Plant, Animal, and Fungal Micronutrient Queuosine Is Salvaged by Members of the DUF2419 Protein Family

Rémi Zallot,[†] Céline Brochier-Armanet,[‡] Kirk W. Gaston,[§] Farhad Forouhar,[‖] Patrick A. Limbach,[§] John F. Hunt,[‖] and Valérie de Crécy-Lagard*,[†,⊥]

[†]Department of Microbiology and Cell Science, University of Florida, Gainesville, Florida 32611, United States

[‡]Université Lyon 1, CNRS, UMR5558, Laboratoire de Biométrie et Biologie Evolutive, Université de Lyon, 69622 Villeurbanne, France

[§]Rieveschl Laboratories for Mass Spectrometry, Department of Chemistry, University of Cincinnati, Cincinnati, Ohio 45221, United States

[‖]Department of Biological Sciences and Northeast Structural Genomics Consortium, Columbia University, New York, New York 10027, United States

[⊥]University of Florida Genetics Institute, Gainesville, Florida 32611, United States

S *Supporting Information*

ABSTRACT: Queuosine (Q) is a modification found at the wobble position of tRNAs with GUN anticodons. Although Q is present in most eukaryotes and bacteria, only bacteria can synthesize Q *de novo*. Eukaryotes acquire queuine (q), the free base of Q, from diet and/or microflora, making q an important but under-recognized micronutrient for plants, animals, and fungi. Eukaryotic type tRNA-guanine transglycosylases (eTGTs) are composed of a catalytic subunit (QTRT1) and a homologous accessory subunit (QTRTD1) forming a complex that catalyzes q insertion into target tRNAs. Phylogenetic analysis of eTGT subunits revealed a patchy distribution pattern in which gene losses occurred independently in different clades. Searches for genes co-distributing with eTGT family members identified DUF2419 as a potential Q salvage protein family. This prediction was experimentally validated in *Schizosaccharomyces pombe* by confirming that Q was present by analyzing tRNA^{Asp} with anticodon GUC purified from wild-type cells and by showing that Q was absent from strains carrying deletions in the QTRT1 or DUF2419 encoding genes. DUF2419 proteins occur in most Eukarya with a few possible cases of horizontal gene transfer to bacteria. The universality of the DUF2419 function was confirmed by complementing the *S. pombe* mutant with the *Zea mays* (maize), human, and *Sphaerobacter thermophilus* homologues. The enzymatic function of this family is yet to be determined, but structural similarity with DNA glycosidases suggests a ribonucleoside hydrolase activity.

Q (Queuosine) derivative originating from the degradation of Q modified tRNA
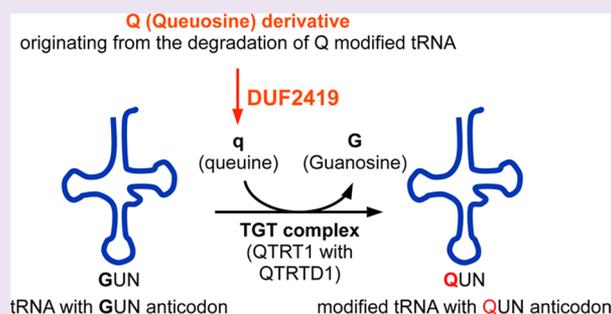
Queuosine (Q) and derivatives are modifications of the wobble base (position 34) of tRNAs with GUN anticodons that incorporate His, Tyr, Asp, or Asn amino acids.[1−3] Even if the Q modification is widely distributed in Bacteria and Eukarya, its origin differs between these two kingdoms;[3] most bacteria synthesize Q *de novo*, whereas all eukaryotes solely rely on salvage from the environment to acquire this complex modification (Figure 1). Bacteria make Q through the synthesis of the 7-aminomethyl-7-deazaguanine base (preQ$_1$) intermediate[4] in five steps from GTP by enzymes encoded by the *folE*, *queD*, *queE*, *queC*, and *queF* genes (Figure 1).[3] PreQ$_1$ is exchanged with guanine at position 34 of the target tRNA by a bacterial type tRNA guanosine(34) transglycosylase (EC 2.4.2.29) (bTGT).[5,6] Two subsequent enzymatic steps catalyzed by QueA (tRNA preQ$_1$(34) S-adenosylmethionine ribosyltransferase-isomerase; EC 2.4.99.17) and QueG (tRNA epoxyqueuosine(34) reductase) produce the final Q nucleoside.[7,8]

Eukaryotes cannot synthesize Q *de novo* and rely on their diet and/or microflora to obtain the Q precursor[9−14] (Figure 1). In contrast with the homodimeric bTGTs[15] that cannot use the queuine base (q) as substrate,[16] the eukaryotic type trans-glycosylases (eTGTs) catalyze the direct insertion of q in target tRNAs[9] (Figure 1). [Note that the standard nomenclature uses "Q" to represent the ribonucleoside and "q" to represent the corresponding base.] The eTGT enzyme is a heterodimer that consists of a catalytic subunit (QTRT1) and a regulatory subunit (QTRTD1), which is probably involved in tRNA binding.[17] Both subunits are homologous to bTGT, but
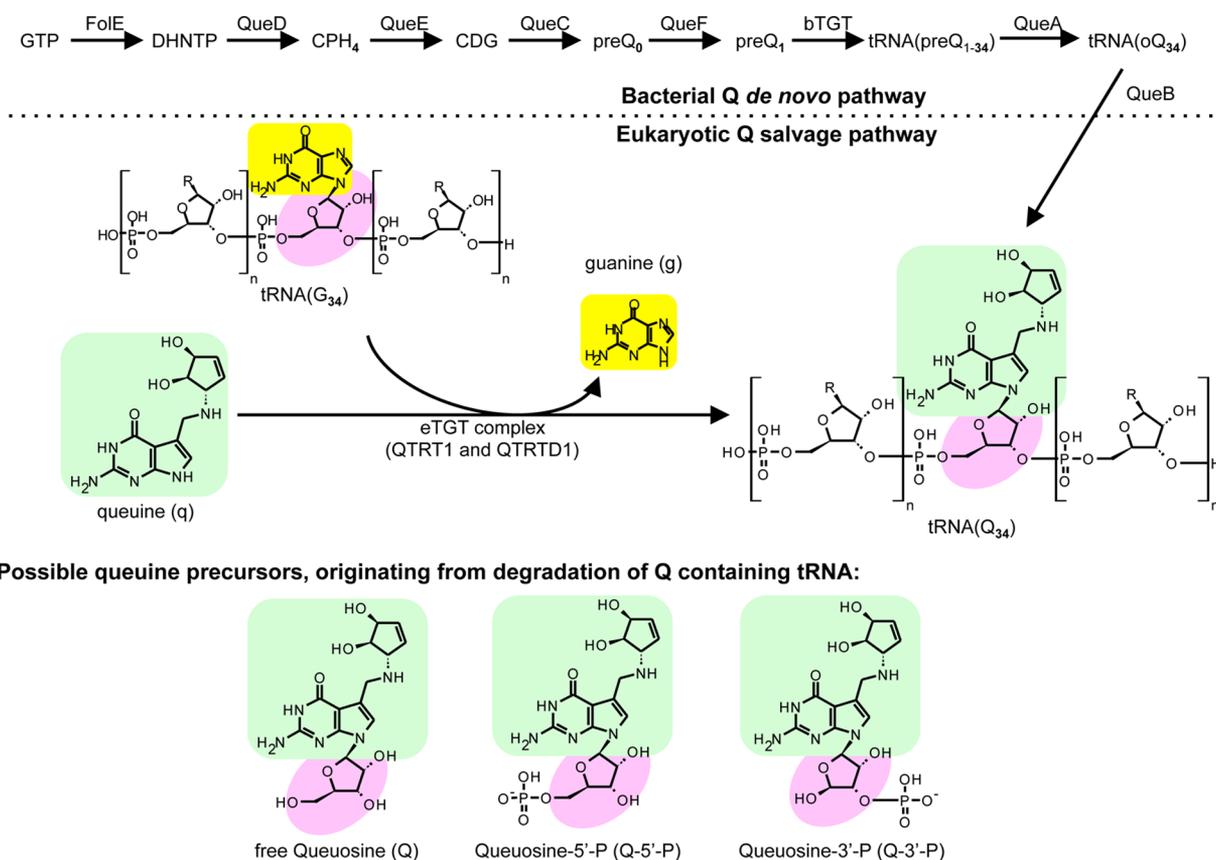
Figure 1. Queuosine structure and salvage pathway. Queuosine (Q) is synthesized *de novo* in bacteria. In eukaryotes, the last step of Q salvage occurs with the eukaryotic tRNA-guanine transglycosylase complex (eTGT) composed of QTRT1 and QTRTD1, responsible for the exchange of guanine (g) from a tRNA containing guanosine at position 34 (tRNA(G$_{34}$)) with queuine (q) to obtain a queuosine containing tRNA (tRNA(Q$_{34}$)). Free queuosine (Q), queuosine-5′-phosphate (Q-5′-P), and queuosine-3′-phosphate (Q-3′-P) have been identified as possible precursors for Q salvage but the q conversion steps are unknown. Other abbreviations are GTP: guanosine-5′-triphosphate; DHNTP: 7,8-dihydroneopterin triphosphate; CPH$_4$: 6-carboxy-5,6,7,8-tetrahydropterin; CDG: 7-carboxy-7-deazaguanine; preQ$_0$: 7-cyano-7-deazaguanine; preQ$_1$: 7-aminomethyl-7-deaza-guanine; tRNA(preQ$_{1-34}$): preQ1 at the position 34 of tRNA; tRNA(oQ$_{34}$): epoxyqueosine at the position 34 of tRNA; FolE: GTP cyclohydrolase-1; QueD: 6-carboxytetrahydropterin synthase; QueE: 7-carboxy-7-deazaguanine synthase; QueC: 7-cyano-7-deazaguanine synthase; QueF: 7-cyano-7-deazaguanine reductase; bTGT: bacterial type tRNA transglycosylase; QueA: tRNA (preQ$_{1-34}$) *S*-adenosylmethionine ribosyltransferase-isomerase; QueG: tRNA epoxyqueuosine (34) reductase.

QTRT1 is more similar to the bacterial enzyme family (40% identity) than QTRTD1 (20% identity).[18] The bTGT protein and the eTGT protein complex are both irreversible enzymes for their natural substrates.[19] While it is well-established that the q base is the preferred substrate for the eTGT complex,[16] the chemical pathway for q salvage has yet to be elucidated. Q-5′-phosphate, Q-3′-phosphate, and Q derived from Q-containing-tRNA have all been postulated as intermediates in the salvage pathway (Figure 1),[20−22] implying the involvement of unidentified but specific nucleosidases,[13,20] but the evidence is always indirect. Similarly, no information is available yet concerning the genes encoding potential transporters and phosphatases involved in Q synthesis from a precursor.

The degree of Q modification in Eukaryotes varies with the isoaccepting tRNA, the tissue, and the developmental stage,[23−25] and the phenotypes caused by the absence of Q in tRNA vary greatly with the organism. Some species such as *Saccharomyces cerevisiae* and *Candida albicans* do not harbor Q in tRNAs and thus do not salvage q.[26,27] The absence of Q leads to no obvious phenotypes in *Clamydomonas reinhardtii*[13] or in *Caenorhabditis elegans* under different stress conditions,[14] even though they incorporate it into tRNA when it is available. By contrast, Q-deficient drosophila are more sensitive to

cadmium stresses,[28] and Q levels in *Dictyostelium discoideum* influence lactate dehydrogenase activity[29] and also aggregation behavior, the latter possibly through the regulation of cyclic-AMP levels.[30] The most dramatic phenotypes were observed in mammals, in which the absence of both Q and tyrosine cause severe symptoms ultimately leading to death,[31] which establishes the Q precursor q as a micronutrient[32] or even a vitamin[24] for these organisms. Tyrosine is a nonessential amino acid that can be synthesized from phenylalanine by phenylalanine hydroxylase (PAH) that requires the biopterin cofactor BH$_4$.[33] It was recently shown that Q is required to protect BH$_4$ from oxidation by an undetermined mechanism.[34] Multiple reports linking the Q modification to defense systems, regulation of metabolism, cell proliferation and malignancy, cell signaling, and cancer have been published recently.[35] The molecular mechanisms underlying these diverse phenotypes are not yet understood; tRNAs lacking Q could be degraded,[36] specific proteins could be mistranslated because Q has been shown to influence codon-anticodon interaction,[37] or eTGT could have roles other than modification of tRNAs.[34,38]

In summary, both the salvage and function of Q in eukaryotes remain elusive, and this hampers understanding of the importance of this micronutrient in human physiology. We

set out to identify the uncharacterized eukaryotic Q-salvage genes using comparative genomic approaches that we had previously used to identify the bacterial *de novo* Q synthesis genes.[39−41]

RESULTS AND DISCUSSION

Phylogenetic Distribution of Known Queosine Metabolism Genes in Eukaryotes. TGT is the only signature enzyme family of the Q pathway. First, for the few eukaryotic organisms that have had tRNAs sequenced, there is a strict correspondence between the presence of a *tgt* gene (bTGT or QTRT1, the catalytic subunit of the eTGT complex) in the genome and the demonstrated presence of Q in tRNAs (Table 1). Furthermore, specific physiological scenarios can make all other genes of the Q pathway dispensable.[42,43] For example, when preQ$_1$ can be salvaged, then the bacterial *queDCEF* genes become dispensable. When q can be salvaged, *queA* and *queG* also become dispensable. Not all eukaryotes harbor Q in tRNAs, and surprising differences are found between closely related species. For example, *Arabidopsis thaliana* lacks Q,[26,44] while it is found in *Triticum* spp. (wheat)[12] (Table 1).

To analyze patterns of gene occurrence related to the Q pathway, we chose a set of 103 nonredundant eukaryotic genomes, 93 from OrthoMCL[45] and an additional 10 from organisms that have been demonstrated experimentally to harbor Q (Table 1). These organisms show an irregular phylogenetic distribution of the genes encoding the two subunits of the eTGT complex, the catalytic subunit QTRT1 and the homologous accessory subunit QTRTD1 (Figure 2A). Ninety-three genomes (90.3%) encode at least one QTRT1 or QTRD1 homologue. However, 31 genomes (30.1%) encode only QTRT1, while 4 genomes (3.9%) encode only QTRDT1 (Figure 2B). Interestingly, QTRT1/QTRD1 homologues are found in most plants but are absent in the model plant *Arabidopsis thaliana*. Further investigation of the *Brassicacae* expressed sequence tag (EST) databases using tBLASTn showed that none have homologues to QTRT1 or QTRDT1, showing that Q is probably globally absent from this phylogenetic order.

The biggest degree of variation in gene content was observed in the Fungi, where no clear pattern of presence or absence of these genes emerged; some ascomycetes such as *Schizosaccharomyces pombe*, *Neurospora crassa*, *Yarrowia lipolytica*, and *Aspergillus oryzae* have QTRT1 and QTRTD1 homologues, but others such as *S. cerevisiae*, *Magnaporthe oryzae*, and *Candida albicans* do not. The genera *Toxoplasma*, *Neospora*, and *Plasmodium* encode only QTRT1, whereas several species such as *Aspergillus nidulans* (*Emericella nidulans*), *Gallus gallus*, and *Ornithorhynchus anatinus* lack QTRT1, the supposed catalytic subunit of eTGT (Figure 2A). Further work will be required to determine whether the generally noncatalytic QTRTD1 subunit has acquired catalytic activity in these organisms or if they lack Q altogether in their tRNA. Overall, the near ubiquitous presence of TGT homologues along most branches of the eukaryotic phylogenetic tree suggests the corresponding gene was present in the last eukaryotic common ancestor (LECA) and that the observed gene absences reflect gene losses that occurred independently in different clades/species during the diversification of Eukarya.

While a set of candidate Q salvage genes have previously been identified by comparative genomic methods in Bacteria, homologous proteins have not been identified in Eukarya. These include transporters and members of a nucleoside

hydrolase family.[46−48] Because the predicted Q salvage genes in Bacteria are restricted to that domain of life, we hypothesized that Q salvage enzymes in Eukarya are likely to be similarly restricted and absent from Bacteria. Archaea are not known to make or salvage Q. While they do encode TGT homologues in their genomes, these enzymes insert 7-cyano-deazaguanine (preQ$_0$) at position 15 of tRNA before its transformation into Archaeosine (G$^+$).[49] Therefore, we expected that the Q salvage enzymes in Eukarya are also likely to be absent from Archaea, although they could potentially encode related enzymes involved in salvage of preQ$_0$.

Identification of DUF2419 as a Candidate Q Salvage Gene by Comparative Genomic Analysis. On the basis of our analysis of the phylogenetic distribution of TGT subunits (Figure 2A and Table 1), we derived a signature phyletic profile for Q salvage genes. We postulated that eukaryotic Q salvage genes should be present in *D. discoideum*, *S. pombe*, *C. elegans*, *Drosophila megalonaster*, *H. sapiens*, and *Rattus norvegicus* but absent from *A. thaliana*, *S. cerevisiae*, *C. albicans*, and the subsets of Bacteria and Archaea present in the OrthoMCL databases. Using this profile as input, we queried the OrthoMCL platform[45] (Supplementary Figure S1A), which identified 21 protein families matching this pattern (Supplementary Figure S1B). As discussed above, QTRT1 and bTGTs are close homologues (∼40% identity)[18] and members of the same orthologous group (OG5_126946), which contains sequences from most bacterial genomes. As expected, this group was not retrieved by our phyletic profile, because it excluded families with bacterial members. In contrast, ortholog group OG5_129559 corresponding to QTRTD1 was retrieved, validating the phyletic profile used for the query (Supplementary Figure S1C). Among the 20 other ortholog groups retrieved, group OG5_130329, corresponding to the domain of unknown function 2419 (DUF2419), had the closest match to the phyletic profile of QTRTD1. The co-distribution of the two families was not perfect, as shown Figure 2B, but it was compelling enough to explore further.

According to Pfam, members of the DUF2419 family are mainly found in eukaryotes (224 of 236 or 94.9% of organisms encoding the domain), as expected for a eukaryotic enzyme involved in Q salvage. The small subset of sequences found in prokaryotes belong to organisms absent from the OrthoMCL database, which would have eliminated the DUF2419 family from the output of our phyletic query. Ten of the prokaryotic homologues are found in Bacteria such as *Sphaerobacter thermophilus* or *Conexibacter woesei* DSM 14684, while two are found in the Archaea *Aeropyrum pernix* (NP_147511.1) and *Pyrolobus fumarii* (YP_004781936.1). Consistent with these archaeal DUF2419 proteins having a diverged enzymatic function compared to the homologous nonarchaeal proteins, their sequences form a distinct and strongly divergent clade in DUF2419 (Supplementary Figure S2). Therefore, the archaeal sequences were not included in our subsequent analyses.

We extended our initial analysis by performing an exhaustive search for DUF2419 homologues in the *nr* database at the NCBI. Among the 374 additional DUF2419 homologues detected, 349 come from eukaryotic organisms, while 25 come from bacterial organisms belonging to a small number of diverse and unrelated taxa. In contrast, DUF2419 homologues are found in all major phyla composing the Eukarya Domain. A global phylogenetic analysis shows that the bacterial sequences form two distinct clusters nested within the eukaryotic sequences in DUF2419 (Supplementary Figure S3).

Table 1. Species for Which Queosine (Q) or Queuine (q) Have Been Demonstrated Present or Absent and Corresponding Accession Numbers[a]

kingdom	species	tRNA	refs	bTGT accession	eTGT accessions		
					QTRT1	QTRTD1	DUF2419 accessions
Organisms with Sequenced Queuosine Containing tRNAs							
Bacteria	Azospirillum lipoferum	Asn	86	YP_005039595.1			
	Bacillus subtilis	Tyr	86	NP_390649.1			
	Escherichia coli	Asp, His, Tyr, Asn	86, 87	NP_414940.1			
	Geobacillus stearothermophilus	Tyr	86	bstearo.Contig544			
	Salmonella typhimurium	His	86	ESE73596.1			
Protozoa	Tetrahymena pyriformis	Tyr	88				
Animals	Bos taurus	Asn, Asp	86		NP_001193621.1	NP_001092378.1	NP_001070446.1
	Drosophila melanogaster	Tyr, His	86, 89		NP_608585.1	NP_648320.1	NP_611573.1
	Homo sapiens	Asn	86		NP_112486.1	NP_078914.1	NP_115683.3
	Ovis aries	His	86		NP_004009348.1	NP_004002979.1	XP_004004160.1
	Rattus norvegicus	Asn, Asp[c], Asp[b]	86		NP_071586.2	NP_001067760.2	NP_001166907.1
	Xenopus laevis	Asp[b]	87		NP_001089529.1	NP_001086589.1	NP_001086304.2
	rabbit (sequence from Oryctolagus cuniculus)	Asp[b]	87		XP_002723326.1	XP_002716877.1	XP_002721316.1
Plant	Triticum aestivum	Tyr	86		CA68945, BQ295388 (EST partial match)		CD866774, CJ623453, CJ955180, CJ588187 (EST partial match)
Other Organisms with Queuosine Containing tRNAs (Not Sequenced)							
Animals	Drosophila megalonaster	Asn, Asp	90		see above	see above	see above
	Lucilia sericata	His, Asn, Asp, Tyr	91				
	Musca domestica	His, Asn, Asp, Tyr	91		XP_005181648.1	XP_005184760.1	XP_005176381.1
	Tenebrio molitor	His, Asn, Asp, Tyr	91				
	Monodelphis domestica	Asp	95		XP_001365875.1	XP_001368404.2	XP_001379226.2
	Caenorhabditis elegans	unspecified (bulk tRNA)	14		NP_502268.1	NP_741662.1	NP_S00924.2
Amoeba	Dictyostelium discoideum	Asn	11		XP_629936.1	XP_643448.1	XP_636598.1
Green algae	Chlamydomonas reinhardtii	unspecified (bulk tRNA)	13		XP_001702820.1	XP_001702426.1	XP_001692426.1
Plant	Chlorella pyrenoidosa (sequence from Chlorella variabilis NC64A)	unspecified (bulk tRNA)	13		IGS.gm_3_00384	IGS.gm_32_00025	IGS.gm_14_00360
	Lupinus luteus	Asn, His, Tyr	92–94				
Other Organisms Where Free Queuine Has Been Identified							
Plant	Cocos nucifera		1				
	Solanum lycopersicum		1		XP_004237947.1	XP_004251975.1	XP_004231913.1
Organisms Where the Absence of Queuosine Has Been Experimentaly Demonstrated							
Fungi	Saccharomyces cerevisiae		26				
	Candida albicans		27				
Plant	Arabidopsis thaliana[d]		44				

[a]Accession numbers were obtained using NCBI BlastP against each organism, with the human protein sequences as input. When no hits were found, NCBI tBlastN was used against EST (eukaryote) or genomic sequences (prokaryotes). An empty entry means that no hits were found. [v]JGI was used for *Chlorella variabilis* (instead of *Chlorella pyrenoidosa*, see: http://genome.jgi-psf.org/ChlNC64A_1/ChlNC64A_1.home.html), and for *Geobacillus stearothermophilus*, sequences were found using tBlastN at http://www.genome.ou.edu/bstearo.html. [b]Modification β,D-mannosylqueuosine detected, implying the existence of the Q modification. [c]Mitochondrial tRNA. [d]Q has never been demonstrated present in *Arabidopsis thaliana*, and Chen et al. suggest that it is absent in *Arabidopsis thaliana* since there are no TGT genes present in this organism, and indeed they could not find it in their bulk tRNA analysis. These authors mention that it could still be there but unstable; however, Q is stable, whereas its glutamylated derivative is not.[96]
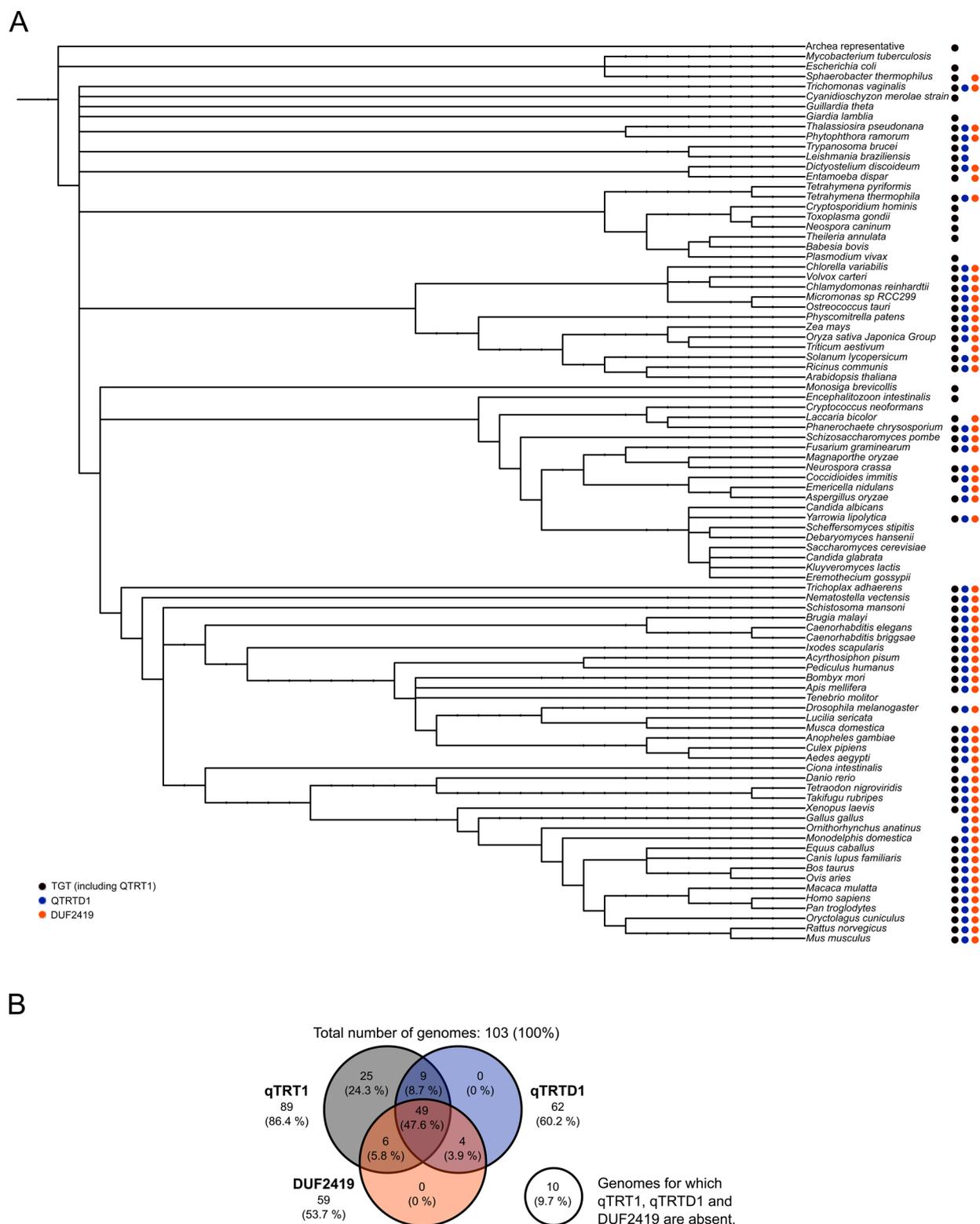
Figure 2. Taxonomic distribution of Q salvage genes. (A) Representative presence/absence pattern of TGT (including QTRT1, as they are in the same orthologous group in orthoMCL), QTRTD1, and DUF2419. Redundancy in clade has been limited. (B) Diagram presenting the full results of the distribution analysis from the data obtained at orthoMCL.org, enriched with NCBI Blast search for organism in which Q has been demonstrated present. Note that these representations are not informative of the number of gene occurrences per genome.

In order to investigate the phylogeny of this family more deeply, we performed a more detailed phylogenetic analysis of a subset of 75 sequences representative of the taxonomic and genetic diversity of DUF2419 homologues. The resulting Bayesian and maximum likelihood trees recovered a monophyletic structure for nearly all eukaryotic phyla (Figure 3), suggesting that DUF2419 is ancient in this Domain of life and may have been present in the LECA. The main exception

Figure 3. Detailed Bayesian phylogeny of DUF2419 (75 sequences, 188 positions). The tree was inferred with MrBayes software. Bacterial sequences are show in black, whereas eukaryotic sequences are shown in color according to their taxonomic affiliation. Supports at branches correspond to posterior probabilities (PP) computed by MrBayes/bootstrap values (BV) computed with PhyML. For clarity, only PP > 0.5 and BV > 50% are shown. The scale bar represents the estimated average number of substitutions per site. Red arrows show sequences that have been used for experimental complementation in *Schizosaccharomyces pombe*.

concerned Amoebozoa. In fact, while Centramoebida emerged at their expected position with other Opisthokonta (i.e., Fungi, Metazoa, and Ichthyosporea (posterior probability (PP = 1.0 and bootstrap value (BV) = 68%), Archamoebae and Mycetozoa grouped with different lineages of Stramenopiles (Figure 3). This grouping suggests that the former acquired DUF2419 from the latter.

Relative to the bacterial sequences, our phylogenetic analysis confirmed the presence of distinct clusters nested within eukaryotic sequences (Figure 3). While cluster 1 was again strongly recovered (PP = 1.0 and BV = 99%), cluster 2 was not significantly supported (PP < 0.5 and BV < 50%), suggesting that it could correspond in fact to two distinct clusters: cluster 2-a (PP = 0.87 and BV = 54%) and cluster 2-b (PP = 0.98 and BV = 57%, respectively; Figure 3). The taxonomic distribution of bacterial sequences within these clusters is totally at odds with the current taxonomy of this Domain, suggesting that it resulted from horizontal gene transfers among unrelated bacterial lineages. This observation combined with the presence of at least two (maybe three) distinct bacterial clusters strongly suggests that bacterial DUF2419 genes were acquired from

distinct eukaryote donors via independent horizontal gene transfers. All of the bacteria that harbor a DUF2419 member also have a TGT homologue. One of these, *C. woesei*, lacks all preQ$_1$ synthesis genes as well as the *queA* and *queG* genes, suggesting that, as in Eukaryotes, q is salvaged in this organism. The substrate binding pockets of the different TGT families have been well characterized,[16] and as shown in Supplementary Figure S4, many bacteria that have a DUF2419 have non-canonical substrate binding residues, suggesting that in these organisms bTGT could recognize q.

Domain analysis of members of the DUF2419 family revealed that in *Caenorhabditis briggsae* the DUF2419 homologue is fused to domains related to PseudoUridine synthase and Archaeosine transglycosylase (PUA) (Supplementary Figure S5A, and accession number XP_002633628).[50] PUA domains are tRNA binding domains found in RNA modification enzymes.[51,52] However, the sequence of this fusion protein failed to yield any significant hits crossing the DUF2419-PUA domain boundary when used to query a database of expressed sequence tags (ESTs) from *C. briggsae*. This observation suggests that the fusion protein sequence

Figure 4. Results of LC−MS/MS analysis of *S. pombe* tRNA extracted from different strains. (A) Quantitation of queuosine (Q) content for different strains used in purified tRNA$^{Asp}_{GUC}$ extracted from WT (left panel), *duf2419Δ* (middle panel), and *duf2419Δ* complemented with *S. pombe* DUF2419 (right panel, *duf2419Δ* + SpDUF2419). UV traces (top) at 254 nm and the extraction ion chromatograms (bottom) expected for Q ([M + H]$^+$ *m/z* 410) are shown. (B) MS/MS fragmentation profile of *m/z* 410 confirming the identification of the nucleoside Q in the WT strain. The blue area highlights the elution time for queuosine. (C) Complementation of the Q$^-$ phenotype analyzed by LC−MS/MS. The ratios of Q/m^{5}C in tRNA$^{Asp}_{GUC}$ purified from the *S. pombe* WT strain and the *duf2419Δ* strain complemented with an empty vector (*duf2419Δ* + Ø), *S. pombe* DUF2419 (*duf2419Δ* + SpDUF2419), human DUF2419 (*duf2419Δ* + HsDUF2419), maize DUF2419 (*duf2419Δ* + ZmDUF2419), or *S. thermophilus* DUF2419 (*duf2419Δ* + StDUF2419) are shown by the filled bars. To control for the amount of tRNA, the m^{5}C content in the complemented strains was compared to the m^{5}C content in the WT control (white bars). The results of a typical experiment are presented.

might represent a misannotation of two different genes physically adjacent to one another on a chromosome. However, functionally associated genes are known to cluster together on the chromosomes of organisms of nematodes, this association still reinforces the possibility of the DUF2419 family being involved in RNA modification.[53]

Finally, to explore other types of functional associations, we used the *S. pombe* DUF21419 gene (SPAC589.05c) as input in the STRING database (Search Tool for the Retrieval of Interacting Genes/Protein; http://string-db.org).[54] The top association detected was the fusion with the PUA domain protein discussed above (Supplementary Figure S5B). All of the additional associations revealed co-expression with proteins involved in nucleoside/nucleotide metabolism.

The combination of phylogenetic distribution, domain fusion analysis, and co-expression analysis strongly supported the

hypothesis that DUF2419 encodes an enzyme involved in Q salvage. Therefore, we set out to test this hypothesis experimentally.

S. pombe tRNA$^{Asp}_{GUC}$ Contains Q at Position 34. *S. pombe* was used as a model organism in this study because it grows easily and has the genes encoding the QTRT1, QTRTD1, and DUF2419 proteins. Furthermore, a comprehensive collection of single-gene knockout mutants is available for this organism.[55] As for other Eukaryotes, the lack of *de novo* Q biosynthesis genes in *S. pombe* implies that the presence of Q is dependent on salvage pathways and that this nucleoside has to be present in the growth media for incorporation into tRNA.

Bactopeptone has been shown to be a source for Q in *Chlorella pyrenoidosa* and *Chlamydomonas reinhardtii*.[13] To test whether bactopeptone can also be a source for Q in *S. pombe*, the WT strain was grown in a commercial YPD preparation

(BD Difco YPD: 1% yeast extract, 2% peptone, 2% dextrose) or in a preparation of 0.5% yeast extract (Difco), 2% bactopeptone (Difco), and 3% dextrose (Fluka). After purification of tRNA$^{Asp}_{GUC}$ from bulk tRNA preparations, it was digested and dephosphorylated to generate ribonucleosides for LC−MS/MS analysis (Figure 4A,B). The 410 m/z ion that corresponds to the protonated molecular weight ([M + H]$^+$) of Q was detected at 21.3 min in tRNA extracted from the WT *S. pombe* strain grown with bactopeptone in the media, whereas it is not detected from a similar tRNA preparation of WT *S. pombe* strain grown in YPD (with a similar amount of tRNA for both growth conditions, Supplementary Figure S6A). On the basis of these mass spectrometric experiments, we concluded that *S. pombe* is able to salvage Q from media containing bactopeptone, similarly to what has been previously described for *C. pyrenoidosa* and *C. reinhardtii*,[13] while YPD does not contain enough of the necessary precursor for Q salvage. Therefore, in all subsequent experiments, we grew *S. pombe* strains in 0.5% yeast extract, 2% bactopeptone, and 3% dextrose.

The Q-containing purified tRNA$^{Asp}_{GUC}$ extracted from WT *S. pombe* was digested with RNase T1 and analyzed by LC−MS/MS to infer the sequence (Figure 5). The resulting digestion products consisted of one sequence containing the anticodon region with the sequence CCU[Q]UCA[m^{5}C]Gp. The total ion chromatogram (TIC) depicts total signal from the eluted digestion products, while the extracted ion chromatogram (XIC) for m/z 1002.5 depicts the elution time of this sequence of interest (Figure 5). The mass spectra collected at this time point reveal three digestion products that elute at a similar time: CCU[Q]UCA[m^{5}C]Gp (m/z 1504.25^{-2} and 1002.58^{-3}), AAUCCCGp, and UACACAAG>p (where >p indicates a cyclic phosphate). Each ion was selected for tandem mass spectrometry by collision-induced dissociation (CID), generating product ions that match the expected ions for these sequences. This analysis provides information about the location of modifications in the RNA sequence, but modifications such as methylations can be found on multiple locations of the base or the sugar. To precisely identify the chemical structrure of the modifications, the isolated tRNA^{Asp} was also digested to nucleosides and analyzed by LC−UV−MS/MS (Supplementary Figure S7). Six modifications were observed from nucleoside digestions of this tRNA, including Q. In this chromatographic analysis, Q was not resolved from adenosine, as indicated by the XIC for the corresponding molecular ion ([M + H]$^+$ m/z 410) (Supplementary Figure S7). Because representative nucleoside fragment ions for Q were not observed in the mass spectrum, tandem mass spectrometry was used to verify the identity. As previously reported,[56] the Q base fragments upon CID to produce an ion lacking part of this modification (m/z 295.00) and a product ion (m/z 163.08) that results in the breaking of the glycosidic bond of the m/z 295.00 product (Supplementary Figure S7). Figure 6 shows the final sequence of *S. pombe* tRNA$^{Asp}_{GUC}$ as deduced from these detailed mass spectral analyses.

DUF2419 Family Is Involved in Q Salvage. To confirm that *S. pombe* salvages Q in an eTGT-dependent manner, we compared the Q content of tRNA$^{Asp}_{GUC}$ purified from WT *S. pombe* cells *versus* *qtrt1Δ* (deletion in SPAC1687.19c) derivatives. Whereas the amount of tRNA was similar in both strains, the Q content of tRNA$^{Asp}_{GUC}$ from the *qtrt1Δ* strain was below the detection limit of our experimental setup
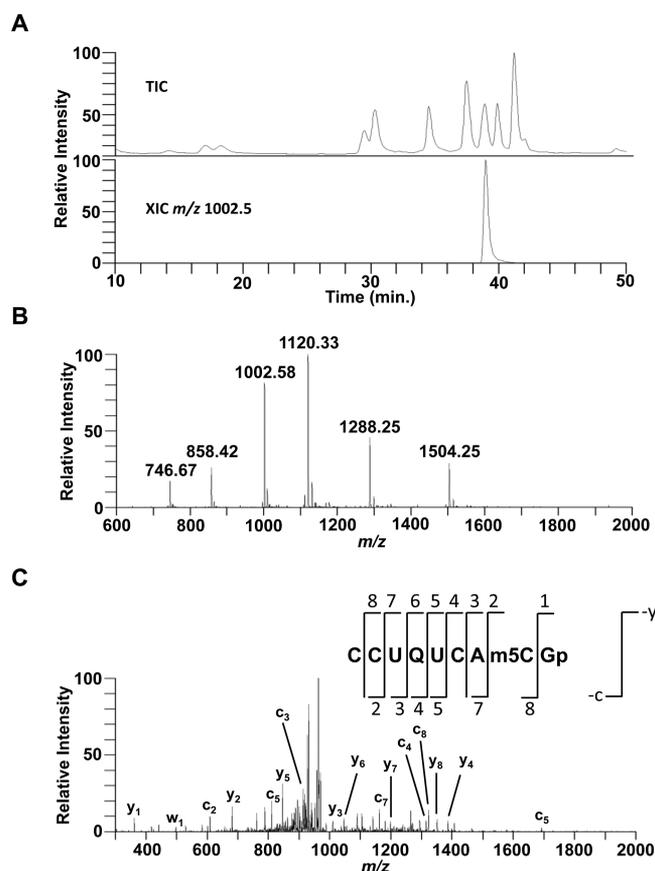
Figure 5. *Schizosaccharomyces pombe* tRNA^{Asp} contains queuosine at position 34. (A) LC−MS/MS was performed on T1 digested tRNA producing a total ion chromatogram (TIC). (B) A digestion product (m/z 1002.6) was found eluting at 39.0 min, as shown by the extracted ion chromatogram (XIC). The MS spectra at this time point depicts a signal from three oligonucleotides, CCU[Q]UCA[m^{5}C]Gp (m/z 1504.25^{-2} and 1002.58^{-3}), AAUCCCGp (m/z 1120.33^{-2} and 746.67^{-3}), and UACACAAG>p (m/z 1288.25^{-2} and 858.42^{-3}). (C) Collision induced dissociation of m/z 1002.58 produces the nearly all expected −c and −y ions for the sequence CCU[Q]UCA[m^{5}C]Gp.

(Supplementary Figure S6B), validating both the role of eTGT in Q salvage in *S. pombe* and our experimental approach.

Having confirmed that *S. pombe* salvages Q, we set out to test if DUF2419 was involved in this process. If DUF2419 is involved in an early step of Q salvage, then tRNA purified from a *duf2419Δ* strain of *S. pombe* should lack the Q ribonucleoside. Figure 4A and Supplementary Figure S6B show that tRNA extracted from the DUF2419 knockout strain lacked the 410 m/z ion, which corresponds to the protonated molecular weight ([M + H]$^+$) of Q detected mass spectrometrically at 20.3 min in the WT strain. This biochemical phenotype was complemented by expressing the *S. pombe duf2419* gene *in trans* (Figure 4A), indicating that Q salvage in *S. pombe* involves DUF2419.

Having shown that the *S. pombe* DUF2419 protein is involved in Q salvage, we tested whether the DUF2419 genes from human, maize, and *S. thermophilus* complemented the Q$^-$ phenotype of the *duf2419Δ S. pombe* strain. Figure 4B and Supplementary Figure S6 show successful complementation by these DUF2149 genes, albeit with different degrees of efficiency. Because the highly diverse members of the DUF2419 family tested (indicated by red arrows in Figure 3) all function in Q salvage, our complementation experiments
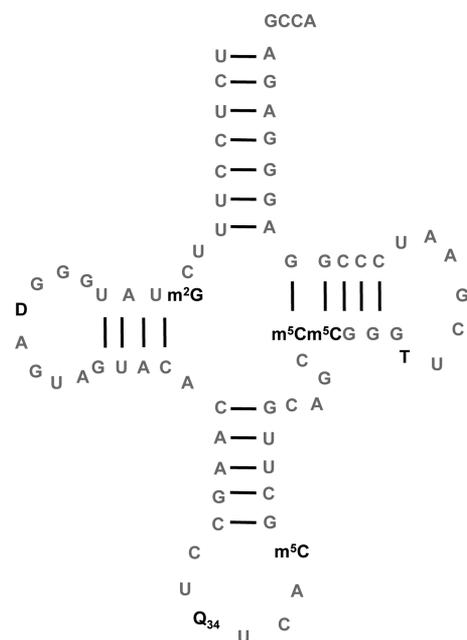
Figure 6. Sequence of *S. pombe* tRNA^{Asp} $_{GUC}$.

Figure 7. Homoloogy model for the structure of the active-site in the DUF2419 protein from *Sphaerobacter thermophilus* (StDUF2419). The stereo ribbon diagram shows superposition of the homology model generated by PHYRE2 for StDUF2419 (the product of the Sthe_2331 gene, shown in cyan) with the template structure used to generate this model, 8-oxoguanine (8OG) DNA glycosylase from *Pyrobaculum aerophilum* (PDB 1XQO, shown in yellow). The 8OG molecule and the residues that interact with it in this structure are shown in ball-and-stick representation, along with the equivalent residues in the homology model of StDUF2419. Carbon atoms from the proteins are shown in the same color as the protein backbones, while carbon atoms from 8OG are shown in purple. Oxygen and nitrogen atoms are shown in red and blue, respectively. The 3′ hydroxyl group of 8OG is labeled, while the magenta sphere indicates the location of the N7 atom of 8OG, which corresponds to the site of modification in queuosine.

suggests that this family is isofunctional at least in eukaryotes and the few bacteria harboring this gene.

DUF2419 Has Homology to a DNA Glycosylase. Possible biochemical activities for a Q salvage enzyme include functioning as a nucleoside transporter, a phosphatase, or a ribonucleoside hydrolase or playing some kind of accessory role in eTGT-dependent base exchange. Because standard sequence-analysis methods fail to detect any transmembrane α-helices in DUF2419,[57,58] it is unlikely that it plays a direct role in transmembrane transport. Also, purified eTGT has been demonstrated to exchange guanine with q in target tRNAs *in vitro* without any additional proteins.[17] An accessory role is therefore unlikely, although it cannot be rigorously excluded because such a function could be required *in vivo* and not *in vitro*.

A strong clue regarding the potential function of DUF2419 was provided by automated sequence profiling and structural modeling methods implemented in the program PHYRE2.[59] This program uses sequence profile–profile alignments to find potentially homologous proteins with an experimentally characterized structure and constructs 3-dimensional models for each candidate using homology-modeling methods. Using the DUF2419 sequence from *Sphaerobacter thermophilus* (StDUF2419) as input, PHYRE2 identified 8-oxoguanine (8OG) DNA glycosylase from *Pyrobaculum aerophilum*[60] (PDB id: 1XQO) as the highest scoring hit, with a confidence score of 94.4% for alignment of residues 199−252 and 274−308 in the query domain (Supplementary Figure S8A), which have 14% identity to the corresponding residues in the structural template. This enzyme belongs to a family that cleaves oxidized guanines in DNA.[61] According to the SCOP database,[62] it has AgoG-like fold within the DNA glycosylase fold superfamily.

The partial structural model for DUF2419 generated using this template includes a set of strongly conserved residues in DUF2419 (Supplementary Figure S8B) that align with residues directly contacting the 8OG product in structure 1XQO. After computational minimization of the homology model from PHYRE2, structural superposition shows that residues D231,

K233, D298, and W302 in StDUF2419 correspond to active-site residues D172, R174, D218, and W222 in 1XQO (Figure 7 and Supplementary Figure S8B). This analysis suggests that these residues in DUF2419 are involved in recognition of a guanine-like substrate. The homology model shows a steric clash with the ribose group of 8OG by the side chains of residues I203 and F229 in StDUF2419, which replace residues F144 and P170 in 1XQO, suggesting a difference in the stereochemistry of substrate binding in DUF2419. Intriguingly, two invariant basic residues in DUF21419 (K199 and R200 in StDUF2419) are located adjacent to this site in the homology model, suggesting that its physiological substrate may contain a phosphate group. Moreover, the N7 atom of 8OG is located in a region where PHYRE2 was unable to build a homology model for StDUF2419, suggesting a significant difference in the chemical structure of its substrate in the vicinity of this atom, which is the site of modification of the guanine base in queuosine.

On the basis of these results from homology modeling, we hypothesize that DUF2419 has ribonucleoside hydrolase activity and interacts with queuosine-3′-phosphate as either a substrate or a product (Figure 1).

Conclusion. The work reported in this paper represents another successful example of combining integrative data mining with experimental validation to discover the function of previously uncharacterized or "unknown" proteins.[43,63] Automated annotation based primarily on analysis of sequence similarity failed to predict a function for the DUF2419 family. Guided by biochemical insight, we used several comparative genomics tools to propose that proteins in this family are functionally involved in Q salvage. Subsequently, we analyzed a series of *S. pombe* knockout strains and demonstrated that a DUF2419 knockout strain does not incorporate Q into tRNA, while complementation of the knockout with DUF2419 homologues from a variety of species restores Q incorporation. Combining these experimental results with a state-of-the-art

homology modeling analysis led us to hypothesize that DUF2419 proteins have a Q-related ribonucleoside hydrolase activity. Biochemical and structural studies are currently underway to test this hypothesis and provide direct evidence whether Q, Q-5′P, or Q-3′-P are substrates.

DUF2419 is also an example of a protein family that is not found in *A. thaliana* or *S. cerevisiae* but is found in other related species such as *Z. maize* and in *S. pombe*, reinforcing the need for using different models even from the same phylum. The results reported in this paper open a new area of investigation related to the pathway for Q salvage in plants and raise questions concerning the ultimate source of Q for these organisms. While it was demonstrated recently that Q synthesis is required for effective nitrogen-fixing symbiosis of *Sinorhizobium meliloti* with its host plant *Medicago truncatula*,[64] Q salvage could have additional roles in plant physiology. Furthermore, our biochemical *trans*-complementation experiment on human gene C9orf64, which encodes a protein belonging to DUF2419,[65−67] provides a functional annotation for this human protein.

■ METHODS

Comparative Genomics. The known and published sequences of the characterized bacterial TGT (*E. coli* TGT, AAA24667[68]), and human QTRT1, and QTRTD1 (IPI00215974.2 and IPI00783033.2, respectively[17]) were used as entry points for all database queries. The BLAST tools[69] and resources at NCBI (National Center for Biotechnology Information, http://www.ncbi.nlm.nih.gov) were routinely used. Multiple sequence alignments were built using Multalin (http://multalin.toulouse.inra.fr/multalin/)[70] and clustal omega (https://www.ebi.ac.uk/Tools/msa/clustalo/).[71] Protein domain organization analysis was performed using the Pfam database tools (http://pfam.sanger.ac.uk)[72] and CDD (Conserved Domain Database, http://www.ncbi.nlm.nih.gov/cdd/).[50] Phyletic profile searches were performed on the OrthoMCL-DB platform (http://orthomcl.org/orthomcl/home.do).[45] Interactions between genes/proteins were investigated using the STRING-DB (Search Tool for the Retrieval of Interacting Genes/Proteins, http://string-db.org).[54] Mapping of gene distribution profile to taxonomic trees were generated using the iTOL suite (Interactive Tree Of Life, http://itol.embl.de).[73]

Phyletic Analyses. The survey of the *nr* database at the NCBI by BLASTP (Basic local alignment search tool for protein,[69] default parameter excepted the evalue max which was set at 10^{-10}) using the sequences of *Schizosaccharomyces pombe* 972h- (Q9HDZ9, Fungi), *Homo sapiens* (Q5T6V5, Metazoa), *Glycine max* (NP_001242636, Viridiplantae), *Galdieria sulphuraria* (XP_005706831, Red algae), *Naegleria gruberi* (XP_002680366, Heterolobosea) and *Conexibacter woesei* DSM 14684 (YP_003392877, Actinobacteria) as seed allowed identifying 385 DUF2419 (25 bacterial and 260 eukaryotic homologues, respectively). The use of other seeds did not allow retrieving more sequences. The retrieved sequences were aligned using MAFFT v.7 (Multiple sequence alignment based on fast Fourier transform, using the default parameters).[74] The resulting alignment was visually inspected using ED, the alignment editor of the MUST package (Management Utilities for Sequences and Trees).[75] At this step, 11 eukaryotic sequences were discarded because they corresponded to isoforms resulting from alternative splicing or very partial sequences. The 374 remaining sequences were aligned again with MAFFT with the (*-linsi* option). The resulting alignment was trimmed using BMGE (Block Mapping and Gathering with Entropy, using the matrix of substitution: BLOSUM30).[76] The maximum likelihood (ML) phylogeny of these 374 sequences was inferred with PhyML v.3.1[77] with the Le and Gascuel (LG) model,[78] a gamma distribution with four categories (Γ_4) and an estimated alpha parameter to take into account the heterogeneity of evolutionary rates among sites, and the NNI+SPR strategy of topology exploration. The branch robustness of the reconstructed tree was estimated with

the SH-like approach implemented in PhyML. A second phylogenetic analysis restricted to 75 bacterial and eukaryote homologues representative of the genetic and genomic diversity of these two Domains was performed using the Bayesian approach implemented in MrBayes v.3.2.[79] MrBayes was run with a mixed substitution model and a Γ_4 distribution. Four chains were run in parallel for 1,000,000 generations. The first 2000 generations were discarded as burn-in. The remaining trees were sampled every 100 generations to build consensus trees and to compute the posterior probabilities. The ML phylogeny was computed with PhyML with the same parameters than used previously. The branch robustness of the ML tree was tested with the nonparametric bootstrap procedure implemented in PhyML (100 replicates of the original data set).

Structure Analysis. Phyre2[59] (Protein Homology/analogy Recognition Engine V 2.0, http://www.sbg.bio.ic.ac.uk/phyre2/) was used to search for for structural similarity. Visualization and comparison of protein structures and manual docking of ligand were performed using XtalView.[80] CNS (Crystallography and NMR system)[81] was used for energetic minimization. PyMol[82] was used to generate the molecular graphics figure.

Strains, Media, and Growth Conditions. All *S. pombe* haploid strains including wild-type (SP286; h+/, ade6-M216 ura4-D18 leu1-32, reference BG_0000H8), qtrt1Δ:kanMX4 corresponding to the deletion of the gene SPAC1687.19c (Bioneer reference BG_H0432) and duf2419Δ:kanMX4 corresponding to the deletion of the gene SPAC589.05c (Bioneer reference BG_H2044), respectively abbreviated qtrt1Δ and duf2419Δ, were obtained from Bioneer. The cells were grown on yeast extract peptone dextrose agar supplemented with G418 disulfate salt (Sigma, 200 µg/mL) at 30 °C, and the genotypes were checked by PCR (Supplementary Figure S9). For propagation, strains were grown in liquid 1% yeast extract, 2% peptone, and 2% dextrose at 30 °C. *E. coli* GC10 (Genesee Scientific Corporation) was grown in Luria−Bertani medium (LB, Thermo Fisher Scientific Inc.) at 37 °C, solidified when needed with 15 g/L of agar (Thermo Fisher Scientific Inc.). Ampicillin (Amp, 100 µg/mL, Sigma) was added when required.

Cloning and Plasmid Construction for Complementation. The DNA encoding *S. pombe* DUF2419 (abbreviated *SpDuf2419*, SPAC589.05c) was synthesized (without optimization) with added restrictions sites *Xho*I and *Bam*HI at its 5′ and 3′ ends (GenScript USA Inc.). The *Zea mays* DUF2419 (abbreviated *ZmDuf2419* GRMZM5G824534) coding sequence was amplified from the plasmid ZM_BFb0154G03 obtained from the Arizona Genomics Institute with addition of the restrictions sites *Xho*I and *Bam*HI at its 5′ and 3′ ends. Prof. Gaetano T. Montelione (Rutgers University, Piscataway, NJ, USA) provided plasmids containing the synthesized *H. sapiens* DUF2419 (abbreviated *HsDuf2419*, UPF0553 protein C9orf64) coding sequence and *S. thermophilus* DSM 20745 DUF2419 (abbreviated StDUF2419, hypothetical protein Sthe_2331) recoded sequence (GenScript USA Inc.) that were amplified by PCR with the addition of restrictions sites *Xho*I and *Bam*HI at their 5′ and 3′ ends. These four DNA fragments were cloned into the *Xho*I and *Bam*HI of *S. pombe* expression vector PJR1-3XU, provided by Dr. Juan C. Ribas (Instituto de Biología Funcional y Genómica, Salamanca, Spain) under the control of the *nmt1* 3X promoter and followed by the *nmt1* terminator.[83] The four fragments containing the promoter, coding sequences, and terminator were subcloned into pJK148 provided by Prof. Jill B. Keeney (Juniata College, Huntingdon, PA, USA), for genomic integration into the leu1-32 locus of the duf2419Δ:kanMX4 ade6-M216 ura4-D18 leu1-32 strain, after linearization with the *Nru*I restriction enzyme.[84] A corresponding negative control was created by integrating into the leu1-32 the empty vector pJK148 (abbreviated empty). The corresponding strains are duf2419Δ:kanMX4 leu1-32:*SpDuf2419*, duf2419Δ:kanMX4 leu1-32:*ZmDuf2419*, duf2419Δ:kanMX4 leu1-32:*HsDuf2419*, duf2419Δ:kanMX4 leu1-32:StDuf2419, and duf2419Δ:kanMX4 leu1-32:empty vector, abbreviated duf2419Δ + SpDUF2419, duf2419Δ + ZmDUF2419, duf2419Δ + HsDUF2419, duf2419Δ + StDUF2419 and duf2419Δ + Ø respectively. Transformations were realized using the Frozen-EZ yeast transformation II kit (Zymo Research), following manufacturer recommendations.

Transformants were selected on minimal SD Agar Base media supplemented with DO Supplement -Leu (Clontech Laboratories) and 5% Dextrose (Fluka). Clones were validated by PCR with primers specific for pJK148 (Supplementary Figure S10). The plasmids constructed, strains and primers used are listed in Supplementary Table S1.

tRNA Extraction and Purification and Modified Nucleoside Content Analysis. Bulk tRNA was prepared from fresh cells from 750 mL of cultures of the different *S. pombe* strains grown in YPD (1% yeast extract, 2% peptone, 2% dextrose, BD Difco YPD) or 0.5% yeast extract (Difco), 2% bactopeptone (Difco), and 3% dextrose (Fluka) at 30 °C, as described previously.[85] tRNA$^{Asp}_{GUC}$ was extracted from bulk tRNA using a biotinylated primer (5′biotin-GCAAGCGTGACAGGC-TTG-3′, Integrated DNA Technologies) bound to a HiTrap Streptavidin HP column (1 mL, GE Healthcare Life Sciences).[85] Twenty-five micrograms of tRNA$^{Asp}_{GUC}$ was digested to nucleosides that were then separated by LC−MS.[85] To compare tRNA concentrations, we compared the ratio of the levels of the modified bases Q (410 m/z) and m^{5}C (258 m/z) in each sample by integrating the peak area from the extracted ion chromatograms. All tRNA extractions and analysis for Q content were performed at least twice independently.

LC−MS/MS for tRNA Sequence Analysis. Purified *S. pombe* tRNA$^{Asp}_{GUC}$ was digested with RNase T1 (Worthington Biochemical) at the rate of 50 U/μg in 20 mM ammonium acetate pH 6.5 (Sigma-Aldrich) for 2 h at 37 °C. The sample was subsequently applied to a Waters Xbridge C18 3.5 μm 1.0 mm × 50 mm column, and digestion products were separated at 30 μL/min using 200 mM 1,1,1,3,3,3-hexafluoroisopropanol (HFIP, Sigma-Aldrich) and 8.15 mM triethyl-amine (TEA, Sigma-Aldrich) for mobile phase A and 50% A and 50% methanol (Burdick and Jackson) for mobile phase B. The column was equilibrated with 5% A, followed by a gradient of 5% B 0 min, 5% B 5 min, 95% B 45 min, and held at 95% B for 5 min. The eluent was directed into a Thermo Scientific LTQ-XL with a capillary temperature of 275 °C and a spray voltage of 4.0 kV. The sheath gas was set to 25 arbitrary units, auxiliary gas to 14, and sweep gas to 10. Tandem mass spectrometry at a collision energy of 42 (arbitrary units) was used to obtain product ion (i.e., sequence) information on the digestion products in data-dependent mode.

■ ASSOCIATED CONTENT

ⓢ Supporting Information

This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
*E-mail: vcrecy@ufl.edu.

Notes
The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Katze, J. R., Basile, B., and McCloskey, J. A. (1982) Queuine, a modified base incorporated posttranscriptionally into eukaryotic transfer RNA: wide distribution in nature. *Science* 216, 55−56.

(2) Nishimura, S., Shindo-Okada, N., Kasai, H., Kuchino, Y., Noguchi, S., Iigo, M., and Hoshi, A. (1983) Characterization and analysis of oncofetal tRNA and its possible application for cancer diagnosis and therapy. *Recent Results Cancer Res.* 84, 401−412.

(3) El Yacoubi, B., Bailly, M., and de Crécy-Lagard, V. (2012) Biosynthesis and function of posttranscriptional modifications of transfer RNAs. *Annu. Rev. Genet.* 46, 69−95.

(4) Okada, N., Noguchi, S., Nishimura, S., Ohgi, T., Goto, T., Crain, P. F., and McCloskey, J. A. (1978) Structure determination of a nucleoside Q precursor isolated from *E. coli* tRNA: 7-(aminomethyl)-7-deazaguanosine. *Nucleic Acids Res.* 5, 2289−2296.

(5) Noguchi, S., Nishimura, Y., Hirota, Y., and Nishimura, S. (1982) Isolation and characterization of an *Escherichia coli* mutant lacking tRNA-guanine transglycosylase. Function and biosynthesis of queuosine in tRNA. *J. Biol. Chem.* 257, 6544−6550.

(6) Curnow, A. W., and Garcia, G. A. (1994) tRNA-guanine transglycosylase from *Escherichia coli*: recognition of dimeric, unmodified tRNA(Tyr). *Biochimie* 76, 1183−1191.

(7) Slany, R. K., Bösl, M., and Kersten, H. (1994) Transfer and isomerization of the ribose moiety of AdoMet during the biosynthesis of queuosine tRNAs, a new unique reaction catalyzed by the QueA protein from *Escherichia coli*. *Biochimie* 76, 389−393.

(8) Miles, Z. D., McCarty, R. M., Molnar, G., and Bandarian, V. (2011) Discovery of epoxyqueuosine (oQ) reductase reveals parallels between halorespiration and tRNA modification. *Proc. Natl. Acad. Sci. U.S.A.* 108, 7368−7372.

(9) Crain, P. F., Sethi, S. K., Katze, J. R., and McCloskey, J. A. (1980) Structure of an amniotic fluid component, 7-(4,5-cis-dihydroxy-1-cyclopenten-3-ylaminomethyl)-7-deazaguanine (queuine), a substrate for tRNA: guanine transglycosylase. *J. Biol. Chem.* 255, 8405−8407.

(10) Farkas, W. R., and Bruce Jacobson, K. (1980) tRNA-guanine transglycosylase and guanine-accepting transfer RNAs of *Drosophila melanogaster*. *Insect Biochem.* 10, 183−188.

(11) Ott, G., Kersten, H., and Nishimura, S. (1982) *Dictyostelium discoideum*: a useful model system to evaluate the function of queuine and of the Q-family of tRNAs. *FEBS Lett.* 146, 311−314.

(12) Walden, T. L., Howes, N., and Farkas, W. R. (1982) Purification and properties of guanine, queuine-tRNA transglycosylase from wheat germ. *J. Biol. Chem.* 257, 13218−13222.

(13) Kirtland, G. M., Morris, T. D., Moore, P. H., O'Brian, J. J., Edmonds, C. G., McCloskey, J. A., and Katze, J. R. (1988) Novel salvage of queuine from queuosine and absence of queuine synthesis in *Chlorella pyrenoidosa* and *Chlamydomonas reinhardtii*. *J. Bacteriol.* 170, 5633−5641.

(14) Gaur, R., Björk, G. R., Tuck, S., and Varshney, U. (2007) Diet-dependent depletion of queuosine in tRNAs in *Caenorhabditis elegans* does not lead to a developmental block. *J. Biosci.* 32, 747−754.

(15) Stengl, B., Meyer, E. A., Heine, A., Brenk, R., Diederich, F., and Klebe, G. (2007) Crystal structures of tRNA-guanine transglycosylase (TGT) in complex with novel and potent inhibitors unravel pronounced induced-fit adaptations and suggest dimer formation upon substrate binding. *J. Mol. Biol.* 370, 492−511.

(16) Stengl, B., Reuter, K., and 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. *ChemBioChem* 6, 1926−1939.

(17) Chen, Y.-C., Kelly, V. P., Stachura, S. V., and Garcia, G. A. (2010) Characterization of the human tRNA-guanine transglycosylase:

confirmation of the heterodimeric subunit structure. *RNA 16*, 958–968.

(18) Boland, C., Hayes, P., Santa-Maria, I., Nishimura, S., and Kelly, V. P. (2009) Queuosine formation in eukaryotic tRNA occurs via a mitochondria-localized heteromeric transglycosylase. *J. Biol. Chem. 284*, 18218–18227.

(19) Morris, R. C., and Elliott, M. S. (2001) Queuosine modification of tRNA: a case for convergent evolution. *Mol. Genet. Metab. 74*, 147–159.

(20) Gündüz, U., and Katze, J. R. (1984) Queuine salvage in mammalian cells. Evidence that queuine is generated from queuosine 5′-phosphate. *J. Biol. Chem. 259*, 1110–1113.

(21) Katze, J. R., Gündüz, U., Smith, D. L., Cheng, C. S., and McCloskey, J. A. (1984) Evidence that the nucleic acid base queuine is incorporated intact into tRNA by animal cells. *Biochemistry 23*, 1171–1176.

(22) Gündüz, U., and Katze, J. R. (1982) Salvage of the nucleic acid base queuine from queuine-containing TRNA by animal cells. *Biochem. Biophys. Res. Commun. 109*, 159–167.

(23) Beier, H., Barciszewska, M., and Sickinger, H. D. (1984) The molecular basis for the differential translation of TMV RNA in tobacco protoplasts and wheat germ extracts. *EMBO J. 3*, 1091–1096.

(24) Nishimura, S. (1983) Structure, biosynthesis, and function of queuosine in transfer RNA. *Prog. Nucleic Acid Res. Mol. Biol. 28*, 49–73.

(25) Brandmayr, C., Wagner, M., Brückl, T., Globisch, D., Pearson, D., Kneuttinger, A. C., Reiter, V., Hienzsch, A., Koch, S., Thoma, I., Thumbs, P., Michalakis, S., Müller, M., Biel, M., and Carell, T. (2012) Isotope-based analysis of modified tRNA nucleosides correlates modification density with translational efficiency. *Angew. Chem., Int. Ed. 51*, 11162–11165.

(26) Walden, T., Reyniers, J. P., Hiatt, V., and Farkas, W. R. (1982) Yeast cells cannot incorporate queuine into their tRNA. *Proc. Soc. Exp. Biol. Med. 170*, 328–332.

(27) Globisch, D., Pearson, D., Hienzsch, A., Brückl, T., Wagner, M., Thoma, I., Thumbs, P., Reiter, V., Kneuttinger, A. C., Müller, M., Sieber, S. A., and Carell, T. (2011) Systems-based analysis of modified tRNA bases. *Angew. Chem., Int. Ed. 50*, 9739–9742.

(28) Siard, T. J., Jacobson, K. B., and Farkas, W. R. (1991) Queuine metabolism and cadmium toxicity in *Drosophila melanogaster*. *Biofactors 3*, 41–47.

(29) Schachner, E., Aschhoff, H. J., and Kersten, H. (1984) Specific changes in lactate levels, lactate dehydrogenase patterns and cytochrome b559 in *Dictyostelium discoideum* caused by queuine. *Eur. J. Biochem. 139*, 481–487.

(30) Schachner, E., and Kersten, H. (1984) Queuine deficiency and restoration *in Dictyostelium discoideum* and related early developmental changes. *Microbiology 130*, 135–144.

(31) Marks, T., and Farkas, W. R. (1997) Effects of a diet deficient in tyrosine and queuine on germfree mice. *Biochem. Biophys. Res. Commun. 230*, 233–237.

(32) Kersten, H. (1988) The nutrient factor queuine: biosynthesis, occurrence in transfer RNA and function. *Biofactors 1*, 27–29.

(33) Curtius, H. C., Heintel, D., Ghisla, S., Kuster, T., Leimbacher, W., and Niederwieser, A. (1985) Biosynthesis of tetrahydrobiopterin in man. *J. Inherit. Metab. Dis. 8* (Suppl 1), 28–33.

(34) Rakovich, T., Boland, C., Bernstein, I., Chikwana, V. M., Iwata-Reuyl, D., and Kelly, V. P. (2011) Queuosine deficiency in eukaryotes compromises tyrosine production through increased tetrahydrobiopterin oxidation. *J. Biol. Chem. 286*, 19354–19363.

(35) Vinayak, M., and Pathak, C. (2010) Queuosine modification of tRNA: its divergent role in cellular machinery. *Biosci. Rep. 30*, 135–148.

(36) Ott, G., and Kersten, H. (1985) Differential turnover of tRNAs of the queuosine family in *Dictyostelium discoideum* and its possible role in regulation. *Biol. Chem. Hoppe-Seyler 366*, 69–76.

(37) Grosjean, H., de Crécy-Lagard, V., and Marck, C. (2010) Deciphering synonymous codons in the three domains of life: co-evolution with specific tRNA modification enzymes. *FEBS Lett. 584*, 252–264.

(38) Hurt, J. K., Olgen, S., and Garcia, G. A. (2007) Site-specific modification of *Shigella flexneri* virF mRNA by tRNA-guanine transglycosylase in vitro. *Nucleic Acids Res. 35*, 4905–4913.

(39) Reader, J. S., Metzgar, D., Schimmel, P., and de Crécy-Lagard, V. (2004) Identification of four genes necessary for biosynthesis of the modified nucleoside queuosine. *J. Biol. Chem. 279*, 6280–6285.

(40) Van Lanen, S. G., Reader, J. S., Swairjo, M. A., de Crécy-Lagard, V., Lee, B., and Iwata-Reuyl, D. (2005) From cyclohydrolase to oxidoreductase: discovery of nitrile reductase activity in a common fold. *Proc. Natl. Acad. Sci. U.S.A. 102*, 4264–4269.

(41) Phillips, G., El Yacoubi, B., Lyons, B., Alvarez, S., Iwata-Reuyl, D., and de Crécy-Lagard, V. (2008) Biosynthesis of 7-deazaguanosine-modified tRNA nucleosides: a new role for GTP cyclohydrolase I. *J. Bacteriol. 190*, 7876–7884.

(42) De Crécy-Lagard V., and Olson G. (2009) RNA modification subsystems in the SEED database, in *DNA and RNA editing enzymes: comparative structure, mechanism, functions, cellular interactions and evolution*, Landes Bioscience: Austin, Texas, United States, pp 624–628, Appendix 3.

(43) De Crécy-Lagard, V. (2007) Identification of genes encoding tRNA modification enzymes by comparative genomics. *Methods Enzymol. 425*, 153–183.

(44) Chen, P., Jäger, G., and Zheng, B. (2010) Transfer RNA modifications and genes for modifying enzymes in *Arabidopsis thaliana*. *BMC Plant Biol. 10*, 201.

(45) Chen, F., Mackey, A. J., Stoeckert, C. J., and Roos, D. S. (2006) OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. *Nucleic Acids Res. 34*, D363–D368.

(46) Roth, A., Winkler, W. C., Regulski, E. E., Lee, B. W. K., Lim, J., Jona, I., Barrick, J. E., Ritwik, A., Kim, J. N., Welz, R., Iwata-Reuyl, D., and Breaker, R. R. (2007) A riboswitch selective for the queuosine precursor preQ1 contains an unusually small aptamer domain. *Nat. Struct. Mol. Biol. 14*, 308–317.

(47) Meyer, M. M., Roth, A., Chervin, S. M., Garcia, G. A., and Breaker, R. R. (2008) Confirmation of a second natural preQ1 aptamer class in Streptococcaceae bacteria. *RNA 14*, 685–695.

(48) Rodionov, D. A., Hebbeln, P., Eudes, A., ter Beek, J., Rodionova, I. A., Erkens, G. B., Slotboom, D. J., Gelfand, M. S., Osterman, A. L., Hanson, A. D., and Eitinger, T. (2009) A novel class of modular transporters for vitamins in prokaryotes. *J. Bacteriol. 191*, 42–51.

(49) Watanabe, M., Matsuo, M., Tanaka, S., Akimoto, H., Asahi, S., Nishimura, S., Katze, J. R., Hashizume, T., Crain, P. F., McCloskey, J. A., and Okada, N. (1997) Biosynthesis of archaeosine, a novel derivative of 7-deazaguanosine specific to archaeal tRNA, proceeds via a pathway involving base replacement on the tRNA polynucleotide chain. *J. Biol. Chem. 272*, 20146–20151.

(50) Marchler-Bauer, A., Lu, S., Anderson, J. B., Chitsaz, F., Derbyshire, M. K., DeWeese-Scott, C., Fong, J. H., Geer, L. Y., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Jackson, J. D., Ke, Z., Lanczycki, C. J., Lu, F., Marchler, G. H., Mullokandov, M., Omelchenko, M. V., Robertson, C. L., Song, J. S., Thanki, N., Yamashita, R. A., Zhang, D., Zhang, N., Zheng, C., and Bryant, S. H. (2011) CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res. 39*, D225–D229.

(51) Pérez-Arellano, I., Gallego, J., and Cervera, J. (2007) The PUA domain - a structural and functional overview. *FEBS J. 274*, 4972–4984.

(52) Aravind, L., and Koonin, E. V. (1999) Novel predicted RNA-binding domains associated with the translation machinery. *J. Mol. Evol. 48*, 291–302.

(53) Suhre, K. (2007) Inference of gene function based on gene fusion events: the rosetta-stone method. *Methods Mol. Biol. 396*, 31–41.

(54) Franceschini, A., Szklarczyk, D., Frankild, S., Kuhn, M., Simonovic, M., Roth, A., Lin, J., Minguez, P., Bork, P., von Mering, C., and Jensen, L. J. (2013) STRING v9.1: protein-protein interaction

1823

dx.doi.org/10.1021/cb500278k | *ACS Chem. Biol.* 2014, 9, 1812–1825

networks, with increased coverage and integration. *Nucleic Acids Res.* 41, D808−D815.

(55) Kim, D.-U., Hayles, J., Kim, D., Wood, V., Park, H.-O., Won, M., Yoo, H.-S., Duhig, T., Nam, M., Palmer, G., Han, S., Jeffery, L., Baek, S.-T., Lee, H., Shim, Y. S., Lee, M., Kim, L., Heo, K.-S., Noh, E. J., Lee, A.-R., Jang, Y.-J., Chung, K.-S., Choi, S.-J., Park, J.-Y., Park, Y., Kim, H. M., Park, S.-K., Park, H.-J., Kang, E.-J., Kim, H. B., Kang, H.-S., Park, H.-M., Kim, K., Song, K., Song, K. B., Nurse, P., and Hoe, K.-L. (2010) Analysis of a genome-wide set of gene deletions in the fission yeast *Schizosaccharomyces pombe*. *Nat. Biotechnol.* 28, 617−623.

(56) Phillipson, D. W., Edmonds, C. G., Crain, P. F., Smith, D. L., Davis, D. R., and McCloskey, J. A. (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.

(57) Tusnády, G. E., and Simon, I. (2001) The HMMTOP transmembrane topology prediction server. *Bioinformatics* 17, 849−850.

(58) Buchan, D. W. A., Minneci, F., Nugent, T. C. O., Bryson, K., and Jones, D. T. (2013) Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Res.* 41, W349−W357.

(59) Kelley, L. A., and Sternberg, M. J. E. (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nat. Protoc.* 4, 363−371.

(60) Lingaraju, G. M., Sartori, A. A., Kostrewa, D., Prota, A. E., Jiricny, J., and Winkler, F. K. (2005) A DNA glycosylase from *Pyrobaculum aerophilum* with an 8-oxoguanine binding mode and a noncanonical helix-hairpin-helix structure. *Structure* 13, 87−98.

(61) Brooks, S. C., Adhikary, S., Rubinson, E. H., and Eichman, B. F. (2013) Recent advances in the structural mechanisms of DNA glycosylases. *Biochim. Biophys. Acta* 1834, 247−271.

(62) Andreeva, A., Howorth, D., Chandonia, J.-M., Brenner, S. E., Hubbard, T. J. P., Chothia, C., and Murzin, A. G. (2008) Data growth and its impact on the SCOP database: new developments. *Nucleic Acids Res.* 36, D419−D425.

(63) Hanson, A. D., Pribat, A., Waller, J. C., and de Crécy-Lagard, V. (2010) Unknown" proteins and "orphan" enzymes: the missing half of the engineering parts list—and how to find it. *Biochem. J.* 425, 1−11.

(64) Marchetti, M., Capela, D., Poincloux, R., Benmeradi, N., Auriac, M.-C., Le Ru, A., Maridonneau-Parini, I., Batut, J., and Masson-Boivin, C. (2013) Queuosine biosynthesis is required for *Sinorhizobium meliloti*-induced cytoskeletal modifications on HeLa cells and symbiosis with *Medicago truncatula*. *PLoS One* 8, e56043.

(65) Sweetser, D. A., Peniket, A. J., Haaland, C., Blomberg, A. A., Zhang, Y., Zaidi, S. T., Dayyani, F., Zhao, Z., Heerema, N. A., Boultwood, J., Dewald, G. W., Paietta, E., Slovak, M. L., Willman, C. L., Wainscoat, J. S., Bernstein, I. D., and Daly, S. B. (2005) Delineation of the minimal commonly deleted segment and identification of candidate tumor-suppressor genes in del(9q) acute myeloid leukemia. *Genes Chromosomes Cancer* 44, 279−291.

(66) Cai, L., Abe, M., Izumi, S., Imura, M., Yasugi, T., and Ushijima, T. (2007) Identification of PRTFDC1 silencing and aberrant promoter methylation of GPR150, ITGA8 and HOXD11 in ovarian cancers. *Life Sci.* 80, 1458−1465.

(67) Ohi, Y., Qin, H., Hong, C., Blouin, L., Polo, J. M., Guo, T., Qi, Z., Downey, S. L., Manos, P. D., Rossi, D. J., Yu, J., Hebrok, M., Hochedlinger, K., Costello, J. F., Song, J. S., and Ramalho-Santos, M. (2011) Incomplete DNA methylation underlies a transcriptional memory of somatic cells in human iPS cells. *Nat. Cell Biol.* 13, 541−549.

(68) Chen, Y.-C., Brooks, A. F., Goodenough-Lashua, D. M., Kittendorf, J. D., Showalter, H. D., and Garcia, G. a. (2011) Evolution of eukaryal tRNA-guanine transglycosylase: insight gained from the heterocyclic substrate recognition by the wild-type and mutant human and *Escherichia coli* tRNA-guanine transglycosylases. *Nucleic Acids Res.* 39, 2834−2844.

(69) Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389−3402.

(70) Corpet, F. (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* 16, 10881−10890.

(71) Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., and Higgins, D. G. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7, 539.

(72) Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J. E., Gavin, O. L., Gunasekaran, P., Ceric, G., Forslund, K., Holm, L., Sonnhammer, E. L. L., Eddy, S. R., and Bateman, A. (2010) The Pfam protein families database. *Nucleic Acids Res.* 38, D211−D222.

(73) Letunic, I., and Bork, P. (2011) Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. *Nucleic Acids Res.* 39, W475−W478.

(74) Katoh, K., and Standley, D. M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772−780.

(75) Philippe, H. (1993) MUST, a computer package of Management Utilities for Sequences and Trees. *Nucleic Acids Res.* 21, 5264−5272.

(76) Criscuolo, A., and Gribaldo, S. (2010) BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol. Biol.* 10, 210.

(77) Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307−321.

(78) Le, S. Q., and Gascuel, O. (2008) An improved general amino acid replacement matrix. *Mol. Biol. Evol.* 25, 1307−1320.

(79) Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M. A., and Huelsenbeck, J. P. (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61, 539−542.

(80) McRee, D. E. (1999) XtalView/Xfit—A versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.* 125, 156−165.

(81) Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 54, 905−921.

(82) *The PyMOL Molecular Graphics System*, Version 1.3r1 (2010) Schrodinger LLC, Cambridge.

(83) Moreno, M. B., Durán, A., and Ribas, J. C. (2000) A family of multifunctional thiamine-repressible expression vectors for fission yeast. *Yeast* 16, 861−872.

(84) Keeney, J. B., and Boeke, J. D. (1994) Efficient targeted integration at leu1−32 and ura4−294 in *Schizosaccharomyces pombe*. *Genetics* 136, 849−856.

(85) Phillips, G., Swairjo, M. A., Gaston, K. W., Bailly, M., Limbach, P. A., Iwata-Reuyl, D., and de Crécy-Lagard, V. (2012) Diversity of archaeosine synthesis in crenarchaeota. *ACS Chem. Biol.* 7, 300−305.

(86) Jühling, F., Mörl, M., Hartmann, R. K., Sprinzl, M., Stadler, P. F., and Pütz, J. (2009) tRNAdb 2009: compilation of tRNA sequences and tRNA genes. *Nucleic Acids Res.* 37, D159−D162.

(87) Sprinzl, M., Moll, J., Meissner, F., and Hartmann, T. (1985) Compilation of tRNA sequences. *Nucleic Acids Res.* 13 (Suppl), r1−49.

(88) Sprinzl, M., Hartmann, T., Meissner, F., Moll, J., Vorderwulbecke, T., and Vorderwülbecke, T. (1987) Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res.* 15 (Suppl), r53−188.

(89) Altwegg, M., and Kubli, E. (1980) The nucleotide sequence of histidine tRNA gamma of *Drosophila melanogaster*. *Nucleic Acids Res.* 8, 3259−3262.

(90) White, B. N., and Tener, G. M. (1973) Activity of a transfer RNA modifying enzyme during the development of *Drosophila* and its relationship to the su(s) locus. *J. Mol. Biol. 74*, 635−651.

(91) White, B. N., and Lassam, N. J. (1979) An analysis of Q and Q* containing tRNAs during the development of *Lucilia sericata, Musca domestica* and *Tenebrio molitor. Insect Biochem. 9*, 375−378.

(92) Barciszewska, M. Z., and Jones, D. S. (1988) The nucleotide sequences of two *Lupinus luteus* asparagine tRNAs. *Nucleic Acids Res. 16*, 9349.

(93) Barciszewska, M. Z., Keith, G., Dirheimer, G., and Barciszewski, J. (1986) The primary structure of the first plant tRNAHis. *Plant Sci. 47*, 103−107.

(94) Barciszewski, J., Barciszewska, M., Suter, B., and Kubli, E. (1985) Plant tRNA suppressors: In vivo readthrough properties and nucleotide sequence of yellow lupin seeds tRNATyr. *Plant Sci. 40*, 193−196.

(95) Mörl, M., Dörner, M., and Pääbo, S. (1995) C to U editing and modifications during the maturation of the mitochondrial tRNA(Asp) in marsupials. *Nucleic Acids Res. 23*, 3380−3384.

(96) Blaise, M., Becker, H. D., Lapointe, J., Cambillau, C., Giegé, R., and Kern, D. (2005) Glu-Q-tRNA(Asp) synthetase coded by the *yadB* gene, a new paralog of aminoacyl-tRNA synthetase that glutamylates tRNA(Asp) anticodon. *Biochimie 87*, 847−861.