Biogenesis and function of transfer RNA related fragments (tRFs)

Pankaj Kumar¹, Canan Kuscü¹, and Anindya Dutta^{1*}

Anindya Dutta; PO Box 800733, 1340 Jefferson Park Ave Jordan Hall Room 1232; ¹Department of Biochemistry and Molecular Genetics, University of Virginia School of Medicine, Charlottesville, Virginia, USA 22901; Email: ad8q@virginia.edu

* To whom correspondence should be addressed. Tel: [+1-434 - 924 – 2466]; Fax: [+1-434 - 924 – 5069]

Keywords: non-coding small RNA, tRF, tRNA, tRFdb

Abstract:

Non-coding small RNAs arise from the various un-annotated and annotated regions of the genome. When they arise from annotated genes, the non-coding small RNAs are functionally different from the parent genes. This is a brief review of one class of non-coding small RNAs, tRNA related fragments (tRFs), which are generated from tRNA. tRFs have been suggested to play roles in cell proliferation, priming of viral reverse transcriptases, regulation of gene expression, RNA processing, modulation of the DNA damage response, tumor suppression, and neurodegeneration.

Small RNA fragments: degradation products, or functional short RNAs?

The study of small RNAs has increased substantially in recent years owing to their inferred roles in gene regulation. Since the discovery of the first small RNA, *lin-4*, in *Caenorhabditis elegans* [1, 2] a large number of small RNAs have been identified in many organism [3]. Unbiased high throughput sequencing and bioinformatic analysis of short RNA (14-36 nt) have added to the family of small RNAs that are encoded in the genome as small genes (tRNAs, snoRNAs etc.) by identifying a large number of small RNA fragments that arise from processing of longer transcripts. MicroRNAs can also be thought of as small RNA fragments, derived from longer primary-microRNA transcripts, and so it seemed worthwhile to pay attention to other small RNA fragments derived from longer genomically encoded parental transcripts [4-6]. Like microRNAs, some of these short RNA fragments could have functions different from those of the annotated parent genes. Because of the need to differentiate small RNAs that arise simply from degradation of longer transcripts, the search of functional small RNA fragments has focused on small RNA fragments that originate from specific locations in the longer parental transcript, are discrete in size and are abundant. Noncoding RNAs longer than 200 nucleotides have been classified as long

noncoding RNAs. Thus any RNA fragments <200 nucleotides could be a short RNA fragment of interest. However, many of these fragments were discovered while sequencing size-selected short RNAs to discover new microRNAs, so that a greater emphasis has been placed on small RNA fragments that are <40 nucleotides. In the last few years there have been several papers on transfer RNA fragments (tRFs) <40 nucleotides that are derived from tRNA transcripts, and recent studies have implicated these small RNA fragments in specific biological functions such as suppression of gene expression, regulation of apoptosis and trans-generational epigenetic inheritance. Therefore this is an opportune time to review the field and discuss what we know and do not know about the biogenesis and biological functions of tRFs.

Types of transfer RNA fragments and their biogenesis

The tRNA genes are transcribed by RNA polymerase III as precursor tRNAs (pre-tRNA), which contain extra bases at their 5'- and 3'- ends called leader and trailer sequences, respectively [7]. RNase P removes the leader sequence [7, 8] whereas RNase Z removes the trailer sequence by endo-nucleolytic cleavage exactly at the first unpaired base (discriminator base) at the 3'-end of the tRNA [7, 9]. Finally, a non-templated single "CCA" sequence is added to the 3'-ends of the trailer-free tRNAs by the enzyme tRNA nucleotidyl transferase [9]. In vertebrates mature tRNAs are exported to the cytoplasm with the help of a nuclear export receptor (exportin-T in *Xenopus*) and this export requires the mature 5' and 3' end of the tRNA, including the added CCA [10].

High-throughput sequencing and analyses of small RNA fragments in several studies have uncovered a new class of "non-micro-short" RNAs that map to the known tRNA genes and are called tRFs [4, 11-14]. tRFs are classified into various types depending on where they map on the primary or mature tRNA transcript (**Fig. 1**). Following these

conventions a database of tRFs has been developed where each unique tRF has been given a number (name) [15]. One of the earliest discovered classes of tRFs is stress (and starvation) induced tRNA fragments called tiRNA (tiR), or tRNA halves, which are generated by specific cleavage in the anticodon loops of mature tRNAs and therefore are 31-40 bases long. Though they are named as stress fragments, they are also detected under non-stressed conditions [16]. There are two subclasses of tiR based on whether they include the sequence 5' or 3' of the anticodon cleavage site: 5tiR start from 5' end of mature tRNA and end in the anticodon loop, whereas 3tiR start from the anti-codon loop and end at 3' end of mature tRNA [17, 18]. Cleavage is performed by the ribonuclease (RNase (A or T)) angiogenin (RNY1 in yeast) and thus tiR have a 5' hydroxyl rather than a 5' phosphate [12, 19, 20], which is different from the ends generated by Dicer or Rnase III type enzymes, which generate microRNAs and siRNAs [21].

Another class of small RNAs derived from tRNA are those that are 14-30 nucleotides long and map to the ends of the mature or primary tRNA transcripts. These small RNAs, like microRNAs, have a 5' phosphate and a 3' hydroxyl group [12], and have recently received attention because of their size similarity to miRNA. On the basis of their mapped positions, these tRFs are of three types: tRF-5, tRF-3 and tRF-1 [4]. The tRF-5 and tRF-3 are generated from 5' and 3' ends of the mature tRNAs whereas tRF-1 are generated from the 3' ends of the primary tRNA transcripts [4]. tRF-5 are 14-30 bases long and are generated by a cleavage in the D-loop or the stem region between D-loop and anticodon loop of tRNA gene. They are mainly of three specific lengths and are therefore further divided into subclasses of tRF-5s: tRF-5a (14-16 nts), tRF-5b (22-24 nts) and tRF-5c (28-30 nts) [22]. tRF-3s are mainly either ~18 or ~22 nts, based on which they are subclassified as either tRF-3a or tRF-3b. Both the tRF-3 cleavage sites are in the TΨC loop. These tRF subclasses are seen in all analyzed small RNA data sets from humans to yeasts [22]. Note, that although there are

other tRNA fragments also seen in the high throughput sequencing data, suggesting that there is much more heterogeneity in tRFs than our classification suggests, the abundance of these other fragments is usually 2-3 orders of magnitude lower than the highest abundance fragments that are classified as tRFs in the tRF database (tRFdb) [Kumar, 2015 #69].

It has been suggested that some of the tRNA transcripts could be alternately folded like pre-miRNAs and hence may get processed by Drosha and Dicer protein complexes [11, 23-26]. However a global analysis of small RNA sequencing data in wild type and Dicer knock-out cells (or cells with the essential Drosha complex component DGCR8 knocked out) show no decrease in tRF abundance in human, mouse, *Drosophila* or *Schizosaccharomyces pombe* (*S. pombe*) [13, 22, 25]. Thus the role of Dicer and DGCR8 or Drosha in tRF generation appears to be an exception rather than the rule and the enzymes that generally liberate tRF-5 and tRF-3 from mature tRNA transcripts are not known. The 5' end analysis of tRF-3s suggests that some unknown enzyme, which cuts between A/U & A/U nucleotides in the single-stranded loop of the tRNA, generates this class of tRFs [15, 22], which may suggest that endonucleases that cut 3' to A or U residues (RNase A, PhyM or U2) may be involved in their biogenesis. The 3' ends of tRF-5s do not appear to have such a sequence specificity, and only one sub-class, tRF-5a, is generated by cleavage in a single-stranded loop. However, both tRF-5s and tRF-3s have 5'-phosphate and 3'-hydroxyl ends like microRNAs, making it unlikely that they are generated by classical RNase A type endonucleases implicated in tRNA generation that usually leave a 3' phosphate and a 5' hydroxyl residue at the cleavage site.

Analysis of tRNA fragments that originate from primary tRNA trailer sequences indicate that the 5' end of tRF-1 matches with the cut site of RNase Z and the 3' end matches with an RNA polymerase III (RNA pol III) transcription termination signal (UUUUU, UUCUU, GUCUU or AUCUU)[4, 27]. Together, this is strong evidence that this type of tRF is generated

by the endo-nucleolytic cleavage of pre-tRNAs during maturation. The Pol III transcription termination signal occurs at different locations in each pre-tRNA and as such, the length distribution of tRF-1 is not discrete, differing from tRF-5 and tRF-3 [22].

In addition to tiRs and tRFs mentioned above, there may be additional classes of tRNA fragments [28-30]. One contains only the anticodon stem and loop tRNA that we have classified here as tRF-2 [28, 29] (**Fig. 1**). Rarer tRFs have been reported in patients with a rare form of neurodegenerative disease associated with mutations in CLP1, a component of the tRNA splicing machinery. In these patients, the neurons accumulate either introns alone or the 3' exons that looks like a 3tiR {Schaffer, 2014 #88}{Karaca, 2014 #128}. Interesting, transfection of the 3tiR-like fragments appear to be cause neuronal cell death, and so accumulation of this fragment has been suggested as a possible cause of the neurodegeneration.

An important caveat: although we have classified the tRNA fragments in specific classes based on their location in the parental tRNA transcript, it is entirely possible that tRFs in a given class may have diverse functions, and likewise that low abundant fragments that do not conform to the classes of tRFs listed above may emerge with important biological functions.

Subcellular (cytoplasmic vs. nuclear) distribution of tRFs:

tRNA 3' end maturation occurs in the nucleus and hence tRF-1 is expected to be present in the nucleus. However tRF-1s are cytoplasmic, suggesting that they are exported to the cytoplasm by some unknown mechanism [4, 14, 22]. Liao et al. [14] found that tRF-1001, which should be generated in the nucleus during tRNA 3' end maturation, is in the cytoplasm, whereas small RNAs derived from the nuclear box C/D snoRNAs, which are also processed in nucleus, are not abundant in the cytoplasm [14]. Therefore, tRF-1 are likely to

be exported to the cytoplasm following their release in the nucleus. Our recent meta-analysis of small RNA data further suggested that tRF-5s are mostly in the nucleus (in HeLa cell,) whereas tRF-3s and tRF-1s are mostly cytoplasmic [15, 22]. tRNAs can be transported back into the nucleus after export to the cytoplasm [31], and so tRF-5s may be imported into the nucleus by similar mechanisms. The presence of tRF-5s uniquely in the nucleus raises interesting questions as to whether they affect gene regulation by epigenetic mechanisms similar to siRNAs in *S. pombe* or piRNAs in germ cells [32-34].

Comparison with miRNA biogenesis

miRNAs are processed by Dicer, which then hands them off to the Ago proteins [35-37]. Although tRF-5s and -3s are not processed by Dicer, they are surprisingly associated with Ago family of proteins, specifically AGO1, 3 and 4 [15, 22, 26, 38-41]. The Ago Photoactivatable-Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation (PAR-CLIP; see Glossary) and cross-linking ligation and sequencing of hybrids (CLASH; see Glossary) data suggest that tRF-3s (and some tRF-5s) may function like miRNAs by interacting with complementary sequences on target RNAs to recruit them to AGO protein-containing complexes and to regulate the expression or function of these targets. The site of cross-link between the thio-uracil and protein in PAR-CLIP experiments is often mutated from U to C during reverse-transcription and so the site of this mutation indicates how the short RNA (or its target RNA) associates with the Ago protein. PAR-CLIP data show that the tRFs associated with the Ago proteins are preferentially cross-linked just downstream from the 7 base seed sequence at the 5' ends of the tRFs [22]. Furthermore, the PAR-CLIP data also reveal several target RNA fragments in the Ago complex that (a) are complementary to the seed and only the seed of the tRFs, and (b) are cross-linked just upstream from the complementary region, much like targets of miRNAs [22]. Thus the PAR-CLIP data strongly

suggest that the tRFs that enter into AGO complexes are paired with their targets using their seed sequences, and that the anatomy of a tRF-Target-Ago complex is very much like that of a microRNA-Target-Ago complex.

AGO1 CLASH experiments identify chimeric RNAs between microRNAs and their target RNAs created by ligation of the two RNAs while they are associated with each other in the AGO1 protein. We identified many chimeras with tRFs ligated to 3'UTRs of annotated protein coding genes or with long noncoding RNAs. The targets RNA region captured in these chimeras have sequence complementary to the corresponding tRF's seed sequence, as would be expected if the target were recruited into the AGO complex by complementarity with the tRF's seed sequence. Intriguingly, the abundance of tRF-3-target RNA chimeras was 2-3 fold higher than that of miRNA-target RNA chimeras in the AGO1 precipitates, suggesting either that more tRFs than microRNA recruit target RNAs into AGO1 complexes, or that the anatomy of a tRF-Target-AGO1 complex somehow favors the ligation more than the anatomy of a microRNA-Target-AGO1 complex. Thus, like the PAR-CLIP data, the CLASH data unequivocally suggest that tRFs are associated with target RNAs with complementarity to the tRF seed sequences in AGO1 complexes [15, 22].

Although strongly suggestive, these data still leave unanswered whether these tRFs regulate the expression of the target RNAs given such micro-RNA-like interaction with their targets. This will be an important question to address experimentally in the future.

Biological function of tRFs

tRFs have been reported in viruses (Herpesvirus and human immunodeficiency virus-1 (HIV-1)) [40, 42], archaea (*Haloferax volcanii*) [43], bacteria [15, 22], protozoa (*Tetrahymena*) [12, 44], plants (*Hordeum vulgare* L. & rice embryogenic callus) [45, 46], chickens [14], mice [23, 47] and humans [4, 15, 22]. This class of small RNAs thus spans the

entire evolutionary tree, and biological roles have been identified for some tRFs in subsets of organisms [4, 26, 28, 40, 48]. The known functions of tRFs are summarized in Table 1. In the rest of this section we will review relatively well-characterized functions of tRFs from different classes. The tRF name is also mentioned while describing the tRF if an individual tRF is annotated in the tRFdb.

tRF-1s:

tRF-1001 is derived from the 3' end of a pre-tRNA^{Ser-TGA} (chr10:tRNA2-SerTGA) through cleavage by ELAC2/RNaseZ [4], and the expression of tRF-1001 is positively correlated with cell proliferation in prostate cancer cell lines. Most interestingly, tRF-1001 promotes the transition of these cells from the G2 phase of the cell-cycle to the M phase, and this function requires the sequence of the tRF as well as unblocked 5' and 3' ends [4]. However, there was no evidence in PAR-CLIP or CLASH data that any tRF-1, including tRF-1001, associates with Ago proteins suggesting that the G2-M regulatory function of tRF-1001 is probably not due to a gene-regulatory function similar to miRNAs. However, Haussecker *et al.* reported that tRF-1001 (the authors named it cand45) does associate with Argonaute proteins, particularly Ago-3 and Ago-4. Despite this, tRF-1001 overexpression did not repress the expression of a reporter containing the complementary sequence to tRF-1001, with the authors suggesting that this was because of the lack of slicer activity in Ago3 and Ago4. Addition of an oligonucleotide antisense to tRF-1001 (sense to the reporter gene) paradoxically silenced the reporter, because of the entry of the tRF-1001 containing duplex into an Ago2 complex followed by siRNA-like gene repression. This phenomenon was called sense-induced *trans*-silencing (SITS) and has been observed in HEK293 cells and in mouse embryonic cell lines [38]. It should be noted, however, that microRNAs do not use the slicer activity of Ago2 to repress target gene expression, and so there is no particular reason to expect Ago1, Ago3 or Ago4 associated tRFs (or miRNAs, for that matter) to be

deficient in gene silencing.

tRF-3s:

Besides association with Ago proteins, tRF-3s and tRF-5s have also been identified in PAR-CLIP data as being associated with components of P-bodies (our analysis of data in [49]). Since microRNAs often direct their targets to P bodies, which then leads to degradation of the target RNA, it would be interesting to determine whether tRF-3s and tRF-5s, promote the degradation of their target RNAs similar to miRNAs. This is an outstanding question that needs to be addressed experimentally.

In another report a tRF-3 corresponding to nucleotides 57–76 of the host cell tRNA^{Lys3} (tRF-3006 in tRFdb), was shown to bind to the primer-binding-site (PBS) in the genomic RNA of HIV, where it serves as the primer for reverse transcription during the viral life-cycle. tRF-3006 could bind to Ago-2 protein and silence an engineered luciferase reporter target [40]. Later it was shown that the cellular prevalence of this tRF was positively correlated with the replication of HIV [40].

Maute *et al.* demonstrated that tRF-3027, a 22 base tRF-3 from tRNA^{Gly-GCC} (the authors named it CU1276), is abundantly expressed in naïve, germinal center and memory B-cells in humans [26] and is physically associated with Argonaute proteins (Ago1-4) [26]. Furthermore, tRF-3027 was not expressed in transformed B-cell or lymphoma biopsies, suggesting that it is repressed during the malignant transformation of B cells [26]. Most interestingly, transient overexpression of tRF-3027, either by expressing the parental tRNA^{Gly-GCC} or by expressing it as a short-hairpin, repressed a target RNA: the mRNA of the single-stranded DNA binding protein RPA1 that is essential for DNA replication and repair [26]. tRF-3027 represses endogenous RPA1, which suppresses cell proliferation and modulates the DNA damage response [26].

Couvillion *et al.* reported that many 18-22 base long tRF-3s (rather than one

specific tRF-3) in *Tetrahymena* associate with Twi12 (an Ago/Piwi family protein) and have a role in ribosomal RNA maturation [44]. The tRF-3-bound Twi12 assembles with nuclear exonuclease Xrn2 to localize Xrn2 to the nucleus and stimulate its exonuclease activity [44]. Twi12 function depends on tRF-3 binding, and depletion of Twi12 or Xrn2 impairs ribosomal RNA processing [44].

tRF-5s:

By examining the small RNA repertoire in monocytes, Zhang *et al.* identified a new piRNA which is derived from tRNA-Glu^{CTC}(63). They showed that the tRNA-Glu-derived piRNA (tRF-5030c in tRFdb) associates with AGO-like proteins, PIWIL4 and PIWIL1. tRF-5030c-PIWIL4 complex recruits SETDB1, SUV39H1, and HP1beta to the CD1A promoter region to facilitate histone H3K9 methylation, and to repress CD1A expression in monocytes. However, during the maturation of monocytes to dendritic cells, the cytokine IL-4 represses RNA polymerase III and thus tRNA and tRF-5030c production, and this leads to derepression of CD1A, a marker of dendritic cells. It should be noted, however, that the decrease in tRNA-Glu level is much less than the decrease of tRF-5030c, so that IL-4 may have additional effects on the cleavage of the tRNA or the turnover of the tRF.

tiR(tRNA halves):

Several tiRs are specifically and abundantly expressed in a sex hormone-dependent manner in estrogen receptor positive breast and androgen receptor positive prostate cancer cell lines [50]. The sex-hormone induced tiRs were called SHOT-RNAs, and siRNA mediated knockdown of the 5tiRs, decreased cell proliferation, by decreasing the 5tiR, but not the mature tRNA. These tiRs were generated by angiogenin-mediated cleavage of the anticodon loop because (1) the 3tiR had the 5'OH signature and the 5tiR had a cyclic-phosphate that is typical of cleavage by RNase A family of nucleases (like angiogenin), and (2) knockdown of angiogenin greatly decreased the levels of the tiRs. Although the

mechanism by which these sex hormone-specific tiRs promote proliferation of breast and prostate cancer remains to be identified, they could be very specific biomarkers of estrogen or androgen action.

Stress such as oxidative stress, heat shock or UV irradiation induced tiRNAs have been shown to repress translation in a phospho-eIF2 α -independent manner [20]. Here again, knockdown of angiogenin decreased the levels of the tiRs produced by arsenite induced hypoxic stress. It should be noted however, that in all the papers where angiogenin knockdown decreased tiR levels, there was still significant residual tiR, either because there is an alternate RNase that generates the tiR, or because the knockdown of angiogenin was incomplete. Later on it was shown that tiRNAs inhibit protein synthesis and induce the phospho-eIF2 α -independent assembly of stress granules (SGs). The stress-induced tiRs, produced independent of the sex-hormone stimulated pathways, have been implicated by others in translation repression and as anti-apoptotic factors, as reviewed in [51].

In an exciting development, two recent studies demonstrate that 5tiR and tRF-5c are important for epigenetic inheritance [52, 53]. Sperm isolated from mice on a low protein diet contain an excess of specific tRNA fragments that are of the 5tiR and tRF-5c classes (tRF-Gly-CCC, -TCC, and -GCC; tRF-Lys-CTT; and tRF-His-GTG). Interestingly, tRF-Gly-GCC (tRF-5002c in tRFdb) suppresses nearly 70 genes associated with the endogenous retroelement MERVL in embryonic stem (ES) cells and embryos, thus suggesting that tRFs can control gene expression from specific regions of the genome [52]. In the second study, sperm isolated from mice fed on a high-fat diet produced offspring with impaired glucose tolerance. These spermatozoa also had an excess of 5tiRs (30 to 34 base long) from several tRNAs and they acquired the tiRs while passing through the epididymis. Injection of the

5tiRs alone into the early embryo down regulated many genes [53], many of which had promoter regions with complementarity to the 5tiRs that were found in the spermatozoa [53]. This indicates that 5tiRs may regulate gene expression through changing the epigenetic states of the genes and their promoters.

tRF-2s:

Recently, a set of tRF-2s derived from tRNA^{Glu}, tRNA^{Asp}, tRNA^{Gly}, and tRNA^{Tyr} were identified as tumor suppressors in breast cancer [28, 29]. These tRF-2s cover the anticodon stem and loop regions. [28]. This is the first tRF reported to be generated from the anti-codon loop and for which neither the 5' nor 3' ends extend to the start or the end of the mature or primary tRNA (**Fig. 1**). This set of tRF-2s act as tumor suppressors because they bind with YBX1, an RNA binding protein that normally binds to and stabilizes many oncogenic transcripts, and sequester YBX1 from those oncogenic transcripts [28]. tRF2s are induced by hypoxic conditions, which is prevalent in tumors. Supporting the idea of tRF-2s as tumor suppressors, they are also down-regulated in metastatic MDA-LM2 cells, allowing these cells to stabilize the relevant oncogenic transcripts.

Novel tRFs

A novel set of tRNA fragments derived from pre-tRNA has been implicated in neuronal loss [54]. CLP1, an enzyme, which is important for splicing of pre-tRNA intron, was shown to be mutated in neurodegenerative diseases [30]. Patients with hereditary CLP1 mutations develop severe motor-sensory defects, cortical dysgenesis and microcephaly. CLP1(R140H) mutant cannot interact with tRNA endonuclease complex (TSEN), and this leads to defects in splicing of the few tRNAs that have introns. The TSEN cut at the 3'intron-exon junction, but in the absence of CLP1, the 5'OH of the 3' exon is not phosphorylated and the subsequent splicing step blocked. This results in the accumulation of tRNA introns or of 3tiR-like fragments from the 3'exon in the cells of the patients. Surprisingly, the authors

did not see a deficiency of mature tRNAs in the patient fibroblasts. It is still unclear whether the accumulation of tRNA introns or the 3tiR-like fragments or defect in maturation of specific tRNAs in neurons is the cause of the disease, although transfection of the 3tiR-like fragments into neurons caused increased cytotoxicity.

Concluding remarks

The biogenesis of most of the tRF-5s and tRF-3s is unknown. It has been argued that some of the tRF-3s (e.g. tRF-3033, annotated as mmu-miR-1983 in mouse) are generated when a tRNA-Ile-TAT forms alternative secondary structures consisting of a stem-loop hairpin (like pre-miRNA) rather than the standard cloverleaf [23]. Schopman et al. have also argued that some tRNAs could fold into stem-loop hairpin structures rather than as tRNA cloverleaves to serve as a conventional pre-miRNA that generate miRNA-like molecules [24]. Hence it is important to explore the other possible structures of tRNA. Many of the bases in tRNAs are modified during tRNA maturation, and such modification affects processing by angiogenin (ANG) [55, 56]. Indeed, mutations in *Nsun2* or *Dnmt2*, which methylate tRNAs and inhibit cleavage by ANG [57-59], leads to excess production of tiRs that promote neurodegeneration. Thus, complete information on the bases that are modified in a tRNA will give additional insights into the biogenesis of tRFs. Although the analysis of short RNA data from cells and flies without Dicer suggest that tRFs are generated independent of Dicer, there are a few papers in the Literature that show a role of Dicer in producing selected tRFs [11, 26]. There is at least one paper that shows angiogenin, implicated till now in production of tiRs [50], can release tRF-like fragments in *in vitro* reactions on tRNAs [27]. Thus it is also possible that although not utilized for the global production of tRFs, enzymes such as Dicer or angiogenin are used in specific circumstances to produce specific tRFs.

Argonaute protein, specifically Ago-1, -3 and -4, associate with tRF-5s and -3s and

further analysis of Argonaute-associated small RNAs is expected to reveal additional tRF-Ago associations. Many questions remain unanswered, such as the mechanism of the Dicer-independent loading of tRFs onto Argonaute proteins. We expect that tRFs will likely appear in additional unexpected protein complexes. For instance, tiRs have been associated with YBX1 where they act as anti-oncogenic factors, and with cytochrome C, where they act as anti-apoptotic factors [28].

tRFs can also form duplexes with complementary RNAs, acting as an HIV-1 primer-binding site or with base-pairing with target RNAs in Ago. Thus a comprehensive list of RNAs that anneal to tRFs, identified by methods such as CLASH, may reveal new roles of tRFs. It is also conceivable that by occupying RNA binding sites on proteins tRFs will perturb the binding of other transcripts, with widespread consequences [28, 38]. Consistent with this, Dicer PAR-CLIP data shows that Dicer binds with many non-canonical Dicer transcripts including tRFs [60]. This observation is very intriguing considering that Dicer is dispensable for tRF biogenesis, and may suggest a role for tRFs in sequestering Dicer. Finally, of course, the nuclear location of tRF-5s and their association with Ago proteins raises the interesting possibility that some tRFs may have a role in regulation of chromatin.

ACKNOWLEDGMENTS

We thank all members of the Dutta Lab for helpful discussions. This work was supported by P01CA104106 and R01GM84465 to AD. Pankaj Kumar is supported by Department of Defense prostate cancer fellowship (W81XWH-13-1-0088).

Glossary

Cross-linking ligation and sequencing of hybrids (CLASH): is a method to identify RNA-RNA interaction. RNAs associated with a protein (e.g. Argonaute) are stabilized by cross-linking the RNA to the protein by UV radiation. The protein is immunoprecipitated and protein-associated RNAs trimmed so that only the region of the RNA around the protein interaction site survives. The trimmed RNAs present in a single protein molecule are ligated to each other by addition of a ligase to the immunoprecipitated complex. Finally the small RNAs are extracted, reverse transcribed, sequenced and mapped back to the genome to identify RNA-RNA chimeras arising from two RNAs interacting with each other in the immunoprecipitated protein.

Photoactivatable-Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation (PAR-CLIP): is a biochemical method to identify RNAs bound to specific proteins and the binding sites on RNA of the proteins. Cells are grown in the presence of photoactivatable nucleoside (e.g. 4-Thiouridine (4SU)) to incorporate the photoreactive ribonucleoside analogs into nascent RNA transcripts. Irradiation of the cells by UV light of 365 nm induces efficient crosslinking of photo-reactive nucleoside-labeled cellular RNAs to interacting RNA binding proteins. The protein (with cross-linked RNA) is immunoprecipitated, the RNA trimmed to save only the regions protected by the protein, the protein digested and the surviving RNA reverse-transcribed and subjected to high throughput sequencing. The reads are mapped back to the genome to identify the RNAs associated with the protein. The cross-linked 4SU is mis-read as a "C" during reverse transcription of the cross-linked RNAs and so the presence of T to C mutation at specific sites in multiple clones indicates the specific binding site on RNA of the protein.

FIGURE LEGEND

Figure 1: Classification of tRNA derived fragments (tRFs)

Fragments from tRNAs color coded to indicate area of origin. White circles indicate an intron present in some tRNAs (e.g. chr16.trna4-ProAGG) that is normally spliced out by TSEN and CLP1. tRF2 is a class proposed in this review.

Table1: Biological function of tRFs reported in the literature.

Organism	tRF type	Function	Reference	tRNA name	tRFdb
Human	tRF5	Translation inhibition	[61]	tRNA-Gln-CTG	tRF-5021
Human	tRF5	Represses target mRNAs in the cytoplasm and promotes human respiratory syncytial virus replication in cells.	[27, 62]	tRNA-Glu-CTC	tRF-5030
Human	tRF5	The tRF5/PIWIL4 complex recruits SETDB1, SUV39H1, and heterochromatin protein 1b to the CD1A promoter region and facilitates H3K9 methylation. tRF is not generated by ANG.	[63]	tRNA-Glu-CTC	tRF-5030
Mouse	tRF5	The tRNA fragment present in sperm that suppresses genes associated with the endogenous retroelement MERVL in ES cells and embryos.	[52]	tRF-Gly-GCC	tRF-5002
Haloferax volcanii	tRF5	This 26 base-long fragment from the 5' end of tRNA binds to the ribosome and reduces protein synthesis by interfering with peptidyl transferase activity.	[64]	tRNA-Val	NA
Human	tRF3	Binds to the primer-binding-site (PBS) in the genomic RNA of HIV, where it serves as the primer for reverse transcription. The cellular prevalence of this tRF was positively correlated with the proliferation of HIV.	[40]	tRNA-Lys3	tRF-3006
Human	tRF3	Represses endogenous RPA1 (role in DNA replication and repair), which leads to suppression of cell proliferation and modulates the DNA damage response. Its level is decreased during lymphomagenesis.	[26]	tRNA-Gly-GCC	tRF-3027
Tetrahymena	tRF3	Has a role in ribosomal RNA processing. The tRF-3-bound Twi12 assembles with nuclear exonuclease Xrn2 and helps to localize Xrn2 to the nucleus and stimulate its exonuclease activity.	[44]	Group of tRNA genes	NA
<i>E. coli</i>	tRF1	The trailer sequence of tRNA ^{leuZ} anneals to RhyB and RhyA short regulatory RNAs and changes gene expression.	[65]	tRNA ^{leuZ}	NA
Human	tRF1	The expression of tRF-1001 is positively correlated with cell proliferation in prostate cancer cell lines. tRF was also required for G2-M transition in the cell cycle.	[4]	tRNA-Ser-TGA	tRF-1001
Human	5tiR	Inhibits protein synthesis and promotes stress granule formation in a phospho eIF2 α -independent manner. Inhibits translation by displacing the eukaryotic initiation factor eIF4G/A from mRNAs.	[66]	tiR-Ala	NA
Mouse	5tiR & 3tiR	tiRs competitively bind to cytochrome c to protect cells from apoptosis due to osmotic stress.	[67]	Group of tRNAs	NA
Mouse	5tiR	Sperm 5tiR contribute to intergenerational inheritance. Alters expression profile and RNA modifications of many genes. The genes that showed alteration had homology to 5tiR at their promoter regions.	[53]	Group of tRNA	NA
Human	tRF2	Binds with YBX1, an RNA binding protein that normally binds to and stabilizes many oncogenic transcripts. Therefore act as tumor suppressor.	[28]	Group of tRNAs	NA
Human	3tiR	tiR fragments with 5'-OH created by a splicing defect are toxic to the cell and are increased in patients with homozygous mutations in CLP1.	[29, 30]	Group of tRNAs	NA

Foot notes: NA = The tRF is not present in tRFdb; tRNA name = Name of tRNA genes on which tRF maps; tRFdb = transfer RNA Fragments database

REFERENCES

- 1 Lee, R.C., *et al.* (1993) The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 75, 843-854
- 2 Wightman, B., *et al.* (1993) Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. *Cell* 75, 855-862
- 3 Kozomara, A. and Griffiths-Jones, S. (2011) miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res* 39, D152-157
- 4 Lee, Y.S., *et al.* (2009) A novel class of small RNAs: tRNA-derived RNA fragments (tRFs). *Genes Dev* 23, 2639-2649
- 5 Pederson, T. (2010) Regulatory RNAs derived from transfer RNA? *Rna* 16, 1865-1869
- 6 Tuck, A.C. and Tollervey, D. (2011) RNA in pieces. *Trends Genet* 27, 422-432
- 7 Phizicky, E.M. and Hopper, A.K. (2010) tRNA biology charges to the front. *Genes Dev* 24, 1832-1860
- 8 Frank, D.N. and Pace, N.R. (1998) Ribonuclease P: unity and diversity in a tRNA processing ribozyme. *Annu Rev Biochem* 67, 153-180
- 9 Maraia, R.J. and Lamichhane, T.N. (2011) 3' processing of eukaryotic precursor tRNAs. *Wiley Interdiscip Rev RNA* 2, 362-375
- 10 Xiong, Y. and Steitz, T.A. (2006) A story with a good ending: tRNA 3'-end maturation by CCA-adding enzymes. *Curr Opin Struct Biol* 16, 12-17
- 11 Cole, C., *et al.* (2009) Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs. *Rna* 15, 2147-2160
- 12 Couvillion, M.T., *et al.* (2010) A growth-essential *Tetrahymena* Piwi protein carries tRNA fragment cargo. *Genes Dev* 24, 2742-2747
- 13 Li, Z., *et al.* (2012) Extensive terminal and asymmetric processing of small RNAs from rRNAs, snoRNAs, snRNAs, and tRNAs. *Nucleic Acids Res*
- 14 Liao, J.Y., *et al.* (2010) Deep sequencing of human nuclear and cytoplasmic small RNAs reveals an unexpectedly complex subcellular distribution of miRNAs and tRNA 3' trailers. *PLoS One* 5, e10563
- 15 Kumar, P., *et al.* (2015) tRFdb: a database for transfer RNA fragments. *Nucleic Acids Res* 43, D141-145
- 16 Thompson, D.M. and Parker, R. (2009) Stressing out over tRNA cleavage. *Cell* 138, 215-219
- 17 Saikia, M., *et al.* (2012) Genome-wide Identification and Quantitative Analysis of Cleaved tRNA Fragments Induced by Cellular Stress. *J Biol Chem* 287, 42708-42725
- 18 Li, S. and Hu, G.F. (2012) Emerging role of angiogenin in stress response and cell survival under adverse conditions. *J Cell Physiol* 227, 2822-2826
- 19 Fu, H., *et al.* (2009) Stress induces tRNA cleavage by angiogenin in mammalian cells. *FEBS Lett* 583, 437-442
- 20 Yamasaki, S., *et al.* (2009) Angiogenin cleaves tRNA and promotes stress-induced translational repression. *J Cell Biol* 185, 35-42
- 21 Carthew, R.W. and Sontheimer, E.J. (2009) Origins and Mechanisms of miRNAs and siRNAs. *Cell* 136, 642-655

- 22 Kumar, P., *et al.* (2014) Meta-analysis of tRNA derived RNA fragments reveals that they are evolutionarily conserved and associate with AGO proteins to recognize specific RNA targets. *BMC Biol* 12, 78
- 23 Babiarz, J.E., *et al.* (2008) Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes Dev* 22, 2773-2785
- 24 Schopman, N.C., *et al.* (2010) A miRNA-tRNA mix-up: tRNA origin of proposed miRNA. *RNA Biol* 7, 573-576
- 25 Langenberger, D., *et al.* (2013) Dicer-Processed Small RNAs: Rules and Exceptions. *J Exp Zool B Mol Dev Evol* 320, 35-46
- 26 Maute, R.L., *et al.* (2013) tRNA-derived microRNA modulates proliferation and the DNA damage response and is down-regulated in B cell lymphoma. *Proc Natl Acad Sci U S A*
- 27 Wang, Q., *et al.* (2012) Identification and Functional Characterization of tRNA-derived RNA Fragments (tRFs) in Respiratory Syncytial Virus Infection. *Mol Ther*
- 28 Goodarzi, H., *et al.* (2015) Endogenous tRNA-Derived Fragments Suppress Breast Cancer Progression via YBX1 Displacement. *Cell* 161, 790-802
- 29 Schaffer, A.E., *et al.* (2014) CLP1 founder mutation links tRNA splicing and maturation to cerebellar development and neurodegeneration. *Cell* 157, 651-663
- 30 Karaca, E., *et al.* (2014) Human CLP1 mutations alter tRNA biogenesis, affecting both peripheral and central nervous system function. *Cell* 157, 636-650
- 31 Hopper, A.K. and Shaheen, H.H. (2008) A decade of surprises for tRNA nuclear-cytoplasmic dynamics. *Trends Cell Biol* 18, 98-104
- 32 Zofall, M. and Grewal, S.I. (2006) RNAi-mediated heterochromatin assembly in fission yeast. *Cold Spring Harb Symp Quant Biol* 71, 487-496
- 33 Emmerth, S., *et al.* (2010) Nuclear retention of fission yeast dicer is a prerequisite for RNAi-mediated heterochromatin assembly. *Dev Cell* 18, 102-113
- 34 Halic, M. and Moazed, D. (2009) Transposon silencing by piRNAs. *Cell* 138, 1058-1060
- 35 Chendrimada, T.P., *et al.* (2005) TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature* 436, 740-744
- 36 Noland, C.L., *et al.* (2011) siRNA repositioning for guide strand selection by human Dicer complexes. *Mol Cell* 43, 110-121
- 37 Gregory, R.I., *et al.* (2005) Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell* 123, 631-640
- 38 Haussecker, D., *et al.* (2010) Human tRNA-derived small RNAs in the global regulation of RNA silencing. *Rna* 16, 673-695
- 39 Kawaji, H., *et al.* (2008) Hidden layers of human small RNAs. *BMC Genomics* 9, 157
- 40 Yeung, M.L., *et al.* (2009) Pyrosequencing of small non-coding RNAs in HIV-1 infected cells: evidence for the processing of a viral-cellular double-stranded RNA hybrid. *Nucleic Acids Res* 37, 6575-6586
- 41 Burroughs, A.M., *et al.* (2011) Deep-sequencing of human Argonaute-associated small RNAs provides insight into miRNA sorting and reveals Argonaute association with RNA fragments of diverse origin. *RNA Biol* 8, 158-177

42 Reese, T.A., *et al.* (2010) Identification of novel microRNA-like molecules generated from herpesvirus and host tRNA transcripts. *J Virol* 84, 10344-10353

43 Heyer, R., *et al.* (2012) High throughput sequencing reveals a plethora of small RNAs including tRNA derived fragments in *Haloferax volcanii*. *RNA Biol* 9, 1011-1018

44 Couvillion, M.T., *et al.* (2012) A Tetrahymena Piwi Bound to Mature tRNA 3' Fragments Activates the Exonuclease Xrn2 for RNA Processing in the Nucleus. *Mol Cell* 48, 509-520

45 Chen, C.J., *et al.* (2011) Genome-wide discovery and analysis of microRNAs and other small RNAs from rice embryogenic callus. *RNA Biol* 8, 538-547

46 Hackenberg, M., *et al.* (2012) A Comprehensive Expression Profile of MicroRNAs and Other Classes of Non-Coding Small RNAs in Barley Under Phosphorous-Deficient and -Sufficient Conditions. *DNA Res*

47 Li, Z., *et al.* (2012) Extensive terminal and asymmetric processing of small RNAs from rRNAs, snoRNAs, snRNAs, and tRNAs. *Nucleic Acids Res* 40, 6787-6799

48 Gopal, S.K., *et al.* (2015) YBX1/YB-1 induces partial EMT and tumourigenicity through secretion of angiogenic factors into the extracellular microenvironment. *Oncotarget* 6, 13718-13730

49 Hafner, M., *et al.* (2010) Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* 141, 129-141

50 Honda, S., *et al.* (2015) Sex hormone-dependent tRNA halves enhance cell proliferation in breast and prostate cancers. *Proc Natl Acad Sci U S A* 112, E3816-3825

51 Saikia, M. and Hatzoglou, M. (2015) The Many Virtues of tRNA-derived Stress-induced RNAs (tiRNAs): Discovering Novel Mechanisms of Stress Response and Effect on Human Health. *J Biol Chem* 290, 29761-29768

52 Sharma, U., *et al.* (2015) Biogenesis and function of tRNA fragments during sperm maturation and fertilization in mammals. *Science*

53 Chen, Q., *et al.* (2015) Sperm tsRNAs contribute to intergenerational inheritance of an acquired metabolic disorder. *Science*

54 Hanada, T., *et al.* (2013) CLP1 links tRNA metabolism to progressive motor-neuron loss. *Nature* 495, 474-480

55 Thompson, D.M., *et al.* (2008) tRNA cleavage is a conserved response to oxidative stress in eukaryotes. *Rna* 14, 2095-2103

56 Schaefer, M., *et al.* (2010) RNA methylation by Dnmt2 protects transfer RNAs against stress-induced cleavage. *Genes Dev* 24, 1590-1595

57 Abbasi-Moheb, L., *et al.* (2012) Mutations in NSUN2 cause autosomal-recessive intellectual disability. *Am J Hum Genet* 90, 847-855

58 Khan, M.A., *et al.* (2012) Mutation in NSUN2, which encodes an RNA methyltransferase, causes autosomal-recessive intellectual disability. *Am J Hum Genet* 90, 856-863

59 Martinez, F.J., *et al.* (2012) Whole exome sequencing identifies a splicing mutation in NSUN2 as a cause of a Dubowitz-like syndrome. *J Med Genet* 49, 380-385

60 Rybak-Wolf, A., *et al.* (2014) A variety of dicer substrates in human and *C. elegans*. *Cell* 159, 1153-1167

- 61 Sobala, A. and Hutvagner, G. (2013) Small RNAs derived from the 5' end of tRNA can inhibit protein translation in human cells. *RNA Biol* 10, 553-563
- 62 Deng, J., *et al.* (2015) Respiratory Syncytial Virus Utilizes a tRNA Fragment to Suppress Antiviral Responses Through a Novel Targeting Mechanism. *Mol Ther* 23, 1622-1629
- 63 Zhang, X., *et al.* (2016) IL-4 Inhibits the Biogenesis of an Epigenetically Suppressible PIWI-Interacting RNA To Upregulate CD1a Molecules on Monocytes/Dendritic Cells. *J Immunol*
- 64 Gebetsberger, J., *et al.* (2012) tRNA-derived fragments target the ribosome and function as regulatory non-coding RNA in *Haloferax volcanii*. *Archaea* 2012, 260909
- 65 Lalaouna, D., *et al.* (2015) A 3' external transcribed spacer in a tRNA transcript acts as a sponge for small RNAs to prevent transcriptional noise. *Mol Cell* 58, 393-405
- 66 Ivanov, P., *et al.* (2011) Angiogenin-induced tRNA fragments inhibit translation initiation. *Mol Cell* 43, 613-623
- 67 Saikia, M., *et al.* (2014) Angiogenin-cleaved tRNA halves interact with cytochrome c, protecting cells from apoptosis during osmotic stress. *Mol Cell Biol* 34, 2450-2463