tRNA Modification Landscapes in Streptococci: Shared Losses and Clade-Specific Adaptations

Ho-Ching Tiffany Tsui[1], Chi-Kong Chan[2,3], Yifeng Yuan[4], Roba Elias[4], Jingjing Sun[2,3], Virginie Marchand[5], Marshall Jaroch[4&], Guangxin Sun[2,3], Irfan Manzoor[1], Ana Kutchuashvili[6], Yuri Motorin[5], Grazyna Leszczynska[7], Kinda Seaton[4], Kelly C. Rice[4], Manal A. Swairjo[6], Malcolm E. Winkler[1], Peter C. Dedon[2,3] and Valérie de Crécy-Lagard[4,8*]

[1]Department of Biology, Indiana University Bloomington, Bloomington, Indiana, USA

[2]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

[3]Antimicrobial Resistance Interdisciplinary Research Group, Singapore-MIT Alliance for Research and Technology, Singapore.

[4]Department of Microbiology and Cell Science, University of Florida, Gainesville, Florida, USA

[5]Université de Lorraine, SMP IBSLor, EpiRNA-Seq core facility, F-54000 Nancy, France

[6]Department of Chemistry, San Diego State University, California, USA

[7]Institute of Organic Chemistry, Lodz University of Technology, Łódź, Poland

[8]University of Florida Genetics Institute, Gainesville, Florida, USA

*Corresponding author: vcrecy@ufl.edu

[&]Current address: Kihealth Inc., Saint Augustine, Florida, USA

Abstract

tRNA modifications are central to bacterial translational control. Here, we integrated genetics, mass spectrometry, epitranscriptomics, and comparative genomics to map the tRNA modification genes of the Gram-positive pathogens *Streptococcus mutans* and *Streptococcus pneumoniae*. Both species show a marked loss of modifications dependent on Fe–S enzymes, consistent with a broader trend of Fe–S enzyme reduction in Streptococcus central metabolism. In addition, the D, m^1A, m^7G, t^6A, and i^6A modifications were mapped in *S. pneumoniae* tRNAs, and we confirmed that a unique DusB1 enzyme is responsible for the insertion of all the detectable D modifications. We uncovered differences in queuosine (Q) metabolism: while *S. mutans* synthesizes Q de novo, *S. pneumoniae* instead salvages $preQ_1$ and accumulates the epoxy-Q precursor, a strategy shared with multiple other Streptococci as revealed by analysis of Q pathways in 1,599 sequenced streptococcal genomes. Comparative essentiality profiling of modification genes revealed notable differences, including the essentiality of the N□-threonylcarbamoyladenosine (t□A) synthesis enzyme TsaE in *S. pneumoniae* but not in *S. mutans*, which was confirmed by genetic studies. We found that suppressor mutations in *asnS* encoding asparaginyl-tRNA synthetase (AsnRS) restored viability to Δ*tsaE* mutants, albeit with reduced growth. Our finding highlights the functional importance of modifications in the recognition of tRNAs by aminoacyl-tRNA synthetases.

Keywords tRNA modification, Queuosine, t^6A, Dihydrouridine, TsaE, pseudogene

Introduction

Streptococci are a diverse group of Gram-positive bacteria that are common members of the human microbiota, capable of colonizing the skin as well as mucosal membranes of the oral cavity, upper respiratory tract, intestines, and vaginal tract [1]. Clinically important species include *Streptococcus pyogenes* (Group A), responsible for pharyngitis, rheumatic fever, and necrotizing fasciitis[2]; *Streptococcus agalactiae* (Group B), a cause of neonatal infections [3]; *Streptococcus mutans,* a keystone pathogen in the development of dental caries, as well as infective endocarditis in at-risk patient groups [4]; and *Streptococcus pneumoniae*, a leading cause of otitis, pneumonia and meningitis [5]. These bacteria exhibit a wide range of virulence factors and adaptation mechanisms, making them important subjects of study in both basic microbiology and infectious disease research.

Transfer RNAs (tRNAs) are the essential decoding molecules in the translation process, and they undergo extensive post-transcriptional modification. These modifications fine-tune translation efficiency, fidelity, and tRNA stability [6,7]. In pathogenic bacteria, tRNA modifications can modulate stress responses [8,9], antibiotic resistance [9,10], metabolic fluxes [11] and the accurate synthesis of infection-related proteins [12,13] and are hence important in microbial pathogenesis and virulence in many different bacteria (see [14–16] for recent reviews). In addition, a few bacterial tRNA modification enzymes are essential and have no ortholog in humans and are thus being explored as antibacterial targets [17,18].

That said, our understanding of the role of tRNA modifications in pathogens remains limited, as comprehensive mapping of specific tRNA modifications and their corresponding biosynthetic enzymes have been completed in only a few organisms, including the model gram-negative *Escherichia coli* K12 [17]. For decades, the intracellular pathogen *Mycoplasma capricolum* was the only Gram-positive organism with combined experimental and bioinformatic evidence linking tRNA modifications with their genes [17,19]. As discussed previously [17], no single method, analytical or computational, can accurately predict the full set of tRNA modification enzymes in a given organism. However, the combination of liquid chromatography-tandem mass spectrometry (LC-MS/MS), tRNA sequencing, and bioinformatics methods greatly increases the accuracy of the predictions [17]. This type of integrative approach has enabled the mapping of modifications and their genes in three additional gram-positive bacteria: *Bacillus subtilis [20]*, *Mycobacterium tuberculosis* [21] and *Streptomyces albidoflavus* [22]. The initial analysis of tRNA modification genes in *B. subtilis* revealed open questions [20], including the specificities of dihydrouridine synthase enzymes and the identity of the enzymes involved in

cmnm^5(s^2)U deacetylation and nm^5(s^2)U methylation that have now both been resolved [23,24](**Fig. 1, Supplemental Data 1A**).

Dihydrouridine (D) is a widespread modification that is important for tRNA structure and flexibility [25]. Three subfamilies of dihydrouridine synthases with different specificities modify *E. coli* tRNAs. DusA (K05539) modifies positions 20 and 20a, DusB (K05540) position 17, and DusC (K05541) position 16. Only two DusB paralogs are present in *B. subtilis,* DusB1 (BSU00810) and DusB2 (BSU08030), and it was recently shown that in this model organism, DusB1 exhibits multisite enzyme activity, enabling D formation at positions 17, 20, 20a, and 47, while DusB2 specifically catalyzes U to D conversion at positions 20 and 20a [23]. Many pathogenic Bacillota (formerly Firmicutes) have lost one of the two DusB enzymes [26]. Many *Staphylococcus aureus* strains encode a full-length DusB2 and a truncated version DusB2-C, but only the full-length one is active [27], while *M. capricolum* and most *Streptococci* only encode a DusB1 [23,26]. Heterologous expression studies with the Mycoplasma DusB1 (MCAP_0837) enzyme showed that it modifies positions 17, 20, and 20a, but this is yet to be validated genetically in an endogenous host [26].

Mnm^{5}s^{2}U is a common wobble base (position 34) modification found in both Gram$^+$ and Gram$^-$ bacteria that is important for the decoding of NNA/NNG codons, particularly in split codon boxes [7]. The MnmEG complex catalyzes the first step of this complex pathway both in *B. subtilis* and *E. coli* [28,29] producing the cmnm^{5}s^{2}U or nm^{5}s^{2}U precursors. Several studies have linked the absence of MnmG (or MnmE) with pleiotropic phenotypes in many bacteria [30] including virulence defects in different Streptococci [31–34]. The subsequent stages of the pathway are catalyzed by non-orthologous enzymes in the two major model organisms. The *E. coli* FAD-dependent cmnm^5(s^2)U deacetylase [MnmC1/MnmC(o), a domain of the bifunctional MnmC (K15461) protein] is replaced in *B. subtilis* by the radical SAM enzyme MnmL (K07139)[24]. Two different families of methylases finalize the synthesis of mnm^{5}U by methylating nm^{5}U: the *E. coli* MnmC2/MnmD/MnmC(m) type and the *B. subtilis* MnmM type [24,35]. The functional roles of the *S. mutans* MnmL and MnmM orthologs have also been confirmed by genetic studies [24]. The corresponding genes do accumulate inactivating mutations in many *S. pneumoniae* strains that lack mnm^{5}s^{2}U and accumulate the nm^{5}s^{2}U precursor [24]. In *S. mutans,* the only other tRNA modification genes that have been characterized are involved in the synthesis of two other complex modifications of the tRNA Anticodon-Stem-loop (ASL): the wobble base modification Queuosine (Q)[36] at position 34 and the universal N6 adenosine modification -threonylcarbamoyladenosine (t^{6}A) at position 37 [37].

Q is a deazapurine derivative synthesized de novo from GTP in many bacteria in seven catalytic steps [38]. Q precursors such as $preQ_0$, $preQ_1$, or the queuine base (q) can be salvaged in some bacteria [38] using a wide variety of transporters, including the QPTR and QueT/QtrT subgroups of the Energy-coupling factor (ECF) transporter families [38,39]. Pathogenic bacteria that colonize the human host can use salvaged q either directly by mutating the substrate specificity of the bacterial TGT enzyme from $preQ_1$ to q [40], or indirectly by converting salvaged q to $preQ_1$, a reaction catalyzed by a radical-SAM enzyme QueL [40]. Queuine can be generated by cleaving Q with hydrolases such as QueK which has also been characterized in a human pathogen [40]. S*. mutans* can synthesize Q *de novo* using the same pathway as *E. coli,* and all Q synthesis genes have been experimentally validated, including the signature enzyme TGT that inserts the $preQ_1$ precursor into target tRNAs [36]. In addition, we confirmed *S. mutans* can salvage $preQ_0$ and $preQ_1$ using a QueT-dependent ECF transporter [36]. No obvious phenotype was linked to Q absence, but it was found that Q levels in *S. mutans* tRNAs were greatly influenced by media conditions [36]. It remains to be determined whether *S. mutans* or any other Streptococci can recycle or salvage q or Q.

t^6A is a universal modification and its synthesis requires two steps, catalyzed in bacteria by TsaC (or TsaC2) and the TsaBDE complex, respectively [41,42]. The essentiality of the t^6A synthesis genes varies among organisms. They are essential in *E. coli* and *S. aureus* but not in *Deinococcus radiodurans*, *B. subtilis,* or *S. mutans* [43]. In some organisms, t^6A is a determinant for the charging of $tRNA^{Ile}$ by Isoleucyl-tRNA synthetase [41,43] and could also be a determinant for lysidine synthase (TilS), which is one of the rare essential tRNA modification enzymes, as it is required for cognate tRNAs to decode AUA (Ile) codons [43]. However, it seems many organisms can circumvent these t^6A requirements by mechanisms that have not been fully elucidated.

In this work, we combined analytical, genomic, genetic, and bioinformatic approaches to predict the complete sets of tRNA modification enzymes in two model Streptococci (*S. mutans* and *S. pneumoniae)*. Our analysis revealed shared differences with the Gram-positive model *B. subtilis*, notably a consistent loss of iron–sulfur cluster–dependent modifications. We also found that the Q pathway is markedly reduced in *S. pneumoniae* compared to *S. mutans*, a pattern representative of most sequenced Streptococci, as shown by our comparative analysis of ~1,600 genomes. Finally, we demonstrate that, unlike in *S. mutans*, where t^6A is dispensable, this conserved anticodon–stem–loop modification is essential in *S. pneumoniae*, although

suppressor mutations arise at high frequency, pointing to previously unrecognized roles of t^6A in translation.

Methods

Databases resources

Different databases and repositories from the NCBI resources (https://www.ncbi.nlm.nih.gov/) [44] including Pubmed (https://pubmed.ncbi.nlm.nih.gov/) and RefSeq (https://www.ncbi.nlm.nih.gov/refseq/) [45] and COG (https://www.ncbi.nlm.nih.gov/research/cog/) [46], from the EBI resources (https://www.ebi.ac.uk/) [47] including Uniprot (https://www.uniprot.org/) [48] and InterPro (https://www.ebi.ac.uk/interpro/) [49], from the KEGG resources (https://www.genome.jp/kegg/) [50] including the KO (KEGG ORTHOLOGY) (https://www.genome.jp/kegg/ko.html) database, the Protein Data Bank (http://www.rcsb.org/) [51] and the BV-BRC (https://www.bv-brc.org/) [52] resources were routinely used. Subtiwiki (https://subtiwiki.uni-goettingen.de/v5/welcome) was used as the source of rRNA modification genes in Gram-positive [53].

Q synthesis and salvage gene/protein analyses

Genome DNA sequences and protein sequences from 1,599 *Streptococcus* strains were retrieved from the NCBI database in July 2025, applying filters for annotation ('RefSeq') and assembly status ('complete') (**Supplemental data 1F** and **Fig. S1 step 4**). Q proteins and Q genes were searched in them using BLASTp and tBLASTn [54] with a P-value cutoff of 1E-10 and an identity cutoff of 20% (**Fig. S1 step 5**). Query sequences were the Queuosine synthesis and salvage proteins from *B. subtilis* **(Supplemental Data 1A and 1E)**. Pseudogenes were called when matches were found by tBLASTn but not by BLASTp, indicative of internal stops or frameshifts. A hundred and twenty conserved marker proteins were identified and aligned using GTDBtk (v2.3.2) [55] with *Pseudolactococcus* and *Lactococcus* as the outgroup. A maximum likelihood tree of the concatenated marker proteins was generated using IQ-TREE2 (v2.2.2.7) [56], employing the LG+G substitution model and 1,000 fast bootstrap replicates (**Fig. S1 step 6**). The tree was visualized with the mapping of Q protein patterns using iTOL (v7.2.1) [57]. Branches of the same species harboring the same Q proteins were collapsed and represented by triangles, the size of which correlates with the number of species. Species names were

updated when their RefSeq classification disagree with their GTDB classification or their position in the tree (**Fig. S1 step 7**, **Supplemental Data 1E**). Species represented by fewer than three strains or rare Q gene patterns were hidden. [

Multiple sequence alignment of Q proteins (input sequences in **Supplemental Data 1F**) was performed using MAFFT (v7.520) [58] with setting "--maxiterate 1000 --localpair". Sequence logos were generated with the aligned sequences using Weblogo [59]. Alignments of genome sequences was performed using BLASTn (v2.15).

Sequence similarity network analyses

The sequence similarity network (SSN) of inosine/uridine-preferring nucleoside hydrolase (IPR023186) family was generated using EFI-EST (EFI Enzyme Similarity Tool, efi.igb.illinois.edu/efi-est)[60](**Fig. S1 step 8**). 36,956 sequences in the family were retrieved from UniProt and subjected to EFI-EST using the family option. Each node in the network represents one or multiple sequences that share no less than 90% identity. The initial SSN was generated with an Alignment Score Threshold (AST) set such that each connection (edge) represented a sequence identity above 40%. The nodes from *Streptococcus* were colored and visualized using Cytoscape 3.10.1[61]. More SSNs were created by gradually increasing the alignment score cutoff in small increments (usually by 5 AST). This process was repeated until paralogs in the IPR023186 family were separated into different clusters (AST=100). In the final view of the SSN, the shapes of nodes are based on the kingdom of species of the representative sequence of each node. The nodes of *Streptococcus* are filled in yellow. The nodes that have been annotated experimentally as not being QueK are filled in black. The nodes are filled in blue and skyblue when their encoding gene is next to *queT* or other transporter gene, respectively. The borders of nodes are highlighted in red when the representative sequences of each node are from genomes that lack *queDECF* genes. For better visualization, clusters with less than 30 nodes are hidden. The resulting SSN was then subjected to EFI-EST and EFI-GNT for gene neighborhood analysis with default settings. The retrieved genome neighborhood was visualized using Gene Graphics (https://genegraphics.net/)[62]. Identifiers for sequences used in the SSN and in the genome neighborhood analyses are available in **Supplemental Data S1G and S1H**.

Codon usage analyses

The genome distances between each of the 234 *S. pneumoniae* genomes and all other Streptococci present in our genome set were calculated using MASH v2.3[63] (**Fig. S1 step 9**).

26 genome pairs that each share at least 92% genome identity (distance < 0.08) between a *S. pneumoniae* strain and a member from another Streptoccocal clade were chosen (**Supplemental Data S1I**). The coding DNA sequences (CDSs) of the selected genome pairs were retrieved from the NCBI database. After removing pseudogenes, the number of each of the 61 sense codons was calculated per CDS (excluding start codons) (**Fig. S1 step 10**). Then, the genome-wide codon usage for every amino acid was compared using a paired two-tailed t-test (**Supplemental Data S1I**).

Structural analysis

Experimental structures of AsnRS, AspRS and LysRS were obtained from the Protein Data Bank (https://www.rcsb.org/), and the AlphaFold models for *Sp*AsnRS and *Sm*AsnRS were obtained from UniProt (IDs Q8CWQ4 and Q8DTM2, respectively). Structures were aligned and analyzed in Pymol (https://www.pymol.org/) [64]. The structure-based multi-sequence alignment was generated using the PROMALS3D online tool (http://prodata.swmed.edu/promals3d/)[65] and printed in ESPript [66].

Strains and culture conditions

All bacterial strains used in this work are listed in **Table S1**. Strains of *S. mutans* were routinely cultured statically at 37°C in 5% CO_2 in BHI medium (BD Biosciences) with 10 µg/mL erythromycin (Sigma) or 1 mg/mL kanamycin (Sigma) when appropriate. *S. pneumoniae* derivatives were derived from unencapsulated strains IU1824 (D39 Δ*cps rpsL1*) and IU1945 (D39 Δ*cps*), which were derived from the encapsulated serotype-2 D39W progenitor strain IU1690 [67,68]. Strains containing antibiotic markers were constructed by transformation of CSP1-induced competent pneumococcal cells with linear DNA amplicons synthesized by overlapping fusion PCR [69]. Strains containing markerless alleles in native chromosomal loci were constructed using allele replacement via the P_c-[*kan-rpsL*$^+$] (Janus cassette)[70]. Primers and DNA templates used to synthesize different amplicons are listed in **Table S2**. *S. pneumoniae* were grown on plates containing trypticase soy agar II (modified; Becton-Dickinson), and 5% (vol/vol) defibrinated sheep blood (TSAII-BA). Plates were incubated at 37°C in an atmosphere of 5% CO_2. TSAII-BA plates for selections contained antibiotics at concentrations described previously [69]. Bacteria were cultured statically in Becton-Dickinson brain heart infusion (BHI) broth at 37°C in an atmosphere of 5% CO_2, and growth was monitored by OD_{620} as described before [69]. Mutant constructs were confirmed by PCR and

DNA sequencing of chromosomal regions corresponding to the amplicon region used for transformation. Ectopic expression of various genes was achieved with a P_{Zn} zinc-inducible promoter in the ectopic *bga*A site. 1/10 concentration of Mn^{2+} was added with Zn^{2+} to prevent zinc toxicity [69]. 0.1 to 0.2 mM (Zn^{2+}/(1/10) Mn^{2+}) was added to TSAII-BA plates or BHI broth for inducing conditions. Mn^{2+} was added with Zn^{2+} to prevent zinc toxicity [69,71]. Ectopic expression of *asnS* [D121A] was achieved with a constitutive promoter P_{ftsA} at the *bgaA* site [72].

In all experiments, *S. pneumoniae* cells were inoculated from frozen glycerol stocks into BHI broth, serially diluted, and incubated 12–15 h statically at 37°C in an atmosphere of 5% CO_2. For culturing merodiploid strains IU18963 (Δ*tsaE*//P_{Zn}-*tsaE*$^+$) that require Zn^{2+} for expressing *tsaE* from a Zn-dependent promoter (P_{Zn}), 0.1 mM (Zn^{2+}/(1/10) Mn^{2+}) was added to BHI broth in the overnight cultures. The next day, cultures at OD_{620} ≈0.1–0.4 were diluted to OD_{620} ≈0.003 in BHI broth with no additional (Zn^{2+}/(1/10) Mn^{2+}) or the amounts of (Zn^{2+}/(1/10) Mn^{2+}) indicated for each experiment. Doubling time determination was performed by first examining the growth curves on a log scale to determine the time points when growth was in the exponential phase. Doubling times were determined with the exponential growth equation using only data points that exhibit exponential growth using GraphPad Prism version 10.0.0 for Windows, GraphPad Software, Boston, Massachusetts USA, www.graphpad.com. Maximal growth yields were determined by the highest OD_{620} values obtained within 9 h of growth. Cultures were sampled for phase microscopy at early to mid-exponential phase.

S. pneumoniae **Tn-seq transposon library generation and insertion sequencing**

Tn-seq was carried out according to protocols described in [73] using a transposon insertion library generated for WT D39 Δ*cps rpsL1* (IU1824) and WT D39 *cps*$^+$ *rpsL1* (IU1781) cultured with BHI media. In brief, transposon library starter cultures were thawed, and were diluted to OD_{620} ≈ 0.005 in 5 ml of BHI and were grown at 37°C with 5% CO_2 to OD_{620} ≈ 0.4. 5 ml of culture at OD_{620} ≈ 0.4 were used to extract genomic DNA. *MmeI* digested DNA were ligated to adaptors and sequenced on the Illumina NextSeq 500 at the Center for Genomics and Bioinformatics, Indiana University Bloomington. Data were mapped and analyzed, and Tn insertion data were visualized graphically using the Artemis genome browser (version 10.2) [74].

S. pneumoniae **transformation assays**

Transformations were performed as previously described [72,75]. Δ*tsaE*::P_c-*erm,* Δ*tsaE* P_c-*kan-rpsL* amplicons, and positive control Δ*bgaA*::P_c-*erm* amplicon were synthesized by PCR using

the primers and templates listed in Table S2 and contain ≈1 kb of flanking chromosomal DNA. A fusion Δ*tsaE* markerless amplicon or an amplicon from strain IU18963 were used for transformation of IU18886. All transformation experiments were performed using no added DNA as the negative control. The sizes of colonies indicated in **Table 1** were relative to colonies transformed with the WT strain with a control Δ*bgaA* amplicon. For transformations in 0.1 mM or 0.2 mM (Zn^{2+}/(1/10) Mn^{2+}), $ZnCl_2$ and $MnSO_4$ stock solutions were added to transformation mixes and soft agar for plating and spread onto blood plates containing the respective (Zn^{2+}/(1/10) Mn^{2+}) concentrations to induce gene expression under control of the P_{Zn} zinc-inducible promoter in the ectopic *bgaA* site.

Whole-genome DNA sequencing

Whole-genome sequencing was employed to identify suppressor mutations and verify the genomes of the constructed mutants. Strains IU18824, IU18826, E797, and K797 (**Table S2**) containing suppressor mutations that allowed growth of a Δ*tsaE* mutant were isolated as described in Results. Genomic DNA preparation, DNA library construction, Illumina MiSeq DNA sequencing, and bioinformatics analyses were performed as described previously [75]. Genomic DNA from the *S. mutans tsaE* deletion strain (JB409) and its isogenic WT strains (*S. mutans* UA159, Wen laboratory Isolate) were extracted using DNeasy Blood and tissue kit (Qiagen), following the manufacturer's instructions. and was sequenced at the Microbial Genome Sequencing Center (MiGS) Illumina sequencing libraries were prepared using the tagmentation-based and PCR-based Illumina DNA Prep kit and custom IDT 10bp unique dual indices (UDI) with a target insert size of 280 bp. No additional DNA fragmentation or size selection steps were performed. Illumina sequencing was performed on an Illumina NovaSeq X Plus sequencer in one or more multiplexed shared-flow-cell runs, producing 2x151bp paired-end reads. Demultiplexing, quality control and adapter trimming were performed with bcl-convert1 (v4.2.4), a proprietary Illumina software for the conversion of bcl files to basecalls. Illumina generated 2x151bp paired end read data was used as the input for variant calling against the reference (GenBank id: AE014133) Variant calling was carried out using BreSeq1 under default settings [76].

tRNA and rRNA extraction

Bulk tRNA samples from *S. mutans* cells were prepared in previous studies [24,36]. tRNA samples from *S. pneumoniae* derivatives were prepared as described in detail previously (see MW methods in [24]). Overnight cultures (12–15 h) diluted to an initial OD_{620nm} 0.005 and

allowed to grow until an OD_{620nm} of ≈ 0.2. For cultures treated with $PreQ_0$ (Millipore Sigma, AMBH2D6F0DFD) or $PreQ_1$ (Millipore Sigma, product # SML0807), stock solution of 100 µM $PreQ_0$ or $PreQ_1$ in DMSO was added to BHI media to a final concentration of 100 nM at initial growth. The same volume of DMSO was added to the control samples. The purified tRNA samples were stored at −80°C until further analysis.

For rRNA samples, the WT *Streptococcus mutans* UA159 strain was cultured in BHI broth at 37°C in 5% CO_2 until mid-log phase (OD_{600}= 0.6–0.8). Cells were harvested and lysed as described by Chia et al. [77], followed by TRIzol extraction (Invitrogen) and purification using the NucleoBond RNA/DNA kit (Macherey-Nagel) according to the manufacturer's instructions. The resulting preparation was treated with DNase I (NEB) for 10 minutes at 37 °C, and the RNA was then cleaned up using the RNeasy Mini Kit (Qiagen). RNA concentration and purity were assessed by NanoDrop spectrophotometry, and integrity was verified by running a 1 % agarose; The RNA electropherogram and RNA Integrity Number (RIN) were obtained using an Agilent 4200 TapeStation (ICBR, UF).

Phase microscopy and cell measurements

Cultures were sampled for phase microscopy at OD_{620} ≈0.1–0.2 for IU1824 and IU18963, and OD_{620}≈0.03–0.05 for IU18824. 500 µL of culture were centrifuged at 20,000g at room temperature for 3 min. 450 µL of supernatant was removed and the pellet was resuspended in the remaining 50 µL media. 1.5 µL of cells were placed on a slide and covered with a coverslip and observed with phase contrast microscopy using a Nikon Eclipse E400 microscope.

Cell lengths and widths of strains growing in BHI broth were measured from phase-contrast images by using Nikon NIS-Element BR software as described before [69]. Only ovoid-shape predivisional cells were measured. Cells (88 or more) from 2 independent experiments were measured and plotted with scatter plot. P values were obtained by one-way ANOVA analysis by using the nonparametric Kruskal-Wallis test in GraphPad Prism program.

RNA hydrolysis for LC-MS/MS analysis

Purified tRNA and rRNA (3 µg) were hydrolyzed to nucleosides as previously described [78] in a 50-µL enzyme cocktail containing 8 U benzonase (Sigma), 5 U calf intestinal alkaline phosphatase (Sigma), 0.15 U snake venom phosphodiesterase I (Sigma), 5 ng coformycin NCI), 0.1 mM deferoxamine (Sigma), 0.1 mM butylated hydroxytoluene (Sigma), 50 nM internal

standard [^{15}N]$_5$-deoxyadenosine (Cambridge Isotope Laboratories), 2.5 mM $MgCl_2$ (Sigma), and 5 mM Tris-HCl buffer pH 8.0 (Invitrogen). Enzyme control reactions were prepared identically but without the addition of RNAs. The reaction mixtures were gently tapped, briefly spun down, and incubated at 37 °C for 6 h. After hydrolysis, the samples were centrifuged at 3,000 × *g* at 4 °C for 10 min and analyzed by LC-MS/MS.

LC-MS/MS analysis of modified ribonucleosides

The LC-MS/MS analysis of tRNAs extracted from *S. mutans* was described previously [24,36]. For *S. pneumoniae*, a first batch was analyzed as described in [24]. All other LC-MS/MS analyses were performed on an Agilent 1290 Infinity II UHPLC system equipped with an inline diode array detector (DAD) and coupled to an Agilent 6495c triple quadrupole mass spectrometer (MS). For each run, 180 ng of hydrolyzed RNA was injected onto a Waters ACQUITY UPLC BEH C18 column (50 × 2.1 mm i.d., 1.7 µm) equipped with an ACQUITY UPLC BEH C18 VanGuard Pre-column (5 × 2.1 mm i.d., 1.7 µm) and an ACQUITY Column In-line Filter (0.2 µm). To facilitate the detection of low abundance modified ribonucleosides, a second injection of 600 ng of hydrolyzed RNA was performed for each sample. The column was maintained at 30 °C and operated at a flow rate of 0.3 mL/min with mobile phase solvents consisting of Buffer A (0.02% formic acid in water) and Buffer B (0.02% formic acid in 70% aqueous acetonitrile). The gradient of Buffer B was programmed as follows: 0–5 min, 0–1%; 5–7 min, 1–3%; 7–9 min, 3–7%; 9–10 min, 7–10%; 10–12 min, 10–12%; 12–13 min, 12–15%; 13–15 min, 15–20%; 15–16 min, 20–75%; 16–17 min, 75–100%; 17–20 min, held at 100%, 20–21 min, 100–0%; and 21–30 min, re-equilibrated at 0%. The DAD wavelength was set to 260 nm to acquire signals of canonical ribonucleosides. The JetSream ESI source operated in positive-ion mode with the following optimized parameters: drying gas temperature, 200 °C; gas flow, 11 L/min; nebulizer, 20 psi; sheath gas temperature, 300 °C; sheath gas flow, 12 L/min; capillary voltage, 3000 V; and nozzle voltage, 0 V. MS/MS analysis leveraged a dynamic multiple reaction monitoring (dMRM) mode with retention time windows set as ± 2 min and collision energies (CEs) optimized for maximal sensitivity. Retention time of modified ribonucleosides were confirmed with synthetic standards (**Supplemental Data 2A**) and served as a primary criterion for peak identification. Raw peak areas for each modified ribonucleoside were extracted using Agilent QQQ Quantitative Analysis version 10.2 and normalized to the sum of the peak areas for the four canonical ribonucleosides recorded by the DAD. Raw LC-MS/MS data have been deposited to the ProteomeXchange Consortium

(http://proteomecentral.proteomexchange.org) via the PRIDE (http://www.ebi.ac.uk/pride/) partner repository with the dataset identifier PXD069173.

Besides the modified ribonucleosides for which we have synthetic standards, we also monitored a set of putative modifications that lack available synthetic standards in our inventory but have been previously reported in other prokaryotes. Information regarding precursor and product ions for these modifications was obtained from Modomics [79], precedent studies, and the literature, and incorporated into the dMRM method. Identities for these putative modifications were validated using a two-step approach. First, extracted ion chromatograms from the unit-resolution LC-MS/MS analysis of hydrolyzed tRNAs were compared with those from enzyme control reactions; only peaks observed exclusively in the hydrolyzed tRNAs were considered for further analysis. These peaks were subsequently subjected to high-resolution LC-MS analysis using a Thermo Fisher Scientific Dionex Ultimate 3000 UHPLC system coupled to an Orbitrap Q Exactive mass spectrometer equipped with an Ion Max source and a heated ESI (HESI II) sprayer. For each injection, 2 µg of hydrolyzed tRNAs was loaded onto the same analytical column, eluted with the same mobile phase solvents, and subjected to the same LC gradient and column oven temperature as described above. The source and MS parameters are as follows: spray voltage, +4.2 kV; capillary temperature, 320 °C; sheath gas, 50 arbitrary unit (au); auxiliary gas, 15 au; spare gas, 3 au; max spray current, 100 µA; probe heater temperature, 400 °C; and S-Lens RF level, 70 au. Initially, the hydrolyzed tRNAs were analyzed using a targeted selected ion monitoring (SIM) mode with a mass inclusion list of putative modifications identified from the LC-MS/MS analysis. The Automatic Gain Control (AGC) target was set at 20,000, with a maximum injection time of 150 ms. SIM mass spectra were acquired at a resolution of 35,000 full width at half maximum (FWHM) at m/z 200 with an isolation window of 4.0 Da. Peaks with a mass error < 5 ppm were subjected to targeted MS/MS fragmentation. To this end, the AGC target was increased to 500,000, with a maximum injection time of 100 ms. MS/MS spectra were acquired were acquired over an *m/z* range of 50 – 450 at a resolution of 17,500 FWHM at *m/z* 200 with an isolation window of 2.0 Da and CEs between 10 and 20 eV. The acquired spectra were compared to those reported in literature or to fragmentation patterns predicted using the Mass Fragmentation Tool in ChemDraw version 22.0.

Analysis of tRNA modifications by AlkAnilineSeq

tRNA samples (~200 ng) from *S. mutans* and *S. pneumoniae* were subjected to random fragmentation by alkaline hydrolysis in 50 mM sodium-bicarbonate buffer at pH 9.2 and 96°C for 5 min [80]. The reaction was stopped by ethanol precipitation using 3M Na-OAc, pH 5.2 and glycoblue. After centrifugation, the RNA pellet was washed with 80% ethanol and resuspended in nuclease-free water. RNA fragments were de-phosphorylated by Antarctic phosphatase (NEB ref M0289L, USA) at 37°C for 1h and precipitated using 3M Na-OAc, pH 5.2 and glycoblue as previously described [81]. After centrifugation, the RNA pellet was washed with 80% ethanol and the pellet was resuspended in 1M Aniline pH 4.5 and incubated for 15 min at 60°C in the dark [82] The reaction was stopped by ethanol precipitation using 3M Na-OAc, pH 5.2 and glycoblue. The pellet was washed twice with 80% ethanol, dried and resuspended in 3.5 µL of nuclease free water. RNA fragments were converted to sequencing library using the NEBNext® Small RNA Library Prep Set for Illumina® (NEB ref E7330S, USA) following the manufacturer's recommendations. DNA library was quantified using a fluorometer (Qubit 3.0 fluorometer, Invitrogen, USA) and qualified using a High Sensitivity DNA chip on Agilent Bioanalyzer 2100. Libraries were multiplexed and subjected to high-throughput sequencing on an Illumina NextSeq2000 instrument with a 50 bp single-end read mode.

High quality raw sequencing reads (> Q30) were subjected to trimming using Trimmomatic v0.39 [83] with the following parameters: MINLEN:08, STRINGENCY:7, AVGQUAL:30, trimmed reads were used for alignment without further processing. Trimmed reads were aligned to the *S. pneumoniae D39* tRNA sequences extracted from the gtRNAdb [84] using bowtie2 v2.4.4 [85]in end-to-end mode (--no-unal --no-1mm-upfront -D 15 -R 2 -N 0 -L 10 -i S,1,1.15 as other bowtie2 parameters), only uniquely mapped reads in positive orientation were retained for further analysis. 5'-reads' extremities were counted for each RNA position in the reference; all further steps were performed in R/R-studio environment. After a -1 shift in the sequence position, since ligated 5'-P extremity is at the N+1 nucleotide in the RNA sequence, this reads' count was used as the measure for intensity of cleavage at a given position. Four scores were used for analysis of AlkAnilineSeq raw data [80,81]: normalized cleavage (NCleavage), normalized count (NormCount), normalized G count (NormGcount) and stop ratio. Normalized cleavage corresponds to 1000x proportion of reads starting at a given position to the total number of reads mapped to a given RNA sequence. This score is less noisy but also less sensitive than others and is well-suited for detection of major cleavage events in RNA. In contrast, stop ratio (closely derived from ψ-ratio used for analysis of Ψ-seq data [86]) is calculated as a ratio of reads starting at a given position to the total number of reads passing (covering) it. This score is relatively sensitive, but also noisy. The last two scores used,

NormCount and NormGcount use local normalization to the median of cleavage signals in ±5 nt window around of analysed position. NormCount uses all cleavage signals in the window, while signals corresponding to G residues are excluded from calculation of the normalization median for NormGcount. These scores represent the intensity of the cleavage at a given nucleotide compared to the local cleavage background in the adjacent RNA region. None of AlkAnilineSeq scores show linear dependence between the score's value and stoichiometry of RNA modification, stop ratio has linear segment at very low modification levels (<5%), while the three other scores better represent the modification stoichiometry at higher modification levels in RNA (5-50%).

Analysis of tRNA modifications by GLORI

GLORI is based on deamination and includes a glyoxal treatment that serves as a 'protection' step, selectively masking guanosine residues from nitrite (NO_2^-)-induced deamination. This is followed by deprotection and reverse transcription to convert RNA into cDNA. The conditions used were as described in [87]. Briefly, 100 ng of tRNA samples were subjected to glyoxal treatment for 30 min at 50°C, followed by a boric acid treatment for 30 min at 50°C. The protected RNA is then subjected to nitrite-mediated deamination followed by RNA deprotection in a buffer containing 500 mM TEAA and formamide. RNA was then precipitated, end-repaired, and purified before being subjected to library preparation with NEBNext small RNA library following the manufacturer's protocol. The quality and quantity of each library were assessed using a high-sensitivity DNA Chip on a Bioanalyzer 2100 and a Qubit 3.0 fluorometer. High-throughput sequencing of the multiplexed libraries was performed on an Illumina NextSeq 2000 instrument in a 2x50 nt paired-end mode. Raw sequencing reads were inspected with FastQC and adapter sequence was removed by trimmomatic v0.39 [83]. Alignment to the A->G converted tRNA reference sequence (*S. pneumoniae D39* tRNA sequences) was done by Bowtie2. v2.4.4 [70] with slightly relaxed alignment stringency, allowing the retention of 1 nt-mismatched reads. Further analysis was done by *samtools mpileup* utility and counting the mismatch profile at every position in the reference. Since RNA A residues were de-aminated to inosines, sequencing reads are expected to contain only CGT nucleotides and thus align perfectly to A->G converted tRNA reference sequence. However, the residual A residues observed in the sequence data correspond to deamination-resistant modified As (m^6A and all other N6-modified A, m^1A was also detected). In the case of partial modification, both G and A are detected at a given position. Thus, the mismatch G in a reference to A in the sequencing read can be used as a score to measure a molar ratio of modified A in RNA (GtoA score). Raw

AlkAnilineSeq and GLORI data are available at the European Nucleotide Archive under the accession number PRJEB96333.

Results

Combining LC-MS/MS, tRNAseq, and *in silico* analysis to predict tRNA modification genes in two model Streptococci.

Over several years and studies [24,36], we performed LC-MS/MS analysis of bulk tRNA isolated from different *S. mutans* UA919 and *S. pneumoniae* D39 derivatives (13 independent *S. mutans* tRNA samples and 20 independent *S. pneumoniae* tRNA samples (**Table S1** and **Sup data 2BCDE**). By comparing retention times of chromatographic peaks in hydrolyzed tRNAs with those of synthetic standards, our analysis identified 31 modifications in *S. mutans* tRNAs and 32 modifications in *S. pneumoniae* tRNAs (**Supplemental Data 2BC** and **3**). Not all these modified nucleosides are from tRNAs. It is well known that modifications derived from contaminating 16S and 23S rRNA can be detected in high quantities in LC/MS analyses of nucleoside digests from tRNA preparations [20]. In addition, some modifications may originate from media contamination [85,86] or have non-enzymatic origins, such as m^1A, m^3C, or m^7G [88]. Hence, LC/MS needs to be combined with tRNA-Seq and computational predictions to map the modifications on tRNA molecules and establish the links between tRNA modifications and their corresponding genes.

Two complementary strategies were employed to identify the tRNA modification enzymes encoded in the *S. mutans* UA159 genome. First, we used a curated set of 44 *B. subtilis* tRNA modification proteins (**Supplemental Data 1A**) to perform BLAST searches against the predicted proteomes of *S. mutans* UA159 (RefSeq Assembly: GCA_000007465.2) and *S. pneumoniae* D39 (RefSeq Assembly: GCA_000014365.1) to identify homologs. In parallel, we leveraged a curated list of orthologous groups from the COG and KEGG (KO) databases, mapped across ~1,000 bacterial genomes [Reed C.J. and de Crécy-Lagard V. (unpublished) and **Supplemental Data 4**], to assess presence/absence patterns and cross-reference them with the BLAST hits. Integrating these datasets with our LC/MS profiles nucleosides in both *S. mutans* and *S. pneumoniae* (**Fig. 1**; **Supplemental Data 1BCD**).

We infer that several nucleosides detected by LC/MS—including acp^3U, m^2G, $m^{22}G$, m^3C, m^3U, m□C, and m□'□A, originate from rRNA or media contaminations, as no candidate tRNA-modifying enzymes were found for their incorporation. Putative rRNA-modifying enzymes

could be identified for all of these, except for m^3C and acp^3U (**Supplemental Data 2F**). Traces of acp^3U were observed exclusively in *S. pneumoniae*, likely due to contamination from the sheep blood in the TSAII-BA growth medium. Similarly, the traces of Q in *S. pneumoniae* samples are likely from media contamination (see section below). Both m^3C and m^5C were detected in both Streptococci, but no m^4C was detected (**Supplemental Data 3**). The presence of SunL (16S rRNA m^5C967-methyltransferase) homologs in both species (**Supplemental Data 2F)** suggests that m^5C is indeed a rRNA modification. However, m^3C had not been detected in bacterial RNAs to date [79], and homologs of RsmH that generally catalyze the insertion of m^4C in rRNA were identified in both Streptococci (**Supplemental Data 2F**). Hence, in this case, the in silico and the LC/MS data did not match, warranting further investigations. As m□Cm is a conserved bacterial 30S rRNA modification [79] [and both Streptococcal genomes encode orthologs of *E. coli* RsmI (**Supplemental Data 2F**)—an enzyme responsible for 2'-O-methylation of C1402 in 16S rRNA [89]—we revisited our initial analysis and found that m□Cm was indeed present in our tRNA samples (**Supplemental Data 3).** To confirm this presence was due to contaminating rRNA fragments, we quantified m□Cm levels in bulk rRNA and tRNA extracted from *S. mutans* UA159 and indeed found them to be 750-fold higher in rRNA than tRNA (**Supplemental Data 2F and 3)**. The LC/MS analysis of nucleosides derived from *S. mutans* rRNA samples confirmed m^2G, $m^{2,2}G$, m^3U, m□C, and m□'□A were only present in rRNAs, while Gm, Ψ, m^7G, m^5U, and possibly D were found in both rRNAs and tRNAs. This analysis also showed that m^3C was only present in traces in rRNA and was certainly derived from a contamination or a non-enzymatic reaction in our tRNA samples. Finally, no Am was detected in rRNA; hence, we cannot rule out that this modification is enzymatically inserted in tRNA by TrmL or an as-yet-unknown enzyme.

Based on enzyme predictions and the LC/MS data, we estimate that 21 of the 31 nucleosides detected in *S. mutans* and 19 of the 32 in *S. pneumoniae* derive from tRNAs. AlkAnilineSeq and GLORI analyzes allowed us to specifically map the D, m^1A, m^7G, t^6A, m^6A and i^6A modifications in *S. pneumoniae* tRNAs sequences (**Fig. 2C** and **Fig. S2**). Exact mapping of Ψ/Um/Cm/Gm and potentially Am in specific tRNAs would require additional analyses.

The main difference between *B. subtilis* and *S. mutans* is the absence of the ms^2i^6A, ms^2t^6A, and ct^6A-derived modifications and the corresponding enzymes (MiaB, MtaB, and TcdA) (**Fig. 1** and **Supplemental Data 1B**). In addition, the loss of one of the DusB2 orthologs and the absence of the alternate GTP cyclohydrolase IB/MtrA reduce the number of tRNA modification genes from 44 in *B. subtilis* to 39 in *S. mutans* (**Fig. 1** and **Supplemental Data 1B**). The same

gene/modification losses were also observed in *S. pneumoniae,* but a further decay of complex modification pathways, such as the previously reported loss of MnmL [24] and the decay of the Q pathway described below, led to a final count of 33 tRNA modification genes in this respiratory pathogen (**Fig. 1** and **Sup data 1C**).

DusB1 is responsible for all D synthesis at positions 17, 20, and 20a in *S. pneumoniae*

S. mutans and *S. pneumoniae* only encode a DusB1, whereas *B. subtilis* encodes DusB1 and DusB2 ([23] and **Fig, 1**). To determine which positions were modified by DusB1 in *S. pneumoniae,* the corresponding gene (*spd_2016*) was deleted, and AlkAnilineSeq was performed on tRNA extracted from the WT and Δ*dusB1* strains. We found that *dusB1* was responsible for the formation of D17, 20, and 20a, as all three modifications disappeared in the mutant (**Fig. 2AB and Fig. S2**). In the process, we mapped the D residues in 40 *S. pneumoniae* tRNAs **(Fig. S2).** We suggest that the DusB1 proteins of *S. mutans* and *S. pneumoniae* have similar catalytic profiles (**Fig. 1**),

In *B. subtilis,* $tRNA^{Met}_{CAU}$ harbors a D at position 47 which was identified by LC/MS of the T1 fragments tRNA, that D47 can only be seen by AlkAnilineSeq in a *trmB* mutant, so that the m^7G46 signal does not overshadow the D signal [23]. Hence, we cannot formally rule out the presence of D47 in *S. pneumoniae* in this tRNA. Recent analysis of *S. aureus* tRNA modification profiles [27] shows that D47 is not found in any tRNA; hence, we did not include this modification in our predictions (**Fig. 1**), even if the final call would require additional experiments.

During the mapping of D residues by AlkAnilineSeq, we also identified multiple m^7G46 modifications in *S. pneumoniae* tRNAs (**Fig. 2AB and Fig. S2A** and, unexpectedly, several U8 also give an AlkAnilineSeq signal, presumably corresponding to the predicted s^4U8. This observation was further confirmed by analysis of the mismatch score at those positions. As anticipated, all AlkAnilineSeq signals at position 8 were also detected as partial U->C transitions, which is a characteristic feature of s^4U, base pairing with A, but also with G residues during the reverse transcription step (**Fig. S2B**). However, we believe that methods allowing specific s^4U detection [90] should be used to map these modifications comprehensively and thus did not include them in our summary.

***S. pneumoniae* salvages preQ$_1$ and not preQ$_0$, as predicted from metabolic reconstruction and stops at the intermediate oQ.**

Metabolic reconstruction of the *S. pneumoniae* Q synthesis pathway (**Fig. 3A and S3**, **Supplemental Data S1CD**) predicts that, unlike *S. mutans*, which is a Q prototroph, *S. pneumoniae* cannot synthesize preQ$_0$ de novo due to the absence of the *queCDE* genes [38]. However, the presence of *queT* and *queF* homologs suggests that *S. pneumoniae* can salvage preQ$_0$, reduce it to preQ$_1$, and incorporate it into tRNA via the signature enzyme TGT (**Fig. 3A**). Notably, homologs of *queG* and *queH*—two non-orthologous enzymes that catalyze the final step of Q biosynthesis—are missing. This indicates that the pathway in *S. pneumoniae* likely halts at the penultimate step: the formation of epoxyQ-tRNA (oQ-tRNA) from preQ$_1$-tRNA catalyzed by QueA (**Fig. 3A and S3**). Such an incomplete Q pathway is highly unusual, with only one other documented case to date [91].

In our initial LC-MS analysis, *S. pneumoniae* was grown in BHI medium without added preQ$_0$ or preQ$_1$ (**Supplemental Data 2C**). Only trace amounts of Q were detected, likely due to media contamination. In contrast, oQ levels were ~20-30 times higher in *S. pneumoniae* than in *S. mutans* (**Supplemental Data 2BC**). The identity of the oQ modification was confirmed using high-resolution LC-MS with targeted SIM and MS/MS analyses. SIM data showed excellent agreement between the observed m/z (426.1608) and the theoretical m/z (426.1619) for the protonated oQ ion, with a mass error < 3 ppm (**Fig. S4**). MS/MS analysis of m/z 426.1619 yielded major fragment ions at m/z 295.1035 (<1 ppm error) and 163.0616 (<2 ppm error), corresponding to the 7-methyl-7-deazaguanosine and 7-methyl-7-deazaguanine moieties of oQ, confirming its identity.

oQ levels in *S. pneumoniae* remained ~10-fold lower than the Q levels observed in *S. mutans* (**Supplemental Data 2BC**). Because tRNAs were initially extracted from cells grown in undefined media with unknown concentrations of preQ$_0$/preQ$_1$, we repeated the experiment with 100 nM preQ$_0$ supplementation. However, oQ levels were unchanged (**Fig. 3B** and **Supplemental Data 2D**). This was unexpected given the presence of QueT (predicted to transport preQ$_0$/preQ$_1$) [36] and QueF predicted to reduce preQ$_0$ into preQ$_1$ [38]. Upon closer inspection, the *S. pneumoniae* D39 QueF was found to be lack catalytic residues: it is a fusion of an N-terminal QueC fragment and a C-terminal QueF domain lacking the catalytic cysteine essential for thiamide adduct formation with preQ$_0$ (**FigS5**) [92]. We then repeated the experiment with 100 nM preQ$_1$ instead. This led to a substantial (>100-fold) increase in oQ levels and high accumulation of preQ$_1$ in *S. pneumoniae* tRNA, while the trace Q levels

remained unchanged (**Fig. 3B** and **Supplemental Data 2D**). The result was supported by the observation that all TGTs from *Streptococcus* have a substrate binding motif that resembles *E. coli* TGT which takes preQ$_1$ as the substrate (**Fig. S6**).

In summary, unlike *S. mutans* UA159, which encodes a fully functional de novo Q synthesis pathway [36], *S. pneumoniae* D39 has lost the capacity to synthesize preQ$_1$ but can still salvage this precursor to insert it into tRNAs. In this organism, the pathway also stops at the oQ-tRNA intermediate. These findings led us to conduct a comprehensive comparative genomic analysis of Q synthesis and salvage genes in sequenced Streptococci to evaluate the variability of the Q pathway across the entire clade.

Plasticity of Q biosynthesis pathways in Streptococci as seen with numerous gene losses and pseudogene accumulation

We analyzed the distribution of queuosine (Q) synthesis and salvage genes in 1,599 *Streptococcus* genomes, selected based on genome completeness **(Fig. 3CD; Supplemental Data S1E**). 1598 strains are predicted to harbor a deazaguanine modification in tRNA, as they all encode a functional TGT enzyme, the remaining one is likely to be a sequencing error. QueA, the next enzyme in the pathway catalyzing the synthesis of oQ-tRNA from preQ$_1$-tRNA [38] is also highly prevalent, making it unlikely that preQ$_1$ is the final incorporated base, unlike in *Bartonella* species [93]. Fewer than 20% of the species analyzed synthesize Q *de novo*, a group that includes many *Streptococcus thermophilus* isolates, all examined strains of *Streptococcus salivarius*, *S. mutans*, *Streptococcus oralis*, and *Streptococcus mitis* (common inhabitants of oral cavity), and the group D streptococci *S. gallolyticus* and *S. equinus* which inhabit the GI tract of mammals (**Fig. 3CD**). Notably, *Streptococcus* species that synthesize preQ$_1$ never encode both non-orthologous enzymes that can catalyze the hydrolytic reduction/deoxygenation of oQ (QueG or QueH). Instead, they encode only one or the other (**Fig. 3CD**). By contrast, 755 strains—mainly *Streptococcus pyogenes* and *Streptococcus suis*—encode both QueG and QueH, yet all of these are auxotrophic for preQ$_1$. Most Streptococci (~80%) must rely on preQ$_1$ salvage (**Fig. 3C**). In line with the prevalence of preQ$_1$ salvage, orthologs of QueT or QPTR are found in most Streptococci analyzed (**Fig. 3D**). Finally, we found that the termination of the pathway at the oQ step, which we experimentally validated in *S. pneumoniae* D39 (**Fig. 3B**), is predicted to occur in ~15% of the strains analyzed. Many of these harbor *queH* and *queG* pseudogenes, detected in 131 and 39 strains, respectively (**Fig. 3CD**). Gene losses and decay

were also prevalent with the *queF* genes. 116 unique QueF sequences of two types were found in 544 *Streptococcus* strains: canonical QueF and inactive QueCF fusion, like the one found in the *S pneumoniae* D39 strain (**Fig. 3CD** and **Fig. 4 and S5**). All the 220 analyzed *S. pneumoniae* genomes encode defective QueCF fusions (**Fig. 4B**)

We also identified the Streptococci orthologs of QueK enzymes that hydrolyze the Q nucleoside to liberate the queuine (q) (**Supplemental Data 1E**), QueK proteins are members of the superfamily isine/uridine-preferring nucleoside hydrolase (IPR023186) family, but members of the QueK subgroup can be separated from the other members of this large family using Sequence Similarity Networks (SSNs) ([40] and **Fig. S7A**). Indeed, we found that QueK homologs from Streptococci are grouped with the experimentally validated QueK from *Clostridioides difficile* 630 (QueK Cd in **Fig. S7A**). Most genomes in this cluster lack *queDECF* genes but encode TGT. In addition, the majority of *queK* genes in this cluster, including those from *Streptococcus*, are next to genes encoding QueT family transporters (**Fig. S7B**). Hence, it is likely that QueK orthologs in *Streptococcus* function as queuosine hydrolase. However, no QueL orthologs could be identified in the analyzed *Streptococcus* genomes (**Supplemental Data 1E**). QueA orthologs are always present (Fig. 3CD), and the encoded Tgt proteins are all of the canonical preQ$_1$ inserting/bacterial type [94] (Fig. S6). Hence, it is not clear what the fate of the released q base would be in Streptococci. One cannot rule out the possibility that a non-orthologous enzyme yet to be discovered replaces QueL to generate preQ$_1$ in these organisms.

To investigate the potential impact of oQ on synonymous codon usage of the GUN Q-dependent codons (Tyr, His, Asp and Asn), we compared codon usage of 26 of the 234 *S. pneumoniae* strains in our dataset predicted to harbor oQ each to a closely related Streptococci (>92% identity) not belonging to the *S. pneumoniae* clade (**Supplemental Data 1F**). We found that *S. pneumoniae* strains exhibited a higher NAU-to-NAC ratio (**Fig. S8A**), which was primarily attributed to the Tyr and His codon biases (**Fig. S8C-F**). However, this difference could be a result of difference in GC content (Fig. S8B) (**Supplemental Data 1F**) as this can be a driver the choice of the wobble base [95]and more extensive analyses will be required to separate these effects.

Survey of the essentiality of tRNA modification genes in *S. mutans* and *S. pneumoniae* reveals differences between the two clades

Based on published studies [96–100] and TnSeq data generated as described in the methods section, we compiled essentiality data for all predicted *S. mutans* UA150 and *S.*

pneumoniae D39 tRNA modification genes (**Supplemental Data 1BC**). As expected from the literature [17], *trmD* and *tilS* were essential in both species. *IscS1* and *mnmA,* both involved in the synthesis of s^2U34 were also essential in both species, whereas they can be deleted in both *E. coli* and *B. subtilis* [101,102] . Discrepancies between TnSeq data, which often report *mnmA* as essential, including in *E. coli* [103], and targeted deletion studies showing that *mnmA* is frequently dispensable [104] can be explained by the severely impaired growth of *mnmA* mutants [101,102] which leads to their counterselection in pooled populations. It appears that *mnmA* is essential in Streptococci, as directly targeting the genes for deletion in *S. mutans* was unsuccessful (**Supplemental Data 1B).**

The *tadA* gene is essential in *E. coli* but not in organisms such as *B. subtilis* [105,106] or *S. aureus* [Marzi], even if it is not totally clear why the unmodified $tRNA^{Arg}_{ACG}$ can decode all CGN codons in *B. subtilis* and not in *E. coli* [105]. TadA is not essential in *S. mutans,* and the two current studies utilizing two different growth media disagree on the essentiality of the gene in *S. pneumoniae* (**Supplemental Data 1BC**); therefore, additional targeted genetic experiments will be necessary.

If no gene was found to be essential only in *S. mutans*, several genes were essential only in *S. pneumoniae* including *mnmEG* and *tsaBCDE.* MmnEG encoding genes are generally not essential in Streptococci, even if their absence leads to virulence defects [31–34]. We had previously confirmed that t^6A was not essential in *S. mutans* as *tsaB*, *tsaC* and *tsaE* deletions were viable and devoid of t^6A [37]. In contrast, all four t^6A pathway-related genes seem to be essential in *S. pneumoniae* (**Supplemental Data S1C, Fig 5A and S9**).

t^6A synthesis genes are essential in *S. pneumoniae*, but a suppressor mutation in asnS make *tsaE* deletion strains viable.

We confirmed the essentiality of the *tsaE* gene in *S. pneumoniae* by attempting to transform Δ*tsaE* amplicons into the unencapsulated IU1824 (D39 Δ*cps rpsL1*) or IU1945 (D39 Δ*cps rpsL*$^+$) background. No transformants were recovered in the wild-type strains, whereas normal colony formation was observed in a strain expressing *tsaE* ectopically under the control of a zinc-inducible promoter (**Table 1**). WT strains transformed with *tsaE* deletion amplicons did appear after 40 hours of growth, most certainly having acquired suppressor mutations. Growth of Δ*tsaE* clones was affected compared to WT, and the growth defect was complemented by *tsaE* expression in trans (**Fig. 5B and Fig. 6AB**). The Δ*tsaE* cells were slightly smaller but the same

shape as WT cells (**Fig. 6CD**). Whole genome sequencing of four independent Δ*tsaE* strains obtained from two different genetic backgrounds and two deletion markers revealed that all of them had acquired a mutation [*asnS*(D121A)] encoding asparaginyl-tRNA synthetase (AsnRS) (**Table 2**). To confirm that the presence of *asnS*(D121A) mutation can suppress the lethality of Δ*tsaE,* we constructed an isogenic set of strains expressing $asnS^+$ in the native chromosomal locus and an ectopic copy of either $asnS^+$ or *asnS*(D121A). The presence of an ectopic copy of *asnS*(D121A), but not WT *asnS* allows the growth of a Δ*tsaE* strain **(Table 2 and Fig. 6B)**. This result indicates that *asnS*(D121A) is a gain-of-function mutation, consistent with the structural model described below (**Fig. 7**). Analysis of t^6A levels in these strains by LC-MS (**Fig. 6CD**) and tRNA seq (GLORI) (**Fig. 2DE**) showed that t^6A was indeed absent in the Δ*tsaE* strains.

Profiling of the other rRNA/tRNA modifications by LC/MS showed that relative levels of all predicted rRNA modifications increased by 50% while all tRNA modifications (other than t^6A) decreased by 20% (**Supplemental data 2C** and **Fig. 6D**). Based on well-established mechanistic links between cell stress and rRNA degradation [107,108], we believe the growth rate defect of the *tsaE* mutant strains leads to rRNA degradation and contamination of the tRNA preparations, which explains the changes in modification profiles.

Structural modeling suggests a role for t^6A37 in tRNA/AsnRS interactions.

To gain structural insight into the role of Asp121 in *S. pneumoniae* AsnRS (*Sp*AsnRS) function, we constructed a model of its complex with t^6A-modified $tRNA^{Asn}$ (**Fig. 7**). Although crystal structures of AsnRS from several organisms are available, none include bound tRNA. The only available experimental structure of $tRNA^{Asn}$ is that of *Pseudomonas aeruginosa* $tRNA^{Asn}$ (*Pa*$tRNA^{Asn}$) in complex with non-discriminating AspRS (*Pa*ND-AspRS, PDB ID 4WJ4 [109]), but this tRNA lacks the t^6A37 modification. By contrast, recent high-resolution cryo-EM structures of human LysRS bound to ms^2t^6A37-modified human $tRNA^{Lys3}$ (PDB IDs 9DPL and 9DPA [110]) provide insight into how a class IIb synthetase accommodates t^6A37 in the anticodon loop and remodels the loop for recognition.

We used the AlphaFold2 model of *Sp*AsnRS (*Sp*AsnRS) [111], which closely resembles AsnRS crystal structures from diverse organisms, including *Thermus thermophilus* (PDB ID 5ZG8), *Elizabethkingia sp.* (PDB ID 6PQH, *Pyrococcus horikoshii* (PDB ID 1X54 [112]), *Brugia malayi* (PDB ID 2XGT [113]), *Entamoeba histolytica* (PDB ID 3M4Q) and *Homo sapiens* (PDB ID 8H53 [114]). The *Sp*AsnRS model superposes with these orthologs with r.m.s.d. values of 0.945-1.564 Å over 294-333 C_α atoms and aligns well with other class IIb synthetases including

E. coli AspRS (r.m.s.d. 2.23 Å over 269 C_α atoms) and *E. coli* LysRS (r.m.s.d 1.5 Å over 244 C_α atoms). *Pa*tRNA^{Asn} was positioned onto *Sp*AsnRS by superposing *Pa*ND-AspRS bound to tRNA^{Asn} with *Sp*AsnRS (r.m.s.d. 1.8 Å over 261 C_α atoms, 25% sequence identity, **Fig. S10A**). In parallel, the cryo-EM structure of human LysRS/tRNA^{Lys3} was superposed onto *Sp*AsnRS to inform the placement and orientation of the anticodon-binding domain relative to a t^{6}A37-modified anticodon loop (r.m.s.d 5.3 Å over 311 C_α atoms, 18% sequence identity, **Fig. S10B**). Together, these superpositions yielded a composite model of the *Sp*AsnRS/tRNA^{Asn} that integrates the conserved fold of class IIb synthetases with structural evidence for anticodon remodeling and t^{6}A37 accommodation (**Fig. 7A**).

AsnRS, AspRS and LysRS-II are all class IIb synthetases that act on ANN-decoding tRNAs, substrates of the t^{6}A pathway. They share a conserved architecture comprising an N-terminal anticodon binding domain, a C-terminal catalytic domain, and an intervening 25-30 residue linker/hinge region formed by a flexible, proline-rich loop and two short α-helices. This linker articulates the catalytic and anticodon-binding domains, making direct tRNA contacts and positioning the acceptor end in the active site through interactions with the acceptor arm and the D-stem minor groove [110,115–117]. Our model suggests that these interactions are also possible in *Sp*AsnRS/tRNA^{Asn}. For example, Pro112 and His118 in *Sp*AsnRS mimic the D-stem contacts made by Pro214 and Lys221 in human LysRS, respectively (**Fig. S10B**). In addition, Asp121 in *Sp*AsnRS aligns with Glu224 in human LysRS and Glu122 in *Pa*ND-AspRS, each observed to form a hydrogen bond with the 2′-OH of ribose 67 at the base of the acceptor stem, a contact that is largely conserved among class IIb synthetases, consistent with the conservation of this acidic residue across the family (**Fig. 7B**).

Linker residues in LysRS engage the modified base ms^{2}t^{6}A37 in the anticodon loop: Phe218 stacks against, and His217 hydrogen bonds with, the threonylcarbamoyl group of ms^{2}t^{6}A37 (Fig. S10B). Our model places *Sp*AsnRS residues Lys116 and Pro115 in analogous positions, suggesting similar interactions with t^{6}A37 in t^{6}A-modified tRNA^{Asn}. When A37 is unmodified, this anticodon-loop contact is lost. Combined with the Asp121Ala substitution, which disrupts the acceptor-stem contact, the linker may be released to adopt an alternative conformation that maintains D-stem interactions. This flexibility provides a plausible route for *Sp*AsnRS to function with unmodified tRNA^{Asn}.

Based on the above structural analysis and the suppression phenotypes of strains carrying *asnS*(D121A) mutation, we hypothesize the following suppressor mechanism. t^{6}A at position 37

of tRNA acts as a strong positive determinant with WT SpAsnRS, similar to the case of t^6A and *E. coli* IleRS [43]. WT SpAsnRS can form a productive complex with t^6A-modified tRNA, but not with unmodified tRNA. Since t^6A-modified $tRNA^{Asn}$ is absent in a *tsaE* deletion mutant, the lack of Asn-charging leads to lethality of Δ*tsaE* cells. We hypothesize that *Sp*AsnRS(D121A) encoded by the suppressor mutation *asnS*(D121A) allows the interaction of the synthetase with unmodified tRNA, leading to Asn charging for translation. However, this reaction is less efficient than that carried out by WT AsnRS with modified tRNA, resulting in slower growth in the suppressed strain. This model is consistent with the result showing *asnS*(D121A) as a gain-of-function mutation. The presence of AsnRS with this amino acid change is sufficient to allow cell growth, independent of the presence or absence of WT AsnRS.

Discussion

It is also well established that pathogenic bacteria tend to streamline their genomes compared to their environmental ancestors [118]. Accordingly, *S. mutans* and *S. pneumoniae* retain 87% (39/45) and 78% (35/45), respectively, of the tRNA modification genes found in *B. subtilis (***Fig. 1** and **Supplemental Data 1ABC***)*. What was unexpected, however, was that the majority of the lost genes encode Fe-S cluster –requiring enzymes. Specifically, 80% of the missing genes in *S. mutans* and 60% in *S. pneumoniae* fall into this category, including *mnmL*, *queG*, *queE*, *mtaB*, *miaB*, and *tcdA* (**Fig. 1** and **Supplemental data 1ABC**). Previous studies had already noted that many widespread Fe–S-dependent metabolic proteins are absent in *S. mutans* and in other Gram-positive species [119]. In addition, the SUF system responsible for Fe–S cluster assembly is not essential in *S. mutans*, unlike in *B. subtilis*, because the essential Fe–S enzymes of isoprenoid synthesis (IspH and IspG) have been replaced in *S. mutans* by the Fe-S independent mevalonate pathway [120]. Hence, we see that the selective pressure in both these organisms has led to shedding hyper-modifications that require Fe–S enzymes, while still maintaining modifications in the same positions. *S. mutans* is exposed to a variety of reactive oxygen species (ROS) in the oral cavity, such as H_2O_2 production by niche-competing early colonizers of dental plaque biofilm (*S. gordonii*, *S. sanguinus*) [121,122], periodontal inflammation [123], hypothiocyanite production by host salivary lactoperoxidase [124], and use of H_2O_2-containing dental care products. Given that Fe-S clusters are an important source of hydroxyl radical generation via Fenton chemistry during exposure to oxidative stress, it makes sense for *S. mutans* to minimize its reliance on Fe-S containing enzymes required for central physiological functions such as tRNA modification. One such specific example is the Q

synthesis pathway. *S. pneumoniae* has lost the two Fe-S dependent enzymes of this pathway: QueE [125], involved in the preQ$_0$ precursor synthesis, and QueG, involved in the last step, the reduction of oQ. However, we show that the pathway has not been lost and oQ accumulates in *S. pneumoniae* tRNA. Even if the Q modification has been repeatedly and independently lost in many clades [38], all the circa 1599 Streptococci genomes analyzed are predicted to still carry Q, or the Q precursor oQ, in target tRNAs, reinforcing its critical role in translation accuracy in this clade. This comparative work emphasizes the difficulty in predicting active pathways through metabolic reconstruction. Indeed, as we show here, pseudogenes can create false positive hits that can skew the pathway predictions: in this case, we predicted preQ$_0$ salvage in *S. pneumoniae* when in reality, preQ$_0$ is salvaged in this organism because the QueF homolog is not functional (**Fig. 3**).

The translation machinery is a classical target for antibacterial compounds with many known antibiotics targeting one aspect of translation or another (such as aminoacyl-tRNA synthetases, ribosome, translation factors)[126]. One aspect of translation that has not been targeted is tRNA modification, for two main reasons. First, it is only in the last decade that many of the genes/enzymes involved in tRNA modification have been identified at least in one model organism from each kingdom [6]. Second, only a handful of modification genes are essential in most pathogenic bacteria. These are the genes that introduce m^{1}G37, k^{2}C34, and t^{6}A37. TrmD, the m^{1}G37 methylase, was the first to be identified as an antibacterial target because it is very different from its human counterpart (Trm5), and inhibitors have already been identified [18,127]. TilS, the enzyme that introduces the k^{2}C modification, is an obvious target as it is absent in eukaryotes [128]. TilS and TrmD are both essential in both *S. mutans* and *S. pneumoniae* (**Supplemental data S1BC**), hence any antibiotic targeting these enzymes should inhibit the growth of both these pathogens. The situation is not as clear-cut for t^{6}A synthesis.

The differences between the bacterial and eukaryotic t^{6}A synthesis machinery open the possibility of targeting this pathway for antibacterial compounds [41,42]. While the first enzyme of the pathway that produces threonylcarbamoyl adenylate, (TsaC in *E. coli*, Tsc2/Sua5 in yeast, YRDC in human) is very conserved across kingdoms [76], the second step of the pathway, the transfer of the threonylcarbamoyl moiety to the target adenosine, varies greatly (see [43] for review), with only one protein family (TsaD/Kae1/Qri7) in common between the different pathways. Mitochondria just require Qri7 to complete t^{6}A synthesis, whereas Bacteria require TsaB and TsaE in addition to TsaD, and the whole KEOPS complex catalyzes this reaction in all eukaryotes [41,42]. Because of their prokaryotic-specific essentiality, the *tsaBE* genes had been

identified as potential antibacterial targets before their role in t^6A synthesis was even established [129–134] and inhibitors of TsaE were developed based on its ATP-binding capabilities [135]. TsaE is essential in many bacterial pathogens, including nearly all ESKAPE pathogens (**Supplemental Data S1J**); however, it can be deleted in some pathogenic bacteria, such as *S. mutans*. We showed here that *tsaE* is essential in *S. pneumoniae* but a suppressor mutation in the *asnS* gene encoding AsnRS was readily selected, suggesting a role for t^6A in AsnRS function in this organism. The mutation results in a D121A substitution in the coding sequence, which could increase flexibility of a linker region and compensate for the loss of anticodon loop interactions that normally depend on t^6A37. In contrast, *S. mutans* AsnRS (*Sm*AsnRS) retains the wild-type sequence in a *ΔtsaE* strain lacking t^6A, indicating that t^6A37 is not critical for *Sm*AsnRS function. D121 is conserved in *Sm*AsnRS, and the protein shares 80% sequence identity with *Sp*AsnRS, including in the linker region (**Fig. 7B**), with predicted 3D structures that are virtually identical. Notably, *S. pneumoniae* and *S. mutans* $tRNA^{Asn}$ molecules differ by a single base pair (C5:G69 versus G5:C69, respectively; **Fig. S10C**), located adjacent to C67 in our docking model and near the contact point for D121. This difference may influence local geometry at the base of the acceptor stem in ways not evident in current structures, but the structural basis for the different requirements for t^6A37 remains unclear. It is worth noting that in our study only one unique *asnS* mutation within the linker region was obtained among four independent mutants. In comparison, evolution repair studies of *E. coli* growth of *trmD* mutant strains show that growth of *trmD* mutants can be largely restored by single mutations in *proS* that restore aminoacylation of G37-unmodified $tRNA^{Pro}$ [136]. However, the locations of the *proS* mutations were not restricted to any specific domains.

This study once again highlights the importance of integrating analytical, computational, and genetic approaches to accurately define tRNA modification profiles in a given organism [17]. Encouragingly, growing recognition of the biological significance of RNA modifications—combined with improved access to analytical technologies—has led to a surge of studies across taxonomically diverse species [21,22,27,93,137–139]. These datasets now can serve as valuable anchors for bioinformatic and machine learning approaches to predict modification pathways in newly sequenced organisms.

Data accessibility

Raw LC-MS/MS data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE (http://www.ebi.ac.uk/pride/) partner repository with the dataset identifier PXD069173. Raw AlkAnilineSeq and GLORI data

are available at the European Nucleotide Archive (https://www.ebi.ac.uk/ena/browser/home) under the accession number PRJEB96333.

Declaration of AI use

Grammarly (https://www.grammarly.com) and ChatGPT version 4.o (https://chatgpt.com/) were used to edit the grammar and clarity of specific sentences.

Acknowledgments

We thank Patricia Dos Santos and Frederic Barras for their helpful discussions, and Stefano Marzi for sharing unpublished data and providing input on the manuscript.

Funding

Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Numbers R01GM70641 and R35GM156215 to VdC-L, R01GM110588 to MAS, R01ES026856 to PCD, and R35GM131767 to MEW. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. CKC is supported by a Human Frontier Science Program Cross-Disciplinary Fellowship. PCD, CKC, JS, and GS were supported by the National Research Foundation of Singapore through the Singapore-MIT Alliance for Research and Technology Antimicrobial Resistance Interdisciplinary Research Group.

Table 1. Transformation of Δ*tsaE* amplicons into different backgrounds confirms *tsaE* essentiality in *S. pneumoniae* D39 and suppression by *asnS*[D121A] mutation[a]

Recipient strain and condition	Recipient genotype	Amplicon	Number of colonies at 22 h after transformation with deletion amplicons[b]
IU1824 - Zn	WT	Δ*tsaE*::P_c*erm*	0 colony[c,d,e]
IU1824 - Zn		Δ*tsaE*::P_c-*kan-rpsL*	0 colony[c,d,f]
IU1824 - Zn		Δ*bgaA::*P_c*erm* (control)	>500 healthy colonies
IU18816 - Zn	*tsaE*$^+$// Δ*bgaA*::P_{Zn}-*tsaE*	Δ*tsaE*::P_c*erm*	>500 small colonies[g]
IU18816 + 0.2 mM Zn			>500 healthy colonies
IU18886 - Zn	Δ*tsaE*::Pc-*kan-rpsL*// Δ*bgaA*::P_{Zn}-*tsaE*	Δ*tsaE* markerless	0 colony
IU18886 + 0.1 mM Zn			>500 healthy colonies
IU19598	*asnS*$^+$ Δ*bgaA*::P_{ftsA}-*asnS*[D121A]	Δ*tsaE*::P_c*erm*	>500 small colonies[g]
IU19599	*asnS*$^+$ Δ*bgaA*::P_{ftsA}-*asnS*$^+$	Δ*tsaE*::P_c*erm*	Faint colonies in agar

[a] Recipient strains in D39 Δ*cps rpsL1* (IU1824) background and amplicons were obtained as described in **Table S1.**

[b] Transformations and visualization of colonies normalized to 1 mL of transformation mixture were performed as described in *Materials and Methods*. Colony appearance is compared to colonies obtained with WT strain transformed with a positive control amplicon, Δ*bgaA::*P_c*erm.* Numbers of colonies listed are normalized to 1 mL of transformation mix. Transformants were confirmed by PCR reactions. Each transformation experiment was performed 2 or more times independently with similar results.

[c] Occasional suppressor mutants were present.

[d] Colonies remained very small, but uniform-sized upon re-streaking on antibiotic selection plates.

[e] The suppressor strain was stored as IU18824 and whole genome sequenced.

[f] The suppressor strain was stored as IU18826 and whole genome sequenced.

[g] Small colonies are obtained with transformation of the merodiploid strains (IU18816 and IU18886) under non-inducing (no Zn) conditions, probably because of the slight expression of *tsaE* from the P_{zn} promoter.

Table 2. *tsaE* suppressor mutants of *S. pneumoniae* D39 contain *asnS*(D121A)

					Suppressor strains[c]			
Genome position[b]	Mutation	Annotation	Gene	Description	IU18824 (sup1)	IU18826[d] (sup2)	E797 (sup3)	K797 (sup4)
1,384,280	T→G	D121A (GAC→GCC)	*asnS* ←	asparaginyl-tRNA synthetase	100%[e]	100%[e]	100%[e]	100%[e]

(asparaginyl-tRNA synthetase) mutation[a]

[a]*S. pneumoniae* has the highest mutation rate among all studied bacteria with functional DNA repair systems [140] and readily accumulates suppressor mutations of essential genes[69,72,75,141].

[b]Whole-genome sequencing performed as described in Materials and Methods was used to identify suppressor mutations and to verify the genomes of constructed mutants. Genomic DNA preparation, DNA library construction, Illumina MiSeq DNA sequencing, and bioinformatics analyses were performed as described in Materials and Methods.

[c]Strains IU18824, IU18826, E797 and K797 (Table S1) containing suppressor mutations that allowed growth of a Δ*tsaE* mutant were isolated as described in Results.

[d]IU18826 (sup2) has additional R191H (CGT→CAT) mutation in spd_0982 pyrophosphokinase family protein at position 995,156.

[e] Percentage of reads containing the indicated mutation

Figure legends

Figure 1. The comparison of tRNA modifications and corresponding enzymes among *B. subtilis*, *S. mutans*, and *S. pneumoniae*. Modified positions are colored blue when the modification and enzymes are conserved and orthologous in three strains; green when they differ between *Bacillus* and *Streptococcus;* and red when they differ between *S. mutans* and *S. pneumoniae*. Evidence code: M, modification reported; G, gene coding modifying protein or an ortholog of an existing modifying protein; E, experiments that validated modifying gene(s) of the species via in vivo gene knockout or in vitro expression; O, ortholog that was validated experimentally in a different species. References for all proteins listed are given in Supplemental data S1D.

Figure 2. ***S. pneumoniae* tRNA modifications mapped by sequencing-based modification detection methods.** (A) Alkaline Seq analysis of tRNA extracted from WT and Δ*dusB1 S. pneumoniae* Representative AlkAnilineSeq profiles showing signals for D (at positions 17, 20, 20a) and m^7G (at position 46) in the WT and Δ*dusB1* strains. Ncleavage score is shown in the same scale, nucleotide color code and strain identity are shown on the right. AlkAniline-Seq signals are measures for biological triplicate (n=3) and the average value is shown, along with the error bar (calculated using the R ggplot2 package). Individual points for replicates are shown in grey. tRNA residues are numbered in sequential order, so position number may not correspond to conventional tRNA numbering (including 17a, 20a/b, and variable loop); (B) Heatmap for absolute Ncleavage values observed upon AlkAnilineSeq analysis of most modified D sites mapped in S. pneumoniae WT and *ΔdusB1* strains. The color code is shown at the bottom right, and the identity of the D site in tRNA is indicated. Data are shown for all replicates (n=3) for WT and ΔDusB strains; (C) Summary of all AlkAniline-Seq -detected modified residues in *S. pneumoniae* tRNAs. m1A residues are shown in orange, m^7G are in green. Identity of tRNAs harboring these modifications is shown; (D) Representative GLORI profiles detecting m^1A and i^6A residues in *S. pneumoniae* tRNA. The signal in GLORI corresponds to non-deaminated (NO_2-resistant) A residues (m^1A and N6-substuituted A residues). Data for WT and *ΔtsaE* strain are shown. Nucleotide color code is shown on the right. tRNA residues are numbered in sequential order, so position number may not correspond to conventional tRNA numbering (including 17a, 20a/b, and variable loop); (E) Heatmap of the GLORI signals for all detectable modifications (m^1A, m^6A, i^6A and t^6A). The color code is shown on the top. Identity of tRNA modified site is indicated on the right. (F) Summary of all GLORI-detected modified residues in *S. pneumoniae* tRNAs. m^1A residues are shown in light blue, N6-substituted A37

residues are in violet. A*37 indicates N6-substituted A which occurs in the sequence different from contexts AAA (for i^6A) and UAA (for t^6A), most likely these tRNAs harbor m^6A modification. Identity of tRNAs harboring these modifications is shown.

Figure 3. Decay of Q pathway in *S. pneumoniae* and other Streptococci. (A) Predicted Q pathway in *S. pneumoniae* D39, with protein names shown in blue and locus tags in red. The ECF transporters include 4 subunits: S, the substrate-specific transmembrane component (QueT); T, the energy-coupling module; A and A', a pair of ATPase; (B) *S. pneumoniae* salvages $preQ_1$ to synthesize oQ in tRNAs. The wild-type strain IU1824 was cultured in BHI medium supplemented with DMSO, 100 nM $preQ_0$ base, or 100 nM $preQ_1$ base. tRNAs were isolated, enzymatically digested, and analyzed by LC-MS/MS as described in the Materials and Methods section. For each analysis, 600 ng of hydrolyzed tRNAs was injected. The identities of the $preQ_1$ and Q peaks were confirmed by comparing them to the retention times of their synthetic standards. The identity of the oQ peak was confirmed through high-resolution LC-MS analysis, as detailed in the Materials and Methods and Results sections. Relative abundance for each modified ribonucleoside was calculated by normalizing its raw peak area to the sum of UV signals from the four canonical ribonucleosides. Reported averages and standard deviations are based on data from three independent growth experiments; (C) The distribution of different pathways of 7-deazaguanine modification in tRNA in 1,599 *Streptococcus*. The four major pathways are *de novo* Q biosynthesis (red), $preQ_0$ salvage Q modification (blue), $preQ_1$ salvage Q modification (green), and $preQ_1$ salvage oQ modification (yellow), with the others (gray) consisting of Q gene patterns that were found in no more than two strains; (D) The Q gene distribution for each category. The Q biosynthesis and Q precursor transportation is depicted in the top panel. Each arrow represents one reaction catalyzed by the enzyme underneath. Dashed arrows represent Q precursor transmembrane salvage via transporter proteins. The presence and absence of each Q gene is indicated by the corresponding filled and open square. The presence of a pseudogene is indicated by a half-shaded square; (E) The composition (number of strains) of species for each type of Q gene pattern. For better visualization, less abundant species collapsed.

Figure 4. Two types of QueF proteins are found in Streptococcus. (A) The MSA of 116 unique QueF or QueCF from 544 Streptococcus strains. The boxed region indicates the N-terminus that do not align; (B) The corresponding Q genes and species of 544 QueF encoding

Streptococcus strains. RefSeq accessions for all proteins listed are given in **Supplemental Data S1F.**

Figure 5. TsaE is essential in *S.pneumoniae* D39. (A) Mini-Mariner *Malgellan6* Tn-Seq transposon insertion profile for the genome region covering *spd_1740* (*cinA*), *spd_1741* (*lytR*), *spd1742*, *spd_1743* (*tsaE*) and *spd_1744* (*comM*) in the genomes of the unencapsulated WT parent (D39 Δ*cps rpsL1*, IU1824) strain growing exponentially in BHI broth in 5% CO_2. The same WT Tn-seq insertion profile was obtained for encapsulated D39 strain IU1781 grown in BHI broth (data not shown). *In vitro* transposition reactions containing purified genomic DNA, *Magellan6* plasmid DNA, and purified MarC9 mariner transposase, transformation, harvesting of transposon-inserted mutants, growth of pooled insertion libraries exponentially in BHI broth, NextSeq 75 high-output sequencing, and analysis were performed as described in *Materials and Methods* based on [73]. Tn-insertions were recovered in the regions encoding *spd_1740*, *spd_1742* and *spd_1744*, but not in *spd_1741*(*lytR*) and *spd_1743* (*tsaE*). (B) Representative growth curves of the WT parent (IU1824), Δ*tsaE (*sup1*)* (IU18824), and Δ*tsaE* complementation strain (IU18963, Δ*tsaE* markerless//Δ*bgaA::*P_{zn}-*tsaE*). *tsaE* deletion strains containing suppressor mutation *asnS*(D121A) are viable but have higher doubling time and lower yield in BHI growth media, and the growth phenotypes are complemented by ectopic *tsaE* expression. IU18824 was whole-genome-sequenced to contain *asnS*(D121A) mutation. IU1824 and IU18824 were grown overnight in BHI broth with no additional Zn^{2+}/Mn^{2+} addition. IU18963 was grown overnight and day of experiment in BHI broth with additional 0.1 mM (Zn^{2+}/(1/10) Mn^{2+}) to induce expression of *tsaE* from the ectopic *bgaA* site. Overnight cultures were diluted to OD_{620} ≈0.003 in the morning in BHI broth for IU1824 and IU18824 and in BHI broth containing 0.1 mM Zn^{2+}/(1/10)Mn^{2+} for IU18963. Doubling times (DT) and maximal growth yields (OD_{620}) (averages ± SE) were obtained during 9 h of growth from 2 independent experiments. (C) and (D) Δ*tsaE* cells with an *asnS* mutations are smaller but with the same shape as WT strain, and the Δ*tsaE* cell morphology phenotype is complemented by ectopic expression of *tsaE*$^+$ (C) Representative phase-contrast images taken at mid-log growth between 3 to 3.5 h for IU1824 and IU18963 (OD_{620} between 0.1 to 0.2), and between 3.5 to 4.0 h for IU18824 (OD_{620} between 0.03 to 0.05). Scale bar = 1 µm. (D) Scatter plots of cell lengths, widths, aspect ratios (cell length to width) and relative cell volumes of the above strains. Lines represent median and 25 and 75 percentiles. IU18963 was grown with 0.1 mM Zn^{2+} (Zn^{2+}/(1/10) Mn^{2+}) shown in (B). Pvalues were obtained by one-way ANOVA analysis (GraphPad Prism, Kruskal-Wallis test). **** and ns denote $p<0.0001$, and not significant, respectively when compared to WT.

Figure 6. Suppressor mutations in *asnS* gene allow growth of *S. pneumoniae* D39 t^6A-deficent strain. (A) *tsaA* deletion strains containing suppressor mutation *asnS*(D121A) are viable but have longer doubling time and lower yield in BHI growth media. Left panel, representative growth curves of the WT parent (IU1824, D39W *rpsL1* Δ*cps*) and mutants derived from this background, Δ*tsaE (*sup1*)* (IU18824, IU1824 Δ*tsaE::*P_c*-erm* sup1), and Δ*tsaE (*sup2*)* (IU18826, IU1824 Δ*tsaE::*P_c*-kan-rpsL* sup2). Right panel, representative growth curves of the WT parent (IU1945, D39W *rpsL*$^+$ Δ*cps*) and mutants derived from this background Δ*tsaE (*sup3*),* (E797, IU1945 Δ*tsaE::*P_c*-erm* sup3), and Δ*tsaE (*sup4*)* (K797, IU1945 Δ*tsaE::*P_c*-kan-rpsL* sup4). IU18824, IU18826, E797 and K797 were whole-genome-sequenced to contain *asnS*(D121A) mutation. Overnight cultures grown in BHI media were diluted to OD_{620} ≈0.003 in the morning in BHI broth. Doubling times (DT) and maximal growth yields (OD_{620}) (averages ± SE) were obtained during 9h of growth from one experiment. (B) Representative growth curves of IU1690, IU1824, IU19598, IU18824, and IU19699. Expression of *asnS*(D121A) in the native genetic locus or at an ectopic site allows growth of *tsaE* deletion strains. IU1690 is D39W *cps*$^+$ reference strain. IU1824 is D39 Δ*cps rpsL1*, the WT parent strain for IU19598, IU18824, and IU19699. IU19598 is IU1824 Δ*bgaA*::P-*asnS* (D121A), harboring ectopic expression of *asnS*(D121A). IU18824 is IU1824 Δ*tsaE*::P_c*-erm* sup 1 (with suppressor mutation *asnS* [D121A]). IU19699 is IU1824 Δ*bgaA*::P-*asnS* (D121A) Δ*tsaE::*P_c*-erm.* This strain is reconstructed to confirm *asnS* (D121) suppression of Δ*tsaE*::P_c*-erm*. (C) Relative abundance of t^6A in tRNAs isolated from IU1690, IU1824, IU19598, IU18824, and IU19699. tRNAs were isolated, enzymatically digested, and analyzed by LC-MS/MS as described in Methods section. For each analysis, 600 ng of hydrolyzed tRNAs was injected. The identity of the t^6A peak was confirmed by comparing to the retention time of its synthetic standard. Relative abundance of t^6A was calculated by normalizing its raw peak area to the sum of UV signals from the four canonical ribonucleosides. Reported averages and standard deviations are based on data from two independent growth experiments. Differences in t^6A levels among the five strains were assessed using a one-way analysis of variance (ANOVA) at a significance level of $p < 0.05$, followed by Tukey's post-hoc test for multiple comparisons. Statistically significant differences were indicated by ****$p < 0.0001$. Statistical analysis was performed using GraphPad Prism 9. (D) Changes in relative abundance of modified ribonucleosides in tRNAs isolated from IU1690, IU1824, IU19598, IU18824, and IU19699. Each grid cell represents the mean fold-change for each modified ribonucleoside in IU1690, IU19598, IU18824, and IU19699 relative to IU1824, calculated from two biological replicates. An exception applies to mnm^5s^2U, for which fold-changes were calculated relative to IU1690. Grey cells indicate LC-MS/MS signals below limit of

detection. Hierarchical clustering analysis was performed using Morpheus (https://software.broadinstitute.org/morpheus/), employing the Euclidean distance metric and the average linkage method to group both strains and modified ribonucleosides.

Figure 7. Structural model of AsnRS/tRNA^{Asn} complex. (A) model of t^{6}A-modified tRNA^{Asn} bound to *Streptococcus pneumoniae* asparaginyl-tRNA synthetase (*Sp*AsnRS, purple), highlighting the inter-domain linker region (tan), potential interactions of linker residues with C67 in the base of the tRNA acceptor stem and with t^{6}A37. The threonylcarbamoyl moiety of t^{6}A37 is highlighted in cyan. (B) Structure-based multi-sequence alignment of AsnRS, ND-AspRS, and LysRS in the linker region (tan highlight). The conserved Asp/Glu corresponding to Asp121 of *Sp*AsnRS is indicated with a red arrow. Residues putatively involved in D-stem and acceptor arm are indicated with black and blue stars, respectively. PDB IDs for the structures used in generating the alignment are indicated in the sequence headers. AF: AlphaFold model**.**

References

1. Lannes-Costa PS, de Oliveira JSS, da Silva Santos G, Nagao PE. 2021 A current review of pathogenicity determinants of Streptococcus sp. *J Appl Microbiol*. **131**, 1600-1620. (doi:10.1111/jam.15090)

2. Walker MJ, Barnett TC, McArthur JD, Cole JN, Gillen CM, Henningham A, Sriprakash KS, Sanderson-Smith ML, Nizet V. 2014 Disease manifestations and pathogenic mechanisms of group A Streptococcus. *Clin Microbiol Rev* **27**, 264-301. (doi:10.1128/CMR.00101-13)

3. Raabe VN, Shane AL. 2019 Group B Streptococcus (*Streptococcus agalactiae*). *Microbiol Spectr* **7**, 10.1128/microbiolspec.gpp3-0007-2018. (doi:10.1128/microbiolspec.gpp3-0007-2018)

4. Bloch S, Hager-Mair FF, Andrukhov O, Schäffer C. 2024 Oral streptococci: modulators of health and disease. *Front Cell Infect Microbiol*. **14**, 1357631. (doi:10.3389/fcimb.2024.1357631)

5. Lewis ML, Surewaard BGJ. 2018 Neutrophil evasion strategies by *Streptococcus pneumoniae* and *Staphylococcus aureus*. *Cell Tissue Res*. **371**, 89-503. (doi:10.1007/s00441-017-2737-2)

6. El Yacoubi B, Bailly M, de Crécy-Lagard V. 2012 Biosynthesis and function of posttranscriptional modifications of transfer RNAs. *Annu Rev Genet* **46**, 69–95. (doi:10.1146/annurev-genet-110711-155641)

7. Grosjean H, de Crécy-Lagard V, Marck C. 2010 Deciphering synonymous codons in the three domains of life: co-evolution with specific tRNA modification enzymes. *FEBS Lett* **584**, 252.

8. Yared MJ, Marcelot A, Barraud P. 2024 Beyond the Anticodon: tRNA core modifications and their impact on structure, translation and stress adaptation. *Genes (Basel)*. **15**, 374. (doi:10.3390/genes15030374)

9. Fruchard L *et al.* 2024 Aminoglycoside tolerance in *Vibrio cholerae* engages translational reprogramming associated with queuosine tRNA modification. *Elife*, **13**, RP96317. (doi:10.1101/2022.09.26.509455)

10. Babosan A, Fruchard L, Krin E, Carvalho A, Mazel D, Baharoglu Z. 2022 Nonessential tRNA and rRNA modifications impact the bacterial response to sub-MIC antibiotic stress. *microLife* **3**, uqac019. (doi:10.1093/femsml/uqac019)

11. Frommeyer YN *et al.* 2025 tRNA hydroxylation is an epitranscriptomic modulator of metabolic states affecting *Pseudomonas aeruginosa* pathogenicity. *Nucleic Acids Res* **53**, gkaf719. (doi:10.1093/nar/gkaf719)

12. Durand JM, Dagberg B, Uhlin BE, Björk GR. 2000 Transfer RNA modification, temperature and DNA superhelicity have a common target in the regulatory network of the virulence of *Shigella flexneri:* the expression of the *virF* gene. *Mol Microbiol* **35**, 924–935.

13. Sato Y, Takita A, Suzue K, Hashimoto Y, Hiramoto S, Murakami M, Tomita H, Hirakawa H. 2024 TusDCB, a sulfur transferase complex involved in tRNA modification, contributes to UPEC pathogenicity. *Sci Rep* **14**, 8978. (doi:10.1038/s41598-024-59614-2)

14. Antoine L, Bahena-Ceron R, Devi Bunwaree H, Gobry M, Loegler V, Romby P, Marzi S. 2021 RNA modifications in pathogenic bacteria: impact on host adaptation and virulence. *Genes (Basel)* **12**, 1125. (doi:10.3390/genes12081125)

15. Kouvela A, Zaravinos A, Stamatopoulou V. 2021 Adaptor molecules epitranscriptome reprograms bacterial pathogenicity. *Int J Mol Sci* **22**, 8409. (doi:10.3390/ijms22168409)

16. Koh CS, Sarin LP. 2018 Transfer RNA modification and infection – Implications for pathogenicity and host responses. *Biochim Biophysa Acta Gene Reg Mech* **1861**, 419–432. (doi:10.1016/j.bbagrm.2018.01.015)

17. de Crécy-Lagard V, Jaroch M. 2021 Functions of bacterial tRNA modifications: from ubiquity to diversity. *Trends Microbiol* **29**, 41–53. (doi:10.1016/j.tim.2020.06.010)

18. Zhong W *et al.* 2019 Thienopyrimidinone derivatives that inhibit bacterial tRNA (guanine37-N(1))-methyltransferase (TrmD) by restructuring the active site with a tyrosine-flipping mechanism. *J Med Chem* **62**, 7788–7805. (doi:10.1021/acs.jmedchem.9b00582)

19. de Crécy-Lagard V, Marck C, Brochier-Armanet C, Grosjean H. 2007 Comparative RNomics and Modomics in Mollicutes: prediction of gene function and evolutionary implications. *IUBMB Life* **59**, 634–58.

20. de Crécy-Lagard V, Ross RL, Jaroch M, Marchand V, Eisenhart C, Brégeon D, Motorin Y, Limbach PA. 2020 Survey and validation of tRNA modifications and their corresponding genes in *Bacillus subtilis* sp subtilis strain 168. *Biomolecules* **10**, 977. (doi:10.3390/biom10070977)

21. Tomasi FG, Kimura S, Rubin EJ, Waldor MK. 2023 A tRNA modification in Mycobacterium tuberculosis facilitates optimal intracellular growth. *Elife* **12**, RP87146. (doi:10.7554/eLife.87146)

22. Koshla O, Vogt LM, Rydkin O, Sehin Y, Ostash I, Helm M, Ostash B. 2023 Landscape of post-transcriptional tRNA modifications in *Streptomyces albidoflavus* J1074 as portrayed by mass spectrometry and genomic data mining. *J Bacteriol* **205**, e0029422. (doi:10.1128/jb.00294-22)

23. Sudol C *et al.* 2024 Functional redundancy in tRNA dihydrouridylation. *Nucleic Acids Res* **52**, 5880–5894. (doi:10.1093/nar/gkae325)

24. Jaroch M *et al.* 2024 Alternate routes to mnm^5s^2U synthesis in Gram-positive bacteria. *J Bacteriol* **206**, e0045223. (doi:10.1128/jb.00452-23)

25. Bou-Nader C, Montémont H, Guérineau V, Jean-Jean O, Brégeon D, Hamdane D. 2018 Unveiling structural and functional divergences of bacterial tRNA dihydrouridine synthases: perspectives on the evolution scenario. *Nucleic Acids Res* **46**, 1386–1394. (doi:10.1093/nar/gkx1294)

26. Faivre B *et al.* 2021 Dihydrouridine synthesis in tRNAs is under reductive evolution in Mollicutes. *RNA Biol* **18**, 2278-2289. (doi:10.1080/15476286.2021.1899653)

27. Jaramillo-Ponce JR *et al.* 2025 Complete landscape of post-transcriptional modifications in individual tRNAs from *Staphylococcus aureus*. bioRxiv 2025.10.30.685614; (doi:.1101/2025.10.30.685614)

28. Moukadiri I, Villarroya M, Benítez-Páez A, Armengod M-E. 2018 *Bacillus subtilis* exhibits MnmC-like tRNA modification activities. *RNA Biol* **15**, 1167–1173. (doi:10.1080/15476286.2018.1517012)

29. Moukadiri I, Garzon MJ, Björk GR, Armengod ME. 2014 The output of the tRNA modification pathways controlled by the *Escherichia coli* MnmEG and MnmC enzymes depends on the growth conditions and the tRNA species. *Nucleic Acids Res* **42**, 2602–2623. (doi:10.1093/nar/gkt1228)

30. Shippy D, Fadl A. 2014 tRNA Modification Enzymes GidA and MnmE: Potential role in virulence of bacterial pathogens. *Int J Mol Sci* **15**, 18267–18280. (doi:10.3390/ijms151018267)

31. Cho KH, Caparon MG. 2008 tRNA Modification by GidA/MnmE is necessary for *Streptococcus pyogenes* Virulence: a new strategy to make live attenuated strains. *Infect Immun* **76**, 3176–3186. (doi:10.1128/IAI.01721-07)

32. Li D, Shibata Y, Takeshita T, Yamashita Y. 2014 A novel gene involved in the survival of *Streptococcus mutans* under Stress Conditions. *Appl Environ Microbiol* **80**, 97–103. (doi:10.1128/AEM.02549-13)

33. Gao T, Tan M, Liu W, Zhang C, Zhang T, Zheng L, Zhu J, Li L, Zhou R. 2016 GidA, a tRNA modification enzyme, contributes to the growth, and virulence of *Streptococcus suis* Serotype 2. *Front Cell Infect Microbiol* **6**, 44. (doi:10.3389/fcimb.2016.00044)

34. Gao T *et al.* 2020 Proteomic and metabolomic analyses provide insights into the mechanism on arginine metabolism regulated by tRNA modification enzymes GidA and MnmE of *Streptococcus suis*. *Front Cell Infect Microbiol* **10**, 597408. (doi:10.3389/fcimb.2020.597408)

35. Cho G, Lee J, Kim J. 2023 Identification of a novel 5-aminomethyl-2-thiouridine methyltransferase in tRNA modification. *Nucleic Acids Res* **51**, 1971–1983. (doi:10.1093/nar/gkad048)

36. Jaroch M, Savage K, Kuipers P, Bacusmo JM, Hu J, Sun J, Dedon PC, Rice KC, de Crécy-Lagard V. 2025 Environmental control of queuosine levels in *Streptococcus mutans* tRNAs. *Mol Microbiol* **123**, 48–59. (doi:10.1111/mmi.15336)

37. Bacusmo JM *et al.* 2018 The t(6)A modification acts as a positive determinant for the anticodon nuclease PrrC, and is distinctively nonessential in *Streptococcus mutans*. *RNA Biol* **15**, 508–517. (doi:10.1080/15476286.2017.1353861)

38. de Crécy-Lagard V *et al.* 2024 Biosynthesis and function of 7-deazaguanine derivatives in bacteria and phages. *Microbiol Mol Biol Rev* **88**, e0019923. (doi:10.1128/mmbr.00199-23)

39. Quaiyum S, Yuan Y, Kuipers PJ, Martinelli M, Jaroch M, de Crécy-Lagard V. 2024 Deciphering the diversity in bacterial transporters that salvage queuosine precursors. *Epigenomes* **8**, 16. (doi:10.3390/epigenomes8020016)

40. Yuan Y *et al.* 2019 Discovery of novel bacterial queuine salvage enzymes and pathways in human pathogens. *Proc Natl Acad Sci USA* **116**, 19126–19135. (doi:10.1073/pnas.1909604116)

41. Zhang W, Westhof E. 2025 The biology of tRNA t^6a modification and hypermodifications—biogenesis and disease relevance. *J Mol Biol* **437**, 169091. (doi:10.1016/j.jmb.2025.169091)

42. Su C, Jin M, Zhang W. 2022 Conservation and diversification of tRNA t^6A-modifying enzymes across the three domains of life. *Int J Mol Sci* **23**, 13600. (doi:10.3390/ijms232113600)

43. Thiaville PC *et al.* 2015 Essentiality of threonylcarbamoyladenosine (t^6A), a universal tRNA modification, in bacteria. *Mol Microbiol* **98**, 1199–221. (doi:10.1111/mmi.13209)

44. NCBI Resource Coordinators. 2016 Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **44**, D7–D19. (doi:10.1093/nar/gkv1290)

45. O'Leary NA *et al.* 2016 Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res* **44**, D733–D745. (doi:10.1093/nar/gkv1189)

46. Galperin MY, Wolf YI, Makarova KS, Alvarez RV, Landsman D, Koonin E V. 2021 COG database update: Focus on microbial diversity, model organisms, and widespread pathogens. *Nucleic Acids Res* **49**, D274-D281. (doi:10.1093/nar/gkaa1018)

47. Li W, Cowley A, Uludag M, Gur T, McWilliam H, Squizzato S, Park YM, Buso N, Lopez R. 2015 The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Res* **43**, W580-4. (doi:10.1093/nar/gkv279)

48. Consortium U. 2021 UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res* **49**, D480–D489. (doi:10.1093/nar/gkaa1100)

49. Paysan-Lafosse T *et al.* 2023 InterPro in 2022. *Nucleic Acids Res* **51**, D418–D427. (doi:10.1093/nar/gkac993)

50. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. 2016 KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res* **44**, D457–D462. (doi:10.1093/nar/gkv1070)

51. Burley SK *et al.* 2022 RCSB Protein Data Bank (RCSB.org): delivery of experimentally-determined PDB structures alongside one million computed structure models of proteins from artificial intelligence/machine learning. *Nucleic Acids Res* **51**, D488-D508. (doi:10.1093/nar/gkac1077)

52. Olson RD *et al.* 2023 Introducing the Bacterial and Viral Bioinformatics Resource Center (BV-BRC): a resource combining PATRIC, IRD and ViPR. *Nucleic Acids Res* **51**, D678–D689. (doi:10.1093/nar/gkac1003)

53. Elfmann C, Dumann V, van den Berg T, Stülke J. 2025 A new framework for *Subti* Wiki, the database for the model organism *Bacillus subtilis*. *Nucleic Acids Res* **53**, D864–D870. (doi:10.1093/nar/gkae957)

54. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997 Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389–3402. (doi:10.1093/nar/25.17.3389)

55. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. 2019 GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* **36**, 1925-1927 (doi:10.1093/bioinformatics/btz848)

56. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, Lanfear R, Teeling E. 2020 IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol* **37**, 1530-1534. (doi:10.1093/molbev/msaa015)

57. Letunic I, Bork P. 2024 Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res* **52**, W78–W82. (doi:10.1093/nar/gkae268)

58. Katoh K, Standley DM. 2013 MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **30**, 772–780. (doi:10.1093/molbev/mst010)

59. Crooks GE, Hon G, Chandonia J-M, Brenner SE. 2004 WebLogo: A sequence logo generator. *Genome Res* **14**, 1188–1190. (doi:10.1101/gr.849004)

60. Zallot R, Oberg N, Gerlt JA. 2019 The EFI web resource for genomic enzymology tools: leveraging protein, genome, and metagenome databases to discover novel enzymes and metabolic pathways. *Biochemistry* **58**, 4169–4182. (doi:10.1021/acs.biochem.9b00735)

61. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003 Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498–2504.

62. Harrison KJ, de Crécy-Lagard V, Zallot R. 2018 Gene Graphics: A genomic neighborhood data visualization web application. *Bioinformatics* **34**, 1406–1408. (doi:10.1093/bioinformatics/btx793)

63. Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, Koren S, Phillippy AM. 2016 Mash: Fast genome and metagenome distance estimation using MinHash. *Genome Biol* **17**. (doi:10.1186/s13059-016-0997-x)

64. PyMol. In press. The PyMOL Molecular Graphics System, Version 1.8, Schrödinger, LLC.

65. Pei J, Tang M, Grishin N V. 2008 PROMALS3D web server for accurate multiple protein sequence and structure alignments. *Nucleic Acids Res* **36**, W30–W34. (doi:10.1093/nar/gkn322)

66. Gouet P, Courcelle E, Stuart DI, Metoz F. 1999 ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**, 305–308. (doi:10.1093/bioinformatics/15.4.305)

67. Lanie JA, Ng WL, Kazmierczak KM, Andrzejewski TM, Davidsen TM, Wayne KJ, Tettelin H, Glass JI, Winkler ME. 2007 Genome sequence of Avery's virulent serotype 2 strain D39 of Streptococcus pneumoniae and comparison with that of unencapsulated laboratory strain R6. *J Bacteriol* **189**, 38-51. (doi:10.1128/JB.01148-06)

68. Slager J, Aprianto R, Veening J-W. 2018 Deep genome annotation of the opportunistic human pathogen *Streptococcus pneumoniae* D39. *Nucleic Acids Res* **46**, 9971-9989 (doi:10.1093/nar/gky725)

69. Tsui HCT, Zheng JJ, Magallon AN, Ryan JD, Yunck R, Rued BE, Bernhardt TG, Winkler ME. 2016 Suppression of a deletion mutation in the gene encoding essential PBP2b reveals a new lytic transglycosylase involved in peripheral peptidoglycan synthesis in *Streptococcus pneumoniae* D39. *Mol Microbiol* **100**, 1039-65. (doi:10.1111/mmi.13366)

70. Sung CK, Li H, Claverys JP, Morrison DA. 2001 An *rpsL* cassette, Janus, for gene replacement through negative selection in *Streptococcus pneumoniae*. *Appl Environ Microbiol* **67**, 5190-6. (doi:10.1128/aem.67.11.5190-5196.2001)

71. Jacobsen FE, Kazmierczak KM, Lisher JP, Winkler ME, Giedroc DP. 2011 Interplay between manganese and zinc homeostasis in the human pathogen Streptococcus pneumoniae. *Metallomics* **3**, 38-41. (doi:10.1039/c0mt00050g)

72. Rued BE *et al.* 2017 Suppression and synthetic-lethal genetic relationships of Δ*gpsB* mutations indicate that GpsB mediates protein phosphorylation and penicillin-binding protein interactions in *Streptococcus pneumoniae* D39. *Mol Microbiol* **103**, 931–957. (doi:10.1111/mmi.13613)

73. Lamanna MM *et al.* 2022 Roles of RodZ and class A PBP1b in the assembly and regulation of the peripheral peptidoglycan elongasome in ovoid-shaped cells of *Streptococcus pneumoniae* D39. *Mol Microbiol* **118**, 336-368. (doi:10.1111/mmi.14969)

74. Carver TE *et al.* 2005 Decrypting the biochemical function of an essential gene from *Streptococcus pneumoniae* using ThermoFluor technology. *J Biol Chem* **280**, 11704–11712. (doi:10.1074/jbc.M413278200)

75. Tsui HCT *et al.* 2023 Negative regulation of MurZ and MurA underlies the essentiality of GpsB- and StkP-mediated protein phosphorylation in *Streptococcus pneumoniae* D39. *Mol Microbiol* **120**, 351-38 (doi:10.1111/mmi.15122)

76. Deatherage DE, Barrick JE. 2014 Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. *Methods Mol Biol* **1151**, 165-88. (doi:10.1007/978-1-4939-0554-6_12)

77. Chia JS, Lee YY, Huang PT, Chen JY. 2001 Identification of stress-responsive genes in *Streptococcus mutans* by differential display reverse transcription-PCR. *Infect Immun* **69**, 2493-501. (doi:10.1128/IAI.69.4.2493-2501.2001)

78. Hengesbach M *et al.* 2025 Toward standardized epitranscriptome analytics: an inter-laboratory comparison of mass spectrometric detection and quantification of modified ribonucleosides in human RNA. *Nucleic Acids Res* **53**, gkaf895. (doi:10.1093/nar/gkaf895)

79. Boccaletto P *et al.* 2018 MODOMICS: A database of RNA modification pathways. 2017 update. *Nucleic Acids Res* **46**, D303–D307. (doi:10.1093/nar/gkx1030)

80. Marchand V, Bourguignon-Igel V, Helm M, Motorin Y. 2021 Mapping of 7-methylguanosine (m^7G), 3-methylcytidine (m^3C), dihydrouridine (D) and 5-hydroxycytidine (ho^5C) RNA modifications by AlkAniline-Seq. In *Methods Enzymol* **658**, 25-47. (doi:10.1016/bs.mie.2021.06.001)

81. Marchand V *et al.* 2018 AlkAniline-Seq: Profiling of m^7G and m^3C RNA Modifications at Single Nucleotide Resolution. *Angew Chem Int Ed Engl* **57**, 16785-16790. (doi:10.1002/anie.201810946)

82. Peattie DA, Gilbert W. 1980 Chemical probes for higher-order structure in RNA. *Proc Natl AcadSci USA* **77**, 4679–4682.

83. Bolger AM, Lohse M, Usadel B. 2014 Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120. (doi:10.1093/bioinformatics/btu170)

84. Chan PP, Lowe TM. 2016 GtRNAdb 2.0: an expanded database of transfer RNA genes identified in complete and draft genomes. *Nucleic Acids Res* **44**, D184–D189. (doi:10.1093/nar/gkv1309)

85. Langmead B, Salzberg SL. 2012 Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357–359. (doi:10.1038/nmeth.1923)

86. Schwartz S *et al.* 2014 Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. *Cell* **159**, 148–162. (doi:10.1016/j.cell.2014.08.028)

87. Shen W *et al.* 2024 GLORI for absolute quantification of transcriptome-wide m^6A at single-base resolution. *Nat Protoc* **19**, 1252–1287. (doi:10.1038/s41596-023-00937-1)

88. Reichle VF, Weber V, Kellner S. 2018 NAIL-MS in *E. coli* determines the source and fate of methylation in tRNA. *ChemBioChem* **19**, 2575–2583. (doi:10.1002/cbic.201800525)

89. Kimura S, Suzuki T. 2010 Fine-tuning of the ribosomal decoding center by conserved methyl-modifications in the *Escherichia coli 16S* rRNA. *Nucleic Acids Res*. **38,** 1341-52. (doi: 10.1093/nar/gkp1073).

90. Bommisetti P, Bandarian V. 2022 Site-Specific Profiling of 4-thiouridine across transfer RNA genes in *Escherichia coli*. *ACS Omega* **7**, 4011-402. (doi:10.1021/acsomega.1c05071)

91. Phillipson DW, Edmonds CG, Crain PF, Smith DL, Davis DR, McCloskey. 1987 Isolation and structure elucidation of an epoxide derivative of the hypermodified nucleoside queuosine from *Escherichia coli* transfer RNA. *J Biol Chem* **262**, 3462–3471.

92. Chikwana VM, Stec B, Lee BWK, de Crécy-Lagard V, Iwata-Reuyl D, Swairjo MA. 2012 Structural basis of biological nitrile reduction. *J Biol Chem* **287**, 30560–30570. (doi:10.1074/jbc.M112.388538)

93. Quaiyum S, Sun J, Marchand V, Sun G, Reed CJ, Motorin Y, Dedon PC, Minnick MF, de Crécy-Lagard V. 2024 Mapping the tRNA modification landscape of *Bartonella henselae* Houston I and *Bartonella quintana* Toulouse. *Front Microbiol* **15**, 1369018. (doi:10.3389/fmicb.2024.1369018)

94. Stengl B, Reuter K, Klebe G. 2005 Mechanism and substrate specificity of tRNA-Guanine transglycosylases (TGTs): tRNA-modifying enzymes from the three different kingdoms of life share a common catalytic mechanism. *Chem. Bio. Chem.* **6**, 1926–1939.

95. López JL, Lozano MJ, Fabre ML, Lagares A. 2020 Codon usage optimization in the Prokaryotic Tree of Life: how synonymous codons are differentially selected in sequence domains with different expression levels and degrees of conservation. *mBio* **11**, e00766-20. (doi:10.1128/mBio.00766-20)

96. Shields RC, Zeng L, Culp DJ, Burne RA. 2018 Genomewide identification of essential genes and fitness determinants of *Streptococcus mutans* UA159. *mSphere* **3**, e00031-18. (doi:10.1128/mSphere.00031-18)

97. Shields RC, Walker AR, Maricic N, Chakraborty B, Underhill SAM, Burne RA. 2020 Repurposing the *Streptococcus mutans* CRISPR-Cas9 system to understand essential gene function. *PLoS Pathog* **16**, 1008344. (doi:10.1371/journal.ppat.1008344)

98. Quivey Jr RG *et al.* 2015 Functional profiling in *Streptococcus mutans:* construction and examination of a genomic collection of gene deletion mutants. *Mol Oral Microbiol* **30**, 474–495. (doi:10.1111/omi.12107)

99. Janssen AB, Gibson PS, Bravo AM, de Bakker V, Slager J, Veening J-W. 2025 PneumoBrowse 2: an integrated visual platform for curated genome annotation and multiomics data analysis of *Streptococcus pneumoniae*. *Nucleic Acids Res* **53**, D839–D851. (doi:10.1093/nar/gkae923)

100. St. Pierre J, Roberts J, Alam MA, Shields RC. 2024 Construction of an arrayed CRISPRi library as a resource for essential gene function studies in *Streptococcus mutans. Microbiol Spectr* **12,** e0314923. (doi:10.1128/spectrum.03149-23)

101. Black KA, dos Santos PC. 2015 Abbreviated pathway for biosynthesis of 2-thiouridine in *Bacillus subtilis*. *J Bacteriol* **196**, 1952–62. (doi:10.1128/JB.02625-14)

102. Urbonavicius J, Qian Q, Durand JMB, Hagervall TG, Björk GR. 2001 Improvement of reading frame maintenance is a common function for several tRNA modifications. *EMBO J*, **20**, 4863-73.

103. Gurumayum S *et al.* 2021 OGEE v3: Online GEne Essentiality database with increased coverage of organisms and human cell lines. *Nucleic Acids Res* **49**, D998-D1003. (doi:10.1093/nar/gkaa884)

104. Kambampati R, Lauhon CT. 2003 MnmA and IscS are required for in vitro 2-thiouridine biosynthesis in *Escherichia coli*. *Biochemistry* **42**, 1109–1117.

105. Soma A *et al.* 2023 YaaJ, the tRNA-specific adenosine deaminase, is dispensable in *Bacillus subtilis. Genes (Basel)* **14,** 1515. (doi:10.3390/genes14081515)

106. Liao W, Nie W, Ahmad I, Chen G, Zhu B. 2023 The occurrence, characteristics, and adaptation of A-to-I RNA editing in bacteria: A review. *Front Microbiol*. **14**, 143929. (doi:10.3389/fmicb.2023.1143929)

107. Zundel MA, Basturea GN, Deutscher MP. 2009 Initiation of ribosome degradation during starvation in *Escherichia coli*. *RNA* **15**, 977–983. (doi:10.1261/rna.1381309)

108. Cherlin T, Magee R, Jing Y, Pliatsika V, Loher P, Rigoutsos I. 2020 Ribosomal RNA fragmentation into short RNAs (rRFs) is modulated in a sex- and population of origin-specific manner. *BMC Biol* **18**, 38. (doi:10.1186/s12915-020-0763-0)

109. Suzuki T, Nakamura A, Kato K, Söll D, Tanaka I, Sheppard K, Yao M. 2015 Structure of the *Pseudomonas aeruginosa* transamidosome reveals unique aspects of bacterial tRNA-dependent asparagine biosynthesis. *Proc Natl Acad Sci U S A* **112**., 382-7. (doi:10.1073/pnas.1423314112)

110. Devarkar SC, Budding CR, Pathirage C, Kavoor A, Herbert C, Limbach PA, Musier-Forsyth K, Xiong Y. 2025 Structural basis for aminoacylation of cellular modified tRNALys3 by human lysyl-tRNA synthetase. *Nucleic Acids Res* **53**, gkaf114. (doi:10.1093/nar/gkaf114)

111. Jumper J *et al.* 2021 Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589. (doi:10.1038/s41586-021-03819-2)

112. Iwasaki W, Sekine S ichi, Kuroishi C, Kuramitsu S, Shirouzu M, Yokoyama S. 2006 Structural basis of the water-assisted asparagine recognition by Asparaginyl-tRNA synthetase. *J Mol Biol* **360,** 329-42. (doi:10.1016/j.jmb.2006.04.068)

113. Crepin T, Peterson F, Haertlein M, Jensen D, Wang C, Cusack S, Kron M. 2011 A hybrid structural model of the complete *Brugia malayi* cytoplasmic asparaginyl-tRNA synthetase. *J Mol Biol* **405**, 1056-69. (doi:10.1016/j.jmb.2010.11.049)

114. Xie SC *et al.* 2024 Reaction hijacking inhibition of *Plasmodium falciparum* asparagine tRNA synthetase. *Nat Commun* **15**, 937. (doi:10.1038/s41467-024-45224-z)

115. Charron C, Roy H, Blaise M, Giegé R, Kern D. 2003 Non-discriminating and discriminating aspartyl-tRNA synthetases differ in the anticodon-binding domain. *EMBO J* **22**, 1632-43. (doi:10.1093/emboj/cdg148)

116. Sauter C, Lorber B, Cavarelli J, Moras D, Giegé R. 2000 The free yeast aspartyl-tRNA synthetase differs from the tRNA(Asp)-complexed enzyme by structural changes in the catalytic site, hinge region, and anticodon-binding domain. *J Mol Biol* **299**, 1313-24. (doi:10.1006/jmbi.2000.3791)

117. Briand C, Poterszman A, Eiler S, Webster G, Thierry J-C, Moras D. 2000 An intermediate step in the recognition of tRNA Asp by aspartyl-tRNA synthetase. *J Mol Biol* **299**, 1051–1060. (doi:10.1006/jmbi.2000.3819)

118. Weinert LA, Welch JJ. 2017 Why might bacterial pathogens have small genomes? *Trends Ecol Evol*. **32**, 936-947. (doi:10.1016/j.tree.2017.09.006)

119. Dos Santos PC. 2017 Fe-S assembly in Gram-positive bacteria. In *Biochemistry, Biosynthesis and Human Diseases: Biochemistry, Biosynthesis, and Human Diseases*, edited by Tracey Rouault, Berlin, Boston: De Gruyter, 2017, pp. 97-116. (doi:10.1515/9783110479850-005)

120. Ellepola K, Huang X, Riley RP, Bitoun JP, Wen ZT. 2021 *Streptococcus mutans* lacking *sufCDSUB* Is viable, but displays major defects in growth, stress tolerance responses and biofilm formation. *Front Microbiol* **12**, 671533. (doi:10.3389/fmicb.2021.671533)

121. Redanz S, Treerat P, Mu R, Redanz U, Zou Z, Koley D, Merritt J, Kreth J. 2020 Pyruvate secretion by oral streptococci modulates hydrogen peroxide dependent antagonism. *ISME J*. **14,** 1074-1088. (doi:10.1038/s41396-020-0592-8)

122. Huang X, Browngardt CM, Jiang M, Ahn SJ, Burne RA, Nascimento MM. 2018 Diversity in antagonistic interactions between commensal oral Streptococci and *Streptococcus mutans*. *Caries Res* **52,** 88-101. (doi:10.1159/000479091)

123. Jiang Q, Zhao Y, Shui Y, Zhou X, Cheng L, Ren B, Chen Z, Li M. 2021 Interactions between neutrophils and periodontal pathogens in late-onset periodontitis. *Front Cell Infect Microbiol*. **11***,* 627328. (doi:10.3389/fcimb.2021.627328)

124. Sarr D, Tóth E, Gingerich A, Rada B. 2018 Antimicrobial actions of dual oxidases and lactoperoxidase. *J Microbiol* **56**, 73-386. (doi:10.1007/s12275-018-7545-1)

125. Dowling DP, Bruender NA, Young AP, McCarty RM, Bandarian V, Drennan CL. 2014 Radical SAM enzyme QueE defines a new minimal core fold and metal-dependent mechanism. *Nature Chem Biol* **10**, 106–112. (doi:10.1038/nchembio.1426)

126. McCoy LS, Xie Y, Tor Y. 2011 Antibiotics that target protein synthesis. *WIREs RNA* **2**, 209–232. (doi:10.1002/wrna.60)

127. Hill PJ *et al.* 2013 Selective inhibitors of bacterial t-RNA-(N1G37) methyltransferase (TrmD) that demonstrate novel ordering of the lid domain. *J Med Chem* **56**, 7278–7288. (doi:10.1021/jm400718n)

128. Suzuki T, Miyauchi K. 2010 Discovery and characterization of tRNAIle lysidine synthetase (TilS). *FEBS Lett* **584**, 272–277.

129. Arigoni F *et al.* 1998 A genome-based approach for the identification of essential bacterial genes. *Nat Biotechnol* **16**, 851–856.

130. Handford JI, Ize B, Buchanan G, Butland GP, Greenblatt J, Emili A, Palmer T. 2009 Conserved network of proteins essential for bacterial viability. *J Bacteriol* **191**, 4732–4749. (doi:10.1128/jb.00136-09)

131. Allali-Hassani A, Campbell TL, Ho A, Schertzer JW, Brown ED. 2004 Probing the active site of YjeE: a vital *Escherichia coli* protein of unknown function. *Biochem J* **384**, 577–584. (doi:10.1042/BJ20041082)

132. Campbell TL, Ederer CS, Allali-Hassani A, Brown ED. 2007 Isolation of the *rstA* gene as a multicopy suppressor of YjeE, an essential ATPase of unknown function in *Escherichia coli*. *J Bacteriol* **189**, 3318–21. (doi:10.1128/JB.00131-06)

133. Allali-Hassani A, Campbell TL, Ho A, Schertzer JW, Brown ED. 2004 Probing the active site of YjeE: a vital *Escherichia coli* protein of unknown function. *Biochem J* **384**, 577–584. (doi:10.1042/BJ20041082)

134. Freiberg C, Wieland B, Spaltmann F, Ehlert K, Brötz H, Labischinski H. 2001 Identification of novel essential *Escherichia coli* genes conserved among pathogenic bacteria. *J Mol Microbiol Biotechnol* **3,** 483-9.

135. Lerner CG *et al.* 2007 From bacterial genomes to novel antibacterial agents: discovery, characterization, and antibacterial activity of compounds that bind to HI0065 (YjeE) from *Haemophilus influenzae*. *Chem Biol Drug Des* **69**, 395–404. (doi:10.1111/j.1747-0285.2007.00521.x)

136. Clifton BE, Fariz MA, Uechi GI, Laurino P. 2021 Evolutionary repair reveals an unexpected role of the tRNA modification m^1G37 in aminoacylation. *Nucleic Acids Res* **49**, 12467-12485. (doi:10.1093/nar/gkab1067)

137. Sun J *et al.* 2025 tRNA modification profiling reveals epitranscriptome regulatory networks in *Pseudomonas aeruginosa*. *Nucleic Acids Res* **53**, gkaf696. (doi:10.1093/nar/gkaf696)

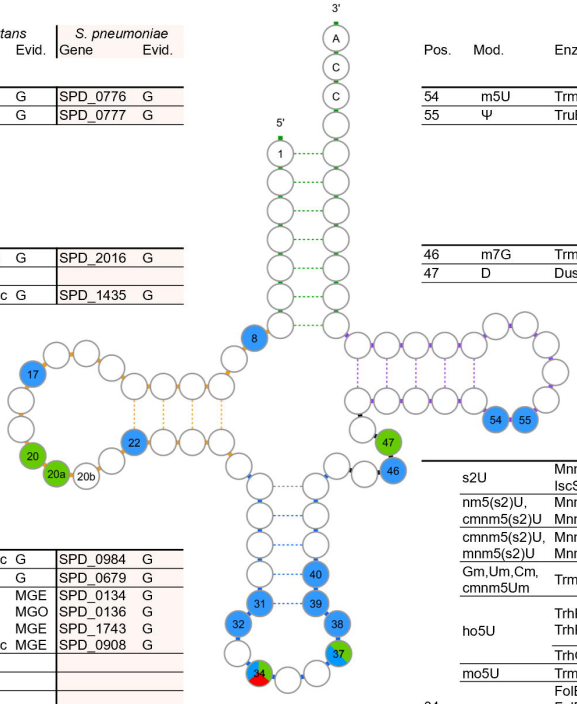
138. Leavitt JS, Moore H, Santangelo TJ, Lowe TM. 2025 The landscape of tRNA modifications in Archaea. bioRxiv [Preprint]. 2025 Sep 9:2025.05.02.651894 (doi:10.1101/2025.05.02.651894)

139. Hardy L *et al.* 2025 The tRNA epitranscriptomic landscape and RNA modification enzymes in *Vibrio cholerae*. *PLoS Genet.* **21**, e1011937. (doi: 10.1371/journal.pgen.1011937)

140. Jiang W *et al.* 2024 Spontaneous mutations and mutational responses to penicillin treatment in the bacterial pathogen *Streptococcus pneumoniae* D39. *Mar Life Sci Technol* **6**, 198–211. (doi:10.1007/s42995-024-00220-6)

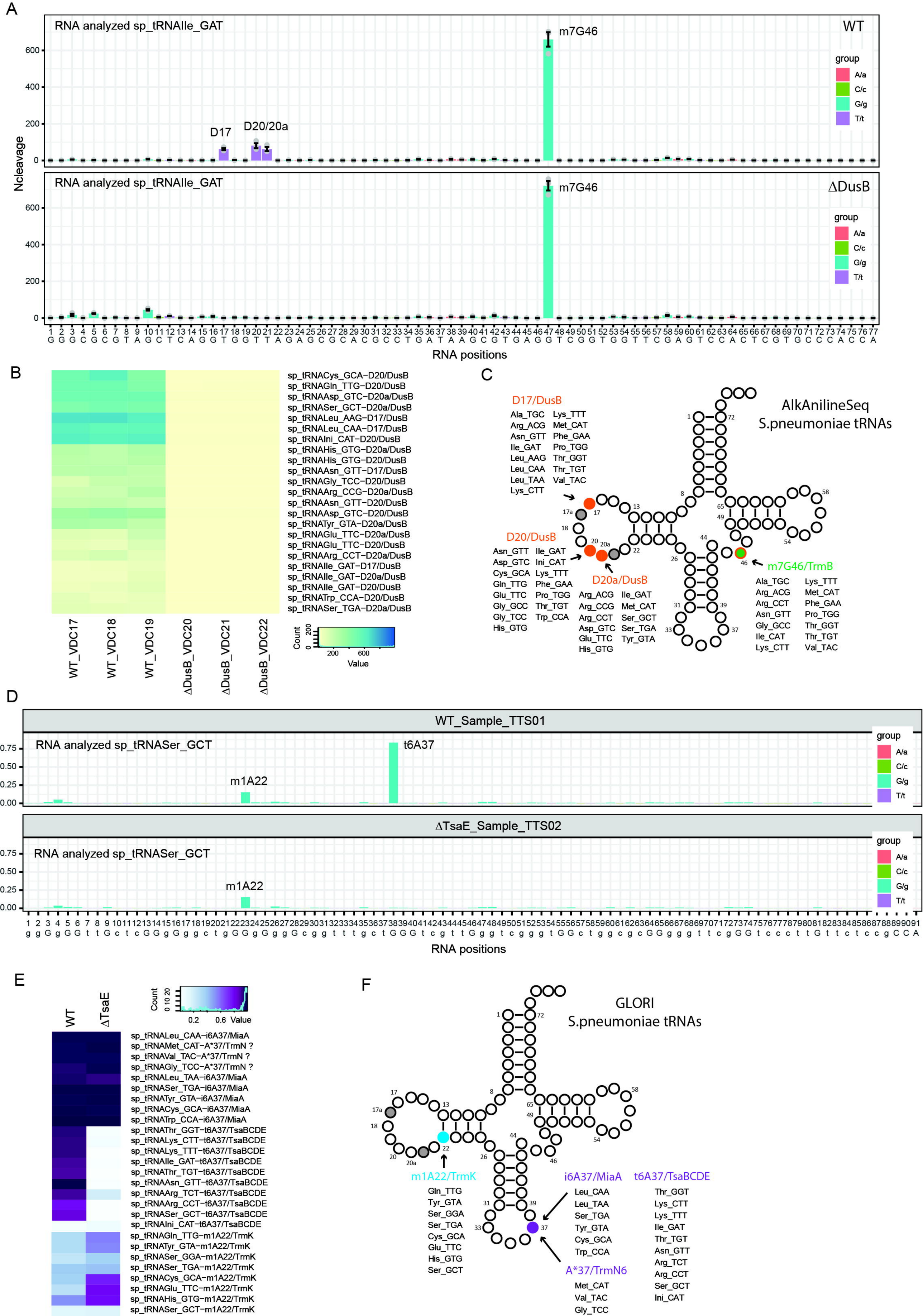
141. Zheng JJ, Perez AJ, Tsui HCT, Massidda O, Winkler ME. 2017 Absence of the KhpA and KhpB (JAG/EloR) RNA-binding proteins suppresses the requirement for PBP2b by overproduction of FtsA in *Streptococcus pneumoniae* D39. *Mol Microbiol* **106**, 793-814. (doi:10.1111/mmi.13847)

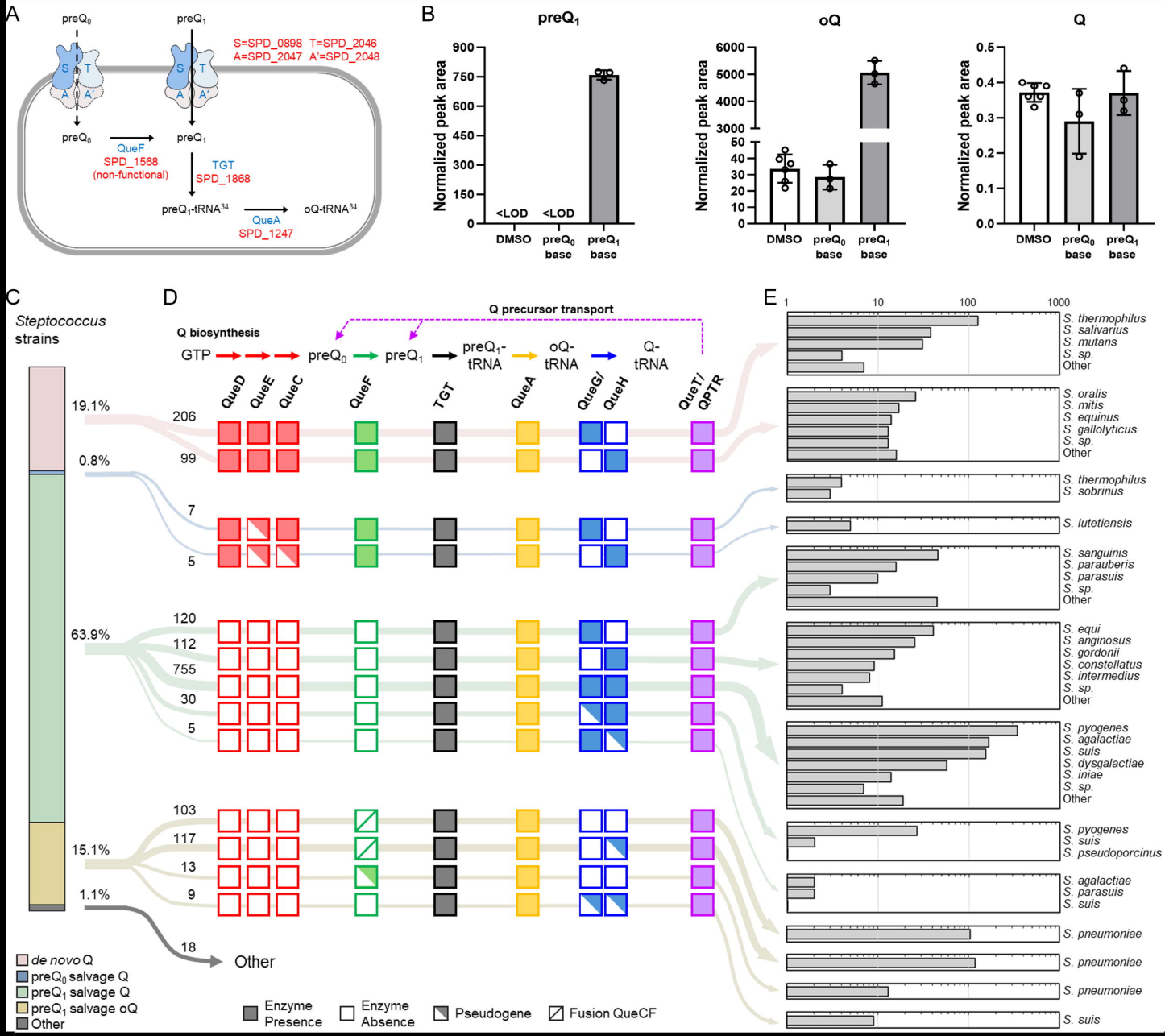
Pos.	Mod.	Enzyme	*B. subtilis* Gene	*B. subtilis* Evid.	*S. mutans* Gene	*S. mutans* Evid.	*S. pneumoniae* Gene	*S. pneumoniae* Evid.
		Acceptor Arm						
8	S4U	NifZ/IscS2	BSU29590	MGE	SMU.841	G	SPD_0776	G
		ThiI	BSU29580	MGE	SMU.842	G	SPD_0777	G
		D arm						
17	D	DusB1	BSU00810	MGE	SMU.187c	G	SPD_2016	G
20,20a	D	DusB2	BSU08030	MGE				
22	m1A	TrmK	BSU25180	MGE	SMU.1464c	G	SPD_1435	G
		Anticodon arm						
31,32	Ψ	YjbO	BSU11620	MGE	SMU.1044c	G	SPD_0984	G
37	m1G	TrmD	BSU16030	MG	SMU.868	G	SPD_0679	G
	t6A	TsaB	BSU05920	MGE	SMU.385	MGE	SPD_0134	G
		TsaD	BSU05940	MGE	SMU.387	MGO	SPD_0136	G
		TsaE	BSU05910	MGE	SMU.409	MGE	SPD_1743	G
		TsaC2	BSU36950	MGE	SMU.1083c	MGE	SPD_0908	G
	ct6A	TcdA	BSU27540	MGO				
	ms2t6A	MtaB	BSU25430	MGE				
	ms2ct6A	TcdA	BSU27540	MGE				
		MtaB	BSU27430	MGE				
	m2A	RlmN	BSU15750	MGO	SMU.522	G	SPD_0669	G
	m6A	TrmN	BSU00340	MO	SMU.621c	G	SPD_1365	G
	i6A	MiaA	BSU17330	MGO	SMU.1477	G	SPD_0583	G
	ms2i6A	MiaB	BSU17010	MGE				
38	Ψ	TruA	BSU01480	MG	SMU.84	G	SPD_1424	G
39	Ψ	TruA	BSU01480	MGE	SMU.84	G	SPD_1424	G
40	Ψ	TruA	BSU01480	MG	SMU.84	G	SPD_1424	G

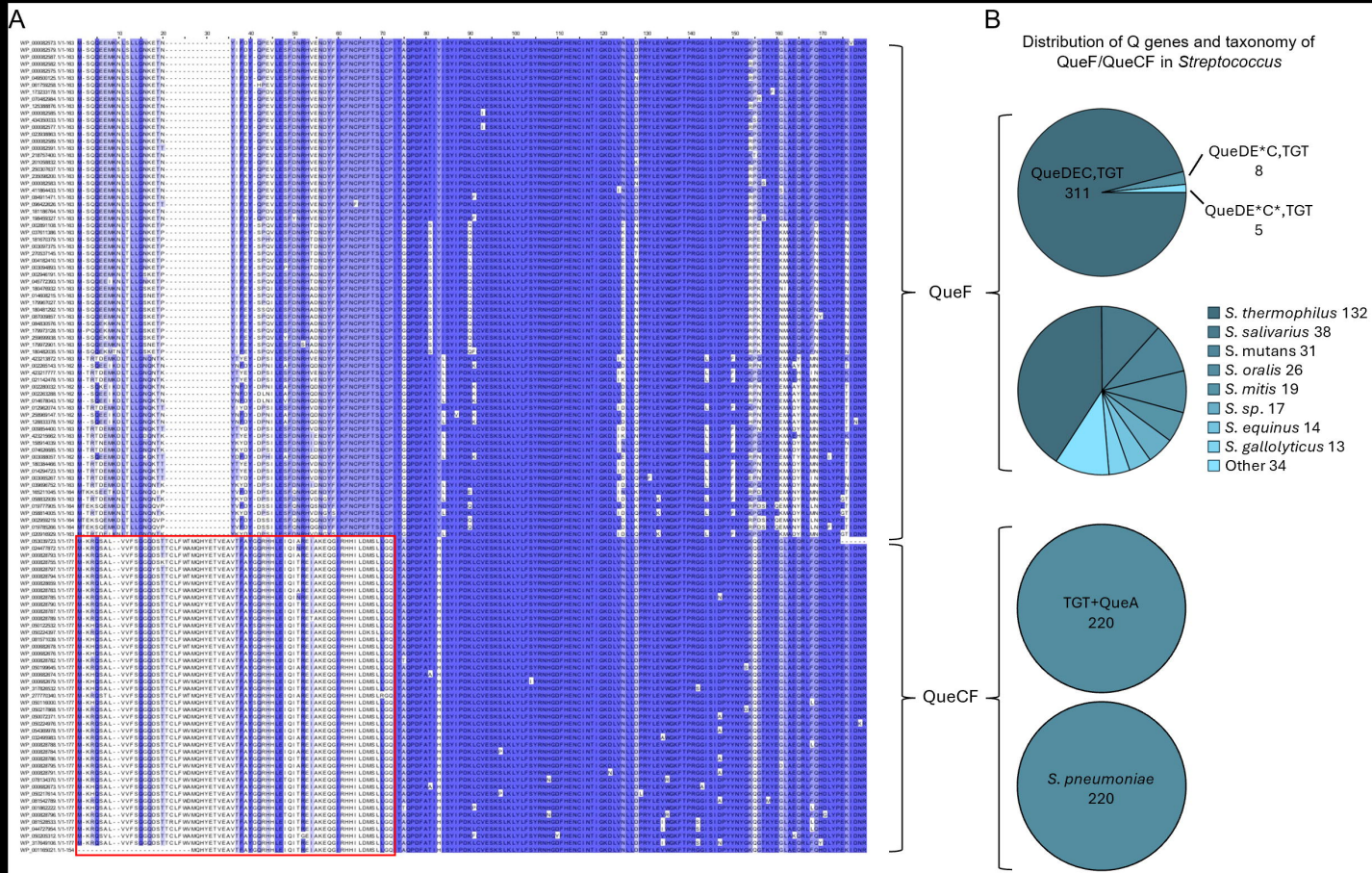


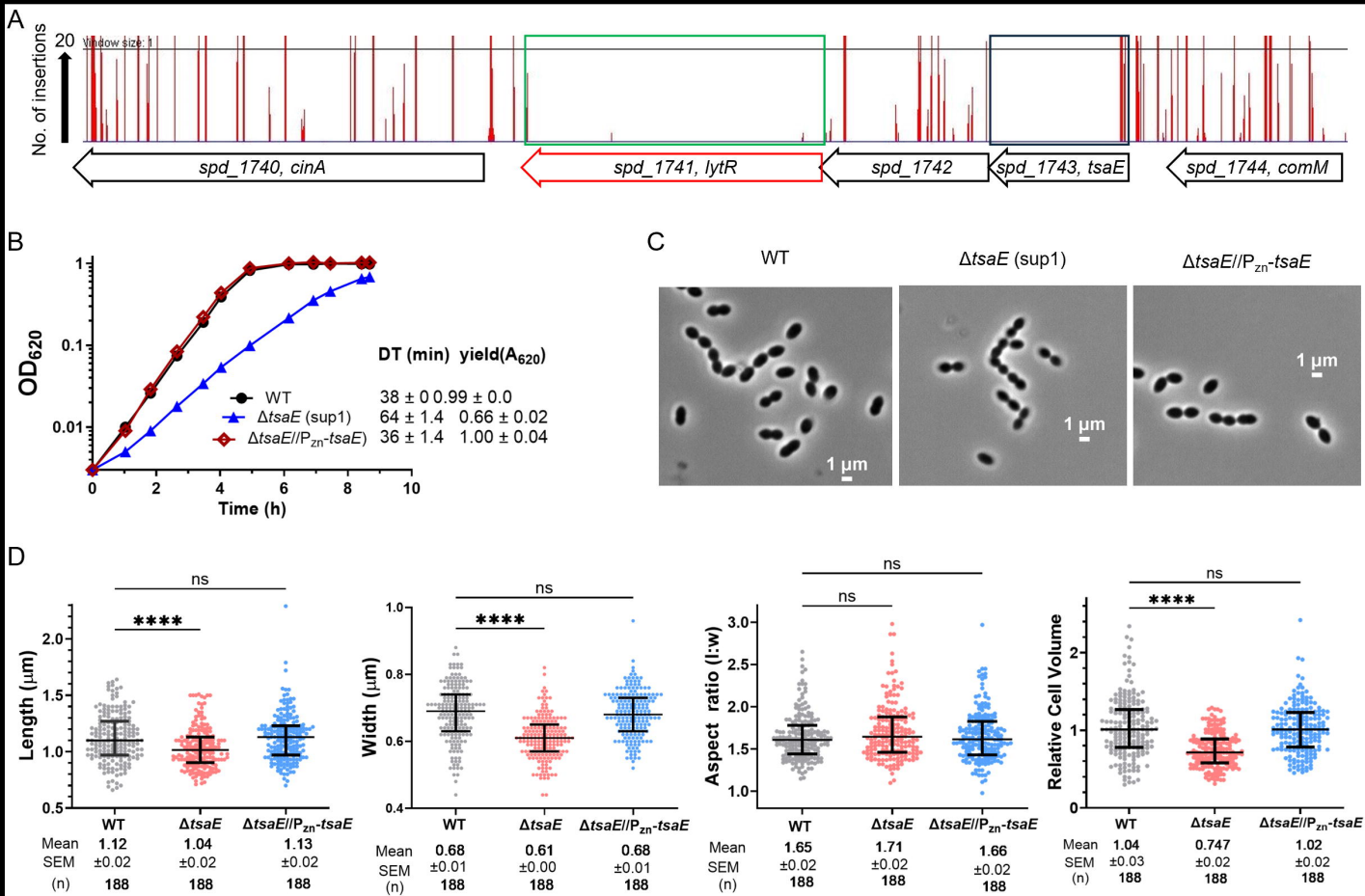
Pos.	Mod.	Enzyme	*B. subtilis* Gene	*B. subtilis* Evid.	*S. mutans* Gene	*S. mutans* Evid.	*S. pneumoniae* Gene	*S. pneumoniae* Evid.
		T arm						
54	m5U	TrmFO	BSU16130	MGE	SMU.1003	G	SPD_0833	G
55	Ψ	TruB	BSU16660	MG	SMU.1144	G	SPD_1070	G
		V arm						
46	m7G	TrmB	BSU29900	MGE	SMU.416	G	SPD_0477	G
47	D	DusB1	BSU00810	MGE				
		Anticodon						
34	s2U	MnmA	BSU27500	MGE	SMU.2143c	G	SPD_0127	G
		IscS1	BSU27510	MGE	SMU.1051	G	SPD_0979	G
	nm5(s2)U, cmnm5(s2)U	MnmE	BSU41020	MGO	SMU.1235	MG	SPD_0902	MG
		MnmG	BSU41010	MGE	SMU.2141	MG	SPD_0129	MG
	cmnm5(s2)U, mnm5(s2)U	MnmL	BSU30480	MGE	SMU.1699c	MGE		
		MnmM	BSU30490	MGE	SMU.1697c	MGE	SPD_0567	MGE
	Gm,Um,Cm, cmnm5Um	TrmL	BSU08930	MG	SMU.1707c	G	SPD_0436	G
	ho5U	TrhP1	BSU27340	MGE	SMU.761	G	SPD_1258,S PD_1256	G,G
		TrhP2	BSU27350	MGE	SMU.759	G	SPD_0704	G
		TrhO	BSU02330	MGE			SPD_0091	G
	mo5U	TrmR	BSU27360	MGE	SMU.647	G	SPD_0867	G
	Q	FolE1	BSU22780	MGE	SMU.968	MGO	SPD_0271	G
		FolE2	BSU03340	MGE		M		
		QueD	BSU13730	MGO	SMU.917c	MGE		
		QueE	BSU13740	MGE	SMU.916c	MGE		
		QueC	BSU13720	MGO	SMU.919c	MGE		
		QueF	BSU13750	MGE	SMU.915c	MGE	SPD_1568	G
		TGT	BSU27710	MGO	SMU.300	MGE	SPD_1868	G
		QueA	BSU27720	MGO	SMU.634	MGE	SPD_1247	G
		QueG	BSU08910	MGE	SMU.1327c	MGE		
		QPTR	BSU21980	MGO		M		
		QueT		M	SMU.1545c	MGO	SPD_0898	G
		QueK/Qng1		M		M		
		QueL		M		M		
	ac4C	TmcAL	BSU15060	MGE	SMU.1791c	G	SPD_1552	G
	k2C	TilS	BSU00670	MGE	SMU.13	G	SPD_0011	G
	I	TadA	BSU00180	MGE	SMU.304	G	SPD_0025	G

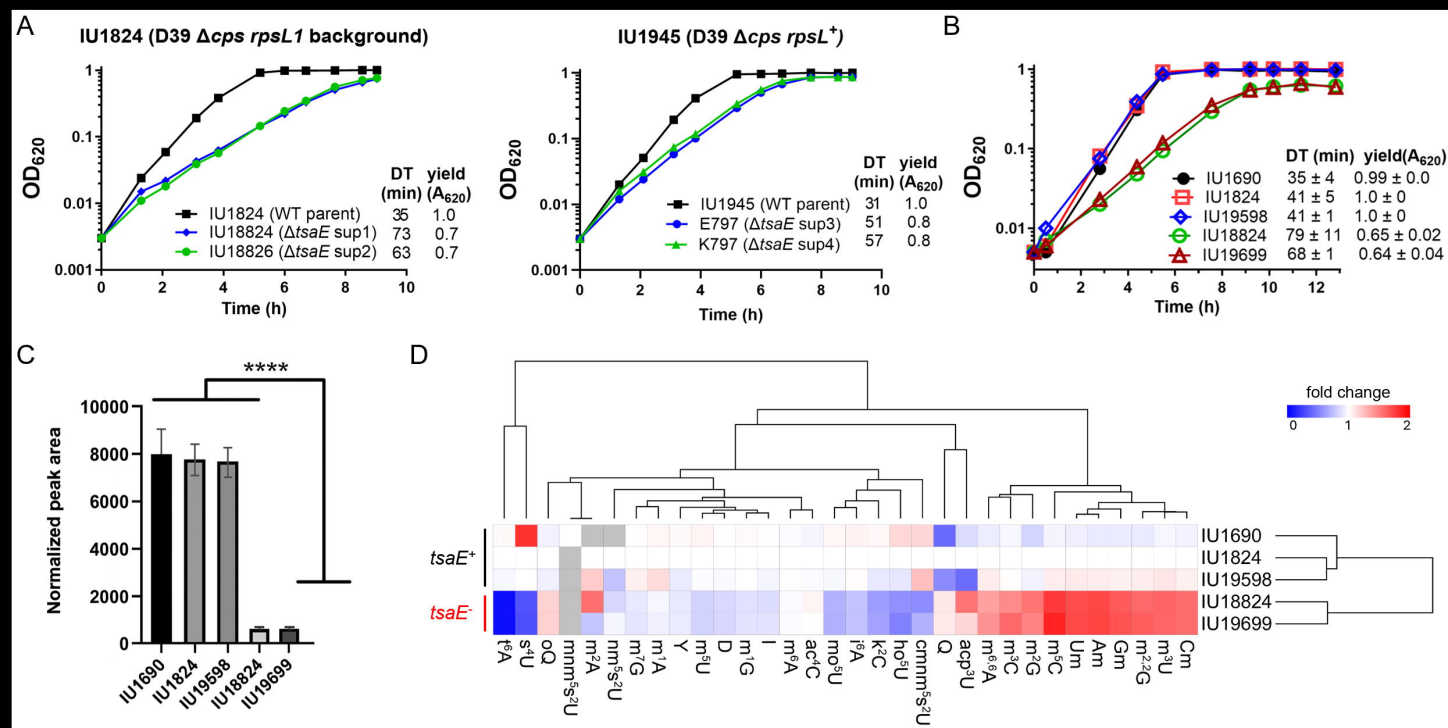
- Same modification and corresponding enzymes among *B. subtilis*, *S. mutans* and *S. pneumoniae*
- Different modification or corresponding enzymes between *Bacillus* and *Streptococcus*
- Different modification or corresponding enzymes between *S. mutans* and *S. pneumoniae*

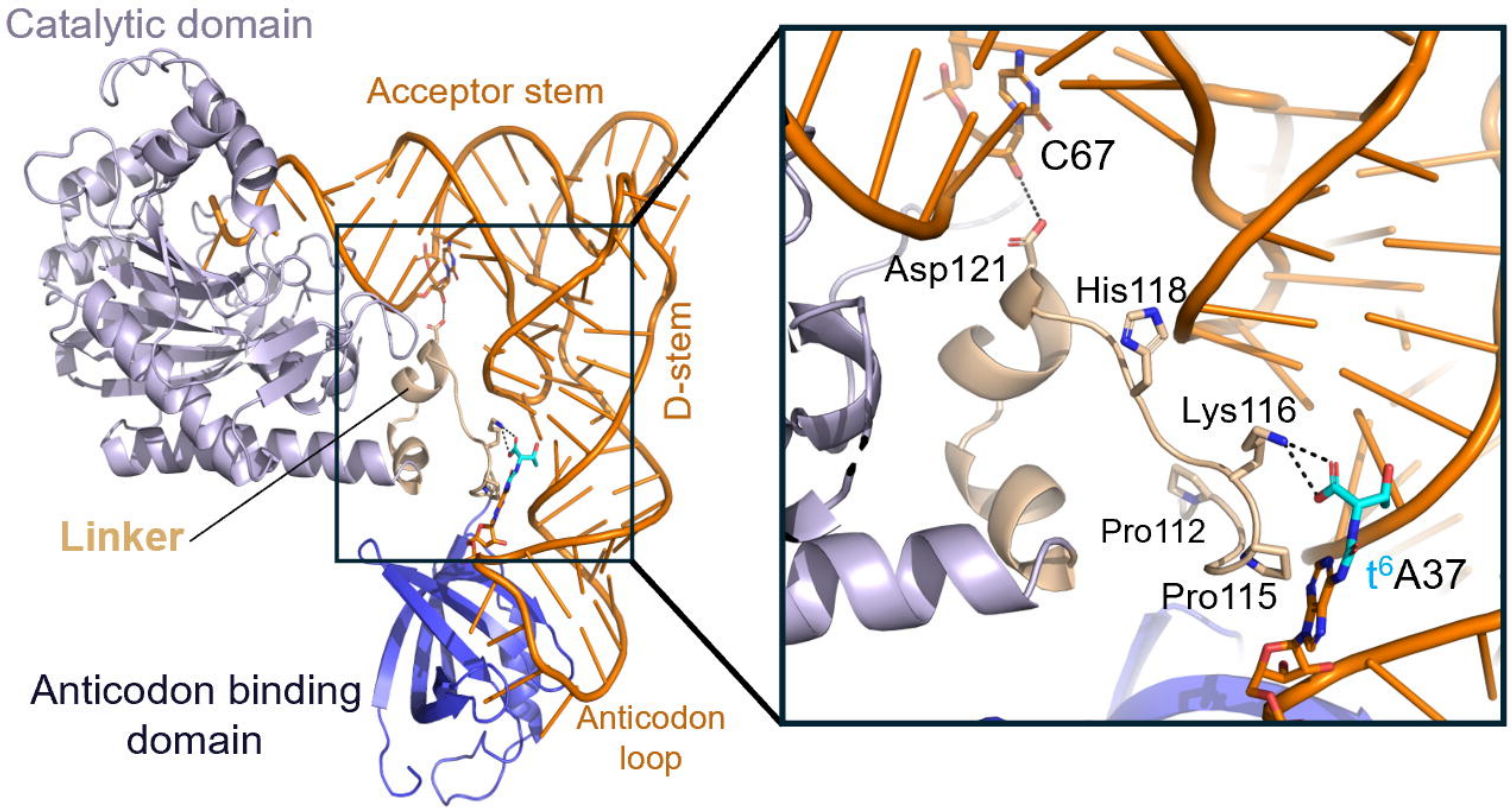

A
RNA analyzed sp_tRNAIle_GAT
WT
m7G46
D17
D20/20a
group
A/a
C/c
G/g
T/t
Ncleavage
ΔDusB
RNA positions
B
sp_tRNACys_GCA−D20/DusB
sp_tRNAGln_TTG−D20/DusB
sp_tRNAAsp_GTC−D20a/DusB
sp_tRNASer_GCT−D20a/DusB
sp_tRNALeu_AAG−D17/DusB
sp_tRNALeu_CAA−D17/DusB
sp_tRNAIni_CAT−D20/DusB
sp_tRNAHis_GTG−D20a/DusB
sp_tRNAHis_GTG−D20/DusB
sp_tRNAAsn_GTT−D17/DusB
sp_tRNAGly_TCC−D20/DusB
sp_tRNAArg_CCG−D20a/DusB
sp_tRNAAsn_GTT−D20/DusB
sp_tRNAAsp_GTC−D20/DusB
sp_tRNATyr_GTA−D20a/DusB
sp_tRNAGlu_TTC−D20a/DusB
sp_tRNAGlu_TTC−D20/DusB
sp_tRNAArg_CCT−D20a/DusB
sp_tRNAIle_GAT−D17/DusB
sp_tRNAIle_GAT−D20a/DusB
sp_tRNAIle_GAT−D20/DusB
sp_tRNATrp_CCA−D20/DusB
sp_tRNASer_TGA−D20a/DusB
WT_VDC17
WT_VDC18
WT_VDC19
ΔDusB_VDC20
ΔDusB_VDC21
ΔDusB_VDC22
Count
Value
C
AlkAnilineSeq
S.pneumoniae tRNAs
D17/DusB
Ala_TGC Lys_TTT
Arg_ACG Met_CAT
Asn_GTT Phe_GAA
Ile_GAT Pro_TGG
Leu_AAG Thr_GGT
Leu_CAA Thr_TGT
Leu_TAA Val_TAC
Lys_CTT
D20/DusB
Asn_GTT Ile_GAT
Asp_GTC Ini_CAT
Cys_GCA Lys_TTT
Gln_TTG Phe_GAA
Glu_TTC Pro_TGG
Gly_GCC Thr_TGT
Gly_TCC Trp_CCA
His_GTG
D20a/DusB
Arg_ACG Ile_GAT
Arg_CCG Met_CAT
Arg_CCT Ser_GCT
Asp_GTC Ser_TGA
Glu_TTC Tyr_GTA
His_GTG
m7G46/TrmB
Ala_TGC Lys_TTT
Arg_ACG Met_CAT
Arg_CCT Phe_GAA
Asn_GTT Pro_TGG
Gly_GCC Thr_GGT
Ile_CAT Thr_TGT
Lys_CTT Val_TAC
D
WT_Sample_TTS01
RNA analyzed sp_tRNASer_GCT
t6A37
m1A22
ΔTsaE_Sample_TTS02
RNA positions
E
WT
ΔTsaE
Count
Value
sp_tRNALeu_CAA−i6A37/MiaA
sp_tRNAMet_CAT−A*37/TrmN ?
sp_tRNAVal_TAC−A*37/TrmN ?
sp_tRNAGly_TCC−A*37/TrmN ?
sp_tRNALeu_TAA−i6A37/MiaA
sp_tRNASer_TGA−i6A37/MiaA
sp_tRNATyr_GTA−i6A37/MiaA
sp_tRNACys_GCA−i6A37/MiaA
sp_tRNATrp_CCA−i6A37/MiaA
sp_tRNAThr_GGT−t6A37/TsaBCDE
sp_tRNALys_CTT−t6A37/TsaBCDE
sp_tRNALys_TTT−t6A37/TsaBCDE
sp_tRNAIle_GAT−t6A37/TsaBCDE
sp_tRNAThr_TGT−t6A37/TsaBCDE
sp_tRNAAsn_GTT−t6A37/TsaBCDE
sp_tRNAArg_TCT−t6A37/TsaBCDE
sp_tRNAArg_CCT−t6A37/TsaBCDE
sp_tRNASer_GCT−t6A37/TsaBCDE
sp_tRNAIni_CAT−t6A37/TsaBCDE
sp_tRNAGln_TTG−m1A22/TrmK
sp_tRNATyr_GTA−m1A22/TrmK
sp_tRNASer_GGA−m1A22/TrmK
sp_tRNASer_TGA−m1A22/TrmK
sp_tRNACys_GCA−m1A22/TrmK
sp_tRNAGlu_TTC−m1A22/TrmK
sp_tRNAHis_GTG−m1A22/TrmK
sp_tRNASer_GCT−m1A22/TrmK
F
GLORI
S.pneumoniae tRNAs
m1A22/TrmK
Gln_TTG
Tyr_GTA
Ser_GGA
Ser_TGA
Cys_GCA
Glu_TTC
His_GTG
Ser_GCT
i6A37/MiaA
Leu_CAA
Leu_TAA
Ser_TGA
Tyr_GTA
Cys_GCA
Trp_CCA
A*37/TrmN6
Met_CAT
Val_TAC
Gly_TCC
t6A37/TsaBCDE
Thr_GGT
Lys_CTT
Lys_TTT
Ile_GAT
Thr_TGT
Asn_GTT
Arg_TCT
Arg_CCT
Ser_GCT
Ini_CAT

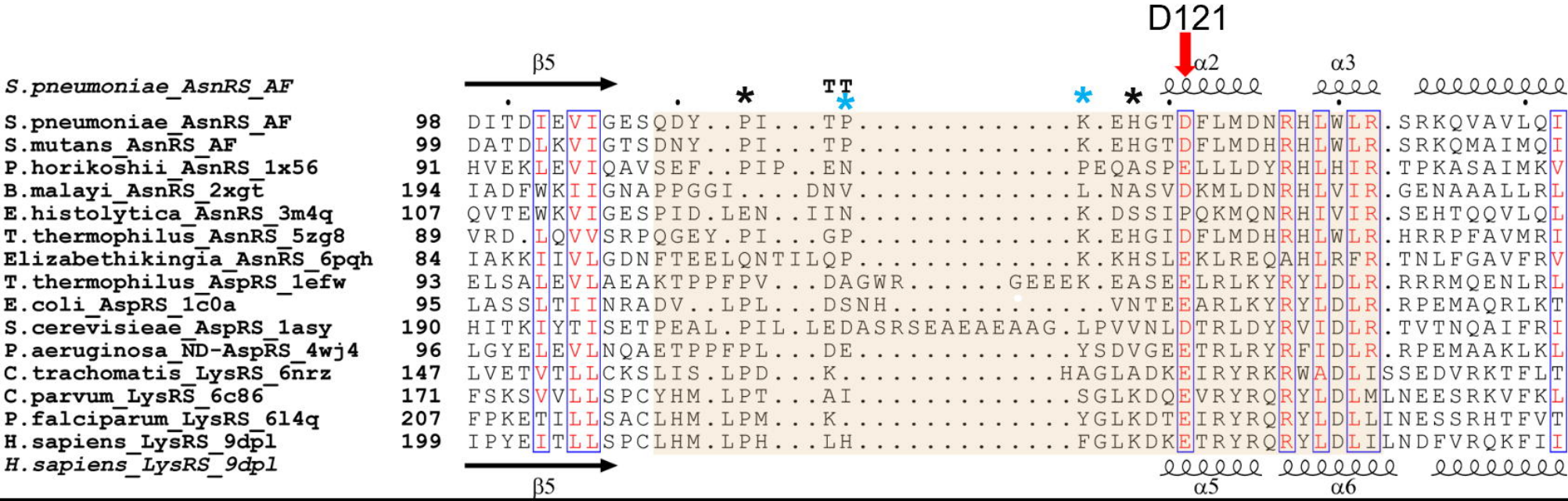

A
preQ0
preQ1
S=SPD_0898 T=SPD_2046
A=SPD_2047 A'=SPD_2048
S T
A A'
QueF
SPD_1568
(non-functional)
TGT
SPD_1868
preQ1-tRNA34
QueA
SPD_1247
oQ-tRNA34
B
preQ1
oQ
Q
Normalized peak area
<LOD
DMSO
preQ0 base
preQ1 base
C
Steptococcus strains
19.1%
0.8%
63.9%
15.1%
1.1%
de novo Q
preQ0 salvage Q
preQ1 salvage Q
preQ1 salvage oQ
Other
D
Q precursor transport
Q biosynthesis
GTP
preQ0
preQ1
preQ1-tRNA
oQ-tRNA
Q-tRNA
QueD
QueE
QueC
QueF
TGT
QueA
QueG/QueH
QueT/QPTR
206
99
7
5
120
112
755
30
5
103
117
13
9
18
Other
Enzyme Presence
Enzyme Absence
Pseudogene
Fusion QueCF
E
1
10
100
1000
S. thermophilus
S. salivarius
S. mutans
S. sp.
Other
S. oralis
S. mitis
S. equinus
S. gallolyticus
S. sp.
Other
S. thermophilus
S. sobrinus
S. lutetiensis
S. sanguinis
S. parauberis
S. parasuis
S. sp.
Other
S. equi
S. anginosus
S. gordonii
S. constellatus
S. intermedius
S. sp.
Other
S. pyogenes
S. agalactiae
S. suis
S. dysgalactiae
S. iniae
S. sp.
Other
S. pyogenes
S. suis
S. pseudoporcinus
S. agalactiae
S. parasuis
S. suis
S. pneumoniae
S. pneumoniae
S. pneumoniae
S. suis


A
B
Distribution of Q genes and taxonomy of QueF/QueCF in Streptococcus
QueF
QueCF
QueDEC,TGT
311
QueDE*C,TGT
8
QueDE*C*,TGT
5
S. thermophilus 132
S. salivarius 38
S. mutans 31
S. oralis 26
S. mitis 19
S. sp. 17
S. equinus 14
S. gallolyticus 13
Other 34
TGT+QueA
220
S. pneumoniae
220


A
No. of insertions
20
Window size: 1
spd_1740, cinA
spd_1741, lytR
spd_1742
spd_1743, tsaE
spd_1744, comM
B
OD620
Time (h)
DT (min) yield(A620)
WT 38 ± 0 0.99 ± 0.0
ΔtsaE (sup1) 64 ± 1.4 0.66 ± 0.02
ΔtsaE//Pzn-tsaE) 36 ± 1.4 1.00 ± 0.04
C
WT
ΔtsaE (sup1)
ΔtsaE//Pzn-tsaE
1 µm
D
Length (µm)
Width (µm)
Aspect ratio (l:w)
Relative Cell Volume
ns
****
WT ΔtsaE ΔtsaE//Pzn-tsaE
Mean 1.12 1.04 1.13
SEM ±0.02 ±0.02 ±0.02
(n) 188 188 188
Mean 0.68 0.61 0.68
SEM ±0.01 ±0.00 ±0.01
(n) 188 188 188
Mean 1.65 1.71 1.66
SEM ±0.02 ±0.02 ±0.02
(n) 188 188 188
Mean 1.04 0.747 1.02
SEM ±0.03 ±0.02 ±0.02
(n) 188 188 188


A
IU1824 (D39 Δcps rpsL1 background)
OD620
Time (h)
IU1824 (WT parent) 35 1.0
IU18824 (ΔtsaE sup1) 73 0.7
IU18826 (ΔtsaE sup2) 63 0.7
DT (min) yield (A620)
IU1945 (D39 Δcps rpsL+)
IU1945 (WT parent) 31 1.0
E797 (ΔtsaE sup3) 51 0.8
K797 (ΔtsaE sup4) 57 0.8
B
IU1690 35 ± 4 0.99 ± 0.0
IU1824 41 ± 5 1.0 ± 0
IU19598 41 ± 1 1.0 ± 0
IU18824 79 ± 11 0.65 ± 0.02
IU19699 68 ± 1 0.64 ± 0.04
C
****
Normalized peak area
IU1690
IU1824
IU19598
IU18824
IU19699
D
fold change
tsaE+
tsaE-
t6A
s4U
oQ
mnm5s2U
m2A
nm5s2U
m7G
m1A
Y
m5U
D
m1G
I
m6A
ac4C
mo5U
i6A
k2C
ho5U
cmnm5s2U
Q
acp3U
m6,6A
m3C
m2G
m5C
Um
Am
Gm
m2,2G
m3U
Cm


Catalytic domain
Acceptor stem
D-stem
Linker
Anticodon binding domain
Anticodon loop
C67
Asp121
His118
Lys116
Pro112
Pro115
t6A37


D121
β5
α2
α3
α5
α6
S.pneumoniae_AsnRS_AF
S.pneumoniae_AsnRS_AF 98 DITDIEVIGESQDY..PI...TP.....K.EHGTDFLMDNRHLWLR.SRKQVAVLQI
S.mutans_AsnRS_AF 99 DATDLKVIGTSDNY..PI...TP.....K.EHGTDFLMDHRHLWLR.SRKQMAIMQI
P.horikoshii_AsnRS_1x56 91 HVEKLEVIQAVSEF..PIP..EN.....PEQASPELLLDYRHLHIR.TPKASAIMKV
B.malayi_AsnRS_2xgt 194 IADFWKIIGNAPPGGI....DNV.....L.NASVDKMLDNRHLVIR.GENAAALLRL
E.histolytica_AsnRS_3m4q 107 QVTEWKVIGESPID.LEN..IIN.....K.DSSIPQKMQNRHIVIR.SEHTQQVLQL
T.thermophilus_AsnRS_5zg8 89 VRD.LQVVSRPQGEY.PI...GP.....K.EHGIDFLMDHRHLWLR.HRRPFAVMRI
Elizabethikingia_AsnRS_6pqh 84 IAKKIIVLGDNFTEELQNTILQP.....K.KHSLEKLREQAHLRFR.TNLFGAVFRV
T.thermophilus_AspRS_1efw 93 ELSALEVLAEAKTPPFPV...DAGWR.....GEEEK.EASEELRLKYRYLDLR.RRRMQENLRL
E.coli_AspRS_1c0a 95 LASSLTIINRADV..LPL...DSNH.....VNTEEARLKYRYLDLR.RPEMAQRLKT
S.cerevisieae_AspRS_1asy 190 HITKIYTISETPEAL.PIL.LEDASRSEAEAEAAG.LPVVNLDTRLDYRVIDLR.TVTNQAIFRI
P.aeruginosa_ND-AspRS_4wj4 96 LGYELEVLNQAETPPFPL...DE.....YSDVGEETRLRYRFIDLR.RPEMAAKLKL
C.trachomatis_LysRS_6nrz 147 LVETVTLLCKSLIS.LPD...K.....HAGLADKEIRYRKRWADLISSEDVRKTFLT
C.parvum_LysRS_6c86 171 FSKSVVLLSPCYHM.LPT...AI.....SGLKDQEVRYRQRYLDLMLNEESRKVFKL
P.falciparum_LysRS_614q 207 FPKETILLSACLHM.LPM...K.....YGLKDTEIRYRQRYLDLLINESSRHTFVT
H.sapiens_LysRS_9dpl 199 IPYEITLLSPCLHM.LPH...LH.....FGLKDKETRYRQRYLDLILNDFVRQKFII
H.sapiens_LysRS_9dpl