3-2024

Advances in Methods for tRNA Sequencing and Quantification

Nigam H. Padhiar

Upendra Katneni

Anton A. Komar

Yuri Motorin

Chava Kimchi-Sarfaty

CellPress

Trends in Genetics

Review

Advances in methods for tRNA sequencing and quantification

Nigam H. Padhiar,[1] Upendra Katneni,[1] Anton A. Komar,[2] Yuri Motorin,[3,*] and Chava Kimchi-Sarfaty[1,*]

In the past decade tRNA sequencing (tRNA-seq) has attracted considerable attention as an important tool for the development of novel approaches to quantify highly modified tRNA species and to propel tRNA research aimed at understanding the cellular physiology and disease and development of tRNA-based therapeutics. Many methods are available to quantify tRNA abundance while accounting for modifications and tRNA charging/acylation. Advances in both library preparation methods and bioinformatic workflows have enabled developments in next-generation sequencing (NGS) workflows. Other approaches forgo NGS applications in favor of hybridization-based approaches. In this review we provide a brief comparative overview of various tRNA quantification approaches, focusing on the advantages and disadvantages of these methods, which together facilitate reliable tRNA quantification.

Highlights

Key challenges in tRNA next-generation sequencing (NGS) library preparations include efficient adapter ligation and overcoming extensive reverse transcriptase-blocking modifications and a highly stable RNA secondary structure.

Computational challenges include accounting for the potential modifications in each tRNA and uniquely mapping each read to a reference, which can comprise hundreds of highly similar tRNA genes.

Several recent applications forgo tRNA-seq in favor of hybridization-based approaches; however, hybridization-based approaches continued to be developed and improved.

Validation approaches remain an issue among tRNA-seq studies – it is difficult to directly compare tRNA NGS datasets given their varied library preparation approaches, and many non-NGS based methods of corroboration are relatively low throughput.

The importance and challenges of tRNA evaluation

tRNAs are implicated in many processes, including translation [1], amino acid metabolism [2,3], metabolite synthesis [4], priming the reverse transcription (RT) of viral RNAs [5], and tRNA cleavage leading to the production of regulatory noncoding RNAs termed tRNA-derived small RNAs (tsRNAs) [6]. Given its multifaceted involvement in cellular physiology, tRNA dysregulation affecting their abundance and function may lead to various human diseases including different types of cancers [7], leukodystrophy [8], cell-cycle arrest [9], and a wide array of diseases related to cellular energetics [7], as well to as diseases of other species such as the deadly rice fungus *Magnaporthe oryzae* [10]. Thus, methods allowing reliable and accurate tRNA quantification are becoming of immense importance for understanding both the normal functioning of the cell and various pathophysiological conditions.

Tools for the quantification of tRNAs can be broadly divided into NGS-based approaches and non-NGS approaches (mostly hybridization). Regardless of the approach, there are many challenges for the accurate assessment of tRNA levels, given the distinctive highly conserved secondary structure and chemical modifications of tRNA molecules. tRNAs are heavily modified, and feature several base and ribose methylations, as well as pseudouridine (ψ) and other complex modifications [11]. These RNA modifications can lead to premature RT stops in the process of complementary DNA (cDNA) synthesis [12,13] in NGS applications and can also hinder hybridization in hybridization-based approaches as a result of impaired Watson–Crick base-pairing with oligonucleotide probes [14,15]. In addition, owing to the variable nature of modifications and the overall diversity of highly similar but still unique tRNA species, which can exceed 400 unique molecules in humans [16], special considerations must be taken into account during quantification to ensure accuracy in differentiating between isodecoders or in mapping to the genome.

In the following we review select protocols in the tRNA quantification space with a particular focus on NGS approaches, and discuss the associated challenges. These include demethylase-thermostable group II intron RT tRNA sequencing (DM-TGIRT-seq) [17], AlkB-facilitated RNA methylation sequencing (ARM-seq) [18], Hydro-tRNAseq [19], Y-shaped adapter-ligated mature tRNA sequencing

[1]Hemostasis Branch 1, Division of Hemostasis, Office of Plasma Protein Therapeutics, Office of Therapeutic Products, Center for Biologics Evaluation and Research, Food and Drug Administration, Silver Spring, MD, USA
[2]Department of Biological, Geological, and Environmental Sciences, Center for Gene Regulation in Health and Disease, Cleveland State University, Cleveland, OH, USA
[3]CNRS-Université de Lorraine, UAR 2008, IBSLor UMR 7365 IMoPA, Nancy, France

*Correspondence:
yuri.motorin@univ-lorraine.fr (Y. Motorin) and Chava.kimchi-sarfaty@fda.hhs.gov (C. Kimchi-Sarfaty).

 https://doi.org/10.1016/j.tig.2023.11.001
Published by Elsevier Ltd.

(YAMAT-seq) [20], long hairpin oligonucleotide-based tRNA high-throughput sequencing (LOTTE-tRNAseq) [21], quantitative mature tRNA sequencing (QuantM-tRNAseq) [12], modification-induced misincorporation tRNA sequencing (mim-tRNAseq) [16], Nano-tRNAseq [22], and adapter-ligated libraries of tRNA-derived sequences (ALL-tRNAseq) [23].

Different library preparation methods in tRNA-seq

In the following we present an overview of library preparation strategies in tRNA-seq; these are summarized in Figure 1 (Key figure).

Unbiased tRNA extraction

The first essential step for tRNA quantification analysis is the isolation of RNA. Depending on the properties of the biological material, extraction of total RNA can be achieved either by general extraction protocols, such as phenol extraction used for bacteria or TRIzol™ extraction that is common for higher eukaryotic cells in culture [24]. Difficult-to-disrupt cells such as *Saccharomyces cerevisiae* require harsher procedures such as acid phenol extraction [25]. For many bacterial cells, simple TRIzol extraction often results in a highly enriched tRNA fraction [26].

Few selected protocols discussed in this review have a specific focus on tRNA in the RNA isolation step. DM-TGIRT-seq, ARM-seq, and ALL-tRNAseq utilize the mirVana™ miRNA isolation kit which is centered on an efficient glass fiber filter (GFF)-based approach. Other protocols initially rely on common TRIzol™ or TRIsure™ extraction and later employ specific tRNA isolation techniques.

Pretreatment of input RNA

An appropriate library preparation protocol plays an essential role in preparing input RNA for tRNA quantification by deep sequencing. In some protocols the input RNA is treated before proceeding with adapter ligation. Hydro-tRNAseq uniquely employs limited alkaline hydrolysis of the purified tRNA to isolate fragments of 19–35 nt that will have a less complex secondary structure and fewer modifications that could induce RT stops [19]. However, fragments carrying RT hindering modifications can be overlooked in this approach owing to generation of short abortive cDNAs (see later) [14,16]. DM-TGIRT-seq, ALL-tRNAseq, and ARM-seq all incorporate demethylation of input tRNA as key steps in their protocols, and the first two use a combination of wild-type *Escherichia coli* alkylation B (wtAlkB) and D135S mutant AlkB, and the latter only uses wtAlkB. This mutant AlkB is able to demethylate *N*2,*N*2-dimethylguanosine ($m^{2,2G}$), which is not an ideal substrate for wtAlkB [27]. Altogether, these modifications improve RT read-through if a highly processive enzyme such as MarathonRT or TGIRT is not used, but will eliminate the potential to detect such tRNA modifications through misincorporation or RT-stop analysis, unless an untreated sample is also sequenced for comparison (which the authors of DM-TGIRT-seq and ARM-seq did). These protocols can also introduce biased representation since only a subset of such modified residues are effectively removed [28].

Adapter ligation strategies

As a result of the maturation pathways of tRNAs, they mostly have 5′-P and 3′-OH extremities [29] that are directly compatible with ligation of adapters required for sequencing. It is important to note that efficient ligation to 3′-OH extremities requires preliminary deacylation of tRNAs because the presence of an attached amino acid at the tRNA 3′-CCA end compromises the ligation of adapter [30]. The adapter ligation protocols in our cohort can be split into four broad categories, as follows: strategy A, separate ligation of 3′ and 5′ adapters; strategy B, 3′ adapter ligation followed by RT, which is then followed by single-stranded (ss)DNA ligation of DNA oligonucleotide to the 3′ extremity of the cDNA after RT; strategy C, 3′ adapter ligation followed by circularization of the cDNA; and strategy D, incorporation of both 5′ and 3′ adapters (and priming sites) into a double-stranded oligonucleotide (Figure 2).

Key figure

Step-by-step library preparation methods for nine tRNA-seq protocols

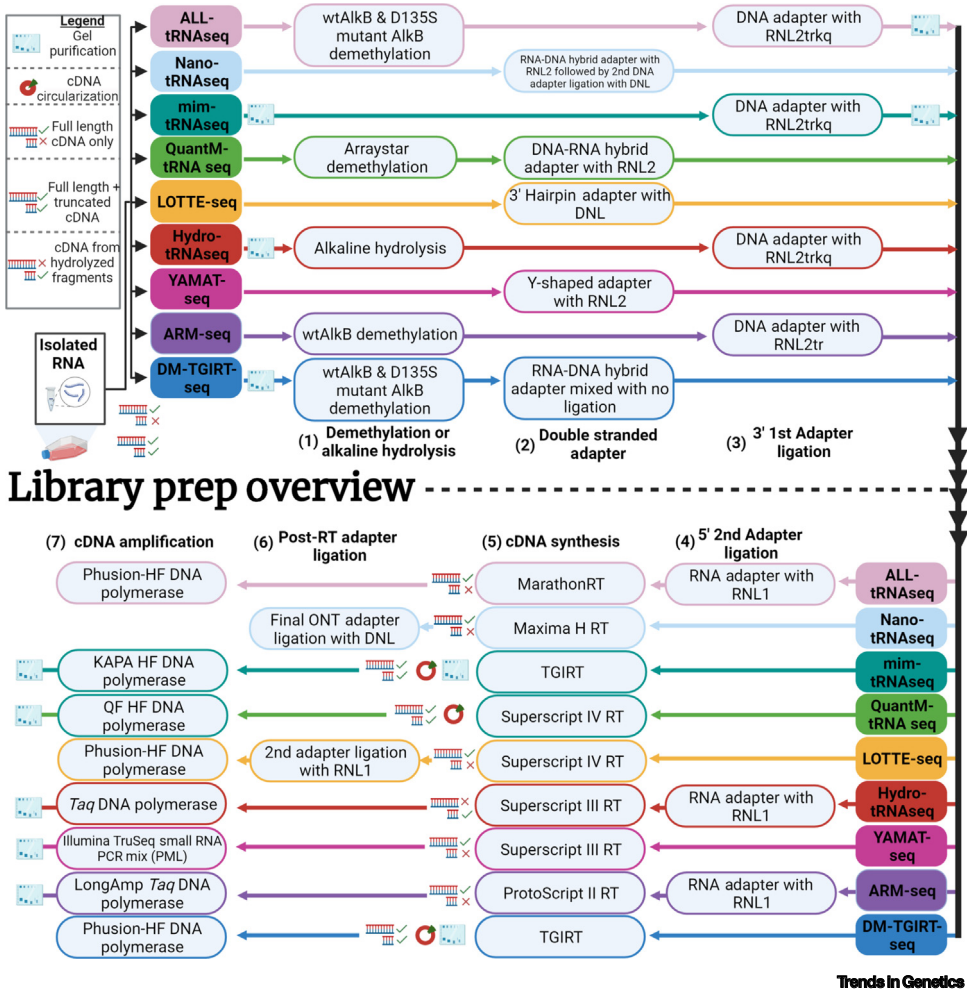

Figure 1. The symbols used to denote gel purification and CircLigase circularization are given in the key (top left). Abbreviations: ALL-tRNAseq, adapter-ligated libraries of tRNA-derived sequences; ARM-seq, AlkB-facilitated RNA methylation sequencing; DM-TGIRT-seq, demethylase-thermostable group II intron RT tRNA sequencing; DNL, T4 DNA ligase; HF, high fidelity; LOTTE-seq, long hairpin oligonucleotide-based tRNA high-throughput sequencing; mim-tRNAseq, modification-induced misincorporation tRNA sequencing; ONT, Oxford Nanopore Technology; QuantM-tRNA seq, quantitative mature tRNA sequencing; RNL1, RNA ligase 1; RNL2tr, truncated RNA ligase 2; RNL2trkq, truncated K227Q mutant RNA ligase 2; RT, reverse transcriptase; wtAlkB, wild-type alkylation B; YAMAT-seq, Y-shaped adapter-ligated mature tRNA sequencing.

ALL-tRNAseq, Hydro-tRNAseq, and ARM-seq employ strategy A; LOTTE-seq uses strategy B; DM-TGIRT-seq, mim-tRNAseq, and QuantM-tRNAseq use strategy C (although DM-TGIRT-seq does not use ligation enzymes, its overall protocol best fits with this strategy); YAMAT-seq uses strategy D; and Nano-tRNAseq uses a combination of strategies B and D because it includes a second primer on a double-stranded oligonucleotide and a final adapter ligation after the RT step. Within these broad categorizations, there are differences in the choice of ligation enzyme, adapter construction, and RT enzyme, the details of which can be found in Figures 1 and 2.

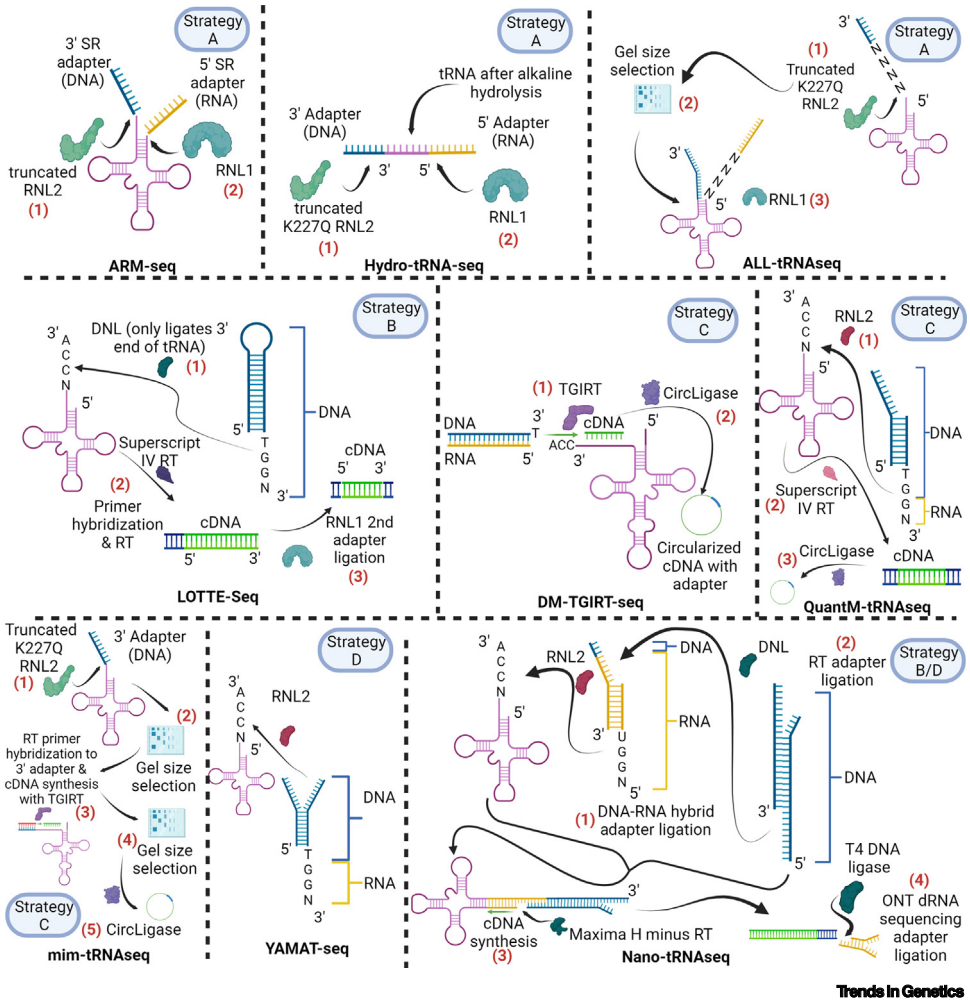

Figure 2. Overview of adapter ligation strategies for each protocol (DM-TGIRT-seq is also included, even though adapters are not technically ligated in this method). In multistep ligation strategies, the steps are numbered in red font. (Strategies A–D) Overall strategies of adapter ligation: (Strategy A) separate ligation of 3′ and 5′ adapters; (Strategy B) 3′ adapter ligation followed by reverse transcription (RT), which is then followed by single-stranded DNA ligation of DNA oligonucleotides to the 3′ extremities of the cDNA; (Strategy C) 3′ adapter ligation followed by circularization of cDNA; and (Strategy D) inclusion of a second primer binding site into the double-stranded oligonucleotide. Abbreviations: ALL-tRNAseq, adapter-ligated libraries of tRNA-derived sequences; ARM-seq, AlkB-facilitated RNA methylation sequencing; DM-TGIRT-seq, demethylase-thermostable group II intron RT tRNA sequencing; DNL, DNA ligase; dRNA, direct RNA sequencing; LOTTE-seq, long hairpin oligonucleotide-based tRNA high-throughput sequencing; mim-tRNAseq, modification-induced misincorporation tRNA sequencing; ONT, Oxford Nanopore Technology; QuantM-tRNA seq, quantitative mature tRNA sequencing; RNL1, RNA ligase 1; RNL2, RNA ligase 2; SR adapter, single-read adapter; YAMAT-seq, Y-shaped adapter-ligated mature tRNA sequencing.

In the vast majority of available protocols, the 3′ DNA adapter is 5′-pre-adenylated (avoiding the necessity of ATP for ligation) and 3′-blocked to avoid the formation of concatemers. Key enzymes for 3′ adapter ligation include T4 RNA ligase 2 (RNL2), truncated T4 RNA ligase 2 (RNL2tr), and truncated K227Q mutant T4 RNA ligase 2 (RNL2trkq) – the truncated form has increased affinity for joining the 5′ ends of adenylated adapters to the 3′ ends of RNA, and the mutated form further reduces unwanted ligation side products [31]. LOTTE-seq and Nano-tRNAseq are the only protocols to make use of DNA ligase (DNL), and the first uses it for ligation of a DNA hairpin adapter to

tRNA, whereas the second uses it to anneal a double-stranded (ds)DNA oligonucleotide containing Oxford Nanopore Technology (ONT) RT adapters to a RNA–DNA hybrid Y-shaped splint adapter (which itself was originally ligated with RNL2). DNL is more specific than RNL, thus reducing the production of undesirable ligation products [32]. A commonality in the protocols using double-stranded adapter ligation strategies (all of which are grouped in different strategies) – YAMAT-seq, QuantM-tRNAseq, Nano-tRNAseq, and LOTTE-seq – is the use of ribonucleotide and/or deoxyribonucleotide overhangs to better hybridize with mature tRNA species. ALL-tRNAseq, while not using double-stranded ligation, uniquely employs a randomized 4 nt 5′ end for its 3′ adapter, which could aid in ligation to tRNA fragments (and not exclusively to mature 3′-CCA tRNAs).

Adapter ligation to tRNA 5′ ends, which is only carried out as a separate step for protocols that employ strategy A, is typically conducted by T4 RNA ligase 1 (RNL1). RNL1 is well suited to ligating ssRNA molecules, which is the structure of the 5′ adapter [33]. Similarly to its 3′ adapter, ALL-tRNAseq also has a randomized tetranucleotide at the 3′ end of its 5′ adapter, for the same purpose as above.

CircLigase circularization (strategy C) and second adapter ligation after RT (strategy B) are alike in that the 3′ end of cDNA is only ligated to an adapter after RT, which has the benefit of enabling incomplete transcription products (as generated by RT stops) to be included in the final library preparation. Notably, mim-tRNAseq and QuantM-tRNAseq both use CircLigase I, whereas DM-TGIRT-seq uses CircLigase II. In at least one study, CircLigase I was found to have higher circularization efficiency than CircLigase II [34]. A different study compared circularization strategies with standard TruSeq™ small RNA preparation, and found that higher numbers of unique reads were observed – this study employed only CircLigase II [35].

Reverse transcription

Since the RT step is almost inevitable in any standard protocol for sequencing library preparation, the presence of modified nucleotides in tRNAs is a well-recognized source of strong bias towards over-representation of less extensively modified tRNA species [22,36]. Moreover, although some tRNA modifications are silent in the RT reaction and only mildly affect cDNA synthesis [such as 5-methylcytosine (m^5C), 7-methylguanosine (m^7G), ψ, and others that do not alter Watson–Crick base-pairing], other tRNA modifications either pause or simply arrest RT primer extension. These are particularly frequent in eukaryotic tRNAs [*N*1-methyladenosine (m^1A), $m^{2,2}G$, 3-methylcytosine (m^3C)] or are common to all living species [m^1G, *N^6-isopentenyl adenosine* (i^6A)/2-methyl-thio-*N^6-isopentenyladenosine* (ms^2i^6A), etc.] [14,37,38]. When such nucleotides are encountered by the RT enzyme, the cDNA extension is either aborted or the enzyme passes through with possible incorporation of a mismatched nucleotide in the cDNA. This RT signature depends not only on the properties of the enzyme [39] but also on the nature of the RT-arresting nucleotide, its sequence context, and the composition of the reaction buffer (namely Mg^{2+}/Mn^{2+} ions) [14,40,41].

The TGIRT enzyme, when included in sequencing library preparation, can facilitate readthrough of tRNA modifications including a subset of Watson–Crick base-pairing modifications [17,42]. In mim-tRNAseq, modifications to the TGIRT reaction conditions were made to further improve its efficiency and reduce the number of premature RT stops [16]. In the recently developed ALL-tRNAseq method, a highly processive group II intron maturase MarathonRT was used which was touted to overperform TGIRT for the sequencing of long and structured RNAs.

Computational methods in tRNA sequencing

In general, tRNA-seq bioinformatic workflows begin with raw FASTQ files, proceed with some preprocessing steps, include curation of a reference, and finally align reads to references. A detailed overview of these steps in each protocol is given in Table 1.

Table 1. Different computational methods in tRNA-seq[a,b]

	DM-TGIRT-seq	ARM-seq	YAMAT-seq	Hydro-tRNAseq	LOTTE-seq	QuantM-tRNAseq	mim-tRNAseq	Nano-tRNAseq	ALL-tRNAseq
Preprocessing – trimming tool	Trimmomatic	'Seqprep' custom script	Not found	Cutadapt	BBDuk	Cutadapt	Cutadapt	N/A	sRNAbench
Preprocessing – additional steps (post-trimming)	Minimum read length of 15 nt	Minimum read length of 15 nt	Information not found	N/A	50–100 nt read length	N/A	Minimum read length of 15 nt	Adapters added to the reference	N/A
Reference curation – database	GtRNAdb tRNAs	tRNAscan-SE output and whole genomes	GtRNAdb genomic tRNAs and whole genome	GtRNAdb tRNAs	tRNAscan-SE output (pre-tRNAs and mature), tRNA-masked whole genome	GtRNAdb	GtRNAdb, MODOMICs modification indexing	GtRNAdb	GtRNAdb, mitotRNAdb
Reference curation – additional steps	3′-CCA added, identical isodecoders consolidated (462 remaining)	3′-CCA and 5′-G added, and introns removed	3′-CCA added (total of 632 genes) in conjunction with whole GRCh37 assembly	GtRNAdb tRNAs were used, supplemented by custom annotation and curation. Both genomic and mature tRNAs were used	3′-CCA tails added	Identical tRNA sequences consolidated	Clustering is applied with a similarity threshold to group reference tRNAs together. A centroid cluster sequence is obtained	Initial set of tRNAs is reduced to 42 sequences corresponding to individual isoacceptors	N/A
Alignment – aligner	Bowtie2	Bowtie2	SHRiMP2	BWA	Segemehl	Bowtie2	GNSNP	BWA	sRNAbench (Bowtie)
Alignment – additional settings/steps	One mismatch allowed.	Default mismatch tolerance used	10% mismatch rate	Hierarchical mapping scheme, with each round of mapping increasing error tolerance (up to two mismatches). These mismatches were used to identify modified residues	80% accuracy threshold for first-round mapping to pre-tRNA and tRNA-masked genome. This was followed by remapping with 85% accuracy to a mature tRNA genome	Min-score value of G,1,8 and seed mismatches set at 1. Isodecoders were quantified by identifying reads with MAPQ >10	Alignment to centroid cluster sequence with –SNP-tolerance flag enabled and 0.1 mismatches per read (outside pre-indexed sites). Remapping after SNP index is updated with new sites from first-round mapping	bwa mem -W13 -k6 -xont2d -T20 was used for alignment, which implies minimum initial seed length (*k*) of 6 and seeded bases discarded if not at least 13 nt in length	Second-round mapping with a Smith–Waterman alignment with a higher mismatch tolerance was used to recover additional potentially modified tRNA reads
Anticodon-level resolution?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Isodecoder resolution?	No	Isodecoder data are presented, but there is no dedicated protocol for isodecoder resolution	No	Isodecoder data are presented, but there is no dedicated protocol for isodecoder resolution	No	Yes	Yes	No	No
% Mapped reads (of total reads)	76.73–82.06%	10.1–13.2%	92.6–95.1%	16.8–60%	97%	90%	87.3–90.6%	58.62%	62%
% Uniquely mapped reads	Unspecified	Unspecified	Unspecified	Unspecified	Unspecified	~40% of mapped	~96.8–100% of mapped	93% of mapped	Unspecified

[a]Preprocessing, reference curation, and alignment steps of bioinformatics pipelines corresponding to tRNAseq protocols. In some cases, incomplete or missing data were noted. Protocols are ordered from left to right in the order that they were published.
[b]Abbreviation: N/A, not available.

CellPress

Preprocessing of raw FASTQ reads

Raw FASTQ reads typically undergo some type of adapter trimming before moving forward in a tRNA-seq analysis pipeline, although the specific tools and criteria used can vary. In summary, the tools used were Trimmomatic [43] (DM-TGIRT-seq), cutadapt [44] (Hydro-tRNAseq, QuantM-tRNAseq, mim-tRNAseq), BBDuk (LOTTE-seq), a custom script entitled Seqprep (ARM-seq), and sRNAbench (All-tRNAseq). YAMAT-seq and Nano-tRNAseq did not specify an adapter trimming or preprocessing protocol. Two publications, which focused on the release of 'fastp' [45] and 'Atria' [46], completed a comparison of two of the aforementioned packages – Trimmomatic and cutadapt (which is used within a wrapper script entitled 'Trim Galore'), alongside many other unmentioned packages. Although cutadapt was found to be faster in both studies, it was reported to have a higher mismatch rate in the fastp study, and a much lower %PPV (positive predictive value) in the Atria study: 40.05% compared with 99.29%.

Perhaps more important than the choice of tool is how it is applied and whether that application is justified. Some studies have pointed out that adapter trimming is not necessary in many cases and should be applied with caution, although these do not mention small-RNA-seq data (which are often considered to include tRNA-seq data) specifically. However, studies which do focus on small RNA adapter trimming emphasize its importance and some include a recommended lower limit of inclusion for post-trimmed reads of ~15 nt [47,48]. Indeed, this is the lower limit followed by two of four tRNA-seq studies which offer detailed information on their adapter trimming protocols – namely DM-TGIRT-seq and ARM-seq. By contrast, mim-tRNAseq is more permissive, and decreases this lower limit to 10 nt. LOTTE-seq is the most restrictive, and has a lower limit of 50 nt.

Reference curation

Before beginning the alignment of reads it is necessary to curate a suitable reference, which can involve varying levels of curation and editing of publicly available genomic data. All the protocols mentioned in this study rely on either a tRNA reference obtained from the application of tRNAscan-SE [49] to a genome or from the GtRNAdb database, which itself is derived from tRNAscan-SE results.

Some protocols have additional references which are used before a mature-tRNA reference – ARM-seq, YAMAT-seq, LOTTE-seq all include the entire human genome as an additional reference, and the latter takes the step of softmasking all tRNA sequences within this genome and concatenating the reference with a pre-tRNA reference. This has the benefit of offering an opportunity to map non-tRNA reads.

Mim-tRNAseq alone incorporates MODOMICS [11] data to annotate their initial reference with modification information.

Alignment as performed by different tRNA-seq protocols

Earlier protocols employed a relatively uncomplicated approach to sequence alignment, and DM-TGIRT-seq, ARM-seq, and QuantM-tRNAseq used Bowtie2 [50] with slight alterations for mismatch tolerance. Similarly, YAMAT-seq has a straightforward mapping approach with SHRiMP2 [51], and specifies a 10% mismatch tolerance. ALL-tRNAseq, Hydro-tRNAseq, and LOTTE-seq all employ a hierarchical mapping approach using Bowtie2, the Burrows–Wheeler aligner (BWA), and Segemehl [52], respectively. In the case of the first two, second-round mapping was modified with an increased mismatch tolerance to recover tRNA reads with more mismatches (modifications). The latter attempted to map first to pre-tRNA and a tRNA-masked genome, followed by mapping to mature tRNA sequences. This step has the benefit of allowing a user to account for non-tRNA reads and potentially assess any contamination within a sample.

Nano-tRNAseq, despite settling on a relatively straightforward single-step mapping approach with BWA, did take the additional step of testing various mapping parameters to optimize their alignment. Finally, mim-tRNA-seq, which uses GSNAP [53] for its aligner, utilizes its unique functionality as an SNP-aware aligner to allow mismatch tolerance at sites specified by MODOMICS data, with an additional 10% mismatch tolerance outside these specific sites. These additional mismatches can be used to add to the library of known tRNA modifications, which can subsequently be used in a 'remapping' step that accounts for previously unknown modification sites. The protocol then applies a unique deconvolution step that allows reads to be separated from their parent cluster into individual tRNA species (if possible).

Hybridization-based approaches to tRNA-seq

Early methods of tRNA quantitation included separation of tRNAs by 2D gel electrophoresis, chromatography, and quantitation by northern blotting [54,55]. Although northern blotting is a low-throughput technique that is cumbersome to perform and requires highly specific probes, it remains a reliable technique for tRNA quantitation and is frequently used to validate newly developed techniques [14,16,56].

tRNA microarrays were the first high-throughput methods developed for tRNA quantitation. These involve hybridization between tRNAs and custom designed sequence-specific DNA probes and detection of binding by either fluorescence or phosphor imaging [57,58]. While offering an improvement over the earlier low-throughput methods, tRNA microarrays have some limitations that are specific to microarray platforms; these include (i) low dynamic range and sensitivity compared with NGS-based methodologies, (ii) low resolution – tRNA microarrays require at least eight nucleotide differences to prevent cross-hybridization, and (iii) the potential requirement for custom probes to detect target species [14]. Further, tRNA microarray protocols that do not include an amplification step will be free from amplification bias; however, they routinely require a substantially higher quantity of the starting material (total RNA or tRNA) in comparison to NGS approaches. Nevertheless, microarray-based approaches continue to be developed further [56,59].

OTTER (oligonucleotide-directed three-prime terminal extension of RNA) [60] and quantitation by microscale thermophoresis (MST) [61] are recently developed hybridization-based tRNA quantitation methods. OTTER employs a reverse primer extension method in which a DNA oligonucleotide specific for an isodecoder/isoacceptor tRNA is hybridized to the 3′-terminal regions of tRNA, and the tRNA/oligonucleotide hybrid is then subjected to DNA polymerization to yield a fluorescent tRNA derivative with a 3′-terminal tetramethylrhodamine-dUTP. Quantitation of fluorescence is used to determine the absolute levels of the target tRNA. While OTTER allows relatively faster quantitation of individual target tRNAs, it is a low-throughput method that relies on the sequence variation in the 3′ region of tRNAs [60]. Quantitation by MST involves hybridization between a constant amount of fluorescently labeled cDNA probes (FCPs) with increasing amounts of denatured tRNAs and quantitation of fluorescence to monitor the thermal mobility of free and tRNA-bound probe. Calculation of the concentration of RNA required to hybridize to 50% of the FCP will yield information on the absolute quantity of the tRNAs in the sample [61]. The advantages and disadvantages of tRNA quantitation by MST are largely comparable with OTTER. Overall, hybridization-based tRNA quantitation methods are hampered by low resolution and dynamic range, the requirement for a high quantity of samples, or low throughput, and are being increasingly replaced by NGS approaches.

tsRNAs and their profiling

tsRNAs, which are a product of tRNA cleavage, have been demonstrated to play a role in multiple biological processes including immune responses, metabolic disorders, and malignancies

by regulating gene expression and translation [62]. Therefore, accurate identification and quantitation of tsRNAs is critical for understanding their roles in normal physiology and disease states.

tsRNAs are broadly categorized into two types: (i) tRNA-derived stress-induced RNAs (tiRNAs), 31–40 nt in length, that result from cleavage of mature tRNA at the anticodon loop by the RNase angiogenin (ANG) under stress conditions, and (ii) tRNA-derived fragments (tRFs), 14–30 nt in length, that are generated through cleavage of pre- or mature tRNAs into tRNA loops by various RNases [63]. The tRNA modifications that are present in the parental tRNAs modulate their cleavage by RNases and are inherited by tsRNAs. Different RNases involved in the generation of tRFs also result in distinct termini including 5′-OH, 3′-P, and 2′,3′-cyclic phosphate (2′3′-cP) modifications [64]. These tRNA modifications and distinct termini interfere with RT read-through and adapter ligation steps, respectively, and affect the identification and quantitation of tsRNAs. Various methods including PANDORA-seq [65], CPA-Seq [66], and ARM-Seq [18] have addressed these limitations by including (i) T4 polynucleotide kinase (T4PNK) to phosphorylate the 5′-OH ends and to convert the 3′-P or 2′3′-cP into 3′-OH, and (ii) AlkB to remove methylation modifications to facilitate RT read-through. In addition, highly processive RT enzymes such as TGIRT and MarathonRT can also be introduced into protocols to facilitate RT read-through and also facilitate the identification of tRNA modifications [23,64].

Validation methods in tRNA quantification and characterization

We focus here on methods to evaluate ligation efficiency, modification data, and read counts. A full summary is given in Table 2.

Ligation efficiency

Ligation efficiency is typically understood to be the proportion of input tRNA that is ligated with adapters. Of the five studies in our cohort which reported ligation efficiency, only one reported this value for each adapter independently – YAMAT-seq reported a 22.8% efficiency for the 5′ adapter and a 91.8% efficiency for the 3′ adapter. This reduced 5′ adapter ligation efficiency is consistent with literature reports about the difficulty of 5′ adapter ligation in small-RNA-seq [67]. LOTTE-seq circumvents this problem of 5′-end ligation to RNA by ligating the second adapter after cDNA synthesis, which occurs after 3′ hairpin adapter ligation with DNL – they report a ligation efficiency with DNL of 80–90%. They also discuss the potential of a unique molecular identifier (UMI) sequence in their adapter; in principle this could be included in other tRNA adapter ligation strategies as well (and is already included in many other low-input RNA-seq library preparations [68]), which would allow quantification of ligation bias or overamplification of artefacts [69]. QuantM-tRNAseq, like LOTTE-seq, implements a double-stranded adapter strategy, although it reports a higher adapter ligation efficiency of 96% using RNL2; however, given the ability of RNL2 to ligate a wider variety of substrates [70], and experiments performed by the LOTTE-seq authors showing the ability of RNL2 to ligate tRNAs missing a 3′-CCA, it is possible that this estimation of ligation efficiency in QuantM-tRNAseq includes tRNA lacking 3′-CCA. Mim-tRNAseq uses a relatively straightforward strategy of 3′ adapter ligation with RNL2 followed by circularization with CircLigase – this resulted in 89–95% ligation efficiency. However, given the reliance on RNL2, this could have the same issue as QuantM-tRNAseq. The final protocol with reported ligation efficiency, Nano-tRNAseq, relies on a multistep protocol involving both RNL2 and DNL, resulting in a final ligation efficiency that we approximated at ~60% based on extended data (Figure 2a,b in [22]). Optimization of reaction conditions showed that increased reaction time and the addition of polyethylene glycol PEG8000 increase ligation efficiency.

Table 2. Validation methods in tRNA quantification and characterization[a]

	DM-TGIRT-seq	ARM-seq	YAMAT-seq	Hydro-tRNAseq	LOTTE-seq	QuantM-tRNAseq	Mim-tRNAseq	Nano-tRNAseq	ALL-tRNAseq
Validation of adapter ligation with ligation efficiency if included	No adapter ligation needed	Ligation efficiency not calculated	Synthetic cytotRNA^{AspGUC} as well as total RNA were tested with a Y-shaped adapter or traditional Illumina adapters – 5′ ligation was improved by 150-fold (22.8% efficiency compared with 0.15%), and 3′ ligation was improved by 1.5-fold (91.8% efficiency vs. 61.3%)	Ligation efficiency was not calculated	DNL ligation was estimated at 80–90% using gels. Gels were also run on tRNA-enriched samples which were either treated with DNL adapter ligation or untreated – only treated samples showed ligation products	Gels were run on tRNA that was either treated with RNL2 and untreated – only treated samples with adapters added showed ligation products. 96% ligation efficiency was calculated	Ligation efficiency was measured by comparing samples treated and untreated with adapters and RNL2 on a gel – ligation efficiency based on band intensity in the range of 89–95%	Ligation products with varying conditions of adapter ligation (changing reaction duration and addition/exclusion of PEG8000) were compared on a gel – overnight reaction duration performed best and gave >60% ligation efficiency. Band intensities were normalized against input unligated tRNA^{Phe}	Ligation efficiency was not calculated
Northern blot or microarray validation	No	No	Yes – cytotRNA^{LysCUU}, cytotRNA^{AlaAGC}, and mttRNA^{ValUAC} were tested – good agreement was noted	No	No	Yes – ten probes for northern blot and 30 probes for microarray testing, which showed good agreement with sequencing	Yes – tRNA^{ArgUCU} and tRNA^{GlyCCC} were tested which showed good agreement with sequencing	No	Yes for four tRNAs, which showed good agreement with sequencing
Primer extension validation of modification data	No	Yes	No	No	No	No	Yes	No	Yes
LC-MS/MS validation of modification data	Yes	No	No	No	No	No	No	Yes	No

(continued on next page)

Table 2. (continued)

	DM-TGIRT-seq	ARM-seq	YAMAT-seq	Hydro-tRNAseq	LOTTE-seq	QuantM-tRNAseq	Mim-tRNAseq	Nano-tRNAseq	ALL-tRNAseq
tRNA abundance versus gene copy number analysis	Yes (human DNA) – poor correlation	No	Yes, positive correlation, R^2 not reported	No	No	No	Yes, R^2 ranged from 0.62 for human induced pluripotent stem cells (iPSCs) to 0.92 for *Saccharomyces cerevisiae*	No	No
Other validation of read counts and modification data	tRNA reads were increased (6.6–15.7 million vs. 2.3–4.7 million) by using gel-purified tRNA versus total RNA. In addition, demethylase treatment gave far fewer RT stops and mismatches compared with untreated samples	By comparing RT stops in methylated versus unmethylated samples, m^1A_{58} modifications were identified in yeast tRNAs, resulting in 94% accuracy versus MODOMICS data	Direct comparison between conventional RNA-seq and YAMAT-seq reads showed that 96% of YAMAT-seq reads mapped to tRNA while only 25% of traditional reads mapped to tRNA	tRNA read content increased by >40% using hydro-tRNA-seq compared with conventional RNA-seq, although the RNA seq-data were not shown. Comparison between SSB PAR-CLIP and hydroseq showed very good correlation for isoacceptor counts ($R = 0.99$)	Modification analysis showed good agreement with literature sources. In addition, comparison of mapping with LOTTE-seq reads and those generated by other protocols showed that ~97% of LOTTE-seq reads mapped to tRNA	Anticodon pools compared across different tissues were consistent (but with differences in individual isodecoders). Analysis of modifications revealed that these were found in regions well known for modifications – new modifications were also identified	Misincorporation analysis of yeast samples (one of which was unable to generate a wybutosine modification at G37 due to a mutation) revealed a distinct difference between the wild-type and the mutant. Many previously annotated modifications were identified, as well as many new ones	Sequencing of wild type and Pus4-deficient *S. cerevisiae* (which has a characteristic ψ modification at position 55 in the T-loop of tRNAs) revealed that this position differed significantly between the wild type and the mutant, supporting the ability of this method ability to identify modifications	Comparisons with MarathonRT and SuperScript III reads showed that it generated a much higher percentage of reads ~75–88 nt in length corresponding to tRNA. In comparisons to DM-tRNA-seq and mim-tRNA-seq data, a higher proportion of full-length tRNAs was identified

[a]Table showing how tRNA-seq data were validated, both using non-NGS data (first row) and other methodologies (second row). If data could not be found for a particular validation strategy, it was assumed that this had not been performed. Protocols are ordered from left to right in the order that they were published.

Comparison with non-NGS methods and other validation

An important, albeit imperfect method for validating tRNA read-count data remains the aforementioned hybridization approach. This approach was used by five protocols, but in all cases the number of probes utilized (QuantM-tRNAseq employed the greatest number of probes) were far outnumbered by the actual number of tRNA species in the sample owing to the low-throughput nature of the technique. However, in all the cases a good agreement between the hybridization- and NGS-derived data was achieved.

Primer extension [71] is another non-NGS method that can be used for validation, as it can be used to identify RT stops. If RT stops are identified at locations corresponding to the NGS output, then this can serve as a support for tRNA modification analysis. This method is featured in three of the protocols, most prominently in ARM-seq.

Mass spectrometry and liquid chromatography (LC-MS) techniques have been used previously many times to detect RNA modifications [72], including in two of the protocols discussed in this paper (DM-TGIRT-seq and Nano-tRNAseq). This can be used to confirm the presence of key tRNA modifications, but (in the methods used in these two protocols) does not give their location in the sequence. Newer MS techniques are able to perform *de novo* tRNA-seq and give modification information – this is known as MS ladder complementation sequencing (MLC-seq), and is a very promising area of emerging research. It requires further development to achieve higher throughput, but is notable for its high precision [73,74].

A variety of other methods can be used to validate tRNA-seq data. Two protocols (mim-tRNAseq and Nano-tRNAseq) employ a mutant biological sample with a known absence of modified tRNA to determine whether this can be detected in a tRNA-seq protocol when compared with the wild type. Other methods compare their results with those derived from more conventional protocols to observe whether their results contain greater quantities of tRNA reads or better RT throughput.

It is important to include some discussion of several challenges associated with direct RNA Nanopore sequencing of tRNAs, which is a very promising technique, but likely requires further validation and troubleshooting to achieve its greatest potential. First, tRNAs are small (<100 nt in length), and this size is not optimal for Nanopore sequencing. Extra adapters ligated to tRNAs can help, but parameters for standard Nanopore basecallers should be adapted to the very short read size. Second, Nanopore sequencing has a low Q-score, ranging from 7 to 12 on average. This low accuracy is not critical for analysis of very long mRNA reads, but becomes crucial for short tRNA molecules, specifically for precise mapping to reference tRNAs. Moreover, a high proportion of RNA modifications in tRNA molecules further increases the error rate, since many RNA modifications are known to affect the ion-current profile and thus basecalling precision. In addition, these basecalling error events may show up not only at the position of modification itself but also at the flanking residues. Altogether, such basecalling errors and the low quality of the sequencing data tremendously affect alignment accuracy, in particular for very complex tRNA pools in higher eukaryotes.

Important biological questions associated with tRNA and costs

At the center of any analysis involving tRNA are the biological questions that a researcher is trying to address – this can include relative tRNA quantitation (is there more of a particular tRNA in sample X than in sample Y), absolute tRNA quantitation (is there more of tRNA A than tRNA B in the same sample), quantitation of tRNA charging, and analysis of tRNA modifications.

Of the protocols included in this study, all are capable to varying extents of both relative and absolute tRNA quantification, although their bioinformatic analyses do not all offer isodecoder-level resolution. At the simplest level, all that is required to obtain isodecoder counts is to map reads to a tRNA genome containing individual isodecoders – in practice, this can lead to excessive multimapping, and a dedicated protocol is therefore ideal for handling isodecoders. Mim-tRNA-seq is the protocol with the most effort dedicated to resolving highly similar isodecoders through its dedicated bioinformatic pipeline which is conveniently available as a conda package. This same protocol is the only one in this paper that offers insight into tRNA charging; however, it does require some important wet laboratory steps, such as isolating RNA under acidic conditions and spike-in addition (for an internal control), as well as the aforementioned conda package for analysis of 3′-CCA quantitation [75]. If one is interested in modification data, then nearly all of the protocols can offer this information, although this will again depend somewhat on the ability of each protocol to provide isodecoder resolution. ALL-tRNAseq, with its use of the highly processive MarathonRT, might miss some modification information, although the use of this same enzyme enables better collection of full-length tRNA reads.

Any discussion of selecting a protocol to meet research needs should also be accompanied by a cost analysis. In Table 3 we present cost estimates for steps that are common to most of the protocols in this paper. However, it is important to consider that this does not include labor or time costs, and can vary greatly depending on the (often changing) prices of reagents that are specific to individual protocols. As an example, the protocol for mim-tRNAseq includes four gel selection steps, whereas ALL-tRNAseq includes only one – this difference can result in the former taking several more days to complete.

A researcher should attempt to perform a cost/benefit analysis when deciding on an appropriate protocol and, importantly, should also feel free to combine pieces of different protocols to serve their specific needs.

Table 3. Typical costs of tRNA sequencing

Step	Option	Cost (USD) per sample processed
Total RNA extraction	Phenol	$1–2
	TRIzol	$2–3
	Commercial kit	$5–15
Isolation of tRNA fraction	Gel purification	$0.5–5 (home-made/precast gels)
	Spin column	$5–10
Library preparation	'Simple' custom protocol	Truncated K227Q mutant T4 ligase 2, $8 App-oligonucleotide, variable ~$10 RT enzyme, $5–15 RNA ligase, $1–2 RNA oligonucleotides, variable $5 PCR oligonucleotides, $5 Total, $35–50
	'Elaborate' custom protocol	Very variable, depends on the cost of the RNA–DNA hybrid oligonucleotides Estimate $50–100
	Commercial kit	$65–75
Sequencing cost	Single-end SR100 sequencing (Illumina NextSeq) 25 million raw reads coverage/sample	$90 (P2 100 cycles kit)
Total cost	Depending on the options used	~$130–225

Concluding remarks

tRNA-seq is an evolving field that has been marked by several notable advances in recent years, both in the generation of tRNA-enriched libraries as well as in bioinformatic analysis of tRNA-seq reads. However, it is also marked by a lack of a single definitive approach for quantifying and analyzing the molecule of interest, and each protocol employs different adapter ligation strategies, varied orders of operation, and heterogeneous bioinformatic pipelines. Even though we have attempted to include as many protocols as possible within the current study to cover a wide range of techniques, it is important to note that additional protocols are available that can offer novel methods – for example, preceding Nano-tRNAseq was another study that examined direct tRNA-seq with Nanopore technology [76], as well as a tRNA study which developed methods to analyze tRNA expression in plants [36]. Comparing results between different approaches is challenging – it is difficult to ascribe differing outputs to any particular difference between protocols because the protocols are so distinct. Future efforts should attempt to compare the wide variety of available tRNA-seq strategies in a single study such that inefficiencies can be eliminated and more reliable quantification of tRNA can be performed. New studies should also attempt to continuously monitor the field for new developments to ensure that analysis strategies are not outdated – for example, the recent release of tModBase [77] could help in expanding modification-aware alignment strategies in the future. We have outlined some current unresolved issues (see Outstanding questions) which aim to address potential future directions in tRNA research.

Outstanding questions

Which adapter ligation strategy results in the least amount of bias while giving the highest possible yield of tRNA?

How can we balance the processivity of RT enzymes – which offers read-through of modified sites and allows more complete sequencing – with the additional information offered by modification-induced mismatch signatures and RT stops? Which RT offers the best balance of these concepts?

While many breakthroughs have been made in the field of modification-aware sequencing, there is still the issue of 'RT-silent' modifications, such as ψ, which would not be detected by current tRNA-seq protocols; what steps should be undertaken to expand the range of modifications that can be detected?

PCR amplification is still an important component of non-ONT techniques, but bias can be introduced at this stage – can UMI adapters be incorporated into additional protocols beyond LOTTE-seq so that this bias can be accounted for?

Acknowledgments

This work was supported by funds from FDA CBER operating funds. This work was also supported by National Heart, Lung, and Blood Institute (NHLBI) grant HL151392. Figures were created with BioRender.com.

Declaration of interests

We have no conflicts of interest to declare.

References

1. Doherty, J. and Guo, M. (2016) Transfer RNA. In *Encyclopedia of Cell Biology* (Bradshaw, R.A. and Stahl, P., eds), pp. 309–340, Academic Press
2. Dong, J. *et al.* (2000) Uncharged tRNA activates GCN2 by displacing the protein kinase moiety from a bipartite tRNA-binding domain. *Mol. Cell* 6, 269–279
3. Pavlova, N.N. *et al.* (2020) Translation in amino-acid-poor environments is limited by tRNA. *eLife* 9, e62307
4. Katz, A. *et al.* (2016) Non-canonical roles of tRNAs and tRNA mimics in bacterial cell biology. *Mol. Microbiol.* 101, 545–558
5. Marquet, R. *et al.* (1995) tRNAs as primer of reverse transcriptases. *Biochimie* 77, 113–124
6. Akiyama, Y. and Ivanov, P. (2023) tRNA-derived RNAs: biogenesis and roles in translational control. *Wiley Interdiscip. Rev. RNA* 14, e1805
7. Orellana, E.A. *et al.* (2022) tRNA dysregulation and disease. *Nat. Rev. Genet.* 23, 651–664
8. Schaffer, A.E. *et al.* (2019) tRNA metabolism and neurodevelopmental disorders. *Ann. Rev. Genomics Hum. Genet.* 20, 359–387
9. Aharon-Hefetz, N. *et al.* (2020) Manipulation of the human tRNA pool reveals distinct tRNA sets that act in cellular proliferation or cell cycle arrest. *eLife* 9, e58461
10. Li, G. *et al.* (2023) Unconventional secretion of *Magnaporthe oryzae* effectors in rice cells is regulated by tRNA modification and codon usage control. *Nat. Microbiol.* 8, 1706–1716
11. Boccaletto, P. *et al.* (2021) MODOMICS: a database of RNA modification pathways. 2021 update. *Nucleic Acids Res.* 50, D231–D235
12. Pinkard, O. *et al.* (2020) Quantitative tRNA-sequencing uncovers metazoan tissue-specific tRNA regulation. *Nat. Commun.* 11, 4104
13. Motorin, Y. *et al.* (2007) Identification of modified residues in RNAs by reverse transcription-based methods. *Methods Enzymol.* 425, 21–53
14. Orioli, A. (2017) tRNA biology in the omics era: stress signalling dynamics and cancer progression. *BioEssays* 39, 1600158
15. Hiley, S.L. *et al.* (2005) Detection and discovery of RNA modifications using microarrays. *Nucleic Acids Res.* 33, e2
16. Behrens, A. *et al.* (2021) High-resolution quantitative profiling of tRNA abundance and modification status in eukaryotes by mim-tRNAseq. *Mol. Cell* 81, 1802–1815
17. Zheng, G. *et al.* (2015) Efficient and quantitative high-throughput tRNA sequencing. *Nat. Methods* 12, 835–837
18. Cozen, A.E. *et al.* (2015) ARM-seq: AlkB-facilitated RNA methylation sequencing reveals a complex landscape of modified tRNA fragments. *Nat. Methods* 12, 879–884
19. Gogakos, T. *et al.* (2017) Characterizing expression and processing of precursor and mature human tRNAs by Hydro-tRNAseq and PAR-CLIP. *Cell Rep.* 20, 1463–1475
20. Shigematsu, M. *et al.* (2017) YAMAT-seq: an efficient method for high-throughput sequencing of mature transfer RNAs. *Nucleic Acids Res.* 45, e70
21. Erber, L. *et al.* (2020) LOTTE-seq (long hairpin oligonucleotide based tRNA high-throughput sequencing): specific selection of tRNAs with 3′-CCA end for high-throughput sequencing. *RNA Biol.* 17, 23–32
22. Lucas, M.C. *et al.* (2023) Quantitative analysis of tRNA abundance and modifications by nanopore RNA sequencing. *Nat. Biotechnol.* Published online April 6, 2023. https://doi.org/10.1038/s41587-023-01743-6
23. Scheepbouwer, C. *et al.* (2023) ALL-tRNAseq enables robust tRNA profiling in tissue samples. *Genes Dev.* 37, 243–257
24. Rio, D.C. *et al.* (2010) Purification of RNA using TRIzol (TRI Reagent). *Cold Spring Harbor Protoc.* 2010 pdb.prot5439
25. Green, M.R. and Sambrook, J. (2021) Total RNA extraction from *Saccharomyces cerevisiae* using hot acid phenol. *Cold Spring Harbor Protoc.* 2021 pdb.prot101691

26. Galvanin, A. *et al.* (2019) Mapping and quantification of tRNA 2′-O-methylation by RiboMethSeq. In *Epitranscriptomics: Methods and Protocols* (Wajapeyee, N. and Gupta, R., eds), pp. 273–295, Springer
27. Wang, Y. *et al.* (2021) A high-throughput screening method for evolving a demethylase enzyme with improved and new functionalities. *Nucleic Acids Res.* 49, e30
28. Clark, W.C. *et al.* (2016) tRNA base methylation identification and quantification via high-throughput sequencing. *RNA* 22, 1771–1784
29. Berg, M.D. and Brandl, C.J. (2021) Transfer RNAs: diversity in form and function. *RNA Biol.* 18, 316–339
30. Czech, A. (2020) Deep sequencing of tRNA's 3′-termini sheds light on CCA-tail integrity and maturation. *RNA* 26, 199–208
31. Viollet, S. *et al.* (2011) T4 RNA ligase 2 truncated active site mutants: improved tools for RNA analysis. *BMC Biotechnol.* 11, 72
32. Kershaw, C.J. and O'Keefe, R.T. (2013) Splint ligation of RNA with T4 DNA ligase. In *Recombinant and In Vitro RNA Synthesis: Methods and Protocols* (Conn, G.L., ed.), pp. 257–269, Humana Press
33. Marintcheva, B. (2018) Viral tools for in vitro manipulations of nucleic acids: molecular cloning. In *Harnessing the Power of Viruses*, pp. 27–67, Academic Press
34. Heyer, E.E. *et al.* (2015) An optimized kit-free method for making strand-specific deep sequencing libraries from RNA fragments. *Nucleic Acids Res.* 43, e2
35. Chu, Y. *et al.* (2015) Intramolecular circularization increases efficiency of RNA sequencing and enables CLIP-Seq of nuclear RNA from human cells. *Nucleic Acids Res.* 43, e75
36. Warren, J.M. *et al.* (2021) Combining tRNA sequencing methods to characterize plant tRNA expression and post-transcriptional modification. *RNA Biol.* 18, 64–78
37. Motorin, Y. and Marchand, V. (2021) Analysis of RNA modifications by second- and third-generation deep sequencing: 2020 update. *Genes (Basel)* 12, 278
38. Motorin, Y. and Helm, M. (2019) Methods for RNA modification mapping using deep sequencing: established and new emerging technologies. *Genes (Basel)* 10, 35
39. Werner, S. *et al.* (2020) Machine learning of reverse transcription signatures of variegated polymerases allows mapping and discrimination of methylated purines in limited transcriptomes. *Nucleic Acids Res.* 48, 3734–3746
40. Kristen, M. *et al.* (2020) Manganese ions individually alter the reverse transcription signature of modified ribonucleosides. *Genes* 11, 950
41. Khoddami, V. *et al.* (2019) Transcriptome-wide profiling of multiple RNA modifications simultaneously at single-base resolution. *Proc. Natl. Acad. Sci. U. S. A.* 116, 6784–6789
42. Qin, Y. *et al.* (2016) High-throughput sequencing of human plasma RNA by using thermostable group II intron reverse transcriptases. *RNA* 22, 111–128
43. Bolger, A.M. *et al.* (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120
44. Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* 17, 10
45. Chen, S. *et al.* (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890
46. Chuan, J. *et al.* (2021) Atria: an ultra-fast and accurate trimmer for adapter and quality trimming. *Gigabyte* 2021, gigabyte31
47. Bezuglov, V. *et al.* (2023) Approaches for sRNA analysis of human RNA-seq data: comparison, benchmarking. *Int. J. Mol. Sci.* 24, 4195
48. Zhong, X. *et al.* (2019) Accurate adapter information is crucial for reproducibility and reusability in small RNA seq studies. *Noncoding RNA* 5, 49
49. Chan, P. *et al.* (2021) tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. *Nucleic Acids Res.* 49, 9077–9096
50. Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359
51. David, M. *et al.* (2011) SHRiMP2: sensitive yet practical short read mapping. *Bioinformatics* 27, 1011–1012
52. Hoffmann, S. *et al.* (2009) Fast mapping of short sequences with mismatches, insertions and deletions using index structures. *PLoS Comput. Biol.* 5, e1000502
53. Wu, T.D. *et al.* (2016) GMAP and GSNAP for genomic sequence alignment: enhancements to speed, accuracy, and functionality. *Methods Mol. Biol.* 1418, 283–334
54. Emilsson, V. and Kurland, C.G. (1990) Growth rate dependence of transfer RNA abundance in *Escherichia coli*. *EMBO J.* 9, 4359–4366
55. Cognat, V.R. *et al.* (2008) On the evolution and expression of *Chlamydomonas reinhardtii* nucleus-encoded transfer RNA genes. *Genetics* 179, 113–123
56. Goodarzi, H. *et al.* (2016) Modulated expression of specific tRNAs drives gene expression and cancer progression. *Cell* 165, 1416–1427
57. Dittmar, K.A. *et al.* (2004) Exploring the regulation of tRNA distribution on the genomic scale. *J. Mol. Biol.* 337, 31–47
58. Grelet, S. *et al.* (2017) SPOt: a novel and streamlined microarray platform for observing cellular tRNA levels. *PLoS One* 12, e0177939
59. Pavon-Eternod, M. *et al.* (2009) tRNA over-expression in breast cancer and functional consequences. *Nucleic Acids Res.* 37, 7268–7280
60. Nagai, A. *et al.* (2021) OTTER, a new method quantifying absolute amounts of tRNAs. *RNA* 27, 628–640
61. Jacob, D. *et al.* (2019) Absolute quantification of noncoding RNA by microscale thermophoresis. *Angew. Chem. Int. Ed.* 58, 9565–9569
62. Liu, B. *et al.* (2021) Deciphering the tRNA-derived small RNAs: origin, development, and future. *Cell Death Dis.* 13, 24
63. Fagan, S.G. *et al.* (2021) tRNA-derived fragments: a new class of non-coding RNA with key roles in nervous system function and dysfunction. *Prog. Neurobiol.* 205, 102118
64. Chen, Q. and Zhou, T. (2023) Emerging functional principles of tRNA-derived small RNAs and other regulatory small RNAs. *J. Biol. Chem.* 299, 105225
65. Shi, J. *et al.* (2021) PANDORA-seq expands the repertoire of regulatory small RNAs by overcoming RNA modifications. *Nat. Cell Biol.* 23, 424–436
66. Wang, H. *et al.* (2021) CPA-seq reveals small ncRNAs with methylated nucleosides and diverse termini. *Cell Discov.* 7, 25
67. Lama, L. *et al.* (2019) Small RNA-seq: the RNA 5′-end adapter ligation problem and how to circumvent it. *J. Biol. Methods* 6, e108
68. You, Y. *et al.* (2021) Benchmarking UMI-based single-cell RNA-seq preprocessing workflows. *Genome Biol.* 22, 339
69. Kivioja, T. *et al.* (2011) Counting absolute numbers of molecules using unique molecular identifiers. *Nat. Methods* 9, 72–74
70. Bullard, R. *et al.* (2006) Direct comparison of nick-joining activity of the nucleic acid ligases from bacteriophage T4. *Biochem. J.* 398, 135–144
71. Carey, M.F. *et al.* (2013) The primer extension assay. *Cold Spring Harbor Protoc.* 2013 pdb.prot071902
72. Thüring, K. *et al.* (2016) Analysis of RNA modifications by liquid chromatography–tandem mass spectrometry. *Methods* 107, 48–56
73. Björkbom, A. *et al.* (2015) Bidirectional direct sequencing of non-canonical RNA by two-dimensional analysis of mass chromatograms. *J. Am. Chem. Soc.* 137, 14430–14438
74. Shi, J. *et al.* (2022) Exploring the expanding universe of small RNAs. *Nat. Cell Biol.* 24, 415–423
75. Behrens, A. and Nedialkova, D.D. (2022) Experimental and computational workflow for the analysis of tRNA pools from eukaryotic cells by mim-tRNAseq. *STAR Protoc.* 3, 101579
76. Thomas, N.K. *et al.* (2021) Direct Nanopore sequencing of individual full length tRNA strands. *ACS Nano* 15, 16642–16653
77. Lei, H.T. *et al.* (2023) tModBase: deciphering the landscape of tRNA modifications and their dynamic changes from epitranscriptome data. *Nucleic Acids Res.* 51, D315–D327