On the Track of the Missing tRNA Genes: A Source of Non-Canonical Functions?

Ricardo Ehrlich[1,2], Marcos Davyt[1], Ignacio López[1], Cora Chalar[1] and Mónica Marín[1]*

[1]Biochemistry-Molecular Biology, Faculty of Science, Universidad de la República, Montevideo, Uruguay, [2]Institut Pasteur de Montevideo, Montevideo, Uruguay

Cellular tRNAs appear today as a diverse population of informative macromolecules with conserved general elements ensuring essential common functions and different and distinctive features securing specific interactions and activities. Their differential expression and the variety of post-transcriptional modifications they are subject to, lead to the existence of complex repertoires of tRNA populations adjusted to defined cellular states. Despite the tRNA-coding genes redundancy in prokaryote and eukaryote genomes, it is surprising to note the absence of genes coding specific translational-active isoacceptors throughout the phylogeny. Through the analysis of different releases of tRNA databases, this review aims to provide a general summary about those "missing tRNA genes." This absence refers to both tRNAs that are not encoded in the genome, as well as others that show critical sequence variations that would prevent their activity as canonical translation adaptor molecules. Notably, while a group of genes are universally missing, others are absent in particular kingdoms. Functional information available allows to hypothesize that the exclusion of isodecoding molecules would be linked to: 1) reduce ambiguities of signals that define the specificity of the interactions in which the tRNAs are involved; 2) ensure the adaptation of the translational apparatus to the cellular state; 3) divert particular tRNA variants from ribosomal protein synthesis to other cellular functions. This leads to consider the "missing tRNA genes" as a source of putative non-canonical tRNA functions and to broaden the concept of adapter molecules in ribosomal-dependent protein synthesis.

Keywords: missing tRNAs, tRNA functions, tRNA modifications, tRNA interactions, non-canonical tRNA functions, tRNA structure

INTRODUCTION

"Though this be madness, yet there is method in it."

(Hamlet Act 2, scene 2)

Over the last few years, tRNA has become the subject of intense research where technological and conceptual advances converge from integrative biology to pathology and biotechnology. An extraordinary volume of work has led to characterize the population of tRNAs in organisms that cover the entire phylogeny, both by analysis of tRNA molecules and genomic data (Chan and Lowe, 2009; Jühling et al., 2009; Lowe and Chan, 2016). Yet, it is interesting to notice that tRNAs, which have been the first sequenced nucleic acid molecule (Holley et al., 1965) and the first solved

three-dimensional structure (Kim et al., 1974), still pose complex challenges centered on their structure and functions, despite the remarkable advances in sequencing techniques and structural studies of nucleic acids (Dittmar et al., 2006; Pang et al., 2014; Ferro and Ignatova, 2015; Zhang et al., 2015; Evans et al., 2017; Shigematsu et al., 2017; Kimura et al., 2020; Pinkard et al., 2020; Zhang et al., 2020). Old questions that focus on the structure-function relationship and fidelity of the interactions and cellular processes in which these molecules participate are thus renewed. Simultaneously, the enormous amount of information that has flourished in the last decade updates evolutionary questions, whose projections acquire new implications in the understanding of numerous biological processes and are related to a vast field of applications (Torres et al., 2014a; Kirchner and Ignatova, 2015; Orioli, 2017; Ho et al., 2018; Tharp et al., 2020).

As central adapter molecules in protein biosynthesis, the diversity of tRNA isoacceptors or isotypes (tRNA genes or tRNA molecules with different anticodon that charge the same amino acid), isodecoders (molecules with the same anticodon but different body structure), and modified states thereof, ensures the transmission of information from a nucleotide to an amino acid sequence. tRNAs also participate in the adjustment of the translation machinery and its kinetics, as well as in co-translational folding of peptides in order to meet the cellular requirements in terms of repertoire and relative amount of proteins (Gingold et al., 2014; Kirchner and Ignatova, 2015; Kirchner et al., 2017; Marín et al., 2017; Rak et al., 2018). Moreover, they play an important role in a number of adaptive processes associated with changes in cellular programs, including major metabolic options, non-ribosomal protein synthesis, and a large number of regulatory mechanisms, that are part of the so-called non-canonical functions.
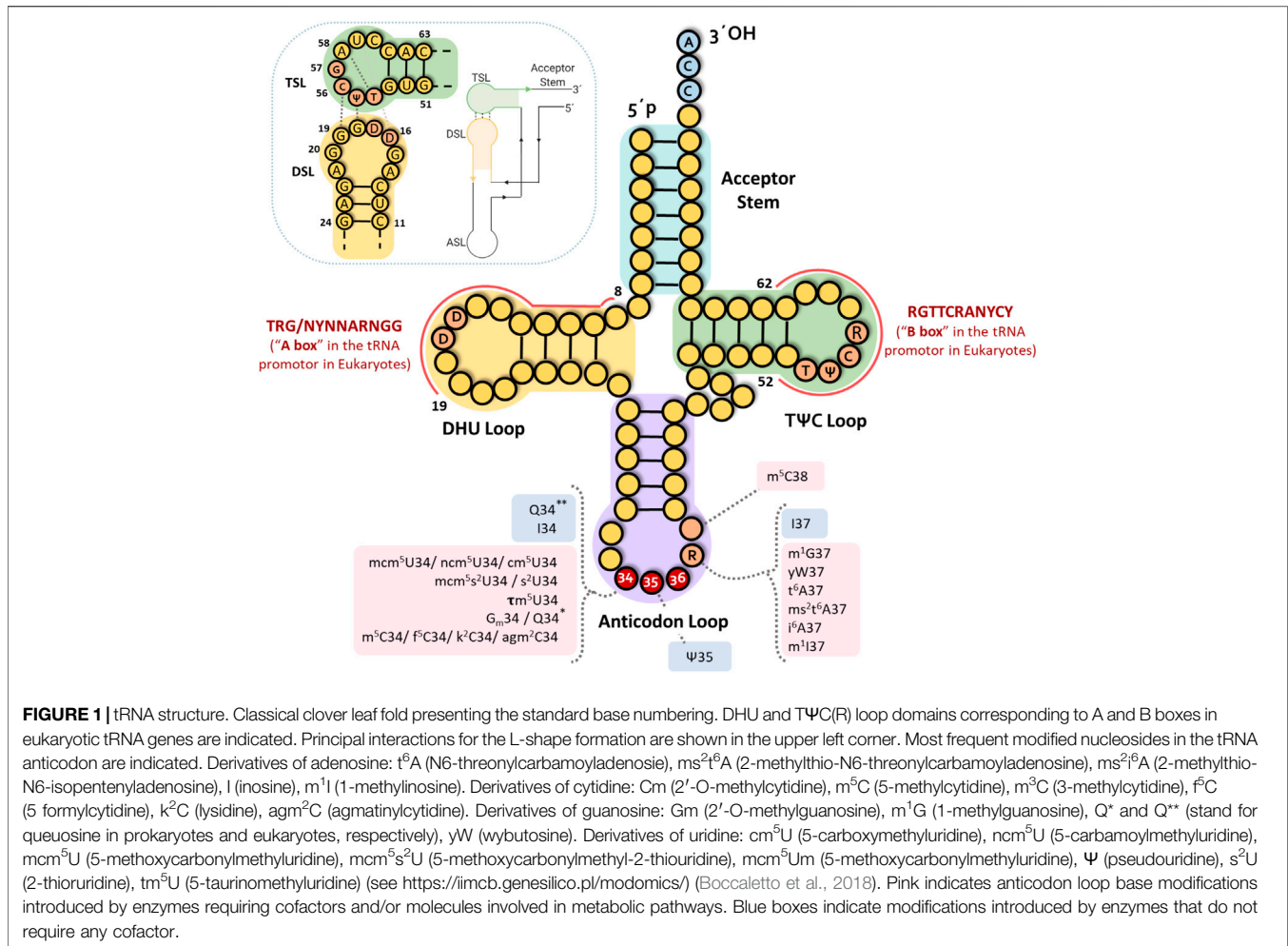
Cellular tRNAs appear today as a diverse population of informative macromolecules with conserved general elements ensuring essential common functions and different and distinctive features securing specific interactions and activities. Their differential expression and the variety of post-transcriptional modifications they are subject to, lead to the existence of complex repertoires of tRNA populations adjusted to defined cellular states (Dittmar et al., 2006; Pavon-Eternod et al., 2009; Gingold et al., 2014; Pang et al., 2014; Goodarzi et al., 2016; Sagi et al., 2016; Kimura et al., 2020; Pinkard et al., 2020). Moreover, an important redundancy has been described in tRNA-coding genes (tDNAs) both in prokaryotes (in some cases up to a hundred) and in eukaryotes (several hundreds) (Goodenbour and Pan, 2006; Fujishima and Kanai, 2014; Chan and Lowe, 2016; Pan, 2018; Rak et al., 2018). As algorithms that identify tRNA genes were improved, the number of tDNAs annotated using high stringency criteria decreased (GtRNAdb http://gtrnadb.ucsc.edu/ data release 18:1, August 2019). This allowed to detect the complete or almost complete absence of genes coding for specific isoacceptors, in either all kingdoms or specifically in bacteria, archaea or eukaryotes. In that sense, expressions drawing the attention to this observation, such as "pseudo-tRNAs," "silent genes," "absent genes," "deleterious,"

"toxic," and "prohibited tRNA species," appeared in the scientific literature (Maraia and Arimbasseri, 2017; Rak et al., 2018; Torres, 2019; Pernod et al., 2020). On the other hand, evidence has also accumulated on the existence of particular species of isoacceptor tRNAs for specific amino acids, defined according to their structural characteristics and anticodon, with different functionalities (Giannouli et al., 2009; Rudinger-Thirion et al., 2011; Rogers et al., 2012; Torres, 2019).

This review is intended to shed light on a particular point: the astonishing absence of genes encoding functional isoacceptor tRNAs carrying particular anticodons throughout the phylogeny. This absence not only refers to tRNAs that are not encoded in the genome but also to others that have critical sequence variations that impair their activity as canonical translation adaptor molecules. Indeed, the analysis of genomic (Lowe and Chan, 2016; Chan and Lowe, 2019) and directly sequenced tRNA databases (Jühling et al., 2009), reveal that prokaryotes and eukaryotes lack an important number of putative tRNAs bearing particular anticodons. This observation has been pointed out by several authors, and has been associated with different aspects including genetic code origin, evolution, and disambiguation mechanisms ensuring the fidelity of the different translation steps (Maraia and Arimbasseri, 2017; Diwan and Agashe, 2018; Rak et al., 2018; Lei and Burton, 2020). Although a systematic and integrated overall perspective about the absence of tDNAs is progressively emerging, much work is still required. This review aims to summarize a general view about the "missing tRNA genes" and to emphasize open questions about tRNA structure-function relationship and their multiple cellular functions.

This work is based on the analysis of different releases of tRNA databases. It should be mentioned that while the absence of a tDNA constitutes a fair evidence, the identification of a putative tRNA coding sequence is not a definite proof of the presence of a tRNA able to be aminoacylated and functionally active in translation. Thus, manual curing was carried out when required, following the criteria summarized by Marck and Grosjean (2002) and Giegé et al. (2012) to consider a sequence to be able or not to fold into a canonical "L-shape" structure. It should also be kept in mind that -with exceptions- there is scarce information available across the phylogeny about the transcription of different genes and the functionality of the encoded molecules. This is because albeit significant recent improvements (Zhang et al., 2015; Kimura et al., 2020), there are still technical restrictions to identify transcription products of specific putative tDNAs due to the presence of numerous post-transcriptional modifications that interfere with sequencing-by-synthesis techniques and hybridization methodologies.

To deeper understand the significance of the absent genes, we will first present a summary of missing tRNA genes for bacteria, archaea, and eukaryotes, as "anticodon charts." Then, we will discuss the selected evolutionary options in light of the recent developments in the field, highlighting sets of excluded genes that appear linked to disambiguation needs, not only related to the recognition of the corresponding codons. These considerations bring forward the idea of further expanding the concept of adaptive molecules, in line with the growing field of non-

FIGURE 1 | tRNA structure. Classical clover leaf fold presenting the standard base numbering. DHU and TΨC(R) loop domains corresponding to A and B boxes in eukaryotic tRNA genes are indicated. Principal interactions for the L-shape formation are shown in the upper left corner. Most frequent modified nucleosides in the tRNA anticodon are indicated. Derivatives of adenosine: t⁶A (N6-threonylcarbamoyladenosie), ms²t⁶A (2-methylthio-N6-threonylcarbamoyladenosine), ms²i⁶A (2-methylthio-N6-isopentenyladenosine), I (inosine), m¹I (1-methylinosine). Derivatives of cytidine: Cm (2′-O-methylcytidine), m⁵C (5-methylcytidine), m³C (3-methylcytidine), f⁵C (5 formylcytidine), k²C (lysidine), agm²C (agmatinylcytidine). Derivatives of guanosine: Gm (2′-O-methylguanosine), m¹G (1-methylguanosine), Q* and Q** (stand for queuosine in prokaryotes and eukaryotes, respectively), yW (wybutosine). Derivatives of uridine: cm⁵U (5-carboxymethyluridine), ncm⁵U (5-carbamoylmethyluridine), mcm⁵U (5-methoxycarbonylmethyluridine), mcm⁵s²U (5-methoxycarbonylmethyl-2-thiouridine), mcm⁵Um (5-methoxycarbonylmethyluridine), Ψ (pseudouridine), s²U (2-thiouridine), tm⁵U (5-taurinomethyluridine) (see https://iimcb.genesilico.pl/modomics/) (Boccaletto et al., 2018). Pink indicates anticodon loop base modifications introduced by enzymes requiring cofactors and/or molecules involved in metabolic pathways. Blue boxes indicate modifications introduced by enzymes that do not require any cofactor.

canonical tRNA functions. In what follows, we adopt the standard numbering for tRNA bases, defining positions 34–36 as corresponding to the anticodon (see **Figure 1**).
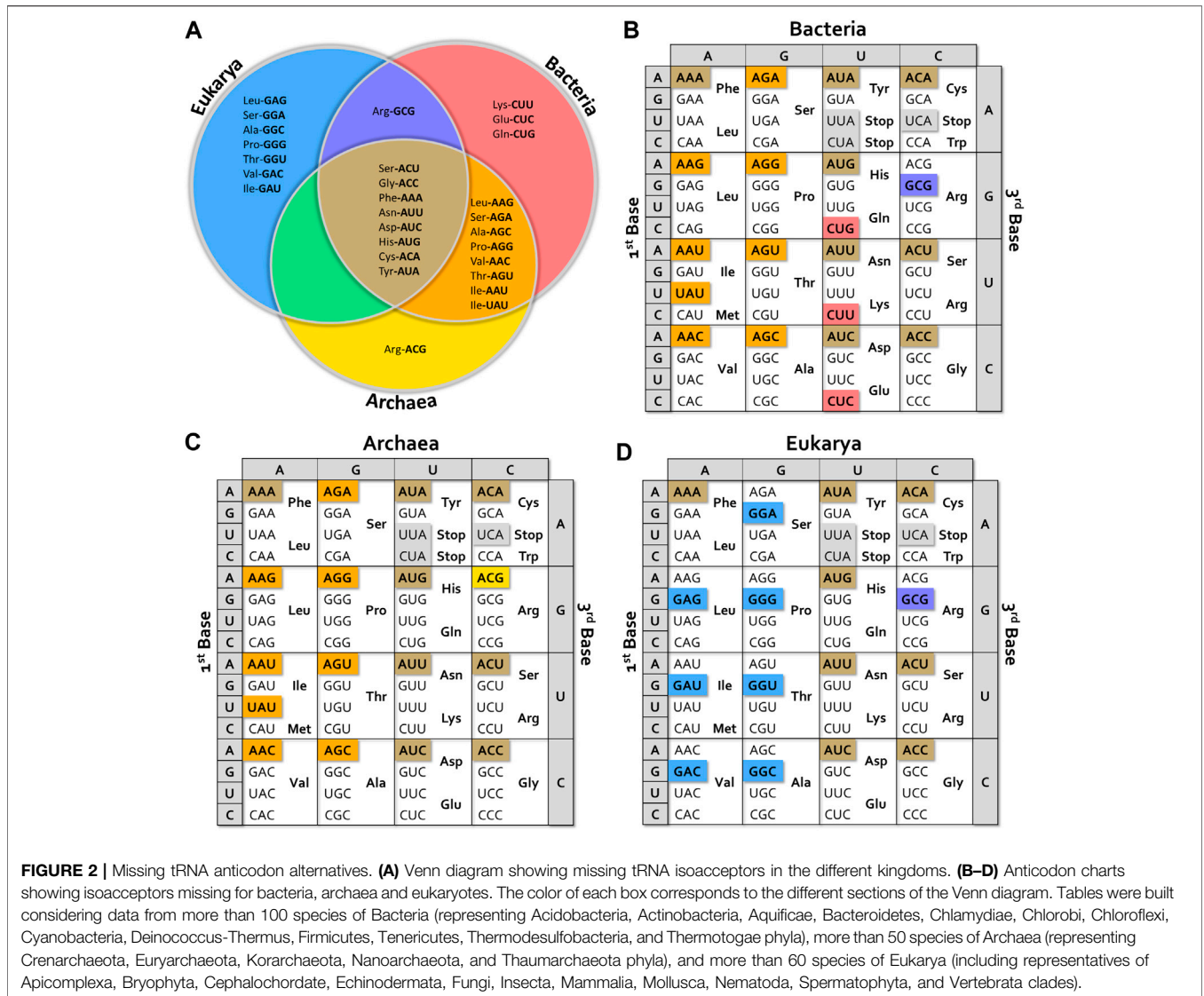
ABSENT ISOACCEPTOR tRNA GENES: A GENERAL DESCRIPTION

Absent isoacceptor tRNAs in bacteria, archaea and eukaryotes (cytoplasmic, mitochondrial and chloroplastic) are summarized as "anticodon charts" in **Figure 2** and **Supplementary Figure S1**, that show that some isoacceptor tRNA genes are universally missing, while others are absent in particular groups. Some interesting exceptions are mentioned below.

As *general facts*, we observe that at least one isoacceptor tRNA per box is lost in all organisms. Eight tRNA genes are universally missing; tRNA^{Gly}ACC and all possible ANN anticodon bearing-tRNAs in split boxes (including those with stop codons), except for tRNA^{Ile} (see **Figure 2A**). No UNN anticodon has disappeared with the exception of UAU in the Ile-Met box in prokaryotes.

In *Bacteria*, the vast majority of species lack tRNA isoacceptors bearing ANN anticodons with the exception of

ArgACG (instead, the ArgGCG is missing), UAU anticodon in the Ile-Met box, and CNN anticodons in His-Gln, Asn-Lys and Asp-Glu split boxes (**Figure 2B**). Most bacterial genomes encode a tRNA^{Ile}CAU. It should be noted that the presence of genes coding for ArgACG tRNA is correlated with the existence of an adenosine-deaminase gene (*tadA*) specific for A in the first position of the anticodon (A34). Thus, the corresponding anticodon is actually ICG. It is interesting to mention some exceptions to this general description. In addition to tRNA^{Arg}ACG, *Leuconostocaceae* (*Leuconostoc* and *Enococcus* sp.) carries genes for tRNA^{Leu}AAG and tRNA^{Thr}AGT instead of the GNN-carrying tRNAs. Some *Firmicutes* species have a gene for a tRNA^{Leu}AAG and all sequenced *Thermotogae* and *Spirochaetes* contain a tRNA^{Arg}GCG instead of tRNA^{Arg}ACG. Interestingly, *Mycoplasma* spp. and other *Mollicutes* additionally lack the gene coding for tRNA^{Arg}CCG (Yokobori et al., 2013). Data available for *bacteriophage*-encoded tRNAs revealed that they lack the same genes absent in *bacteria* (Morgado and Vicente, 2019). Despite the very high genetic variability of phages (Hatfull, 2015; Pope et al., 2015), these tRNA genes are also considered as "prohibited," strongly suggesting that they would be associated with viral spread drawback.

FIGURE 2 | Missing tRNA anticodon alternatives. **(A)** Venn diagram showing missing tRNA isoacceptors in the different kingdoms. **(B–D)** Anticodon charts showing isoacceptors missing for bacteria, archaea and eukaryotes. The color of each box corresponds to the different sections of the Venn diagram. Tables were built considering data from more than 100 species of Bacteria (representing Acidobacteria, Actinobacteria, Aquificae, Bacteroidetes, Chlamydiae, Chlorobi, Chloroflexi, Cyanobacteria, Deinococcus-Thermus, Firmicutes, Tenericutes, Thermodesulfobacteria, and Thermotogae phyla), more than 50 species of Archaea (representing Crenarchaeota, Euryarchaeota, Korarchaeota, Nanoarchaeota, and Thaumarchaeota phyla), and more than 60 species of Eukarya (including representatives of Apicomplexa, Bryophyta, Cephalochordate, Echinodermata, Fungi, Insecta, Mammalia, Mollusca, Nematoda, Spermatophyta, and Vertebrata clades).

In *Archaea* (**Figure 2C**), all ANN anticodon options are missing. As in bacteria, the IleUAU is also lost.

In *Eukarya* (**Figure 2D**), one anticodon per box is lost, regardless of stop codons. These include the eight universally missing ANN anticodons and eight GNN anticodons that correspond to the eight boxes where A34 can be deaminated to inosine, namely the seven degenerated boxes except Gly, plus the Ile-Met box. Among the few exceptions to this general scheme, it is worth mentioning that *Saccharomyces* spp. and *Candida glabrata* have a gene coding for tRNA^Leu^GAG, while no genes coding for tRNA^Leu^AAG were identified. Interestingly, although the tRNA^Leu^GAG is transcribed in *S. cerevisiae*, its deletion did not impair yeast growth. On the other hand, deletion of tRNA^Leu^UAG was lethal (Huang et al., 2012). This raises questions about the cellular role of tRNA^Leu^GAG, for which no conclusive evidence about its aminoacylation is available. Additionally, this also highlights the issue of decoding

LeuCUC codon, for which a superwobble mechanism has been proposed (Huang et al., 2012).

The ***mitochondrial*** tRNA population can be either entirely or partially encoded by the mitochondrial genome or be fully encoded by the nuclear genome (Schneider, 2011; Salinas-Giegé et al., 2015). In the last two cases, a highly specific mitochondrial import process is required to complete the tRNA set (Duchêne et al., 2009; Schneider, 2011; Salinas-Giegé et al., 2015). In any of those contexts, the mitochondrial tRNA population (http://trnadb.bioinf.uni-leipzig.de/ and http://plantrna.ibmp.cnrs.fr/) shows similar traits: 1) only one anticodon is conserved in fully degenerated boxes, being UNN the most commonly used and few exceptions are observed (ArgACG in mitochondria of Cestoda and in some plants, and ThrGGU and GlyGCC in some plants); 2) in split boxes, the sole options are GNN and UNN; 3) Ile and Trp tRNA coding genes present different alternatives: in plants, tRNA Ile anticodons are GAU or CAU, while in animals only option GAU is present;

concerning Trp anticodon, the options are CCA in plants, UCA in animals and in *Saccharomyces* (see **Supplementary Figure S1A**).

It is interesting to recall that plant mitochondrial tRNA populations have been described to be mainly built through nuclear import and gene transfer processes (Maréchal-Drouard et al., 1990; Dietrich et al., 1996; Knie et al., 2015), in particular from other plants, algae, chloroplasts, fungi and bacteria (reviewed in Warren and Sloan, 2020). This complex origin and great diversity underscore the significance of the constraints that have led to the selection of an almost universal set of mitochondrial isodecoder tRNA anticodons with only few variants.

Sequence information about *chloroplast* tRNA is still scarce, hence their frequency cannot be precisely estimated (less than forty entries in databases: PlantRNA database (http://plantrna.ibmp.cnrs.fr/, http://trnadb.bioinf.uni-leipzig.de/) (Michaud et al., 2011; Cognat et al., 2013). However, it is possible to extract a clear general profile for angiosperms, with only small variants in bryophytes, algae and diatom plastids (see **Supplementary Figure S1B**). In a general way, the selected options present common traits with those of mitochondria. G/UNN anticodons have been selected in almost all boxes, except for Leu and Ala where GNN is also absent, and for Arg in the degenerated box where only ACG is present. It should be mentioned that an eukaryotic adenosine deaminase (ADAT) specific for A34 in tRNA^Arg^ACG has been identified in chloroplasts (Karcher and Bock, 2009).

POSSIBLE CLUES FOR THE ABSENCE OF tRNA GENES

As central adaptor molecules in translation, tRNA must ensure the specificity of the interactions with different macromolecules and macromolecular complexes. The shape, conformational properties and optimal steric positioning of chemical groups at the base edges, constitute key elements to ensure the fidelity of the mechanisms in which tRNAs are involved. Very particularly, they have to be recognized by the corresponding aminoacyl-tRNA synthetase and loaded with the correct amino acid. In addition, they must be localized into the ribosome by specific factors and recognize the corresponding mRNA triplets at the ribosome decoding site. Proper accommodation is crucial since tRNAs also contribute to the concerted conformational changes that relocate the aminoacylated tRNA in the A site, dissociation of the corresponding elongation factor, and triggering of the peptidyl-transferase activity. Moreover, they play a role in the concerted movement that leads to the release from A site. Given such vital functions, the absence of particular tRNAs appears as a striking observation. What could be the causes that explain the systematic exclusion of certain tRNA genes? Some possible clues are addressed in the following sections.

The present analysis was centered on the tRNA populations defined only by their anticodon, though other structural elements and several supplementary levels of adjustment are required to ensure a faithful and efficient translation (for a general review see

Agris et al., 2018). Although the anticodon triplet clearly defines the tRNA identity referred to the cognate amino acid and codon recognition, specific interactions in the decoding site further involve the overall anticodon-stem and loop (ASL). In particular, neighboring bases stabilize the codon-anticodon interaction, prevent slippage and frame-shifting, and increase decoding fidelity. Within critical ASL elements, much work has been focused on the role of the 32–38 base-pairing, and the conserved and frequently modified or hypermodified purine at position 37, among others (Konevega et al., 2004; Olejniczak and Uhlenbeck, 2006; Ledoux et al., 2009; Pernod et al., 2020). Interestingly, structural studies showed that initial binding of tRNA to the A site is followed by a rate-limiting rearrangement of the anticodon loop in the ribosome decoding center that is favored by the purine 37, yielding additional interactions with the rRNA (Konevega et al., 2004).

About Database Curation, Exceptions and Exclusions

The systematic absence of particular tRNA genes, related to alternatives in the first anticodon base (base 34, the "wobble position"), suggests that those anticodons would be considered as "deleterious" and then might be cleared off by negative selection. This is an interesting phenomenon to underline, considering the high number of copies and variants of tRNA gene isotypes, especially in eukaryotes. Indeed, a considerable number of putative tRNA coding sequences carrying "prohibited" anticodons could be found in eukaryotic genomes in earlier genomic tRNA databases. Most of them have been eliminated in the last release of the GtRNAdb (http://gtrnadb.ucsc.edu/ data release 18:1, August 2019) based on high stringency algorithms. Even though some tRNAs are still signaled as functional tRNAs, they contain substitutions in other regions (mainly at the level of the TΨCR and DHU loops), that would affect the stability of the "L-shape" and/or the helical fold and stability of different arms (in particular mispairings and/or higher number of G-U base pairs). Except for a few recently reported cases discussed below, there is no information yet about the transcription of most of these sequences. Moreover, in eukaryotes, modifications of the genomic sequences encoding the DHU and TΨCR arms and loops also affect the A and B sites of the tRNA gene promoters, altering their transcription (see **Figure 1**). However, if these sequences are indeed transcribed, they would fold into a non-canonical structure inconsistent with the ribosomal-dependent translation function.

The Principle of the Excluded Purine: A Matter of Three Purines

Purine selection in the first anticodon position appears as one of the major elements in structuring the extant genetic code. Isoacceptors containing A or G are mutually excluded throughout the phylogeny. This was already highlighted by different authors as one of the major anticodon-sparing strategies (Marck and Grosjean, 2002; Maraia and Arimbasseri, 2017). In addition, the non-canonical inosine

nucleoside (product of adenosine deamination) appears as a third option and its role has been recently underlined (Torres et al., 2014b; Rafels-Ybern et al., 2018). It is found in all domains of life at three possible positions on tRNAs: 34 (the wobble position), except for archaea; 37 (following the anticodon); and 57 (at the TΨCR loop). As early proposed by Crick in the wobble hypothesis (Crick, 1966), inosine at the wobble position has an obvious implication on codon-anticodon recognition, since it is capable of pairing with A, U and C, while A34 only pairs with U. Furthermore, it was suggested that I34 introduction might play a major role in driving codon usage-biased translation to shape proteome landscape (Rafels-Ybern et al., 2019).

Adenosine in the first position of the anticodon appears throughout the phylogeny linked to the presence of A34 specific adenosine deaminases. Indeed, no A34-specific ADAT has been identified in archaea. In most bacteria, the only tRNA bearing adenosine at this position is ArgACG, which is in fact transformed to inosine by a homodimeric TadA. The same modification of ArgACG is introduced in chloroplasts by an ADAT imported from the cytoplasm (Delannoy et al., 2009). However, no A34 has been identified yet for mitochondrial tRNAs. In the vast majority of eukaryotes, a heterodimeric ADAT (composed of ADAT2/Tad2 and ADAT3/Tad3) modifies cytosolic tRNA species with ANN anticodons (tRNA^{Arg}ACG, tRNA^{Ala}AGC, tRNA^{Ile}AAU, tRNA^{Leu}AAG, tRNA^{Pro}AGG, tRNA^{Ser}AGA, tRNA^{Thr}AGU and tRNA^{Val}AAC) (Torres et al., 2014b; Rafels-Ybern et al., 2018).

One More Step on the Purine Conflict: The Q Link Between Bacteria and Eukaryotes

An additional complexity level in the purine conflict is introduced by the replacement of guanosine in Tyr, His, Asn and Asp tRNAs (all found in split boxes) with the hypermodified queuine base (Q) at the position 34 of tRNA anticodons. Queuine is a 7-deaza-guanosine derivative, synthesized by eubacteria and salvaged by eukaryotes for its incorporation into tRNA (Fergus et al., 2015; Müller et al., 2019). In higher eukaryotes, queuine is further modified in tRNA^{Tyr} and tRNA^{Asp} by the addition of a galactose and a mannose sugar, respectively (Fergus et al., 2015). The base-exchange reaction resulting in Q incorporation is carried out by the tRNA guanine transglycosylase (TGT). Albeit TGTs are present in all kingdoms, conserve structural traits, and share the catalytic mechanism, archaea TGT do not modify the first anticodon position and instead incorporate a queuine analogue, archaeosine, at position 15 in the D-loop of many tRNAs (Phillips and de Crécy-Lagard, 2011; Turner et al., 2020).

The Purine Conflict: An Information or a Structural Problem?

Interestingly, in all extant organisms, A and G are mutually excluded at the first position of the anticodon within tRNAs carrying the same amino acid (except tRNA^{Ser} in eukaryotes). This means that tRNAs isoacceptors bearing the other purine at this position are somehow excluded. In prokaryotes, guanosine is the only purine used at the wobble position, except for tRNA^{Arg} in bacteria. In eukaryotes instead, the choice of the purine at this position is diverse. Eukaryotes use eight tRNAs with inosine

(seven in the cases of *Saccharomyces* spp. and *C. glabrata* mentioned in previous sections), four have queuine and four harbor guanosine. The four eukaryotic tRNAs maintaining guanosine at position 34 are: PheGAA, GlyGCC, CysGCA and SerGCU (in the split box shared with Arg). In organisms in which the tRNA sequences were determined, the corresponding G34 appears to be modified in the following cases (http://trnadb.bioinf.uni-leipzig.de/): one of the two tRNA^{Gly}GCC reported for *S. cerevisiae*, and all tRNA^{Phe}GAA (a 2MeO modification of G34 is almost universally present, except an unknown G modification reported for *Mus musculus*, and the absence of modifications in *Mycoplasma capricolum*, *Rhodospirillum rubrum* and archaea). So far, no modifications of G34 have been described for tRNA^{Cys}GCA and tRNA^{Ser}GCU. All these data suggest that the choice of the purine at position 34 is critical, as confirmed by a recent report studying the use and putative toxicity of G34 in eukaryotes (Pernod et al., 2020).

What could be the physical basis for the purine conflict that leads to the exclusion of one of the three options? Undoubtedly, interaction alternatives allowed by inosine in synonymous codons recognition is a major element. However, this may not be the only cause if bacteria and archaea tRNAs are considered, as this modification is only present in tRNA^{Arg}ACG. Most tRNA modifications are directly linked to the metabolic state through the requirement of key cofactor molecules, from which S-adenosylmethionine as chemical group donor and thiol source is the main paradigm (Helm and Alfonzo, 2014; Danchin et al., 2020). However, adenosine deamination, Q/G transglycosylation (in eukaryotes) and pseudouridylation (that also involves a transglycosylation mechanism) (Veerareddygari et al., 2016) are not directly related to the metabolic state of the cell. Hence, the purine conflict should be associated with other general common constraints.

An attractive alternative would be to consider the shape of tRNA accepted by the sites in the ribosome, and the specific interactions that take place there and that are mostly conserved in all kingdoms of life (Fox, 2010; Melnikov et al., 2012; Petrov et al., 2014). Another critical factor would be the role of different tautomeric and protonation properties and states of the three purines in defining ambiguity levels admitted in the recognition process in the ribosome. Nucleic acid bases can adopt multiple tautomeric forms due to the presence of multiple solvent-exchangeable protons (reviewed in Shukla and Leszczynski, 2013; Singh et al., 2015) and this plays a key role in the specificity of the interactions in which they are involved (Westhof, 2014; Agris et al., 2018).

Crystallographic studies showed that the ribosomal grip around the triplet codon/anticodon fits sterically better with the dimensions and volume of a standard RNA helix that is recognized through the shallow minor groove (Murphy and Ramakrishnan, 2004; Demeshkina et al., 2012). This helix may contain Watson–Crick-like pairs involving particularly-protonated bases, tautomeric and modification states, as well as *anti-syn* conformation options, resulting in different H-bonding alternatives at their edges (Westhof et al., 2014). There is now extensive structural information about the codon/anticodon interaction in the ribosome showing the

critical role of modified bases (Westhof, 2014; Westhof et al., 2014). So far, only some evidence has been presented regarding variants of the tautomeric state of the bases in the wobble position. In addition, for I34-bearing anticodons, it was reported that while I-C base-pair stericity was similar to the canonical G-C base pair, the purine-purine I-A pairing has an increased width that requires a change in the geometry of the anticodon to fit the conserved interactions at the decoding center (Murphy and Ramakrishnan, 2004).

Additionally, it was recently communicated that the I34-A3 base-pair is tightly dependent on the anticodon loop conformation and the modification state of neighboring bases (Vangaveti et al., 2020). Alltogether, these observations highlight the role of base tautomerism in base pairing schemes related to the steric and geometric constraints imposed by the ribosome as originally discussed by Topal & Fresco (Topal and Fresco, 1976).

Too Many Partners for a Common Language: Split Box Informative Problems
The Ile-Met Box Dilemma
The genetic code establishes that three codons correspond to Ile and one to Met. Their box can be considered either as originally completely degenerated with all four triplets corresponding to Ile and invaded by Met, or as split box invaded by Ile leaving only one option for initiator or elongator Met (Lei and Burton, 2020). Notably, functional tRNAs in this box are different in eukaryotes and prokaryotes.

In the former, the anticodon GAU is missing, AAU yields IAU, and UAU anticodon is modified to ΨAU, ensuring the recognition of AUC, AUU and AUA codons. In prokaryotes, however, AAU and UAU anticodons are lost, and CAU anticodon is shared by tRNAs for Ile and Met. Nonetheless, tRNA^{Ile}CAU and tRNA^{Met}CAU are perfectly discriminated by the corresponding synthetases, and specifically recognize Ile and Met codons, respectively. The mechanism ensuring discrimination is different in bacteria and archaea. In both cases, a post-transcriptional modification of cytosine takes place in the anticodon of the tRNA^{Ile}CAU (Suzuki and Numata, 2014). In almost all bacteria, C34 is modified with a lysidine (k^{2}C) residue catalyzed by a tRNA^{Ile} lysidine synthetase (TilS) (Soma et al., 2003). In archaea, C34 in IleCAU anticodon is modified with agmatine, yielding 2-agmatinylcytidine (agm^{2}C) (Mandal et al., 2010). Agmatine is an intermediate metabolite in the polyamines pathway, generated from arginine via decarboxylation by the enzyme arginine decarboxylase (TiaS).

Conjugation of lysine or agmatine to the C2 carbon of cytosine by deoxidization induces a tautomeric change of cytosine from enamine to imine, with eventual protonation of N3. Thus, these modifications completely alter the proton donor-acceptor pattern of cytosine, preventing base-pairing with G and enabling that with A (Voorhees et al., 2013), avoiding MetAUG codon misreading. Interestingly, TilS and TiaS enzymes belong to non-related protein families that modify the wobble cytidine by distinct catalytic mechanisms: TilS activates the C2 carbon by adenylation and TiaS activates it by phosphorylation (Numata, 2015). However, their activity has the same physical and

physiological effect. Hence, this constitutes an interesting case of convergent evolution since the decoding system for AUA codons would have appeared after the separation of bacteria and archaea from their common ancestor (Suzuki and Numata, 2014). In addition, it is a notable case of storing different information in the anticodon of the tRNA, which allows, on one hand, disambiguation to read the mRNA and, on the other, proper recognition and acylation by the cognate aminoacyl-tRNA synthetase.

Charging Amide Amino Acids in Prokaryotes
Isodecoder tRNAs selected within NUN split boxes differ between bacteria and archaea. As in general cases, ANN anticodons have not been retained. In addition, CUN anticodon is also missing in bacteria, yielding a single decoder tRNA species for Gln, Asn, Lys, Asp and Glu amino acids. Interestingly, most bacterial and all known archaeal genomes do not encode glutaminyl-tRNA synthetase (GlnRS) and a large number of prokaryotes do not have an asparaginyl-tRNA synthetase (AsnRS) (Tumbula et al., 2000; Sheppard and Söll, 2008; Nureki et al., 2010) even though they have a tRNA^{Gln}UUG and a tRNA^{Asn}GUU. In those cases, charging the corresponding amino acid on the cognate tRNA occurs by an indirect pathway: the tRNA is first charged with Glu or Asp by glutamyl-tRNA synthetase (GluRS) or aspartyl-tRNA synthetase (AspRS) respectively, and then amidated by specific amidotransferases (Sheppard et al., 2008). There are two tRNA-dependent amidotransferases (AdT) described: the heterotrimeric GatCAB in both archaea and bacteria, and the heterodimeric GatDE only in archaea missing AsnRS. GatCAB is required to catalyze the conversion of Glu-tRNA^{Gln} and/or Asp-tRNA^{Asn} into Gln-tRNA^{Gln} and/or Asn-tRNA^{Asn}. In a similar way, GatDE is involved in the Gln-tRNA^{Gln} formation (for a review see Sheppard and Söll, 2008; Sheppard et al., 2008). Therefore, AdTs must discriminate their mischarged tRNA substrates from the cognate aa-tRNA species. Interestingly, the co-crystal structure of GatDE with tRNA^{Gln} in *M. thermautotrophicus* showed that AdTs would achieve this discrimination without directly recognizing the anticodon of their tRNA substrates (Oshikane et al., 2006), suggesting a possible involvement of different constraints dispersed in the tRNA structure.

The two-step process to introduce amide amino acids raises complex evolutionary questions while providing interesting clues (Sheppard and Söll, 2008; Di Giulio, 2020). On the one hand, they have been associated with the general metabolic roles of these amino acids in various prokaryotes (Sheppard and Söll, 2008). On the other, they must ensure translation fidelity in cases of anticodon pairs UUC/UUG (Glu/Gln) and GUC/GUU (Asp/Asn) where a supplementary discrimination step is required, complementing the codon-anticodon recognition.

The post-aminoacylation amidation mechanism opens up the possibility that unmodified Asp-tRNA^{Asn} and Glu-tRNA^{Gln} can reach the decoding site and introduce an error. However, it has been demonstrated that the elongation factor EF-Tu discriminates misacylated from correctly amidated tRNA. This not only sheds light on an additional mechanism for ensuring translation fidelity, but also reveals that the elongation factor

constitutes a further partner to read the selected coding system (Roy et al., 2007; Eargle et al., 2008).

Cys-Trp-Stop, Tyr-Stop, Phe-Leu and Ser-Arg Split Boxes

In the Cys-Trp box, the anticodon option CysACA is universally absent, and there is a single option for Cys and for Trp. The fourth triplet, corresponding to the opal termination codon (UGA) can be decoded by the rare tRNA^{Sec}UCA carrying selenocysteine, the infrequent 21st amino acid. Selenocysteine does not have a dedicated synthetase in any of the three kingdoms. Sec-tRNA^{Sec} is synthesized by the conversion of serine through a multistep process in a Sec-specific tRNA-dependent manner. tRNA^{Sec} is first aminoacylated with Ser by seryl-tRNA synthetase (SerRS) to produce Ser-tRNA^{Sec}. The following step is species-dependent: in bacteria, Sec synthetase (SelA) converts Ser-tRNA^{Sec} to Sec-tRNA^{Sec} in a single-step reaction; in contrast, archaea and eukaryotes carry on the synthesis through an intermediate step where the Ser moiety is first phosphorylated (Yuan et al., 2010; Wang et al., 2015; Holman et al., 2017).

In a number of methanogenic archaea, a cysteinyl-tRNA synthetase (CysRS) is either absent or dispensable. In those cases there is a tRNA-dependent indirect pathway in two steps. The tRNA^{Cys} is initially aminoacylated with O-phosphoserine (Sep) by O-phosphoseryl-tRNA synthetase (SepRS). The Sep moiety is subsequently transformed to a tRNA-bound cysteine by Sep-tRNA:Cys-tRNA synthetase (SepCysS) (Sauerwald et al., 2005). Finally, concerning this box, it should be reminded that in animals and *Saccharomyces* mitochondria, the Trp anticodon is UCA and UAG, respectively, and that they universally correspond to opal and amber stop codons.

In the Tyr box, the AUA option is universally absent, GUA is modified to QUA in bacteria and eukarya, and there are two stop codons: the amber UAG and the ochre UAA. The amber codon is read by a rare tRNA^{Pyr}CUA, up to now only described in *Methanosarcinaceae* archaea and in bacterial phyla *Clostridia* and *δ-proteobacteria*. This tRNA is charged with pyrrolysine - the uncommon 22nd amino acid—by a dedicated pyrrolysine-tRNA synthetase (Yuan et al., 2010; Wan et al., 2014).

In the Phe-Leu and Ser-Arg split boxes, the option A34 is universally missing, leaving a 1-2 distribution for the two amino acids.

As briefly summarized above, the information available on the tRNAs corresponding to split boxes, strongly suggests that many constraints and options have accompanied the evolution of the translational apparatus ensuring its fidelity, and are condensed in their structures as a kind of "molecular algorithm."

The Mystery of tRNA Gly Anticodons

The intriguing absence of an A in the wobble position of Gly degenerated box in all organisms has already been a particular focus of attention (Novoa et al., 2012; Maraia and Arimbasseri, 2017). It is interesting to note that the human isoacceptor tRNA^{Gly}GCC, with its anticodon experimentally changed to either ACC or ICC, was shown to be recognized and loaded with Gly by glycyl-tRNA synthetase (GlyRS) (Saint-Léger et al.,

2016). This strongly suggests that in this case, the purine at position 34 would not play a critical role in the discrimination by the cognate synthetase. Furthermore, they showed that the anticodon loop of the different chimeric tRNA^{Gly} did not interfere with binding but prevented deamination by hADAT. Complementary molecular dynamics studies showed that the introduction of an A34 destabilized the structure of tRNA^{Gly} anticodon. This led to propose the challenging hypothesis that a kind of "signal saturation" would have been reached limiting the evolution of extant genomic populations of tRNAs (Saint-Léger et al., 2016).

It is interesting to link these observations with the idea that tRNA^{Gly} would be the first tRNA, possibly initially loaded by a ribozyme and that GlyRS would be the primordial aminoacyl-tRNA synthetase (Lei and Burton, 2020).

Diacritic Marks in tRNA Languages: Disambiguation and Link With Cellular State

Giegé et al. (2012) have described the modified bases of tRNA as true diacritical marks. The critical role of modified bases of ribonucleic acids, very particularly those of tRNAs, have been recognized since the early work of Grosjean and Björk (Björk and Neidhardt, 1975; Grosjean et al., 1995, Grosjean et al., 2010; Grosjean, 2015). Nowadays, they have become the subject of massive work in light of the large diversity of functions in which these molecules are involved (Phizicky and Hopper, 2015; Tuorto and Lyko, 2016; Agris et al., 2017; Marín et al., 2017; Barraud and Tisné, 2019). Particularly, there is an increasing association between the modified status of tRNAs and a number of pathologies (Torres et al., 2014a; Huang et al., 2018; Pereira et al., 2018; Santos et al., 2019). Within the scope of this review, we would like to emphasize that the enzymes involved in a large number of post-transcriptional modifications of tRNA (methylation, thiolation and incorporation of amino acid derivatives, dihydrouridine synthase, among others) require cofactors that constitute key molecules in metabolic pathways. Within them, it is worth mentioning S-adenosylmethionine (AdoMet), folate and NADPH (see **Figure 1**) (Helm and Alfonzo, 2014; Danchin et al., 2020).

The vast majority of the modifications in the base at position 34 and in the corresponding ribose (excluding I and Q in eukaryotes), very particularly U34, as well as in the neighboring bases of the anticodon, require the participation of some of the above-mentioned cofactors. It seems relevant to highlight the high metabolic cost of modifying U34, that relies on several enzymatic steps including methylation and thiolation processes (El Yacoubi et al., 2012; Karlsborn et al., 2014; Ranjan and Rodnina, 2016; Boccaletto et al., 2018), or the biosynthesis of the modified guanine Yw base, which may require up to six AdoMet molecules (Young and Bandarian, 2013). All this indicates the close link between the state of these tRNA modifications and the cellular metabolic context. On the contrary, deamination of A34 in bacteria and eukaryotes, and introduction of Q34 in eukaryotes, seem to be essentially linked to disambiguation requirements in translation. In bacteria, queuosine biosynthesis is an expensive metabolic process,
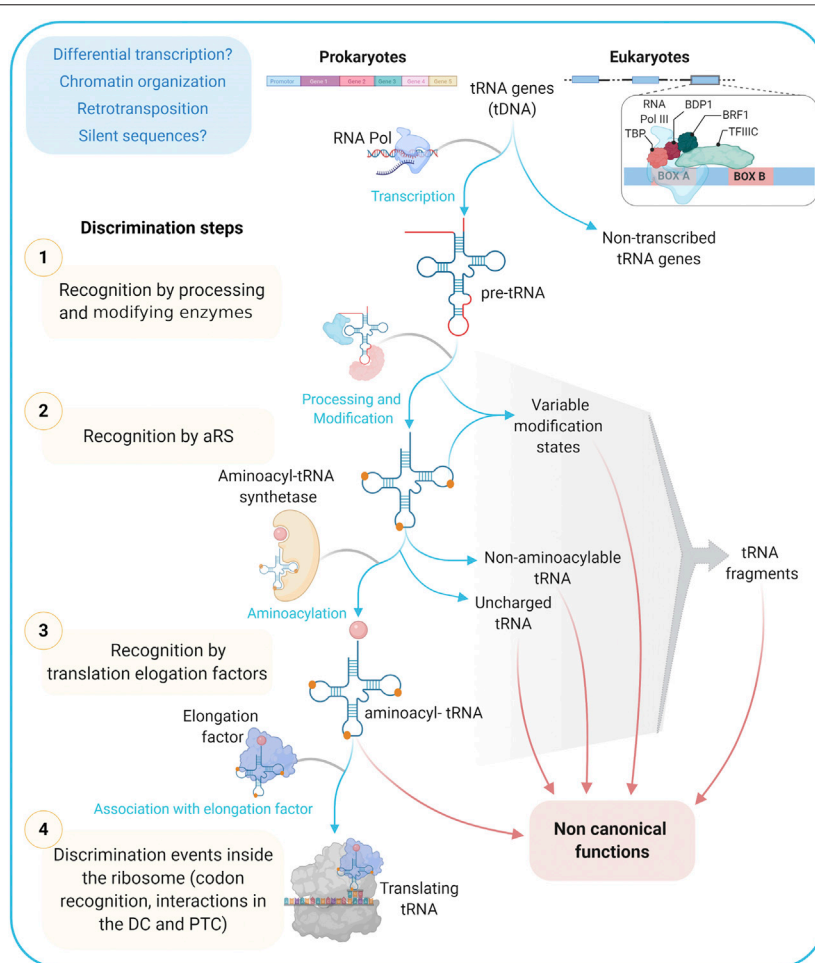
FIGURE 3 | tRNA informative layers. From transcription to codon recognition in the ribosome, tRNAs should go through many discrimination steps to ensure the translation fidelity. Alternative paths from these steps/checkpoints may provide a wide repertoire of tRNAs with functions different to the ribosomal protein synthesis and that are together referred to as non-canonical ones. *TBP:* TATA Binding Protein, Transcription Factor (TF); *BDP1:* B Double Prime 1, TFIIIB subunit also known as TFIIIB150; *BRF1:* TFIIIB 90 kDa subunit, *TFIIIC:* Transcription Factor IIIC; *aRS:* Aminoacyl tRNA Synthetase; *DC:* Decoding Center; *PTC:* Peptidyl-Transferase Center.

requiring AdoMet, THF, NADPH and GTP (Slany and Kersten, 1994; Phillips et al., 2008).

It is interesting to feature the close link between the set of processes involved in one-carbon metabolism with the translation apparatus through tRNA modifications, that has been nicely reviewed elsewhere (Helm and Alfonzo, 2014; Danchin et al., 2020).

Missing tRNAs or Large Reserves of Functional Alternatives?

The role of the high number of tRNA genes is still unclear (Iben and Maraia, 2014; Maraia and Arimbasseri, 2017; Pan, 2018; Torres, 2019) and it was suggested that an important part of them are actually silent (Torres, 2019). Interestingly, given the large size of the cellular tRNA population (which in higher eukaryotes is estimated to be around tens of millions of copies), it is expected that they participate in a large number of

interactions with other macromolecules with a high degree of specificity in the context of several cellular processes besides protein synthesis (Pan, 2018). Half a century has passed since the first reports on extra-ribosomal roles of tRNAs (Soffer et al., 1969; Leibowitz and Soffer, 1970), and many evidence has accumulated on the so-called "non-canonical" functions of tRNAs to differentiate them from their key role in ribosomal protein biosynthesis (Raina and Ibba, 2014; Katz et al., 2016; Schimmel, 2018; Su et al., 2020). The list continues to grow, most particularly with the increasing description of cellular properties of tRNA fragments that have received much attention during the last years. All this leads to think about the significance of the tRNA population in the cell in a whole new way. Therefore, it is appealing to establish a correlation between the high number of putative tRNA gene sequences coding for tRNAs with non-canonical structures and the diversity of functions described. This poses semantic problems referred to the nomenclature of the genes, and still

logistical problems linked to the search and identification of their possible products.

A primary classification into canonical and non-canonical tRNA genes can be attempted based on the structure of both the transcript and the promoter. Canonical genes code for molecules recognized and loaded by the cognate aminoacyl-tRNA synthetase and involved in ribosomal biosynthesis of proteins. Instead, non-canonical genes, apart from the obvious non-transcribed sequences (see further below), would include: 1) sequences coding for molecules with high scores of structure conservation that are aminoacylated by the cognate synthetase, but that will contribute to non-ribosomal synthesis, 2) genes coding for molecules capable of adopting alternative folds that fulfill protein synthesis-unrelated functions (see **Figure 3**).

Among tRNA roles outside ribosomal protein synthesis, the following have been previously highlighted: nutrient sensing, transcription regulation, retroelement insertion, translation kinetics and protein folding, stress response, immune response, apoptosis inhibition, peptidic antibiotic biosynthesis, bacterial wall biosynthesis, post-translational protein modification, membrane lipid modification, retroviral replication, mitochondrial ribosome assembly, and mitochondrial DNA replication (Seligmann, 2010; Raina and Ibba, 2014; Katz et al., 2016; Balasubramaniam et al., 2017; Bogenhagen et al., 2018; Su et al., 2020; and this issue). Also, several roles have been described for tRNA fragments, among which the following stand out: gene silencing, translation regulation, transposable element regulation, noncoding RNA regulation, cell differentiation, cell proliferation and cancer, host defense, stress response, apoptosis, and epigenetic inheritance (Magee and Rigoutsos, 2020; Polacek and Ivanov, 2020; Su et al., 2020; and this issue). In most cases, the information about the precise isotype, state of post-transcriptional modification, recognition by the cognate aminoacyl-tRNA synthetase, and aminoacylation state is still scarce, leaving the door open for new research avenues.

Here we describe a few illustrative examples of non-canonical functions for which many layers of evidence exist. In the available *Staphylococcus aureus* sp. genomes, seven different genes have been annotated to encode tRNA^{Gly} isoacceptors. One of them, a tRNA^{Gly}UCC, has a very low score in the tRNA database and presents variations in the DHU and TΨCR loops. Notably, though it is efficiently charged by the cognate glycyl-tRNA synthetase, it is not efficiently recognized by the elongation factor EF-Tu and is rather involved in the pentaglycine bridge synthesis during the bacterial wall formation (Giannouli et al., 2009). Interestingly, two other tRNA^{Gly}UCC putative genes carrying similar variants in the TΨCR loop exist in the *S. aureus* genome.

In humans, a low-expressed and non-aminoacylable tRNA^{Asp}GUC was reported to play a particular role in the regulation of aspartyl-tRNA synthetase gene transcription through direct binding to an Alu sequence in the 3′UTR (Rudinger-Thirion et al., 2011). This molecule is similar to other human tRNA^{Asp} sequences competent to adopt the classical cloverleaf tRNA structure, although it has significant changes in the TΨCR stem and loop domain. Furthermore, the

authors also demonstrated that it is not able to fold into a canonical tertiary tRNA structure.

More recently, by analyzing human tRNA-Seq datasets through a bioinformatics strategy, it was shown that changes in tRNA gene expression not only include the relative abundance of mature tRNA but also significant changes in immature tRNA sequences and tRNA fragments. Importantly, this has been associated with different biological functions unrelated to protein synthesis (Torres, 2019). This approach paves the way to delve into the analysis of the biological significance of non-canonical tRNA sequences. Such non-canonical tRNAs are usually encoded by tDNAs that have been mostly (but not completely) removed from the last databases due to the use of high-stringency algorithms, mainly based on the occurrence of alterations in DHU and TΨCR loops and/or mispairings. This is exemplified by tRNA^{Ala}AGC putative coding sequences in human and *Danio rerio* genomes, which dropped from 29 and 92 in 2015 to 26 and 6 in 2019, respectively. Nevertheless, sequences with variations in the TΨCR loop still remain. This strongly suggests the importance of keeping the discussion open.

The above-described cases correspond to tRNA genes carrying anticodons that are not absent, i.e., that there are other isodecoders for that particular anticodon able to perform the canonical adaptor function. Nevertheless, altogether, this data emphasizes that although several non-canonical tRNA gene sequences previously labeled as "missing genes" would be effectively silent, many others might have important activities, either similar to those described here, or even alternative functions that wait to be discovered (Pan, 2018; Torres, 2019).

tDNA Genes and Pseudogenes: Adapter Roles in Chromatin Structure?

Unlike prokaryotes, where tRNAs are clustered in polycistronic genes and/or included in the rDNA locus, eukaryotic tRNA genes are largely monocistronic and are considered as independent units scattered throughout the genome. They constitute real middle repetitive DNA elements and part of them could be considered as pseudogenes, either silent or poorly expressed (Torres, 2019). However, they could have soundly impacts from a structural perspective.

Indeed, there is growing evidence underscoring the role of tDNAs in chromatin organization, independent of their capacity to be transcribed or not. tRNAs are transcribed by a dedicated RNA polymerase (RNA pol III) and their promoters are bound by several specific factors, including TFIIIB, TFIIIA and TFIIIC. The latter is a multi-subunit complex which recognizes the two conserved A and B elements located in the transcribed sequence overlapping the DHU and TΨCG tRNA loops respectively (Arimbasseri and Maraia, 2016; Arimbasseri, 2018) (see **Figure 1**). In particular, the critical binding of TFIIIC to B box sequences, both in active promoters and in scattered sequences in the genome, has been pointed out as an essential contributor to chromatin architecture (Noma et al., 2006; Wallrath and Geyer, 2006). Furthermore, it has been described that tDNA domains contain binding sites for numerous chromatin-binding proteins such as cohesins and

condensins complexes, as well as some SMC (structural maintenance of chromatin) proteins (Hamdani et al., 2019) that are involved in the organization of high order chromosomes, as well as in the structure of chromatin that changes the accessibility of DNA to regulatory factors, thus impacting on gene expression (Hübner and Spector, 2010). Moreover, tDNA sequences have also been implicated in several other activities at the chromatin level. Among them stand out: nucleosome positioning, insulator role blocking the spreading of heterochromatin domains, replication fork pausing that mediates long-range interactions defining chromosome architecture, and recombination events (Noma et al., 2006; Wallrath and Geyer, 2006; Arimbasseri and Maraia, 2016; Ciesla et al., 2018; Shukla and Bhargava, 2018; Hamdani et al., 2019; Torres, 2019). Interestingly, tDNA genomic loci have also been described as targets for transposon insertion (Cheung et al., 2018).

In turn, and constituting a feed-back loop, the genomic context and the local structure of chromatin would define the bases for the differential expression of tRNA genes in eukaryotes, stressing the vital role of chromatin organization in tRNA-mediated responses (Arimbasseri and Maraia, 2016; Sagi et al., 2016; Ciesla et al., 2018; Shukla and Bhargava, 2018).

CONCLUSION: READING THE PALIMPSEST

The challenging problem of the origin of the genetic code and the explanation of why amino acids are encoded by a different number of synonymous codons is still open (Di Giulio, 2005; Ribas de Pouplana et al., 2017; Lei and Burton, 2020) as well as it is the problem of why certain anticodons were excluded throughout evolution. Furthermore, the multiplicity of functions of these molecules adds an additional level of complexity.

There is a general conception that despite tRNAs are among the most ancient and highly conserved molecules, they are poor phylogenetic markers because they are short, and often subject to horizontal gene transfer and recombination events (Widmann et al., 2010). However, they constitute precious material for deciphering early events in the origin of life (Widmann et al., 2010; Ribas de Pouplana et al., 2017; Lei and Burton, 2020). The structure of each individual molecule and that of the entire cellular tRNA population, could be seen as a historical summary of the origin and evolution of the storage and expression of biological information, but also of the regulatory mechanisms that ensure precise adaptation between metabolism, changes of the cell program and environmental conditions.

tRNA has emerged as an adaptor molecule, associated with the first events of the establishment of life (Di Giulio, 2005). Progressively, the successful adaptor structure would have acquired new functions involving very diverse, but at the same time specific, intermolecular interactions. Without going into the possible temporal succession of evolutionary events, it is clear that the reached structure was acquiring new biological roles, both in prokaryotes and eukaryotes, accumulating signals that ensure the fidelity of the molecular interactions. This important diversity of functions fulfilled by tRNAs leads to an expansion of Crick's concept of adapter molecule (Ribas de Pouplana and Dedon, 2014).

In bacteria and archaea, the same set of tRNAs could perform canonical and non-canonical functions, as multifunctional molecules. In eukaryotes, the significant increase in the number of tRNA genes, their different genomic organization, and the existence of a dedicated RNA polymerase, would be accompanied by the appearance of specially dedicated tRNA molecules for extra-ribosomal functions, with "admitted" and "prohibited" anticodons, increasing the activities in which tRNA molecules are involved.

Canonical and non-canonical functions require the reduction of ambiguity of signals that define the specific interactions in which tRNA molecules are involved. This necessary disambiguation would be at the origin of the exclusion of particular isodecoding options. Besides the basic structural elements that ensure the fidelity of the interactions, a level of modulation of the signals deposited in the tRNA structure by post-transcriptional modifications must be also considered.

In conclusion, tRNA genes and molecules appear as a remarkable example of molecular palimpsests and evolutionary puzzles. As in the work of an archaeologist, the layers of information should be delicately removed one after the other. This advance in their decryption might now be accomplished by the use of state-of-the-art technologies.

AUTHOR CONTRIBUTIONS

RE and MM developed the idea for the article. RE performed the majority of the literature search. MD, IL, CC and MM contributed additional topics of interest to be included. RE and MD prepared the figures. RE, MD, IL, CC and MM wrote the manuscript. All authors read and approved the submitted version of the manuscript.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2021.643701/full#supplementary-material.

REFERENCES

Agris, P. F., Eruysal, E. R., Narendran, A., Väre, V. Y. P., Vangaveti, S., and Ranganathan, S. V. (2018). Celebrating wobble decoding: half a century and still much is new. *RNA Biol.* 15, 537–553. doi:10.1080/15476286.2017.1356562

Agris, P. F., Narendran, A., Sarachan, K., Väre, V. Y. P., and Eruysal, E. (2017). The importance of being modified: the role of RNA modifications in translational fidelity. *Enzymes* 41, 1–50. doi:10.1016/bs.enz.2017.03.005

Arimbasseri, A. G., and Maraia, R. J. (2016). RNA polymerase III advances: structural and tRNA functional views. *Trends Biochem. Sci.* 41, 546–559. doi:10.1016/j.tibs.2016.03.003

Arimbasseri, G. A. (2018). Interactions between RNAP III transcription machinery and tRNA processing factors. *Biochim. Biophys. Acta Gene Regul. Mech.* 1861, 354–360. doi:10.1016/j.bbagrm.2018.02.003

Balasubramaniam, M., Reis, R. J. S., Ayyadevara, S., Wang, X., Ganne, A., and Khaidakov, M. (2017). Involvement of tRNAs in replication of human mitochondrial DNA and modifying effects of telomerase. *Mech. Ageing Dev.* 166, 55–63. doi:10.1016/j.mad.2017.07.004

Barraud, P., and Tisné, C. (2019). To be or not to be modified: miscellaneous aspects influencing nucleotide modifications in tRNAs. *IUBMB Life* 71, 1126–1140. doi:10.1002/iub.2041

Björk, G. R., and Neidhardt, F. C. (1975). Physiological and biochemical studies on the function of 5-methyluridine in the transfer ribonucleic acid of *Escherichia coli*. *J. Bacteriol.* 124, 99–111. doi:10.1128/JB.124.1.99-111.1975

Boccaletto, P., Machnicka, M. A., Purta, E., Piatkowski, P., Baginski, B., Wirecki, T. K., et al. (2018). MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic Acids Res.* 46, D303–D307. doi:10.1093/nar/gkx1030

Bogenhagen, D. F., Ostermeyer-Fay, A. G., Haley, J. D., and Garcia-Diaz, M. (2018). Kinetics and mechanism of mammalian mitochondrial ribosome assembly. *Cell Rep.* 22, 1935–1944. doi:10.1016/j.celrep.2018.01.066

Chan, P. P., and Lowe, T. M. (2009). GtRNAdb: a database of transfer RNA genes detected in genomic sequence. *Nucleic Acids Res.* 37, D93–D97. doi:10.1093/nar/gkn787

Chan, P. P., and Lowe, T. M. (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

Chan, P. P., and Lowe, T. M. (2019). tRNAscan-SE: searching for tRNA genes in genomic sequences. *Methods Mol. Biol.* 1962, 1–14. doi:10.1007/978-1-4939-9173-0_1

Cheung, S., Manhas, S., and Measday, V. (2018). Retrotransposon targeting to RNA polymerase III-transcribed genes. *Mobile DNA* 9, 14. doi:10.1186/s13100-018-0119-2

Ciesla, M., Skowronek, E., and Boguta, M. (2018). Function of TFIIIC, RNA polymerase III initiation factor, in activation and repression of tRNA gene transcription. *Nucleic Acids Res.* 46, 9444–9455. doi:10.1093/nar/gky656

Cognat, V., Pawlak, G., Duchêne, A. M., Daujat, M., Gigant, A., Salinas, T., et al. (2013). PlantRNA, a database for tRNAs of photosynthetic eukaryotes. *Nucleic Acids Res.* 41, D273–D279. doi:10.1093/nar/gks935

Crick, F. H. (1966). Codon--anticodon pairing: the wobble hypothesis. *J. Mol. Biol.* 19 (2), 548–555. doi:10.1016/s0022-2836(66)80022-0

Danchin, A., Sekowska, A., and You, C. (2020). One-carbon metabolism, folate, zinc and translation. *Microb. Biotechnol.* 13, 899–925. doi:10.1111/1751-7915.13550

Delannoy, E., Le Ret, M., Faivre-Nitschke, E., Estavillo, G. M., Bergdoll, M., Taylor, N. L., et al. (2009). Arabidopsis tRNA adenosine deaminase arginine edits the wobble nucleotide of chloroplast tRNAArg(ACG) and is essential for efficient chloroplast translation. *Plant Cell* 21, 2058–2071. doi:10.1105/tpc.109.066654

Demeshkina, N., Jenner, L., Westhof, E., Yusupov, M., and Yusupova, G. (2012). A new understanding of the decoding principle on the ribosome. *Nature* 484, 256–259. doi:10.1038/nature10913

Dietrich, A., Small, I., Cosset, A., Weil, J. H., and Maréchal-Drouard, L. (1996). Editing and import: strategies for providing plant mitochondria with a complete set of functional transfer RNAs. *Biochimie* 78, 518–529. doi:10.1016/0300-9084(96)84758-4

Di Giulio, M. (2005). The origin of the genetic code: theories and their relationships, a review. *Biosystems* 80, 175–184. doi:10.1016/j.biosystems.2004.11.005

Di Giulio, M. (2020). The phylogenetic distribution of the glutaminyl-tRNA synthetase and Glu-tRNAGln amidotransferase in the fundamental lineages would imply that the ancestor of archaea, that of eukaryotes and LUCA were progenotes. *Biosystems* 196, 104174. doi:10.1016/j.biosystems.2020.104174

Dittmar, K. A., Goodenbour, J. M., and Pan, T. (2006). Tissue-specific differences in human transfer RNA expression. *PLoS Genet.* 2, e221. doi:10.1371/journal.pgen.0020221

Diwan, G. D., and Agashe, D. (2018). Wobbling forth and drifting back: the evolutionary history and impact of bacterial tRNA modifications. *Mol. Biol. Evol.* 35, 2046–2059. doi:10.1093/molbev/msy110

Duchêne, A.-M., Pujol, C., and Maréchal-Drouard, L. (2009). Import of tRNAs and aminoacyl-tRNA synthetases into mitochondria. *Curr. Genet.* 55, 1–18. doi:10.1007/s00294-008-0223-9

Eargle, J., Black, A. A., Sethi, A., Trabuco, L. G., and Luthey-Schulten, Z. (2008). Dynamics of recognition between tRNA and elongation factor Tu. *J. Mol. Biol.* 377, 1382–1405. doi:10.1016/j.jmb.2008.01.073

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. doi:10.1146/annurev-genet-110711-155641

Evans, M. E., Clark, W. C., Zheng, G., and Pan, T. (2017). Determination of tRNA aminoacylation levels by high-throughput sequencing. *Nucleic Acids Res.* 45, e133. doi:10.1093/nar/gkx514

Fergus, C., Barnes, D., Alqasem, M. A., and Kelly, V. P. (2015). The queuine micronutrient: charting a course from microbe to man. *Nutrients* 7, 2897–2929. doi:10.3390/nu7042897

Ferro, I., and Ignatova, Z. (2015). Quantifying the "escapers" among RNA species. *Biochem. Soc. Trans.* 43, 1215–1220. doi:10.1042/BST20150158

Fox, G. E. (2010). Origin and evolution of the ribosome. *Cold Spring Harb. Perspect. Biol.* 2, a003483. doi:10.1101/cshperspect.a003483

Fujishima, K., and Kanai, A. (2014). tRNA gene diversity in the three domains of life. *Front. Genet.* 5, 142. doi:10.3389/fgene.2014.00142

Giannouli, S., Kyritsis, A., Malissovas, N., Becker, H. D., and Stathopoulos, C. (2009). On the role of an unusual tRNAGly isoacceptor in *Staphylococcus aureus*. *Biochimie* 91, 344–351. doi:10.1016/j.biochi.2008.10.009

Giegé, R., Jühling, F., Pütz, J., Stadler, P., Sauter, C., and Florentz, C. (2012). Structure of transfer RNAs: similarity and variability. *Wiley Interdiscip. Rev. RNA* 3 (1), 37–61. doi:10.1002/wrna.103

Gingold, H., Tehler, D., Christoffersen, N. R., Nielsen, M. M., Asmar, F., Kooistra, S. M., et al. (2014). A dual program for translation regulation in cellular proliferation and differentiation. *Cell* 158, 1281–1292. doi:10.1016/j.cell.2014.08.011

Goodarzi, H., Nguyen, H. C. B., Zhang, S., Dill, B. D., Molina, H., and Tavazoie, S. F. (2016). Modulated expression of specific tRNAs drives gene expression and cancer progression. *Cell* 165, 1416–1427. doi:10.1016/j.cell.2016.05.046

Goodenbour, J. M., and Pan, T. (2006). Diversity of tRNA genes in eukaryotes. *Nucleic Acids Res.* 34, 6137–6146. doi:10.1093/nar/gkl725

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. doi:10.1016/j.febslet.2009.11.052

Grosjean, H., Sprinzl, M., and Steinberg, S. (1995). Posttranscriptionally modified nucleosides in transfer RNA: their locations and frequencies. *Biochimie* 77, 139–141. doi:10.1016/0300-9084(96)88117-x

Grosjean, H. (2015). RNA modification: the golden period 1995–2015. *RNA* 21, 625–626. doi:10.1261/rna.049866.115

Hamdani, O., Dhillon, N., Hsieh, T. S., Fujita, T., Ocampo, J., Kirkland, J. G., et al. (2019). tRNA genes affect chromosome structure and function via local effects. *Mol. Cell. Biol.* 39 (8), e00432–18. doi:10.1128/MCB.00432-18

Hatfull, G. F. (2015). Dark matter of the biosphere: the amazing world of bacteriophage diversity. *J. Virol.* 89, 8107–8110. doi:10.1128/JVI.01340-15

Helm, M., and Alfonzo, J. D. (2014). Posttranscriptional RNA modifications: playing metabolic games in a cell's chemical Legoland. *Chem. Biol.* 21, 174–185. doi:10.1016/j.chembiol.2013.10.015

Ho, J. M., Bakkalbasi, E., Söll, D., and Miller, C. A. (2018). Drugging tRNA aminoacylation. *RNA Biol.* 15, 667–677. doi:10.1080/15476286.2018.1429879

Holley, R. W., Apgar, J., Everett, G. A., Madison, J. T., Marquisee, M., Merrill, S. H., et al. (1965). Structure of a ribonucleic acid. *Science* 147, 1462–1465. doi:10.1126/science.147.3664.1462

Holman, K. M., Puppala, A. K., Lee, J. W., Lee, H., and Simonović, M. (2017). Insights into substrate promiscuity of human seryl-tRNA synthetase. *RNA* 23, 1685–1699. doi:10.1261/rna.061069.117

Huang, Q., Yao, P., Eriani, G., and Wang, E. D. (2012). *In vivo* identification of essential nucleotides in tRNALeu to its functions by using a constructed yeast tRNALeu knockout strain. *Nucleic Acids Res.* 40, 10463–10477. doi:10.1093/nar/gks783

Huang, S. Q., Sun, B., Xiong, Z. P., Shu, Y., Zhou, H. H., Zhang, W., et al. (2018). The dysregulation of tRNAs and tRNA derivatives in cancer. *J. Exp. Clin. Cancer Res.* 37, 101. doi:10.1186/s13046-018-0745-z

Hübner, M. R., and Spector, D. L. (2010). Chromatin dynamics. *Annu. Rev. Biophys.* 39, 471–489. doi:10.1146/annurev.biophys.093008.131348

Iben, J. R., and Maraia, R. J. (2014). tRNA gene copy number variation in humans. *Gene* 536, 376–384. doi:10.1016/j.gene.2013.11.049

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. doi:10.1093/nar/gkn772

Karcher, D., and Bock, R. (2009). Identification of the chloroplast adenosine-to-inosine tRNA editing enzyme. *RNA* 15, 1251–1257. doi:10.1261/rna.1600609

Karlsborn, T., Tükenmez, H., Mahmud, A. K., Xu, F., Xu, H., and Byström, A. S. (2014). Elongator, a conserved complex required for wobble uridine modifications in eukaryotes. *RNA Biol.* 11, 1519–1528. doi:10.4161/15476286.2014.992276

Katz, A., Elgamal, S., Rajkovic, A., and Ibba, M. (2016). Non-canonical roles of tRNAs and tRNA mimics in bacterial cell biology. *Mol. Microbiol.* 101, 545–558. doi:10.1111/mmi.13419

Kim, S. H., Suddath, F. L., Quigley, G. J., McPherson, A., Sussman, J. L., Wang, A. H., et al. (1974). Three-dimensional tertiary structure of yeast phenylalanine transfer RNA. *Science* 185, 435–440. doi:10.1126/science.185.4149.435

Kimura, S., Dedon, P. C., and Waldor, M. K. (2020). Comparative tRNA sequencing and RNA mass spectrometry for surveying tRNA modifications. *Nat. Chem. Biol.* 16, 964–972. doi:10.1038/s41589-020-0558-1

Kirchner, S., Cai, Z., Rauscher, R., Kastelic, N., Anding, M., Czech, A., et al. (2017). Alteration of protein function by a silent polymorphism linked to tRNA abundance. *PLoS Biol.* 15, e2000779. doi:10.1371/journal.pbio.2000779

Kirchner, S., and Ignatova, Z. (2015). Emerging roles of tRNA in adaptive translation, signalling dynamics and disease. *Nat. Rev. Genet.* 16, 98–112. doi:10.1038/nrg3861

Knie, N., Polsakiewicz, M., and Knoop, V. (2015). Horizontal gene transfer of chlamydial-like tRNA genes into early vascular plant mitochondria. *Mol. Biol. Evol.* 32, 629–634. doi:10.1093/molbev/msu324

Konevega, A. L., Soboleva, N. G., Makhno, V. I., Semenkov, Y. P., Wintermeyer, W., Rodnina, M. V., et al. (2004). Purine bases at position 37 of tRNA stabilize codon-anticodon interaction in the ribosomal A site by stacking and Mg^{2+}-dependent interactions. *RNA* 10, 90–101. doi:10.1261/rna.5142404

Ledoux, S., Olejniczak, M., and Uhlenbeck, O. C. (2009). A sequence element that tunes *Escherichia coli* tRNA(Ala)(GGC) to ensure accurate decoding. *Nat. Struct. Mol. Biol.* 16, 359–364. doi:10.1038/nsmb.1581

Lei, L., and Burton, Z. F. (2020). Evolution of life on Earth: tRNA, aminoacyl-tRNA synthetases and the genetic code. *Life* 10 (3), 21. doi:10.3390/life10030021

Leibowitz, M. J., and Soffer, R. L. (1970). Enzymatic modification of proteins. 3. purification and properties of a leucyl, phenylalanyl transfer ribonucleic acid protein transferase from *Escherichia coli*. *J. Biol. Chem.* 245, 2066–2073. doi:10.1016/s0021-9258(18)63207-7

Lowe, T. M., and Chan, P. P. (2016). tRNAscan-SE on-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res.* 44, W54–W57. doi:10.1093/nar/gkw413

Magee, R., and Rigoutsos, I. (2020). On the expanding roles of tRNA fragments in modulating cell behavior. *Nucleic Acids Res.* 48, 9433–9448. doi:10.1093/nar/gkaa657

Mandal, D., Köhrer, C., Su, D., Russell, S. P., Krivos, K., Castleberry, C. M., et al. (2010). Agmatidine, a modified cytidine in the anticodon of archaeal tRNA(Ile), base pairs with adenosine but not with guanosine. *Proc. Natl. Acad. Sci. U.S.A.* 107, 2872–2877. doi:10.1073/pnas.0914869107

Maraia, R. J., and Arimbasseri, A. G. (2017). Factors that shape eukaryotic tRNAomes: processing, modification and anticodon–codon use. *Biomolecules* 7 (1), 26. doi:10.3390/biom7010026

Marck, C., and Grosjean, H. (2002). tRNomics: analysis of tRNA genes from 50 genomes of Eukarya, Archaea, and Bacteria reveals anticodon-sparing strategies and domain-specific features. *RNA* 8, 1189–1232. doi:10.1017/s1355838202022021

Maréchal-Drouard, L., Guillemaut, P., Cosset, A., Arbogast, M., Weber, F., Weil, J. H., et al. (1990). Transfer RNAs of potato (*Solanum tuberosum*) mitochondria have different genetic origins. *Nucleic Acids Res.* 18, 3689–3696. doi:10.1093/nar/18.13.3689

Marín, M., Fernández-Calero, T., and Ehrlich, R. (2017). Protein folding and tRNA Biology. *Biophys. Rev.* 9, 573–588. doi:10.1007/s12551-017-0322-2

Melnikov, S., Ben-Shem, A., Garreau de Loubresse, N., Jenner, L., Yusupova, G., and Yusupov, M. (2012). One core, two shells: bacterial and eukaryotic ribosomes. *Nat. Struct. Mol. Biol.* 19, 560–567. doi:10.1038/nsmb.2313

Michaud, M., Cognat, V., Duchêne, A. M., and Maréchal-Drouard, L. (2011). A global picture of tRNA genes in plant genomes. *Plant J.* 66, 80–93. doi:10.1111/j.1365-313X.2011.04490.x

Morgado, S., and Vicente, A. (2019). Global *in-silico* scenario of tRNA genes and their organization in virus genomes. *Viruses* 11, 180. doi:10.3390/v11020180

Müller, M., Legrand, C., Tuorto, F., Kelly, V. P., Atlasi, Y., Lyko, F., et al. (2019). Queuine links translational control in eukaryotes to a micronutrient from bacteria. *Nucleic Acids Res.* 47, 3711–3727. doi:10.1093/nar/gkz063

Murphy, F. V., and Ramakrishnan, V. (2004). Structure of a purine-purine wobble base pair in the decoding center of the ribosome. *Nat. Struct. Mol. Biol.* 11, 1251–1252. doi:10.1038/nsmb866

Noma, K., Cam, H. P., Maraia, R. J., and Grewal, S. I. (2006). A role for TFIIIC transcription factor complex in genome organization. *Cell* 125, 859–872. doi:10.1016/j.cell.2006.04.028

Novoa, E. M., Pavon-Eternod, M., Pan, T., and Ribas de Pouplana, L. (2012). A role for tRNA modifications in genome structure and codon usage. *Cell* 149, 202–213. doi:10.1016/j.cell.2012.01.050

Numata, T. (2015). Mechanisms of the tRNA wobble cytidine modification essential for AUA codon decoding in prokaryotes. *Biosci. Biotechnol. Biochem.* 79, 347–353. doi:10.1080/09168451.2014.975185

Nureki, O., O'Donoghue, P., Watanabe, N., Ohmori, A., Oshikane, H., Araiso, Y., et al. (2010). Structure of an archaeal non-discriminating glutamyl-tRNA synthetase: a missing link in the evolution of Gln-tRNAGln formation. *Nucleic Acids Res.* 38, 7286–7297. doi:10.1093/nar/gkq605

Olejniczak, M., and Uhlenbeck, O. C. (2006). tRNA residues that have coevolved with their anticodon to ensure uniform and accurate codon recognition. *Biochimie* 88, 943–950. doi:10.1016/j.biochi.2006.06.005

Orioli, A. (2017). tRNA biology in the omics era: stress signalling dynamics and cancer progression. *Bioessays* 39 (3). doi:10.1002/bies.201600158

Oshikane, H., Sheppard, K., Fukai, S., Nakamura, Y., Ishitani, R., Numata, T., et al. (2006). Structural basis of RNA-dependent recruitment of glutamine to the genetic code. *Science* 312, 1950–1954. doi:10.1126/science.1128470

Pan, T. (2018). Modifications and functional genomics of human transfer RNA. *Cell Res.* 28, 395–404. doi:10.1038/s41422-018-0013-y

Pang, Y. L., Abo, R., Levine, S. S., and Dedon, P. C. (2014). Diverse cell stresses induce unique patterns of tRNA up- and down-regulation: tRNA-seq for quantifying changes in tRNA copy number. *Nucleic Acids Res.* 42, e170. doi:10.1093/nar/gku945

Pavon-Eternod, M., Gomes, S., Geslain, R., Dai, Q., Rosner, M. R., and Pan, T. (2009). tRNA over-expression in breast cancer and functional consequences. *Nucleic Acids Res.* 37, 7268–7280. doi:10.1093/nar/gkp787

Pereira, M., Francisco, S., Varanda, A., Santos, M., Santos, M., and Soares, A. (2018). Impact of tRNA modifications and tRNA-modifying enzymes on proteostasis and human disease. *Int. J. Mol. Sci.* 19, 3738. doi:10.3390/ijms19123738

Pernod, K., Schaeffer, L., Chicher, J., Hok, E., Rick, C., Geslain, R., et al. (2020). The nature of the purine at position 34 in tRNAs of 4-codon boxes is correlated with nucleotides at positions 32 and 38 to maintain decoding fidelity. *Nucleic Acids Res.* 48, 6170. doi:10.1093/nar/gkaa221

Petrov, A. S., Bernier, C. R., Hsiao, C., Norris, A. M., Kovacs, N. A., Waterbury, C. C., et al. (2014). Evolution of the ribosome at atomic resolution. *Proc. Natl. Acad. Sci. U.S.A.* 111, 10251–10256. doi:10.1073/pnas.1407205111

Phillips, G., and de Crécy-Lagard, V. (2011). Biosynthesis and function of tRNA modifications in Archaea. *Curr. Opin. Microbiol.* 14, 335–341. doi:10.1016/j.mib.2011.03.001

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. doi:10.1128/JB.00874-08

Phizicky, E. M., and Hopper, A. K. (2015). tRNA processing, modification, and subcellular dynamics: past, present, and future. *RNA* 21, 483–485. doi:10.1261/rna.049932.115

Pinkard, O., McFarland, S., Sweet, T., and Coller, J. (2020). Quantitative tRNA-sequencing uncovers metazoan tissue-specific tRNA regulation. *Nat. Commun.* 11, 4104. doi:10.1038/s41467-020-17879-x

Polacek, N., and Ivanov, P. (2020). The regulatory world of tRNA fragments beyond canonical tRNA Biology. *RNA Biol.* 17, 1057–1059. doi:10.1080/15476286.2020.1785196

Pope, W. H., Bowman, C. A., Russell, D. A., Jacobs-Sera, D., Asai, D. J., Cresawn, S. G., et al. (2015). Whole genome comparison of a large collection of mycobacteriophages reveals a continuum of phage genetic diversity. *Elife* 4, e06416. doi:10.7554/eLife.06416

Rafels-Ybern, À., Torres, A. G., Camacho, N., Herencia-Ropero, A., Roura Frigolé, H., Wulff, T. F., et al. (2019). The expansion of inosine at the wobble position of tRNAs, and its role in the evolution of proteomes. *Mol. Biol. Evol.* 36, 650–662. doi:10.1093/molbev/msy245

Rafels-Ybern, À., Torres, A. G., Grau-Bove, X., Ruiz-Trillo, I., and Ribas de Pouplana, L. (2018). Codon adaptation to tRNAs with Inosine modification at position 34 is widespread among Eukaryotes and present in two bacterial phyla. *RNA Biol.* 15, 500–507. doi:10.1080/15476286.2017.1358348

Raina, M., and Ibba, M. (2014). tRNAs as regulators of biological processes. *Front. Genet.* 5, 171. doi:10.3389/fgene.2014.00171

Rak, R., Dahan, O., and Pilpel, Y. (2018). Repertoires of tRNAs: the couplers of genomics and proteomics. *Annu. Rev. Cell Dev. Biol.* 34, 239–264. doi:10.1146/annurev-cellbio-100617-062754

Ranjan, N., and Rodnina, M. V. (2016). tRNA wobble modifications and protein homeostasis. *Translation* 4, e1143076. doi:10.1080/21690731.2016.1143076

Ribas de Pouplana, L., and Dedon, P. C. (2014). More than an adaptor molecule: the emerging role of tRNA in cell signaling and disease. *FEBS Lett.* 588, 4267. doi:10.1016/j.febslet.2014.10.022

Ribas de Pouplana, L., Torres, A. G., and Rafels-Ybern, À. (2017). What froze the genetic code?. *Life* 7 (2), 14. doi:10.3390/life7020014

Rogers, T. E., Ataide, S. F., Dare, K., Katz, A., Seveau, S., Roy, H., et al. (2012). A pseudo-tRNA modulates antibiotic resistance in *Bacillus cereus*. *PLoS One* 7, e41248. doi:10.1371/journal.pone.0041248

Roy, H., Becker, H. D., Mazauric, M. H., and Kern, D. (2007). Structural elements defining elongation factor Tu mediated suppression of codon ambiguity. *Nucleic Acids Res.* 35, 3420–3430. doi:10.1093/nar/gkm211

Rudinger-Thirion, J., Lescure, A., Paulus, C., and Frugier, M. (2011). Misfolded human tRNA isodecoder binds and neutralizes a 3' UTR-embedded Alu element. *Proc. Natl. Acad. Sci. U.S.A.* 108, E794–E802. doi:10.1073/pnas.1103698108

Sagi, D., Rak, R., Gingold, H., Adir, I., Maayan, G., Dahan, O., et al. (2016). Tissue- and time-specific expression of otherwise identical tRNA genes. *PLoS Genet.* 12, e1006264. doi:10.1371/journal.pgen.1006264

Saint-Léger, A., Bello, C., Dans, P. D., Torres, A. G., Novoa, E. M., Camacho, N., et al. (2016). Saturation of recognition elements blocks evolution of new tRNA identities. *Sci. Adv.* 2, e1501860. doi:10.1126/sciadv.1501860

Salinas-Giegé, T., Giegé, R., and Giegé, P. (2015). tRNA Biology in mitochondria. *Int. J. Mol. Sci.* 16, 4518–4559. doi:10.3390/ijms16034518

Santos, M., Fidalgo, A., Varanda, A. S., Oliveira, C., and Santos, M. A. S. (2019). tRNA deregulation and its consequences in cancer. *Trends Mol. Med.* 25, 853–865. doi:10.1016/j.molmed.2019.05.011

Sauerwald, A., Zhu, W., Major, T. A., Roy, H., Palioura, S., Jahn, D., et al. (2005). RNA-dependent cysteine biosynthesis in archaea. *Science* 307, 1969–1972. doi:10.1126/science.1108329

Schimmel, P. (2018). The emerging complexity of the tRNA world: mammalian tRNAs beyond protein synthesis. *Nat. Rev. Mol. Cell Biol.* 19, 45–58. doi:10.1038/nrm.2017.77

Schneider, A. (2011). Mitochondrial tRNA import and its consequences for mitochondrial translation. *Annu. Rev. Biochem.* 80, 1033–1053. doi:10.1146/annurev-biochem-060109-092838

Seligmann, H. (2010). Mitochondrial tRNAs as light strand replication origins: similarity between anticodon loops and the loop of the light strand replication origin predicts initiation of DNA replication. *Biosystems* 99, 85–93. doi:10.1016/j.biosystems.2009.09.003

Sheppard, K., and Söll, D. (2008). On the evolution of the tRNA-dependent amidotransferases, GatCAB and GatDE. *J. Mol. Biol.* 377, 831–844. doi:10.1016/j.jmb.2008.01.016

Sheppard, K., Yuan, J., Hohn, M. J., Jester, B., Devine, K. M., and Söll, D. (2008). From one amino acid to another: tRNA-dependent amino acid biosynthesis. *Nucleic Acids Res.* 36, 1813–1825. doi:10.1093/nar/gkn015

Shigematsu, M., Honda, S., Loher, P., Telonis, A. G., Rigoutsos, I., and Kirino, Y. (2017). YAMAT-seq: an efficient method for high-throughput sequencing of mature transfer RNAs. *Nucleic Acids Res.* 45, e70. doi:10.1093/nar/gkx005

Shukla, A., and Bhargava, P. (2018). Regulation of tRNA gene transcription by the chromatin structure and nucleosome dynamics. *Biochim. Biophys. Acta Gene Regul. Mech.* 1861, 295–309. doi:10.1016/j.bbagrm.2017.11.008

Shukla, M. K., and Leszczynski, J. (2013). Tautomerism in nucleic acid bases and base pairs: a brief overview. *WIREs Comput. Mol. Sci.* 3, 637–649. doi:10.1002/wcms.1145

Singh, V., Fedeles, B. I., and Essigmann, J. M. (2015). Role of tautomerism in RNA Biochemistry. *RNA* 21, 1–13. doi:10.1261/rna.048371.114

Slany, R. K., and Kersten, H. (1994). Genes, enzymes and coenzymes of queuosine biosynthesis in procaryotes. *Biochimie* 76, 1178–1182. doi:10.1016/0300-9084(94)90047-7

Soffer, R. L., Horinishi, H., and Leibowitz, M. J. (1969). The aminoacyl tRNA-protein transferases. *Cold Spring Harb. Symp. Quant. Biol.* 34, 529–533. doi:10.1101/sqb.1969.034.01.060

Soma, A., Ikeuchi, Y., Kanemasa, S., Kobayashi, K., Ogasawara, N., Ote, T., et al. (2003). An RNA-modifying enzyme that governs both the codon and amino acid specificities of isoleucine tRNA. *Mol. Cell* 12, 689–698. doi:10.1016/S1097-2765(03)00346-0

Su, Z., Wilson, B., Kumar, P., and Dutta, A. (2020). Noncanonical roles of tRNAs: tRNA fragments and beyond. *Annu. Rev. Genet.* 54, 47. doi:10.1146/annurev-genet-022620-101840

Suzuki, T., and Numata, T. (2014). Convergent evolution of AUA decoding in bacteria and archaea. *RNA Biol.* 11, 1586–1596. doi:10.4161/15476286.2014.992281

Tharp, J. M., Ad, O., Amikura, K., Ward, F. R., Garcia, E. M., Cate, J. H. D., et al. (2020). Initiation of protein synthesis with non-canonical amino acids *in vivo*. *Angew. Chem. Int. Ed. Engl.* 59, 3122–3126. doi:10.1002/anie.201914671

Topal, M. D., and Fresco, J. R. (1976). Complementary base pairing and the origin of substitution mutations. *Nature* 263, 285–289. doi:10.1038/263285a0

Torres, A. G., Batlle, E., and Ribas de Pouplana, L. (2014a). Role of tRNA modifications in human diseases. *Trends Mol. Med.* 20, 306–314. doi:10.1016/j.molmed.2014.01.008

Torres, A. G., Piñeyro, D., Filonava, L., Stracker, T. H., Batlle, E., and Ribas de Pouplana, L. (2014b). A-to-I editing on tRNAs: biochemical, biological and evolutionary implications. *FEBS Lett.* 588, 4279–4286. doi:10.1016/j.febslet.2014.09.025

Torres, A. G. (2019). Enjoy the silence: nearly half of human tRNA genes are silent. *Bioinform. Biol. Insights* 13, 1177932219868454. doi:10.1177/1177932219868454

Tumbula, D. L., Becker, H. D., Chang, W. Z., and Söll, D. (2000). Domain-specific recruitment of amide amino acids for protein synthesis. *Nature* 407, 106–110. doi:10.1038/35024120

Tuorto, F., and Lyko, F. (2016). Genome recoding by tRNA modifications. *Open Biol.* 6 (12), 160287. doi:10.1098/rsob.160287

Turner, B., Burkhart, B. W., Weidenbach, K., Ross, R., Limbach, P. A., Schmitz, R. A., et al. (2020). Archaeosine modification of archaeal tRNA: role in structural stabilization. *J. Bacteriol.* 202 (8), e00748–19. doi:10.1128/JB.00748-19

Vangaveti, S., Cantara, W. A., Spears, J. L., DeMirci, H., Murphy, F. V., Ranganathan, S. V., et al. (2020). A structural basis for restricted codon recognition mediated by 2-thiocytidine in tRNA containing a wobble position inosine. *J. Mol. Biol.* 432, 913–929. doi:10.1016/j.jmb.2019.12.016

Veerareddygari, G. R., Singh, S. K., and Mueller, E. G. (2016). The pseudouridine synthases proceed through a glycal intermediate. *J. Am. Chem. Soc.* 138, 7852–7855. doi:10.1021/jacs.6b04491

Voorhees, R. M., Mandal, D., Neubauer, C., Köhrer, C., RajBhandary, U. L., and Ramakrishnan, V. (2013). The structural basis for specific decoding of AUA by isoleucine tRNA on the ribosome. *Nat. Struct. Mol. Biol.* 20, 641–643. doi:10.1038/nsmb.2545

Wallrath, L. L., and Geyer, P. K. (2006). TFIIIC boxes in the genome. *Cell* 125, 829–831. doi:10.1016/j.cell.2006.05.016

Wan, W., Tharp, J. M., and Liu, W. R. (2014). Pyrrolysyl-tRNA Synthetase: an ordinary enzyme but an outstanding genetic code expansion tool. *Biochim. Biophys. Acta* 1844, 1059–1070. doi:10.1016/j.bbapap.2014.03.002

Wang, C., Guo, Y., Tian, Q., Jia, Q., Gao, Y., Zhang, Q., et al. (2015). SerRS-tRNASec complex structures reveal mechanism of the first step in selenocysteine biosynthesis. *Nucleic Acids Res.* 43, 10534–10545. doi:10.1093/nar/gkv996

Warren, J. M., and Sloan, D. B. (2020). Interchangeable parts: the evolutionarily dynamic tRNA population in plant mitochondria. *Mitochondrion* 52, 144–156. doi:10.1016/j.mito.2020.03.007

Westhof, E., Yusupov, M., and Yusupova, G. (2014). Recognition of Watson-Crick base pairs: constraints and limits due to geometric selection and tautomerism. *F1000Prime Rep.* 6, 19. doi:10.12703/P6-19

Westhof, E. (2014). Isostericity and tautomerism of base pairs in nucleic acids. *FEBS Lett.* 588, 2464–2469. doi:10.1016/j.febslet.2014.06.031

Widmann, J., Harris, J. K., Lozupone, C., Wolfson, A., and Knight, R. (2010). Stable tRNA-based phylogenies using only 76 nucleotides. *RNA* 16, 1469–1477. doi:10.1261/rna.726010

Yokobori, S., Kitamura, A., Grosjean, H., and Bessho, Y. (2013). Life without tRNAArg-adenosine deaminase TadA: evolutionary consequences of decoding the four CGN codons as arginine in Mycoplasmas and other Mollicutes. *Nucl. Acids Res.* 41 (13), 6531–6543. doi:10.1093/nar/gkt356

Young, A. P., and Bandarian, V. (2013). Radical mediated ring formation in the biosynthesis of the hypermodified tRNA base wybutosine. *Curr. Opin. Chem. Biol.* 17, 613–618. doi:10.1016/j.cbpa.2013.05.035

Yuan, J., O'Donoghue, P., Ambrogelly, A., Gundllapalli, S., Sherrer, R. L., Palioura, S., et al. (2010). Distinct genetic code expansion strategies for selenocysteine and pyrrolysine are reflected in different aminoacyl-tRNA formation systems. *FEBS Lett.* 584, 342–349. doi:10.1016/j.febslet.2009.11.005

Zhang, N., Shi, S., Wang, X., Ni, W., Yuan, X., Duan, J., et al. (2020). Direct sequencing of tRNA by 2D-HELS-AA MS seq reveals its different isoforms and dynamic base modifications. *ACS Chem. Biol.* 15, 1464–1472. doi:10.1021/acschembio.0c00119

Zhang, W., Yu, Y., Hertwig, F., Thierry-Mieg, J., Zhang, W., Thierry-Mieg, D., et al. (2015). Comparison of RNA-seq and microarray-based models for clinical endpoint prediction. *Genome Biol.* 16, 133. doi:10.1186/s13059-015-0694-1