The natural history of transfer RNA and its interactions with the ribosome

Gustavo Caetano-Anollés1* and Feng-Jie Sun2

1 Evolutionary Bioinformatics Laboratory, Department of Crop Sciences, University of Illinois, Urbana-Champaign, IL, USA
2 School of Science and Technology, Georgia Gwinnett College, Lawrenceville, GA, USA

*Correspondence: gca@illinois.edu

Edited by:
Akio Kanai, Keio University, Japan
Reviewed by:
Kosuke Fujishima, NASA Ames Research Center, USA
Savio Torres Farias, Universidade Federal da Paraíba, Brazil

Keywords: structure, phylogenetic analysis, sequence, non-coding RNA, translation, ribosome, origin of life

Transfer RNA (tRNA) is undoubtedly the most central and one of the oldest molecules of the cell. Without it genetics and coded protein synthesis are impossible. The crucial specificities responsible for the genetic code and accurate translation are by far entrusted to interactions between tRNA and translation proteins, fundamentally aminoacyl-tRNA synthetase (aaRS) enzymes and elongation factor (EF) switches (Yadavalli and Ibba, 2012). Discrimination mediated by aaRSs and EFs against misincorporated tRNA and amino acids is at least 20 times more stringent than ribosomal recognition, editing, and other proofreading mechanisms (Reynolds et al., 2010). The fact that crucial genetic code specificities in highly selective interactions with protein enzymes do not involve the ribosomal ribonucleoprotein biosynthetic machinery challenges the “replacers first” origin of life scenario of an ancient RNA world (Caetano-Anollés and Seufferheld, 2013). It also highlights the central functional, mechanistic, and evolutionary roles of tRNA and its recognition determinants, which enable coevolution between nucleic acids and proteins. These coevolutionary relationships are compatible with a late origin of the ribosome in its mechanism and not in protein biosynthesis, which was inferred from the computational analysis of thousands of RNAs and proteomes (Harish and Caetano-Anollés, 2012). These analyses showed tight coevolution of ribosomal RNA (rRNA) and ribosomal proteins (r-proteins). While these relationships delimit molecular makeup when organisms use translation to negotiate growth and viability amidst environmental change, coevolution also constrains recruitment of the canonical L-shaped structure of the tRNA molecule into a multiplicity of modern functions. These new functions include the synthesis of antibiotics, bacterial cell wall peptidoglycans and tetapyrroles, modification of bacterial membrane lipids, protein turnover, and the synthesis of other aminoacyl-tRNA molecules (Francklyn and Minajigi, 2010). Here we unfold coevolutionary relationships between tRNA substructures and translation proteins that embody crucial protein-nucleic acid interactions. We focus on a series of computational biology analyses of the structure and conformational diversity of tRNAs and their interacting proteins that provide information about the history of structural accretion of this “adaptor” molecule. Using this information, we place tRNA history within the framework of an evolutionary timeline of protein domain evolution, uncovering the natural history of tRNA within the context of the geological record.

tRNA MOLECULES ARE OLD AND EVOLVE BY ACCRETION OF STRUCTURAL PARTS

When studying the organismal distribution of a catalog of over a thousand RNA families describing the modern RNA world, tRNA was found to be one of only five families that were universally present (Hoeppner et al., 2012). These families showed a strong vertical evolutionary trace and included rRNA and ribonuclease P (RNase P) RNA, which are present (with exceptions; e.g., Randau et al., 2008) in all studied cellular organisms and are minimally affected by horizontal gene transfer. We note however that RNA-free RNase P (Gutmann et al., 2012; Taschner et al., 2012) can challenge RNase P RNA ancestry (Sun and Caetano-Anollés, 2010). The ubiquity of tRNA in the cellular lineages of life and its central molecular role provide strong support to the very early origin of the molecule, prompting the study of the origin and evolution of the tRNA molecule using information in its sequence and structure (Fitch and Upper, 1987; Eigen et al., 1989; Di Giulio, 1994; Sun and Caetano-Anollés, 2008a; Farias, 2013). A computational analysis of the history of tRNA based on the structure of thousands of molecules revealed that tRNAs evolve by accretion of component parts (substructures) and that the “top half” of tRNA that includes the acceptor stem is more ancient than the “bottom half” with its anticodon arm (Sun and Caetano-Anollés, 2008a; reviewed in Sun and Caetano-Anollés, 2008b) (Figure 1A). While other models of evolutionary growth of the tRNA molecule have been proposed (Di Giulio, 2012), phylogenetic reconstructions are compatible with biochemical evidence of molecular recognition that makes amino acid charging ancestral and molecularly distant (∼70 Å) to codon recognition, which locate to more modern regions of tRNA (Caetano-Anollés et al., 2013). These findings revive the “genomic tag” hypothesis in which tRNA harbored ancestral genomic information and the derived bottom half provided genetic code specificity (Weiner and Maizels, 1987).
FIGURE 1 | The natural history of tRNA inferred from nucleic acid-protein interactions and structural phylogenomics. (A) The history of tRNA portrays the history of its interactions with cognate aminocacyl-tRNA synthetase (aaRS) protein enzymes. This is exemplified by the domains of the tRNA and cysteinyl-tRNA synthetase binary complex (PDB entry 1U0B), which are colored according to their age. The ancient “top half” of tRNA embeds a “operational code” in the identity elements of the acceptor arm that interact with the catalytic domain of aaRSs through classes I and II modes of tRNA recognition. The evolutionarily recent “bottom half” of tRNA holds the standard code in identity elements of the anticodon loop that interact with anticodon-binding domains of aaRSs. (B) Flow diagram showing the retrodiction strategy used to build phylogenetic trees of RNA molecules (ToMs) and associated trees of substructures (ToSs), and trees of protein domains (ToDs). The structures of RNA molecules are first decomposed into substructures. Structural features of substructures such as helical stem tracts and unpaired regions are coded as phylogenetic characters and assigned character states according to an evolutionary model that polarizes character transformation toward an increase in conformational order (character argumentation). Coded characters (s) are arranged in data matrices, which can be transposed. Phylogenetic analysis using maximum parsimony optimality criteria generates rooted ToMs and ToSs. A census of domain structures in proteomes of hundreds of completely sequenced organisms is used to compose data matrices, which are then used to build ToDs. Elements of the matrix (g) represent genomic abundances of domain structures in proteomes, defined at different levels of classification of domain structure (e.g., SCOP folds, superfamilies, and families). They are converted into multistate phylogenetic characters with character states transforming according to linearly ordered and reversible pathways. Embedded in the trees of nucleic acids and proteins are timelines that assign age to molecular structures and associated functions. (C) The natural history of tRNA and rRNA overlap when they are mapped onto a timeline of protein domain history. A tree of tRNA substructures (ToS) was derived from statistical phylogenetic characters that define a molecular morphospace (the Shannon entropy of the base-pairing probability matrix, base-pairing propensity and mean length of stem structures) in 571 tRNA molecules. The optimal most parsimonious tree (43,281 steps; consistency index = 0.853, retention index = 0.654, rescaled consistency index = 0.557, g1 = −1.033) was recovered from a branch-and-bound search. The most basal subtree of a ToS describing the evolution of the rRNA core (Harish and Caetano-Anollés, 2012) is also shown. Both trees are anchored to the geological record via an evolutionary timeline of first appearance of protein domains that are capable of establishing crucial interactions with the RNA molecules (see description in the main text). AC, anticodon; PTC, peptidyl transferase center.
PHYLOGENOMIC RETRODICTI
UNCOVERS COEVOLUTION BETWEEN
tRNA SUBSTRUCTURES AND
INTERACTING aaRS PROTEIN
DOMAINS

In the studies mentioned above, phylo-
genetic analysis of nucleic acid structure
was directly derived from structural topol-
ogy and the thermodynamics of tRNA
(Caetano-Anollés, 2002a,b; Sun et al.,
2007; Sun and Caetano-Anollés, 2008a),
taking unique advantage of links that exist
between secondary structure and confor-
mation, dynamics, and adaptation (Bailor
et al., 2010). Specifically, a census of geo-
metrical features that describe the length
and topology of tRNA substructures (such
as stem and non-paired segments) or sta-
tistical features describing their stability
and conformational diversity were ana-
lyzed with modern phylogenetic methods
to produce phylogenetic trees of molecules
(ToMs) and trees of substructures (ToSs)
that portray the history of the system
(molecules) or its component parts (sub-
structures), respectively. Figure 1C shows
a ToS that describes the evolution of stem
substructures of the tRNA molecule
and of early evolving stem substructures
of rRNA. The trees that are produced
are rooted using a phylogenetic process
model that complies with Weston’s gen-
erality criterion. The model automatically
roots the trees by assuming conforma-
tional stability increases in evolution as
structures become canalized (Sun et al.,
2010). The validity of polarization and
rooting depends on the axiomatic com-
ponent of character transformation, which
is falsifiable and supported by consider-
able evidence (e.g., thermodynamic and
phylogenetic; Sun et al., 2010).

While ToSs are powerful retrodiction
statements that unfold history of
RNA accretion (Sun and Caetano-Anollés,
2008a,b,c, 2009, 2010; Sun et al., 2007;
Harish and Caetano-Anollés, 2012), the
gradual appearance of protein domains
in evolutionary history can be inferred
from phylogenetic trees of domains
(ToDs) (Figure 1B) (Caetano-Anollés and
Caetano-Anollés, 2003) and can illus-
trate the establishment of intermolecu-
lar interactions in evolution. Domains are
structural and evolutionary units of pro-
teins that are highly conserved
(Caetano-Anollés et al., 2009). The evolutionary
accumulation of these units unfolds recur-
rence patterns that encompass the entire
history of proteins and can be mined with
suitable phylogenomic methods. ToDs are
derived from a structural census of protein
domains in the proteomes of hundreds
to thousands of genomes that have been
completely sequenced. The fold structures
domains are defined using the different
levels of structural abstraction of the
accepted classification gold standards, the
SCOP (Murzin et al., 1995) or CATH
(Orengo et al., 1997) databases. Timelines
of domain innovation are then derived
directly from the trees taking advan-
tage of their highly imbalanced nature.
Imbalance unfolds when the splitting of
lineages depends on an evolving “heri-
table” trait (Heard, 1996). In our case,
the evolving trait is the gradual accu-
cumulation of domains in proteomes and
the semipunctuated discovery of new fold
structures (made evident for example in
simulations; Zeldovich et al., 2007). The
predictive power of ToDs is consider-
able (Caetano-Anollés and Seufferheld,
2013) and central for the history of
tRNA, as ToDs have established the evo-
lutionary history of aaRS domain struc-
tures and their associated coevolving tRNA
molecules (Caetano-Anollés et al., 2013).
The timeline of evolutionary appearance
of folds families revealed the early emer-
gence of the “operational” RNA code
linked to the specificities of synthetases
that were homologous to the catalytic
domains of modern TyrRS and SerRS pro-
tein enzymes. These archaic synthetases
interacted with the “top half” of tRNA
and were capable of peptide bond for-
mation and aminoacylation (Caetano-
Anollés et al., 2013). The timeline also
showed the late implementation of the
standard genetic code with the late appear-
ance of anticondon-binding domains that
interacted with the “bottom half” of tRNA.
Figure 1A shows a representative
aaRS enzyme and the tight coevolution-
ary link between aaRS domains and tRNA
arms. Remarkably, structural phyloge-
nomic retrodictions indicate that genet-
ics arose through episodes of structural
recruitment as an exacting mechanism
that favored flexibility and folding of the
emergent proteins (Caetano-Anollés et al.,
2013). These enhancements of phenotypic
robustness matched evolutionary trends of
folding speed in proteins (Debes et al.,
2013) and are compatible with recent simu-
lations of the origin of the genetic code
(Jee et al., 2013).

ABUNDANCE OF PROTEIN DOMAINS
IN PROTEOMES FOLLOWS AN
E VOLUTIONARY CLOCK

The history of RNA does not repres-
ent a phylogenetic statement that applies
to the entire world of RNA molecules.
Consequently, it cannot be placed within
a global historical context. In contrast,
the history of protein domains inferred
from ToDs follows a global molecular clock
of fold structures that spans 3.8 billion years
(Gy) of evolution (Wang et al., 2011).
Traditionally, molecular clocks are based
on rates of change in protein or nucleic
acid sequences, which are limited by his-
storical information existing in the individ-
ual protein or nucleic acid molecules being
studied (Zuckerkandl and Pauling, 1965;
Ayala et al., 1998). These clocks are there-
fore constrained by the highly dynamic
nature of sequence change, including the
problems of mutational saturation and
rate heterogeneity (heterotachy). In con-
trast, molecular structures exhibit char-
acteristics of recurrent change that are
much more stable. The clocks of domain
structures were calibrated by associating
diagnostic domain structures with mul-
tiple geological ages derived from the
study of fossils and microfossils, geochem-
ical, biochemical, and biomarker data.
Remarkably, excellent linear correlations
between the ages of domain structures at
fold and fold superfamly levels of SCOP
and geological timescales were identified
and used to time fundamental evolution-
ary events (Wang et al., 2011). These
events included the rise of planetary oxy-
gen and episodes of organizational diversi-
fication (Wang et al., 2011; Kim et al.,
2012).

THE CLOVERLEAF STRUCTURE OF tRNA
UNFOLDS EARLY IN EVOLUTION, PRIOR
TO THE APPEARANCE OF A
FUNCTIONAL RIBOSOMAL
MACHINERY

Assuming that the age of interactions
that are established between RNA and pro-
teins is the age of the interacting com-
ponents, we tracked the appearance of
domains in ribonucleoprotein complexes
along the evolutionary timeline and used the molecular clock of folds to link interactions to a geological timescale (Figure 1C). The catalytic domains of classes I and II aaRS enzymes (belonging to SCOP families d.104.I.1 and c.26.1.1, respectively) are the first to appear in the timeline ~3.7 Gy ago (Caetano-Anollés et al., 2013). These domains harbor pre-transfer and post-transfer editing and trans-editing activities. The most ancient of these editing structures, present in the catalytic domains of TyrRS, SerRS, and LeuRS, involve interactions with the oldest type II cognate tRNAs, which harbor a long variable loop necessary for tRNA recognition (Sun and Caetano-Anollés, 2008c). While the evolutionary significance of the variable loop in tRNA-aaRS interactions is unclear (Sun and Caetano-Anollés, 2008c), its late evolutionary appearance could simply represent the shift or recruitment of an archaic interacting region of the molecule. Interactions of tRNA with the “ValRS/LleRS/LeuRS editing” domain (SCOP family b.51.1.1) (Hale et al., 1997) suggest the D arm was already present ~3.3 Gy ago, which is derived compared to the acceptor stem (Sun and Caetano-Anollés, 2008a). The late appearance of anticodon-binding domains (beginning with SCOP family c.51.1.1) in well over half of aaRSs ~3 Gy ago confirms that the full “bottom half” of tRNA and its anticodon loop identity elements unfolded completely before the onset of planetary oxygenation and cellular diversification ~2.9 Gy ago.

Comparing the natural history of tRNA (Sun and Caetano-Anollés, 2008a) and the ribosome (Harish and Caetano-Anollés, 2012) within the framework of the interacting proteins shows the remarkable functional connection of the cloverleaf structure and ribosomal functionality (Figure 1C). The origin of r-proteins in interaction with helix 44 (the ribosomal ratchet) of the small subunit (SSU) tRNA occurred 3.3–3.4 Gy ago once the tRNA molecule unfolded its anticodon arm. This manifests in the pivotal role of one of the two earliest r-proteins, S12, in tRNA selection (anticipated by Ogle and Ramakrishnan, 2005), which is mediated by a bonding network connecting two sites in S12 to the anticodon and the CCA arm of the tRNA-elongation factor bound state (Li et al., 2008). Similarly, the full cloverleaf structure of tRNA was already present when the ribosomal peptidyl transferase center (PTC) responsible for modern protein synthesis appeared in the emerging domain V of the large subunit of tRNA 2.8–3.1 Gy ago. This is an expected outcome since the structurally mature 70–80 Å-long and 20–25 Å-wide tRNA molecule must traverse a path of ~100 Å and physically span the intersubunit interface of the ribosomal core for the ensemble to be fully functional (Agirrezabala and Frank, 2009). Remarkably, this late development of the ribosomal core coincided with the appearance of pathways of amino acid (Kim et al., 2012) and purine nucleotide biosynthesis (Caetano-Anollés and Caetano-Anollés, 2013). This suggests that tRNA and ribosomal functionality (anticodon loop recognition, decoding, protein biosynthesis) and modern metabolic pathways for amino acids and nucleotides developed concurrently, supporting the co-evolution theory of the genetic code (Wong, 2005).

CONCLUSION

The natural and overlapping history of tRNA and rRNA reveals that: (1) the tRNA cloverleaf structure unfolded prior to the appearance of a fully functional ribosomal core, (2) the primordial role of tRNA, originally linked to archaic dipeptide-forming synthetases, was coopted into modern translation functions once anticodon-loop specificities appeared concurrently with the PTC, and (3) the emergence of modern genetics unfolded relatively quickly in a period of 0.3–0.5 Gy, starting with anticodon-loop recognition and once the cloverleaf structure had formed.

REFERENCES

Agirrezabala, X., and Frank, J. (2009). Elongation in translation as a dynamic interaction among the ribosome, tRNA, and elongation factors EF-G and EF-Tu. Q. Rev. Biophys. 42, 159–200. doi: 10.1017/S0033585309990060

Ayala, F. J., Rzhetsky, A., and Ayala, F. J. (1998). Origin of the metazoan phase: molecular clocks confirm paleontological studies. Proc. Natl. Acad. Sci. U.S.A. 95, 606–611. doi: 10.1073/pnas.95.2.606

Bailor, M. H., Sun, X., and Al-Hashimi, H. M. (2010). Topology links RNA secondary structure with global conformation, dynamics, and adaptation. Science 327, 202–206. doi: 10.1126/science.1181085

Caetano-Anollés, G. (2002a). Tracing the evolution of RNA structure in ribosomes. Nucleic Acids Res. 30, 2575–2587. doi: 10.1093/nar/30.11.2575

Caetano-Anollés, G. (2002b). Evolved RNA secondary structure and the rooting of the universal tree of life. J. Mol. Evol. 54, 333–345. doi: 10.1007/s00239-001-0048-3

Caetano-Anollés, G., and Caetano-Anollés, D. (2003). An evolutionarily structured universe of protein architecture. Genome Res. 13, 1563–1571. doi: 10.1101/gr.1161903

Di Giulio, M. (1994). The phylogeny of tRNA molecules and the origin of the genetic code. Origin. Life Evol. Biosph. 24, 425–434. doi: 10.1007/BF01582018

Di Giulio, M. (2012). The origin of the tRNA molecule: independent data favor a specific model of its evolution. Biochimie 94, 1464–1466. doi: 10.1016/j.bioch.2012.01.014

Eigen, M., Lindemann, B. F., Tietze, M., Winkler-Oswatitsch, R., Dress, A., and von Haeseler, A. (1989). How old is the genetic code? Science 244, 673–679. doi: 10.1126/science.2497522

Farias, S. T. (2013). Suggested phylogeny of tRNAs based on the construction of ancestral sequences. J. Theor. Biol. 335, 245–248. doi: 10.1016/j.jtib.2013.06.033

Fitch, W. M., and Upper, K. (1987). The phylogeny of tRNA sequences provides evidence for ambiguity reduction in the origin of the genetic code. Cold Spring Harb. Symp. Quant. Biol. 52, 759–767. doi: 10.1101/SQB.1987.52.01.085

Franky, C. S., and Minajigi, A. (2010). tRNA as active chemical scaffold for diverse chemical transformations. FEBS Lett. 584, 366–375. doi: 10.1016/j.febslet.2009.11.045

Gutmann, B., Gobert, A., and Giegé, P. (2012). PRORP proteins support RNAP activity in both organelles and the nucleus in arabidopsis. Genes Dev. 26, 1022–1027. doi: 10.1101/gad.189514.112

Hale, S. P., Auld, D. S., Schmidt, E., and Schimmel, P. (1997). Discrete determinants in transfer RNA for editing and aminoclaylation. Science 276, 1250–1252. doi: 10.1126/science.276.5316.1250

Harish, A., and Caetano-Anollés, G. (2012). Ribosomal history reveals origins of modern
protein synthesis. *PLoS ONE* 7:e32276. doi: 10.1371/journal.pone.0032276

Heard, S. B. (1996). Patterns of phylogenetic tree balance with variable or evolving speciation rates. *Evolution* 50, 2141–2148. doi: 10.2307/2410685

Hoepfner, M. P., Gardner P. P., and Poole, A. M. (2012). Comparative analysis of RNA families reveals distinct repertoire trees for each domain of life. *PLoS Comput. Biol.* 8:e1002752. doi: 10.1371/journal.pcbi.1002752

Jee, J., Sundstrom, A., Massey, S. E., and Mishra, B. (2013). What can information-asymmetric games tell us about the context of Crick’s ‘frozen accident’? *J. R. Soc. Interface* 10:20130614. doi: 10.1098/rsif.2013.0614

Kim, K. M., Qin, T., Jiang, Y.-Y., Xiong, M., Caetano-Anollés, D., Zhang, H.-Y., et al. (2012). Protein domain structure uncovers the origin of aero-

bic metabolism and the rise of planetary oxygen. *Structure* 20, 67–76. doi: 10.1016/j.str.2011.11.003

Li, W., Agirrezabala, X., Lei, J., Bouakaz, L., Brunelle, J. L., Ortiz-Meoz, R. F., et al. (2008). Recognition of aminocyl-tRNA: a common molecular mechanism revealed by cryo-EM. *EMBO J.* 27, 3322–3331. doi: 10.1038/emboj.2008.243

Muzrin, A. G., Brenner, S. E., Hubbard, T., and Chothia, C. (1995). SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* 247, 536–540. doi: 10.1016/S0022-2836(95)80134-2

Ogle, J. M., and Ramakrishnan, V. (2005). Structural insights into translational fidelity. *Annu. Rev. Biochem.* 74, 129–177. doi: 10.1146/annurev.biochem.74.061903.155440

Orengo, C. A., Michie, A., Jones, S., Jones, D. T., Swindells, M., and Thornton, J. M. (1997). CATH–a hierarchical classification of protein domain structures. *Structure* 5, 1093–1109. doi: 10.1016/S0969-2126(97)00260-8

Randau, L., Schroder, I., and Soll, D. (2008). Life without RNase P. *Nature* 453, 120–123. doi: 10.1038/nature06833

Reynolds, N. M., Lazazzera, B. A., and Ibba, M. (2010). Cellular mechanisms that control mis-

translation. *Nat. Rev. Microbiol.* 8, 849–856. doi: 10.1038/nrmicro2472

Sun, F.-J., and Caetano-Anollés, G. (2008a). The origin and evolution of tRNA inferred from phyloge-

netic analysis of structure. *J. Mol. Evol.* 66, 21–35. doi: 10.1007/s00239-007-9050-8

Sun, F.-J., and Caetano-Anollés, G. (2008b). Transfer RNA and the origin of diversified life. *Sci. Prog.* 91, 265–284. doi: 10.1384/003685008X360650

Sun, F.-J., and Caetano-Anollés, G. (2008c). Evolutionary patterns in the sequence and structure of transfer RNA: a window into early translation and the genetic code. *PLoS ONE* 3:e2799. doi: 10.1371/journal.pone.0002799

Sun, F.-J., and Caetano-Anollés, G. (2009). The evolution-

ary history of the structure of SS riboso-

mal RNA. *J. Mol. Evol.* 69, 430–443. doi: 10.1007/s00239-009-9264-z

Sun, F.-J., and Caetano-Anollés, G. (2010). The ancient history of the structure of ribonuclease P and the early origins of archaea. *BMC Bioinformatics* 11:153. doi: 10.1186/1471-2105-11-153

Sun, F.-J., Fleurdepine, S., Bousquet-Antonelli, C., Caetano-Anollés, G., and Deragon, J.-M. (2007). Common evolutionary trends for SINE RNA structures. *Trends Genet.* 23, 26–33. doi: 10.1016/j.tig.2006.11.005

Sun, F.-J., Harish, A., and Caetano-Anollés, G. (2010) “Phylogenetic utility of RNA structure: evolution’s arrow and emergence of modern biochemistry and diversified life,” in *Evolutionary Bioinformatics and Systems Biology*, ed G. Caetano-Anollés (Hoboken, NJ: Wiley-Blackwell), 329–360.

Taschner, A., Weber, C., Buzet, A., Hartmann, R. K., Hartig, A., and Rossmanith, W. (2012). Nuclear RNA P of *Trypanosoma brucei*: a single protein in place of the multicomponent RNA-protein complex. *Cell Rep.* 2, 19–25. doi: 10.1016/j.celrep.2012.05.021

Wang, M., Jiang, Y.-Y., Kim, K. M., Qu, G., Ji, H.-F., Mittenthal, J. E., et al. (2011). A universal molecular clock of protein folds and its power in tracing the early history of aerobic metabolism and planet oxygenation. *Mol. Biol. Evol.* 28, 567–582. doi: 10.1093/molbev/msq232

Weiner, A. M., and Maizels, N. (1987). tRNA-like structures tag the 3′ ends of genomic RNA molecules for replication: implications for the origin of protein synthesis. *Proc. Natl. Acad. Sci. U.S.A.* 84, 7383–7387. doi: 10.1073/pnas.84.21.7383

Wong, J. T. (2005). Coevolution theory of the genetic code at age thirty. *Bioessays* 27, 416–425. doi: 10.1002/bies.20208

Yadavalli, S. S., and Ibba, M. (2012). Quality control in aminoacyl-tRNA synthesis: its role in translational fidelity. *Adv. Protein Chem. Struct. Biol.* 86, 1–43. doi: 10.1016/B978-0-12-386497-0.00001-3

Zeldovich, K. B., Chen, P., Shakhnovich, B. E., and Shakhnovich, E. I. (2007). A first-principles model of early evolution: emergence of gene families, species and preferred protein folds. *PLoS Comput. Biol.* 3:e139. doi: 10.1371/journal.pcbi.100

30139

Zuckerkandl, E., and Pauling, L. (1965). “Evolutionary divergence and convergence in proteins,” in *Evolving Genes and Proteins*, eds V. Bryson and H. J. Vogel (New York, NY: Academic Press), 97–166.

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

Received: 30 March 2014; accepted: 22 April 2014; published online: 09 May 2014.

Citation: Caetano-Anollés G and Sun F-J (2014) The origin and evolution of tRNA. *Front. Genet.* 5:127. doi: 10.3389/fgene.2014.00127

This article was submitted to Non-Coding RNA, a section of the journal Frontiers in Genetics.

Copyright © 2014 Caetano-Anollés and Sun. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licen-

ser are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduc-

tion is permitted which does not comply with these terms.