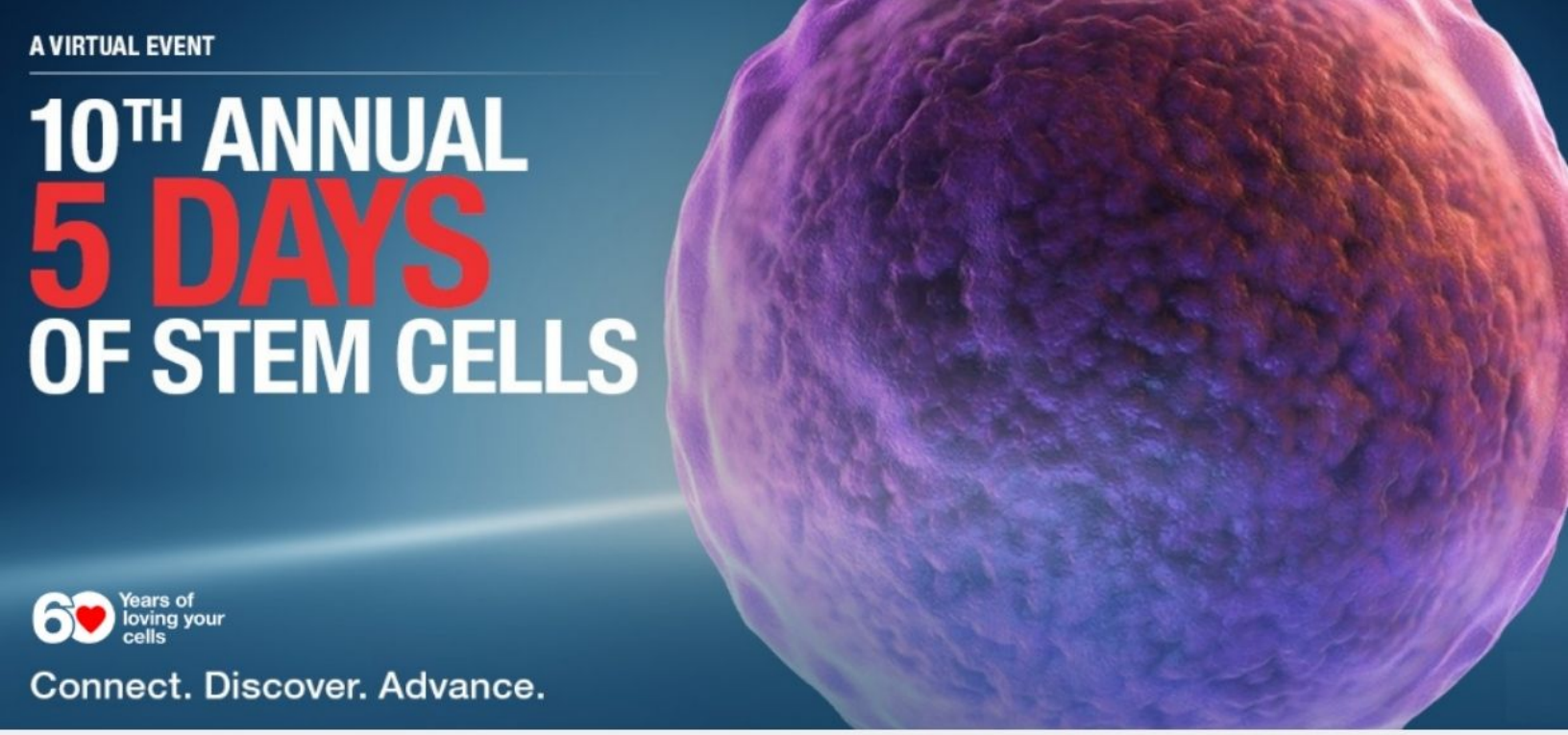

Join us for the world's leading virtual Stem Cell event

October 31 – November 4

The program includes various sessions in:

- 3D culture/organoid
- Cell therapies
- Stem cell best practices
- Reprogramming & gene editing
- Workflow optimization
- Future of stem cells

What's in it for you?

- Gain and share insights and discuss industry trends.
- Celebrate Gibco's 60th anniversary: play and win prizes
- Network with your peers directly
- Receive a certificate of attendance

Thermofisher.com/5daysofstemcells

WILEY CURRENT PROTOCOLS A Wiley Brand gibco ThermoFisher SCIENTIFIC

Cloning and Sequencing Eukaryotic Small RNAs

Olivia J. Crocker,[1,2,3] Natalie A. Trigg,[1,2,3] and Colin C. Conine[1,2,4]

[1]Division of Neonatology, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania
[2]Departments of Genetics and Pediatrics - Penn Epigenetics Institute, Institute of Regenerative Medicine, and Center for Research on Reproduction and Women's Health, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania
[3]These authors contributed equally and are listed alphabetically.
[4]Corresponding author: *conine@upenn.edu*

Published in the Molecular Biology section

Small RNAs are ubiquitous regulators of gene expression that participate in nearly all aspects of physiology in a wide range of organisms. There are many different classes of eukaryotic small RNAs that play regulatory roles at every level of gene expression, including transcription, RNA stability, and translation. While eukaryotic small RNAs display diverse functions across and within classes, they are generally grouped functionally based on the machinery required for their biogenesis, the effector proteins they associate with, and their molecular characteristics. The development of techniques to clone and sequence small RNAs has been critical for their identification, yet the ligation-dependent addition of RNA adapters and the use of reverse transcriptase to generate cDNA in traditional library preparation protocols can be unsuitable to detect certain small RNA subtypes. In particular, 3′ or 5′ chemical modifications that are characteristic of specific types of small RNAs can impede the ligation-dependent addition of RNA adapters, while internal RNA modifications can interfere with accurate reverse transcription. The inability to clone certain small RNA subtypes with traditional protocols results in an inaccurate assessment of small RNA abundance and diversity, where some RNAs appear over-represented and others are not detected. This overview aims to guide users on how to design small RNA cloning workflows in eukaryotes to more accurately capture specific small RNAs of interest. Hence, we discuss the molecular biology underlying the identification and quantitation of small RNAs, explore the limitations of commonly used protocols, and detail the alternative approaches that can be used to enrich specific small RNA classes.

Keywords: microRNAs • piRNAs • RNAi • RNA-seq • small RNA • tRNA-fragments

How to cite this article:
Crocker, O. J., Trigg, N. A., & Conine, C. C. (2022). Cloning and sequencing eukaryotic small RNAs. *Current Protocols*, *2,* e495. doi: 10.1002/cpz1.495

INTRODUCTION

Small RNAs play important roles in gene expression regulation in all domains of life (Cech & Steitz, 2014; Gelsinger & Diruggiero, 2018). While the term is rather arbitrary, "small RNA" generally refers to 18-40 nucleotide (nt), non-coding regulatory RNAs in eukaryotes (Kim, Han, & Siomi, 2009). The distinction is particular to eukaryotic organisms, as small RNAs in prokaryotes are often significantly larger (up to 500 nt) (Carrier, Lalaouna, & Massé, 2018).

CURRENT PROTOCOLS
A Wiley Brand
Current Protocols e495, Volume 2
Published in Wiley Online Library (wileyonlinelibrary.com).
doi: 10.1002/cpz1.495

Eukaryotic small RNAs predominately associate with effector proteins called Argonautes (AGO) and function in RNA interference (RNAi)-related pathways (microRNAs, piRNAs, and endo-siRNAs). More recently, AGO-independent small RNA pathways have been described, such as those that produce transfer RNA- and ribosomal RNA-fragments. In this overview, we focus on methods to identify and quantitate eukaryotic small RNAs and, thus, "small RNA" will reference the definition presented above throughout the review.

The first small RNA described, *lin-4*, was discovered by Victor Ambros, Rosalind Lee, and colleagues using genetic screens in the nematode *C. elegans* (Lee, Feinbaum, & Ambros, 1993). While the *lin-4* gene was found to produce a 22-nucleotide RNA with partial complementarity to the mRNA produced by the *lin-14* gene (which genetically interacts with *lin-4*), the function of the *lin-4* small RNA was unclear at the time. Notably, it was speculated that *lin-4* could regulate the translation of *lin-14*, providing an accurate prediction of its eventually determined function as a microRNA (miRNA) (Lee et al., 1993; Wightman, Ha, & Ruvkun, 1993). Five years later, the compounding discoveries of RNAi and another highly conserved miRNA (*let-7*) in *C. elegans,* along with the discovery of small-interfering RNAs (siRNAs) in plants, established that small RNAs were not part of a niche form of gene expression regulation but rather ubiquitous regulators of cellular physiology in eukaryotes (Fire et al., 1998; Hamilton & Baulcombe, 1999; Pasquinelli et al., 2000; Reinhart et al., 2000). The 21st century saw an explosion of small RNA research, which was made possible by the development of new techniques to fractionate and clone small RNAs, along with the advent of high-throughput sequencing technologies. These advances have led to the discovery of a myriad of novel small RNA sequences and regulatory functions in different organisms, as well as distinctive small RNA classes with unique molecular characteristics (Girard, Sachidanandam, Hannon, & Carmell, 2006; Lee, Shibata, Malhotra, & Dutta, 2009). While high-throughput sequencing has revealed a whole complement of small RNA subtypes, with an array of biological functions, the molecular characteristics of different small RNA classes (including internal and 5′ and 3′ end RNA modifications) demand careful experimental design in order to capture certain small RNAs of interest that may not be detected using standard protocols and workflows. To address these issues, techniques have been—and are continuing to be—developed to chemically modify and, ultimately, capture all of the small RNAs in a biological sample for detection by sequencing.

In this overview, we discuss current methods of cloning eukaryotic small RNAs, and the known limitations and bias when designing small RNA-sequencing experiments using traditional cloning protocols, that is, those that depend on 5′ and 3′ adapter ligation and traditional reverse transcriptases. Each class of small RNA harbors unique characteristics that need to be considered to select the most appropriate cloning protocol. Therefore, we first discuss the biogenesis pathways of each small RNA class and describe how the molecular characteristics of certain small RNAs can limit their detection. We then review the current methods and workflows that incorporate this information to successfully clone and sequence different classes of small RNAs. Lastly, we consider novel methods and the future direction of small RNA cloning and sequencing.

CLASSES OF SMALL RNAs

siRNAs and Endo-siRNAs

RNA interference (RNAi) refers to the mechanism by which double-stranded RNA (dsRNA) can direct sequence-specific suppression of gene expression. Both exogenously introduced and endogenously produced dsRNA are recognized in the cytoplasm and cleaved by the RNase-III enzyme Dicer into shorter dsRNA molecules called small-interfering RNAs (siRNAs), which are ∼20 nt long and have symmetrical 2 nt 3′ overhangs (Bernstein, Caudy, Hammond, & Hannon, 2001). One of the strands of this siRNA duplex is then loaded into the small RNA effector protein AGO (Fig. 1). This complex is referred to as the RNA-Induced Silencing Complex (RISC), which can target, through base pairing, mRNAs with perfect complementarity to the siRNA sequence, thereby initiating endonucleolytic cleavage of the target RNA using the slicer activity of AGO (Ghildiyal & Zamore, 2009). Since the discovery by Andrew Fire, Craig Mello, and colleagues in 1998 that experimentally administered exogenous dsRNA can trigger RNAi in *C. elegans,* it has been found that organisms also utilize the RNAi-pathway endogenously to control gene expression to regulate all aspects of cellular physiology (Kim et al.,

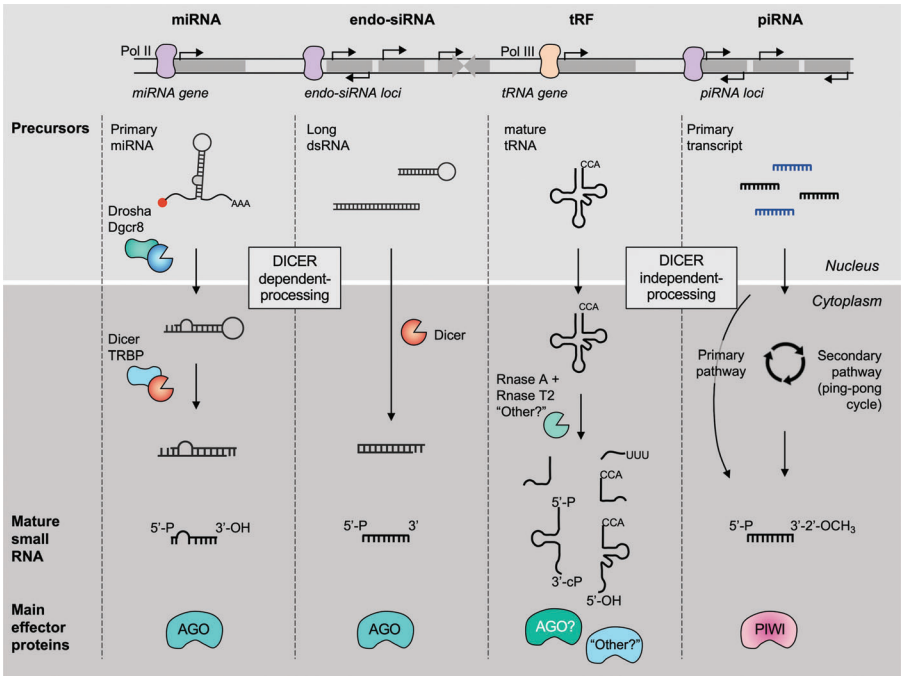

Figure 1 Biogenesis of small RNAs in animals. Small RNA biogenesis begins with the transcription of RNAs that can act as precursors for small RNAs, which are subsequently exported to the cytoplasm (in some cases after processing in the nucleus for example, microRNAs [miRNAs]) and processed into mature small RNAs, including miRNAs, endogenous small interfering RNAs (endo-siRNAs), transfer-RNA (tRNA) fragments (tRFs), and piRNAs. The mature small RNAs are then loaded into their respective effector proteins to elicit their given cellular function. As a result of the different enzymes involved in the processing of each small RNA subclass, the 5′ and 3′ ends harbor different chemical structures. Abbreviations: P, phosphate; OH, hydroxyl; 2′-OCH_3, 2′-*O*-methylation; cP, cyclic phosphate.

2009; Mello & Conte, 2004). Additionally, the ability of artificially introduced exogenous dsRNA to direct gene silencing in cells and organisms has proven to be a powerful tool that has been harnessed experimentally in most organisms to identify the function of specific genes via reverse genetics (Wolters & Mackeigan, 2008).

A different class of small RNAs, termed endo-siRNAs, are derived from endogenous sources of dsRNA precursors that can originate from a variety of sources, including transposon transcripts and sense-antisense transcripts derived from convergent transcription events, pseudogenes, and long stem-loop structures (Okamura & Lai, 2008). The production of endo-siRNAs in flies (*D. melanogaster*) and mammals relies on Dicer processing dsRNAs into siRNAs that are then loaded into AGO (Fig. 1). While a single Dicer enzyme is involved in both miRNA and endo-siRNA biogenesis in mice, flies have two separate pathways for miRNA and endo-siRNA biogenesis, each with a specific Dicer enzyme (Czech et al., 2008) (see also next section). Conversely, in *C. elegans,* endo-siRNAs are synthesized by an RNA-dependent RNA-polymerase (RdRP) from mRNA templates, resulting in antisense small RNAs that are then loaded into AGO proteins. Together, endo-siRNAs and Argonautes (>20) regulate nearly all aspects of the worm's life cycle by regulating gene expression via a wide variety of pathways and mechanisms (Ketting & Cochella, 2021). Although the biogenesis of endo-siRNAs differs by species, endo-siRNAs perform conserved biological functions, including repressing transposons, modulating chromatin organization, and regulating gene expression (Cecere & Grishok, 2014; Morris, Chan, Jacobsen, & Looney, 2004).

miRNAs

miRNAs encompass the most extensively studied class of small regulatory RNAs. In animals, and unlike endo-siRNAs, their biogenesis requires the action of two RNase III-endonucleases, Drosha and Dicer (Ha & Kim, 2014) (Fig. 1). The canonical pathway of miRNA biogenesis begins with

transcription of the miRNA gene by RNA polymerase II, generating a primary-miRNA (pri-miRNA) transcript that encodes a palindromic sequence that results in the generation of a double-stranded hairpin RNA. The hairpin structure is then recognized and bound in the nucleus by a dsRNA binding (DRB) protein, Dgcr8. Dgcr8 positions Drosha on the pri-miRNA to ensure an accurate and efficient Drosha-catalyzed cleavage, which frees the stem-loop of the hairpin and produces the precursor-miRNA (pre-miRNA). The liberated pre-miRNA is then bound by Exportin-5 (XPO5) for export to the cytoplasm. In the cytoplasm, the pre-miRNA is bound by a second DRB protein, TRBP (transactivation response element RNA-binding protein), which accurately positions Dicer on the pre-miRNA stem-loop to ensure the efficient Dicer-catalyzed liberation of the miRNA/miRNA* duplex. This duplex features 2-nt overhangs at both 3′ ends, with both molecules containing a 5′ monophosphate and a 3′ -OH. The miRNA/miRNA* duplex is subsequently loaded into AGO, which retains one duplex strand (the so-called miRNA guide strand), to form an activated miRISC (miRNA induced silencing complex), while simultaneously releasing the other strand (the miRNA* passenger strand) (Fig. 1). The miRNA guide strand provides the sequence specificity needed to direct the assembled complex to target transcripts, analogous to RISC (Ha & Kim, 2014). In contrast to RISC, however, miRISC target recognition does not require perfect complementarity between the entire ∼22 nt loaded miRNA and the mRNA target. Rather, the specificity of binding is dominated by the six-nucleotide sequence situated from positions 2-7 of the 5′ termini of the miRNA, referred to as the "seed" sequence (Bartel, 2009). It should be noted, however, that other types of targeting including supplemental base-paring outside of the seed sequence, and "centered-site" seed targeting can also occur (Shin et al., 2010). In mammals, miRNA binding sites almost exclusively lie in the 3′ untranslated region (3′ UTR) of target transcripts, with the regulation of gene expression mediated by miRISC binding at these sites. Such regulation is commonly achieved by destabilization of the target mRNA, but can also happen by repressing translation (Jonas & Izaurralde, 2015). Conversely, plant miRNAs have high complementarity to their target genes and, since full mRNA:target miRNA complementarity induces AGO endonuclease activity, almost always regulate gene expression by mRNA cleavage of the target gene (Brodersen & Voinnet, 2009). While repression of gene expression remains the most comprehensively evidenced action of miRNAs, these small RNAs have also been demonstrated to stimulate target gene expression (Vasudevan, 2012).

piRNAs

PIWI-interacting RNAs (piRNAs) are a class of endogenously derived small RNAs that are primarily expressed in the germline of metazoans (Iwasaki, Siomi, & Siomi, 2015). Originally identified in *Drosophila melanogaster* (Aravin et al., 2001), piRNAs play a crucial role in maintaining the fidelity of gametogenesis and fertility by protecting the germline against invasive "non-self" nucleic acid sequences such as transposons or viruses. To prevent transposon mobilization in the germline, piRNAs engage in a variety of processes, including heterochromatin formation, DNA methylation, and, most prominently, PIWI-directed degradation of target RNAs (Ozata, Gainetdinov, Zoch, O'Carroll, & Zamore, 2019). To regulate the level of "non-self" RNA targets in the germline, piRNAs associate with the PIWI subfamily of AGO proteins and direct endonuclease-mediated transcript degradation. While piRNAs, miRNAs, and endo-siRNAs all share the ability to modulate target RNA stability, piRNAs differ from these two other types of regulatory RNAs in several important ways. Both miRNAs and endo-siRNAs are produced from dsRNA precursors, while piRNAs are produced from ssRNA precursors through a separate Dicer-independent biogenesis pathway (Kim et al., 2009) (Fig. 1). Further, while endo-siRNAs rely on perfect complementarity to their targets and miRNAs require complementarity of a seed sequence to a target, piRNA targeting exhibits seed sequence pairing supported by additional base pairing (yet it should be noted that the rules of piRNA targeting are still being uncovered) (Anzelon et al., 2021).

While piRNAs have similar functions across species, their sequences are highly divergent, suggesting that piRNAs are involved in an evolutionary host-pathogen "arms race," and must rapidly evolve to counteract transposon activity and spreading throughout the genome (Parhad & Theurkauf, 2019). piRNAs can be characterized into two general categories based on their biogenesis: primary piRNAs, which are processed from a directly

transcribed precursor RNA, and secondary piRNAs, which are produced by a "ping-pong" amplification process (Iwasaki et al., 2015). The MIWI-associated pachytene piRNAs, which are expressed during the meiotic stages of mouse spermatogenesis, are an example of primary piRNAs (Beyret, Liu, & Lin, 2012). 21U-RNAs in *C. elegans* are another type of primary piRNA. These are directly transcribed from their genomic loci, processed, and then loaded into a PIWI protein (PRG-1) (Batista et al., 2008). The "ping-pong" amplification of secondary piRNAs occurs in germ cells to silence expressed transposons. It uses primary piRNAs for the cleavage of antisense transposon transcripts, which then generate antisense secondary piRNAs that can then target the sense transposon transcript. This thus creates a feed-forward loop of post-transcriptional silencing (Brennecke et al., 2007; Gunawardane et al., 2007). In keeping with the observation that piRNA biogenesis is evolutionary divergent across organisms while still exhibiting similar mechanisms of transcriptional regulation, the nematode *C. elegans* does not engage in ping-pong amplification of piRNAs, but does exhibit similar activity through the generation of secondary siRNAs by RdRPs from primary piRNAs (21U-RNAs), thereby representing a similar mechanism of "amplification" of primary piRNAs (Das et al., 2008). In addition to their role in post-transcriptional regulation of transposons, piRNAs and their PIWI effectors have been demonstrated to affect the transcription of transposons by influencing the deposition of H3K9me3 and DNA methylation at their genomic loci (Pezic, Manakov, Sachidanandam, & Aravin, 2014).

piRNAs are 21-35 nt in length and are typically characterized by a 5′- monophosphate and a 2′ *O*-methyl-modified (2-OCH_3) 3′ end (Kirino & Mourelatos, 2007; Ohara et al., 2007) (Fig. 1). This 2′ *O*-methylation mark is added by the methyltransferase HEN1 during maturation. Interestingly, while the presence of the 2′-OCH_3 modification among piRNAs can interfere with standard ligation-dependent cloning procedures, this chemical modification can be leveraged to specifically enrich for this type of RNA during library preparation (see below) (Dard-Dascot et al., 2018; Munafó & Robb, 2010).

Other Small RNAs

While investigations of endo-siRNAs, miRNAs, and piRNAs represent the majority of research on small RNA biogenesis and function, the advancement of sequencing technologies and bioinformatics have uncovered additional classes of small RNAs, including some derived from tRNAs and rRNAs that were previously thought to be random degradation products (Lee et al., 2009). The interrogation of the biogenesis and functions of tRNA fragments and rRNA fragments is an exciting new area of small RNA research, and are thus briefly described below.

tRNA fragments (tRFs)

tRNA fragments, also known as tRFs or tsRNAs, are small RNA fragments generated by cleavage of tRNAs. These tRFs have been implicated in cancer, neurodegenerative disease, viral infection, male germline maturation, and aging (Kumar, Kuscu, & Dutta, 2016; Magee & Rigoutsos, 2020; Soares & Santos, 2017). While the biogenesis of these RNAs is poorly understood, tRFs have been demonstrated to play roles in RNA silencing, translation regulation, and epigenetic inheritance, as reviewed elsewhere (Kumar et al., 2016; Magee & Rigoutsos, 2020; Soares & Santos, 2017). Stress-mediated cleavage of tRNAs into 30-40 nt tRNA halves has been demonstrated to occur in response to hypoxia, starvation, viral infection, arsenite, heat shock, and heavy-metal toxicity in mammalian cell culture and yeast. While the RNase A endonuclease Angiogenin cleaves some tRNAs into tRFs in mammalian cell culture (Fu et al., 2009), the stress-mediated biogenesis of most tRF species is not dependent on Angiogenin, suggesting that multiple factors contribute to the cleavage of specific tRNAs (Su, Kuscu, Malik, Shibata, & Dutta, 2019). For example, in yeast, the RNase T2 endonuclease RNY1p mediates tRNA cleavage in response to oxidative stress (Thompson & Parker, 2009). Stress-derived tRNA fragments are generally defined as being 3′ - or 5′ -tRNA halves, and arise as result of a cleavage event occurring within or proximal to the anticodon loop. While these endonucleases have been implicated in the stress-mediated biogenesis of tRFs, the factors directing the endogenous biogenesis of tRFs in either the physiological context of germline maturation or the pathological context of disease are poorly characterized. These endogenous tRFs include tRNA halves, but can also include smaller fragments characterized by cleavage sites within the other loops of the tRNA "clover-leaf" structure.

The most well characterized example of tRFs regulating cellular function is their role in the inhibition of translation under

stress conditions. Angiogenin-induced tRNA halves, such as the $tRNA^{Ala}$ 5′ half, have been shown to promote stress granule assembly in human cells, thereby inhibiting translation by segregating mRNA transcripts and translation initiation factors away from the ribosome (Emara et al., 2010; Lyons, Achorn, Kedersha, Anderson, & Ivanov, 2016; Yamasaki, Ivanov, Hu, & Anderson, 2009). Additionally, 5′ tRFs from $tRNA^{Pro}$, $tRNA^{Val}$, and $tRNA^{Gln}$ can directly interact with the ribosome, resulting in the attenuation of translation due to ribosome stalling or competition with mRNA (Gebetsberger, Wyss, Mleczko, Reuther, & Polacek, 2017; Gonskikh et al., 2020; Sobala & Hutvagner, 2013). Further, 5′ tRFs from $tRNA^{Gln}$ have been demonstrated to bind to and disrupt the stability of the Multisynthetase Complex, leading to inhibition of ribosome maturation and global translational repression (Keam, Sobala, Ten Have, & Hutvagner, 2017; Mleczko, Celichowski, & Bąkowska-Żywicka, 2018). Interestingly, a $tRNA^{Thr}$ 3′ half has been implicated in promoting—rather than inhibiting—translation following nutrient deprivation in the parasite *T. brucei*, suggesting that tRFs play diverse roles in translational regulation (Fricker et al., 2019). Under normal, non-stress conditions, endogenously expressed tRFs have also been linked to retrotransposon regulation, particularly in the context of the early embryo and pluripotent stem cells (Schorn, Gutbrod, Leblanc, & Martienssen, 2017; Sharma et al., 2016). Some tRFs control LTR-retrotransposon transcription by interfering with the tRNA primer binding site that is essential for reverse transcription of the retrotransposon (Schorn et al., 2017). Other reports indicate that a Gly-GCC-derived tRF can regulate MERVL-retroelement-associated genes by influencing the levels of the U7 snoRNA, which, in turn, regulates histone gene levels and influences chromatin state (Boskovic, Bing, Kaymak, & Rando, 2020). tRFs have also been suggested to participate in post-transcriptional gene silencing through mechanisms similar to miRNAs or siRNAs (Chen et al., 2016; Gustafsson et al., 2022; Molla-Herman et al., 2020; Peng et al., 2012; Sharma et al., 2016, 2018; Shin et al., 2021; Zhang et al., 2018).

Notably, because mature tRNAs contain a high percentage of modified nucleotides (∼17% on average across all three domains of life) and tRFs appear to be derived from mature tRNAs, tRFs are predicted to contain a high level of internal RNA modifications relative to other small RNA species (Gustafsson et al., 2022; Jackman & Alfonzo, 2013). These RNA modifications on tRFs can prevent efficient cDNA generation by standard reverse transcriptases, thereby resulting in reduced detection of these small RNA species in small RNA sequencing datasets prepared with standard methods. Furthermore, as will be discussed below, the endonucleolytic cleavage of tRNAs into tRFs by RNase A and RNase T2 enzymes can leave distinct 2′-3′ -cyclic phosphate groups on tRFs' 3′ termini and dephosphorylate the 5′ termini, both of which impede the ligation of ssRNA adapters during small RNA cloning. This limitation in ligation of adapters to tRFs can result in reduced detection of these small RNAs when using standard protocols. It was previously unclear whether the enrichment of particular tRF species in small RNA datasets reflects the endogenous abundance of tRFs or whether the chemical or physical structure of tRFs constrains efficient cloning of certain fragments. Recent work to understand the difficulties in cloning tRFs has demonstrated that additional enzymatic treatments and the use of alternative reverse transcriptase enzymes must be used to capture the full breadth of tRF species, as discussed in Section: Small RNA Characteristics Affecting Standard Protocols.

rRNA fragments (rRFs)

Ribosomal RNA fragments (rRFs) have also been identified as abundant small RNAs in sequencing datasets. These include fragments derived from all six rRNAs, four nuclear rRNAs (18S, 5.8S, 28S, 5S) and two mitochondrial rRNAs (12S, 16S), with highly abundant rRFs found to originate from specific "hotspots" for each rRNA (Cherlin et al., 2020). Interestingly, like tRFs, rRFs specifically accumulate in mammalian sperm undergoing the final steps of maturation in the epididymis. Accounting for approximately 60% of the rRF population in sperm, are those derived from the 28S-rRNA precursor (Chu et al., 2017; Hua et al., 2019). rRNAs also undergo nucleotide modifications at approximately 2% of nucleotides, with 2′-*O*-methylation of ribose (at any nucleotide) and isomerization of uridine to psuedouridine (Ψ) being the most prevalent types of modifications in eukaryotes (Sloan et al., 2017). In yeast, 55 2′-*O*-methylation sites and 45 Ψ sites have been identified, while, in humans, hundreds of each modification have been reported (Birkedal et al., 2015; Lestrade & Weber, 2006; Piekna-Przybylska, Decatur, & Fournier, 2008; Taoka et al., 2016). These

modifications can represent a potential barrier for small RNA cloning, as modified bases can prevent canonical reverse transcription (RT) enzymes from synthesizing cDNA through these modifications, or lead to misincorporated bases and errors in sequencing data. Indeed, recent efforts to circumvent these issues have been developed, such as a small RNA cloning protocol called PANDORA-seq that removes modified nucleotides and has revealed an extensive repertoire of rRFs in diverse cell-types at previously unappreciated levels (see below) (Shi et al., 2021).

METHODS TO CLONE AND SEQUENCE SMALL RNAs

Small RNA cloning techniques have been used to identify and quantitate small RNAs since their discovery, over 20 years ago. Initially, these techniques were low throughput, relying on the purification of small RNAs by size selection, followed by the ligation of adapters, the generation of cDNA from the ligated molecules, and finally, identification by Sanger sequencing, to randomly identify hundreds, or at most thousands, of captured small RNA sequences (Ambros, Lee, Lavanway, Williams, & Jewell, 2003; Lagos-Quintana, Rauhut, Lendeckel, & Tuschl, 2001; Lagos-Quintana, Rauhut, Meyer, Borkhardt, & Tuschl, 2003; Lau, Lim, Weinstein, & Bartel, 2001; Lee & Ambros, 2001; Lim et al., 2003; Llave, Kasschau, Rector, & Carrington, 2002; Reinhart, Weinstein, Rhoades, Bartel, & Bartel, 2002). The throughput of these techniques was higher than that of northern blotting, the standard small RNA biochemical detection technique at the time, and allowed for the identification of new small RNA species without any prior knowledge of their sequence. The full diversity of small RNA populations, however, would not begin to be revealed until the introduction of high-throughput sequencing. High-throughput small RNA-sequencing was first employed to study plant small RNAs and, soon after, to document new small RNA species in *C. elegans* (Lu, Tej, Luo, Haudenschild, & Green, 2005; Ruby et al., 2006). The adoption of high-throughput sequencing revolutionized small RNA cloning, as evidenced by the identification of thousands of new small RNA species in *C. elegans* through the generation of 400,000 short-read sequencing reads from cloned small RNA libraries (Ruby et al., 2006). The first small RNA-seq protocols used a similar strategy to that mentioned above, involving the gel purification of small RNAs, followed by ligation of adapter sequences, cDNA synthesis, and finally, the addition of sequencing primers to the cDNA during PCR amplification. The resulting small RNA library was subsequently sequenced and then analyzed by computational pipelines to identify and quantify small RNA reads. Since these original studies, there has been a steady progression of small RNA cloning and sequencing techniques developed to address the limitations of the traditional approach, as discussed below.

Selection and Purification of Small RNAs

The first obstacle in cloning and quantitating small RNAs is separating them from the rest of the RNA present in the cell. Small RNA can be purified or enriched from total RNA samples via various approaches, including denaturing polyacrylamide gel electrophoresis (PAGE) purification, removal of high molecular RNAs by polyethylene glycol (PEG), filtration of RNA species based on molecular weight, or the use of commercial kits (e.g., mirVana miRNA isolation kit) (Fig. 2) (Guo et al., 2014). While enrichment of small RNAs from the total RNA pool increases the likelihood of successfully cloning small RNAs, and can also be more cost efficient (as less reagents are required than with total RNA input), such methods are subject to a substantial loss of material during the purification process. Further, methods used for small RNA purification (e.g., PAGE) can be laborious. Therefore, total RNA input is commonly adopted for library generation. In cases where total RNA is used as an input and small RNAs are not first purified, size selection by non-denaturing PAGE can be performed following cDNA synthesis to purify the pool of cDNAs representing cloned, adapter-ligated small RNA constructs. Nevertheless, obtaining RNA with as little degradation as possible is critical for quality small RNA sequencing and data analysis.

Ligation-Dependent Methods

One of the most common procedures used to clone small RNA libraries is a ligation-dependent method that relies on T4 RNA ligases to attach 3′ and 5′ ssRNA adapters that provide an anchor to prime cDNA generation and to amplify the ligated molecules via PCR (Fig. 2 – Ligation-dependent). 3′ adapter ligation occurs through an ATP-independent mechanism mediated by a truncated T4 RNA ligase 2 (1-249). This truncated T4 RNA ligase

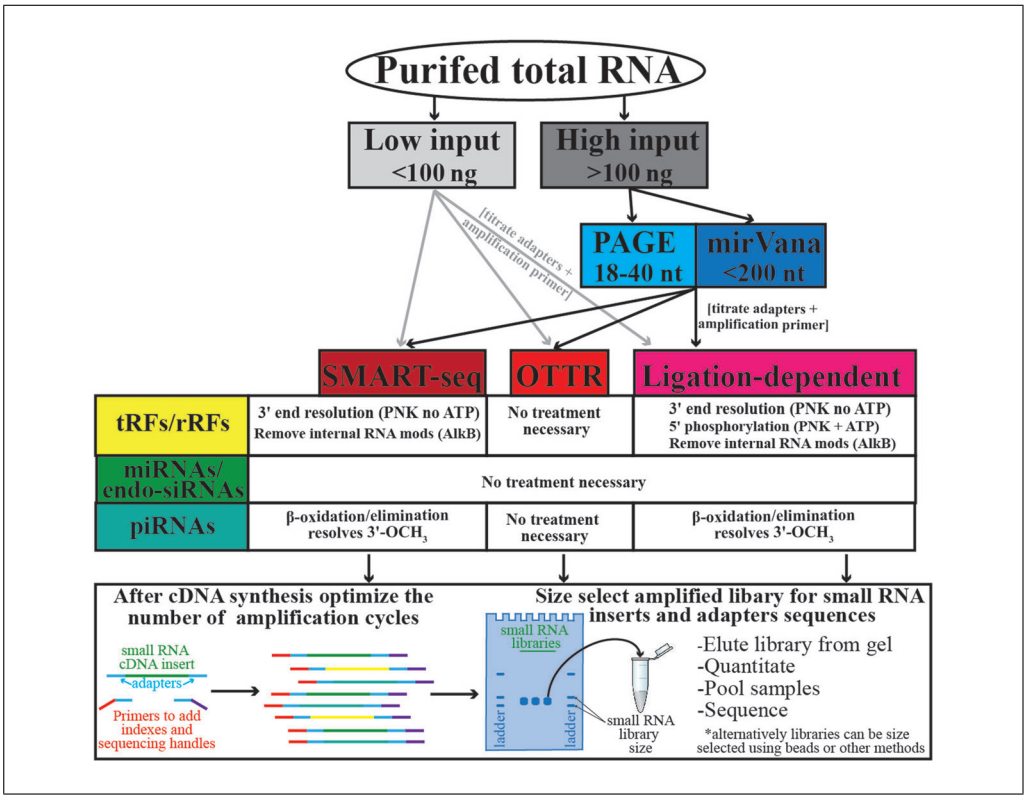

Figure 2 Flowchart for small RNA cloning and sequencing experiments. When performing small-RNA-sequencing experiments, the amount of input RNA should be assessed first, to determine the protocol and whether size selection of the small RNA pool can be effectively performed (Section: Selection and purification of small RNAs). If using a ligation-dependent approach, it is important to consider titrating adapters and amplification primers with the amount of RNA input. Excess adapters and primers can generate sequenced molecules without any small RNA insert, which decreases the usable data in small RNA-seq experiments. Generally, three different protocols are available to generate small RNA libraries, namely, those centered around 5′- and 3′-ligation (ligation-dependent; Section: Ligation-dependent methods), protocols that bypass ligation (ligation independent e.g., SMART-seq, Section: Ligation-independent methods), and a recently developed method using an alternate reverse transcriptase (OTTR; Section: Ligation-independent methods). Each of these protocols have limitations that can depend on the small RNA of interest and are discussed throughout this overview. Further, due to the varying chemical signatures of the 5′- and 3′-ends of different small RNAs, certain enzymatic treatments or library generation protocols should be considered (e.g., OTTR) to ensure small RNAs are efficiently cloned. Following the generation of complementary DNA (cDNA), the number of amplification cycles should be optimized before size selection of the amplified libraries (by non-denaturing PAGE, bead selection or other methods) and pooling of samples ready for sequencing.

catalyzes the formation of a phosphodiester bond between the 5′-monophosphate of the 3′ ssRNA adapter and the 3′-hydroxylated end of a small RNA, resulting in a 3′-adapter-ligated small RNA (Viollet, Fuchs, Munafo, Zhuang, & Robb, 2011). The 3′ ssRNA adapter must be pre-adenylated for this ligation reaction to occur, because T4 RNA ligase requires an adenylated substrate to act as the nucleophile during phosphodiester bond formation and the truncated T4 ligase lacks the activity to adenylate its substrates. Conversely, 5′ adapter ligation utilizes wildtype T4 RNA ligase 2, which, once activated by ATP, forms a new phosphodiester bond between the 5′-monophosphate of the 3′ ligation product and the 3′ -hydroxylated end of the 5′ ssRNA adapter. The wildtype T4 ligase retains its ability to adenylate substrates, allowing it to adenylate the 3′ ligation product (5′ end of cloned small RNA) and generate an adenylated (activated) precursor ligation product. This precursor ligation product can then perform a nucleophilic attack on the 5′ adapter, thereby creating a small RNA with adapters on both ends. After the adapters are attached, cDNA can be generated using reverse transcriptase and a primer that binds the 3′ adapter sequence. Finally, the cDNA is amplified via PCR using primers binding the sequences added by the

5′ and 3′ adapters to generate sufficient DNA to sequence.

The truncated T4 RNA ligase is preferred over the wildtype T4 RNA ligase for 3′ adapter ligation because the wildtype T4 RNA ligase's adenylyltransferase activity results in off-target adenylation of small RNA substrates. These adenylated small RNA substrates can then act as nucleophiles during T4 ligase-mediated phosphodiester bond formation. The 5′ -monophosphate of an adenylated small RNA substrate can form a phosphodiester bond with its own 3′ -terminal hydroxyl group or the 3′ -terminal hydroxyl group of another small RNA, resulting in the formation of circularized and dimerized byproducts that reduce the yield of the 3′ adapter ligation step. Truncated T4 RNA ligase 2 limits the production of these byproducts because it is incapable of adenylating small RNA substrates and is, therefore, preferred for the 3′ ligation step.

The application of two-adapter ligation methods, while common, is accompanied by several caveats, notably the introduction of ligation bias resulting from differences in the 5′ and 3′ terminal nucleotides. This bias leads to sequencing results that do not accurately reflect the "true" cellular levels of small RNAs present in a sample (Raabe, Tang, Brosius, & Rozhdestvensky, 2014). While bias introduced by ligation should not prevent small RNA quantification comparisons between samples subjected to the same library preparation protocol, this does influence the ability to accurately perform other comparisons across and within sequencing experiments. For example, this bias influences the ability to assess the levels of miRNA and miRNA* strand in the same sample, or expression of small RNAs from two datasets employing different library preparation approaches (Sorefan et al., 2012). Methods aiming to mitigate ligation bias have been developed, including the use of adapters with randomized bases at the ligation junction (Table 1). This technology addresses ligation biases by using a pool of adapter sequences to increase the chance of efficient ligation of any given small RNA. Indeed, this protocol can robustly detect miRNAs that are not captured by using a single set of adapter sequences (Dard-Dascot et al., 2018; Giraldez et al., 2018). Comparative studies have demonstrated that the use of randomized adapters continually outperforms the traditional two-adapter ligation strategy (NEXTFLEX small RNA-seq Kit v3, PerkinElmer) (Dard-Dascot et al., 2018; Wright et al., 2019).

Ligation-Independent Methods

In addition to using adapters with randomized bases at the ligation junction, a promising approach to alleviate ligation bias is to bypass ligation altogether. Some ligation-independent methods employ artificial polyadenylation of the 3′ end of the small RNA by poly(A) polymerase to create an anchor sequence for reverse transcriptase to synthesize cDNA. Original iterations of this protocol involved polyadenylation of the 3′ end and subsequent ligation of the 5′ end adapter. However, ligation at the 5′ end is often inefficient and particularly challenging. Hence, more recently, novel small RNA library preparations have replaced 5′ ligation with the template switching activity of the MMLV (Moloney murine leukemia virus) reverse transcriptase (Wulf et al., 2019; Zhu, Machleder, Chenchik, Li, & Siebert, 2001). During first-strand synthesis, the terminal transferase activity of this reverse transcriptase adds cytidine residues at the 5′ end of each cDNA molecule, serving as 5′-end adapter. Ultimately, these activities generate double-stranded cDNA with 5′ and 3′ adapter sequences attached to the adenylated small RNA template (Fig. 2 – SMART-seq). In line with the motivation that drove the development of this method, the main advantage lies in the ligation independence, which removes bias to allow for better small RNA quantification. Additionally, these methods are highly sensitive, generating high quality libraries from small amounts of input RNA (as little as 1 ng). Since this method of library preparation facilitates the use of low quantities of input RNA, it has been frequently used to make single-cell mRNA–seq libraries; however, other strategies have been employed for single-cell small RNA-seq studies thus far (Ramsköld et al., 2012). While providing advantages in regard to ligation bias and low input requirement, the MMLV template switching method is not without disadvantages. The untemplated adenosines added to the 3′-end and the cytidines added to the 5′-end make it difficult to distinguish native nucleotides from those that are added during the cloning procedure, and can contribute to discarded reads during sequence mapping and the general inability to accurately determine the exact termini of cloned sequences. Currently, two companies, Diagenode (D-plex Small RNA-seq Kit) and Takara Bio (SMARTer smRNA-seq Kit) manufacture kits that offer ligation free methods (Table 1).

Recently, a new cloning technique, Ordered Two-Template Relay (OTTR), has been

Table 1 Commercially Available Protocols for Small RNA Isolation and Cloning

Technology	Product name	Catalog number	Company	Notes
***mir*Vana miRNA isolation kit**	*mir*Vana miRNA isolation kit	AM1560	Invitrogen, ThermoFisher	Enriches for small RNA (<200nt)
Two-adapter ligation	TruSeq small RNA library prep kit	RS-200-0012 (set A indexes) RS-200-0024 (set B indexes) RS-200-0036 (set C indexes RS-200-0048 (set D indexes) – all 24 reactions	Illumina	
	NEBNext multiplex small RNA library prep kit	E7300S (24 reactions) E7300L (96 reactions)	New England Biolabs	
	CleanTag® small RNA library prep kit	L-3206	TriLink BioTechnologies	
	Small RNA-seq library prep kit	052.08 (8 preps) 052.24 (24 preps) 052.96 (96 preps)	Lexogen	
	NEXTFLEX small RNA-seq kit v3	NOVA-5132-05 (8 reactions) NOVA-5132-06 (48 reactions)	PerkinElmer	Utilizes randomized adapters
	Small RNA library prep kit	63600 (Indexes 1-24) 63620 (Indexes 25-48)	Norgen Biotek Corp.	
	MGIEasy small RNA library prep kit	940-000196-00	MGI	
	QIAseq miRNA library kit	331502 (12 reactions) 331505 (96 reactions)	Qiagen	Integrated UMIs (12 bp)
Ligation-independent (polyadenylation and template switching)	SMARTer smRNA-seq kit	635029 (12 reactions) 635030 (48 reactions) 635031 (96 reactions)	Takara Bio	
	D-plex small RNA-seq kit	C05030001	Diagenode	

developed that also bypasses the requirement of adapter ligations to clone small RNAs (Upton et al., 2021). OTTR uses a N-terminal truncated R2 RT enzyme, termed BoMoC, encoded by a non-LTR retroelement from *Bombyx mori* (silk moth), which can jump from one synthesis template to another. Hence, BoMoC is used to add both 5′ and 3′ adapters and generate cDNA from small RNAs in a single tube reaction (Fig. 2 - OTTR). Recent data suggests that OTTR can outperform all current ligation-dependent and independent protocols in recovering accurate levels of miRNAs using a standardized miRNA library (Upton et al., 2021). Additionally, OTTR has been used to assay the levels of tRFs in both yeast and mammalian spermatozoa, demonstrating the true representation of different RNAs in these samples (see below). While the BoMoC enzyme is not currently commercially produced, when available for purchase, this technique could soon become the gold standard for small RNA cloning protocols.

UMIs and Quantitative Counting

Small RNAs perform their cellular functions and regulate gene expression by interacting with other biomolecules. Thus, the functions of small RNAs are dependent on their cellular or subcellular concentration and stoichiometry relative to their interactors, in addition to the binding affinities (K_d) of these interactions. To identify small RNAs and regulatory targets that are functional

in a given cell or tissue, it can be useful to quantitate, in an absolute manner, the number of molecules present. However, typical small RNA-sequencing experiments only quantify the small RNAs relative to one another because standard cloning procedures produce libraries that represent a sampling of the RNAs that are present in a sample. Additionally, most current protocols for making libraries for small RNA-seq rely on amplification by PCR of the cDNA generated from the small RNA pool. Amplification can result in "jackpotting" events from sampling of cDNA representing small RNAs. For example, if a rare or low abundant RNA is efficiently amplified during early PCR cycles, it can become overrepresented in the final sequencing data. Because the entire sequence of the small RNA is determined in a single read during standard short-read sequencing, as opposed to mRNA-seq, where fragments of the RNA are sequenced (which can be used to differentiate unique molecules from amplification products), it is impossible to determine if multiple reads from the same small RNA sequence represent unique molecules or PCR amplification products.

To address the issue of duplicate reads generated by PCR amplification, randomized sequences present in the adapters can be added to the cloned small RNAs during ligation or cDNA synthesis (Fu, Wu, Beane, Zamore, & Weng, 2018). These randomized sequences, referred to as Unique Molecular Identifiers (UMIs), "barcode" individual molecules prior to PCR amplification. During computational analysis after sequencing, reads with the same UMI are then used to infer PCR duplicates and, thus, are collapsed into the same count for that small RNA. From these analyses, UMIs can be used to determine the relative number of molecules in a given sample (Kivioja et al., 2011). If the number of cells the RNA has been collected from is known, and the concentration of the purified RNA is accurately determined prior to small RNA cloning, then read counts determined with UMIs can be used to infer the number of molecules of that RNA present in a cell (Ziegenhain, Hendriks, Hagemann-Jensen, & Sandberg, 2022). However, it should be noted that these types of analysis assume that all the cells in a sample have an equivalent number of small RNA molecules. It is currently unclear if this assumption is true, as single-cell small RNA-seq techniques are in nascent stages of development and optimization (see below).

To accurately quantify the absolute number of small RNA molecules in a sample, spike-in RNA standards can be used (Locati et al., 2015; Lutzmayer, Enugutti, & Nodine, 2017). In this approach, RNAs with a known sequence (and not present among the small RNAs being sequenced) are added to total RNA in a proportion relative to the numbers of cells used as input material. For example, if total RNA is isolated from 1 million cells, 1 million molecules of RNA standard A, 10 million molecules of RNA standard B, and 100 million molecules of RNA standard C are spiked into the purified RNA. After cloning and sequencing, the number of reads mapped to each of these standards represent 1, 10, and 100 molecules, respectively. Thus, a standard curve is generated, which should span the magnitude of small RNA abundance present in the sample being studied, to absolutely quantitate the number of molecules of a given small RNA, based on the number of reads mapped to that sequence (Fahlgren et al., 2009; Locati et al., 2015; Sharma et al., 2016). While spike-ins are an effective way to absolutely quantify small RNA abundance, they also are sensitive to the same biases from ligations and PCR amplification that afflict the cloning of small RNAs (Sharma et al., 2016). Accordingly, the most effective and accurate absolute quantitation approach comes from pairing the incorporation of both UMIs and spike-in standards to small RNA cloning protocols (Gainetdinov et al., 2021).

SMALL RNA CHARACTERISTICS AFFECTING STANDARD PROTOCOLS

Internal RNA modifications and the chemical structure of the 5′ or 3′ end of the mature small RNA of interest can both affect its detection via small RNA cloning by interfering with cDNA synthesis and ligation, respectively.

Most small RNAs undergo a number of post-transcriptional processing steps, including 5′ and 3′ end processing during RNA maturation or turnover, which differs for each small RNA class (Ji & Chen, 2012) and can influence the "clonability" of certain types of small RNAs when using standard cloning procedures involving 5′ and 3′ ligation. For example, a class of *C. elegans* endo-siRNAs, 22G-RNAs, carry a 5′-triphosphate that is incompatible with 5′ ligation (Pak & Fire, 2007). Additionally, modifications made to the ribose backbone of terminal RNA nucleotides can interfere with several steps

in library construction. miRNAs in plants and piRNAs throughout the metazoan lineage harbor a 2′-*O*-methylation of the ribose moiety of the 3′ terminal nucleotide. While this modification affords these small RNAs increased stability, it severely impacts 3′ ligation efficiency (Munafó & Robb, 2010). Therefore, cloning of these small RNAs requires strategies to generate cloneable products prior to ligation. Indeed, several enzymatic and chemical treatments have been developed to remove these modifications to improve their capture during cloning procedures (Table 2).

Further, cellular RNAs, including some small RNAs, are decorated with an elaborate array of RNA modifications, such as (but not limited to) methylation (including m^1A, m^3C, m^1G, and $m^2{}_2G$), pseudouridine (Ψ), 8-oxo-7,8-dihydroguanosine (8-oxoG), and N^4-acetylcytidine (ac^4C), all of which should be considered when cloning and generating small RNA-seq libraries. Structural RNAs such as tRNAs and rRNAs are particularly enriched for RNA modifications (Jackman & Alfonzo, 2013). For example, tRNAs are modified at ∼17% of their nucleotides, with >100 different types of modifications, including some that are ubiquitous across all tRNAs and some that are specific for different isoacceptors. rRNAs, on the other hand, carry modifications on ∼2% of their nucleotides (Cherlin et al., 2020). These modifications play an important role in RNA biogenesis, stability, and function and, therefore, have become an important area of research (Oberbauer & Schaefer, 2018). RNA modifications, however, present a challenge for RNA-seq protocols that involve generating cDNA (which encompass nearly all current methods apart from direct RNA-sequencing; see below), as the canonical reverse transcriptases used for cDNA generation cannot effectively synthesize DNA from modified bases, consequently resulting in inefficient or incomplete conversion to cDNA. Thus, RNA modifications can lead to the mis-incorporation of complementary nucleotides during cDNA synthesis or even block elongation altogether, thereby generating polymorphisms and truncated sequencing reads, respectively. Currently, two main approaches are taken to overcome the challenges of sequencing highly modified small RNAs: (a) enzymatic pretreatment to remove the modifications or (b) the use of alternative RT enzymes that are able to read through such modifications. In the following sub-sections, we first address the enzymatic or chemical pre-treatments that can be used to "resolve" small RNAs with modified 5′ or 3′ ends for adapter ligation, and then detail the enzymatic pretreatments and non-canonical RT enzymes that can be used to detect small RNAs with internal RNA modifications.

Endo-siRNAs in Worms – RppH Treatment

One class of small RNA with low cloning efficiency using standard methods is a type of endo-siRNA called 22G-RNAs. 22G-RNAs are specific to nematodes and are transcribed by RNA-dependent RNA polymerases (RdRPs) using mRNAs as a template; subsequently, they are antisense to and target their cognate mRNAs. Synthesis by RdRP leaves a triphosphate group at the 5′ end of the 22G-RNA (Billi, Fischer, & Kim, 2014). As discussed above, ligation of a 5′ adapter during small RNA cloning involves a nucleophilic attack of the 5′-monophosphate of the small RNA to the 3′ of a ssRNA 5′ adapter. 22G-RNAs have reduced cloning efficiency in traditional small RNA cloning procedures because the 5′-triphosphate group of a 22G-RNA does not complete this reaction as efficiently as a 5′-monophosphate group. As such, this endo-siRNA subclass is underrepresented in small RNA libraries prepared using traditional cloning procedures.

To increase the clonability of 22G-RNAs, small RNA samples can be enzymatically pre-treated to transform the 5′ triphosphate to a cloneable monophosphate. Tobacco Acid Pyrophosphate (TAP) was the first enzyme used to hydrolyze the phosphodiester bonds of the 5′-triphosphate to create a cloneable 5′-monophosphate end (Gu et al., 2009). However, after TAP was commercially discontinued, the primary enzyme currently used is RNA 5′ Pyrophosphohydrolase (RppH) (Almeida, De Jesus Domingues, Lukas, Mendez-Lago, & Ketting, 2019).

3′-Terminal 2′-O-Methylation of piRNAs – Oxidation of 3′ Ends

piRNAs are another example of a small RNA class exhibiting limited cloning efficiency under standard small RNA cloning procedures due to their chemical characteristics. As discussed above, piRNAs have a 2′-*O*-methyl-modified (2′-OCH_3) 3′ terminus. The 3′ methylation is thought to protect the piRNA from 3′-5′ degradation and 3′ uridylation (tailing with uridine), and is highly conserved across animal species (Pastore, Hertz, Price, & Tang 2021). Notably, HEN1-mediated 3′ methylation of small RNAs is present in

Table 2 Description of the Chemical Modifications Displayed by Different Small RNA Subtypes, How those Modifications Interfere With the Cloning Process, and the Methods That Can be Used to Either Improve Cloning or Specifically Enrich for Certain Subtypes

Location	Chemical modification	Small RNA subtype example	Interferes with	Recommended treatment	Treatment molecular action	Objective
5′ end	5′-Triphosphate group	22G-RNAs (*C. elegans* endo-siRNA subtype)	5′ adapter ligation	RNA 5′ Pyrophosphohydrolase (RppH)	Dephosphorylation of 5′-triphosphate to 5′-monophosphate	Improved detection of 22G-RNAs
	5′-hydroxyl	tRNA fragments (tRFs) – 3′ cleavage products	5′ adapter ligation	T4 polynucleotide kinase (PNK) (3′ phosphatase minus) with ATP	Phosphorylation of 5′-hydroxyl	Improved detection of 3′ tRFs
3′ end	2′-3′-cyclic phosphate (cP)	tRNA fragments (tRFs) – 5′ cleavage products	3′ adapter ligation	T4 polynucleotide kinase (PNK) No ATP	3′-cP removal (resolved to 3′-hydroxyl)	Improved detection of 5′ tRFs
				(1) Acid hydrolysis (typically with HCl) followed by (2) phosphatase treatment	(1) Hydrolysis of 3′-cP to 3′-phosphate (2) Dephosphorylation of 3′-phosphate to 3′ -hydroxyl	Improved detection of 5′ tRFs
				cP-RNA-seq: (1) phosphatase treatment followed by (2) periodate oxidation and (3) subsequent PNK treatment	(1) Removal of 3′-phosphate from all RNAs (2) 3′ cleavage of non-cP RNAs (resulting in degradation) (3) 3′ -cP removal	Enrichment of cyclic phosphate containing small RNAs (mainly 5′ tRFs)

(*Continued*)

Table 2 Description of the Chemical Modifications Displayed by Different Small RNA Subtypes, How those Modifications Interfere With the Cloning Process, and the Methods That Can be Used to Either Improve Cloning or Specifically Enrich for Certain Subtypes, *continued*

Location	Chemical modification	Small RNA subtype example	Interferes with	Recommended treatment	Treatment molecular action	Objective
				Arabidopsis thaliana tRNA ligase (AtRNL) treatment	Selective ligation of 3′ adapter to a 3′-cP	Enrichment of cyclic phosphate containing small RNAs (mainly 5′ tRFs)
	2′-O-methylation	piRNAs (animals), miRNAs and siRNAs (plants), Ago2-associated siRNAs (*D. melanogaster*)	3′ adapter ligation	Periodate oxidation followed by β-elimination	3′ cleavage of non-methylated RNAs (resulting in degradation)	Enrichment of piRNAs (animals), miRNAs and siRNAs (plants), or Ago2-associated siRNAs (*D. melanogaster*)
Internal RNA modification	1-methyladenosine, 3-methylcytidine, and 1-methylguanosine)	tRNA fragments (tRFs) and rRNA fragments (rRFs)	cDNA synthesis	a-ketoglutarate-dependent hydroxylase (AlkB) treatment (ARM-seq)	RNA demethylation	Detection of methylated small RNAs (mainly tRFs and rRFs)
				PANDORA-seq combines AlkB treatment with PNK treatment	RNA demethylation and 3′ -cP removal	Detection of methylated small RNAs also containing cyclic phosphates (mainly tRFs and rRFs)
	Internal RNA modifications and RNAs with unknown termini status	tRNA fragments (tRFs) and rRNA fragments (rRFs)	cDNA synthesis	Ordered Two-Template Relay OTTR	Template jumping of reverse transcriptase to bypass ligation-based adapter addition and cDNA synthesis through modified RNA bases	Capture RNAs with complicated termini and modified RNAs without enzymatic treatment

plants, but, there, the small RNA substrates are miRNAs and siRNAs rather than piRNAs (Yang, Ebright, Yu, & Chen, 2006; Yu et al., 2005). Similarly, *Drosophila melanogaster* Ago2-associated siRNAs are substrates of the *Drosophila* HEN1 homolog in addition to piRNAs (Horwich et al., 2007; Kingston & Bartel, 2021). This *O*-methyl-modified 3′ terminus is known to interfere with 3′ adapter ligation, thereby limiting the representation of piRNAs and other similarly methylated RNAs in small RNA libraries (Dard-Dascot et al., 2018; Munafó & Robb, 2010).

The 2′ *O*-methylation modification of piRNAs endows these RNAs with a greater resistance to oxidation compared to non-methylated RNAs. Therefore, small RNA libraries can be enriched for piRNAs by performing periodate oxidation and ß-elimination reactions on small RNA samples prior to initiating the cloning procedure (Fig. 2) (Huang, Yoshitake, & Asakawa, 2021; Kurth & Mochizuki, 2009; Roovers et al., 2015). Periodate oxidation of RNAs with a free 2′-OH group followed by β-elimination results in nucleotide-by-nucleotide RNA degradation at the 3′ end. In these reactions, RNAs with 2′-OH groups at their 3′ ends are largely degraded, while piRNAs with 2′-OCH_3 groups are protected and, therefore, enriched for downstream cloning procedures. Importantly, while performing periodate oxidation and β-elimination in animal samples enriches for piRNAs, performing the same treatment on plant or fly samples enriches for other types of small RNAs that contain a 2′-OCH_3 group in these organisms, namely miRNAs and endo-siRNAs.

2′-3′ Cyclic Phosphates - PNK Treatment

A 2′-3′ -cyclic phosphate (2′-3′-cP) is primarily found on the 3′ termini of RNAs generated from cleavage by some ribonucleases and is present as a stable modification on certain cellular RNAs (Licht, Medenbach, Lührmann, Kambach, & Bindereif, 2008; Sporn, Lazarus, Smith, & Henderson, 1969). The endoribonucleolytic cleavage of mature tRNAs by RNaseA and RNaseT2 enzymes produces tRFs with a 2′-3′-cP at the 3′ end of the 5′ cleavage product, which can interfere with small RNA cloning protocols (Shigematsu, Kawamura, & Kirino, 2018). Since the enzymes commonly used for ligation of RNAs require substrates with a 3′ hydroxyl group, 3′ terminal cyclic phosphates need to be resolved for recovery of these RNAs during cloning. In the absence of ATP, T4 polynucleotide kinase (PNK), through its phosphatase activity, resolves and removes cyclic phosphates at the 3′ end of RNAs to facilitate subsequent ligation (Fig. 2). Indeed, PNK treatment prior to small RNA cloning has proven an effective method in uncovering an unappreciated population of tRFs in male germ cells (Gustafsson et al., 2022; Sharma et al., 2018). Further, such enzymatic treatment in differentiated mouse embryonic stem cells revealed a modest increase in tRF 5′ cleavage products (Krishna et al., 2019). Alternatively, cyclic phosphates can also be resolved prior to ligation via treatment with an acid, such as hydrochloric acid, that first hydrolyzes the 2′-3′-cP to a 3′-monophosphate, combined with subsequent treatment with a phosphatase (commonly calf intestine phosphatase), which removes the 3′-P (Lund & Dahlberg, 1992). Endoribonucleolytic cleavage can also result in RNAs lacking a 5′-monophosphate, which would thus evade cloning; this includes the 3′ tRF cleavage product. However, in the presence of ATP, PNK end treatment can also be utilized to catalyze the phosphorylation of 5′ -OH to generate 5′ -monophosphate, which can be successfully ligated and cloned (Fig. 2).

If small RNAs containing cyclic phosphates are of specific interest, protocols have been developed using total cellular RNA that can selectively capture and sequence these RNAs. One such method, called cP-RNA-seq, involves treatment with a phosphatase to remove 3′-monophosphate from all RNAs, which is then followed by periodate oxidation. The oxidation step cleaves all 3′ ends with a free 3′-OH, leaving only cP-containing RNAs, which can then be subjected to cP removal via T4 PNK treatment to allow exclusive ligation and amplification of cP-containing RNAs (Honda, Morichika, & Kirino, 2016). An alternative method to selectively capture cP-containing RNAs uses the *Arabidopsis thaliana* tRNA ligase, AtRNL. This enzyme exclusively ligates 3′ adapters to cP-RNAs, exhibiting substrate specificity to 3′-cP termini over 3′-OH or 3′-P termini (Schutz, Hesselberth, & Fields, 2010).

tRFs and tRNA Modifications-Enzymatic Treatment for Removal of Modifications

Owing to the enrichment of RNA modifications in tRNAs and rRNAs, small RNAs generated from these RNAs (tRFs and rRFs) encompass the most highly modified small RNAs (Jackman & Alfonzo, 2013; Lopez

Sanchez, Cipullo, Gopalakrishna, Khawaja, & Rorbach, 2020; Phizicky & Hopper, 2010). In particular, the prevalence of RNA modifications on tRFs is hypothesized to drive the apparent enrichment of 5′ tRF fragments in small RNA-seq datasets produced from mammalian sperm by limiting the detection of highly modified 3′ fragments (Sharma et al., 2018; Zhang et al., 2018). While 3′ fragments are vastly underrepresented in most small-RNA seq data, their presence is detectable by northern blot, indicating that there is a more equal endogenous representation of these 5′ and 3′ fragments than what is represented by sequencing. Therefore, in order to capture the entire complement of tRFs and rRFs while performing small RNA-seq, one approach to clone these RNAs efficiently is the enzymatic removal of RNA modifications. Treatment of RNA with the *Escherichia coli* dealkylating enzyme α-ketoglutarate-dependent hydroxylase (AlkB) removes RNA methylation (specifically 1-methyladenosine, 3-methylcytidine, and 1-methlyguanosine), allowing for successful cDNA generation of methylated small RNAs (Delaney & Essigmann, 2004; Hrabeta-Robinson, Marcus, Cozen, Phizicky, & Lowe, 2017). Indeed, AlkB-facilitated RNA methylation sequencing (ARM-seq) has revealed previously undetected methylated small RNAs, mainly tRFs (Fig. 2) (Cozen et al., 2015). Conveniently, not only does this technique uncover otherwise undetected small RNAs, but comparison of sequencing results with untreated RNA facilitates the characterization of methylation patterns of specific RNAs. To detect certain small RNAs that harbor RNA modifications as well as incompatible 5′ and 3′-ends, cloning techniques may need to incorporate multiple steps to permit efficient ligation and reverse transcription. Indeed, a new RNA-sequencing strategy combines different enzymatic treatments, namely AlkB and T4 PNK, to remove small RNA methylation and cyclic phosphates prior to library preparation (Shi et al., 2021), respectively. Coupled with improved bioinformatic pipelines, this strategy, termed panoramic RNA display by overcoming RNA modification aborted sequencing (PANDORA-seq), has successfully identified abundant, modified small RNAs (mainly tRFs and rRFs) that were previously not detected (Shi et al., 2021). As more protocols are developed that are capable of sequencing previously unappreciated pools of small RNAs, such as tRFs and rRFs, the small RNA landscape of many cells will need to be revisited.

While the presence of RNA modifications generally presents a problem in generating a full-length cDNA that accurately represents the entirety of the RNA molecule, it can be used as a method for identifying the position of modified RNA bases by analyzing the endpoints of truncated cDNA products. Indeed, truncated cDNA can be quantitated to determine when RT synthesis stops as a result of a presumed RNA modification (Kuksa et al., 2017). However, due to their short length, difficulty arises in unambiguously determining a truncated product; therefore, such analyses have yet to be adopted to uncover RNA modifications present in small RNAs.

Alternative Reverse Transcriptase Enzymes

Recent work utilizing OTTR has revealed that the enrichment of 5′ versus 3′ tRFs in budding yeast and mammalian sperm is less pronounced than previously reported (Fig. 2) (Gustafsson et al., 2022). Using OTTR on mouse sperm small RNAs revealed that not only is there near equal distribution of both 5′ and 3′ fragments for all tRFs, but that fragments of most tRNAs are equally present, in contrast to previous results suggesting the selective accumulation of specific tRFs. While part of the success of OTTR in recovering tRFs is due to bypassing ligation, it also appears that the BoMoC RT can more efficiently synthesize cDNA from modified RNAs. Enzymatically resolving 3′ cyclic phosphates and adding 5′ monophosphates to mouse sperm small RNAs only partially recovers the small RNAs revealed by OTTR, indicating that other aspects of this technique are required for recovering the observed complement of tRFs during small RNA cloning (Gustafsson et al., 2022). RT enzymes can display varying levels of sensitivity to cDNA elongation or nucleotide incorporation when presented with a modified RNA substrate. This aspect of RT function can be taken advantage of when cloning small RNAs with modified nucleotides (Upton et al., 2021). Interestingly, tRF sequencing reads that do not match the annotated genome are located precisely at modified tRNA bases, allowing for the determination of RNA modification status through RNA-sequencing (Gustafsson et al., 2022).

NOVEL METHODS

Since the small RNA revolution at the turn of the millennium, methods used to discover and detect these RNAs have evolved from the

detection of single species using northern blotting to global analyses using high-throughput sequencing technology. As discussed, the latter has evolved from traditional ligation and adapter-dependent library preparation techniques that amplify cDNA sequences from cloned small RNAs to include a variety of RNA treatments (highlighted throughout this overview) that allow the capture of previously uncloneable RNAs and further bypass the ligation steps altogether. These methods, however, all still require the addition of adapter sequences to the 5′ and 3′ ends, which can inherently introduce biases to the cloned and sequenced small RNAs. Recently, methods using Nanopore sequencing have been developed that allow direct sequencing of native full-length mRNAs (Garalde et al., 2018). To date, this approach has been successfully used to identify and quantitate mRNAs in a variety of biological contexts. Small RNAs, however, with their limited sequence diversity and short length, create unique problems for Nanopore sequencing and, further, these methods typically require large amounts of input material (>500 ng purified mRNA), which is difficult to obtain for these species of RNAs. These issues have now begun to be resolved with the recently described direct-sequencing of miRNAs using the Nanopore-induced phase-shift sequencing (NIPSS) approach (Zhang et al., 2020). As Nanopore technology has rapidly evolved to include new flavors of direct single molecule sequencing, including direct RNA-sequencing (Garalde et al., 2018) and the detection of nucleotide modifications (Garalde et al., 2018; Liu et al., 2019), it seems promising that direct sequencing of small RNAs will be possible in the near future. Additionally, the direct analysis of RNA sequence by mass spectrometry is a developing field of study. Small RNAs are a good initial candidate for these techniques, as their limited complexity and length simplify these analyses compared to mRNAs and other RNAs (Wein et al., 2020; Zhang et al., 2019).

Another future direction of small RNA biology and sequencing is single-cell analysis. Single-cell mRNA-seq has been used in thousands of papers over the past decade, on a vast array of different tissues and species (Kolodziejczyk, Kim, Svensson, Marioni, & Teichmann, 2015; Saliba, Westermann, Gorski, & Vogel, 2014). However, single-cell small RNA-seq has lagged significantly in comparison, due to the lack of a "molecular handle" (e.g., polyadenylation in mRNAs) that can be used to generate and amplify cDNA. Over the past several years, however, a handful of papers have reported the development of techniques to sequence small RNAs from single cells (Faridani et al., 2016; Hagemann-Jensen, Abdullayev, Sandberg, & Faridani, 2018; Hücker et al., 2021). These techniques all employ ligation-based approaches and typically involve blocking rRNAs from being cloned, followed by size selection of miRNA-sized insert cDNA clones after PCR amplification. While these studies are promising, they have currently only been performed on small numbers of cells (hundreds or less in a given experiment), while single-cell mRNA-seq is routinely performed on thousands of cells. Consequently, there is still much to be learned about the biology of small RNAs in single cells. Interestingly, two studies have either split single-cell lysates or have purified polyadenylated mRNA and small RNAs separately, to perform small RNA and mRNA-seq from the same cell (Wang et al., 2019; Xiao et al., 2018). As predicted, the data from these studies provide some evidence that the expression of a miRNA is anti-correlated with its mRNA targets.

GENERAL CONSIDERATIONS FOR SMALL RNA-SEQ EXPERIMENTS

The first consideration for any small RNA-seq experiment should be the amount of material that can be obtained (Fig. 2). If plentiful amounts of RNA can be purified (total RNA >100 ng), nearly any of the protocols reviewed above can be used. However, in case of studying samples from which it may be difficult to obtain adequate quantities, less RNA can certainly be used, with some considerations. In these cases, and if using a ligation-based approach, it is important to titrate the amount of adapter used, as to not overwhelm the library with adapter-to-adapter products that will dominate the library after PCR amplification. Alternatively, non-ligation-based methods such as direct, artificial polyadenylation and cDNA generation can also be used on low-input samples. Additionally, for very low input samples, rather than size selecting small RNAs from total RNA and then creating cDNA libraries, it can be advantageous to use total RNA input and size selection for the adapter-ligated small RNA. A final consideration for input concentration is the amount of PCR cycles that are required to amplify a library for sequencing. It is best to perform a series of different PCR amplification cycles and optimize for the minimal number of cycles

that will produce the desired size band (representing the size of an adapter-ligated small RNA) in a non-denaturing PAGE gel without overamplifying (generating bulge products) the small RNA library.

The next important consideration is the type of small RNA of interest. For miRNAs that have 5′-monophosphates and 3′ hydroxyl groups, most standard cloning procedures will be effective. However, if a user is interested in small RNAs with terminal modifications, RNAs lacking 5′-monophosphates and 3′ hydroxyl termini, or small RNAs with internal RNA modifications, then enzymatic or other treatments to the RNA highlighted in this overview should be performed (Fig. 2). For example, to ensure that small RNAs of interest are efficiently cloned, RNAs cleaved by endonucleases such as tRFs should have their 5′ ends phosphorylated and 3′-cyclic phosphates resolved through different PNK treatments with and without ATP, respectively. For 2′ -*O*-methylated piRNAs, these should be enriched via oxidation of non-methylated small RNAs, while RNAs that have internal modifications such as tRFs and rRNA-fragments should be either treated with enzymes that remove these groups or users should employ alternative RT enzymes during cDNA synthesis (Fig. 2).

The goal of any small RNA-seq experiment should also be assessed prior to the initiation of the experiment with regards to the type of quantitative information that one aims to produce. If relative information about small RNA levels between control and experimental conditions (mutant strain, environmental change, etc.) is all that is necessary, then cloning and sequencing alone can be performed. However, if absolute quantitation of the concentration of small RNA molecules present in a sample is essential, protocols using UMIs, spike-in standards, or both should be used. Absolute quantification cannot be done in retrospect; therefore, if this information is required, it must be planned for in advance.

While this overview focuses on the technical/experimental aspects of small RNA-seq, the computational analysis of the data should also be considered when planning an experiment. While details of computational analysis are reviewed elsewhere (Buschmann et al., 2016; Ilnytskyy & Bilichak, 2017), we briefly highlight here some important aspects. Small RNA-seq data requires analyses outside of typical mRNA-seq pipelines. First, because small RNAs map outside of annotated transcriptomic sequences (mRNAs), the data must also be mapped to transcript files that are annotated for the known miRNAs, piRNAs, tRNAs, and rRNAs of the species of interest. While this permits the quantitation of known small RNAs, it does not; however, allow for the discovery of novel small RNA species. To identify novel small RNA species, sequencing data must be mapped to the genome without biasing towards already annotated transcriptomic data, allowing unique reads (outside of known transcripts) to be identified by scanning genome browsers or with novel computational programs to find these RNAs (Friedlander, Mackowiak, Li, & Chen, 2012; Handzlik, Tastsoglou, Vlachos, & Hatzigeorgiou, 2020). For organisms that produce small RNAs by RdRPs that are antisense to target RNAs, it is also important to map the sequencing data not only to the transcriptome/genome but also to the sequences representing the antisense of the transcriptome. These mapping strategies are commonly used for small RNA-seq analysis in *C. elegans* and plants (Gu et al., 2009).

After mapping of the data, and to compare samples against one another (and between experiments), the small RNA-seq counts must be normalized. Unlike mRNA-seq, the entire length of the small RNA is sequenced in a single read. Therefore, it is not necessary to normalize the data to the length of the RNA as is required for mRNAs. Thus, small RNAs can be normalized to all genome-matching reads to generate reads per million (rpm) for all the quantitated small RNAs. However, if an experimental treatment or mutant leads to large scale changes in small RNAs, normalization to all genome matching reads can skew that data, leading to an apparent increase in small RNAs that are not affected. For example, in the mouse (and other animals) testis, mutations in PIWI-AGOs lead to the loss of piRNAs, the predominant small RNA in that tissue. Therefore, the small RNAs that remain in the mutant testis, predominantly miRNAs, appear enriched compared to the control due to the loss of this major source of small RNA. Instead, reads can be normalized to individual classes of RNAs; for example, miRNAs would be normalized to the total number of miRNA reads. Another way to avoid pitfalls associated with normalizing small RNA-seq data to total reads is to normalize the data to standards within the data that are not changing between conditions. These statistical models include scaling methods such as Trimmed Mean Method (TMM) normalization and Lowess

normalization, both of which generally assume that the range of data is the same across samples and that the changes in small RNA expression are proportional to the signal intensity. Conversely, non-scaling techniques such as quantile normalization assume that the distribution of signal intensity does not change across samples. Broadly, scaling and non-scaling methods use different statistical approaches to evaluate and interpret the stochastic noise present across samples in order to identify bona-fide changes of RNA abundance between conditions. Garmire & Subramanian provide a thorough evaluation of a variety of statistical methods on microRNA datasets, ultimately concluding that Lowess normalization and quantile normalization outperform alternative methods (Garmire & Subramaniam, 2012, 2013). Finally, RNA spike-ins of known sequence between samples of different experimental conditions can be used to normalize small RNA-sequencing data.

CLOSING REMARKS

Small RNAs participate in nearly all aspects of biology as pivotal regulators of gene expression in eukaryotes. Since their initial discovery 30 years ago, the biology and technology related to small RNAs have been rapidly unveiled and developed, respectively. Undoubtedly, the ability to sequence cellular (and extracellular) small RNAs efficiently and reliably is a powerful tool for understanding and discovering small RNA biology. However, because of the unique molecular characteristics exhibited by small RNAs, specific procedures need to be considered and implemented to effectively analyze different classes of small RNAs. As new varieties and functions of small RNAs are discovered, cloning techniques must continue to adapt to accurately capture and assess the small RNAs present in a tissue, cell, or biological sample.

Acknowledgments

The authors thank Grace Lee and Nick Galambos for critical reading of the manuscript. This work was supported by a Pew Biomedical Scholars Award to C.C.C.

Author Contributions

Olivia J. Crocker: conceptualization; investigation; visualization; writing — original draft preparation; writing — review & editing. **Natalie A. Trigg:** conceptualization; investigation; visualization; writing — original draft preparation; writing — review & editing. **Colin C. Conine:** conceptualization; investigation; visualization; writing — original draft preparation; writing — review & editing; funding acquisition; project administration.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

LITERATURE CITED

Almeida, M. V., De Jesus Domingues, A. M., Lukas, H., Mendez-Lago, M., & Ketting, R. F. (2019). RppH can faithfully replace TAP to allow cloning of 5′-triphosphate carrying small RNAs. *MethodsX*, *6*, 265–272. doi: 10.1016/j.mex.2019.01.009

Ambros, V., Lee, R. C., Lavanway, A., Williams, P. T., & Jewell, D. (2003). MicroRNAs and other tiny endogenous RNAs in C. elegans. *Current Biology: CB*, *13*, 807–818. doi: 10.1016/s0960-9822(03)00287-2

Anzelon, T. A., Chowdhury, S., Hughes, S. M., Xiao, Y., Lander, G. C., & Macrae, I. J. (2021). Structural basis for piRNA targeting. *Nature*, *597*, 285–289. doi: 10.1038/s41586-021-03856-x

Aravin, A. A., Naumova, N. M., Tulin, A. V., Vagin, V. V., Rozovsky, Y. M., & Gvozdev, V. A. (2001). Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the D. melanogaster germline. *Current Biology: CB*, *11*, 1017–1027.

Bartel, D. P. (2009). MicroRNAs: Target recognition and regulatory functions. *Cell*, *136*, 215–233. doi: 10.1016/j.cell.2009.01.002

Batista, P. J., Ruby, J. G., Claycomb, J. M., Chiang, R., Fahlgren, N., Kasschau, K. D., … Mello, C. C. (2008). PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in C. elegans. *Molecular Cell*, *31*, 67–78. doi: 10.1016/j.molcel.2008.06.002

Bernstein, E., Caudy, A. A., Hammond, S. M., & Hannon, G. J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature*, *409*, 363–366. doi: 10.1038/35053110

Beyret, E., Liu, N., & Lin, H. (2012). piRNA biogenesis during adult spermatogenesis in mice is independent of the ping-pong mechanism. *Cell Research*, *22*, 1429–1439. doi: 10.1038/cr.2012.120

Billi, A. C., Fischer, S. E., & Kim, J. K. (2014). Endogenous RNAi pathways in C. elegans. *WormBook*, 1–49. doi: 10.1895/wormbook.1.170.1

Birkedal, U., Christensen-Dalsgaard, M., Krogh, N., Sabarinathan, R., Gorodkin, J., & Nielsen, H. (2015). Profiling of ribose methylations in RNA by high-throughput sequencing. *Angewandte Chemie - International Edition in English*, *54*, n/a–n/a, doi: 10.1002/anie.201408362

Boskovic, A., Bing, X. Y., Kaymak, E., & Rando, O. J. (2020). Control of noncoding RNA production and histone levels by a 5′ tRNA fragment. *Genes & Development*, *34*, 118–131. doi: 10.1101/gad.332783.119

Brennecke, J., Aravin, A. A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., & Hannon, G. J. (2007). Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell*, *128*, 1089–1103.

Brodersen, P., & Voinnet, O. (2009). Revisiting the principles of microRNA target recognition and mode of action. *Nature Reviews. Molecular Cell Biology*, *10*, 141–148. doi: 10.1038/nrm2619

Buschmann, D., Haberberger, A., Kirchner, B., Spornraft, M., Riedmaier, I., Schelling, G., & Pfaffl, M. W. (2016). Toward reliable biomarker signatures in the age of liquid biopsies - how to standardize the small RNA-Seq workflow. *Nucleic Acids Research*, *44*, 5995–6018. doi: 10.1093/nar/gkw545

Carrier, M. C., Lalaouna, D., & Massé, E. (2018). Broadening the Definition of bacterial small rnas: Characteristics and mechanisms of action. *Annual Review of Microbiology*, *72*, 141–161. doi: 10.1146/annurev-micro-090817-062607

Cecere, G., & Grishok, A. (2014). A nuclear perspective on RNAi pathways in metazoans. *Biochimica et Biophysica Acta*, *1839*, 223–233. doi: 10.1016/j.bbagrm.2013.11.009

Cech, T R., & Steitz, J A. (2014). The noncoding RNA revolution-trashing old rules to forge new ones. *Cell*, *157*, 77–94. doi: 10.1016/j.cell.2014.03.008

Chen, Q., Yan, M., Cao, Z., Li, X., Zhang, Y., Shi, J., … Zhou, Q. (2016). Sperm tsRNAs contribute to intergenerational inheritance of an acquired metabolic disorder. *Science (New York)*, *351*, 397–400. doi: 10.1126/science.aad7977

Cherlin, T., Magee, R., Jing, Y., Pliatsika, V., Loher, P., & Rigoutsos, I. (2020). Ribosomal RNA fragmentation into short RNAs (rRFs) is modulated in a sex- and population of origin-specific manner. *BMC Biology*, *18*, 38, doi: 10.1186/s12915-020-0763-0

Chu, C., Yu, L., Wu, B., Ma, L., Gou, L. - T., He, M., … Zhang, Y. (2017). A sequence of 28S rRNA-derived small RNAs is enriched in mature sperm and various somatic tissues and possibly associates with inflammation. *Journal Molecular Cell Biology*, *9*, 256–259. doi: 10.1093/jmcb/mjx016

Cozen, A. E., Quartley, E., Holmes, A. D., Hrabeta-Robinson, E., Phizicky, E. M., & Lowe, T. M. (2015). ARM-seq: AlkB-facilitated RNA methylation sequencing reveals a complex landscape of modified tRNA fragments. *Nature Methods*, *12*, 879–884. doi: 10.1038/nmeth.3508

Czech, B., Malone, C. D., Zhou, R., Stark, A., Schlingeheyde, C., Dus, M., … Brennecke, J. (2008). An endogenous small interfering RNA pathway in Drosophila. *Nature*, *453*, 798–802.

Dard-Dascot, C., Naquin, D., D'aubenton-Carafa, Y., Alix, K., Thermes, C., & Van Dijk, E. (2018). Systematic comparison of small RNA library preparation protocols for next-generation sequencing. *BMC Genomics*, *19*, 118, doi: 10.1186/s12864-018-4491-6

Das, P. P., Bagijn, M. P., Goldstein, L. D., Woolford, J. R., Lehrbach, N. J., Sapetschnig, A., … Miska, E. A. (2008). Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the Caenorhabditis elegans germline. *Molecular Cell*, *31*, 79–90. doi: 10.1016/j.molcel.2008.06.003

Delaney, J. C., & Essigmann, J. M. (2004). Mutagenesis, genotoxicity, and repair of 1-methyladenine, 3-alkylcytosines, 1-methylguanine, and 3-methylthymine in alkB Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 14051–14056. doi: 10.1073/pnas.0403489101

Emara, M. M., Ivanov, P., Hickman, T., Dawra, N., Tisdale, S., Kedersha, N., … Anderson, P. (2010). Angiogenin-induced tRNA-derived stress-induced RNAs promote stress-induced stress granule assembly. *The Journal of Biological Chemistry*, *285*, 10959–10968. doi: 10.1074/jbc.M109.077560

Fahlgren, N., Sullivan, C. M., Kasschau, K. D., Chapman, E. J., Cumbie, J. S., Montgomery, T. A., … Carrington, J. C. (2009). Computational and analytical framework for small RNA profiling by high-throughput sequencing. *RNA*, *15*, 992–1002. doi: 10.1261/rna.1473809

Faridani, O. R., Abdullayev, I., Hagemann-Jensen, M., Schell, J. P., Lanner, F., & Sandberg, R. (2016). Single-cell sequencing of the small-RNA transcriptome. *Nature Biotechnology*, *34*, 1264–1266. doi: 10.1038/nbt.3701

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., & Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature*, *391*, 806–811. doi: 10.1038/35888

Fricker, R., Brogli, R., Luidalepp, H., Wyss, L., Fasnacht, M., Joss, O., … Polacek, N. (2019). A tRNA half modulates translation as stress response in Trypanosoma brucei. *Nature Communications*, *10*, 118, doi: 10.1038/s41467-018-07949-6

Friedlander, M. R., Mackowiak, S. D., Li, N., & Chen, W. (2012). & Rajewsky, N. miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Research*, *40*, 37–52. doi: 10.1093/nar/gkr688

Fu, H., Feng, J., Liu, Q., Sun, F., Tie, Y., Zhu, J., … Zheng, X. (2009). Stress induces tRNA cleavage by angiogenin in mammalian cells. *FEBS Letters*, *583*, 437–442. doi: 10.1016/j.febslet.2008.12.043

Fu, Y., Wu, P. - H., Beane, T., Zamore, P. D., & Weng, Z. (2018). Elimination of PCR duplicates in RNA-seq and small RNA-seq using unique

molecular identifiers. *BMC Genomics*, *19*, 531, doi: 10.1186/s12864-018-4933-1

Gainetdinov, I., Colpan, C., Cecchini, K., Arif, A., Jouravleva, K., Albosta, P., … Zamore, P. D. (2021). Terminal modification, sequence, length, and PIWI-protein identity determine piRNA stability. *Molecular Cell*, *81*, 4826–4842.e8. doi: 10.1016/j.molcel.2021.09.012

Garalde, D. R., Snell, E. A., Jachimowicz, D., Sipos, B., Lloyd, J. H., Bruce, M., … Turner, D. J. (2018). Highly parallel direct RNA sequencing on an array of nanopores. *Nature Methods*, *15*, 201–206. doi: 10.1038/nmeth.4577

Garmire, L. X., & Subramaniam, S. (2012). Evaluation of normalization methods in mammalian microRNA-Seq data. *RNA*, *18*, 1279–1288. doi: 10.1261/rna.030916.111

Garmire, L. X., & Subramaniam, S. (2013). The poor performance of TMM on microRNA-Seq. *RNA*, *19*, 735–736. doi: 10.1261/rna.039271.113

Gebetsberger, J., Wyss, L., Mleczko, A. M., Reuther, J., & Polacek, N. (2017). A tRNA-derived fragment competes with mRNA for ribosome binding and regulates translation during stress. *RNA Biology*, *14*, 1364–1373. doi: 10.1080/15476286.2016.1257470

Gelsinger, D., & Diruggiero, J. (2018). The Non-Coding Regulatory RNA Revolution in Archaea. *Genes (Basel)*, *141*, doi: 10.3390/genes9030141

Ghildiyal, M., & Zamore, P. D. (2009). Small silencing RNAs: An expanding universe. *Nature Reviews. Genetics*, *10*, 94–108.

Giraldez, M. D., Spengler, R. M., Etheridge, A., Godoy, P. M., Barczak, A. J., Srinivasan, S., … Tewari, M. (2018). Comprehensive multicenter assessment of small RNA-seq methods for quantitative miRNA profiling. *Nature Biotechnology*, *36*, 746–757. doi: 10.1038/nbt.4183

Girard, A., Sachidanandam, R., Hannon, G. J., & Carmell, M. A. (2006). A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature*, *442*, 199–202.

Gonskikh, Y., Gerstl, M., Kos, M., Borth, N., Schosserer, M., Grillari, J., & Polacek, N. (2020). Modulation of mammalian translation by a ribosome-associated tRNA half. *RNA Biology*, *17*, 1125–1136. doi: 10.1080/15476286.2020.1744296

Gu, W., Shirayama, M., Conte, D., Vasale, J., Batista, P. J., Claycomb, J. M., … Mello, C. C. (2009). Distinct argonaute-mediated 22G-RNA pathways direct genome surveillance in the C. elegans germline. *Molecular cell*, *36*, 231–244. doi: 10.1016/j.molcel.2009.09.020

Gunawardane, L. S., Saito, K., Nishida, K. M., Miyoshi, K., Kawamura, Y., Nagami, T., … Siomi, M. C. (2007). A slicer-mediated mechanism for repeat-associated siRNA 5′ end formation in Drosophila. *Science (New York)*, *315*, 1587–1590. doi: 10.1126/science.1140494

Guo, Y., Bosompem, A., Zhong, X., Clark, T., Shyr, Y., & Kim, A. S. (2014). A comparison of microRNA sequencing reproducibility and noise reduction using mirVana and TRIzol isolation methods. *International Journal of Computational Biology and Drug Design*, *7*, 102–112. doi: 10.1504/IJCBDD.2014.061642

Gustafsson, H. T., Galan, C., Yu, T., Upton, H. E., Ferguson, L., Kaymak, E., … Rando, O. J. (2022). Deep sequencing of yeast and mouse tRNAs and tRNA fragments using OTTR. *bioRxiv*, 2022.2002.2004.479139. doi: 10.1101/2022.02.04.479139

Ha, M., & Kim, V. N. (2014). Regulation of microRNA biogenesis. *Nature Review Molecular Cell Biology*, *15*, 509–524. doi: 10.1038/nrm3838

Hagemann-Jensen, M., Abdullayev, I., Sandberg, R., & Faridani, O. R. (2018). Small-seq for single-cell small-RNA sequencing. *Nature Protocols*, *13*, 2407–2424. doi: 10.1038/s41596-018-0049-y

Hamilton, A. J., & Baulcombe, D. C. (1999). A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science (New York)*, *286*, 950–952. doi: 10.1126/science.286.5441.950

Handzlik, J. E., Tastsoglou, S., Vlachos, I. S., & Hatzigeorgiou, A. G. (2020). Manatee: Detection and quantification of small non-coding RNAs from next-generation sequencing data. *Science Report*, *10*, 705, doi: 10.1038/s41598-020-57495-9

Honda, S., Morichika, K., & Kirino, Y. (2016). Selective amplification and sequencing of cyclic phosphate-containing RNAs by the cP-RNA-seq method. *Nature Protocols*, *11*, 476–489. doi: 10.1038/nprot.2016.025

Horwich, M. D., Li, C., Matranga, C., Vagin, V., Farley, G., Wang, P., & Zamore, P. D. (2007). The Drosophila RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. *Current Biology: CB*, *17*, 1265–1272. doi: 10.1016/j.cub.2007.06.030

Hrabeta-Robinson, E., Marcus, E., Cozen, A. E., Phizicky, E. M., & Lowe, T. M. (2017). High-Throughput Small RNA Sequencing Enhanced by AlkB-Facilitated RNA de-Methylation (ARM-Seq). *Methods Molecular Biology*, *1562*, 231–243. doi: 10.1007/978-1-4939-6807-7_15

Hua, M., Liu, W., Chen, Y., Zhang, F., Xu, B., Liu, S., … Wu, L. (2019). Identification of small noncoding RNAs as sperm quality biomarkers for in vitro fertilization. *Cell Discovery*, *5*, 20, doi: 10.1038/s41421-019-0087-9

Huang, S., Yoshitake, K., & Asakawa, S. (2021). A review of discovery profiling of PIWI-interacting RNAs and their diverse functions in metazoans. *International Journal of Molecullar Science*, *22*, 11166. doi: 10.3390/ijms222011166

Hücker, S. M., Fehlmann, T., Werno, C., Weidele, K., Lüke, F., Schlenska-Lange, A., … Kirsch, S. (2021). Single-cell microRNA sequencing

method comparison and application to cell lines and circulating lung tumor cells. *Nature Communications*, *12*, 4316, doi: 10.1038/s41467-021-24611-w

Ilnytskyy, S., & Bilichak, A. (2017). Bioinformatics Analysis of Small RNA Transcriptomes: The Detailed Workflow. *Methods Molecular Biology*, *1456*, 197–224. doi: 10.1007/978-1-4899-7708-3_16

Iwasaki, Y. W., Siomi, M. C., & Siomi, H. (2015). PIWI-Interacting RNA: Its Biogenesis and Functions. *Annual Review of Biochemistry*, *84*, 405–433. doi: 10.1146/annurev-biochem-060614-034258

Jackman, J. E., & Alfonzo, J. D. (2013). Transfer RNA modifications: Nature's combinatorial chemistry playground. *Wiley Interdisciplinary Reviews. RNA*, *4*, 35–48. doi: 10.1002/wrna.1144

Ji, L., & Chen, X. (2012). Regulation of small RNA stability: Methylation and beyond. *Cell Research*, *22*, 624–636. doi: 10.1038/cr.2012.36

Jonas, S., & Izaurralde, E. (2015). Towards a molecular understanding of microRNA-mediated gene silencing. *Nature Reviews. Genetics*, *16*, 421–433. doi: 10.1038/nrg3965

Keam, S. P., Sobala, A., Ten Have, S., & Hutvagner, G. (2017). tRNA-derived RNA fragments associate with human multisynthetase complex (MSC) and modulate ribosomal protein translation. *Journal of Proteome Research*, *16*, 413–420. doi: 10.1021/acs.jproteome.6b00267

Ketting, R. F., & Cochella, L. (2021). Concepts and functions of small RNA pathways in C. elegans. *Current Topics in Developmental Biology*, *144*, 45–89. doi: 10.1016/bs.ctdb.2020.08.002

Kim, V. N., Han, J., & Siomi, M. C. (2009). Biogenesis of small RNAs in animals. *Nature Reviews. Molecular Cell Biology*, *10*, 126–139. doi: 10.1038/nrm2632

Kingston, E. R., & Bartel, D. P. (2021). Ago2 protects Drosophila siRNAs and microRNAs from target-directed degradation, even in the absence of 2′-O-methylation. *RNA*, *27*, 710–724. doi: 10.1261/rna.078746.121

Kirino, Y., & Mourelatos, Z. (2007). Mouse Piwi-interacting RNAs are 2′-O-methylated at their 3′ termini. *Nature Structural & Molecular Biology*, *14*, 347–348. doi: 10.1038/nsmb1218

Kivioja, T., Vähärautio, A., Karlsson, K., Bonke, M., Enge, M., Linnarsson, S., & Taipale, J. (2011). Counting absolute numbers of molecules using unique molecular identifiers. *Nature Methods*, *9*, 72–74. doi: 10.1038/nmeth.1778

Kolodziejczyk, A. A., Kim, J. K., Svensson, V., Marioni, J. C., & Teichmann, S. A. (2015). The technology and biology of single-cell RNA sequencing. *Molecular Cell*, *58*, 610–620. doi: 10.1016/j.molcel.2015.04.005

Krishna, S., Yim, D. G., Lakshmanan, V., Tirumalai, V., Koh, J. L., Park, J. E., … Dasgupta, R. (2019). Dynamic expression of tRNA-derived small RNAs define cellular states. *EMBO Reports*, *20*, e47789, doi: 10.15252/embr.201947789

Kuksa, P. P., Leung, Y. Y., Vandivier, L. E., Anderson, Z., Gregory, B. D., & Wang, L. S. (2017). In silico identification of RNA modifications from high-throughput sequencing data using HAMR. *Methods Molecular Biology*, *1562*, 211–229. doi: 10.1007/978-1-4939-6807-7_14

Kumar, P., Kuscu, C., & Dutta, A. (2016). Biogenesis and function of transfer RNA-related fragments (tRFs). *Trends in Biochemical Sciences*, *41*, 679–689. doi: 10.1016/j.tibs.2016.05.004

Kurth, H. M., & Mochizuki, K. (2009). 2′-O-methylation stabilizes Piwi-associated small RNAs and ensures DNA elimination in Tetrahymena. *RNA*, *15*, 675–685. doi: 10.1261/rna.1455509

Lagos-Quintana, M., Rauhut, R., Lendeckel, W., & Tuschl, T. (2001). Identification of novel genes coding for small expressed RNAs. *Science (New York)*, *294*, 853–858. doi: 10.1126/science.1064921

Lagos-Quintana, M., Rauhut, R., Meyer, J., Borkhardt, A., & Tuschl, T. (2003). New microRNAs from mouse and human. *RNA*, *9*, 175–179. doi: 10.1261/rna.2146903

Lau, N. C., Lim, L. P., Weinstein, E. G., & Bartel, D. P. (2001). An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. *Science (New York)*, *294*, 858–862. doi: 10.1126/science.1065062

Lee, R. C., & Ambros, V. (2001). An extensive class of small RNAs in Caenorhabditis elegans. *Science (New York)*, *294*, 862–864. doi: 10.1126/science.1065329

Lee, R. C., Feinbaum, R. L., & Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell*, *75*, 843–854. doi: 10.1016/0092-8674(93)90529-y

Lee, Y. S., Shibata, Y., Malhotra, A., & Dutta, A. (2009). A novel class of small RNAs: TRNA-derived RNA fragments (tRFs). *Genes & Development*, *23*, 2639–2649.

Lestrade, L., & Weber, M. J. (2006). snoRNA-LBME-db, a comprehensive database of human H/ACA and C/D box snoRNAs. *Nucleic Acids Research*, *34*, D158–D162. doi: 10.1093/nar/gkj002

Licht, K., Medenbach, J., Lührmann, R., Kambach, C., & Bindereif, A. (2008). 3′-cyclic phosphorylation of U6 snRNA leads to recruitment of recycling factor p110 through LSm proteins. *RNA*, *14*, 1532–1538. doi: 10.1261/rna.1129608

Lim, L. P., Lau, N. C., Weinstein, E. G., Abdelhakim, A., Yekta, S., Rhoades, M. W., … Bartel, D. P. (2003). The microRNAs of Caenorhabditis elegans. *Genes & Development*, *17*, 991–1008. doi: 10.1101/gad.1074403

Liu, H., Begik, O., Lucas, M. C., Ramirez, J. M., Mason, C. E., Wiener, D., … Novoa, E. M. (2019). Accurate detection of m(6)A RNA modifications in native RNA sequences. *Nature Communications*,

10, 4079, doi: 10.1038/s41467-019-11713-9

Llave, C., Kasschau, K. D., Rector, M. A., & Carrington, J. C. (2002). Endogenous and silencing-associated small RNAs in plants. *Plant Cell*, *14*, 1605–1619. doi: 10.1105/tpc.003210

Locati, M. D., Terpstra, I., De Leeuw, W. C., Kuzak, M., Rauwerda, H., Ensink, W. A., … Dekker, R. J. (2015). Improving small RNA-seq by using a synthetic spike-in set for size-range quality control together with a set for data normalization. *Nucleic Acids Research*, *43*, e89, doi: 10.1093/nar/gkv303

Lopez Sanchez, M. I. G., Cipullo, M., Gopalakrishna, S., Khawaja, A., & Rorbach, J. (2020). Methylation of ribosomal RNA: A mitochondrial perspective. *Frontiers in Genetics*, *11*, 761, doi: 10.3389/fgene.2020.00761

Lu, C., Tej, S. S., Luo, S., Haudenschild, C. D., Meyers, B. C., & Green, P. J. (2005). Elucidation of the small RNA component of the transcriptome. *Science (New York)*, *309*, 1567–1569. doi: 10.1126/science.1114112

Lund, E., & Dahlberg, J. E. (1992). Cyclic 2′,3′-phosphates and nontemplated nucleotides at the 3′ end of spliceosomal U6 small nuclear RNA's. *Science (New York)*, *255*, 327–330. doi: 10.1126/science.1549778

Lutzmayer, S., Enugutti, B., & Nodine, M. D. (2017). Novel small RNA spike-in oligonucleotides enable absolute normalization of small RNA-Seq data. *Science Report*, *7*, 5913, doi: 10.1038/s41598-017-06174-3

Lyons, S. M., Achorn, C., Kedersha, N. L., Anderson, P. J., & Ivanov, P. (2016). YB-1 regulates tiRNA-induced Stress Granule formation but not translational repression. *Nucleic Acids Research*, *44*, 6949–6960. doi: 10.1093/nar/gkw418

Magee, R., & Rigoutsos, I. (2020). On the expanding roles of tRNA fragments in modulating cell behavior. *Nucleic Acids Research*, *48*, 9433–9448. doi: 10.1093/nar/gkaa657

Mello, C. C., & Conte, D. (2004). Revealing the world of RNA interference. *Nature*, *431*, 338–342. doi: 10.1038/nature02872

Mleczko, A. M., Celichowski, P., & Bąkowska-Żywicka, K. (2018). Transfer RNA-derived fragments target and regulate ribosome-associated aminoacyl-transfer RNA synthetases. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, *1861*, 647–656. doi: 10.1016/j.bbagrm.2018.06.001

Molla-Herman, A., Angelova, M. T., Ginestet, M., Carré, C., Antoniewski, C., & Huynh, J. - R. (2020). tRNA fragments populations analysis in mutants affecting tRNAs processing and tRNA methylation. *Frontiers in Genetics*, *11*, 518949, doi: 10.3389/fgene.2020.518949

Morris, K. V., Chan, S. W. - L., Jacobsen, S. E., & Looney, D. J. (2004). Small interfering RNA-induced transcriptional gene silencing in human cells. *Science (New York)*, *305*, 1289–1292. doi: 10.1126/science.1101372

Munafó, D. B., & Robb, G. B. (2010). Optimization of enzymatic reaction conditions for generating representative pools of cDNA from small RNA. *RNA*, *16*, 2537–2552. doi: 10.1261/rna.2242610

Oberbauer, V., & Schaefer, M. (2018). tRNA-derived small RNAs: Biogenesis, modification, function and potential impact on human disease development. *Genes (Basel)*, *9*, 607, doi: 10.3390/genes9120607

Ohara, T., Sakaguchi, Y., Suzuki, T., Ueda, H., Miyauchi, K., & Suzuki, T. (2007). The 3′ termini of mouse piwi-interacting RNAs are 2′-O-methylated. *Nature Structural & Molecular Biology*, *14*, 349–350. doi: 10.1038/nsmb1220

Okamura, K., & Lai, E. C. (2008). Endogenous small interfering RNAs in animals. *Nature reviews. Molecular Cell Biology*, *9*, 673–678. doi: 10.1038/nrm2479

Ozata, D. M., Gainetdinov, I., Zoch, A., O'carroll, D., & Zamore, P. D. (2019). PIWI-interacting RNAs: Small RNAs with big functions. *Nature Reviews. Genetics*, *20*, 89–108. doi: 10.1038/s41576-018-0073-3

Pak, J., & Fire, A. (2007). Distinct populations of primary and secondary effectors during RNAi in C. elegans. *Science (New York)*, *315*, 241–244. doi: 10.1126/science.1132839

Parhad, S. S., & Theurkauf, W. E. (2019). Rapid evolution and conserved function of the piRNA pathway. *Open Biology*, *9*, 180181, doi: 10.1098/rsob.180181

Pasquinelli, A. E., Reinhart, B. J., Slack, F., Martindale, M. Q., Kuroda, M. I., Maller, B., … Ruvkun, G. (2000). Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature*, *408*, 86–89. doi: 10.1038/35040556

Pastore, B., Hertz, H. L., Price, I. F., & Tang, W. (2021). pre-piRNA trimming and 2′-O-methylation protect piRNAs from 3′ tailing and degradation in C. elegans. *Cell Reports*, *36*, 109640, doi: 10.1016/j.celrep.2021.109640

Peng, H., Shi, J., Zhang, Y., Zhang, H., Liao, S., Li, W., … Duan, E. (2012). A novel class of tRNA-derived small RNAs extremely enriched in mature mouse sperm. *Cell Research*, *22*, 1609–1612.

Pezic, D., Manakov, S. A., Sachidanandam, R., & Aravin, A. A. (2014). piRNA pathway targets active LINE1 elements to establish the repressive H3K9me3 mark in germ cells. *Genes & Development*, *28*, 1410–1428. doi: 10.1101/gad.240895.114

Phizicky, E. M., & Hopper, A. K. (2010). tRNA biology charges to the front. *Genes & Development*, *24*, 1832–1860. doi: 10.1101/gad.1956510

Piekna-Przybylska, D., Decatur, W. A., & Fournier, M. J. (2008). The 3D rRNA modification maps database: With interactive tools for ribosome analysis. *Nucleic Acids Research*, *36*, D178–D183. doi: 10.1093/nar/gkm855

Raabe, C. A., Tang, T. - H., Brosius, J., & Rozhdestvensky, T. S. (2014). Biases in small

RNA deep sequencing data. *Nucleic Acids Research*, *42*, 1414–1426. doi: 10.1093/nar/gkt1021

Ramsköld, D., Luo, S., Wang, Y. - C., Li, R., Deng, Q., Faridani, O. R., … Sandberg, R. (2012). Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nature Biotechnology*, *30*, 777–782. doi: 10.1038/nbt.2282

Reinhart, B. J., Weinstein, E. G., Rhoades, M. W., Bartel, B., & Bartel, D. P. (2002). MicroRNAs in plants. *Genes & Development*, *16*, 1616–1626. doi: 10.1101/gad.1004402

Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., … Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature*, *403*, 901–906. doi: 10.1038/35002607

Roovers, E. F., Rosenkranz, D., Mahdipour, M., Han, C.-T., He, N., Chuva De Sousa Lopes, S. M., … Ketting, R. F. (2015). Piwi proteins and piRNAs in mammalian oocytes and early embryos. *Cell Reports*, *10*, 2069–2082. doi: 10.1016/j.celrep.2015.02.062

Ruby, J. G., Jan, C., Player, C., Axtell, M. J., Lee, W., Nusbaum, C., … Bartel, D. P. (2006). Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in C. elegans. *Cell*, *127*, 1193–1207. doi: 10.1016/j.cell.2006.10.040

Saliba, A.-E., Westermann, A. J., Gorski, S. A., & Vogel, J. (2014). Single-cell RNA-seq: Advances and future challenges. *Nucleic Acids Research*, *42*, 8845–8860. doi: 10.1093/nar/gku555

Schorn, A. J., Gutbrod, M. J., Leblanc, C., & Martienssen, R. (2017). LTR-retrotransposon control by tRNA-derived small RNAs. *Cell*, *170*, 61–71.e11. doi: 10.1016/j.cell.2017.06.013

Schutz, K., Hesselberth, J. R., & Fields, S. (2010). Capture and sequence analysis of RNAs with terminal 2′,3′-cyclic phosphates. *RNA*, *16*, 621–631. doi: 10.1261/rna.1934910

Sharma, U., Conine, C. C., Shea, J. M., Boskovic, A., Derr, A. G., Bing, X. Y., … Rando, O. J. (2016). Biogenesis and function of tRNA fragments during sperm maturation and fertilization in mammals. *Science (New York)*, *351*, 391–396. doi: 10.1126/science.aad6780

Sharma, U., Sun, F., Conine, C. C., Reichholf, B., Kukreja, S., Herzog, V. A., … Rando, O. J. (2018). Small RNAs are trafficked from the epididymis to developing mammalian sperm. *Developmental Cell*, *46*, 481–494.e6. doi: 10.1016/j.devcel.2018.06.023

Shi, J., Zhang, Y., Tan, D., Zhang, X., Yan, M., Zhang, Y., … Chen, Q. (2021). PANDORA-seq expands the repertoire of regulatory small RNAs by overcoming RNA modifications. *Nature Cell Biology*, *23*, 424–436. doi: 10.1038/s41556-021-00652-7

Shigematsu, M., Kawamura, T., & Kirino, Y. (2018). Generation of 2′,3′-cyclic phosphate-containing rnas as a hidden layer of the transcriptome. *Frontiers in Genetics*, *9*, 562, doi: 10.3389/fgene.2018.00562

Shin, C., Nam, J. - W., Farh, K. K. H., Chiang, H. R., Shkumatava, A., & Bartel, D. P. (2010). Expanding the microRNA targeting code: Functional sites with centered pairing. *Molecular Cell*, *38*, 789–802. doi: 10.1016/j.molcel.2010.06.005

Shin, G., Koo, H. J., Seo, M., Lee, S. -J. V., Nam, H. G., & Jung, G. Y. (2021). Transfer RNA-derived fragments in aging Caenorhabditis elegans originate from abundant homologous gene copies. *Science Report*, *11*, 12304, doi: 10.1038/s41598-021-91724-z

Sloan, K. E., Warda, A. S., Sharma, S., Entian, K. - D., Lafontaine, D. L. J., & Bohnsack, M. T. (2017). Tuning the ribosome: The influence of rRNA modification on eukaryotic ribosome biogenesis and function. *RNA Biology*, *14*, 1138–1152. doi: 10.1080/15476286.2016.1259781

Soares, A. R., & Santos, M. (2017). Discovery and function of transfer RNA-derived fragments and their role in disease. *Wiley Interdisciplinary Reviews. RNA*, *8*, e1423. doi: 10.1002/wrna.1423

Sobala, A., & Hutvagner, G. (2013). Small RNAs derived from the 5′ end of tRNA can inhibit protein translation in human cells. *RNA Biology*, *10*, 553–563. doi: 10.4161/rna.24285

Sorefan, K., Pais, H., Hall, A. E., Kozomara, A., Griffiths-Jones, S., Moulton, V., & Dalmay, T. (2012). Reducing ligation bias of small RNAs in libraries for next generation sequencing. *Silence*, *3*, 4, doi: 10.1186/1758-907X-3-4

Sporn, M. B., Lazarus, H. M., Smith, J. M., & Henderson, W. R. (1969). Studies on nuclear exoribonucleases. 3. Isolation and properties of the enzyme from normal and malignant tissues of the mouse. *Biochemistry*, *8*, 1698–1706. doi: 10.1021/bi00832a053

Su, Z., Kuscu, C., Malik, A., Shibata, E., & Dutta, A. (2019). Angiogenin generates specific stress-induced tRNA halves and is not involved in tRF-3-mediated gene silencing. *The Journal of Biological Chemistry*, *294*, 16930–16941. doi: 10.1074/jbc.RA119.009272

Taoka, M., Nobe, Y., Yamaki, Y., Yamauchi, Y., Ishikawa, H., Takahashi, N., … Isobe, T. (2016). The complete chemical structure of Saccharomyces cerevisiae rRNA: Partial pseudouridylation of U2345 in 25S rRNA by snoRNA snR9. *Nucleic Acids Research*, *44*, 8951–8961. doi: 10.1093/nar/gkw564

Thompson, D. M., & Parker, R. (2009). The RNase Rny1p cleaves tRNAs and promotes cell death during oxidative stress in Saccharomyces cerevisiae. *The Journal of Cell Biology*, *185*, 43–50. doi: 10.1083/jcb.200811119

Upton, H. E., Ferguson, L., Temoche-Diaz, M. M., Liu, X. - M., Pimentel, S. C., Ingolia, N. T., … Collins, K. (2021). Low-bias ncRNA libraries using ordered two-template relay: Serial template jumping by a modified retroelement reverse transcriptase. *Proceedings of the National*

Academy of Sciences of the United States of America, *118*, e2107900118. doi: 10.1073/pnas.2107900118

Vasudevan, S. (2012). Posttranscriptional upregulation by microRNAs. *Wiley Interdiscip Rev RNA*, *3*, 311–330. doi: 10.1002/wrna.121

Viollet, S., Fuchs, R. T., Munafo, D. B., Zhuang, F., & Robb, G. B. (2011). T4 RNA ligase 2 truncated active site mutants: Improved tools for RNA analysis. *BMC Biotechnology*, *11*, 72, doi: 10.1186/1472-6750-11-72

Wang, N., Zheng, J., Chen, Z., Liu, Y., Dura, B., Kwak, M., … Lu, J. (2019). Single-cell microRNA-mRNA co-sequencing reveals non-genetic heterogeneity and mechanisms of microRNA regulation. *Nature Communications*, *10*, 95, doi: 10.1038/s41467-018-07981-6

Wein, S., Andrews, B., Sachsenberg, T., Santos-Rosa, H., Kohlbacher, O., Kouzarides, T., … Weisser, H. (2020). A computational platform for high-throughput analysis of RNA sequences and modifications by mass spectrometry. *Nature Communications*, *11*, 926, doi: 10.1038/s41467-020-14665-7

Wightman, B., Ha, I., & Ruvkun, G. (1993). Post-transcriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. *Cell*, *75*, 855–862. doi: 10.1016/0092-8674(93)90530-4

Wolters, N. M., & Mackeigan, J. P. (2008). From sequence to function: Using RNAi to elucidate mechanisms of human disease. *Cell Death Differ*, *15*, 809–819. doi: 10.1038/sj.cdd.4402311

Wright, C., Rajpurohit, A., Burke, E. E., Williams, C., Collado-Torres, L., Kimos, M., … Shin, J. H. (2019). Comprehensive assessment of multiple biases in small RNA sequencing reveals significant differences in the performance of widely used methods. *BMC Genomics*, *20*, 513, doi: 10.1186/s12864-019-5870-3

Wulf, M. G., Maguire, S., Humbert, P., Dai, N., Bei, Y., Nichols, N. M., … Guan, S. (2019). Non-templated addition and template switching by Moloney murine leukemia virus (MMLV)-based reverse transcriptases co-occur and compete with each other. *The Journal of Biological Chemistry*, *294*, 18220–18231. doi: 10.1074/jbc.RA119.010676

Xiao, Z., Cheng, G., Jiao, Y., Pan, C., Li, R., Jia, D., … Jia, J. (2018). Holo-Seq: Single-cell sequencing of holo-transcriptome. *Genome Biology*, *19*, 163, doi: 10.1186/s13059-018-1553-7

Yamasaki, S., Ivanov, P., Hu, G. -F., & Anderson, P. (2009). Angiogenin cleaves tRNA and promotes stress-induced translational repression. *The Journal of Cell Biology*, *185*, 35–42. doi: 10.1083/jcb.200811106

Yang, Z., Ebright, Y. W., Yu, B., & Chen, X. (2006). HEN1 recognizes 21-24 nt small RNA duplexes and deposits a methyl group onto the 2′ OH of the 3′ terminal nucleotide. *Nucleic Acids Research*, *34*, 667–675. doi: 10.1093/nar/gkj474

Yu, B., Yang, Z., Li, J., Minakhina, S., Yang, M., Padgett, R. W., … Chen, X. (2005). Methylation as a crucial step in plant microRNA biogenesis. *Science (New York)*, *307*, 932–935. doi: 10.1126/science.1107130

Zhang, J., Yan, S., Chang, L., Guo, W., Wang, Y., Wang, Y., … Huang, S. (2020). Direct microRNA sequencing using nanopore-induced phase-shift sequencing. *iScience*, *23*, 100916, doi: 10.1016/j.isci.2020.100916

Zhang, N., Shi, S., Jia, T. Z., Ziegler, A., Yoo, B., Yuan, X., … Zhang, S. (2019). A general LC-MS-based RNA sequencing method for direct analysis of multiple-base modifications in RNA mixtures. *Nucleic Acids Research*, *47*, e125, doi: 10.1093/nar/gkz731

Zhang, Y., Zhang, X., Shi, J., Tuorto, F., Li, X., Liu, Y., … Chen, Q. (2018). Dnmt2 mediates intergenerational transmission of paternally acquired metabolic disorders through sperm small non-coding RNAs. *Nature Cell Biology*, *20*, 535–540. doi: 10.1038/s41556-018-0087-2

Zhu, Y. Y., Machleder, E. M., Chenchik, A., Li, R., & Siebert, P. D. (2001). Reverse transcriptase template switching: A SMART approach for full-length cDNA library construction. *Biotechniques*, *30*, 892–897. doi: 10.2144/01304pf02

Ziegenhain, C., Hendriks, G.-J., Hagemann-Jensen, M., & Sandberg, R. (2022). Molecular spikes: A gold standard for single-cell RNA counting. *Nature Methods*, *19*, 560–566. doi: 10.1038/s41592-022-01446-x