




Review

Small RNA-Sequencing: Approaches and Considerations for miRNA Analysis

Sarka Benesova ^{1,2}, Mikael Kubista ^{1,3} and Lukas Valihrach ^{1,*}

¹ Laboratory of Gene Expression, Institute of Biotechnology, CAS, BIOCEV, 252 50 Vestec, Czech Republic; sarka.benesova@ibt.cas.cz (S.B.); mikael.kubista@tataa.com (M.K.)

² Laboratory of Informatics and Chemistry, Faculty of Chemical Technology, University of Chemistry and Technology, 166 28 Prague, Czech Republic

³ TATAA Biocenter AB, 411 03 Gothenburg, Sweden

* Correspondence: lukas.valihrach@ibt.cas.cz

Abstract: MicroRNAs (miRNAs) are a class of small RNA molecules that have an important regulatory role in multiple physiological and pathological processes. Their disease-specific profiles and presence in biofluids are properties that enable miRNAs to be employed as non-invasive biomarkers. In the past decades, several methods have been developed for miRNA analysis, including small RNA sequencing (RNA-seq). Small RNA-seq enables genome-wide profiling and analysis of known, as well as novel, miRNA variants. Moreover, its high sensitivity allows for profiling of low input samples such as liquid biopsies, which have now found applications in diagnostics and prognostics. Still, due to technical bias and the limited ability to capture the true miRNA representation, its potential remains unfulfilled. The introduction of many new small RNA-seq approaches that tried to minimize this bias, has led to the existence of the many small RNA-seq protocols seen today. Here, we review all current approaches to cDNA library construction used during the small RNA-seq workflow, with particular focus on their implementation in commercially available protocols. We provide an overview of each protocol and discuss their applicability. We also review recent benchmarking studies comparing each protocol's performance and summarize the major conclusions that can be gathered from their usage. The result documents variable performance of the protocols and highlights their different applications in miRNA research. Taken together, our review provides a comprehensive overview of all the current small RNA-seq approaches, summarizes their strengths and weaknesses, and provides guidelines for their applications in miRNA research.



Citation: Benesova, S.; Kubista, M.; Valihrach, L. Small RNA-Sequencing: Approaches and Considerations for miRNA Analysis. *Diagnostics* **2021**, *11*, 964. <https://doi.org/10.3390/diagnostics11060964>

Academic Editor: Anastasia Malek

Received: 31 March 2021

Accepted: 24 May 2021

Published: 27 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: small RNA-seq; miRNA; diagnostics

1. Introduction

MicroRNAs (miRNAs) represent a class of short (~22 nucleotides) non-coding RNA molecules that are well conserved across various species. They are known mainly for their post-transcriptional regulation of gene expression via degradation and translational repression [1]. However, there is a growing body of evidence that they can also act as an expression activator upon specific conditions such as cell cycle arrest [2–4]. Moreover, the localization of mature miRNA formed in the nucleus has led to the discovery of their ability to regulate gene expression even at the transcriptional level [5].

MiRNA genes are located both within defined transcription units as well as in intergenic regions, where they create clusters transcribed as polycistronic RNA [6]. Their biogenesis involves transcription by RNA polymerase II resulting in structured primary miRNA (pri-miRNA), which is processed by the Microprocessor complex into a precursor miRNA (pre-miRNA). Pre-miRNA is subsequently exported to the cytoplasm and cleaved by the Dicer protein complex, creating the duplex miRNA. One strand of miRNA duplex is loaded into the Argonaut (AGO) protein, creating the RNA-induced silencing complex (RISC), which then binds to the mRNA using the complementary seed sequence

of the loaded mature miRNA [7]. Alternative cleavage by the Microprocessor complex and the dicer leads to the production of isomiRs, biologically active miRNA isoforms with alternative 3'- and/or 5'-end compared to the canonical mature sequence [8].

Since the discovery of the first miRNA *lin-4* in nematodes, more than 38,000 miRNAs have been detected in 271 organisms, including 1917 human miRNAs (miRBase v22) [9]. It has been proposed that more than 60% of human protein-coding genes can be directly post-transcriptionally regulated by miRNAs [10]. As a single miRNA can regulate up to hundreds of genes, miRNAs are supposed to act as hub genes in regulatory networks and are thought to play a role in almost all biological processes. In the field of cancer research, miRNAs became the subject of numerous studies, which demonstrated their active role in the disease and showed cancer-specific gene expression [11–14]. The presence of cancer-specific miRNA profiles, together with the presence and stability of miRNAs in various biofluids have made them useful biomarkers allowing for non-invasive cancer diagnosis and prognosis [15,16].

Microarrays, reverse transcription-quantitative PCR (RT-qPCR) and RNA-sequencing (RNA-seq) are the three most common approaches for miRNA profiling. For a long time, microarray analysis represented the most commonly used high-throughput technology for miRNA profiling [12]. However, this hybridization-based technique does not allow for absolute quantification, identification of novel miRNAs, and separate detection of canonical miRNAs and their isomiRs, of which the importance has only started to be recognized recently [17–19]. Moreover, microarray-based methods usually require high quantities of input RNA. This may be a limiting factor for the non-invasive analysis of liquid biopsies, where the concentration of miRNAs is rather low [20]. Compared to the microarray, RT-qPCR allows for absolute quantification, has a broad dynamic range, high sensitivity and allows for the analysis of low RNA inputs [21]. For its non-demanding instrumentation and availability of expertise in many laboratories, RT-qPCR stands out as a gold standard method for targeted miRNA analysis. RNA-seq represents the last major approach for miRNA analysis. The main advantage of small RNA-seq is its non-targeted character, allowing for simultaneous detection of novel miRNAs, detection of isomiRs as well as other small RNA species. The disadvantages are represented by the complex experimental workflow which introduces various types of biases and demanding computational analysis. To minimize these disadvantages, there have been substantial technological developments within the past five years resulting in the introduction of several new small RNA-seq protocols (Table 1). However, a comprehensive overview and comparison of all the methods, which would serve as a guideline for new users, is missing. Although several comparative studies testing the performance of the new small RNA-seq protocols have appeared in recent years, they have not been properly reviewed yet. To fill the gap, we summarized the recent knowledge on new approaches for small RNA-seq analysis as well as reviewed all recent comparative studies. Here, we aim to provide a comprehensive overview of the small RNA-seq technology that will facilitate to any new user the ability to make an informed selection of the best-suited protocol for a particular experimental question. We hope that our efforts will promote the standardization of the technology and its broader applications in the near future.

Table 1. Overview of commercially available protocols for small RNA-seq analysis.

Technology	Product Name	Company	Reference
Original two-adaptor ligation	Small RNA-Seq Library Prep Kit	Lexogen GmbH, Vienna, Austria,	not available
	Small RNA Library Prep Kit	Norgen Biotek Corp., Thorold, ON, Canada	not available
	TruSeq Small RNA Library Prep Kit	Illumina, San Diego, CA, USA	not available
	TailorMix miRNA Sample Preparation Kit	SeqMatic, Fremont, CA, USA	not available
	NEBNext Multiplex Small RNA Library Prep Set	New England Biolabs, Ipswich, MA, USA	not available
	CleanTag Small RNA Library Prep Kit	TriLink BioTechnologies, Inc., San Diego, CA, USA	[22]
	ScriptMiner Library preparation Technology for Small RNA	Cambio Ltd., Cambridge, UK	[23]
Randomized adaptors	NEXTflex Small RNA Sequencing Kit	PerkinElmer, Waltham, MA, USA	[24]
Single adaptor ligation and circularization	RealSeq-AC Kit	Somagenics, Santa Cruz, CA, USA	[25]
	RealSeq-biofluids Kit	Somagenics, Santa Cruz, CA, USA	not available
	SMARTer microRNA-Seq Kit	Takara Bio, Shiga, Japan	not available
UMI	TrueQuant SmallRNA Seq Kit for Ultra Low Input	GenXPro GmbH, Frankfurt Main, Germany	not available
	QIAseq miRNA Library Kit (12bp UMIs)	Qiagen, Hilden, Germany	not available
Polyadenylation and template switching	SMARTer smRNA-seq Kit	Takara Bio, Shiga, Japan	not available
	CATS Small RNA-seq Kit	Diagenode, Liege, Belgium	not available
Sequencing of hybridization probes	HTG EdgeSeq miRNA Whole Transcriptome Assay	HTG Molecular Diagnostics, Inc., Tuscon, AZ, USA	not available
Hybridization based techniques without NGS readout	FirePlex miRNA assays	Abcam, Cambridge, UK	not available
	nCounter miRNA Expression Panels	NanoString, Seattle, WA, USA	[26]

2. Small RNA-seq Technology Overview

RNA-seq has undoubtedly revolutionized the characterization of the small transcriptome, enabling high-throughput profiling and discovery of novel forms of short non-coding RNAs (miRNAs, piRNAs, tRNAs, siRNAs, snoRNA, etc.) [27]. The first application was demonstrated in 2005, when small RNA-seq was used to profile various types of small RNAs of *Arabidopsis thaliana* [28]. A year later, small RNA-seq identified 18 new miRNAs in *Caenorhabditis elegans* [29]. Since then, there have been more than 700 human miRNA expression profiling studies, which used small RNA-seq and were deposited on Gene Expression Omnibus (GEO) [30]. With the development of single-cell RNA-seq, it is now possible to analyze miRNAs even in individual cells [31].

The small RNA-seq workflow involves three main steps: (i) isolation of RNA; (ii) cDNA library construction; and (iii) sequencing [32]. Small RNA may be isolated using conventional total RNA extraction methods followed by length separation or binding of small RNAs to specific proteins. As small RNA enrichment may lead to losses [33] and/or isolation-specific biases [34,35], analysis of total RNA is an alternative and usually preferred strategy. For example, recent studies reported a higher proportion of reads mapping to small RNA/miRNAs and a higher number of detected miRNAs with protocol isolating total RNA from human plasma than with small RNA enrichment protocol [34,36]. The library construction starts with reverse transcription (RT) of small RNAs, producing complementary DNA (cDNA) for amplification and sequencing. As miRNA molecules are of short length, they are extended by ligation or polyadenylation, which introduce primer-binding sites for RT and subsequent amplification. The extension step, especially the ligation, is considered as the most prominent source of bias in small RNA-seq [32,37,38]. Different affinity of adaptors to target molecules causes artificial changes in the true small RNA abundances, which confounds the analysis. The amplification step of cDNA molecules represents another source of bias, which originates from the differing efficiencies of the PCR while amplifying molecules of different lengths and secondary structures [39]. This leads again to artificial shifts in the detected signals. This PCR bias is well known in the field of mRNA sequencing, and is usually mitigated through the use of unique molecular identifiers (UMIs) [40,41]. UMIs allow for the distinguishing of reads generated from an

identical molecule that was amplified by PCR. The last step in the small RNA-seq workflow is sequencing. Although this step might be associated with certain types of biases (lane-, flow-cell effects), these are mostly considered as minor ones having a subtle effect on the overall bias level [42].

During the past years, the major technical development in the small RNA-seq was focused on the reduction of the ligation and PCR bias. At present, there are 15 commercial protocols available, which cover three distinct approaches: (i) original two-adaptor ligation; (ii) improved two-adaptor ligation; and (iii) ligation-free (polyadenylation-based) approach (Table 1). Besides these methods, there is a group of alternative technologies relying on the hybridization of miRNAs to target probes that completely avoid the ligation or polyadenylation step and in some cases also the PCR step and NGS readout. These techniques usually offer a simplified workflow that is suitable for routine and automatized analysis. On the other hand, they are limited in the spectrum of analyzable miRNAs (given by the set of predesigned probes), which limits their discovery potential.

2.1. Two-Adaptor Ligation-Based Methods

The traditional two-adaptor ligation approach utilizes the ligation of the pre-adenylated 3'-end adaptor followed by the ligation of the 5'-end adaptor (Figure 1). Although this approach has been successfully applied in the majority of small RNA-seq studies (for historical reasons), its application is associated with several issues, including: (i) ligation bias; (ii) preferences to certain small RNA classes; and (iii) formation of adaptor dimers.

First, the ligation reaction is a major source of bias that negatively affects miRNA quantification in traditional two-adaptor small RNA-seq protocols [37,43,44]. Consequently, measured profiles do not reflect true miRNA abundances, the majority of miRNAs are systematically over- or under-estimated and some may be completely undetected. The primary cause of this phenomenon was shown to be secondary structures of both ligation substrates, i.e., miRNAs and adaptors [38,45]. The secondary structure of the adaptor defines the preference to certain miRNAs and vice versa. This results in some miRNAs ligating with high efficiency whereas others with substantially lower, leading to bias in the data. The issue might be diminished by the application of the improved two-adaptor ligation methods (see next chapter). Apart from that, there is also a possibility for in silico correction of measured expression values [46]. The model uses the calculation of bias factor, which is learned for each miRNA on an equimolar pool of synthetic miRNAs and is subsequently applied for the correction of real data. Although promising, the correction is problematic for low input samples where few miRNAs consume most of the reads. This can result in some of the low abundant miRNAs being completely missed (dropouts) with the inability for correction.

Secondly, as the method utilizes T4 RNA ligases that recognize the hydroxyl group on the 3'-end and the phosphate group on the 5'-end of miRNAs, it enables profiling of all RNA species with such modifications. This includes long non-coding RNAs, which need to be removed by bead-based size selection or polyacrylamide gel electrophoresis (PAGE) purification. Other small RNAs with ligase-compatible 5'- and 3'-end modifications and similar lengths, such as piRNAs, are retained in the library and analyzed [42], whereas others with ligase-incompatible modification will be missing. For example, small RNAs with trimethyl-guanosine cap on 5'-end such as small nuclear RNAs (snRNAs) are not captured and sequenced. For profiling of these missing small RNA classes, there are protocols such as ScriptMiner™ (Cambio Ltd.) utilizing enzymes, which degrade caps and polyphosphate groups on the 5'-end and make these molecules analyzable [23,47]. Another approach to capture small RNAs regardless of their 5'-end modifications is to add the 5'-adaptor after RT and target the cDNA molecules [48,49]. Lastly, the recently-introduced phospho-RNA-seq method utilizes T4-polynucleotide kinase, which is used for phosphorylation and dephosphorylation in order to capture small RNAs regardless of their end modifications [50]. Focused on 3'-end, utilization of T4 Rnl2tr ligase in optimized conditions improves capturing of small RNA species with 2'-O-methyl on 3'-end [51].

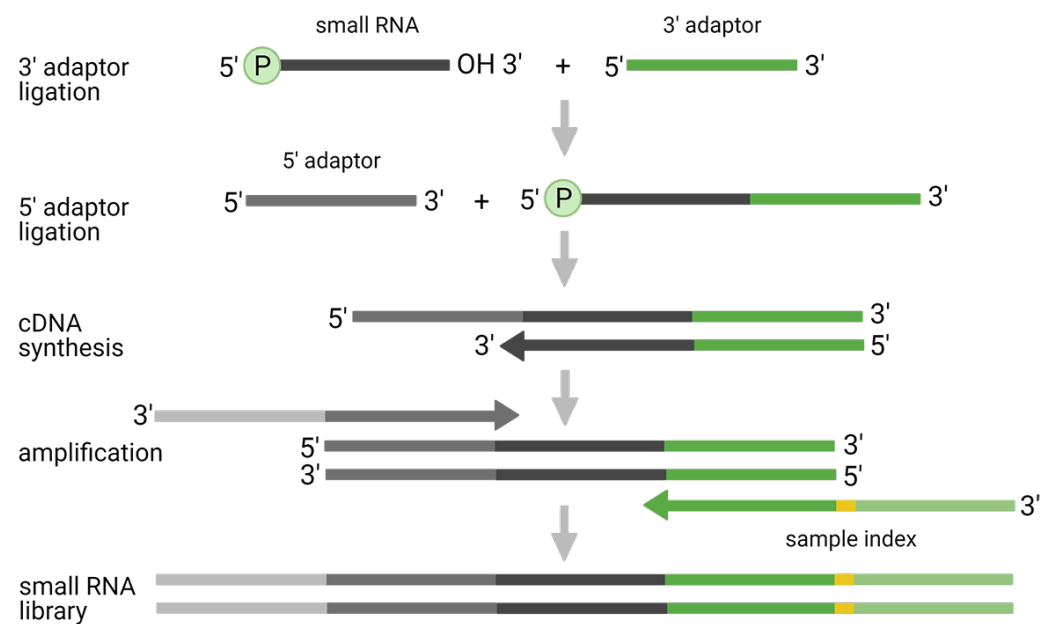


Figure 1. Original two-adaptor ligation protocol for small RNA-seq analysis.

The final problematic aspect of the traditional two-adaptor approach is the formation of adaptor dimers, which may consume a substantial proportion of the sequenced reads when not removed properly. Especially in low input samples, where libraries are prepared using a higher number of PCR cyclers, this may cause complete loss of lowly abundant small RNAs. There are multiple approaches how to remove or limit the formation of adaptor dimers. These include: (i) bead-based or PAGE removal of adaptors from libraries [52]; (ii) enzymatic removal of 3'-adaptor excess limiting the formation of dimers [23,47]; (iii) usage of chemically-modified adaptors which do not ligate without insert [22]; (iv) annealing of RT primer in equimolar concentration to 3'-adaptor before ligation of 5'-end adaptor which blocks RT of dimers [53]; (v) blocking RT of adaptor dimers by adding oligonucleotide with complementary sequence [54]; (vi) and usage of CRISPR/Cas9 carrying sgRNAs specific to adaptor dimers which leads to their selective depletion [55]. The last approach can be used also for the removal of highly abundant miRNAs, which would consume a large proportion of the reads, but are not of interest. An alternative method for the removal of highly abundant miRNAs is to block them as a substrate for ligase by the addition of oligonucleotide complementary to the undesired miRNA [56,57]. Some of the aforementioned approaches to prevent adaptor-dimers are a part of commercially available ligation-based protocols, including the original two-adaptor ligation-based (Table 1). Moreover, there are other improvements such as PEG addition, specific adaptor concentrations, and specific ligation temperature, which increase the yield of ligation and may contribute to the mitigation of overall bias [22].

2.2. Improved Two-Adaptor Ligation-Based Methods

The improved two-adaptor ligation approaches aim to prevent or mitigate the ligation or PCR bias. Currently, there are three approaches available: (i) ligation of two adaptors with randomized nucleotides (Figure 2A); (ii) ligation of single adaptor and subsequent circularization (Figure 2B); and (iii) ligation of two adaptors with UMIs (Figure 2C).

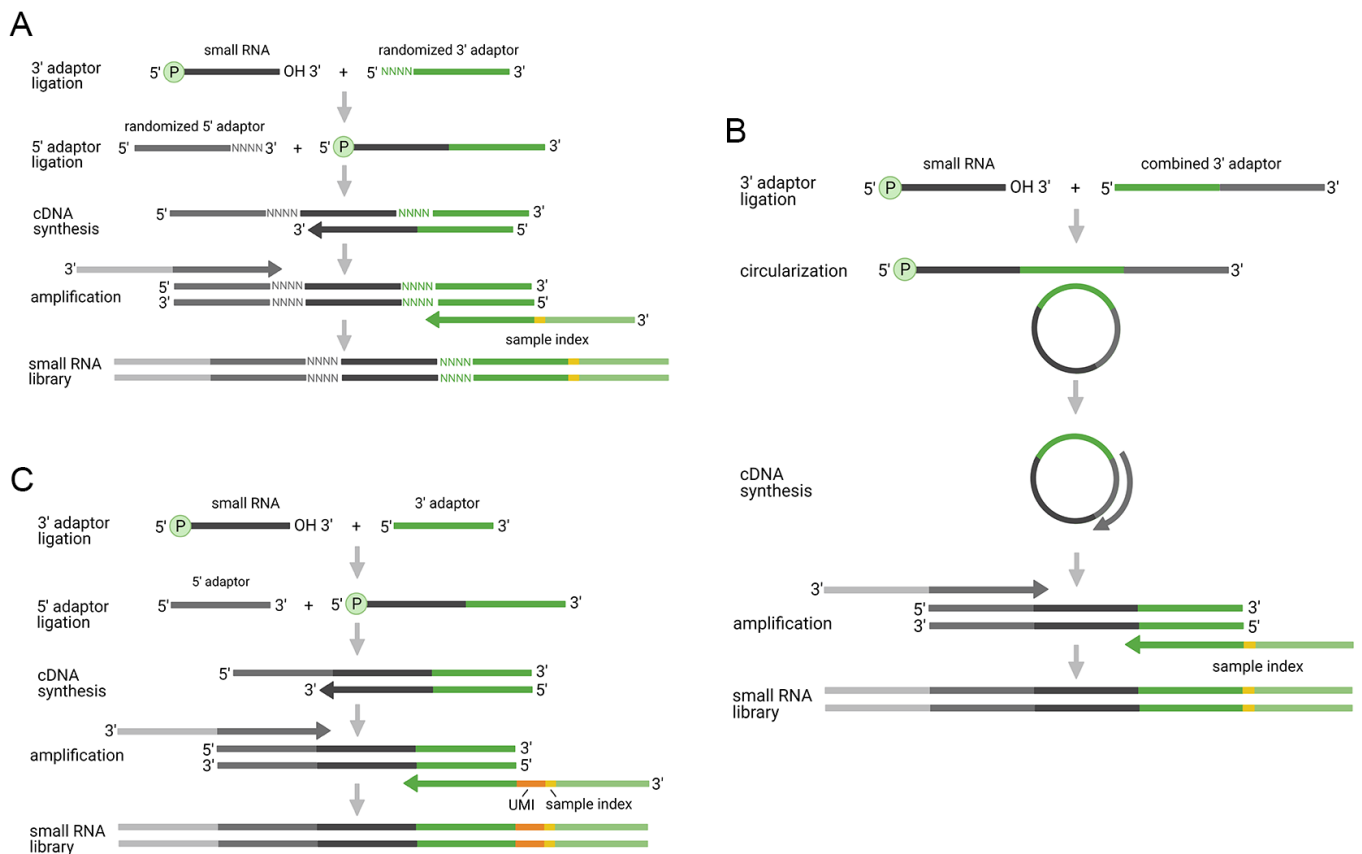


Figure 2. Improved two-adaptor ligation-based methods for small RNA-seq analysis. The reduction of ligation or PCR bias is achieved by: (A) randomized adaptors; (B) single adaptor-ligation and circularization; and (C) UMIs.

The superior performance of methods utilizing randomized adaptors over the traditional two-adaptor approach is frequently reported in the literature [58–60], including our own data [61]. Currently, there is only a single commercially-available protocol—NEXTflex provided by PerkinElmer. However, the superior performance of the randomized-adaptor approach was reported using an in-house protocol as well [50]. The application of adaptors that contain a random oligonucleotide part was originally proposed to study the cause of ligation bias on the 3'-end of miRNAs [62]. The study showed that ligases do not prefer any miRNAs according to their primary sequence, but they are rather biased by secondary structures presented in the adaptors and targeted small RNAs. Thus, the inclusion of random parts in the adaptors, efficiently decreased ligation bias. Similarly, the usage of “High Definition” (HD) adaptors containing four random bases in both adaptors led to a substantial decrease of ligation bias [44]. Of note, a subsequent study complemented the data and demonstrated that the inclusion of randomized parts diminished the ligation bias regardless of the distance to ligation junction [38]. The finding that miRNAs preferentially ligate to the adaptors which increase the potential to create secondary structures motivated the development of a new approach for optimal adaptor design [38]. Using this model, adaptors are designed to include randomized oligonucleotides and complementary regions between 5'- and 3'-adaptors, which induce structures preferred by ligase. Lastly, a recent study showed that small RNA capture efficiency can be further improved by a combination of randomized adaptors with the addition of nucleotide modifications and optimized adaptor concentration [63]. Noteworthy, randomized nucleotides can be used as UMIs as well and mitigate the PCR bias, as has been demonstrated in a recent study [60]. The utilization of 4 + 4 random nucleotides from both 3'- and 5'-end adaptors led to a substantial decrease of PCR bias, which was estimated using an equimolar pool of 962 miRNAs. However, the experimental design did not account for the total number of

molecules and the limiting number of unique combinations given by $4 + 4$ nucleotides (4^8), which caused over-deduplication and over-estimation of PCR bias [64]. Therefore, the use of randomized adaptors as UMIs should be made with caution and should be considered with respect to the concentration of the ligated RNAs, especially those presented in high copy numbers.

The ligation of a single 3'-end adaptor and subsequent circularization represents another promising approach that has been recently introduced to mitigate ligation bias [25]. The decrease of ligation bias is primarily achieved by intramolecular circularization, which is more efficient than intermolecular 5'-ligation. To prevent self-circularization of the 3'-adaptor, its 3'-end is blocked by a phosphate group. After successful ligation of the 3'-adaptor, the miRNA-adaptor construct is purified, dephosphorylated, circularized and RT is performed on the circularized molecules. Although the system is well designed to mitigate 5'-end bias, 3'-end bias is supposed to be still present. This is reflected by contradictory results in the literature. Whereas the original and recent benchmarking studies showed less ligation bias than the protocol based on randomized adaptors [25,65], ours and others data showed rather average to poor performance [61,66]. However, the contradictory results were achieved by different chemistries. The positive data were reported using the Somagenics RealSeq-AC Kit intended for tissue samples, whereas later two further reports used Somagenics RealSeq-biofluids Kit or SMARTer microRNA-Seq Kit from Takara Bio. Further assessment of protocol-specific parameters and their impact on the level of ligation bias is therefore needed to provide further insight.

The last group of methods focused on the mitigation of PCR bias. This is achieved via the utilization of UMIs that may be introduced in one or both adaptor sequences, alternatively in the RT primer. There is an ongoing debate on the importance of PCR bias. Where early reports highlighted a minor role of PCR bias [38,43,67], recent studies called this in question [41,60]. The amount of starting material and sequencing depth was shown to influence PCR duplicates frequency and their removal increased reproducibility of the data [41]. Currently, two-adaptor ligation protocols with UMIs are commercially available from Qiagen and GenXPro [68]. The QIAseq miRNA Library Kit from Qiagen includes adaptors containing 12bp UMIs, which produce enough unique sequences even for high abundant miRNAs, therefore, deduplication does not result in their under-estimation. In comparison, the protocol performs worse than approaches using randomized adaptors, but better than original two-adaptor ligation protocols [65,66]. Our own recent study showed a similar outcome, scoring the QIAseq protocol close to other ligation mitigating approaches, suggesting that also other protocol-specific factors might play a role in its improved performance [61].

2.3. Ligation Free Methods

Ligation-free methods include protocols based on polyadenylation with template switching (Figure 3) and probe-based techniques, which enable targeted profiling of known miRNAs. Although the second group of methods does not rely, in all cases, on RNA-seq, they were included in this chapter as they represent an interesting alternative to traditional small RNA-seq. They offer a simplified workflow and computational analysis, while bypassing the biases associated with the ligation and PCR step.

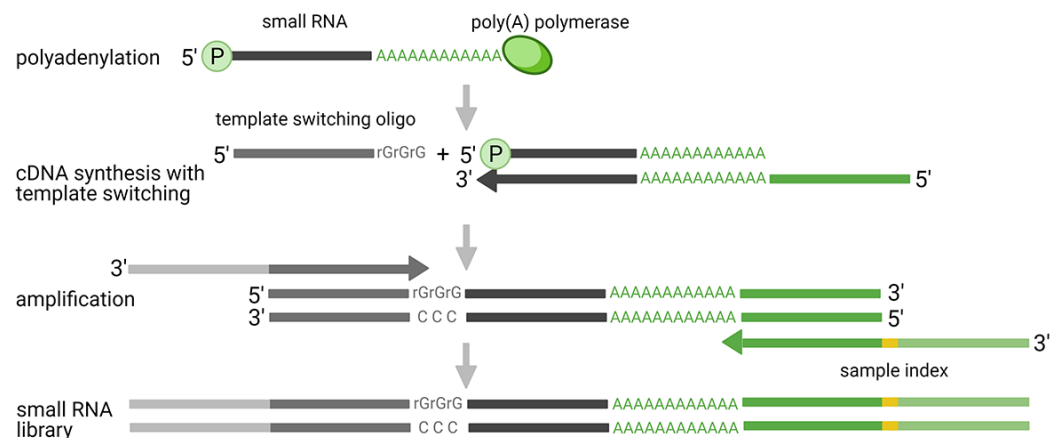


Figure 3. Polyadenylation and template switching mechanism applied in small RNA-seq analysis.

Historically, the polyadenylation method was introduced together with the original two-adaptor ligation approach and contributed to the discovery of various mammalian miRNAs [69]. The first protocol employed polyadenylation of the 3'-end and subsequent ligation of the 5'-end adaptor. Nowadays, the 5'-ligation was replaced by the addition of the 5'-adaptor using template-switching activity of reverse transcriptase (RTase), a mechanism originally introduced for full-length reverse transcription of mRNAs [70]. This technology uses Moloney murine leukemia virus (MMLV) RTase, which adds non-templated cytidines at the 5'-end of each cDNA molecule. The added stretch of nucleotides consequently serves as a binding site for 5'-end adaptor. Currently, there are two companies providing this type of solution for small RNA analysis: Diagenode and Takara Bio. The major advantage of these protocols is the independence on ligation, which ensures more accurate miRNAs quantification compared to ligation approaches. However, there are also disadvantages, which need to be considered. In recent comparative studies, both of the protocols showed that a high proportion of reads did not align to miRNA sequences, but to other small RNA classes, and also that a high proportion of reads did not map at all [65,66]. Consequently, the low miRNA mapping rate might increase the sequencing costs (to achieve the desired coverage of miRNAs) or result in an insufficient number of miRNA reads, thereby limiting further analysis. On the other hand, the signals from other small RNA classes may provide extra information and lead to new discoveries. Another aspect for consideration is that polyadenylation does not add a specific number of adenosines to the template. Therefore, it is not possible to distinguish native nucleotides from those artificially added and determine the exact 3'-end of sequenced miRNAs. Similarly, the template switching activity of RTase is supposed to add three non-templated cytidines. However, the exact number may vary, which complicates recognition of the 5'-end of miRNAs. These biases cause problems during the mapping of sequencing reads to the reference, which results in a high proportion of discarded reads. Moreover, the problems with recognition of exact 5'- and 3'-end impair quantification of isomiRs and other closely related molecules.

The application of small RNA-seq methods is especially advantageous where novel forms of small RNA/miRNAs are of interest. However, these methods suffer from protocol-specific biases that do not allow for the direct cross-study comparison and complicate meta-analysis. Moreover, the complex workflow negatively affects their applicability for clinical diagnostics where rapid turnaround time and robust protocols are desirable. To answer these requests, several modern hybridization methods have been recently developed, including Nanostring nCounter, Abcam FirePlex and EdgeSeq from HTG Molecular Diagnostics. Nanostring nCounter utilizes direct molecular barcoding and digital detection of targeted molecules by color-coded probes [26,71]. Unique oligonucleotide tags (miRtags) are ligated to the 3'-end of targeted miRNA, whereas the specificity of the reaction is ensured by bridge oligonucleotide with complementary sequence to the targeted miRNA and miRtag. Created complexes are immobilized and after several washing steps, color-coded

barcodes are counted. Each barcode corresponds to a specific miRNA target, which allows its reliable identification. Currently, the technology allows for the quantification of over 800 highly curated human, mouse and rat miRNAs. FirePlex miRNA assay from Abcam leverages the hydrogel-based microparticle technology [72]. First, miRNAs are bounded to the central miRNA-specific region of hydrogel particles, which contain two end regions with different fluorescent intensities serving as barcodes. Bounded miRNAs are ligated to universal adaptors and eluted from hydrogel particles. Subsequently, miRNAs are amplified using biotinylated primers complementary to the adaptors. After amplification, biotinylated miRNA targets are re-hybridized to hydrogel particles and a flow cytometer is used for detection of reporters specific to biotin. This method allows for quantification of 65 miRNAs in a single reaction, different disease-specific panels are available, and the workflow enables analysis of crude biofluid samples, eliminating bias introduced by isolation [34,35,73]. The analysis of crude biofluid samples is also possible with EdgeSeq from HTG Molecular Diagnostics. The protocol utilizes hybridization probes designed for specific binding of 2083 miRNAs annotated in the miRBase v20 database. After hybridization, unbound RNA and probes are digested by S1 nuclease. Targeted miRNAs are further released from capturing probes, these are amplified and sequenced using Illumina sequencing [74]. The complete process of library preparation is automatized, which saves hands-on time and eliminates errors. Moreover, the computational analysis is standardized leading to reliable results that are comparable across laboratories.

3. Benchmarking Studies

Over the past five years, several new approaches for small RNA profiling have been introduced and the majority of them transformed into commercial products (Table 1). This naturally raised an interest to compare their performance in various experimental settings, using different types of samples and evaluation metrics. The tested samples include plant or human total RNA, RNA isolated from immune cells or cell cultures, human serum and plasma samples for their frequent use in diagnostics, and lastly equimolar mixture of artificial miRNAs (miRXplore) allowing quantification of the ligation and PCR bias. The basic parameters, which were evaluated in most of the benchmarking studies, include: (i) yield—the proportion of mapped reads; (ii) reproducibility—variation of technical replicates; (iii) specificity—ability to detect only true targeted sequences; (iv) sensitivity—ability to detect targeted sequences considering given sequencing depth and minimal number of targeted miRNA reads; (v) accuracy—proportion of over- or under-estimated sequences; and (vi) ability to detect differentially expressed RNAs. Additionally, protocols were evaluated in terms of the workflow complexity (hands-on time, number of pipetting steps, robustness), ability to analyze low RNA input samples, level of multiplexing, and total costs.

We conducted a thorough literature search and identified twelve comparative studies that compared at least two different approaches for small RNA-analysis. The results are summarized in Table 2, alongside major conclusions. Only three studies, including our own, compared all available approaches for small RNA-seq [25,61,65]. In the study of Barberan-Soler et al. [25], the authors introduced the circularization approach and compared it with remaining protocols. The circularization protocol outperformed others in most of the tested metrics. It greatly reduced ligation bias and enabled the identification of the largest variety of miRNAs found in human brain RNA. The reduced bias allowed also for a robust miRNA quantification regardless of RNA input levels. The study of Herbert et al. [65] confirmed the result from the previous study, demonstrating a high accuracy and sensitivity of the circularization approach. The comparison was conducted on total RNA isolated from mesenchymal stem cells and miRXplore and contained all major RNA-seq approaches. Of note, both of the studies used an early-release beta version of the Somagenics RealSeq-AC protocol. In our recent study, we used Somagenics RealSeq-Biofluids protocol and observed a rather moderate reduction of ligation bias [61]. The non-optimal performance of another circularization protocol SMARTer microRNA-Seq Kit (Takara Bio) was reported by the study of Heinicke et al. [66], where the majority of

reads were lost during the mapping procedure. Together, the results showed that other protocols, or sample-specific factors, might play a role and therefore require further investigation. The study of Herbert et al. [65] included also the hybridization-based approach nCounter (NanoString) and showed its low correlation with other RNA-seq based methods. This is in accordance with a recent study comparing RNA-seq with all modern hybridization platforms, demonstrating substantial protocol-specific biases limiting inter-platform comparison [75]. Lastly, our recent study compared all available RNA-seq approaches for miRNA analysis in human plasma, including the hybridization-based EdgeSeq protocol from HTG Molecular Diagnostics [61]. As others, we showed that the performance and spectrum of captured miRNAs is greatly affected by the library preparation protocol. In concordance with Godoy et al. [75], we documented a superior performance of EdgeSeq for accuracy, mapping rate and sensitivity, but it lacked in specificity. The most balanced performance was shown by randomized approach reaching well to average results in the majority of parameters. The similar conclusion was reached in the majority of other comparative studies (even though limited in the spectrum of compared methods) [58–60,75,76]. This emphasizes the robustness of the randomized protocol, which has been also recently selected by the Extracellular RNA Communication Consortium (ERCC) as a standard for their RNA-seq based studies [59,77]. Surprisingly, a good performance was achieved by QIAseq, although we were not able to confirm the beneficial effect of UMIs. Again, this is in concordance with the study of Herbert et al. [65], where QIAseq clustered with randomized and circularization approach that was the best rated in this comparison. Noteworthy, some studies performed using low input samples reported an even better QIAseq performance than achieved by using the randomized approach [36,78], suggesting the influence of template concentration. Although polyadenylation-based methods were shown to be more accurate than traditional ligation-based methods, their yield and overall mapping statistics are substantially lower [61,66]. A recent study showed that the proportion of mapped reads obtained using the Diagenode CATS protocol did not even reach units of percentage [66], even though such dramatic loss of reads was probably a consequence of the uniform mapping strategy used across all protocols in the study. Lastly, all the studies agreed on poor performance when using the traditional two-adaptor ligation protocols that introduce the largest proportion of bias that negatively affects most of the other metrics [36,60,76].

To sum up, the approach using randomized adaptors seems to be the most robust approach currently available. The status of this well-established method will probably even increase in the future, as some of the research consortia (ERRC) selected its application as a standard for small RNA-Seq studies [59,77]. For routine application, targeted approach EdgeSeq provides superior performance, although some protocol-specific negatives (lower specificity, expensive instrumentation) need to be considered before its employment.

Table 2. Overview of recently published benchmarking studies of small RNA-seq protocols. Some protocols were benchmarked when offered by different companies, therefore, the names may vary from the Table 1.

Benchmarking Study	Included Approaches	Commercial Protocols Included	Sample Types	Conclusion
Dard-Dascot, C. et al. BMC Genomics 19, 1–16 (2018) [58].	Original two-adaptor ligation	TruSeq Small RNA Library Prep Kit (Illumina) + in house modifications	Synthetic RNAs	Protocol utilizing randomized adaptors performed best on both human and plant miRNAs. Addition of polyethylene glycol (PEG) and usage of chimeric DNA-RNA 5' adaptor with random nucleotides led to improved performance.
	Randomized adaptors	NEXTflex (Bioo Scientific) + in house modifications	<i>Arabidopsis thaliana</i> total RNA	
	Polyadenylation and template switching	SMARTer (Clontech) and CATS (Diagenode)	Oilseed rape total RNA HeLa cells total RNA	
Coenen-Stass, A. M. L. et al. RNA Biol. 15, 1133–1145 (2018) [78].	Original two-adaptor ligation	NEBNext Multiplex Small RNA Library Prep Set (New England Biolabs)	Human plasma miRNAs	Protocols utilizing randomized adaptors and UMIs performed well across all measured characteristics and showed the least sequence bias.
	Randomized adaptors	NEXTflex (Bioo Scientific)	Human serum miRNAs	
	Polyadenylation and template switching UMI	SMARTer (Clontech) QIAseq miRNA library Kit (Qiagen)	miRXPlore universal reference	
Giraldez, M. D. et al. Nat. Biotechnol. 36, 746–757 (2018) [59].	Original two-adaptor ligation	NEBNext Multiplex Small RNA Library Prep Set (New England Biolabs)	Human plasma total RNA	In-house adjusted protocol with randomized adaptors showed the least bias.
	Randomized adaptors	NEXTflex (Bioo Scientific) + in house modifications	Equimolar and ratiometric pool of synthetic RNAs	Optimization of ligation temperature and PEG helped to reduce bias.
Yeri, A. et al. BMC Genomics 19, 1–15 (2018) [76].	Original two-adaptor ligation	TruSeq Small RNA Library Prep Kit (Illumina)	Human brain total RNA	Protocol utilizing randomized adaptors detected the highest number of miRNAs and had the highest correlation with results from ligation free methods, but it required user experience to be performed consistently.
	Randomized adaptors	NEBNext Multiplex Small RNA Library Prep Set (New England Biolabs) NEXTflex (Bioo Scientific) + in house modifications	Human liver total RNA Human placenta total RNA	
	Sequencing of hybridization probes Hybridization-based technique	EdgeSeq (HTG Molecular Diagnostics) FirePlex (Abcam)	Human plasma total RNA	
Barberán-Soler, S. et al. Genome Biol. 19, 105 (2018) [25].	Original two-adaptor ligation	NEBNext Multiplex Small RNA Library Prep Set (New England Biolabs)	Human brain total RNA	Single adaptor ligation and circularization protocol detected the the largest spectrum of miRNAs and showed the least bias.
	UMI	QIAseq miRNA library Kit (Qiagen)	miRXPlore universal reference	
	Randomized adaptors	NEXTflex (Bioo Scientific)		
	Polyadenylation and template switching Single adaptor ligation and circularization	SMARTer smRNA-seq Kit (Takara Bio) Beta version of RealSeq-AC Kit (Somagenics)		

Table 2. Cont.

Benchmarking Study	Included Approaches	Commercial Protocols Included	Sample Types	Conclusion
Godoy, P. M. et al. Cell Rep. 29, 4212–4222.e5 (2019) [75].	Randomized adaptors	TruSeq Small RNA Library Prep Kit (Illumina) + in house modifications	Equimolar and ratio metric pool of synthetic RNAs	Small RNA-seq showed better specificity, ability to detect expected differential expression, but higher level of bias.
	Sequencing of hybridization probes Hybridization based techniques	EdgeSeq (HTG Molecular Diagnostics) FirePlex (Abcam) + nCounter (NanoString)	Human plasma total RNA	
Wong, R. K. Y. et al. BMC Genomics 20, 1–12 (2019) [36].	Original two-adaptor ligation	CleanTag Small RNA Library Prep Kit (TriLink BioTechnologies)	Human plasma total RNA	Protocol utilizing UMIs detected the highest number of miRNAs and correlated most closely to RT-qPCR validation data.
	Randomized adaptors UMI	NEXTflex (Bioo Scientific) QIAseq miRNA library Kit (Qiagen)		
Wright, C. et al. BMC Genomics 20, 513 (2019) [60].	Original two-adaptor ligation	TruSeq Small RNA Library Prep Kit (Illumina)	Human brain total RNA	Protocol utilizing randomized adaptors performed the best. Authors suggested usage of random nucleotides as UMIs to reduce PCR bias.
	Randomized adaptors Polyadenylation and template switching	NEXTflex (Bioo Scientific) SMARTer (Clontech)	miRXplore universal reference	
Heinicke, F. et al. RNA Biol. 17, 75–86 (2020) [66].	Original two-adaptor ligation	Small RNA-Seq Library Prep Kit (Lexogen) + TailorMix miRNA Sample Preparation Kit (SeqMatic)	Equimolar and ratio metric pool of synthetic RNAs	Best performance was shown by protocol utilizing UMIs. Polyadenylation and circularization protocols showed a poor yield of miRNAs reads and were not considered for further analysis.
	Single adaptor ligation and circularization Polyadenylation and template switching UMI	CleanTag Small RNA Library Prep Kit (TriLink BioTechnologies) SMARTer microRNA-seq Kit (Takara Bio) CATS (Diagenode) QIAseq miRNA library Kit (Qiagen)	Human CD8+ T cells total RNA	
Herbert, Z. T. et al. J. Biomol. Tech. 31, 47–56 (2020) [65].	All available approaches to small RNA-seq	Small RNA-Seq Library Prep Kit (Lexogen) + TruSeq Small RNA Library Prep Kit (Illumina)	miRXplore universal reference	All methods showed high reproducibility and there was no protocol outperforming others across all the metrics.
		NEBNext Multiplex Small RNA Library Prep Set (New England Biolabs) CleanTag Small RNA Library Prep Kit (TriLink BioTechnologies) NEXTflex (PerkinElmer) SMARTer smRNA-seq Kit + CATS (Diagenode) RealSeq-AC Kit (Somagenics) QIAseq miRNA library Kit (Qiagen) Hybridization-based method	Mesenchymal stem cells total RNA nCounter (NanoString)	

Table 2. Cont.

Benchmarking Study	Included Approaches	Commercial Protocols Included	Sample Types	Conclusion
Baldrich, P. et al. bioRxiv (2020) [79]	Original two-adaptor ligation	NEBNext Multiplex Small RNA Library Prep Set (New England Biolabs)	Maize anthers total RNA	Protocols using randomized adaptors and two-adaptor ligation generated highest number of reads mapping to miRNAs and phasiRNAs.
		TruSeq Small RNA Library Prep Kit (Illumina)		
		CleanTag Small RNA Library Prep Kit (TriLink BioTechnologies)		
	Randomized adaptors	NEXTflex v2, NEXTflex v3 (Bioo Scientific)		
	Single adaptor ligation and circularization	RealSeq-AC Kit (Somagenics)		
	Polyadenylation and template switching	SMARTer smRNA-seq Kit (Takara Bio)		
Androvic, P. et al. bioRxiv (2021) [61].	All available approaches to small RNA-seq	Small RNA-Seq Library Prep Kit (Lexogen) + Small RNA Library Prep Kit (Norgen)	Human plasma total RNA	Protocols utilizing randomized adaptors and UMIs had the best overall performance. Hybridization-based methods showed the highest sensitivity and low specificity.
		QIaseq miRNA library Kit (Qiagen)	miRXplore universal reference	
		NEXTflex (Bioo Scientific)		
		SMARTer smRNA-seq Kit (Takara Bio)		
	RealSeq-Biofluids Kit (Somagenics)			
	Hybridization-based method	EdgeSeq (HTG Molecular Diagnostics)		

4. Considerations and Future Prospects

In the majority of comparative studies, none of the protocols outperformed others across all tested metrics. The results rather documented their variable performance and highlighted their different applications in miRNA research. When selecting a small RNA-seq protocol for the new study, the choice of the protocol should be therefore considered with respect to target-, sample-, and protocol-specific features, requirements on the analysis and expected outcomes. Here, we provide a simplified decision diagram for protocol selection (Figure 4) and summarize the basic aspects that need to be reflected.

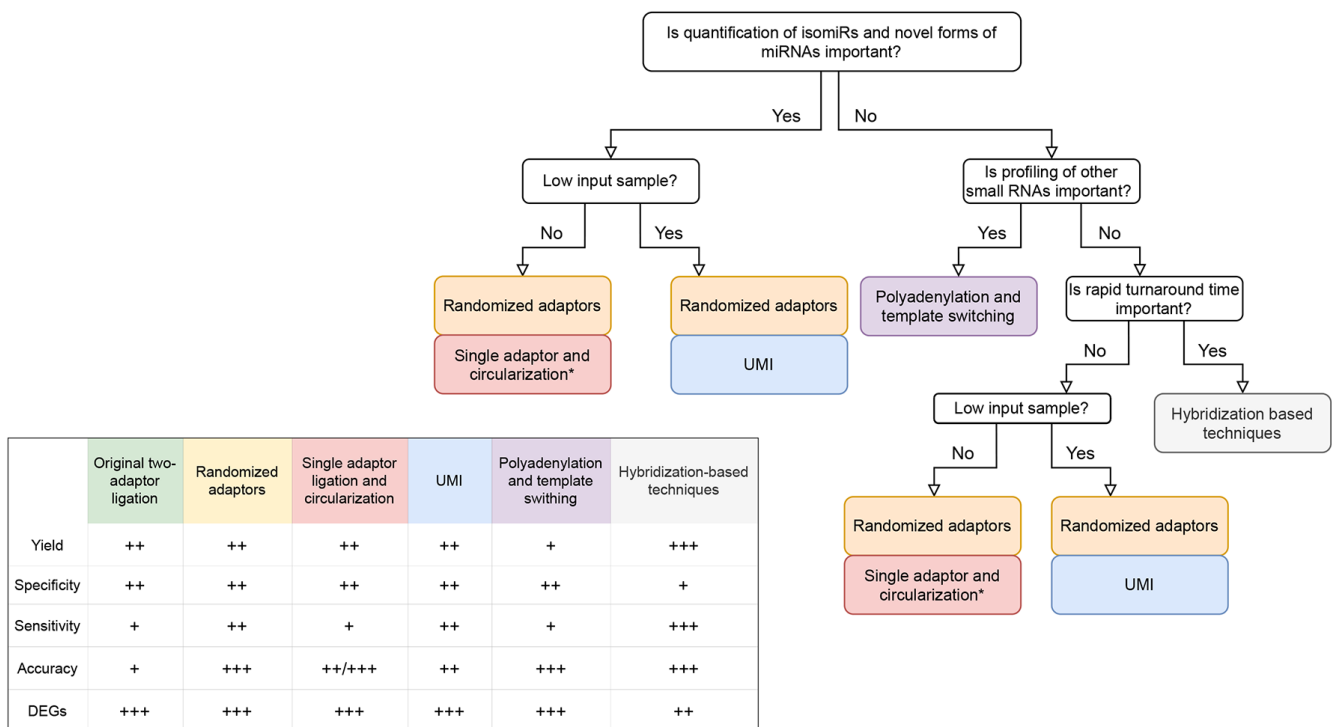


Figure 4. Selection of platform for small RNA-seq analysis. Decision diagram guides an informed choice of suitable protocol. Of note, other non-depicted experiment-specific criteria need to be further considered. Performance metrics represent a consensus of recent benchmarking studies listed in Table 2. +; poor, ++; average, +++; good, DEGs; Differential expression, *; Somagenics RealSeq-AC Kit.

(i) First, the choice of a protocol is given by the molecule of interest. miRNAs are enriched in the majority of ligation-based protocols, whereas other small RNAs along with miRNAs are captured by polyadenylation methods. As polyadenylation suffers from a low mapping rate and inability to distinguish isomiRs, protocol adjustments (specialized enzymes, ligation of cDNA) might be applied in the ligation-based protocols to capture also other small RNA classes [49,80].

(ii) Sample input requirement is another protocol-specific aspect for consideration. Although the majority of current methods supports low-input applications, including analysis of liquid biopsies, their performance varies broadly [59,61,75,76,78]. Using a plasma sample, protocols using randomized adaptors [59], alternatively UMIs (QIAspec specifically) [36] showed good performance. For routine analysis of crude biofluids sample, EdgeSeq platform outperformed other competitors [76]. More details on miRNA analysis in low-input samples might be found elsewhere [21,42].

(iii) After considering target- and sample-related factors, the accuracy of measurement is a key aspect in small RNA-seq analysis. As discussed, ligation and PCR biases might significantly confound the analysis and negatively influence the overall performance. To mitigate the bias while preserving good performance, hybridization-based approaches

(EdgeSeq specifically) or randomized adaptors-based approaches are the methods of choice [61]. On rich RNA samples, the superior performance was documented for the circularization approach [25,65].

(iv) A common goal of small RNA-seq studies, is finding differential abundant miRNAs, the reproducibility of measurement needs to be considered as well. From this perspective, all available small RNA-seq methods demonstrated low within-protocol variability, allowing for precise differential expression analysis [75,76]. Although reproducible, one should keep in mind that in many cases the data suffered from strong protocol-specific bias, which limits any inter-protocol comparison.

(v) Lastly, the practical aspects of protocol selection are represented by the level of multiplexing, the complexity of library preparation or costs of measurement. The majority of protocols offers 96 indexes allowing the pooling of the same number of samples and their analysis in a single sequencing run. However, some protocols still offer a limiting level of multiplexing (48 or 24), which requests multiple sequencing runs and increases total costs in the case of analysis of large sample cohorts. From the view of the complexity of library preparation, EdgeSeq automated platform stands out over competitors, making the process of library preparation easy, fast and robust. This is however accompanied by expensive instrumentation and a higher cost of reagents.

In addition to existing protocols and platforms, there is ongoing research introducing novel improvements to small RNA-seq. When incorporated into current protocols, these advances possess a good promise to further increase the performance and applicability of small RNA-seq analysis. For example, the combination of optimized ligation temperature, the addition of polyethylene glycol (PEG) and 12bp-long randomized adaptors usable as UMIs have been demonstrated to achieve higher performance than commercially available NEXTflex [63]. Similarly, the use of 20% PEG during the ligation steps was shown to increase the sensitivity of the two-adaptor ligation protocol [81]. Another possible improvement is to use adaptors with randomized overhangs, which binds to the complementary ends of miRNAs. The utilization of this approach was shown to have less technical bias and significantly higher sensitivity than NEXTflex [82,83]. Moreover, it enables for profiling of small RNA species with 2'-O-methyl on 3'-end and therefore might be used for profiling of plant miRNAs [80]. Appealing is also the introduction of a higher level of multiplexing [84] that would allow for the analysis of large sample cohorts, employing the power and cost-effectiveness of modern sequencers. Apart from the improvement in wet-lab protocols, there is a possibility for the development of new computational methods for bias correction. The first step in this direction has been done in the study of Baroin-Tourancheau et al. [46], where authors reduced bias using a correction factor learned on the synthetic miRXplore standard. More advanced models, employing large sets of publicly available datasets or benchmarking studies could provide further progress in this field and finally allow such needed cross-study comparison. Lastly, we envision development in the data analysis pipelines. Currently, the miRNA analysis represents a complex process involving a large number of steps and analysis tools, with more than 1000 currently available [85,86]. This requires the assistance of experienced bioinformaticians and complicates the wide application of the technology. The development of user-friendly analysis platforms will help overcome this obstacle, promote standardization of the technology and its transfer to routine clinical practice.

5. Conclusions

Here, we reviewed the current state of small RNA-seq technology for analysis of miRNAs. Undoubtedly, small RNA-seq technology underwent a dramatic development in recent years. The major changes were targeted on the reduction of technical bias and simplification of the analysis workflow. The success rate of these efforts varies which is reflected by the diverse performance of small RNA-seq protocols in comparative studies. Overall, none of the protocols are superior to others and experiment-specific factors need to be carefully considered when selecting a protocol for a new study. To guide the selection,

we summarized all major advancements in the field and provided our perspective on their applicability for miRNA profiling. We believe that small RNA-seq holds the promise to become a leading technology for the analysis of miRNAs and once standardized, its applicability will grow even more in popularity.

Author Contributions: Original draft preparation, S.B.; review and editing, L.V. and M.K.; visualization, S.B.; supervision, L.V.; project administration, L.V. and M.K.; funding acquisition, L.V. and M.K. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by Czech Science Foundation P303/18/21942S, Czech Health Research Council NU21-08-00286 and institutional support RVO 86652036.

Acknowledgments: The manuscript was revised for English grammar by Ravindra Naraine. His help is greatly appreciated.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bartel, D.P. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *Cell* **2004**, *116*, 281–297. [[CrossRef](#)]
2. Vasudevan, S.; Tong, Y.; Steitz, J.A. Switching from Repression to Activation: microRNAs Can Up-Regulate Translation. *Science* **2007**, *318*, 1931–1934. [[CrossRef](#)] [[PubMed](#)]
3. Vasudevan, S. Posttranscriptional Upregulation by MicroRNAs. *Wiley Interdiscip. Rev. RNA* **2012**, *3*, 311–330. [[CrossRef](#)]
4. O'Brien, J.; Hayder, H.; Zayed, Y.; Peng, C. Overview of microRNA biogenesis, mechanisms of actions, and circulation. *Front. Endocrinol.* **2018**, *9*, 1–12. [[CrossRef](#)] [[PubMed](#)]
5. Catalanotto, C.; Cogoni, C.; Zardo, G. MicroRNA in control of gene expression: An overview of nuclear functions. *Int. J. Mol. Sci.* **2016**, *17*, 1721. [[CrossRef](#)] [[PubMed](#)]
6. Kim, V.N.; Nam, J.W. Genomics of microRNA. *Trends Genet.* **2006**, *22*, 165–173. [[CrossRef](#)] [[PubMed](#)]
7. Gebert, L.F.; MacRae, I.J. Regulation of microRNA function in animals. *Nat. Rev. Mol. Cell Biol.* **2019**, *20*, 21–37. [[CrossRef](#)] [[PubMed](#)]
8. Neilsen, C.T.; Goodall, G.J.; Bracken, C.P. IsomiRs—The overlooked repertoire in the dynamic microRNAome. *Trends Genet.* **2012**, *28*, 544–549. [[CrossRef](#)]
9. Kozomara, A.; Birgaoanu, M.; Griffiths-Jones, S. MiRBase: From microRNA sequences to function. *Nucleic Acids Res.* **2019**, *47*, D155–D162. [[CrossRef](#)]
10. Friedman, R.C.; Farh, K.K.H.; Burge, C.B.; Bartel, D.P. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* **2009**, *19*, 92–105. [[CrossRef](#)]
11. Lu, J.; Getz, G.; Miska, E.A.; Alvarez-Saavedra, E.; Lamb, J.; Peck, D.; Sweet-Cordero, A.; Ebert, B.L.; Mak, R.H.; Ferrando, A.A.; et al. MicroRNA expression profiles classify human cancers. *Nature* **2005**, *435*, 834–838. [[CrossRef](#)]
12. Calin, G.A.; Croce, C.M. MicroRNA signatures in human cancers. *Nat. Rev. Cancer* **2006**, *6*, 857–866. [[CrossRef](#)]
13. Ruan, K.; Fang, X.; Ouyang, G. MicroRNAs: Novel regulators in the hallmarks of human cancer. *Cancer Lett.* **2009**, *285*, 116–126. [[CrossRef](#)]
14. Anfossi, S.; Babayan, A.; Pantel, K.; Calin, G.A. Clinical utility of circulating non-coding RNAs—An update. *Nat. Rev. Clin. Oncol.* **2018**, *15*, 541–563. [[CrossRef](#)] [[PubMed](#)]
15. Cortez, M.A.; Bueso-Ramos, C.; Ferdin, J.; Lopez-Berestein, G.; Sood, A.K.; Calin, G.A. MicroRNAs in body fluids—the mix of hormones and biomarkers. *Nat. Rev. Clin. Oncol.* **2011**, *8*, 467–477. [[CrossRef](#)] [[PubMed](#)]
16. Pritchard, C.C.; Cheng, H.H.; Tewari, M. MicroRNA profiling: Approaches and considerations. *Nat. Rev. Genet.* **2012**, *13*, 358–369. [[CrossRef](#)] [[PubMed](#)]
17. Koppers-Lalic, D.; Hackenberg, M.; Menezes, R.D. Non invasive prostate cancer detection by measuring miRNA variants (isomiRs) in urine extracellular vesicles. *Oncotarget* **2016**, *7*, 22566. [[CrossRef](#)]
18. Telonis, A.G.; Loher, P.; Jing, Y.; Londin, E.; Rigoutsos, I. Beyond the one-locus-one-miRNA paradigm: MicroRNA isoforms enable deeper insights into breast cancer heterogeneity. *Nucleic Acids Res.* **2015**, *43*, 9158–9175. [[CrossRef](#)]
19. Telonis, A.G.; Magee, R.; Loher, P.; Chervoneva, I.; Londin, E.; Rigoutsos, I. Knowledge about the presence or absence of miRNA isoforms (isomiRs) can successfully discriminate amongst 32 TCGA cancer types. *Nucleic Acids Res.* **2017**, *45*, 2973–2985. [[CrossRef](#)]
20. Ono, S.; Lam, S.; Nagahara, M.; Hoon, D. Circulating microRNA Biomarkers as Liquid Biopsy for Cancer Patients: Pros and Cons of Current Assays. *J. Clin. Med.* **2015**, *4*, 1890–1907. [[CrossRef](#)]
21. Valihrach, L.; Androvic, P.; Kubista, M. Circulating miRNA analysis for cancer diagnostics and therapy. *Mol. Asp. Med.* **2020**, *72*, 1–19. [[CrossRef](#)] [[PubMed](#)]
22. Shore, S.; Henderson, J.M.; Lebedev, A.; Salcedo, M.P.; Zon, G.; McCaffrey, A.P.; Paul, N.; Hogrefe, R.I. Small RNA library preparation method for next-generation sequencing using chemical modifications to prevent adapter dimer formation. *PLoS ONE* **2016**, *11*, e0167009. [[CrossRef](#)] [[PubMed](#)]

23. Pease, J. Small-RNA sequencing libraries with greatly reduced adaptor-dimer background. *Nat. Methods* **2011**, *8*, iii–iv. [[CrossRef](#)]
24. Baran-Gale, J.; Lisa Kurtz, C.; Erdos, M.R.; Sison, C.; Young, A.; Fannin, E.E.; Chines, P.S.; Sethupathy, P. Addressing bias in small RNA library preparation for sequencing: A new protocol recovers microRNAs that evade capture by current methods. *Front. Genet.* **2015**, *6*, 1–9. [[CrossRef](#)] [[PubMed](#)]
25. Barberán-Soler, S.; Vo, J.M.; Hogans, R.E.; Dallas, A.; Johnston, B.H.; Kazakov, S.A. Decreasing miRNA sequencing bias using a single adapter and circularization approach. *Genome Biol.* **2018**, *19*, 105. [[CrossRef](#)]
26. Dennis, D.; Rhodes, M.; Maclean, K. *Targeted miRNA Discovery and Validation—Using the nCounter® Platform*; NanoString Technologies—WHITE PAPER—nCounter PanCancer Immune Profiling Panel; NanoString Technologies, Inc.: Seattle, WA, USA, 2015; pp. 1–8.
27. Wang, Z.; Gerstein, M.; Snyder, M. RNA-Seq: A revolutionary tool for transcriptomics in Western Equatoria State. *Nat. Rev. Genet.* **2009**, *10*, 57. [[CrossRef](#)]
28. Lu, C.; Tej, S.S.; Luo, S.; Haudenschild, C.D.; Meyers, B.C.; Green, P.J. Genetics: Elucidation of the small RNA component of the transcriptome. *Science* **2005**, *309*, 1567–1569. [[CrossRef](#)]
29. Ruby, J.G.; Jan, C.; Player, C.; Axtell, M.J.; Lee, W.; Nusbaum, C.; Ge, H.; Bartel, D.P. Large-Scale Sequencing Reveals 21U-RNAs and Additional MicroRNAs and Endogenous siRNAs in *C. elegans*. *Cell* **2006**, *127*, 1193–1207. [[CrossRef](#)]
30. Edgar, R.; Domrachev, M.; Lash, A. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* **2002**, *30*, 207–210. [[CrossRef](#)]
31. Chakraborty, C.; Bhattacharya, M.; Agoramoorthy, G. Single-cell sequencing of miRNAs: A modified technology. *Cell Biol. Int.* **2020**, *44*, 1773–1780. [[CrossRef](#)]
32. Hafner, M.; Landgraf, P.; Ludwig, J.; Rice, A.; Ojo, T.; Lin, C.; Holoch, D.; Lim, C.; Tuschl, T. Identification of microRNAs and other small regulatory RNAs using cDNA library sequencing. *Methods* **2008**, *44*, 3–12. [[CrossRef](#)]
33. Redshaw, N.; Wilkes, T.; Whale, A.; Cowen, S.; Huggett, J.; Foy, C.A. A comparison of miRNA isolation and RT-qPCR technologies and their effects on quantification accuracy and repeatability. *BioTechniques* **2013**, *54*, 155–164. [[CrossRef](#)]
34. Guo, Y.; Vickers, K.; Xiong, Y.; Zhao, S.; Sheng, Q.; Zhang, P.; Zhou, W.; Flynn, C.R. Comprehensive evaluation of extracellular small RNA isolation methods from serum in high throughput sequencing. *BMC Genom.* **2017**, *18*, 1–9. [[CrossRef](#)]
35. Srinivasan, S.; Yeri, A.; Cheah, P.S.; Chung, A.; Dehoff, P.; Filant, J.; Laurent, C.D.; Laurent, L.D.; Magee, R.; Moeller, C.; et al. Small RNA sequencing across diverse biofluids identifies optimal methods for exRNA isolation. *Cell* **2019**, *177*, 446–462. [[CrossRef](#)] [[PubMed](#)]
36. Wong, R.K.; MacMahon, M.; Woodside, J.V.; Simpson, D.A. A comparison of RNA extraction and sequencing protocols for detection of small RNAs in plasma. *BMC Genom.* **2019**, *20*, 1–12. [[CrossRef](#)]
37. Raabe, C.A.; Tang, T.H.; Brosius, J.; Rozhdestvensky, T.S. Biases in small RNA deep sequencing data. *Nucleic Acids Res.* **2014**, *42*, 1414–1426. [[CrossRef](#)]
38. Fuchs, R.T.; Sun, Z.; Zhuang, F.; Robb, G.B. Bias in ligation-based small RNA sequencing library construction is determined by adaptor and RNA structure. *PLoS ONE* **2015**, *10*, e0126049. [[CrossRef](#)] [[PubMed](#)]
39. Kalle, E.; Kubista, M.; Rensing, C. Multi-template polymerase chain reaction. *Biomol. Detect. Quantif.* **2014**, *2*, 11–29. [[CrossRef](#)] [[PubMed](#)]
40. Kivioja, T.; Vähärautio, A.; Karlsson, K.; Bonke, M.; Enge, M.; Linnarsson, S.; Taipale, J. Counting absolute numbers of molecules using unique molecular identifiers. *Nat. Methods* **2012**, *9*, 72–74. [[CrossRef](#)] [[PubMed](#)]
41. Fu, Y.; Wu, P.H.; Beane, T.; Zamore, P.D.; Weng, Z. Elimination of PCR duplicates in RNA-seq and small RNA-seq using unique molecular identifiers. *BMC Genom.* **2018**, *19*, 531. [[CrossRef](#)] [[PubMed](#)]
42. Buschmann, D.; Haberberger, A.; Kirchner, B.; Spornraft, M.; Riedmaier, I.; Schelling, G.; Pfaffl, M.W. Toward reliable biomarker signatures in the age of liquid biopsies—How to standardize the small RNA-Seq workflow. *Nucleic Acids Res.* **2016**, *44*, 5995–6018. [[CrossRef](#)] [[PubMed](#)]
43. Hafner, M.; Renwick, N.; Brown, M.; Mihailović, A.; Holoch, D.; Lin, C.; Pena, J.T.; Nusbaum, J.D.; Morozov, P.; Ludwig, J.; et al. RNA-ligase-dependent biases in miRNA representation in deep-sequenced small RNA cDNA libraries. *RNA* **2011**, *17*, 1697–1712. [[CrossRef](#)] [[PubMed](#)]
44. Sorefan, K.; Pais, H.; Hall, A.E.; Kozomara, A.; Griffiths-Jones, S.; Moulton, V.; Dalmay, T. Reducing ligation bias of small RNAs in libraries for next generation sequencing. *Silence* **2012**, *3*, 1–11. [[CrossRef](#)]
45. Zhuang, F.; Fuchs, R.T.; Robb, G.B. Small RNA expression profiling by high-throughput sequencing: Implications of enzymatic manipulation. *J. Nucleic Acids* **2012**, *2012*, 360358. [[CrossRef](#)]
46. Baroin-Tourancheau, A.; Jaszczyszyn, Y.; Benigni, X.; Amar, L. Evaluating and Correcting Inherent Bias of microRNA Expression in Illumina Sequencing Analysis. *Front. Mol. Biosci.* **2019**, *6*, 17. [[CrossRef](#)]
47. Xu, P.; Billmeier, M.; Mohorianu, I.; Green, D.; Fraser, W.D.; Dalmay, T. An improved protocol for small RNA library construction using High Definition adapters. *Methods Next Gener. Seq.* **2015**, *2*, 1–10. [[CrossRef](#)]
48. Pak, J.; Fire, A. Distinct Populations of Primary and Secondary Effectors During RNAi in *C. elegans*. *Science* **2007**, *315*, 241–244. [[CrossRef](#)] [[PubMed](#)]
49. Kugelberg, U.; Nätt, D.; Skog, S.; Kutter, C.; Öst, A. 5'XP sRNA-seq: Efficient identification of transcripts with and without 5' phosphorylation reveals evolutionary conserved small RNA. *RNA Biol.* **2021**, 1–12. [[CrossRef](#)]

50. Giraldez, M.D.; Spengler, R.M.; Etheridge, A.; Goicochea, A.J.; Tuck, M.; Choi, S.W.; Galas, D.J.; Tewari, M. Phospho-RNA-seq: A modified small RNA-seq method that reveals circulating mRNA and lncRNA fragments as potential biomarkers in human plasma. *EMBO J.* **2019**, *38*, 1–14. [[CrossRef](#)] [[PubMed](#)]
51. Munafó, D.B.; Robb, G.B. Optimization of enzymatic reaction conditions for generating representative pools of cDNA from small RNA. *RNA* **2010**, *16*, 2537–2552. [[CrossRef](#)]
52. Head, S.R.; Kiyomi Komori, H.; LaMere, S.A.; Whisenant, T.; Van Nieuwerburgh, F.; Salomon, D.R.; Ordoukhanian, P. Library construction for next-generation sequencing: Overviews and challenges. *BioTechniques* **2014**, *56*, 61–77. [[CrossRef](#)] [[PubMed](#)]
53. Vigneault, F.; Ter-Ovanesyan, D.; Alon, S.; Eminaga, S.; Christodoulou, D.C.; Seidman, J.G.; Eisenberg, E.; Church, G.M. High-throughput multiplex sequencing of miRNA. *Curr. Protoc. Hum. Genet.* **2012**, *73*, 1–10. [[CrossRef](#)]
54. Kawano, M.; Kawazu, C.; Lizio, M.; Kawaji, H.; Carninci, P.; Suzuki, H.; Hayashizaki, Y. Reduction of non-insert sequence reads by dimer eliminator LNA oligonucleotide for small RNA deep sequencing. *BioTechniques* **2010**, *49*, 751–754. [[CrossRef](#)]
55. Hardigan, A.A.; Roberts, B.S.; Moore, D.E.; Ramaker, R.C.; Jones, A.L.; Myers, R.M. CRISPR/Cas9-targeted removal of unwanted sequences from small-RNA sequencing libraries. *Nucleic Acids Res.* **2019**, *47*, e84. [[CrossRef](#)] [[PubMed](#)]
56. Wickersheim, M.L.; Blumenstiel, J.P. Terminator oligo blocking efficiently eliminates rRNA from Drosophila small RNA sequencing libraries. *Biotechniques* **2013**, *55*, 269–272. [[CrossRef](#)] [[PubMed](#)]
57. Roberts, B.S.; Hardigan, A.A.; Kirby, M.K.; Fitz-Gerald, M.B.; Wilcox, C.M.; Kimberly, R.P.; Myers, R.M. Blocking of targeted microRNAs from next-generation sequencing libraries. *Nucleic Acids Res.* **2015**, *43*, 1–8. [[CrossRef](#)] [[PubMed](#)]
58. Dard-Dascot, C.; Naquin, D.; D'Aubenton-Carafa, Y.; Alix, K.; Thermes, C.; van Dijk, E. Systematic comparison of small RNA library preparation protocols for next-generation sequencing. *BMC Genom.* **2018**, *19*, 1–16. [[CrossRef](#)] [[PubMed](#)]
59. Giraldez, M.D.; Spengler, R.M.; Etheridge, A.; Godoy, P.M.; Barczak, A.J.; Srinivasan, S.; De Hoff, P.L.; Tanriverdi, K.; Courtright, A.; Lu, S.; et al. Comprehensive multi-center assessment of small RNA-seq methods for quantitative miRNA profiling. *Nat. Biotechnol.* **2018**, *36*, 746–757. [[CrossRef](#)] [[PubMed](#)]
60. Wright, C.; Rajpurohit, A.; Burke, E.E.; Williams, C.; Collado-Torres, L.; Kimos, M.; Brandon, N.J.; Cross, A.J.; Jaffe, A.E.; Weinberger, D.R.; et al. Comprehensive assessment of multiple biases in small RNA sequencing reveals significant differences in the performance of widely used methods. *BMC Genom.* **2019**, *20*, 513. [[CrossRef](#)]
61. Androvic, P.; Benesova, S.; Rohlova, E.; Kubista, M.; Valihrach, L. Small RNA-sequencing for Analysis of Circulating miRNAs: Benchmark Study. *bioRxiv* **2021**. [[CrossRef](#)]
62. Zhuang, F.; Fuchs, R.T.; Sun, Z.; Zheng, Y.; Robb, G.B. Structural bias in T4 RNA ligase-mediated 3'-adapter ligation. *Nucleic Acids Res.* **2012**, *40*, e54. [[CrossRef](#)] [[PubMed](#)]
63. Belair, C.D.; Hu, T.; Chu, B.; Freimer, J.W.; Cooperberg, M.R.; Blleloch, R.H. High-throughput, Efficient, and Unbiased Capture of Small RNAs from Low-input Samples for Sequencing. *Sci. Rep.* **2019**, *9*, 5–12. [[CrossRef](#)] [[PubMed](#)]
64. Saunders, K.; Bert, A.G.; Dredge, B.K.; Toubia, J.; Gregory, P.A.; Pillman, K.A.; Goodall, G.J.; Bracken, C.P. Insufficiently complex unique-molecular identifiers (UMIs) distort small RNA sequencing. *Sci. Rep.* **2020**, *10*, 1–9. [[CrossRef](#)] [[PubMed](#)]
65. Herbert, Z.T.; Thimmapuram, J.; Xie, S.; Kershner, J.P.; Kolling, F.W.; Ringelberg, C.S.; Leclerc, A.; Alekseyev, Y.O.; Fan, J.; Podnar, J.W.; et al. Multisite evaluation of next-generation methods for small RNA quantification. *J. Biomol. Tech.* **2020**, *31*, 47–56. [[CrossRef](#)]
66. Heinicke, F.; Zhong, X.; Zucknick, M.; Breidenbach, J.; Sundaram, A.Y.; Flåm, S.T.; Leithaug, M.; Dalland, M.; Farmer, A.; Henderson, J.M.; et al. Systematic assessment of commercially available low-input miRNA library preparation kits. *RNA Biol.* **2020**, *17*, 75–86. [[CrossRef](#)] [[PubMed](#)]
67. Jayaprakash, A.D.; Jabado, O.; Brown, B.D.; Sachidanandam, R. Identification and remediation of biases in the activity of RNA ligases in small-RNA deep sequencing. *Nucleic Acids Res.* **2011**, *39*, 1–12. [[CrossRef](#)] [[PubMed](#)]
68. Lipps, C.; Northe, P.; Figueiredo, R.; Rohde, M.; Brahmmer, A.; Krämer-Albers, E.M.; Liebetrau, C.; Wiedenroth, C.B.; Mayer, E.; Kriechbaum, S.D.; et al. Non-invasive approach for evaluation of pulmonary hypertension using extracellular vesicle-associated small non-coding RNA. *Biomolecules* **2019**, *9*, 666. [[CrossRef](#)]
69. Berezikov, E.; Van Tetering, G.; Verheul, M.; Van De Belt, J.; Van Laake, L.; Vos, J.; Verloop, R.; Van De Wetering, M.; Guryev, V.; Takada, S.; et al. Many novel mammalian microRNA candidates identified by extensive cloning and RAKE analysis. *Genome Res.* **2006**, *16*, 1289–1298. [[CrossRef](#)] [[PubMed](#)]
70. Zhu, Y.Y.; Machleder, E.M.; Chenchik, A.; Li, R.; Siebert, P.D. Reverse transcriptase template switching: A SMART™ approach for full-length cDNA library construction. *BioTechniques* **2001**, *30*, 892–897. [[CrossRef](#)]
71. Geiss, G.K.; Bumgarner, R.E.; Birditt, B.; Dahl, T.; Dowidar, N.; Dunaway, D.L.; Fell, H.P.; Ferree, S.; George, R.D.; Grogan, T.; et al. Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat. Biotechnol.* **2008**, *26*, 317–325. [[CrossRef](#)]
72. Chapin, S.C.; Appleyard, D.C.; Pregibon, D.C.; Doyle, P.S. Rapid microRNA profiling on encoded gel microparticles. *Angew. Chem. Int. Ed.* **2011**, *50*, 2289–2293. [[CrossRef](#)]
73. El-Khoury, V.; Pierson, S.; Kaoma, T.; Bernardin, F.; Berchem, G. Assessing cellular and circulating miRNA recovery: The impact of the RNA isolation method and the quantity of input material. *Sci. Rep.* **2016**, *6*, 1–14. [[CrossRef](#)]
74. Girard, L.; Rodriguez-Canales, J.; Behrens, C.; Thompson, D.M.; Botros, I.W.; Tang, H.; Xie, Y.; Rekhtman, N.; Travis, W.D.; Wistuba, I.I.; et al. An Expression Signature as an Aid to the Histologic Classification of Non-Small Cell Lung Cancer. *Clin. Cancer Res.* **2016**, *22*, 4880–4889. [[CrossRef](#)]

75. Godoy, P.M.; Barczak, A.J.; DeHoff, P.; Srinivasan, S.; Etheridge, A.; Galas, D.; Das, S.; Erle, D.J.; Laurent, L.C. Comparison of Reproducibility, Accuracy, Sensitivity, and Specificity of miRNA Quantification Platforms. *Cell Rep.* **2019**, *29*, 4212–4222.e5. [[CrossRef](#)] [[PubMed](#)]
76. Yeri, A.; Courtright, A.; Danielson, K.; Hutchins, E.; Alsop, E.; Carlson, E.; Hsieh, M.; Ziegler, O.; Das, A.; Shah, R.V.; et al. Evaluation of commercially available small RNAseq library preparation kits using low input RNA. *BMC Genom.* **2018**, *19*, 1–15. [[CrossRef](#)] [[PubMed](#)]
77. Das, S.; Abdel-Mageed, A.B.; Adamidi, C.; Adelson, P.D.; Akat, K.M.; Alsop, E.; Ansel, K.M.; Arango, J.; Aronin, N.; Avsaroglu, S.K.; et al. The Extracellular RNA Communication Consortium: Establishing Foundational Knowledge and Technologies for Extracellular RNA Research. *Cell* **2019**, *177*, 231–242. [[CrossRef](#)] [[PubMed](#)]
78. Coenen-Stass, A.M.L.; Magen, I.; Brooks, T.; Ben-Dov, I.Z.; Greensmith, L.; Hornstein, E.; Fratta, P. Evaluation of methodologies for microRNA biomarker detection by next generation sequencing. *RNA Biol.* **2018**, *15*, 1133–1145. [[CrossRef](#)] [[PubMed](#)]
79. Baldrich, P.; Tamim, S.; Mathioni, S.; Meyers, B. Ligation bias is a major contributor to nonstoichiometric abundances of secondary siRNAs and impacts analyses of microRNAs. *bioRxiv* **2020**. [[CrossRef](#)]
80. Maguire, S.; Lohman, G.J.; Guan, S. A low-bias and sensitive small RNA library preparation method using randomized splint ligation. *Nucleic Acids Res.* **2020**, *48*, 1–14. [[CrossRef](#)] [[PubMed](#)]
81. Kim, H.; Kim, J.; Kim, K.; Chang, H.; You, K.V.; Kim, N. Bias-minimized quantification of microRNA reveals widespread alternative processing and 3 end modification. *Nucleic Acids Res.* **2019**, *47*, 2630–2640. [[CrossRef](#)]
82. Raine, A.; Manlig, E.; Wahlberg, P.; Syvänen, A.C.; Nordlund, J. SPLINTED: Ligation Adapter Tagging (SPLAT), a novel library preparation method for whole genome bisulphite sequencing. *Nucleic Acids Res.* **2017**, *45*. [[CrossRef](#)] [[PubMed](#)]
83. Wu, J.; Dai, W.; Wu, L.; Wang, J. SALP, a new single-stranded DNA library preparation method especially useful for the high-throughput characterization of chromatin openness states. *BMC Genom.* **2018**, *19*, 1–12. [[CrossRef](#)] [[PubMed](#)]
84. Persson, H.; Søkilde, R.; Pirona, A.C.; Rovira, C. Preparation of highly multiplexed small RNA sequencing libraries. *BioTechniques* **2017**, *63*, 57–64. [[CrossRef](#)] [[PubMed](#)]
85. Chen, L.; Heikkinen, L.; Wang, C.L.; Yang, Y.; Knott, K.E.; Wong, G. MiRToolsGallery: A tag-based and rankable microRNA bioinformatics resources database portal. *Database* **2018**, *2018*, 1–10. [[CrossRef](#)] [[PubMed](#)]
86. Chen, L.; Heikkinen, L.; Wang, C.; Yang, Y.; Sun, H.; Wong, G. Trends in the development of miRNA bioinformatics tools. *Brief. Bioinform.* **2019**, *20*, 1836–1852. [[CrossRef](#)] [[PubMed](#)]