This is the peer reviewed version of the following article: Sobala Andrew and Hutvagner Gyorgy 2011, 'Transfer RNA-derived fragments: origins, processing, and functions', Wiley, vol. 2, no. November/December, pp. 853-862. which has been published in final form at http://dx.doi.org/10.1002/wrna.96 This article may be used for non-commercial purposes in accordance With Wiley Terms and Conditions for self-archiving'

tRNA derived fragments (tRFs): origins, processing and functions.

Andrew Sobala<sup>1</sup> and Gyorgy Hutvagner<sup>1,2,\*</sup>

1: Wellcome Trust Centre for Gene Regulation and Expression, Dundee University,

Dundee, UK

2: Centre for Health Technologies, University of Technology, Sydney, Australia

\*: Correspondence to be sent: <u>g.hutvagner@dundee.ac.uk</u>

Key words: tRNA, small RNA, tRF, tRNA fragment, Dicer, Argonaute

Running title: small RNAs from tRNAs

#### **Abstract**

Deep sequencing approaches have revealed multiple types of small RNAs with known and unknown functions. In this review we focus on recently identified group of small RNAs that are derived from tRNAs, tRNA fragments (tRFs). We review the mechanism of their processing and their functions in mammalian cells, and highlight points of possible cross-talk between tRFs and the canonical small RNA pathway characterized by siRNAs, miRNAs and piRNAs. We also propose a nomenclature that is based on their processing characteristics.

## Introduction

Gene expression studies have gone through a clear paradigm shift. The revolutionary advances in gene expression profiling techniques, such as the development of new generation sequencing, have revealed additional layers of gene regulatory mechanisms that are mediated by varieties of non-coding RNAs. By now it is clear that eukaryotic genomes produce a similar amount of non-coding RNAs as mRNA that can be translated into proteins. Non-coding RNAs (apart from those such as rRNAs, tRNA, snRNAs and snoRNAs) that are smaller than 200 nucleotides are categorized as small RNAs; however, this does not reflect similarities in their functions. Indeed, it defines a group of RNAs that have a wide range of both verified and suggested roles in gene regulation.

## The still-expanding realm of small regulatory RNAs

The first non-structural regulatory small RNA was discovered in 1993. It was *lin-4*, a microRNA (miRNA, originally temporal RNA) identified in *C. elegans* using a genetic screen <sup>1</sup>. In the last two decades immense numbers of miRNAs have been identified in almost all eukaryotic organisms (the miRBase depository, version 16.0, contains more than 17000 entries) and it is now obvious that they have pivotal roles in regulating gene expression in almost all biological processes. Studies examining the mechanism of RNA interference led to

the discovery of the second class of small regulatory RNAs, called small interfering RNAs (siRNAs) <sup>2</sup>. Originally discovered as a small RNA species from external origins (transgenes, viruses), it is now clear that plants and animals generate a variety of siRNAs from endogenous sources to regulate different steps of gene expression (reviewed in reference 3). Another class of small RNAs was discovered in *D. melanogaster* and mouse germlines that seems to be responsible for germline integrity and fertility (piRNAs) <sup>4 5 6</sup>. The functions and biogenesis of these three small regulatory RNAs are usually reviewed together. The reason for this is that it is very likely that the machines that process and confer function to these three types of small regulatory RNAs have similar evolutionary origins. For instance, conserved RNase III enzymes are involved in the processing of miRNAs and siRNAs in every eukaryotic organism in which they have been described. But the key connection between all these small RNAs is that they are associated with a member of the Argonaute (Ago) protein family (with only one known exception) <sup>7</sup>, which is required for their functionality and some cases for their processing.

The development of revolutionary deep sequencing technologies to identify small RNAs resulted in a rapid increase of catalogued miRNAs, siRNAs and piRNAs from many organisms. These approaches utilized the key structural characteristics of miRNAs and siRNAs, such as the terminal nucleotides having a 5′ monophosphate and 3′ hydroxyl as a result of RNase III cleavage. Also, the development of antibodies against the Argonaute protein family helped to sequence enriched small RNA fractions associated with these proteins and also contributed to the better understanding of the targeted transcriptome of these types of small regulatory RNAs.

## The siRNA/miRNA pathway

siRNAs and miRNAs can be produced from a large variety of cellular sources of

double stranded RNA. siRNAs can be generated from sources such as overlapping sense-antisense transcription, inverted genomic repeats of transposons and natural antisense transcripts (reviewed in reference 3). miRNAs are produced from hairpin structures, that can be formed from miRNA "genes" after processing by the RNase III enzyme Drosha <sup>8</sup>, or from intron lariats after debranching <sup>9 10 11</sup>. However, from this point all siRNAs and miRNAs are produced by similar mechanisms: the double stranded RNA is cleaved into 20-22 nt fragments by the RNase III enzyme Dicer <sup>12 13</sup>, of which one strand (the "guide strand") is incorporated into an effector Argonaute protein <sup>14 15</sup>.

Most organisms have multiple Argonaute proteins that can load siRNAs/miRNAs; there are four in humans (reviewed in reference 16). Argonaute proteins have two main roles: some have endonuclease ("slicer") activity and can catalyze the sequence-specific cleavage of complementary targets. Of the human proteins, Ago2 and Ago3 have conserved catalytic residues but only Ago2 has been shown to have slicer activity. Argonaute proteins without slicer activity act as sequence specific scaffolds, whereby they exert their effects by recruiting other proteins to targeted mRNA to cause responses such as translational inhibition, mRNA degradation and transcriptional silencing (reviewed in reference 17).

# Non-coding RNAs can be processed and enter the miRNA/siRNA pathway

Recent studies showed that abundant non-coding RNAs with known functions can be processed by enzymes that generate miRNAs/siRNAs and/or can be incorporated into the effector Argonaute protein. snoRNAs were the first example demonstrated to be associated with the miRNA pathway. So far, one snoRNA has been shown to produce a small RNA that functions as a canonical miRNA and regulates endogenous gene expression <sup>18</sup>. However, the relationship between snoRNAs and miRNAs may be more general since many other small RNAs derived from snoRNA have been identified, their production dependent on the miRNA

machinery, and are awaiting the identification of cellular targets <sup>19</sup>. Also, evolutionary connections between miRNAs and snoRNA have been proposed based on the suggestion that the structures of many pri-miRNAs also contain well-defined snoRNA elements <sup>20</sup>.

rRNAs also can be associated with the miRNA/siRNA pathways, mainly in stressed cells. In *N. crassa*, UV damage initiates the generation of small RNAs derived from rRNAs (qiRNAs) that are processed by the siRNA machinery and loaded into QDE-2, an Argonaute homologue <sup>21</sup>. These RNAs interfere with the ribosome and decrease translation as a response to UV damage. In *S. pombe* small RNAs generated from rRNAs are loaded to Ago1, without processing by the normal siRNA machinery, when the TRAMP/exosome complex is compromised <sup>22</sup>.

### **tRNAs**

A properly folded tRNA has three stem loops: the D loop, anticodon loop, and T loop (Figure 1). tRNAs go through several stages of maturation before becoming competent for translation (reviewed in reference 23). The order of these events is not well defined, and likely varies between organisms and tRNA isotypes. A primary transcript (pre-tRNA) is spliced if necessary, and the 5' and 3' ends are trimmed by the endonucleases RNase P and RNase Z respectively <sup>24 25</sup>. The 3' 'CCA' acceptor stem must be added as this is not encoded in eukaryotes <sup>26</sup>. Also, the tRNA undergoes many modification events to generate non-canonical bases. Modification is involved in ensuring the correct secondary and tertiary structures, and modifications have been variously described to have structural functions in both loosening secondary structure where extensive base-pairing would otherwise occur, in order to ensure cloverleaf formation, and in forming tertiary interactions. Overall, correctly modified tRNAs show lower conformational flexibility and higher thermal stability than their unmodified forms (reviewed in reference 27). Finally, tRNAs undergo aminoacylation and

are exported from the nucleus to become a competent substrate for translation <sup>28</sup>.

Whilst full-length tRNA has a well-defined role in protein translation, multiple groups have cloned and sequenced shorter tRNA fragments (reviewed in reference 29). The fragments identified so far can be classified into those that were generated by cleavage at the anticodon loop to produce longer RNA species of around 35 nt, and those that are around 20 nt (often corresponding to a cleavage in the D or T loops), which is conspicuously similar to the size of siRNAs and miRNAs. Of the former category, both 5' and 3' halves have been cloned. Of the latter category, three types have been cloned: those from the 5' end of tRNAs, those from the 3' end of mature tRNAs, and those from the 3' end of pre-tRNAs (Figure 1).

### tRNA halves

Accumulation of RNAs of 30-35 nucleotides that correspond to tRNA halves were first reported in *Tetrahymena thermophila* <sup>30</sup>, and subsequently shown to be a conserved response to oxidative stress in a wide variety of eukaryotes <sup>31</sup>. The nuclease responsible for tRNA half formation was discovered to be angiogenin <sup>32</sup>. Fascinatingly, a number of tRNAs (including tRNA<sup>Asp(GTC)</sup>, tRNA<sup>Val(AAC)</sup> and tRNA<sup>Gly(GCC)</sup>) can be methylated by Dnmt2 and this protects these tRNAs from cleavage in stress conditions <sup>33</sup>. This layer of specificity to which tRNAs are cleaved may cause different tRNAs to be cleaved in different situations, raising the possibility that the spectrum of cleaved tRNA fragments can contain information.

What is the biological effect of half-tRNA fragments? The Anderson group report that isolated 5' half-tRNA fragments can cause eIF2α-independent translational repression when transfected into cells <sup>34</sup>; however, these experiments were done with size-fractionated whole cell RNA and may contain species that are not tRNA-derived. They also reported that 5' half-tRNA fragments cause stress granule formation <sup>35</sup> and this result was obtained both with size-fractionated RNA and synthetic oligos. Half-tRNAs have also been shown to accumulate in

cytoplasmic granules in trypanosomes <sup>36</sup>.

When angiogenin cleaves tRNA, it is likely that the immediate *in vivo* product is a base paired tRNA-like molecule with a nick in the anticodon loop. Some interesting experiments have been done in plants that show that nicked tRNA (produced *in vitro* by incubation with low concentrations of RNase A) inhibits translation in wheat germ extract, but tRNA fragments (produced by denaturation after RNase A treatment but before adding to the in vitro reaction) do not <sup>37</sup>. The fact that the Anderson group did their experiments with transfected single strands, and that they only saw an effect with 5' halves and not 3' halves, suggests that half-tRNAs can also have a biological effect when separated from their cognate partner.

## **Small tRNA fragments (tRFs)**

A number of reports have been published in the last few years detailing the sequencing and analysis of tRNA-derived small RNAs of approximately 20 nt in size, and these are summarised in Table 1. The small tRNA-derived molecules reported lend themselves to a natural classification scheme, which is based on what part of tRNA or pre-tRNA from which they are derived (see also Figure 1).

The first class of short tRNA fragment (tRF) are derived from the 5' end of the molecule and are formed by a cleavage in the D loop. These have been reported in mammalian cells, plants and fission yeast <sup>22 38 39 40 41</sup>. The second class of tRFs are those derived from the 3' end of a mature tRNA by cleavage in the T loop <sup>38 42 43 44 45</sup>. As mature tRNA end in the acceptor trinucleotide CCA, these RNAs also characteristically end CCA. The final class of tRFs are those derived from the 3' end of a pre-tRNA <sup>38 44 46 47</sup>. In most cases these begin directly after the 3' terminus of the mature tRNA, and end in a series of U residues that are residual from RNA polymerase III transcription run-off.

The nomenclature used to refer to different types of tRF varies in the literature. We propose, and will use in this review, a scheme based on the part of the molecule from which they are derived: those from the 5' end of a tRNA are 5' tRFs, those from the 3' end of a mature tRNA are 3' CCA tRFs (as they end -CCA), and those from the 3' end of a pre-tRNA are 3' U tRFs (as they end in a series of uridine residues) (Figure 1). This scheme is sufficiently descriptive to be expanded if necessary. 3' CCA tRFs were previously called Type I and 3' U tRFs were previously called Type II<sup>44</sup>.

Table 1. Instances of tRFs reported in the literature

Cell	Length	tRF type	Nuclease
line/organism			
HepG2 (liver	22 nt	3' CCA	Not
carcinoma)			investigated
S. pombe	23 nt	5'	Dicer
			independent
Mouse ES	21 nt	3' U	Dicer
HIV-1 infected	18 nt	3' CCA	Dicer
MT4 cells			
LNCaP and C4-2	18 – 22 nt	5'	Not
(prostate		3' CCA	investigated
carcinoma)		3' U	RNase Z
HeLa	19 nt	5'	Dicer
Phosphate	19 nt	5'	Not
starved A.			investigated
thaliana roots			
5-8F	19 nt	3' U	Not
	line/organism  HepG2 (liver carcinoma)  S. pombe  Mouse ES  HIV-1 infected  MT4 cells  LNCaP and C4-2 (prostate carcinoma)  HeLa  Phosphate  starved A.  thaliana roots	line/organism  HepG2 (liver 22 nt carcinoma)  S. pombe 23 nt  Mouse ES 21 nt  HIV-1 infected 18 nt  MT4 cells  LNCaP and C4-2 18 – 22 nt  (prostate carcinoma)  HeLa 19 nt  Phosphate 19 nt  starved A.  thaliana roots	HepG2 (liver   22 nt   3' CCA   carcinoma)

	(nasopharyngeal			investigated
	carcinoma)			
44	HEK293 and	20 – 22 nt	3' CCA	Dicer
	HCT116		3' U	RNase Z
45	Tetrahymena	~ 23 nt	3' CCA	Not
				investigated

In some studies the nuclease involved in producing tRFs was investigated. The production of 5' tRFs is reported as Dicer dependent in mammals, although not in yeast; differences in length support the conclusion that different biogenesis mechanisms are at work (in mammals, the tRFs are 19 nt long, whereas in yeast the one that has been reported is 23 nt long) <sup>22</sup> <sup>39</sup>. 3' CCA tRFs have also been reported to be generated by Dicer <sup>43</sup> <sup>44</sup>. In *Tetrahymena*, although the nuclease identity was not proven, the 3' CCA tRFs have 5' monophosphates that are characteristic of Dicer processing, and this contrasts with the 5' hydroxyls found on longer half-tRNAs produced by starvation-induced cleavage <sup>45</sup>. 3' U tRFs are normally Dicer-independent and produced by RNase Z cleavage of a pre-tRNA transcript <sup>38</sup> <sup>44</sup>, although one 3' U tRF has been reported that is produced when RNase Z does not act on the transcript, but instead the pre-tRNA is predicted to fold into a bulged hairpin and a 3' U tRF is produced by a Dicer dependent mechanism <sup>46</sup>.

### The tRNA source of tRFs

tRNAs could be diverted from their normal processing pathway at a number of points and undergo cleavage to form a tRF (Figure 2). Clearly 3' U tRFs are formed from the pre-tRNA as they contain the polyuridine tract left over from transcription, but it has been pointed out that as 3' U tRFs (and, indeed, all described tRFs) are almost exclusively

cytoplasmic they may be formed by a cytoplasmic pool of RNase Z<sup>38</sup> <sup>44</sup> <sup>48</sup>. Otherwise, they may undergo rapid nuclear export after formation.

On the other hand, 3' CCA tRFs must be generated from tRNA after CCA addition and 5' tRFs may be generated at any stage after 5' end formation by RNase P. In one case, 5' tRFs were identified to be produced after the base-pairing of tRNA<sup>Lys</sup> with the HIV-1 primer binding site (PBS), raising the possibility that, at least in certain cases, tRFs may be produced when a tRNA basepairs with another RNA molecule <sup>43</sup>. However, this is probably an exceptional case caused by the involvement of tRNA<sup>Lys</sup> in making double stranded RNA in HIV replication; it is likely that in most cases tRFs are produced by the cleavage of a single tRNA molecule.

As one tRF has been reported to be generated by Dicer from a potentially misfolded pre-tRNA substrate, where the pre-tRNA effectively folds into a hairpin <sup>46</sup>, it is intriguing to speculate if this is a general mechanism. In this case, incorrectly formed tRNAs would be cleaved into tRFs. If this process is efficient, it would effectively destroy the aberrant tRNAs and even if inefficient, would still produce tRFs which may act as intracellular signaling molecules. These incorrectly formed tRNAs could be pre-tRNAs that undergo premature nuclear export, or hypomodified tRNAs. In both cases, they would be unlikely to fold into a correct cloverleaf structure.

# **Biological function of tRFs**

Do tRFs have *bona fide* biological functions, or are they merely the products of tRNA degradation or nuclease off-target effects? If they do have functions, how diverse are these functions given their diverse processing mechanisms? Several lines of evidence point towards regulated production, suggesting that they may be functional RNA species. First, the sequencing abundance of different types of tRF does not correlate with the number of parent

tRNA gene copies <sup>39 40 42</sup> with the exception of those found in *Tetrahymena* <sup>45</sup>. Second, they are all produced by cleavage at specific bases, and third, whilst tRFs corresponding to the 5' and 3' ends of tRNA have been reported, those corresponding to the middle (incorporating the anticodon loop) have not.

In one case, a tRF has been shown to target RNAs similarly to siRNAs. HIV-PBS tRF associates with Ago2, and causes RNA cleavage of complementary sequences <sup>43</sup>. As the HIV RNA genome contains a PBS sequence, this is an intriguing example of a mammalian cell employing RNAi and tRFs in virus defence.

Other tRFs have also been shown to be associated with Argonautes therefore they have a potential to function as siRNAs or miRNAs, although endogenous targets have not been demonstrated yet. Haussecker and colleagues<sup>44</sup> investigated the propensity of 3' tRFs (both types) to associate with Argonaute proteins and have an effect on luciferase reporter genes; their results were surprising. They found that both types of 3' tRF associated with Argonaute proteins, but often more effectively with Ago3 and Ago4 than Ago1 or Ago2. They found that 3' CCA tRFs had a moderate effect on reporter transgene silencing, but 3' U tRFs did not. However, upon cotransfection of a small RNA complementary to a 3' U tRF, they found that the tRF preferentially associated with Ago2 and silenced a reporter transgene, a phenomenon they termed Sense-Induced Transgene Silencing (SITS). This is in stark contrast with results normally obtained in the miRNA field where sequences complementary to miRNAs relieve repression. Haussecker and colleagues suggest that the double stranded perfect RNA helix produced when a sense strand is present causes the more efficient loading into Ago2, which is consistent with *in vitro* studies showing that Ago2's slicer activity causes more efficient loading of perfect duplexes <sup>49</sup>.

Serendipitously, Lee and colleagues<sup>38</sup> and Haussecker and colleagues<sup>44</sup> chose to characterize the same tRF, called cand45 by Haussecker et al. and tRF-1001 by Lee et al.

Whereas Lee et al. did not look at Argonaute association of tRF-1001, they did find that its knockdown by siRNA decreased cell proliferation and re-addition increased proliferation. Hence Lee et al. found a function for 3' U tRFs in a situation where Haussecker et al. showed them to be primarily associated with Ago3 and Ago4. Could tRF association with Ago3/4 be necessary for cell proliferation? In these experiments, the level of mature and pre-tRNA does not change, suggesting a role of tRFs independent of modulating the level of their parental tRNA.

It has also been shown that 3' CCA tRFs associate with the PIWI protein Twi12 in *Tetrahymena* <sup>45</sup>, which does not contain the residues necessary for RNA cleavage. The function of this is unclear, as is whether tRF fragments associate with PIWI proteins in mammalian cells; it is possible that some proteins, such as Twi12 and human Ago3 and Ago4, have evolved to mop up tRF fragments when they are not needed. Alternatively their association with these types of small RNA may be a clue to a function which, in the case of mammalian Ago1, 3 and 4 for instance, is not clear yet.

In contrast to 3' tRFs, 5' tRFs show only inefficient association with Argonautes <sup>39</sup> and do not exhibit detectable silencing effects (AS and GH, unpublished observations). Their functions remain to be fully elucidated.

# Regulation of tRF production

To date, no systematic quantitative analysis of tRF abundance across different cell types has been done. Most of the studies on tRFs have been done on cancer-derived cell lines or highly proliferative ES cells. Interestingly, various researchers have noticed that the abundance of tRFs is correlated to cell proliferation; increasing media serum concentration elevates the amount of 3' and 5' tRFs <sup>38 44</sup> and overexpressing the tRNA transcription factor Brf1, a situation known to lead to increased cell proliferation <sup>50</sup>, also increases tRF levels <sup>44</sup>.

### **Future directions**

In the almost 20 years since its inception, the small RNA field has concentrated on the siRNA, miRNA and piRNA small RNA classes. In recent years the unbiased sequencing of small RNA fractions has further widened the horizon of small RNA research revealing the existence of many novel small RNA species such as 5′-capped promoter-associated small RNAs (PASRs) <sup>51</sup>, transcription initiation RNAs (tiRNAs) <sup>52</sup>, transcription start site-associated RNAs (TSSa RNAs) <sup>53</sup>, splice site RNAs (spliRNAs) <sup>54</sup>, and short RNA products from rRNA <sup>22</sup>, snoRNA <sup>18</sup> and tRNA as covered in this review. It is very unlikely that the databases of verified small non coding RNAs are complete. Bioinformatic studies still predict many more miRNAs and other small RNAs <sup>55</sup> <sup>56</sup> that may so far have avoided identification due to either not being present in cells that have been subjected to deep sequencing or not been sequenced due to technical issues <sup>57</sup>. However, the most important task in the small regulatory RNA field is to go beyond the cataloguing of this huge arsenal of small RNAs and to associate them with functions, as is already being carried out in the case of siRNAs, miRNAs, and piRNAs.

tRFs could have a variety of functions. Where they are incorporated into Argonaute proteins, they would be expected to inhibit gene expression of specific mRNAs by either translation inhibition and mRNA destabilization, or by endonucleolytic cleavage. If this is the case, targets for these RNAs can be found by the use of bioinformatic or systemic approaches such as transcription profiling and proteomics. Alternatively, could an Argonaute protein loaded with a tRF be directed to interact with mature tRNAs directly? If so, it could conceivably affect tRNA maturation or result in Argonaute proteins being brought to the ribosome in a general, non-messenger specific way. This would be extremely interesting since Ago2 originally was described as a co-factor for eIF2 and its activity was shown to stimulate ternary complex formation <sup>58 59 60</sup>. In accordance with this full length mitochondrial

tRNA<sup>Met</sup> has been described to associate with Ago2 <sup>61</sup>.

Where tRFs are not Argonaute-associated they may have other functions outside of the canonical small RNA pathways that still need to be elucidated. tRNAs have already been shown to have diverse functions outside translation; for example, they act anti-apoptotically by directly binding cytochrome c <sup>62</sup>. Possibly tRFs can act as RNA aptamers for as-yet uncharacterized ligands. A recent review insightfully points out that the very 5' and 3' ends of tRNAs that are represented as tRFs specify aminoacylation specificity in the full length molecule, and presumably coevolved with tRNA aminoacyl synthetases and the anticodon sequence. It also notes that the 5' and 3' CCA tRFs are to some degree complementary, and may form double stranded RNA *in vivo* if the local concentration is high enough <sup>29</sup>.

Accumulation of small tRNAs and rRNAs and their association with the siRNA/miRNA machinery seems to be characteristic of highly proliferative cells and cells with aberrant cell division characteristics. Indeed, in *S. pombe* an active mechanism was identified in which rRNAs and tRFs were excluded from entering the siRNA pathway by the TRAMP complex <sup>22</sup>. The impairment of this complex resulted in replacing centromere associated short RNAs with tRFs and short rRNAs in Ago1 that eventually led to cell death. Does this mechanism exist in higher eukaryotes? If yes, how do cancer cells accumulate this type of RNA and overcome their potential cell toxicity? One possibility is that Argonautes other than Ago2 mop up these RNA fragments providing a buffering capacity that help to avoid cell death in organisms that encode multiple Argonaute proteins.

Accurately profiling tRFs could be technically challenging. It is known that current deep sequencing methods show some degree of sequence bias depending on the method used for library generation <sup>63</sup>. Additionally, any tRFs derived from mature tRNAs would be expected to be base modified. Some tRNA base modifications, such as N¹-methyl-A, significantly inhibit reverse transcriptases <sup>64</sup> and this would have to be addressed in any

sequencing technology used for quantification. Mapping modified nucleotides in tRFs is also essential to understand their function, as any base-pairing function will be significantly altered by the presence of altered nucleotides. It has also been shown that modifications of the terminal nucleotides of small RNAs could be characteristic of what Argonaute they are loaded onto <sup>65</sup> <sup>66</sup>, so where tRFs are derived from mature tRNA sequence the different preferences of tRFs for different Argonautes may be a consequence of the possibly diverse modifications of the different types of tRFs.

Whatever the answer to these questions, it seems possible that tRFs could have very different roles to those small RNAs discovered so far. Defining their effects on gene regulation and other cellular processes will no doubt lead to a better understanding of the highly diverse, still expanding universe of small RNAs.

# Acknowledgements

This work was supported by the Wellcome Trust Career Development Program, the Wellcome Trust GRE Centre, and the European FP6 consortium SIROCCO. A.S is funded by a BBSRC Doctoral Training Fund.

# **Figure Legends**

Figure 1. Proposed nomenclature of small (less than 30nt) tRNA fragments (tRFs). 3'U tRFs are processed from pre-tRNAs and consist of the sequence between the RNAse Z cleavage site and the PolIII run-off poly(U) tract. Mature tRNAs can generate two types of tRFs: one is processed from the 5'end (5'tRFs) and one from the 3'end that contains the added CCA sequence (3'CCA tRFs).

Figure 2. Possible processing and maturation of tRFs. 3´U tRFs could be processed from pre-tRNAs by RNase Z and promptly exported into the cytoplasm where they incorporate into Argonaute complexes. Alternatively, pre-tRNAs may go through premature rapid nuclear export and be processed by a cytoplasmic RNase. tRNA processed by RNAse P and Z with the 3' terminal CCA added may be prematurely exported into the cytoplasm and a lack of base modifications may cause them to form a Dicer accessible structure. Alternatively, mature aminoacylated tRNAs may enter the small RNA pathway via Dicer processing. Dicer could produce two types of tRFs from these tRNAs: 5'tRFs and 3'CCA tRFs. 3' CCA tRFs can incorporate into Argonaute proteins. Open triangle: RNAse Z cleavage site, filled triangle: RNase P cleavage site. Red arrows denote possible premature nuclear export.

## References

Lee, R C, Feinbaum, R L, and Ambros, V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell*, 75, 5 (Dec 3, 1993), 843-54.

Hamilton, A J and Baulcombe, D C. A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science (New York, NY)*, 286, 5441 (Oct 29, 1999), 950-2.

Ghildiyal, Megha and Zamore, Phillip D. Small silencing RNAs: an expanding universe. *Nature reviews Genetics*, 10, 2 (Feb 1, 2009), 94-108. 10.1038/nrg2504.

Aravin, Alexei A, Naumova, N M, Tulin, A V, Vagin, Vasily V, Rozovsky, Y M,

and Gvozdev, Vladimir. Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the D. melanogaster germline. *Current biology : CB*, 11, 13 (Jul 10, 2001), 1017-27.

Aravin, Alexei A, Gaidatzis, Dimos, Pfeffer, Sébastien et al. A novel class of small RNAs bind to MILI protein in mouse testes. *Nature*, 442, 7099 (Jul 13, 2006), 203-7. 10.1038/nature04916.

Girard, Angelique, Sachidanandam, Ravi, Hannon, Gregory J, and Carmell, Michelle A. A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature*, 442, 7099 (Jul 13, 2006), 199-202. 10.1038/nature04917.

Eiring, Anna M, Harb, Jason G, Neviani, Paolo et al. miR-328 functions as an RNA decoy to modulate hnRNP E2 regulation of mRNA translation in leukemic blasts. *Cell*, 140, 5 (Mar 5, 2010), 652-65. 10.1016/j.cell.2010.01.007.

Lee, Yoontae, Ahn, Chiyoung, Han, Jinju et al. The nuclear RNase III Drosha initiates microRNA processing. *Nature*, 425, 6956 (Sep 25, 2003), 415-9. 10.1038/nature01957.

Ruby, J Graham, Jan, Calvin H, and Bartel, David P. Intronic microRNA precursors that bypass Drosha processing. *Nature*, 448, 7149 (Jul 5, 2007), 83-6. 10.1038/nature05983.

Okamura, Katsutomo, Hagen, Joshua W, Duan, Hong, Tyler, David M, and Lai, 0 Eric C. The mirtron pathway generates microRNA-class regulatory RNAs in Drosophila. *Cell*, 130, 1 (Jul 13, 2007), 89-100. 10.1016/j.cell.2007.06.028.

Berezikov, Eugene, Chung, Wei-Jen, Willis, Jason, Cuppen, Edwin, and Lai, Eric 1 C. Mammalian mirtron genes. *Molecular Cell*, 28, 2 (Oct 26, 2007), 328-36. 10.1016/j.molcel.2007.09.028.

Hutvagner, Gyorgy, McLachlan, J, Pasquinelli, Amy E, Bálint, E, Tuschl, Thomas,
and Zamore, Phillip D. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science (New York, NY)*, 293, 5531 (Aug 3, 2001), 834-8. 10.1126/science.1062961.

Grishok, A, Pasquinelli, Amy E, Conte, D et al. Genes and mechanisms related to 3 RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. *Cell*, 106, 1 (Jul 13, 2001), 23-34.

Martinez, Javier, Patkaniowska, Agnieszka, Urlaub, Henning, Lührmann, 4 Reinhard, and Tuschl, Thomas. Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell*, 110, 5 (Sep 6, 2002), 563-74.

Schwarz, Dianne S, Hutvágner, György, Du, Tingting, Xu, Zuoshang, Aronin,

Neil, and Zamore, Phillip D. Asymmetry in the assembly of the RNAi enzyme complex.

Cell, 115, 2 (Oct 17, 2003), 199-208.

Hutvagner, Gyorgy and Simard, Martin J. Argonaute proteins: key players in RNA 6 silencing. *Nature reviews Molecular cell biology*, 9, 1 (Jan 1, 2008), 22-32. 10.1038/nrm2321.

Huntzinger, Eric and Izaurralde, Elisa. Gene silencing by microRNAs: 7 contributions of translational repression and mRNA decay. *Nature reviews Genetics*, 12, 2 (Feb 1, 2011), 99-110. 10.1038/nrg2936.

Ender, Christine, Krek, Azra, Friedländer, Marc R et al. A Human snoRNA with 8 MicroRNA-Like Functions. *Molecular Cell*, 32, 4 (Nov 21, 2008), 519-28. 10.1016/j.molcel.2008.10.017.

Taft, Ryan J, Glazov, Evgeny A, Lassmann, Timo, Hayashizaki, Yoshihide, 9 Carninci, Piero, and Mattick, John S. Small RNAs derived from snoRNAs. *RNA* (*New*  York, NY), 15, 7 (Jul 1, 2009), 1233-40. 10.1261/rna.1528909.

Scott, Michelle S, Avolio, Fabio, Ono, Motoharu, Lamond, Angus I, and Barton,

O Geoffrey J. Human miRNA precursors with box H/ACA snoRNA features. *PLoS*computational biology, 5, 9 (Sep 1, 2009), e1000507. 10.1371/journal.pcbi.1000507.

Lee, Heng-Chi, Chang, Shwu-Shin, Choudhary, Swati, Aalto, Antti P, Maiti, Mekhala, Bamford, Dennis H, and Liu, Yi. qiRNA is a new type of small interfering RNA induced by DNA damage. *Nature*, 459, 7244 (May 14, 2009), 274-7. 10.1038/nature08041.

Bühler, Marc, Spies, Noah, Bartel, David P, and Moazed, Danesh. TRAMP2 mediated RNA surveillance prevents spurious entry of RNAs into the
Schizosaccharomyces pombe siRNA pathway. *Nature structural & molecular biology*, 15,
10 (Oct 1, 2008), 1015-23. 10.1038/nsmb.1481.

Li, Hong. Complexes of tRNA and maturation enzymes: shaping up for 3 translation. *Current opinion in structural biology*, 17, 3 (Jun 1, 2007), 293-301. 10.1016/j.sbi.2007.05.002.

Robertson, H D, Altman, S, and Smith, J D. Purification and properties of a 4 specific Escherichia coli ribonuclease which cleaves a tyrosine transfer ribonucleic acid presursor. *The Journal of biological chemistry*, 247, 16 (Aug 25, 1972), 5243-51.

Schiffer, Steffen, Rösch, Sylvia, and Marchfelder, Anita. Assigning a function to a 5 conserved group of proteins: the tRNA 3'-processing enzymes. *The EMBO journal*, 21, 11 (Jun 3, 2002), 2769-77. 10.1093/emboj/21.11.2769.

Aebi, M, Kirchner, G, Chen, J Y, Vijayraghavan, U, Jacobson, A, Martin, N C, and Abelson, J. Isolation of a temperature-sensitive mutant with an altered tRNA nucleotidyltransferase and cloning of the gene encoding tRNA nucleotidyltransferase in

the yeast Saccharomyces cerevisiae. *The Journal of biological chemistry*, 265, 27 (Sep 25, 1990), 16216-20.

Helm, Mark. Post-transcriptional nucleotide modification and alternative folding of RNA. *Nucleic acids research*, 34, 2 (Jan 1, 2006), 721-33. 10.1093/nar/gkj471.

Lund, E and Dahlberg, J E. Proofreading and aminoacylation of tRNAs before 8 export from the nucleus. *Science*, 282, 5396 (Dec 11, 1998), 2082-5.

Pederson, Thoru. Regulatory RNAs derived from transfer RNA? *RNA* (*New York*, 9 *NY*), 16, 10 (Oct 1, 2010), 1865-9. 10.1261/rna.2266510.

Lee, Suzanne R and Collins, Kathleen. Starvation-induced cleavage of the tRNA on anticodon loop in Tetrahymena thermophila. *The Journal of biological chemistry*, 280, 52 (Dec 30, 2005), 42744-9. 10.1074/jbc.M510356200.

Thompson, Debrah M, Lu, Cheng, Green, Pamela J, and Parker, Roy. tRNA 1 cleavage is a conserved response to oxidative stress in eukaryotes. *RNA (New York, NY)*, 14, 10 (Oct 1, 2008), 2095-103. 10.1261/rna.1232808.

Fu, Hanjiang, Feng, Junjun, Liu, Qin et al. Stress induces tRNA cleavage by 2 angiogenin in mammalian cells. *FEBS letters*, 583, 2 (Jan 22, 2009), 437-42. 10.1016/j.febslet.2008.12.043.

Schaefer, Matthias, Pollex, Tim, Hanna, Katharina, Tuorto, Francesca, 3 Meusburger, Madeleine, Helm, Mark, and Lyko, Frank. RNA methylation by Dnmt2 protects transfer RNAs against stress-induced cleavage. *Genes & development*, 24, 15 (Aug 1, 2010), 1590-5. 10.1101/gad.586710.

Yamasaki, Satoshi, Ivanov, Pavel, Hu, Guo-Fu, and Anderson, Paul. Angiogenin 4 cleaves tRNA and promotes stress-induced translational repression. *The Journal of Cell Biology*, 185, 1 (Apr 6, 2009), 35-42. 10.1083/jcb.200811106.

Emara, Mohamed M, Ivanov, Pavel, Hickman, Tyler et al. Angiogenin-induced 5 tiRNAs promote stress-induced stress granule assembly. *The Journal of biological chemistry* (Feb 3, 2010). 10.1074/jbc.M109.077560.

Garcia-Silva, Maria Rosa, Frugier, Magali, Tosar, Juan Pablo et al. A population of tRNA-derived small RNAs is actively produced in Trypanosoma cruzi and recruited to specific cytoplasmic granules. *Molecular and biochemical parasitology*, 171, 2 (Jun 1, 2010), 64-73. 10.1016/j.molbiopara.2010.02.003.

Zhang, Shoudong, Sun, Li, and Kragler, Friedrich. The phloem-delivered RNA pool contains small noncoding RNAs and interferes with translation. *Plant physiology*, 150, 1 (May 1, 2009), 378-87. 10.1104/pp.108.134767.

Lee, Yong Sun, Shibata, Yoshiyuki, Malhotra, Ankit, and Dutta, Anindya. A novel 8 class of small RNAs: tRNA-derived RNA fragments (tRFs). *Genes & development*, 23, 22 (Nov 15, 2009), 2639-49. 10.1101/gad.1837609.

Cole, Christian, Sobala, Andrew, Lu, Cheng et al. Filtering of deep sequencing 9 data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs. *RNA* (*New York, NY*), 15, 12 (Dec 1, 2009), 2147-60. 10.1261/rna.1738409.

Hsieh, Li-Ching, Lin, Shu-I, Shih, Arthur Chun-Chieh et al. Uncovering small 0 RNA-mediated responses to phosphate deficiency in Arabidopsis by deep sequencing. *Plant physiology*, 151, 4 (Dec 1, 2009), 2120-32. 10.1104/pp.109.147280.

Burroughs, Alexander Maxwell, Ando, Yoshinari, Hoon, Michiel Laurens de,

1 Tomaru, Yasuhiro, Suzuki, Harukazu, Hayashizaki, Yoshihide, and Daub, Carsten Olivier.

Deep-sequencing of human Argonaute-associated small RNAs provides insight into miRNA sorting and reveals Argonaute association with RNA fragments of diverse origin.

RNA biology, 8, 1 (Jan 1, 2011).

Kawaji, Hideya, Nakamura, Mari, Takahashi, Yukari et al. Hidden layers of 2 human small RNAs. *BMC genomics*, 9 (Jan 1, 2008), 157. 10.1186/1471-2164-9-157.

Yeung, Man Lung, Bennasser, Yamina, Watashi, Koichi, Le, Shu-Yun, Houzet, 3 Laurent, and Jeang, Kuan-Teh. 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 research*, 37, 19 (Oct 1, 2009), 6575-86. 10.1093/nar/gkp707.

Haussecker, Dirk, Huang, Yong, Lau, Ashley, Parameswaran, Poornima, Fire,
4 Andrew Z, and Kay, Mark A. Human tRNA-derived small RNAs in the global regulation of RNA silencing. *RNA* (*New York*, *NY*) (Feb 24, 2010). 10.1261/rna.2000810.

Couvillion, Mary T, Sachidanandam, Ravi, and Collins, Kathleen. A growth-5 essential Tetrahymena Piwi protein carries tRNA fragment cargo. *Genes & Development*, 24, 24 (Dec 15, 2010), 2742-7. 10.1101/gad.1996210.

Babiarz, Joshua E, Ruby, J Graham, Wang, Yangming, Bartel, David P, and Blelloch, Robert. Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes & development*, 22, 20 (Oct 15, 2008), 2773-85. 10.1101/gad.1705308.

Liao, Jian-You, Ma, Li-Ming, Guo, Yan-Hua et al. Deep sequencing of human nuclear and cytoplasmic small RNAs reveals an unexpectedly complex subcellular distribution of miRNAs and tRNA 3' trailers. *PloS one*, 5, 5 (Jan 1, 2010), e10563. 10.1371/journal.pone.0010563.

Elbarbary, Reyad A, Takaku, Hiroaki, Uchiumi, Naoto et al. Modulation of gene 8 expression by human cytosolic tRNase Z(L) through 5'-half-tRNA. *PloS one*, 4, 6 (Jan 1, 2009), e5908. 10.1371/journal.pone.0005908.

Yoda, Mayuko, Kawamata, Tomoko, Paroo, Zain, Ye, Xuecheng, Iwasaki,

9 Shintaro, Liu, Qinghua, and Tomari, Yukihide. ATP-dependent human RISC assembly pathways. *Nature structural & molecular biology*, 17, 1 (Jan 1, 2010), 17-23. 10.1038/nsmb.1733.

Marshall, Lynne, Kenneth, Niall S, and White, Robert J. Elevated tRNA(iMet) 0 synthesis can drive cell proliferation and oncogenic transformation. *Cell*, 133, 1 (Apr 4, 2008), 78-89. 10.1016/j.cell.2008.02.035.

Affymetrix ENCODE Transcriptome Project and Cold Spring Harbor Laboratory

1 ENCODE Transcriptome Project. Post-transcriptional processing generates a diversity of
5'-modified long and short RNAs. *Nature*, 457, 7232 (Feb 19, 2009), 1028-32.
10.1038/nature07759.

Taft, Ryan J, Glazov, Evgeny A, Cloonan, Nicole et al. Tiny RNAs associated with transcription start sites in animals. *Nature genetics*, 41, 5 (May 1, 2009), 572-8. 10.1038/ng.312.

Seila, Amy C, Calabrese, J Mauro, Levine, Stuart S et al. Divergent transcription 3 from active promoters. *Science (New York, NY)*, 322, 5909 (Dec 19, 2008), 1849-51. 10.1126/science.1162253.

Taft, Ryan J, Simons, Cas, Nahkuri, Satu et al. Nuclear-localized tiny RNAs are associated with transcription initiation and splice sites in metazoans. *Nature structural & molecular biology*, 17, 8 (Aug 1, 2010), 1030-4. 10.1038/nsmb.1841.

Miranda, Kevin C, Huynh, Tien, Tay, Yvonne et al. A pattern-based method for 5 the identification of MicroRNA binding sites and their corresponding heteroduplexes. *Cell*, 126, 6 (Sep 22, 2006), 1203-17. 10.1016/j.cell.2006.07.031.

Cserzo, Miklos, Turu, Gabor, Varnai, Peter, and Hunyady, Laszlo. Relating 6 underrepresented genomic DNA patterns and tiRNAs: the rule behind the observation and

beyond. Biology direct, 5 (Jan 1, 2010), 56. 10.1186/1745-6150-5-56.

Moore, Hayley C, Johnston, Michael, Nicol, Samantha M, Bourdon, Jean7 Christophe, Thompson, Alastair M, Hutvagner, Gyorgy, and Fuller-Pace, Frances V. An evolutionarily conserved, alternatively spliced, intron in the p68/DDX5 DEAD-box RNA helicase gene encodes a novel miRNA. *RNA* (*New York*, *NY*) (Feb 23, 2011). 10.1261/rna.2591611.

Dasgupta, A, Das, A, Roy, R, Ralston, R, Majumdar, A, and Gupta, N K. Protein synthesis in rabbit reticulocytes XXI. Purification and properties of a protein factor (Co-EIF-1) which stimulates Met-tRNAf binding to EIF-1. *The Journal of biological chemistry*, 253, 17 (Sep 10, 1978), 6054-9.

Ghosh-Dastidar, P, Yaghmai, B, Das, A, Das, H K, and Gupta, N K. Protein synthesis in rabbit reticulocytes. Demonstration of the requirements for eIF-2 and Co-eIF-2A for peptide chain initiation using immune sera. *The Journal of biological chemistry*, 255, 2 (Jan 25, 1980), 365-8.

Chakravarty, I, Bagchi, M K, Roy, R, Banerjee, A C, and Gupta, N K. Protein synthesis in rabbit reticulocytes. Purification and properties of an Mr 80,000 polypeptide (Co-eIF-2A80) with Co-eIF-2A activity. *The Journal of biological chemistry*, 260, 11 (Jun 10, 1985), 6945-9.

Maniataki, Elisavet and Mourelatos, Zissimos. Human mitochondrial tRNAMet is exported to the cytoplasm and associates with the Argonaute 2 protein. *RNA (New York, NY)*, 11, 6 (Jun 1, 2005), 849-52. 10.1261/rna.2210805.

Mei, Yide, Yong, Jeongsik, Liu, Hongtu, Shi, Yigong, Meinkoth, Judy, Dreyfuss,
Gideon, and Yang, Xiaolu. tRNA binds to cytochrome c and inhibits caspase activation.
Molecular cell, 37, 5 (Mar 12, 2010), 668-78. 10.1016/j.molcel.2010.01.023.

Linsen, Sam E V, de Wit, Elzo, Janssens, Georges et al. Limitations and 3 possibilities of small RNA digital gene expression profiling. *Nature Methods*, 6, 7 (Jul 1, 2009), 474-6. 10.1038/nmeth0709-474.

Saikia, M, Fu, Y, Pavon-Eternod, M, He, C, and Pan, T. Genome-wide analysis of M1-methyl-adenosine modification in human tRNAs. *RNA (New York, NY)*, 16, 7 (Jul 1, 2010), 1317-1327. 10.1261/rna.2057810.

Tomari, Yukihide, Du, Tingting, and Zamore, Phillip D. Sorting of Drosophila 5 small silencing RNAs. *Cell*, 130, 2 (Jul 27, 2007), 299-308. 10.1016/j.cell.2007.05.057.

Förstemann, Klaus, Horwich, Michael D, Wee, Liangmeng, Tomari, Yukihide, and Zamore, Phillip D. Drosophila microRNAs are sorted into functionally distinct argonaute complexes after production by dicer-1. *Cell*, 130, 2 (Jul 27, 2007), 287-97. 10.1016/j.cell.2007.05.056.

Thompson, Debrah M and Parker, Roy. Stressing out over tRNA cleavage. *Cell*, 7 138, 2 (Jul 23, 2009), 215-9. 10.1016/j.cell.2009.07.001.

Suhasini, Avvaru N and Sirdeshmukh, Ravi. Transfer RNA cleavages by onconase 8 reveal unusual cleavage sites. *The Journal of biological chemistry*, 281, 18 (May 5, 2006), 12201-9. 10.1074/jbc.M504488200.

Shigematsu, M, Ogawa, T, Kido, A, Kitamoto, H, Hidaka, M, and Masaki, H.

9 Cellular and transcriptional responses of yeast to the cleavage of cytosolic tRNAs induced by colicin D. *Yeast (Chichester, England)* (Oct 29, 2009). 10.1002/yea.1725.

Serebrov, V, Clarke, R J, Gross, H J, and Kisselev, L. Mg2+-induced tRNA 0 folding. *Biochemistry*, 40, 22 (Jun 5, 2001), 6688-98.

Phizicky, Eric M and Hopper, Anita K. tRNA biology charges to the front. *Genes*1 & development, 24, 17 (Sep 1, 2010), 1832-60. 10.1101/gad.1956510.

Madore, E, Florentz, C, Giegé, R, and Lapointe, J. Magnesium-dependent 2 alternative foldings of active and inactive Escherichia coli tRNA(Glu) revealed by chemical probing. *Nucleic Acids Research*, 27, 17 (Sep 1, 1999), 3583-8.

Grosjean, H, Sprinzl, M, and Steinberg, S. Posttranscriptionally modified nucleosides in transfer RNA: their locations and frequencies. *Biochimie*, 77, 1-2 (1995), 139-41.

Alexandrov, Andrei, Chernyakov, Irina, Gu, Weifeng, Hiley, Shawna L, Hughes, 4 Timothy R, Grayhack, Elizabeth J, and Phizicky, Eric M. Rapid tRNA decay can result from lack of nonessential modifications. *Molecular cell*, 21, 1 (Jan 6, 2006), 87-96. 10.1016/j.molcel.2005.10.036.

Tolia, N and Joshua-Tor, L. Slicer and the argonautes. *Nat Chem Biol*, 3, 1 (Jan 1, 5 2007), 36-43.

Mei, Yide, Yong, Jeongsik, Liu, Hongtu, Shi, Yigong, Meinkoth, Judy, Dreyfuss,
Gideon, and Yang, Xiaolu. tRNA binds to cytochrome c and inhibits caspase activation.
Molecular cell, 37, 5 (Mar 12, 2010), 668-78. 10.1016/j.molcel.2010.01.023.