcribed spacer (5'-ETS) region mark the five octanucleotide sequences that are putative transcription start sites. The internal transcribed region (ITS2) is also indicated. The BHB (bulge-helix-bulge) motif positions are indicated with the black ticks in the operon. Coordinates on top of the main chromosome (NC_002607.1), in base pairs. **(B)** 16S BHB structural motif with the most abundant circRNA junction sequence identified below. The gray box in the background highlights the BHB motif, and black arrows mark the canonical processing sites. **(C)** RT-PCR validations of rRNA circRNAs. Convergent arrows represent reactions made with convergent primers (expected to amplify both linear and circular targets), while divergent arrows show reactions with divergent primers (expected to amplify only circular targets). Blue arrows (left) indicate the expected linear product, while red arrows (right) indicate the expected bands for circRNA junction amplification. The multiple arrows in the 5S validation gel indicate expected band sizes for circRNA_0098 and circRNA_0108; multiple arrows in the 5S validation gel indicate different bands made by rolling circle amplification. gDNA = genomic DNA; RNA = reactions made with cDNA amplified from total RNA; RNA + RNase R = reactions made with cDNA made from RNA treated with RNase R. Uncropped images can be found in [Figure S4](#).

3.3 *Halobacterium salinarum* has novel circRNAs

3.3.1 circRNAs associated with the rRNA operon

Besides the canonical rRNA circRNAs, we also found many other smaller circRNAs associated with the rRNA operon. All the circRNAs annotated as “intergenic” ([Figure 1C](#), [Table S3](#)) are associated with the rRNA operon upstream of the 16S rRNA or 23S rRNA genes in the 5' external transcribed spacer 5'-ETS) and internal transcribed spacer 2 (ITS2) ([Figure 2A](#)). In the 5'-ETS region, there are five octanucleotide sequences (TGCGAACG) that are putative transcription start sites ([Chant and Dennis, 1986](#)). Four of the eight circRNAs in this region start in these sites, and all end at the BHB processing site ([Figure 2A](#)). This might suggest that circularization occurs between the 5' end of the operon transcript and the 3' left after BHB endonuclease cleavage. These circRNAs were validated using RT-PCR ([Figure 2C](#), [Figure S4](#)). In the ITS2 region, all the circRNAs start at the BHB site and end near the start of the 23S gene ([Figure 2A](#), [Figure S2](#)).

Inside the 16S and 23S genes, the circRNAs accumulate in their 3' portions ([Figure 2A](#), [Figures S5, S6](#)). In the 16S, most circRNAs are in the 3'M and 3'm domains ([Figure S5](#)). In the 23S, the circRNAs are concentrated in the VI domain, predominantly in helices 94-101 ([Figure S6](#)).

3.3.2 Circular tRNAs

We found two whole circular tRNAs in addition to the circular tRNA introns: circRNA_0047 in tRNA-Met (VNG_RS02185) and circRNA_0068 in tRNA-Leu (VNG_RS05810) ([Figure S8](#)). Their circularization junctions do not have all bases aligned in the reference genome; both have either C, CC, or CCA extra bases between the two halves of the alignment. The sequence “CCA” is formed in the circularization junction in all cases. This exact sequence is added to the 3' end of the tRNA in its maturation process ([Clouet-d'Orval et al., 2018](#)), suggesting circularization occurs after this processing step.

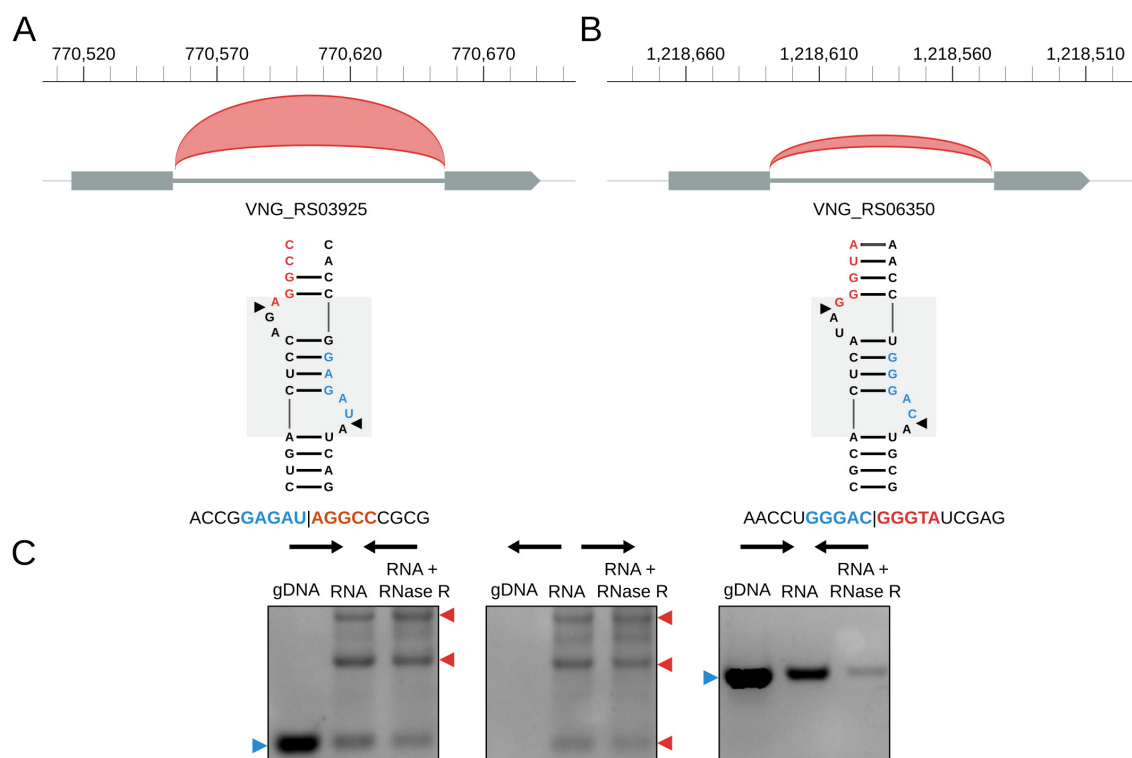


FIGURE 3

Halobacterium salinarum circRNAs in tRNA genes. (A) circRNA_0059 (red arc) in tRNA-Trp intron and its associated BHB structural motif. The gray box in the background highlights the BHB motif, and black arrows mark the canonical processing sites. Below is a representative read of this circRNA. Coordinates on top of the main chromosome (NC_002607.1), in base pairs. (B) circRNA_0074 (red arc) in tRNA-Met intron and its associated BHB structural motif. The gray box in the background highlights the BHB motif, and black arrows mark the canonical processing sites. Below is a representative read of this circRNA. Coordinates on top of the main chromosome (NC_002607.1), in base pairs. (C) RT-PCR validation of the circRNA associated with the tRNA-Trp intron. The left panel is a reaction made with convergent primers expected to amplify both linear and circular templates. The middle panel is the reaction made with divergent primers expected to amplify only circular products. The right panel is made with primers that amplify a linear product that is digested by RNase R. Blue arrows (left) indicate the expected linear product, while red arrows (right) indicate the expected bands for circRNA junction amplification. The uncut agarose gel is in Figure S7.

3.3.3 Circular isoforms of IS200/IS605-associated ωRNAs

Finally, we identified novel circRNAs associated with IS in *H. salinarum*. We found two circRNAs (circRNA_0012 and circRNA_0397, Table S3) associated with the IS200/IS605 family and their ωRNAs (originally named sense overlapping transcripts, sotRNAs, by Gomes-Filho et al., 2015) (Figures 4A, B) and one circRNA in an ISH3/IS4 (Table S3).

H. salinarum has 10 ωRNAs with a conserved structure among them (Figure S9). They seem to have characteristic features of previously identified and characterized ωRNAs, such as a pseudoknot between the second hairpin loop and the TEM sequence and a right-end hairpin (Figures 4C,D, Figures S9, S10) (He et al., 2015; Nakagawa et al., 2023; Sasnauskas et al., 2023; Žedaveinyte et al., 2024).

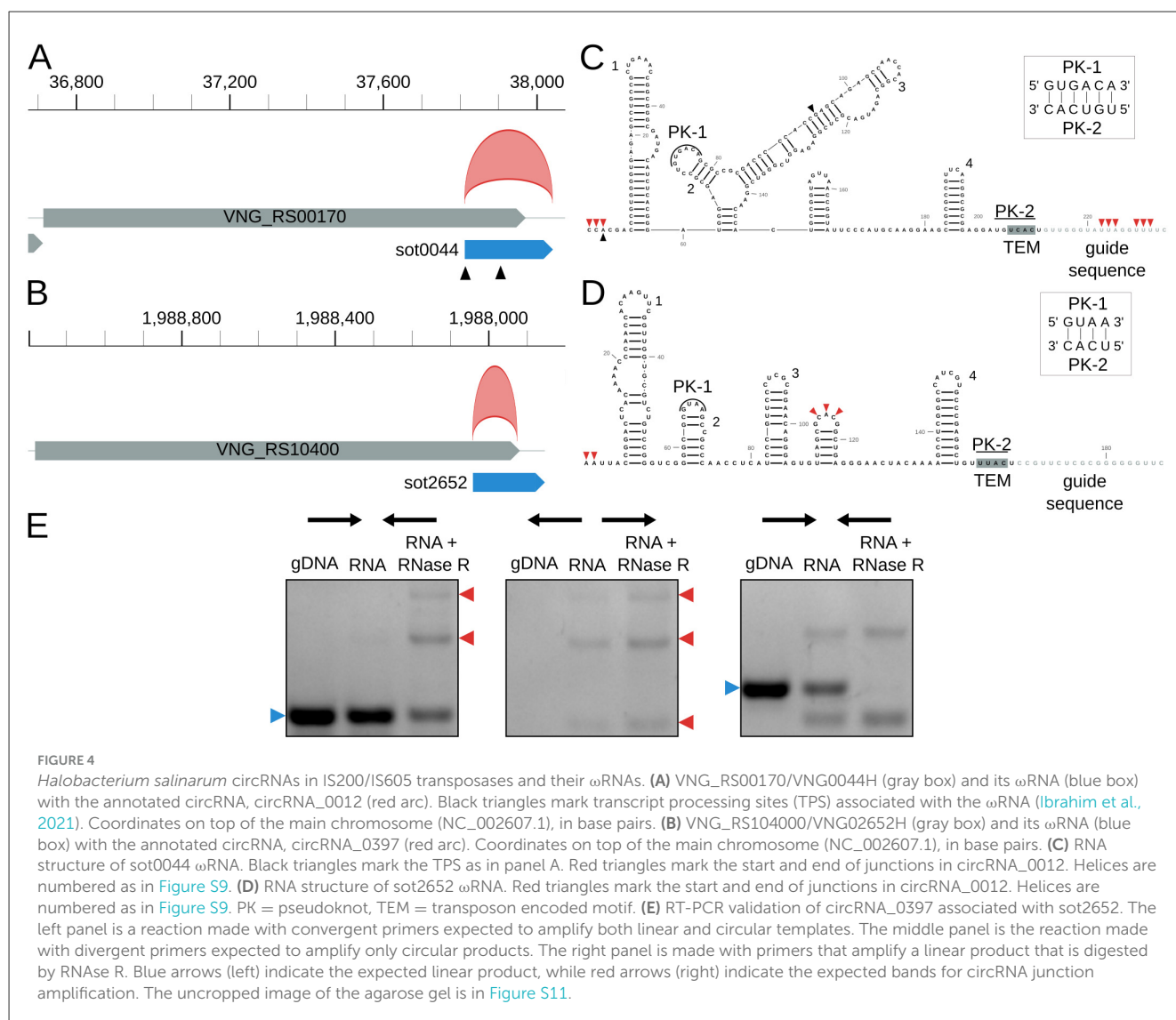
We mapped the circRNAs to the sot0044 and sot2652 ωRNA structures, and we could not find a BHB structural motif associated with them (Figures 4C, D). circRNA_0397 is in the 5' portion of the RNA, while circRNA_0012 encompasses the whole RNA, including part of the predicted guide sequence. circRNA_0012 coordinates also match a previously annotated RNA processing site (Ibrahim et al., 2021). Both start at the 5' end of their corresponding ωRNAs.

We validated circRNA_0397 using RT-PCR (Figure 4E, Figure S11).

3.4 *H. salinarum* circular ωRNAs have a growth-dependent expression pattern

In eukaryotes, circRNAs can be expressed in a tissue and condition-dependent manner (Salzman et al., 2012; Jeck et al., 2013; Memczak et al., 2013). We searched for circular ωRNAs in other *H. salinarum* RNA-Seq data to understand their expression patterns. We reanalyzed regular RNA-Seq data of *H. salinarum* in different stages of growth (López García de Lomana et al., 2020). Lomana and colleagues sequenced RNA from early (T1, O.D.₆₀₀ = 0.2), mid- (T2, O.D.₆₀₀ = 0.5), and late (T3, O.D.₆₀₀ = 0.8) exponential phases, and from the stationary phase, at 40.8h of growth (T4).

We compared the expression profile of the reads that aligned in the circularization junction, the “circular reads”, and the reads that aligned normally in the ωRNA locus, “total reads”, as they comprise both the reads from the linear isoform and reads that came from the circRNA but not the circularization junction (Figure 5). All



the circRNAs analyzed were more expressed in the later stages of growth (Figure 5). They all present different expression patterns and levels among themselves and between the total and circular reads (Figure 5): (i) the total RNA expression is higher in T3, lower in T4, while the circRNA has a peak in T4 (sot0013 and sot0044); (ii) both classes roughly have the same expression pattern (sot6181 and sot6361); and (iii) the total RNA count is fairly constant along the growth curve while the circRNA expression rises from T2 to T4 (sot2652). It is also worth noting that the expression levels of the circular ωRNA do not always correlate with the expression for the whole locus; for example, the highest expressed ωRNA is sot0013 in T3, but the highest expressed circRNA is from sot6181, which has the lowest total read count (Figure 5, Table S4).

It is known that circRNAs can be stable molecules that outlast their linear counterparts (Ashwal-Fluss et al., 2014). This stability could explain why circular ωRNAs are more present in the latter stages of cell growth. However, the counts for each transcript type are in orders of magnitude of difference (Table S4), and in the cell, the circRNAs probably account for a small fraction of the transcripts from ωRNAs. In *Bacillus altitudinis*, circular DucS RNA

abundance increases in the latter stages of growth while one of its linear isoforms decreases (He et al., 2023). Together, these two results are the only examples, to our knowledge, of condition-specific expression of prokaryotic circRNAs.

3.5 circRNAs are conserved among archaea

In our reanalysis of archaeal circRNA-Seq, we found the same classes of circRNAs we found in *H. salinarum*, in addition to some other already known archaeal circRNAs (Figure 1C, Table S3).

In the circRNA-Seq data, we found circRNAs associated with the rRNA operon and the canonical circular pre-16S and pre-23S in all organisms (*H. volcanii*, *S. solfataricus*, and *S. acidocaldarius*) except *P. abyssi* (Figure 6, Figure S12). We also found small circRNAs in the rRNA operon in *P. abyssi*, *S. solfataricus*, and *S. acidocaldarius*. In *P. abyssi*, these circRNAs are concentrated in the 3' end of the 23S gene, as in *H. salinarum*. In *S. solfataricus*, the circRNAs inside the 16S and 23S genes do not seem to be concentrated in any region in particular. In *S. acidocaldarius*, the

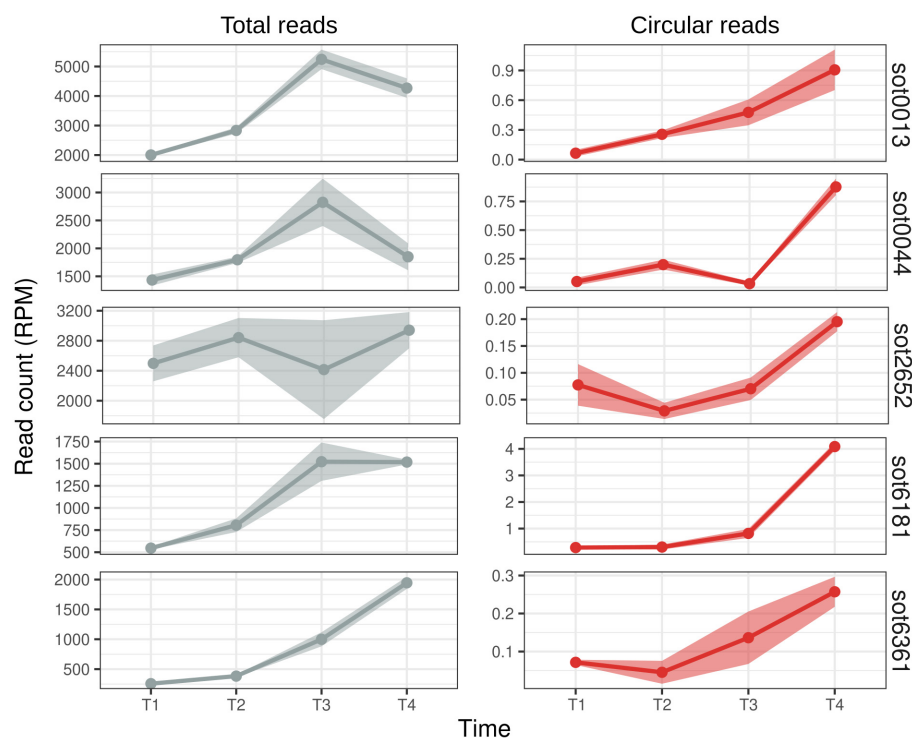


FIGURE 5

Expression profile of *Halobacterium salinarum* ω RNAs during growth. Read counts (normalized in reads per million - RPM) along the growth curve (T1 = early exponential, T2 = mid-exponential, T3 = late exponential, and T4 = stationary) of *H. salinarum* ω RNAs. Total reads are the reads that aligned normally in the ω RNA locus (left, gray), and circular reads are the reads from the circularization junction (right, red). The mean and standard error of the read counts of 3 biological replicates are shown.

small 16S circRNAs are concentrated in its 5' portion, while in the 23S, they seem to accumulate in domain I. *H. volcanii* has only one small circRNA in the rRNA operon (circRNA_0340, Table S3). We believe that *H. volcanii* is different from the other organisms in this aspect because this was the only dataset that did not have a small RNA enrichment step before sequencing. Both *S. solfataricus* and *S. acidocaldarius* have circRNAs in the 5S rRNA (Figure 6, Figure S13, Table S3).

We identified the circular tRNA-Trp intron for all organisms reanalyzed, besides other circular tRNA introns for *H. volcanii*, *S. solfataricus*, and *S. acidocaldarius* (Figure 6). In *H. volcanii*, we identified whole circular tRNAs, as we did for *H. salinarum* (Figure 6, Figure S14). These circRNAs also contain some bases in their circularization junctions that do not align in the reference genome, forming the sequence "CCA".

We also retrieved some previously annotated circRNAs in these archaea. We identified circular C/D box RNAs in the three thermophile species, with this type of circRNA being the most abundant in *P. abyssi* (Figure 1C, Table S3). We also identified circRNAs associated with the signal recognition particle (SRP) in *H. volcanii* and *S. solfataricus*, and with RNase P in *S. solfataricus* (Figure 6, Table S3), as was identified in *S. solfataricus* and *S. acidocaldarius* in the original study (Danan et al., 2012).

We found circRNAs associated with IS200/IS605 in *S. solfataricus* and *S. acidocaldarius* (Figure 6, Table S3). In *S. acidocaldarius*, the circRNAs are associated with the 3' of the transposase gene, while in *S. solfataricus*, they are in the 5' end. In

both species, the RNA-Seq coverage in the transposase is consistent with the presence of an ω RNA (Figure S14). *S. solfataricus*, besides circRNAs in two IS200/IS605 transposases, has circRNAs associated with three IS not classified yet (ISNCY) (Table S3).

Motivated by the finding of circular ω RNAs in other archaea and that circRNAs appear to be conserved among different species, we expanded our search for circRNAs in representative species of major archaea groups (Figure 6). Using regular RNA-Seq data, we searched for the same types of circRNAs we identified in RNase R-treated data to avoid false positives, and used 10 reads supporting the circularization junction as a cutoff to consider a circRNA significant.

The most consistent classes of circRNAs identified in all datasets are circRNAs in rRNA and tRNA (especially tRNA-Trp intron), indicating circularization is an important or at least a conserved part of processing these transcripts (Figure 6). Indeed, in the circRNA-Seq datasets, these circRNAs were the most abundant in most organisms studied (Figure 1C). We found whole circular tRNAs only in haloarchaea species (*H. salinarum*, *H. volcanii*, and *Haloferax mediterranei*), indicating this could be haloarchaea-specific processing (Figure 6). We found circular C/D box RNAs (not considering tRNA introns) only in thermophile species (Figure 6), which supports the hypothesis that the circularization of these transcripts is important at high temperatures (Starostina et al., 2004).

We found circRNAs associated with IS200/IS605 in *H. mediterranei*, *H. volcanii*, and *Haloarubrum lacusprofundi*

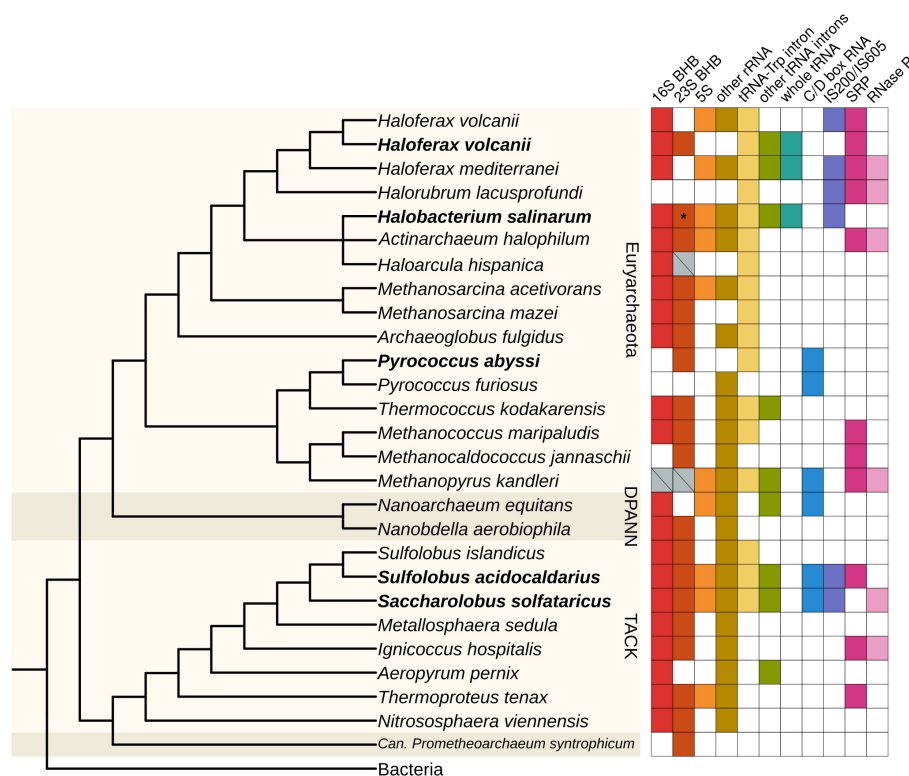


FIGURE 6

Distribution of circRNAs across archaeal species. The presence of circRNAs in representatives of major archaeal groups was identified by searching for the indicated classes of circRNAs in RNA-Seq data. Organisms with circRNA-Seq data are highlighted in bold. Gray crossed boxes in the “16S BHB” and “23S BHB” categories indicate we could find circRNAs encompassing the whole rRNA gene, being putative circRNAs associated with a BHB motif, but we could not identify or predict the exact BHB structure. *H. volcanii* is duplicated because we used two different datasets to identify circRNAs for this analysis (see Table S1). The asterisk indicates that *H. salinarum* circRNA associated with the 23S BHB does not have 10 reads or more to support it, but it was validated with RT-PCR. The tree was constructed based on the phylogeny from (Liu et al., 2021), with missing organisms added as a polytomy; the Bacteria group were used as an outgroup. BHB, bulge-helix-bulge; SRP, signal recognition particle.

(Figure 6). These three species are haloarchaea belonging to the Haloferacales order. In the RFAM database, the RFAM families RF02656 and RF02657 (still bearing the original sotrRNA terminology) from *H. salinarum* ω RNAs were expanded to other Halobacteria, indicating that ω RNAs, and probably their circular isoforms, might be abundant in these groups.

4 Discussion

4.1 MonArch consistently identifies archaeal circRNAs

MonArch, our computational pipeline for finding circRNAs in RNA-Seq data, showed good results in recovering known archaeal circRNAs in all circRNA-Seq datasets and in our reanalysis compared with published results (Figures 1B, C). MonArch assumes that reads from the circularization junction do not align regularly in the genome but in a chiastic manner. Even though simple, this approach has proven effective in searching for prokaryotic circRNAs.

MonArch does not need major data pre-processing before using it to annotate circRNAs. Other eukaryotic-focused tools

need the RNA-Seq data to be aligned by specific tools before the identification of the circRNAs. MonArch can handle raw RNA-Seq reads as well as reads that have been processed in some manner, as long as they are in FASTA format. In this study, we chose to first trim adapters from the RNA-seq reads since it helped MonArch identify circRNAs downstream. We also used unaligned reads from an RNA-Seq read aligner in our tests. This could save some time in processing large datasets, but the filters in MonArch itself can identify circRNAs with confidence (Figure S1A), making it not mandatory. This approach resulted in the same circRNAs reported here but with fewer reads supporting each one (results not shown).

MonArch also does not need any other information on the organism studied besides its reference genome. Published tools often need genome annotation or splicing annotation for the organism. This seems to be useful in the precise annotation of high-confidence circRNAs in eukaryotes (Vromman et al., 2023), but it can be a hindrance to use in prokaryotes that do not have splicing events or well-annotated genomes. However, even tools that support *de novo* circRNA annotation cannot identify prokaryotic circRNAs adequately. We have tested CIRI2 (Gao et al., 2015, 2018) and circRNA_finder (Westholm et al., 2014) with default parameters in the *H. salinarum* circRNA-Seq dataset. CIRI2 only identified one circRNA, while circRNA_finder identified none.

Another study also tested CIRI2 in their *S. solfataricus* RNA-Seq data, and it could not retrieve all the circRNAs they validated (Bathke et al., 2020).

However, it is important to stress that MonArch was made to analyze prokaryotic RNA-Seq data. It was never tested or optimized for larger eukaryotic datasets and genomes. Besides, it only searches for the circularization junction signature and will not necessarily benefit from paired-end sequencing information when reads come from different sides of the junction but do not contain it. Finally, given its simplicity and standard tools, it could be a prohibitively slow approach for facility-level workloads.

4.2 Biogenesis of circular ω RNAs

In archaea, it is well-known that circRNAs can be generated by processing the RNA at the bulge-helix-bulge (BHB) RNA motif by EndA, an endonuclease, followed by ligation of the ends (Clouet-d'Orval et al., 2018; Qi et al., 2020; Schwarz et al., 2020; Grünberger et al., 2023). The ligation reaction could be performed by RtcB or other RNA ligases since Pab1020, a ligase from the Rnl3 family, has been found to circularize RNAs in *P. abyssi* (Becker et al., 2017). However, we did not find the BHB motif associated with all *H. salinarum* circRNAs, in particular the circular ω RNAs (Figures 4C, D). This suggests these circRNAs may have a different biogenesis pathway independent of the BHB motif and EndA endonuclease, as was proposed before for other archaeal circRNAs (Danan et al., 2012; Becker et al., 2017). This is not surprising, since all domains of life seem to circularize their RNAs by different pathways, and even RNAs of one domain have different means to do so (with the pairing of flanking Alu repeats or binding of RNA-binding proteins in Eukarya (Kristensen et al., 2019), and self-splicing introns (Hausner et al., 2014; Roth et al., 2021) or the independent mechanism of DncS circularization in Bacteria (He et al., 2023), for example).

TnpBs can process their own ω RNAs at the 5' end, according to a survey of 59 TnpB orthologs with an *in vitro* transcription and translation system (Nety et al., 2023). Ibrahim and colleagues also observed transcriptional processing sites (TPS) at IS200/IS605 of different prokaryotic organisms, further suggesting that processing of the ω RNA is a widespread phenomenon (Ibrahim et al., 2021).

Gomes-Filho and colleagues' Northern-Blot experiments suggest *H. salinarum* ω RNAs/sotRNAs are processed from the primary transcript (Gomes-Filho et al., 2015). Also, its ω RNAs are enriched for TPS, especially at their 5' end (Ibrahim et al., 2021; Lorenzetti et al., 2023). These suggest the processing of the ω RNA from the *tnpB* RNA in *H. salinarum*. Given that all of the *H. salinarum* circular ω RNAs we studied here (identified in the RNase R-treated and the growth curve datasets) start at the 5' end of their cognate linear ω RNA, it is reasonable to assume TnpB could process the transcript for circularization. However, it remains unclear if TnpB or other RNA nuclease processes the RNA at the 3' end. In the case of the circularization of the whole or majority of ω RNA (as circRNA_0012, Figures 4A, C), it could be the case that circularization occurs between the newly processed 5' end of the ω RNA and the 3' of the transcript. It is still unknown

which RNA ligase could ligate the circRNA 5' and 3' ends for these transcripts.

4.3 Possible functions of circular ω RNAs

We found novel circRNAs associated with IS200/IS605 and their ω RNAs in *H. salinarum* and other halophilic archaea, as well as in *S. solfataricus* and *S. acidocaldarius*. The interest in IS200/IS605 transposases and their ω RNA has been increasing in the last few years due to their evolutionary relationship to Cas9 and Cas12 (Kapitonov et al., 2015; Shmakov et al., 2017; Altae-Tran et al., 2021) and the promise of a new and more compact gene-editing tool (Li et al., 2024; Xiang et al., 2024). The circRNAs associated with the ω RNAs are a new piece in this puzzle with unexplored functions and biogenesis pathways.

The growth-dependent expression of the circular ω RNAs might suggest some function or regulation in the cell. Bacterial circular DncS regulates the level of its linear counterpart: in later stages of growth, the circular isoform has increased expression, while the level of the linear isoform decays (He et al., 2023). *H. salinarum* sot0013 and sot0044 ω RNAs have an expression pattern that could suggest this kind of regulation (Figure 5), but more orthogonal approaches are necessary to check this hypothesis. *H. salinarum* *tnpBs* and their ω RNAs have inverse expression profiles (Gomes-Filho et al., 2015), and its TnpB proteins are either detected at very low levels or not detected at all by mass spectrometry (Lorenzetti et al., 2023).

H. salinarum mobilome and transposition are post-transcriptionally regulated by different mechanisms (Lorenzetti et al., 2023). We searched for *H. salinarum* circular ω RNAs in different RNA-Seq datasets besides the growth curve shown in Figure 5 (low salinity, different genetic backgrounds), but we either (i) did not find evidence for circular ω RNAs or (ii) could not see an expression profile for the total reads that was different for the circular ones in these datasets (data not shown). This could be due to the lack of RNase R enrichment and consequent inability to consistently detect circRNAs and to the actual absence of circRNA regulation in the conditions investigated.

Transposition and transposases are usually kept at low levels in the cell to avoid deleterious effects on the host (Ellis and Haniford, 2016). This is achieved by a myriad of mechanisms at different information levels (Nagy and Chandler, 2004). Specifically in the IS200/IS605 family, both the transposition by TnpA and the DNA cleavage by TnpB seem to be post-transcriptionally regulated (Ellis et al., 2015; Nety et al., 2023). Post-transcriptional regulation fine-tunes gene expression in a condition-dependent manner allowing rapid adaptation to stress (Martínez and Vadyvaloo, 2014; Papenfort and Melamed, 2023) through RNA processing, stability, and turnover (Shine et al., 2024). It has been proposed that circRNAs could also be part of the RNA degradation process (Danan et al., 2012), and IS200/IS605 circRNAs (one of which we reidentified here—circRNA_7379 in Table S3) interact with exosome proteins in *S. solfataricus* (Bathke et al., 2020).

circRNAs have the exact same sequence as their linear counterpart (except for the circularization junction), and yet they can have different conformations and structures (Chen, 2016),

which could lead to different functions. The prediction of the circular or linear ω RNA in Vienna RNAfold (Gruber et al., 2008) did not significantly change the structure shown in Figures 4C, D. Since Cryo-EM experiments have shown that 2D structure predictors could not accurately reproduce the experimentally found structure for the ω RNA (Nakagawa et al., 2023; Sasnauskas et al., 2023), it is reasonable to assume that this does not necessarily mean that the circular and linear ω RNA have the same structure.

In eukaryotes, some circRNAs form short imperfect duplexes absent in their linear counterparts that can bind to proteins and modulate their activity. For example, circRNAs can regulate the innate immune response by binding to the dsRNA-binding PKR (protein kinase R) (Liu et al., 2019) or the DNA-binding protein cGAS (cyclic GMP-AMP synthase) to avoid its self-DNA activity (Xia et al., 2018). If circular and linear ω RNAs have different structures, the circRNA could have a similar role with TnpB, serving as a decoy to avoid the protein activity when unnecessary. The interaction with exosome proteins in *S. solfataricus* may suggest a role in degradation or degradation signaling for these circRNAs; in eukaryotes, circRNA conformation and interaction partners are signals to modulate its turnover (Liu et al., 2024).

Yet, the actual functions and structural conformation of circular ω RNAs remain to be properly elucidated. The first steps in this endeavor could be to determine the interaction partners of both the circular and linear ω RNA, and eventually experimentally determine their structures. Classical genetic and functional approaches are also important, but careful planning and controls are necessary to address the challenge of overlapping transcripts (in this case, the *tnpB* mRNA and the linear and circular forms of the ω RNA). Studies on the circularization mechanisms of the ω RNA would also be of great help for informing the genetic manipulations in these loci.

4.4 Conservation of circRNAs among archaea

Using circRNA-Seq data, we found circRNAs associated with the same classes of transcripts in all organisms studied (rRNA, tRNA, IS, C/D box RNA, and other small RNAs) (Figure 1C). This is evidence that circRNAs are conserved among this group of organisms since they are associated with the same transcripts. More than that, in our expanded search for circRNAs in archaea representatives, we could find at least one circRNA in each of them, even without RNase R enrichment (Figure 6). Our search identified the novel circular ω RNAs in different phylogenetically distant species, indicating this could be a widespread or conserved phenomenon.

It is important to point out that our results on the presence of circRNAs do not necessarily result only from phylogenetic relationships. Even though we searched only for circRNAs annotated in circRNA-Seq data in regular RNA-Seq datasets to avoid false positives, the latter were not treated with RNase R and probably did not consider circRNA annotation when they were made. This leads to a great variability in results and what can be detected according to how the data was generated. For example, the sequencing of small RNAs may prevent the identification of

circular pre-16S and pre-23S, very established and known circular intermediates in archaea. The opposite may also be true, with total RNA sequencing hindering smaller circRNA detection. The identification of circular ω RNAs in one *H. volcanii* dataset and not the other (total RNA with no rRNA depletion, treated with RNase R vs. RNA-Seq of small RNAs) (Figure 6) exemplifies this limitation. The sequencing depth may also influence whether less abundant circRNAs are detected, especially without RNase R treatment.

Considering this, we do not believe that our results represent a final landscape of circRNA presence in archaea. If a circRNA is absent in one of the organisms in this analysis, it does not necessarily mean it does not exist. Our analysis is a first effort, using already published data, to detect archaeal circRNAs in a more comprehensive and systematic way. We believe that many more circRNAs are to be identified in Archaea (even more examples of circular ω RNAs); we show here that circRNAs are present in all groups investigated, making future circRNA searches promising.

4.5 Concluding remarks

Here, we have identified and described in detail *H. salinarum* circRNAs using MonArch, a computational pipeline we developed for circRNA identification in archaeal RNA-Seq data. We could find known circRNAs as rRNA and tRNA processing intermediates, as well as new circRNAs associated with IS200/IS605 and their non-coding RNAs, ω RNAs. We also showed that these circRNAs are expressed in a growth-dependent manner in *H. salinarum*, being one of the few examples of circRNA conditional expression in prokaryotes. These circRNAs could be interesting new pieces in the TnpB/ ω RNA system.

With our extensive reanalysis of RNA-Seq data, we showed that the same classes of transcripts are circularized in archaea. circRNAs seem to be prevalent and conserved in this group of organisms, maybe more than previously appreciated. The circular ω RNAs also seem to be conserved, being present in haloarchaea and two Sulfolobales species, two phylogenetically distant archaeal groups.

With this study, we contributed to expanding the yet scarce knowledge of prokaryotic circRNAs. We hope our results incentivize the search and characterization of circRNAs in other archaea to gain a better understanding of these molecules in the third domain of life and in prokaryotes as a whole.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary material.

Author contributions

BP: Validation, Conceptualization, Data curation, Writing – review & editing, Methodology, Writing – original draft,

Investigation, Visualization, Formal analysis, Software. VF-S: Writing – review & editing, Visualization, Software, Formal analysis, Investigation, Writing – original draft, Data curation, Methodology. LZ: Visualization, Formal analysis, Investigation, Methodology, Writing – review & editing, Conceptualization. RV: Software, Writing – original draft, Investigation, Resources, Writing – review & editing, Formal analysis, Methodology, Data curation, Visualization, Supervision, Conceptualization. TK: Resources, Visualization, Validation, Formal analysis, Funding acquisition, Project administration, Writing – original draft, Supervision, Data curation, Investigation, Writing – review & editing, Methodology, Conceptualization.

Funding

This study was supported by São Paulo Research Foundation (FAPESP) grant 2009/09532-0. BP was supported by FAPESP (2022/00308-4 and 2024/13402-4) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brazil (CAPES - Finance Code 001) fellowships. VF-S was supported by FAPESP fellowship 2018/25329-9. LZ was supported by FAPESP fellowship 2011/07487-7.

Acknowledgments

This research was developed with HPC resources provided by the Information Technology Superintendence (HPC-STI) of the University of São Paulo. Sequencing was performed at Hemocentro USP-RP facility. We thank Silvia Helena Epifânio for technical assistance.

References

- Altae-Tran, H., Kannan, S., Demircioglu, F. E., Oshiro, R., Nety, S. P., McKay, L. J., et al. (2021). The widespread IS200/IS605 transposon family encodes diverse programmable RNA-guided endonucleases. *Science* 374, 57–65. doi: 10.1126/science.abj6856
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2
- Ashwal-Fluss, R., Meyer, M., Pamudurti, N. R., Ivanov, A., Bartok, O., Hanan, M., et al. (2014). circRNA biogenesis competes with pre-mRNA splicing. *Mol. Cell* 56, 55–66. doi: 10.1016/j.molcel.2014.08.019
- Bach, D.-H., Lee, S. K., and Sood, A. K. (2019). Circular RNAs in cancer. *Mol. Ther. Nucleic Acids* 16, 118–129. doi: 10.1016/j.omtn.2019.02.005
- Bathke, J., Gauernack, A. S., Rupp, O., Weber, L., Preusser, C., Lechner, M., et al. (2020). iCLIP analysis of RNA substrates of the archaeal exosome. *BMC Genomics* 21:797. doi: 10.1186/s12864-020-07200-x
- Becker, H. F., Héliou, A., Djaout, K., Lestini, R., Regnier, M., and Myllykallio, H. (2017). High-throughput sequencing reveals circular substrates for an archaeal RNA ligase. *RNA Biol.* 14, 1075–1085. doi: 10.1080/15476286.2017.1302640
- Bernier, C. R., Petrov, A. S., Waterbury, C. C., Jett, J., Li, F., Freil, L. E., et al. (2014). RiboVision suite for visualization and analysis of ribosomes. *Faraday Discuss.* 169, 195–207. doi: 10.1039/C3FD00126A
- Birkedal, U., Beckert, B., Wilson, D. N., and Nielsen, H. (2020). The 23S ribosomal RNA from is circularly permuted. *Front. Microbiol.* 11:582022. doi: 10.3389/fmicb.2020.582022
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170
- Brooks, A. N., Reiss, D. J., Allard, A., Wu, W.-J., Salvanha, D. M., Plaisier, C. L., et al. (2014). A system-level model for the microbial regulatory genome. *Mol. Syst. Biol.* 10:740. doi: 10.15252/msb.20145160
- Cai, H., Li, Y., Niringiyumukiza, J. D., Su, P., and Xiang, W. (2019). Circular RNA involvement in aging: an emerging player with great potential. *Mech. Ageing Dev.* 178, 16–24. doi: 10.1016/j.mad.2018.11.002
- Chant, J., and Dennis, P. (1986). Archaeobacteria: transcription and processing of ribosomal RNA sequences in *Halobacterium cutirubrum*. *EMBO J.* 5, 1091–1097. doi: 10.1002/j.1460-2075.1986.tb04327.x
- Chen, L., Wang, C., Sun, H., Wang, J., Liang, Y., Wang, Y., et al. (2021). The bioinformatics toolbox for circRNA discovery and analysis. *Brief Bioinform.* 22, 1706–1728. doi: 10.1093/bib/bbaa001
- Chen, L.-L. (2016). The biogenesis and emerging roles of circular RNAs. *Nat. Rev. Mol. Cell Biol.* 17, 205–211. doi: 10.1038/nrm.2015.32
- Chen, Y.-J., Chen, C.-Y., Mai, T.-L., Chuang, C.-F., Chen, Y.-C., Gupta, S. K., et al. (2020). Genome-wide, integrative analysis of circular RNA dysregulation and the corresponding circular RNA-microRNA-mRNA regulatory axes in autism. *Genome Res.* 30, 375–391. doi: 10.1101/gr.255463.119
- Clouet d'Orval, B., Bortolin, M. L., Gaspin, C., and Bachellerie, J. P. (2001). Box C/D RNA guides for the ribose methylation of archaeal tRNAs. The tRNA^{Trp} intron guides the formation of two ribose-methylated nucleosides in the mature tRNA^{Trp}. *Nucleic Acids Res.* 29, 4518–4529. doi: 10.1093/nar/29.22.4518

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2025.1641342/full#supplementary-material>

- Clouet-d'Orval, B., Batista, M., Bouvier, M., Quentin, Y., Fichant, G., Marchfelder, A., et al. (2018). Insights into RNA-processing pathways and associated RNA-degrading enzymes in Archaea. *FEMS Microbiol. Rev.* 42, 579–613. doi: 10.1093/femsre/fuy016
- Cocquerelle, C., Daubersies, P., Majérus, M. A., Kerckaert, J. P., and Bailleul, B. (1992). Splicing with inverted order of exons occurs proximal to large introns. *EMBO J.* 11, 1095–1098. doi: 10.1002/j.1460-2075.1992.tb05148.x
- Cocquerelle, C., Mascréz, B., Hétiuin, D., and Bailleul, B. (1993). Mis-splicing yields circular RNA molecules. *FASEB J.* 7, 155–160. doi: 10.1096/fasebj.7.1.7678559
- Danan, M., Schwartz, S., Edelheit, S., and Sorek, R. (2012). Transcriptome-wide discovery of circular RNAs in Archaea. *Nucleic Acids Res.* 40, 3131–3142. doi: 10.1093/nar/gkr1009
- DasSarma, S. (1993). Identification and analysis of the gas vesicle gene cluster on an unstable plasmid of *Halobacterium halobium*. *Experientia* 49, 482–486. doi: 10.1007/BF01955149
- de Almeida, J. P. P., Vêncio, R. Z. N., Lorenzetti, A. P. R., Caten, F. T., Gomes-Filho, J. V., and Koide, T. (2019). The primary antisense transcriptome of NRC-1. *Genes* 10:280. doi: 10.3390/genes10040280
- Dodbele, S., Mutlu, N., and Wilusz, J. E. (2021). Best practices to ensure robust investigation of circular RNAs: pitfalls and tips. *EMBO Rep.* 22:e52072. doi: 10.15252/embr.202052072
- Du, W. W., Zhang, C., Yang, W., Yong, T., Awan, F. M., and Yang, B. B. (2017). Identifying and characterizing circRNA-protein interaction. *Theranostics* 7, 4183–4191. doi: 10.7150/thno.21299
- Dulmage, K. A., Darnell, C. L., Vreugdenhil, A., and Schmid, A. K. (2018). Copy number variation is associated with gene expression change in archaea. *Microb. Genom.* 4:e000210. doi: 10.1099/mgen.0.000210
- Dyall-Smith, M. (2009). *The Haloarchaeal Handbook: Protocols for Haloarchaeal Genetics (Version 7.2)*. Available online at: <http://www.haloarchaea.com/resources/haloarchaeal/> (Accessed June 1, 2025).
- Ellis, M. J., and Haniford, D. B. (2016). Riboregulation of bacterial and archaeal transposition. *Wiley Interdiscip. Rev. RNA* 7, 382–398. doi: 10.1002/wrna.1341
- Ellis, M. J., Trussler, R. S., and Haniford, D. B. (2015). A cis-encoded sRNA, Hfq and mRNA secondary structure act independently to suppress IS200 transposition. *Nucleic Acids Res.* 43, 6511–6527. doi: 10.1093/nar/gkv584
- Eme, L., Spang, A., Lombard, J., Stairs, C. W., and Ettema, T. J. G. (2018). Archaea and the origin of eukaryotes. *Nat. Rev. Microbiol.* 16:120. doi: 10.1038/nrmicro.2017.154
- Gao, Y., Wang, J., and Zhao, F. (2015). CIRI: an efficient and unbiased algorithm for de novo circular RNA identification. *Genome Biol.* 16:4. doi: 10.1186/s13059-014-0571-3
- Gao, Y., Wang, J., Zheng, Y., Zhang, J., Chen, S., and Zhao, F. (2016). Comprehensive identification of internal structure and alternative splicing events in circular RNAs. *Nat. Commun.* 7:12060. doi: 10.1038/ncomms12060
- Gao, Y., Zhang, J., and Zhao, F. (2018). Circular RNA identification based on multiple seed matching. *Brief Bioinform.* 19, 803–810. doi: 10.1093/bib/bbx014
- Gomes-Filho, J. V., Zaramela, L. S., Italiani, V. C. da S., Baliga, N. S., Vêncio, R. Z. N., and Koide, T. (2015). Sense overlapping transcripts in IS1341-type transposase genes are functional non-coding RNAs in archaea. *RNA Biol.* 12, 490–500. doi: 10.1080/15476286.2015.1019998
- Gruber, A. R., Lorenz, R., Bernhart, S. H., Neuböck, R., and Hofacker, I. L. (2008). The Vienna RNA websuite. *Nucleic Acids Res.* 36, W70–W74. doi: 10.1093/nar/gkn188
- Grünberger, F., Jüttner, M., Knüppel, R., Ferreira-Cerca, S., and Grohmann, D. (2023). Nanopore-based RNA sequencing deciphers the formation, processing, and modification steps of rRNA intermediates in archaea. *RNA* 29, 1255–1273. doi: 10.1261/rna.079636.123
- Hansen, T. B., Jensen, T. I., Clausen, B. H., Bramsen, J. B., Finsen, B., Damgaard, C. K., et al. (2013). Natural RNA circles function as efficient microRNA sponges. *Nature* 495, 384–388. doi: 10.1038/nature11993
- Hansen, T. B., Venø, M. T., Damgaard, C. K., and Kjems, J. (2016). Comparison of circular RNA prediction tools. *Nucleic Acids Res.* 44:e58. doi: 10.1093/nar/gkv1458
- Hausner, G., Hafez, M., and Edgell, D. R. (2014). Bacterial group I introns: mobile RNA catalysts. *Mobile DNA* 5, 1–12. doi: 10.1186/1759-8753-5-8
- He, S., Corneloup, A., Guynet, C., Lavatine, L., Caumont-Sarcos, A., Siguier, P., et al. (2015). The IS200/IS605 family and “peel and paste” single-strand transposition mechanism. *Microbiol. Spectr.* 3. doi: 10.1128/microbiolspec.MDNA3-0039-2014
- He, T.-T., Xu, Y.-F., Li, X., Wang, X., Li, J.-Y., Ou-Yang, D., et al. (2023). A linear and circular dual-conformation noncoding RNA involved in oxidative stress tolerance in *Bacillus altitudinis*. *Nat. Commun.* 14:5722. doi: 10.1038/s41467-023-41491-4
- Hsu, M. T., and Coca-Prados, M. (1979). Electron microscopic evidence for the circular form of RNA in the cytoplasm of eukaryotic cells. *Nature* 280, 339–340. doi: 10.1038/280339a0
- Ibrahim, A. G. A. E.-R., Vêncio, R. Z. N., Lorenzetti, A. P. R., and Koide, T. (2021). *Halobacterium salinarum* and *Haloferax volcanii* comparative transcriptomics reveals conserved transcriptional processing sites. *Genes* 12:1018. doi: 10.3390/genes12071018
- Jeck, W. R., and Sharpless, N. E. (2014). Detecting and characterizing circular RNAs. *Nat. Biotechnol.* 32, 453–461. doi: 10.1038/nbt.2890
- Jeck, W. R., Sorrentino, J. A., Wang, K., Slevin, M. K., Burd, C. E., Liu, J., et al. (2013). Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA* 19, 141–157. doi: 10.1261/rna.035667.112
- Johnson, P. Z., and Simon, A. E. (2023). RNAcanvas: interactive drawing and exploration of nucleic acid structures. *Nucleic Acids Res.* 51, W501–W508. doi: 10.1093/nar/gkad302
- Kapitonov, V. V., Makarova, K. S., and Koonin, E. V. (2015). ISC, a novel group of bacterial and archaeal DNA transposons that encode Cas9 homologs. *J. Bacteriol.* 198, 797–807. doi: 10.1128/JB.00783-15
- Karvelis, T., Druteika, G., Bigelyte, G., Budre, K., Zedaveinyte, R., Silanskas, A., et al. (2021). Transposon-associated TnpB is a programmable RNA-guided DNA endonuclease. *Nature* 599, 692–696. doi: 10.1038/s41586-021-04058-1
- Kichenaradja, P., Siguier, P., Pérochon, J., and Chandler, M. (2010). ISbrowser: an extension of ISfinder for visualizing insertion sequences in prokaryotic genomes. *Nucleic Acids Res.* 38, D62–D68. doi: 10.1093/nar/gkp947
- Kjems, J., and Garrett, R. A. (1988). Novel splicing mechanism for the ribosomal RNA intron in the archaeobacterium *Desulfurococcus mobilis*. *Cell* 54, 693–703. doi: 10.1016/S0092-8674(88)80014-X
- Kos, A., Dijkema, R., Arnberg, A. C., van der Meide, P. H., and Schellekens, H. (1986). The hepatitis delta (delta) virus possesses a circular RNA. *Nature* 323, 558–560. doi: 10.1038/323558a0
- Kristensen, L. S., Andersen, M. S., Stagsted, L. V. W., Ebbesen, K. K., Hansen, T. B., and Kjems, J. (2019). The biogenesis, biology and characterization of circular RNAs. *Nat. Rev. Genet.* 20, 675–691. doi: 10.1038/s41576-019-0158-7
- Lasda, E., and Parker, R. (2014). Circular RNAs: diversity of form and function. *RNA* 20, 1829–1842. doi: 10.1261/rna.047126.114
- Legnini, I., Di Timoteo, G., Rossi, F., Morlando, M., Briganti, F., Sthandier, O., et al. (2017). Circ-ZNF609 is a circular RNA that can be translated and functions in myogenesis. *Mol. Cell* 66, 22–37.e9. doi: 10.1016/j.molcel.2017.02.017
- Li, Z., Guo, R., Sun, X., Li, G., Shao, Z., Huo, X., et al. (2024). Engineering a transposon-associated TnpB-ωRNA system for efficient gene editing and phenotypic correction of a tyrosinaemia mouse model. *Nat. Commun.* 15:831. doi: 10.1038/s41467-024-45197-z
- Liao, Y., Smyth, G. K., and Shi, W. (2019). The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res.* 47, e47. doi: 10.1093/nar/gkz114
- Liu, C.-X., and Chen, L.-L. (2022). Circular RNAs: characterization, cellular roles, and applications. *Cell* 185, 2016–2034. doi: 10.1016/j.cell.2022.04.021
- Liu, C.-X., Li, X., Nan, F., Jiang, S., Gao, X., Guo, S.-K., et al. (2019). Structure and degradation of circular RNAs regulate PKR activation in innate immunity. *Cell* 177, 865–880.e21. doi: 10.1016/j.cell.2019.03.046
- Liu, C.-X., Yang, L., and Chen, L.-L. (2024). Dynamic conformation: marching toward circular RNA function and application. *Mol. Cell* 84, 3596–3609. doi: 10.1016/j.molcel.2024.08.020
- Liu, Y., Makarova, K. S., Huang, W.-C., Wolf, Y. I., Nikolskaya, A. N., Zhang, X., et al. (2021). Expanded diversity of Asgard archaea and their relationships with eukaryotes. *Nature* 593, 553–557. doi: 10.1038/s41586-021-03494-3
- López García de Lomana, A., Kusebauch, U., Raman, A. V., Pan, M., Turkarslan, S., Lorenzetti, A. P. R., et al. (2020). Selective translation of low abundance and upregulated transcripts in *Halobacterium salinarum*. *mSystems* 5, e00329–20. doi: 10.1128/mSystems.00329-20
- Lorenzetti, A. P. R., Kusebauch, U., Zaramela, L. S., Wu, W.-J., de Almeida, J. P. P., Turkarslan, S., et al. (2023). A genome-scale atlas reveals complex interplay of transcription and translation in an Archaeon. *mSystems* 8:e0081622. doi: 10.1128/mSystems.00816-22
- Lykke-Andersen, J., and Garrett, R. A. (1994). Structural characteristics of the stable RNA introns of archaeal hyperthermophiles and their splicing junctions. *J. Mol. Biol.* 243, 846–855. doi: 10.1006/jmbi.1994.1687
- Martínez, L. C., and Vadyvaloo, V. (2014). Mechanisms of post-transcriptional gene regulation in bacterial biofilms. *Front. Cell Infect. Microbiol.* 4:38. doi: 10.3389/fcimb.2014.00038
- Memczak, S., Jens, M., Elefsinioti, A., Torti, F., Krueger, J., Rybak, A., et al. (2013). Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* 495, 333–338. doi: 10.1038/nature11928
- Nagy, Z., and Chandler, M. (2004). Regulation of transposition in bacteria. *Res. Microbiol.* 155, 387–398. doi: 10.1016/j.resmic.2004.01.008

- Nahand, J. S., Jamshidi, S., Hamblin, M. R., Mahjoubin-Tehran, M., Vosough, M., Jamali, M., et al. (2020). Circular RNAs: new epigenetic signatures in viral infections. *Front. Microbiol.* 11:1853. doi: 10.3389/fmicb.2020.01853
- Nakagawa, R., Hirano, H., Omura, S. N., Nety, S., Kannan, S., Altae-Tran, H., et al. (2023). Cryo-EM structure of the transposon-associated TnpB enzyme. *Nature* 616, 390–397. doi: 10.1038/s41586-023-05933-9
- Nety, S. P., Altae-Tran, H., Kannan, S., Demircioglu, F. E., Faure, G., Hirano, S., et al. (2023). The transposon-encoded protein TnpB processes its own mRNA into ω RNA for guided nuclease activity. *CRISPR J.* 6, 232–242. doi: 10.1089/crispr.2023.0015
- Nigro, J. M., Cho, K. R., Fearon, E. R., Kern, S. E., Ruppert, J. M., Oliner, J. D., et al. (1991). Scrambled exons. *Cell* 64, 607–613. doi: 10.1016/0092-8674(91)90244-S
- Omer, A. D., Lowe, T. M., Russell, A. G., Ebhardt, H., Eddy, S. R., and Dennis, P. P. (2000). Homologs of small nucleolar RNAs in Archaea. *Science* 288, 517–522. doi: 10.1126/science.288.5465.517
- Orell, A., Tripp, V., Aliaga-Tobar, V., Albers, S.-V., Maracaja-Coutinho, V., and Randau, L. (2018). A regulatory RNA is involved in RNA duplex formation and biofilm regulation in *Sulfolobus acidocaldarius*. *Nucleic Acids Res.* 46, 4794–4806. doi: 10.1093/nar/gky144
- Papenfert, K., and Melamed, S. (2023). Small RNAs, large networks: posttranscriptional regulons in gram-negative bacteria. *Annu. Rev. Microbiol.* 77, 23–43. doi: 10.1146/annurev-micro-041320-025836
- Patop, I. L., Wüst, S., and Kadener, S. (2019). Past, present, and future of circRNAs. *EMBO J.* 38:e100836. doi: 10.15252/embj.2018100836
- Pfeiffer, F., Broicher, A., Gillich, T., Klee, K., Mejia, J., Rampp, M., et al. (2008). Genome information management and integrated data analysis with HaloLex. *Arch. Microbiol.* 190, 281–299. doi: 10.1007/s00203-008-0389-z
- Pfeiffer, F., Marchfelder, A., Habermann, B., and Dyll-Smith, M. L. (2019). The genome sequence of the halobacterium salinarum type strain is closely related to that of laboratory strains NRC-1 and R1. *Microbiol. Resour. Announc.* 8, e00429–19. doi: 10.1128/MRA.00429-19
- Qi, L., Li, J., Jia, J., Yue, L., and Dong, X. (2020). Comprehensive analysis of the pre-ribosomal RNA maturation pathway in a methanococcus exposes the conserved circularization and linearization mode in archaea. *RNA Biol.* 17, 1427–1441. doi: 10.1080/15476286.2020.1771946
- Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. doi: 10.1093/bioinformatics/btq033
- Raden, M., Ali, S. M., Alkhnbashi, O. S., Busch, A., Costa, F., Davis, J. A., et al. (2018). Freiburg RNA tools: a central online resource for RNA-focused research and teaching. *Nucleic Acids Res.* 46, W25–W29. doi: 10.1093/nar/gky329
- Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., et al. (2011). Integrative genomics viewer. *Nat. Biotechnol.* 29:24. doi: 10.1038/nbt.1754
- Roth, A., Weinberg, Z., Vanderschuren, K., Murdock, M. H., and Breaker, R. R. (2021). Natural circularly permuted group II introns in bacteria produce RNA circles. *iScience* 24:103431. doi: 10.1016/j.isci.2021.103431
- Rybak-Wolf, A., Stottmeister, C., Glažar, P., Jens, M., Pino, N., Giusti, S., et al. (2015). Circular RNAs in the mammalian brain are highly abundant, conserved, and dynamically expressed. *Mol. Cell* 58, 870–885. doi: 10.1016/j.molcel.2015.03.027
- Salzman, J., Gawad, C., Wang, P. L., Lacayo, N., and Brown, P. O. (2012). Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. *PLoS One* 7:e30733. doi: 10.1371/journal.pone.0030733
- Sanger, H. L., Klotz, G., Riesner, D., Gross, H. J., and Kleinschmidt, A. K. (1976). Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures. *Proc. Natl. Acad. Sci. U. S. A.* 73, 3852–3856. doi: 10.1073/pnas.73.11.3852
- Sasnauskas, G., Tamulaitiene, G., Druteika, G., Carabias, A., Silanskas, A., Kazlauskas, D., et al. (2023). TnpB structure reveals minimal functional core of Cas12 nuclease family. *Nature* 616, 384–389. doi: 10.1038/s41586-023-05826-x
- Sato, K., Kato, Y., Hamada, M., Akutsu, T., and Asai, K. (2011). IPknot: fast and accurate prediction of RNA secondary structures with pseudoknots using integer programming. *Bioinformatics* 27, i85–93. doi: 10.1093/bioinformatics/btr215
- Schwarz, T. S., Berkemer, S. J., Bernhart, S. H., Weiß, M., Ferreira-Cerca, S., Stadler, P. F., et al. (2020). Splicing endonuclease is an important player in rRNA and tRNA maturation in Archaea. *Front. Microbiol.* 11:594838. doi: 10.3389/fmicb.2020.594838
- Shine, M., Gordon, J., Schärffen, L., Zigackova, D., Herzel, L., and Neugebauer, K. M. (2024). Co-transcriptional gene regulation in eukaryotes and prokaryotes. *Nat. Rev. Mol. Cell Biol.* 25, 534–554. doi: 10.1038/s41580-024-00706-2
- Shmakov, S., Smargon, A., Scott, D., Cox, D., Pyzocha, N., Yan, W., et al. (2017). Diversity and evolution of class 2 CRISPR-Cas systems. *Nat. Rev. Microbiol.* 15, 169–182. doi: 10.1038/nrmicro.2016.184
- Siguier, P., Perochon, J., Lestrade, L., Mahillon, J., and Chandler, M. (2006). ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Res.* 34, D32–D36. doi: 10.1093/nar/gkj014
- Soma, A., Onodera, A., Sugahara, J., Kanai, A., Yachie, N., Tomita, M., et al. (2007). Permuted tRNA genes expressed via a circular RNA intermediate in *Cyanidioschyzon merolae*. *Science* 318, 450–453. doi: 10.1126/science.1145718
- Starostina, N. G., Marshburn, S., Johnson, L. S., Eddy, S. R., Terns, R. M., and Terns, M. P. (2004). Circular box C/D RNAs in *Pyrococcus furiosus*. *Proc. Natl. Acad. Sci. U. S. A.* 101, 14097–14101. doi: 10.1073/pnas.0403520101
- Sweeney, B. A., Hoksza, D., Nawrocki, E. P., Ribas, C. E., Madeira, F., Cannone, J. J., et al. (2021). R2DT is a framework for predicting and visualising RNA secondary structure using templates. *Nat. Commun.* 12:3494. doi: 10.1038/s41467-021-23555-5
- Tang, T.-H., Polacek, N., Zywicki, M., Huber, H., Brugger, K., Garrett, R., et al. (2005). Identification of novel non-coding RNAs as potential antisense regulators in the archaeon *Sulfolobus solfataricus*. *Mol. Microbiol.* 55, 469–481. doi: 10.1111/j.1365-2958.2004.04428.x
- Ten-Caten, F., Vêncio, R. Z. N., Lorenzetti, A. P. R., Zaramela, L. S., Santana, A. C., and Koide, T. (2018). Internal RNAs overlapping coding sequences can drive the production of alternative proteins in archaea. *RNA Biol.* 15, 1119–1132. doi: 10.1080/15476286.2018.1509661
- Toffano-Nioche, C., Ott, A., Crozat, E., Nguyen, A. N., Zytnecki, M., Leclerc, F., et al. (2013). RNA at 92 °C: the non-coding transcriptome of the hyperthermophilic archaeon *Pyrococcus abyssi*. *RNA Biol.* 10, 1211–1220. doi: 10.4161/rna.25567
- Vromman, M., Anckaert, J., Bortoluzzi, S., Buratin, A., Chen, C.-Y., Chu, Q., et al. (2023). Large-scale benchmarking of circRNA detection tools reveals large differences in sensitivity but not in precision. *Nat. Methods* 20, 1159–1169. doi: 10.1038/s41592-023-01944-6
- Weisel, J., Wagner, S., and Klug, G. (2010). The Nop5-L7A-fibrillarin RNP complex and a novel box C/D containing sRNA of *Halobacterium salinarum* NRC-1. *Biochem. Biophys. Res. Commun.* 394, 542–547. doi: 10.1016/j.bbrc.2010.03.012
- Westholm, J. O., Miura, P., Olson, S., Shenker, S., Joseph, B., Sanfilippo, P., et al. (2014). Genome-wide analysis of drosophila circular RNAs reveals their structural and sequence properties and age-dependent neural accumulation. *Cell Rep.* 9, 1966–1980. doi: 10.1016/j.celrep.2014.10.062
- Will, S., Joshi, T., Hofacker, I. L., Stadler, P. F., and Backofen, R. (2012). LocARNA-P: accurate boundary prediction and improved detection of structural RNAs. *RNA* 18, 900–914. doi: 10.1261/rna.029041.111
- Williams, T. A., Cox, C. J., Foster, P. G., Szöllosi, G. J., and Embley, T. M. (2020). Phylogenomics provides robust support for a two-domains tree of life. *Nat. Ecol. Evol.* 4, 138–147. doi: 10.1038/s41559-019-1040-x
- Xia, P., Wang, S., Ye, B., Du, Y., Li, C., Xiong, Z., et al. (2018). A circular RNA protects dormant hematopoietic stem cells from DNA sensor cGAS-mediated exhaustion. *Immunity* 48, 688–701.e7. doi: 10.1016/j.immuni.2018.03.016
- Xiang, G., Li, Y., Sun, J., Huo, Y., Cao, S., Cao, Y., et al. (2024). Evolutionary mining and functional characterization of TnpB nucleases identify efficient miniature genome editors. *Nat. Biotechnol.* 42, 745–757. doi: 10.1038/s41587-023-01857-x
- Yang, Y., Fan, X., Mao, M., Song, X., Wu, P., Zhang, Y., et al. (2017). Extensive translation of circular RNAs driven by N-methyladenosine. *Cell Res.* 27, 626–641. doi: 10.1038/cr.2017.31
- Zago, M. A., Dennis, P. P., and Omer, A. D. (2005). The expanding world of small RNAs in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Mol. Microbiol.* 55, 1812–1828. doi: 10.1111/j.1365-2958.2005.04505.x
- Zaramela, L. S., Vêncio, R. Z. N., ten-Caten, F., Baliga, N. S., and Koide, T. (2014). Transcription start site associated RNAs (TSSaRNAs) are ubiquitous in all domains of life. *PLoS One* 9:e107680. doi: 10.1371/journal.pone.0107680
- Žedaveinyte, R., Meers, C., Le, H. C., Mortman, E. E., Tang, S., Lampe, G. D., et al. (2024). Antagonistic conflict between transposon-encoded introns and guide RNAs. *Science* 385:eadm8189. doi: 10.1126/science.adm8189
- Zhang, X.-O., Dong, R., Zhang, Y., Zhang, J.-L., Luo, Z., Zhang, J., et al. (2016). Diverse alternative back-splicing and alternative splicing landscape of circular RNAs. *Genome Res.* 26, 1277–1287. doi: 10.1101/gr.202895.115
- Zhang, Y., Zhao, Y., Liu, Y., Wang, M., Yu, W., and Zhang, L. (2020). Exploring the regulatory roles of circular RNAs in Alzheimer's disease. *Transl. Neurodegener.* 9:35. doi: 10.1186/s40035-020-00216-z