



Published in final edited form as:

Trends Cell Biol. 2018 October ; 28(10): 793–806. doi:10.1016/j.tcb.2018.05.006.

Tie-Break: Host and Retrotransposons Play tRNA

Andrea J. Schorn¹ and Rob Martienssen^{*,1,2}

¹Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA;

²Howard Hughes Medical Institute, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA

Abstract

tRNA fragments (tRF) are a class of small, regulatory RNAs with diverse functions. 3'-derived tRFs perfectly match long terminal repeat (LTR)-retroelements which use the 3'-end of tRNAs to prime reverse transcription. Recent work has shown that tRFs target LTR-retroviruses and -transposons for the RNA interference (RNAi) pathway and also inhibit mobility by blocking reverse transcription. The highly conserved tRNA primer binding site (PBS) in LTR-retroelements is a unique target for 3'-tRFs to recognize and block abundant but diverse LTR-retrotransposons that become transcriptionally active during epigenetic reprogramming in development and disease. 3'-tRFs are processed from full-length tRNAs under yet unknown conditions and potentially protect many cell types. tRFs appear as an ancient link between RNAi, transposons, and genome stability.

Keywords

tRNA; small RNA; tRF; LTR-retrotransposon; retrovirus; PBS

LTR-Retroelements and tRNA Fragments Come Together

tRNA fragments (tRFs) have been found in an increasing number of tissues and organisms in all kingdoms of life [1]. Originally disregarded as mere degradation products of abundant structural RNAs, it has become clear that tRFs are regulatory small RNAs with biological functions. tRFs have been shown to bind AGO and PIWI proteins and to guide RNAi-mediated silencing of complementary targets. 3'-tRFs have perfect sequence complementarity to long terminal repeat (LTR) retroelements which use the 3'-end of tRNAs to prime reverse transcription. LTR-retroelements in eukaryotes replicate through an RNA intermediate and include both, endogenous LTR-retrotransposons, as well as exogenous, infectious LTR-retroviruses. LTR-retrotransposons in mammals are also called endogenous retroviruses (ERVs) because of their phylogenetic similarity to LTR-retroviruses (Box 1). While transposon and virus mobility is hazardous to the host, their sequences and promoter activity have provided important regulatory elements for development and tissue identity [2–4].

*Correspondence: martiens@cshl.edu.

LTR-retrotransposons are very abundant in higher eukaryotes and become transcriptionally active during epigenetic reprogramming in development and disease. While piRNAs protect the germline during reprogramming [5] and endogenous siRNAs safeguard the genome in oocytes [6, 7], it was unknown what restricts transposition in tissues that lack these small RNA classes but experience high transposon expression, such as the preimplantation embryo. 3'-tRFs are abundantly expressed in these tissues and are capable of recognizing potentially mobile LTR-retrotransposons at their highly conserved tRNA primer binding site (PBS). Thus the dependency of LTR-retroelements on host tRNA becomes their Achilles heel, and tRF-targeting is a potentially highly conserved mechanism of small RNA-mediated transposon control. Insights into tRF function, tRNA metabolism and potent impacts on retroelement mobility have emerged recently. tRF production is tightly regulated, and cells are much more sensitive to tRNA levels than previously appreciated. Here, we review what is known about tRF-mediated control of LTR-retroelements and its impact on genome integrity. We discuss evolutionary implications as well as the many open questions concerning biogenesis and mechanism of this novel small RNA-regulation.

LTR-Retroelements and tRNA Go a Long Way - tRNAs Prime Reverse Transcription

Retroelements encode RNA-dependent DNA polymerase or reverse transcriptase (RT) which probably enabled the transition from the ancient RNA world to today's DNA-based life forms [8]. They can be divided into LTR- and non-LTR retroelements (Box 1). Non-LTR retroelements nick the genomic target DNA and use that free 3'-end to prime reverse transcription. In contrast, LTR-retroelements encode proteins that form viral particles and reverse transcribe in a protected, subcellular environment in the cytoplasm, where an extrachromosomal DNA copy can be generated without invoking DNA repair. That necessitates bringing along an RT primer, and retroviral RT enzymes are most efficient with an RNA primer [9]. LTR-retroelements have solved this problem by recruiting cytoplasmic, host tRNAs to prime RT (Figure 1). There are a few exceptions: the Tf1/Sushi family is self-priming [10] and a subgroup of Copia elements is primed by truncated tRNAs, e.g. a 39 nt long 5'-fragment of tRNA-iMet [11] and a leucine tRNA missing the last 5 nt at its 3'-end [12]. However, the overwhelming majority of LTR-retroviruses and -transposons are primed by full-length tRNAs (Table 1). While short RNA or DNA pieces work as primers *in vitro*, *in vivo* processivity of the RT enzyme requires contact with the full-length, structured tRNA for successful priming and elongation [13, 14]. Reverse transcription is initiated at the tRNA primer binding site (PBS) and proceeds through several strand transfer events until a full copy of the transposon is ready for integration (Figure 1). During this process, autonomous coding retrotransposons frequently reverse transcribe RNA from non-coding family members with similar LTRs that "hitchhike" the replication machinery. In fact, a majority of observed novel insertions are non-coding transposons that get carried over by a few active, coding master copies in the genome [15, 16].

tRNA Fragments - Throwing a Wrench in the Works

tRNA genes are transcribed by RNA polymerase III (POLIII) into a precursor RNA (pre-tRNA) that undergoes several maturation steps: a leader and a trailer sequence are cleaved off by RNaseP and RNaseZ, respectively, and a CCA-trinucleotide is added at the 3'-end. RNA modifications are introduced and stabilize the L-shaped tertiary structure that is distorted during retroelement priming. Several cleavage products of mature tRNAs have been reported (Box 2). Interestingly, 3'- and 5'-tRFs are able to guide RNaseP and an RNaseZ isoform, respectively, *in vivo* to cleave mRNA that base-pairs with them in a tRNA-like fold [17, 18]. 5'-derived fragments and halves have been found in exosomes and are upregulated during stress response, cancer and neurological disorders [1, 19–22]. 5'-halves and full-length tRNAs exert anti-apoptotic effects when bound to cytoplasmic cytochrome C during stress [23, 24]. Conversely, 5'-halves and internal fragments from the anticodon loop can inhibit initiation of translation and act as tumor suppressors through displacement of the oncogenic RNA-binding protein YBX-1 [25, 26]. Because they inhibit translation, assessing the function of 5'-fragments in reporter gene silencing assays needs caution. Similarly, studies of 3'-fragments are possibly confounded when using antisense oligos which necessarily contain parts of the 5'-tRNA sequence.

In contrast to 5'-tRF, little was known about the biogenesis or function of 3'-derived fragments until recently. 3'-tRFs from mature tRNAs include the post-transcriptional 3'-CCA tail and have been detected in prokaryotes, fungi, protists, plants, and animals [1]. They come in two sizes, 17–19 and 22 nt, and carry 1-methyladenosine and pseudouridine at positions 19 and 22, respectively, counting from the mature 3' tRNA end. Similar to 5'-tRFs, 3'-tRFs have the capacity to inhibit gene expression, as demonstrated by miRNA reporter assays in which a complementary target sequence was present in the 3'-terminal UTR [27–30]. Furthermore, both 3'- and 5'-tRFs can bind to AGO and PIWI proteins supporting the idea that they act in small RNA-mediated silencing [31–33]. For example, a 22 nt 3'-tRF from glycine tRNA (Gly-GGC) acts as a miRNA *in vivo* and targets RPA1 to modulate the DNA damage response [34]. In the unicellular eukaryote, *Tetrahymena thermophila*, 18 nt 3'-tRFs bind to the PIWI protein TWI12, upon which it translocates to the nucleus and activates the XRN2 exonuclease involved in RNA decay [35]. 3'-tRFs antisense to the PBS of retroelements have been noticed as very abundant in mouse and human and previously been proposed to regulate LTR-retroelements [27, 30, 36–38], although multi-mapping reads are often discarded in small RNA sequencing studies. Indeed, retrotransposition of the two most active murine LTR-retroelements (Box 1 and Table 1), IAP and MusD/ETn, is strongly inhibited by 3'-tRFs. 22 nt 3'-tRFs silence ERVs through an RNAi-like mechanism, while 18 nt 3'-tRFs specifically interfere with reverse transcription and retrotransposon mobility [39]. Endogenous 22 nt 3'-tRFs reduce RNA and protein levels of autonomous, coding LTR-retroelements that are translated (Figure 2). 18 nt 3'-tRFs do not change RNA or protein levels but instead block reverse transcription, as evidenced by strongly decreased RNaseH intermediates following first strand (–)ssDNA synthesis [39]. In accordance with this, HIV infection induced an 18 nt 3'-tRFs from its cognate lysine tRNA primer that inhibited retrovirus replication [27]. This leads to a model in which any replication-competent LTR-element with a functional PBS can be blocked, at the latest, at the reverse transcription step.

This small RNA-mediated transposon control is able to recognize not only autonomous, but also abundant non-coding, non-autonomous hitchhiker elements that escape post-transcriptional gene silencing during translation [39].

18 nt 3'-tRFs are highly expressed in tissues that are challenged by LTR-retroelement expression or infection. In mammalian embryos, LTR-retrotransposons are usually repressed by histone H3 lysine-9 trimethylation (H3K9me3) but become transcriptionally active during developmental epigenetic reprogramming [40]. Loss of H3K9me3 in mouse embryonic stem cells (ESC) leads to strong derepression of IAP and ETn [41, 42]. 3'-tRF matching IAP and ETn are highly enriched in these cells, suggesting that 3'-tRFs protect the pre-implantation embryo from transposon damage during epigenetic reprogramming [39], similar to piRNAs in the germline [5]. 3'-tRF levels are also high in trophoblast stem cells that have low H3K9me3, high ERV expression, and have co-opted ERVs in many ways [39, 43, 44]. Interestingly, increasing HIV titers triggers production of lysine 3'-tRFs against HIV [27], and T cells infected with HTLV-1 upregulate proline 3'-tRFs targeting the virus [36]. All these findings suggest that viral defense could have evolved from a much more widespread role for tRFs in inhibiting LTR-retroelements, and that 3'-tRFs are able to serve as an innate immune response. Unlike other small RNA defense mechanisms that require an increase in copy number of the invader over time, tRFs are readily available in naive cells that have not been previously exposed to infection or horizontal transfer. 3'-tRFs from non-primer tRNAs are also expressed in cells [30, 31, 37, 39] and it is possible that 3'-tRFs are generated at random before those matching the PBS of LTR-retroelements become enriched and utilized for silencing.

Several 5'-derived tRFs have been proposed to regulate LTR-retrotransposons. Recently, 5'-tRFs (Gly-GGC) in mouse sperm have been implicated in paternal epigenetic inheritance and expression changes of genes driven by MERVL LTR-promoter activity in the offspring [45]. It will be interesting to see if other 5'-tRFs have an effect on transposon promoter activity and how this suppression is achieved, as sequence complementarity of the Gly-GGC 5'-tRF to MERVL elements is not evident. Given the prevalence of reverse transcribed tRNA pseudogenes and that genomic tRNA halves often overlap retrotransposons [46], it is probable that certain 5'-tRFs have some sequence complementarity to non-PBS targets within transposons and guide RNAi, as recently reported in *Arabidopsis thaliana* [47]. Similarly, there are dozens of miRNAs that originate from and target transposons [48–50] and there are miRNAs that are derived from pre-tRNA [51, 52]. Interestingly, there is also an overlap between piRNAs and tRF sequences that are expressed both in somatic [32, 53] and in germline tissues (our unpublished results) [54]. This suggests some of those tRNA-derived piRNAs (td-piRNA) might well target transposon sequences as do piRNAs. 5'-tRFs, miRNA, and piRNA typically target transposons at random sites along the mRNA, while 3'-tRFs recognize the highly conserved 18 nt PBS located outside of promoter and coding sequences.

Transposons and viruses frequently evade defense mechanisms through mutation, but mutations within the PBS abrogate reverse transcription and therefore disable the element. In plants, the target site of the highly conserved miR845 family overlaps the PBS of LTR-retrotransposons, which are primed by tRNA-iMet [55, 56]. In some plants these miRNA

precursors are derived from pre-tRNA-iMet genes by a sequence inversion, while in others (including *Arabidopsis*) they are derived from retrotransposons. miR845 is specifically expressed in *Arabidopsis* pollen, where it targets the PBS of hundreds of retrotransposons, and triggers an RNAi response that impacts the expression of imprinted genes and mediates chromosome dosage response [55]. 3'-tRFs have the ability to target all LTR-retroelements that are replication-competent by recognizing the highly conserved PBS. In addition, they come without off-target effects on transposon-derived genes and should therefore be especially useful at developmental stages that are driven by transposon-derived sequences for timed expression and imprinting [2, 57].

tRF Biogenesis - the Who? What? When? Where? Why?

Where & When

To better understand tRF-mediated inhibition, it will be important to learn more about subcellular localization of tRFs and their opportunities to interact with retroelements. While 5'- and 3'-halves are derived from aminoacylated tRNAs [20], most 3'-tRFs that have been sequenced to date were captured in ligation-dependent sequencing without prior treatment and hence had a free 3'-hydroxyl RNA end. Mature tRNAs are cytoplasmic and their 3'-end is free of the covalently attached amino acid before they are loaded by aminoacyl-synthetases but also after translation at the ribosome when they get recycled for another round of translation [58]. Hence, 3'-tRFs that include the CCA-tail of mature tRNAs are thought to be cytoplasmic and could be cleaved from full-length tRNAs before aminoacylation or after translation. 22 nt 3'-tRFs induce post-transcriptional gene silencing and hence bind to and inhibit the retroelement mRNA during translation in the cytoplasm [28, 31, 34, 39]. Silencing by 22 nt tRFs is strong even in the presence of mismatches, reminiscent of miRNA-mediated silencing that tolerates multiple mismatches with the target. 17–19 nt 3'-tRFs on the other hand interfere with retrotransposition later in the life cycle of the element, during reverse transcription of the unspliced “genomic” virus RNA template (Figure 2) [39]. A fraction of the LTR-retroelement transcript is spliced into mRNA for translation and its freshly synthesized proteins play an active part in selection of the primer tRNA and possibly tRF binding. In general, RT enzymes have evolved to bind to one specific tRNA with high affinity and recruit this tRNA to the viral particle for reverse transcription [59]. This primer tRNA is specifically enriched in the viral particle even in the absence of viral genomic RNA demonstrating that its recruitment is by protein-RNA interactions rather than an RNA-RNA interaction of PBS and tRNA [59]. HIV and MuLV only use cytoplasmic tRNAs and replication is maximal with those tRNAs that have passed rigorous cellular quality control and are capable of loading the correct amino acid [60]. Retrovirus particles are built of GAG proteins on their surface and contain 2 copies of unspliced virus genomic RNA and ~50 molecules of tRNA, very few of which are non-specific tRNAs [61]. In HIV, tRNA selection is mediated through the interaction of GAG proteins and the Lysyl-tRNA synthetase [62]. Other examples of LTR-retroviruses show high affinity of tRNA directly for RT [63]. In principle, 18 nt 3'-tRFs could inhibit tRNA recruitment to viral particles, interfere with maturation of the RT enzyme, or inhibit reverse transcription within the viral particle. Although RT priming of both coding MusD and non-coding ETn depend on the MusD RT enzyme and tRNA Lys-AAA, inhibition by 3'-tRFs

does not affect both elements to the same extent indicating recruitment of enzymes and primer tRNA were not affected but instead sequence complementarity at the PBS is decisive [39]. Consistent with inhibition by 18 nt 3'-tRFs occurring within the virus particle, 18 nt proline 3'-tRFs matching the PBS of HTLV-1 have been found enriched in retroviral particles [35]. Inhibition of priming could either happen during initiation as a consequence of tRNA and tRF competition or during elongation of (–)ssDNA due to a failed allosteric transition of RT to the processive phase of DNA synthesis [13, 64]. It is important to keep in mind, that while short RNA or DNA oligos are able to prime purified RT *in vitro*, *in vivo* tRNA primer recruitment as well as priming and elongation by RT require interaction with several residues along the full-length, structured tRNA [13]. tRNAs with perfect match to the PBS but truncated or mutated outside the acceptor 3'-arm have been shown to be orders of magnitude less processive than full-length tRNA [14] resulting in a *de facto* suppression of reverse transcription *in vivo* when tRNAs and tRFs compete [27, 39].

What & When

Infectious viruses like HIV and HTLV-1 trigger a rise of 18 nt 3'-tRF (tRF-3a) level from their cognate primer tRNA [27, 36]. However, mouse ERV expression in human HeLa cells did not trigger significant amounts of tRFs matching those ERVs, even though mouse cells with high ERV expression contain high levels of matching tRF-3a [39]. It is conceivable that intracellular transposons fail to evoke the same response as exogenous, infectious viruses which trigger a series of cellular immune responses. Initially, tRF production might be a consequence of the epigenetic state of cells and POLIII precursor transcription. For example, tRF production could be a consequence of low H3K9me3 levels and tRNA metabolism [65]. Of course it is ideal for transposons to replicate in highly proliferating cells: high levels of tRNAs for translation of viral proteins and RT priming, abundant replication forks, and breakdown of the nuclear membrane for genome integration all play in favor of transposition [59, 66, 67]. Transposition rates can be limited by host tRNA expression [68] and oncogenic LTR-retroviruses like RSV and MuLV are known to induce expression of their primer tRNA [69]. While tRNA amounts in the cell directly affect transposon translation and primer availability for transposition, tRNA processing can also indirectly affect transposon expression [70]. Defects in tRNA processing in *Drosophila* result in increased replication stress, misregulation of neighboring piRNA clusters, and therefore expression of transposons [71]. Given that subtle changes in tRNA level can promote cancer [72], it seems safe to assume that cleavage into tRFs is tightly regulated by the cell to inhibit retrotransposition but not to disable translation.

tRNA transcription by POLIII, editing, folding, and quality control ultimately feed into the pool of mature tRNA available for 3'-tRF production [73]. The Lupus Autoantigen (La) protein chaperones nascent POLIII transcripts and functions as a gatekeeper for pre-tRNAs ending up in the RNAi pathway [74]. La knockdown results in an increase of AGO1–4 binding to 21–23 nt fragments from immature precursor tRNAs compared to abundant 18 nt tRFs from mature tRNAs in wild-type cells. One example is the tRNA-Ile precursor that is either processed and CCA-tailed to result in tRNA-Ile and tRFs or “misfolded” into a miRNA precursor and processed by the double-stranded endoribonuclease DICER into miR-1983 [52, 74]. Expression of full-length tRNAs does not seem to directly correlate with

tRF levels [30], but tRNA quantification by cDNA sequencing is hampered by post-transcriptional RNA modifications that inhibit RT, and has only recently become feasible [72, 75, 76].

Who & Why

While there is some consensus on the biogenesis of tRNA-halves (Box 2), little is certain about the biogenesis of tRFs. AGO and PIWI family proteins bind tRFs and act as effector molecules in many studies that showed tRF-mediated silencing, but their role in cleaving tRNA is less clear. The main cargo of the *Tetrahymena* PIWI protein Twi12 are tRF-3a, yet it has no slicing activity [33]. Several endonucleases cleave tRNA *in vitro*, yet the endonucleases that cut mature tRNAs into tRFs still need to be identified *in vivo*. One idea is that Lys-AAA tRF-3a may be processed by DICER from double-stranded RNA formed by hybridization of the HIV PBS and the tRNA primer [27]. Biogenesis of individual 5'- and 3'-tRFs have been shown to be DICER-dependent [27, 28, 34, 47, 52]. At the same time, most tRFs are still produced in Dicer (and Drosha) knockouts in mouse ESC and *Schizosaccharomyces pombe* [31, 37], as well as in *Arabidopsis*, *Tetrahymena*, or *Drosophila melanogaster* [31, 77, 78] though redundancy of Dicer-like genes has not been examined in the latter. The ratio of tRF-3a and tRF-3b is highly tissue-specific [39], suggesting those 18 and 22 nt 3'-tRFs may be produced by two different endonucleases under varying conditions. Alternatively, the longer tRF-3b could be an initial cleavage product that is trimmed by an exonuclease upon specific stimuli.

It is tempting to speculate that an endonuclease of the RNAi pathway evolved to cleave tRNAs into 3'-tRF to counteract RT activity. Both the tertiary structure of mature, L-shaped tRNA as well as the tRNA minihelix, thought to be an evolutionary precursor [58] offer a perfect hairpin RNA substrate for RNAi. The minihelix and the upper arm of today's L-shape consist of the acceptor stem and the T Ψ C-loop (Figure 3), so they contain 3'-tRFs and the first ~7 nt of the 5'-tRNA end [58]. tRNA minihelices are functional substrates for the CCA-adding enzyme and its quality control mechanism [79]. Specific nucleotides in the acceptor stem portion of the minihelix are necessary and sufficient for aminoacylation by specific tRNA synthetases, supporting the idea of a primordial code by tRNA minihelices [58, 80]. Intriguingly, several of the proteins involved in 3'-tRF mediated retroelement silencing carry an RNaseH domain: RT has an RNaseH domain to digest the virus RNA template, and the PIWI domain in AGO/PIWI is closely related to retroviral RNaseH [81]. Both, AGO and PIWI proteins, function in "digestion" of virus RNA, and AGO proteins are related to components of the eukaryotic Initiation Factor eIF2 that carries the initiator tRNA for translation [82]. tRF-silencing of LTR-retroelements points to the interaction and co-evolution of ancient and antagonistic players, namely tRNAs and their cleavage products. For example, translation is aided by eIF2 and tRNAs while RNAi is guided by AGO and tRFs; reverse transcription is primed by tRNAs while it is inhibited by tRFs. Thus, tRNA could be an ancient RNAi substrate that links small-RNA mediated silencing and transposon control.

Concluding Remarks and Future Perspectives

Pieces of tRNA were first noticed during phage infection of bacteria [83, 84], and factors involved in tRF-mediated silencing may be traced back to the origins of RNAi in the defense against mobile repetitive elements. Viruses and tRNAs as well as RT and AGO/PIWI family proteins which bind tRNAs and tRFs have an ancient and intricate relationship [58, 81]. Several retroelements and long non-coding RNAs encode tRNA genes or terminate in a tRNA-like structure with the 3'-CCA trinucleotide [67, 85]. Most importantly, tRNAs prime reverse transcription of retroelements at the PBS, and so 3'-tRFs enable eukaryotes to recognize mobile LTR-retroelements by their PBS without affecting domesticated transposon-derived genes and promoters. Recognition is followed by post-transcriptional silencing and RT inhibition. 3'-tRFs might also guide transcriptional silencing like other small RNAs in yeast, animals and plants [86]. Indeed, td-piRNAs guide HIWI2 for SETDB1-dependent H3K9me3 deposition [87] and heterochromatin silencing of ERVs nucleates at the PBS [88]. Several zinc finger proteins (ZFP) bind to retroelements, and initiate H3K9me3 heterochromatin through interaction with KAP1 and SETDB1 [89]. However, not all spreading of H3K9me3 starting at the PBS can be explained by ZFPs as only some of them bind to the PBS. Studying the function of tRFs is complicated by the fact that tRNA metabolism is essential and the use of antisense oligos is problematic, as anti-3'-tRFs overlap the potential (miRNA) seed sequence of the corresponding 5'-tRF. It will be important to understand how similar tRF targeting rules are to those of miRNA-mediated silencing, and the extent to which the RNAi pathway is involved in biogenesis and function of tRFs (see Outstanding Questions).

One prediction is that LTR-retroelements which evade tRF targeting by using a degenerate, non-canonical PBS are more successful - until they lose the ability to bind tRNA at all and are unable to move. Can we see this signature in genomes and use it to discriminate active from non-active elements? The rate at which retrotransposons co-mobilize non-autonomous elements adds another level of complexity to define "success" of an element. In agreement with this idea, the autonomous MusD6 retrotransposon that is exposed to both post-transcriptional silencing by tRF-3b and RT inhibition by tRF-3a, has two mismatches with its cognate tRNA or 3'-tRF. Its non-coding partner element ETnIIbeta has a perfect PBS for optimal tRNA priming - perhaps because it goes unrecognized by 22 nt tRF-3b and is "only" exposed to 18 nt tRF-3a. Similarly, IAPs when released from epigenetic silencing in mouse ESC exhibit lower expression level [42] if they have a perfect match to a tRNA - and a tRF. This could be due to suppression by 3'-tRFs in cells that lose heterochromatin at retrotransposons.

Potential target sites of 3'-tRFs are also found at the PBS of human ERVs [30, 37]. Cancer cells show elevated levels of transposon expression [90] and abundant 3'-tRFs with perfect complementarity to HERVs are found in many cancer cell lines. While tRFs protect stem cells from transposition, they might stabilize cancer cells and have oncogenic properties. The proposed role of AGO and PIWI in cancer [91] suggests abundant, somatic tRFs might have an important function during reprogramming and transposon reactivation in disease.

Acknowledgements

We thank Benjamin Roche and Michael Gutbrod for discussion and reading of the manuscript. Research in the Martienssen lab is supported by Howard Hughes Medical Institute, the National Science Foundation, the National Institute of General Medical Sciences (GM067014 and GM076396) and the Cold Spring Harbor Laboratory Robertson Research Fund. We apologize to authors whose references were omitted due to space constraints.

Glossary

AGO/PIWI

Argonaute and related PIWI (*P-element induced wimpy testis*) proteins mediate silencing of complementary sequences guided by small RNAs and contain a PIWI domain that resembles an RNaseH domain.

DICER

The RNaseIII endoribonuclease *DICER* cleaves double-stranded RNA into short double-stranded RNA fragments that guide RNAi

epigenetic reprogramming

Epigenetic reprogramming erases and resets DNA and histone modifications to redefine active and inactive genes in the genome. This process enables cells to become pluripotent during development but can also be a consequence of malignant transformation in cancer.

isoacceptor

tRNAs that are loaded with the same amino acid but have distinct sequences, including the anticodon, are called *isoacceptors*

isodecoder

tRNAs that differ in sequence but have the same anticodon sequence are called *isodecoder*.

LTR

long terminal repeats flank the coding or non-coding sequence of LTR-retroelements and contain promoter as well as poly-A signals

miRNA

microRNAs base-pair with complementary target RNA and induce cleavage of the target or inhibit its translation. They are processed from imperfectly double-stranded hairpin RNA

PBS

LTR-retroelements use tRNAs to initiate reverse transcription at their *primer binding site* downstream of the 5'-LTR and upstream of the first open reading frame

piRNA

PIWI-interacting RNA and PIWI proteins are typically but not exclusively expressed in germline tissues and frequently target transposable element transcripts for degradation but can also induce heterochromatic silencing.

tRNA

transfer RNA are ~70–120 nucleotide adaptor molecules that translate mRNA into protein at the ribosome. The acceptor stem sequence and the anticodon determine which amino acid gets attached to decode a specific codon RNA triplet in the mRNA

tRF

tRNA fragments are cleaved from immature precursor tRNA or from mature tRNA and have diverse biological functions.

RT

Reverse transcriptase is an RNA-dependent DNA polymerase that enables retroelements to copy their RNA transcript into a DNA copy for insertion into the genome

RNAi

RNA interference occurs when small RNAs bound to AGO/PIWI proteins interfere with expression or translation of complementary target RNA at the transcriptional or post-transcriptional level

RNase H

The *Ribonuclease H* family of endonucleases cleaves RNA in RNA-DNA hybrids. In LTR-retroelements, RNaseH degrades the virus RNA after reverse transcription; in the cell it resolves RNA-DNA loops during transcription

RNA world

The *RNA world* hypothesizes an original, primitive form of life completely executed by the enzymatic potential of RNA without DNA or proteins

siRNA

When bound to AGO/PIWI proteins small interfering *RNAs* base-pair with complementary target RNA and induce cleavage of the target or inhibit its translation. In many organisms, they also guide heterochromatin formation via modification of DNA and histones

virus particle

Proteins of the group-antigen (Gag) gene of LTR-retroelements form subcellular compartments called *virus-like or virus particles* in which reverse transcription takes place

References

1. Kumar P et al. (2016) Biogenesis and Function of Transfer RNA-Related Fragments (tRFs). *Trends Biochem Sci* 41 (8), 679–89. [PubMed: 27263052]
2. Peaston AE et al. (2004) Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Dev Cell* 7 (4), 597–606. [PubMed: 15469847]
3. Feschotte C (2008) Transposable elements and the evolution of regulatory networks. *Nat Rev Genet* 9 (5), 397–405. [PubMed: 18368054]
4. Faulkner GJ et al. (2009) The regulated retrotransposon transcriptome of mammalian cells. *Nat Genet* 41 (5), 563–71. [PubMed: 19377475]
5. Siomi MC et al. (2011) PIWI-interacting small RNAs: the vanguard of genome defence. *Nat Rev Mol Cell Biol* 12 (4), 246–58. [PubMed: 21427766]
6. Watanabe T et al. (2008) Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* 453 (7194), 539–43. [PubMed: 18404146]

7. Tam OH et al. (2008) Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature* 453 (7194), 534–8. [PubMed: 18404147]
8. Boeke JD (2003) The unusual phylogenetic distribution of retrotransposons: a hypothesis. *Genome Res* 13 (9), 1975–83. [PubMed: 12952870]
9. Baltimore D and Smoler D (1971) Primer requirement and template specificity of the DNA polymerase of RNA tumor viruses. *Proc Natl Acad Sci U S A* 68 (7), 1507–11. [PubMed: 4327004]
10. Levin HL (1995) A novel mechanism of self-primed reverse transcription defines a new family of retroelements. *Mol Cell Biol* 15 (6), 3310–7. [PubMed: 7760826]
11. Kikuchi Y et al. (1986) Unusual priming mechanism of RNA-directed DNA synthesis in copia retrovirus-like particles of *Drosophila*. *Nature* 323 (6091), 824–6. [PubMed: 2430190]
12. Saigo K (1986) A potential primer for reverse transcription of *mdg3*, a *Drosophila* copia-like element, is a leucine tRNA lacking its 3' terminal 5 bases. *Nucleic Acids Res* 14 (10), 4370. [PubMed: 3635844]
13. Le Grice SF (2003) “In the beginning”: initiation of minus strand DNA synthesis in retroviruses and LTR-containing retrotransposons. *Biochemistry* 42 (49), 14349–55. [PubMed: 14661945]
14. Keeney JB et al. (1995) Multiple molecular determinants for retrotransposition in a primer tRNA. *Mol Cell Biol* 15 (1), 217–26. [PubMed: 7528326]
15. Ribet D et al. (2004) An active murine transposon family pair: retrotransposition of “master” *MusD* copies and *ETn* trans-mobilization. *Genome Res* 14 (11), 2261–7. [PubMed: 15479948]
16. Nellaker C et al. (2012) The genomic landscape shaped by selection on transposable elements across 18 mouse strains. *Genome Biol* 13 (6), R45. [PubMed: 22703977]
17. Plehn-Dujowich D and Altman S (1998) Effective inhibition of influenza virus production in cultured cells by external guide sequences and ribonuclease P. *Proc Natl Acad Sci U S A* 95 (13), 7327–32. [PubMed: 9636148]
18. Elbarbary RA et al. (2009) Modulation of gene expression by human cytosolic tRNase Z(L) through 5'-half-tRNA. *PLoS One* 4 (6), e5908. [PubMed: 19526060]
19. Nolte-t Hoen EN et al. (2012) Deep sequencing of RNA from immune cell-derived vesicles uncovers the selective incorporation of small non-coding RNA biotypes with potential regulatory functions. *Nucleic Acids Res* 40 (18), 9272–85. [PubMed: 22821563]
20. 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 (29), E3816–25. [PubMed: 26124144]
21. Chakraborty SK et al. (2015) Extracellular vesicle-mediated transfer of processed and functional RNY5 RNA. *RNA* 21 (11), 1966–79. [PubMed: 26392588]
22. Shurtleff MJ et al. (2017) Broad role for YBX1 in defining the small noncoding RNA composition of exosomes. *Proc Natl Acad Sci U S A* 114 (43), E8987–E8995. [PubMed: 29073095]
23. Mei Y et al. (2010) tRNA binds to cytochrome c and inhibits caspase activation. *Mol Cell* 37 (5), 668–78. [PubMed: 20227371]
24. Saikia M et al. (2014) Angiogenin-cleaved tRNA halves interact with cytochrome c, protecting cells from apoptosis during osmotic stress. *Mol Cell Biol* 34 (13), 2450–63. [PubMed: 24752898]
25. Goodarzi H et al. (2015) Endogenous tRNA-Derived Fragments Suppress Breast Cancer Progression via YBX1 Displacement. *Cell* 161 (4), 790–802. [PubMed: 25957686]
26. Ivanov P et al. (2011) Angiogenin-induced tRNA fragments inhibit translation initiation. *Mol Cell* 43 (4), 613–23. [PubMed: 21855800]
27. Yeung ML 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 (19), 6575–86. [PubMed: 19729508]
28. Haussecker D et al. (2010) Human tRNA-derived small RNAs in the global regulation of RNA silencing. *RNA* 16 (4), 673–95. [PubMed: 20181738]
29. Cole C et al. (2009) Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs. *RNA* 15 (12), 2147–60. [PubMed: 19850906]
30. Kawaji H et al. (2008) Hidden layers of human small RNAs. *BMC Genomics* 9, 157. [PubMed: 18402656]

31. 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. [PubMed: 25270025]
32. Keam SP et al. (2014) The human Piwi protein Hiwi2 associates with tRNA-derived piRNAs in somatic cells. *Nucleic Acids Res* 42 (14), 8984–95. [PubMed: 25038252]
33. Couvillion MT et al. (2010) A growth-essential *Tetrahymena* Piwi protein carries tRNA fragment cargo. *Genes Dev* 24 (24), 2742–7. [PubMed: 21106669]
34. Maute RL 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* 110 (4), 1404–9. [PubMed: 23297232]
35. Couvillion MT 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 (4), 509–20. [PubMed: 23084833]
36. Ruggiero K et al. (2014) Small noncoding RNAs in cells transformed by human T-cell leukemia virus type 1: a role for a tRNA fragment as a primer for reverse transcriptase. *J Virol* 88 (7), 3612–22. [PubMed: 24403582]
37. Li Z et al. (2012) Extensive terminal and asymmetric processing of small RNAs from rRNAs, snoRNAs, snRNAs, and tRNAs. *Nucleic Acids Res* 40 (14), 6787–99. [PubMed: 22492706]
38. Schorn A and Martienssen RA, F1000Prime Recommendation of [Couvillion MT et al., *Genes Dev* 2010, 24(24):2742–7]., [PubMed: 21106669]
39. Schorn AJ et al. (2017) LTR-Retrotransposon Control by tRNA-Derived Small RNAs. *Cell* 170 (1), 61–71 e11. [PubMed: 28666125]
40. Leung DC and Lorincz MC (2012) Silencing of endogenous retroviruses: when and why do histone marks predominate? *Trends Biochem Sci* 37 (4), 127–33. [PubMed: 22178137]
41. Matsui T et al. (2010) Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. *Nature* 464 (7290), 927–31. [PubMed: 20164836]
42. Rowe HM et al. (2010) KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature* 463 (7278), 237–40. [PubMed: 20075919]
43. Chuong EB et al. (2013) Endogenous retroviruses function as species-specific enhancer elements in the placenta. *Nat Genet* 45 (3), 325–9. [PubMed: 23396136]
44. Hemberger M (2010) Genetic-epigenetic intersection in trophoblast differentiation: implications for extraembryonic tissue function. *Epigenetics* 5 (1), 24–9. [PubMed: 20083894]
45. Sharma U et al. (2016) Biogenesis and function of tRNA fragments during sperm maturation and fertilization in mammals. *Science* 351 (6271), 391–6. [PubMed: 26721685]
46. Zuo Z et al. (2013) Genome-wide analysis reveals origin of transfer RNA genes from tRNA halves. *Mol Biol Evol* 30 (9), 2087–98. [PubMed: 23744908]
47. Martinez G et al. (2017) tRNA-derived small RNAs target transposable element transcripts. *Nucleic Acids Res.*
48. Smalheiser NR and Torvik VI (2005) Mammalian microRNAs derived from genomic repeats. *Trends Genet* 21 (6), 322–6. [PubMed: 15922829]
49. Piriyaopongsa J et al. (2007) Origin and evolution of human microRNAs from transposable elements. *Genetics* 176 (2), 1323–37. [PubMed: 17435244]
50. Creasey KM et al. (2014) miRNAs trigger widespread epigenetically activated siRNAs from transposons in *Arabidopsis*. *Nature* 508 (7496), 411–5. [PubMed: 24670663]
51. Venkatesh T et al. (2016) tRFs: miRNAs in disguise. *Gene* 579 (2), 133–8. [PubMed: 26743126]
52. Babiarz JE et al. (2008) Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes Dev* 22 (20), 2773–85. [PubMed: 18923076]
53. Tosar JP et al. (2018) Non-coding RNA fragments account for the majority of annotated piRNAs expressed in somatic non-gonadal tissues. *Communications Biology* 1 (1), 2. [PubMed: 30271890]
54. Honda S et al. (2017) The biogenesis pathway of tRNA-derived piRNAs in *Bombyx* germ cells. *Nucleic Acids Res* 45 (15), 9108–9120. [PubMed: 28645172]

55. Borges F et al. (2018) Transposon-derived small RNAs triggered by miR845 mediate genome dosage response in Arabidopsis. *Nat Genet* 50 (2), 186–192. [PubMed: 29335544]
56. Surbanovski N et al. (2016) A highly specific microRNA-mediated mechanism silences LTR retrotransposons of strawberry. *Plant J* 85 (1), 70–82. [PubMed: 26611654]
57. Rebollo R et al. (2012) Transposable elements: an abundant and natural source of regulatory sequences for host genes. *Annu Rev Genet* 46, 21–42. [PubMed: 22905872]
58. Schimmel P (2018) The emerging complexity of the tRNA world: mammalian tRNAs beyond protein synthesis. *Nat Rev Mol Cell Biol* 19 (1), 45–58. [PubMed: 28875994]
59. Coffin JM et al. (1997) The Interactions of Retroviruses and their Hosts In *Retroviruses* (Coffin JM et al. eds).
60. Kelly NJ et al. (2003) Selection of retroviral reverse transcription primer is coordinated with tRNA biogenesis. *J Virol* 77 (16), 8695–701. [PubMed: 12885888]
61. Telesnitsky A and Wolin SL (2016) The Host RNAs in Retroviral Particles. *Viruses* 8 (8).
62. Kleiman L et al. (2010) Formation of the tRNALys packaging complex in HIV-1. *FEBS Lett* 584 (2), 359–65. [PubMed: 19914238]
63. Panet A et al. (1975) Specific binding of tryptophan transfer RNA to avian myeloblastosis virus RNA-dependent DNA polymerase (reverse transcriptase). *Proc Natl Acad Sci U S A* 72 (7), 2535–9. [PubMed: 52156]
64. Liu S et al. (2010) Initiation complex dynamics direct the transitions between distinct phases of early HIV reverse transcription. *Nat Struct Mol Biol* 17 (12), 1453–60. [PubMed: 21102446]
65. Barski A et al. (2010) Pol II and its associated epigenetic marks are present at Pol III-transcribed noncoding RNA genes. *Nat Struct Mol Biol* 17 (5), 629–34. [PubMed: 20418881]
66. Zaratiegui M (2017) Cross-Regulation between Transposable Elements and Host DNA Replication. *Viruses* 9 (3).
67. Albers S and Czech A (2016) Exploiting tRNAs to Boost Virulence. *Life (Basel)* 6 (1).
68. Xu H and Boeke JD (1990) Host genes that influence transposition in yeast: the abundance of a rare tRNA regulates Ty1 transposition frequency. *Proc Natl Acad Sci U S A* 87 (21), 8360–4. [PubMed: 2172984]
69. Esposito F et al. (1985) Pseudouridine excretion and transfer RNA primers for reverse transcriptase in tumors of retroviral origin. *Cancer Res* 45 (12 Pt 1), 6260–3. [PubMed: 2415241]
70. Genencher B et al. (2018) Mutations in Cytosine-5 tRNA Methyltransferases Impact Mobile Element Expression and Genome Stability at Specific DNA Repeats. *Cell Reports* 22 (7), 1861–1874. [PubMed: 29444437]
71. Molla-Herman A et al. (2015) tRNA processing defects induce replication stress and Chk2-dependent disruption of piRNA transcription. *EMBO J* 34 (24), 3009–27. [PubMed: 26471728]
72. Goodarzi H et al. (2016) Modulated Expression of Specific tRNAs Drives Gene Expression and Cancer Progression. *Cell* 165 (6), 1416–1427. [PubMed: 27259150]
73. Megel C et al. (2015) Surveillance and cleavage of eukaryotic tRNAs. *Int J Mol Sci* 16 (1), 1873–93. [PubMed: 25599528]
74. Hasler D et al. (2016) The Lupus Autoantigen La Prevents Mis-channeling of tRNA Fragments into the Human MicroRNA Pathway. *Mol Cell* 63 (1), 110–24. [PubMed: 27345152]
75. Gogakos T et al. (2017) Characterizing Expression and Processing of Precursor and Mature Human tRNAs by Hydro-tRNAseq and PAR-CLIP. *Cell Rep* 20 (6), 1463–1475. [PubMed: 28793268]
76. Zheng G et al. (2015) Efficient and quantitative high-throughput tRNA sequencing. *Nat Methods* 12 (9), 835–837. [PubMed: 26214130]
77. Alves CS et al. (2017) Genome-wide identification and characterization of tRNA-derived RNA fragments in land plants. *Plant Mol Biol* 93 (1–2), 35–48. [PubMed: 27681945]
78. Couvillion MT et al. (2009) Sequence, biogenesis, and function of diverse small RNA classes bound to the Piwi family proteins of *Tetrahymena thermophila*. *Genes Dev* 23 (17), 2016–32. [PubMed: 19656801]
79. Kuhn CD et al. (2015) On-enzyme refolding permits small RNA and tRNA surveillance by the CCA-adding enzyme. *Cell* 160 (4), 644–58. [PubMed: 25640237]

80. Schimmel P and Ribas de Pouplana L (1995) Transfer RNA: from minihelix to genetic code. *Cell* 81 (7), 983–6. [PubMed: 7600584]
81. Moelling K et al. (2017) RNase H As Gene Modifier, Driver of Evolution and Antiviral Defense. *Front Microbiol* 8, 1745. [PubMed: 28959243]
82. Doi N et al. (2003) Short-interfering-RNA-mediated gene silencing in mammalian cells requires Dicer and eIF2C translation initiation factors. *Curr Biol* 13 (1), 41–6. [PubMed: 12526743]
83. David M et al. (1982) Bacteriophage T4-induced anticodon-loop nuclease detected in a host strain restrictive to RNA ligase mutants. *Proc Natl Acad Sci U S A* 79 (23), 7097–101. [PubMed: 6296815]
84. Kano-Sueoka T and Sueoka N (1968) Characterization of a modified leucyl-tRNA of *Escherichia coli* after bacteriophage T2 infection. *J Mol Biol* 37 (3), 475–91. [PubMed: 4888151]
85. Wilusz JE et al. (2008) 3' end processing of a long nuclear-retained noncoding RNA yields a tRNA-like cytoplasmic RNA. *Cell* 135 (5), 919–32. [PubMed: 19041754]
86. Castel SE and Martienssen RA (2013) RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. *Nat Rev Genet* 14 (2), 100–12. [PubMed: 23329111]
87. Zhang X et al. (2016) IL-4 Inhibits the Biogenesis of an Epigenetically Suppressive PIWI-Interacting RNA To Upregulate CD1a Molecules on Monocytes/Dendritic Cells. *J Immunol* 196 (4), 1591–603. [PubMed: 26755820]
88. Rebollo R et al. (2011) Retrotransposon-induced heterochromatin spreading in the mouse revealed by insertional polymorphisms. *PLoS Genet* 7 (9), e1002301. [PubMed: 21980304]
89. Ecco G et al. (2017) KRAB zinc finger proteins. *Development* 144 (15), 2719–2729. [PubMed: 28765213]
90. Burns KH (2017) Transposable elements in cancer. *Nat Rev Cancer* 17 (7), 415–424. [PubMed: 28642606]
91. Ross RJ et al. (2014) PIWI proteins and PIWI-interacting RNAs in the soma. *Nature* 505 (7483), 353–359. [PubMed: 24429634]
92. Ono M and Ohishi H (1983) Long terminal repeat sequences of intracisternal A particle genes in the Syrian hamster genome: identification of tRNAPhe as a putative primer tRNA. *Nucleic Acids Res* 11 (20), 7169–79. [PubMed: 6314280]
93. Wolf G et al. (2015) The KRAB zinc finger protein ZFP809 is required to initiate epigenetic silencing of endogenous retroviruses. *Genes Dev* 29 (5), 538–54. [PubMed: 25737282]
94. Baust C et al. (2003) Structure and expression of mobile ETnII retroelements and their coding-competent MusD relatives in the mouse. *J Virol* 77 (21), 11448–58. [PubMed: 14557630]
95. Benit L et al. (1997) Cloning of a new murine endogenous retrovirus, MuERV-L, with strong similarity to the human HERV-L element and with a gag coding sequence closely related to the Fv1 restriction gene. *J Virol* 71 (7), 5652–7. [PubMed: 9188643]
96. Peters G and Glover C (1980) tRNA's and priming of RNA-directed DNA synthesis in mouse mammary tumor virus. *J Virol* 35 (1), 31–40. [PubMed: 6157839]
97. Harada F et al. (1979) The primer tRNA for Moloney murine leukemia virus DNA synthesis. Nucleotide sequence and aminoacylation of tRNAPro. *J Biol Chem* 254 (21), 10979–85. [PubMed: 115865]
98. Colicelli J and Goff SP (1986) Isolation of a recombinant murine leukemia virus utilizing a new primer tRNA. *J Virol* 57 (1), 37–45. [PubMed: 2416955]
99. Barat C et al. (1989) HIV-1 reverse transcriptase specifically interacts with the anticodon domain of its cognate primer tRNA. *EMBO J* 8 (11), 3279–85. [PubMed: 2479543]
100. Dahlberg JE et al. (1974) Transcription of DNA from the 70S RNA of Rous sarcoma virus. I. Identification of a specific 4S RNA which serves as primer. *J Virol* 13 (5), 1126–33. [PubMed: 4132919]
101. Slotkin RK and Martienssen R (2007) Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet* 8 (4), 272–85. [PubMed: 17363976]
102. Kloc A et al. (2008) RNA interference guides histone modification during the S phase of chromosomal replication. *Curr Biol* 18 (7), 490–5. [PubMed: 18394897]

103. Dewannieux M et al. (2003) LINE-mediated retrotransposition of marked Alu sequences. *Nat Genet* 35 (1), 41–8. [PubMed: 12897783]
104. Garcia-Perez JL et al. (2007) Distinct mechanisms for trans-mediated mobilization of cellular RNAs by the LINE-1 reverse transcriptase. *Genome Res* 17 (5), 602–11. [PubMed: 17416749]
105. Curcio MJ and Derbyshire KM (2003) The outs and ins of transposition: from mu to kangaroo. *Nat Rev Mol Cell Biol* 4 (11), 865–77. [PubMed: 14682279]
106. Chen Q et al. (2016) Sperm tsRNAs contribute to intergenerational inheritance of an acquired metabolic disorder. *Science* 351 (6271), 397–400. [PubMed: 26721680]
107. Kim HK et al. (2017) A transfer-RNA-derived small RNA regulates ribosome biogenesis. *Nature* 552 (7683), 57–62. [PubMed: 29186115]

Outstanding Questions Box

- Which endonucleases cut 3'-tRFs from mature tRNAs? How is the highly tissue-specific ratio of 18 *versus* 22 nt 3'-tRFs (tRF-3a/tRF-3b) achieved?
- How does 5'-tRF production relate to 3'-tRF production? Does tRNA expression level affect tRF cleavage rate?
- Are there host proteins that stabilize 18 nt 3'-tRFs in viral particles and mediate interaction with the genomic virus RNA or do tRF-3a directly bind to viral proteins when they inhibit RT?
- Are 3'-tRFs able to guide transcriptional silencing by H3K9me3 at the PBS?
- What triggers 3'-tRNAs cleavage? What are the costs for proliferation for the cell?
- Do 3'-tRFs regulate LTR-retroelements in all eukaryotes and can they be used to restrict virus infection?
- When do tRFs interact with AGO/PIWI proteins and which other RNAi factors are involved? Were tRNAs an early substrate of the RNAi machinery?

Highlights

- tRNA fragments (tRF) are highly abundant in stem cells and other cell types undergoing epigenetic reprogramming
- tRFs bind to AGO and PIWI proteins and guide repression of complementary mRNAs similar to piRNA and endo-siRNA
- 3'-tRFs include the post-transcriptional 3'-CCA tail of mature tRNAs and target the highly conserved tRNA primer binding site (PBS) of LTR-retrotransposons
- 3'-tRFs therefore allow the host to discriminate self from non-self without repressing domesticated transposon genes and promoters
- 3'-tRFs inhibit retrotransposition of murine endogenous retroviruses as well as replication of human LTR-retroviruses and most likely any LTR-retroelement which has a PBS and is potentially mobile
- 18 nt 3'-tRFs (tRF-3a) specifically block reverse transcription; 22 nt 3'-tRFs (tRF-3b) induce RNAi

Box 1:**Transposable Element Classes and Families**

Prokaryotes, Archaea, and Eukaryotes all have transposable elements. The genomes of complex organisms contain large amounts of transposon sequences (Figure I) which can drive transcription or induce heterochromatin at neighboring genes depending on developmental and cell cycle cues [101, 102]. Transposons direct chromosome organization and imprinting, but also large gene regulatory networks due to their high copy numbers. At the same time, their mobility is mutagenic and needs to be tightly controlled by the host. Usually, DNA and histone modifications repress their activity but during epigenetic reprogramming transposons are released and transcribed which is when small RNA-mediated silencing mechanisms become crucial to limit transposition.

Class II, DNA transposons, cut and paste their DNA into a new genomic locus, while class I, RNA transposons or retrotransposons, use reverse transcriptase (RT) to copy their RNA into DNA and paste it into the genome. Non-LTR retrotransposons nick the target genomic DNA and use that 3'-end to prime reverse transcription. With a few exceptions, LTR-retroelements use the 3'-end of tRNAs to prime reverse transcription and are therefore susceptible to targeting by 3'-tRNA fragments. Which families of transposons are most prevalent in the genome depends on the organism. For example, transposon sequences in prokaryotes, zebrafish (*Danio rerio*) and worm (*Caenorhabditis elegans*) are dominated by DNA-transposons. *Arabidopsis thaliana* has similar amounts of DNA- and RNA-transposons, whereas in mammals, flies (*Drosophila melanogaster*), and maize (*Zea mays*) RNA-transposons comprise the vast majority of repetitive sequences. Abundant transposon families can be "old" and inactivated by mutations, while less abundant elements that entered the host more recently may be highly active, such as P-element DNA-transposons in flies. No active DNA-transposons have been found in mouse and humans, instead RNA-transposons contribute to ongoing mutagenesis. Active transposons recognize and co-mobilize non-coding family members that lack one or all open reading frames (Figure I), and are termed 'non-autonomous'. Non-autonomous LTR-retroelements still require an intact tRNA primer binding site (PBS) to pass reverse transcription. Oftentimes, a few autonomous, coding copies in the genome mobilize large numbers of non-autonomous, non-coding elements. An example is the murine MusD transposon moving non-coding ETn elements (Table 1) [15], but also human LINE transposons carrying over e.g. SINEs that lack an enzymatic machinery [103, 104]. For an overview on mechanisms of transposon mobility please refer to [105].

Figure I. Genomic Abundance and Domain Structure of Retroelements.

(A) The proportion of retrotransposons in the genome of *Homo sapiens*, *Mus musculus*, and *Zea mays* are shown. (B) Retroelements can be divided into non-LTR (DNA-target primed) and LTR-elements (tRNA-primed reverse transcription). An example of a non-LTR retrotransposon, an active human long interspersed nuclear element (LINE) (top), and the most prevalent superfamilies of LTR-retroelements (bottom) are shown along with their genomic structures. LTR-retrotransposons of the ubiquitous Copia and Gypsy superfamilies, first observed in *Drosophila*, differ in the order of domains within the Pol gene; Gypsy elements can also contain an envelope gene. Endogenous retroviruses (ERV) in mammals are closely related to LTR-retroviruses with the major difference being that most ERVs have lost their envelope gene or accumulated mutations. Thus, ERVs cannot leave the cell and are not infectious, but instead form virus-like particles within the cell. ERVs derived from betaretroviruses (depicted here, e.g. IAP, MusD) have a separate PR ORF and comprise the most active transposon superfamily in mice. LTR-retroviruses (e.g. Lentiviruses) often have several additional ORFs and differ in whether they have the PR domain in one ORF with Gag, with Pol, or separate. Rectangles represent open reading frames (ORF); ORF names are annotated below rectangles, domains in rectangles. EN: endonuclease, RT: reverse transcriptase, LTR: long terminal repeat, Gag: group-antigen, PR: protease, RH: RNaseH, IN: Integrase, Pol: polymerase, Env: envelope. Components are not drawn to scale.

Box 2:**tRNA Fragments, Binding Partners, and Biological Targets**

tRF nomenclature [1] is illustrated in Figure II. Internal tRF-2 fragments bind to YBX-1 and inhibit translation [25], as do 5'-halves cleaved by Angiogenin (ANG) [26]. ANG or RNY1 (*Saccharomyces cerevisiae*) cleavage is often stress-induced and results in a 2',3'-cyclic phosphate at the 3'-end of 5'-halves which are also called tiRNAs [20]. Under stress conditions mitochondria release cytochrome C (CYC) binding 5'-halves and counteract apoptosis [24]. PIWI-bound 5'-fragments or halves can be ANG-independent or ANG-dependent, are 25–35 nt in length [54, 87] and are often called tRNA-derived piRNAs (td-piRNA). They bind human HIWI and HIWI2 [32, 87] as well as *Bombyx mori* Siwi and BmAgo3 [54]. Select 5'-tRFs are paternally inherited through sperm and change gene expression in the offspring by a yet unknown mechanism [45, 106]. tRF-5b fragments bind to all four human AGO proteins (AGO3, 4 > AGO1 > AGO2) and guide AGO1 to complementary target mRNAs obeying seed-sequence rules of miRNA, which means that base-pairing of the seed sequence nucleotides 2–8 with their target is important and suggests an overlap with canonical miRNA pathways [31]. In *Arabidopsis*, tRF-5b fragments have been reported to guide cleavage of Gypsy mRNAs in pollen, though in this case while tolerating mismatches in the seed sequence [47]. In humans, Leu-CTG tRF-3b binds two ribosomal protein mRNAs enhancing ribosome biosynthesis [107], while DICER- and AGO-dependent Gly-GGC tRF-3b downregulates replication protein RPA1 [34]. 3'-tRFs include the post-transcriptional 3'-CCA tail of mature tRNAs and therefore target replication-competent LTR-retroelements at their tRNA primer binding site (PBS). tRF-3b fragments reduce RNA and protein levels of LTR-retrotransposons in a miRNA-like fashion, while shorter tRF-3a fragments block reverse transcription [39]. tRF-3a bind AGO proteins [28, 31] and in the ciliate *Tetrahymena*, binding of the PIWI protein TWI12 triggers translocation to the nucleus and activation of XRN2 exonuclease-mediated RNA decay [35]. Cleavage of tRFs by DICER is an exception rather than the rule and tRNA halves can still be observed after ANG knock-down [1]. Many published small RNA sequencing data sets have been size selected for 21–22nt miRNA during library preparation and hence exclude tRFs <19 nt, CCA-tails and RNA modifications from their mature, parent tRNA and require extra consideration during sequence analysis of tRFs.

Figure II. tRNA Fragment Biogenesis and Function.

Fragments from mature tRNAs include (counter-clockwise) 19–24 nt 5'-tRFs (tRF-5b), 25–34 nt 5'-tRFs (tRF-5c), ~35 nt 5'-halves, ~30 nt internal fragments (tRF-2), 22 nt 3'-tRFs (tRF-3b) and 17–19 nt 3'-tRFs (tRF-3a). Cleavage is mediated by endonucleases (pac-man symbols) that are largely unknown (“?”), however, a few individual fragments of the tRF-5b, tRF-3b, and tRF-3a type depend on the RNaseIII-type endoribonuclease DICER. 5'-halves are generated by Angiogenin (ANG, or RNY1 in yeast) of the RNaseA family. After cleavage, halves and fragments exert their functions through binding to effector molecules. Known binding partners depicted here are AGO, PIWI, cytochrome c (CYC) released from mitochondria, and YBX-1. YBX-1 binding of 5'-halves and tRF-2 inhibits translation at the ribosome (R). Ribosome biosynthesis can be increased (plus symbols) by tRF-3b regulation. tRF-5 and tRF-3 fragments guide recognition of complementary RNA by AGO and PIWI proteins that cleave target mRNA or stall its translation. RPA1 and other specific mRNA targets are discussed in Box 2. tRF-3 fragments inhibit LTR-retroelements through recognition at their tRNA primer binding site (PBS). Short tRF-3a fragments inhibit priming by reverse transcriptase (RT) in retroviral particles and interact with AGO/PIWI proteins. *Tetrahymena* PIWI loaded with tRF-3a triggers XRN2-mediated RNA decay in the nucleus. Fragments of the tRF-5c type are inherited through sperm. For further details on functions of tRFs and references please see Box text.

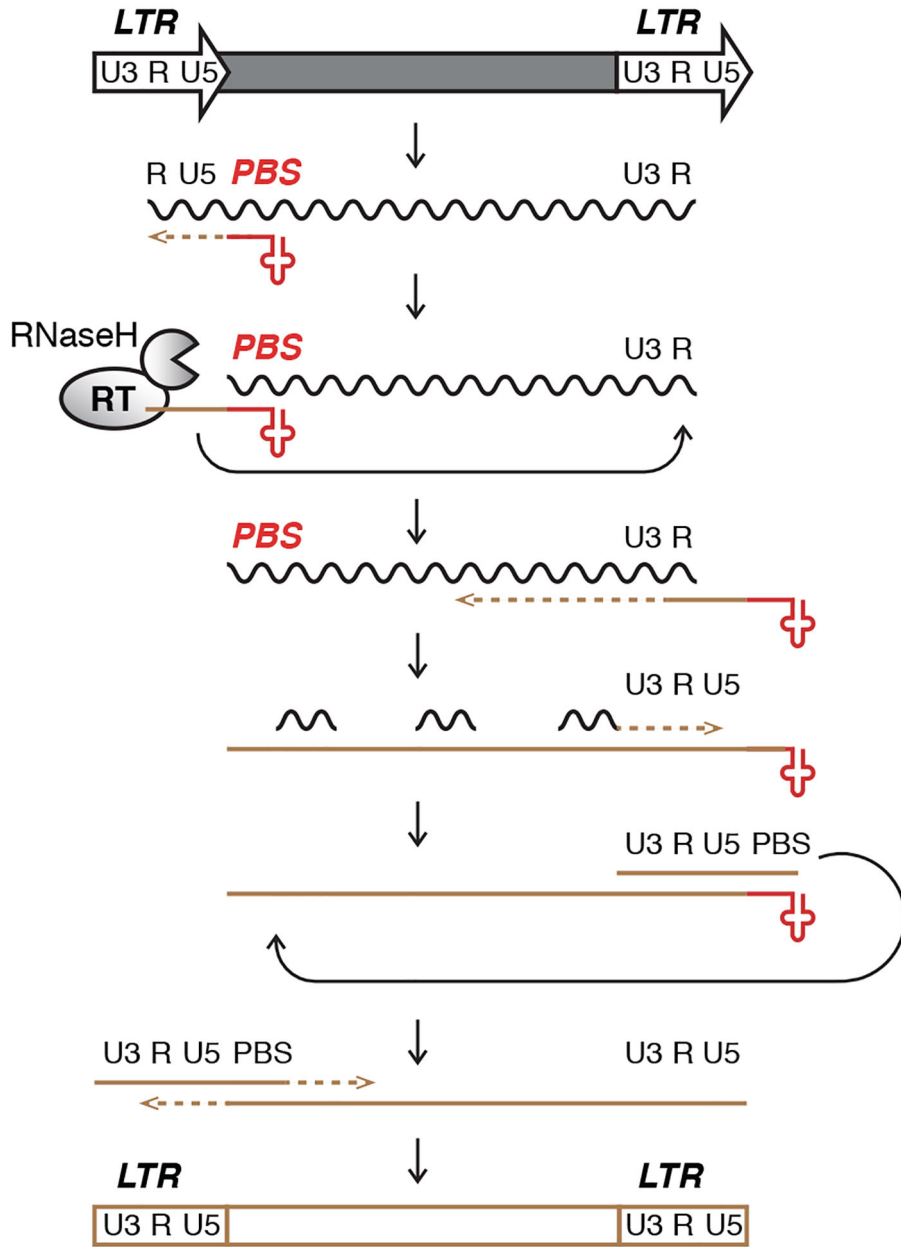


Figure 1. Reverse Transcription of LTR-Retrotransposons and -Viruses. Long terminal repeats (LTR) encode promoter elements and termination signals. The RNA transcript contains a region repeated at either end (R), a segment unique to the 5' end of the RNA (U5), and a segment only included at the 3'-end of the RNA (U3). The 3'-end of cellular tRNAs (red cloverleaf) primes reverse transcription by hybridizing to the primer binding site (PBS). While this segment is being copied into first-strand cDNA (brown line), also called (-)ssDNA (minus strong stop DNA), the RNaseH activity of RT degrades the template RNA. The elongating cDNA is transferred to the 3'-end of the retrotransposon transcript by hybridizing to the R region. The remaining RNA is partially degraded by RNaseH leaving behind primers for second-strand cDNA synthesis. After another transfer

event, first- and second-strand synthesis are completed to result in a full-length, double-stranded retroviral DNA that will be integrated into the host genome.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

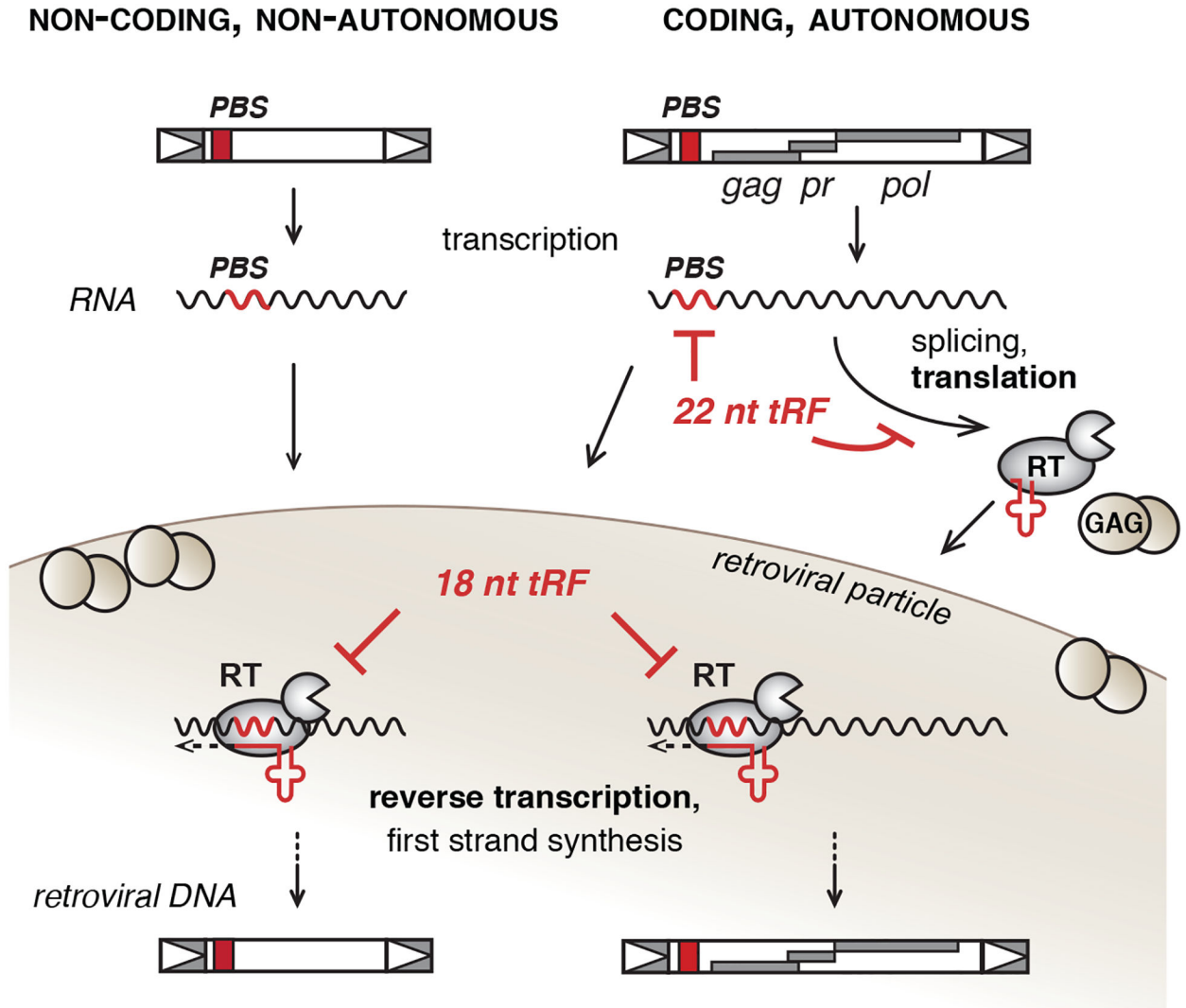


Figure 2. Model of LTR-Retroelement Silencing by 3'-tRFs.

LTR-retroelements comprise LTR-retrotransposons and LTR-retroviruses. In the absence of epigenetic suppression, non-coding and coding LTR-retroelements are transcribed. Elements that encode the enzymatic machinery for retroviral replication are autonomous and in the first step their RNA is spliced and translated. 22 nt 3'-tRFs (tRF3-b) target coding-competent LTR-retroelements at the level of retroviral protein production through post-transcriptional gene silencing. Retroviral gene products of the autonomous, coding element constitute all components of the retroviral particle and recruit virus “genomic” unspliced RNA template of both coding and non-coding family members as well as host-dependent co-factors like tRNAs to the particle. 18 nt 3'-tRFs (tRF3-a) interfere with reverse transcription at the level of first-strand synthesis and therefore inhibit both coding and noncoding, non-autonomous retrotransposons. 3'-tRFs specifically promote silencing of any potentially mobile LTR-retroelement that has maintained a functional PBS. PBS, primer binding site; RT, reverse transcriptase.

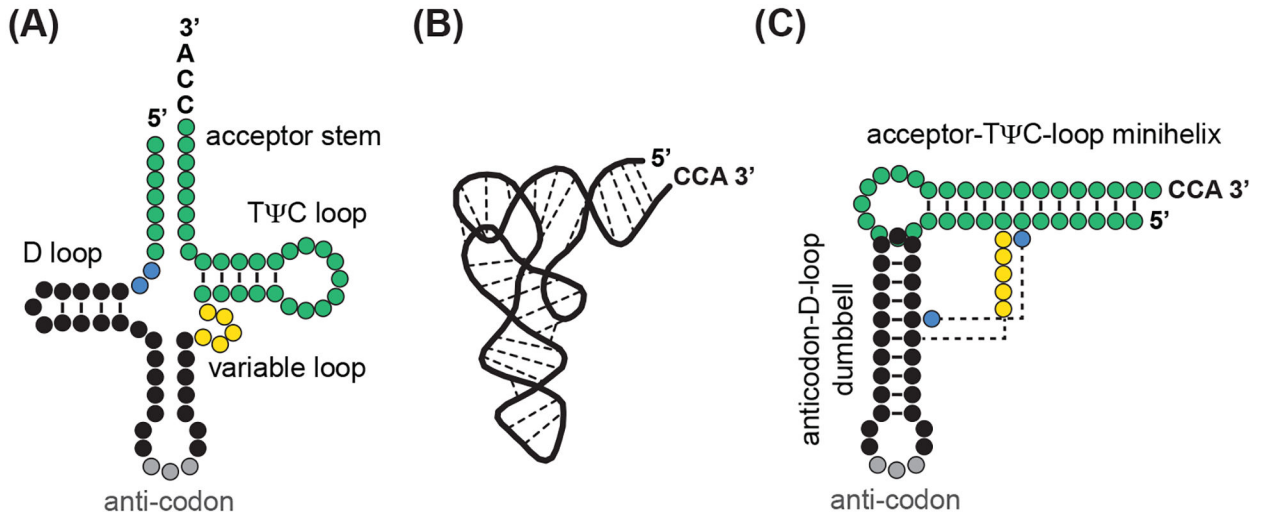


Figure 3. tRNA Structure and Domains

Three alternative representations of the same tRNA, here of tRNA Lys-AAA with the anticodon-sequence UUU, illustrate its functional domains and tertiary structure. (A) The typical cloverleaf diagram provides an overview of the major domains and loops. The acceptor stem has the post-transcriptional CCA-tag at its 3'-end that will be covalently linked with the amino-acid prior to translation. (B) The three-dimensional L-shaped structure of tRNA is composed of two double-stranded helices: (C) the anticodon-D-loop (black/grey) that serves proof-reading functions during translation at the ribosome and the upper minihelix (green) consisting of the acceptor-stem and the TΨC-loop. (Figure modified from [58]).

Table 1.Examples of tRNAs Priming LTR-Retroviruses/Transposons.^a

tRNA-codon primer	retro-transposon/virus family	18 nt 3'end of tRNA sequence complementary to PBS	Reference	virus genus
Phe-TTY	IAP (Intracisternal A-Particle)	TCCCGGGTTTCGGCACCA	[92]	Betaretrovirus
Phe-TTY	MMERVK10C (Mus Musculus ERV using tRNA ^{Lys} type 10C)	TCCCGGGTTTCGGCACCA	[93]	Betaretrovirus
Lys-AAA	ETn (Early Transposon)	GTCCCTGTTTCGGGCGCCA	[94]	Betaretrovirus
Lys-AAA	MusD (Mus Musculus typ D provirus), ETn family	GTCCCTGTTTCGGGCGCCA	[94]	Betaretrovirus
Leu-TTA	MERVL (mouse ERV-L)	ACCCCACTTCTGGTACCA	[95]	Spumavirus
Lys-AAA	MMTV (mouse mammary tumor virus)	GTCCCTGTTTCGGGCGCCA	[96]	Betaretrovirus
Pro-CCN	MMuLV (or MMLV, Moloney Murine Leukemia Virus)	ATCCCGGACGAGCCCCCA	[97]	Gamma-retrovirus
Gln-CAG	MMuLV	ATCTCGGTGGAACCTCCA	[98]	Gamma-retrovirus
Lys-AAA	HIV (Human Immunodeficiency Virus)	GTCCCTGTTTCGGGCGCCA	[99]	Lentivirus
Pro-CCN	HTLV-1 (human T-cell leukemia virus type 1)	ATCCCGGACGAGCCCCCA	[36]	Deltaretrovirus
Trp-TGG	RSV (Rous Sarcoma Virus)	ATCACGTCGGGGTCACCA	[100]	Alpha-retrovirus

^aThe original nomenclature of endogenous retrovirus families in mouse and humans was based on their cognate tRNA primer, e.g. ERV-K(Lys), ERV-L(Leu), but beware of many outliers within all groups. For example, IAP (ERV-K family) uses tRNA-Phe and even though MMERVK10C is called “*Mus musculus* ERV using tRNA-Lys type 10C” its full-length genomic copies have a PBS binding site for tRNA-Phe. A more comprehensive list of mouse ERV PBS sequences can be found in [93]. Members of one transposon/virus family often have polymorphic PBS sequences with mutations compared to the ideal tRNA binding site. Note that tRNAs with a specific codon can include different isodecoders with variable 3'-ends and stem sequences, indicated by numbers in addition to the codon triplet letters [58]. For example, retrotransposons of the betaretrovirus family like IAP, ETn/MusD, MMTV, and HIV are primed by the Lys³-AAA isodecoder, MMuLV by isodecoders Pro^{1,2}. For details on the concept of isodecoders, isoacceptors, and fundamentals of tRNA biology please refer to [58]. Viruses can also adapt to use alternative tRNAs for priming, an example is a MMuLV element that binds Gln-CAG [98]. Endogenous, intracellular retroviruses are shaded grey and tRNA primers are predicted from their PBS sequence. For infectious retroviruses with an extracellular life cycle, references point to experimentally validated primer tRNA from viral particles or bound to RT. The Lowe UCSC tRNA database offers a curated collection of genomic tRNA sequences in mouse and human but is currently lacking some isoacceptors, while the more comprehensive RepeatMasker annotation (at repeatmasker.org or the Table Browser of UCSC) includes pseudogenes and degenerate tRNA sequences.