Chapter 26

Analysis of Small RNA Sequencing Data in Plants

Vanika Garg and Rajeev K. Varshney

Abstract

Over the past decades, next-generation sequencing (NGS) has been employed extensively for investigating the regulatory mechanisms of small RNAs. Several bioinformatics tools are available for aiding biologists to extract meaningful information from enormous amounts of data generated by NGS platforms. This chapter describes a detailed methodology for analyzing small RNA sequencing data using different open source tools. We elaborate on various steps involved in analysis, from processing the raw sequencing reads to identifying miRNAs, their targets, and differential expression studies.

Key words Small RNA sequencing, miRNAs, Bioinformatics, Differential expression, miRDeep-P

1 Introduction

Small RNAs (sRNA) are typically 18–34 nucleotides (nts) long non-coding molecules known to play a pivotal role in posttranscriptional gene expression regulation. Next-generation sequencing (NGS) is a powerful tool used to identify sRNAs in many plant species. NGS has several advantages over microarray techniques as it allows the discovery of novel sRNAs and has a better signal to noise ratio than microarrays. The reduced cost of sequencing and the availability of reagent kits from commercial companies have made the preparation and sequencing of libraries from sRNA fragments a routine process. However, NGS platforms generate an enormous amount of data whose management is quite a tedious task. The analysis of this data requires a lot of computational and statistical knowledge to obtain meaningful information and address complex biological questions.

This chapter introduces the researchers to different aspects of the small RNA sequencing (sRNA-seq) data analysis. Bioinformatics analysis of sRNA-seq data differs from standard RNA-seq protocols (Fig. 1). The sRNA-seq data analysis begins with filtration of low-quality data, removal of adapter sequences, followed by mapping of filtered data onto the ribosomal RNA (rRNA), transfer

David Edwards (ed.), *Plant Bioinformatics: Methods and Protocols*,
Methods in Molecular Biology, vol. 2443, https://doi.org/10.1007/978-1-0716-2067-0_26, © The Author(s) 2022

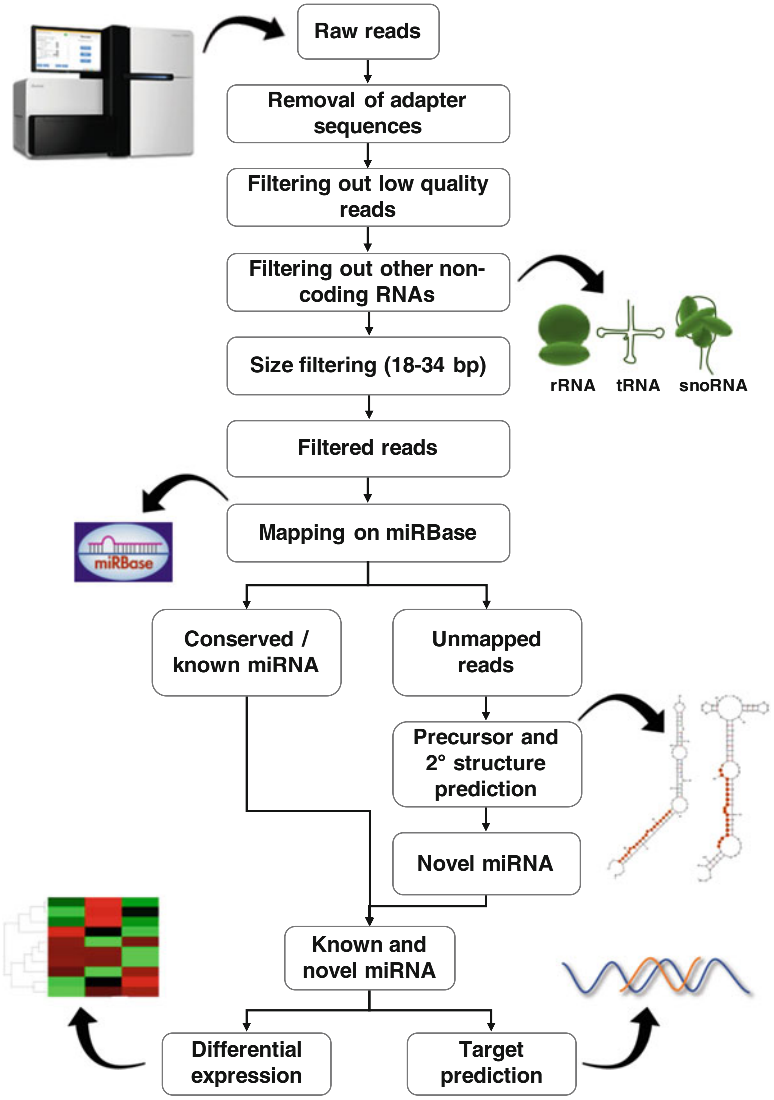


Fig. 1 Schematic outline depicting the different steps involved in analyzing the small RNA sequencing data

RNA (tRNA), small nuclear RNA (snRNA), and small nucleolar RNA (snoRNA) using short read aligners. The reads mapping to these RNAs are discarded as they are not the products of dicer-like (DCL) protein activity and are thought to have very little likelihood of being involved in small RNA pathways. The filtered reads are further mapped to miRBase to identify the conserved miRNAs, and the unmapped reads are processed to identify novel miRNAs using various tools such as miRDeep-P [1], ShortStack [2], miRPlant [3], MIReNA [4], and miRkwood [5]. These tools focus on miRNA prediction by considering the essential features of miRNAs like length, structural features, DCL cleavage, and their high conservation among species.

2 Materials

2.1 Workstation

The analysis presented here requires, for best results, a 64-bit version of the Linux operating system with at least 4 Gb RAM.

2.2 Small RNA Sequencing Data

Raw reads from a small RNA sequencing experiment in Fastq format. Publicly available data can be downloaded from the Sequence Read Archive (SRA) under the National Centre for Biotechnology Information (NCBI; https://www.ncbi.nlm.nih.gov/sra). For this demonstration, we have selected an experiment on chickpea [6]. We selected four samples (SRR12847935, SRR12847937, SRR12847941, and SRR12847943) comprising chickpea genotypes resistant and susceptible to Ascochyta blight grown under control and Ascochyta blight inoculated conditions.

2.3 Reference Genome

The species reference genome sequence in the FASTA format. In this case, we downloaded the chickpea (*C. arietinum*) CDC-Frontier reference genome from NCBI [7].

2.4 Software and Tools

The workflow will require the following tools. The installation instructions of tools can be found on their respective websites.

1. *Trimmomatic v0.39* (http://www.usadellab.org/cms/?page=trimmomatic [8]).
2. *Cutadapt v2.10* (https://cutadapt.readthedocs.io/en/stable/guide.html [9]).
3. *FASTX-Toolkit v0.0.14* (http://hannonlab.cshl.edu/fastx_toolkit/).
4. *Bowtie v1.2.2* (http://bowtie-bio.sourceforge.net/ [10]).
5. *ViennaRNA Package v2.4.15* (http://www.tbi.univie.ac.at/~ivo/RNA/ [11]).
6. *randfold v2.0* (https://github.com/erbon7/randfold),
7. *miRDeep-P v1.3* (https://sourceforge.net/projects/mirdp/ [1]),
8. *R v4.0.3* (https://cran.r-project.org/).
9. *DESeq2 v1.28.1* (https://bioconductor.org/packages/release/bioc/html/DESeq2.html [12]).
10. *SAMtools v1.10-2* (https://github.com/samtools/ [13]).

The workflow demonstrated below applies to single-end sRNA sequencing data generated from the Illumina sequencing platform.

3 Methods

3.1 Data Preprocessing

The raw reads obtained from sRNA-seq are first subjected to quality check, including removal of low-quality reads and adapter sequences (*see* **Note 1**). This step can be performed using trimmomatic with the following command:

```
> java -jar trimmomatic-0.39.jar SE -threads 2 SRR12847935.fq.
gz SRR12847935.cleaned.fq ILLUMINACLIP:illumina.fa:2:30:10 SL
IDINGWINDOW:10:20
```

Here, the file illumina.fa contains the adapter sequences and is provided with trimmomatic. The ILLUMINACLIP option instructs trimmomatic to cut adapters and other Illumina-specific sequences from the reads. The SLIDINGWINDOW option specifies the window size and the Phred-quality score to trim low-quality bases. The number of threads "-threads" can be increased based on the number of cores available on the user's machine to speed up this step. Further, the reads with poly-A tail are trimmed, and reads shorter than 18 nt and longer than 34 nt are discarded using Cutadapt (*see* **Note 2**):

```
> cutadapt -a "A{20}" -m 18 -M 34 -o SRR12847935.cleaned.
polyAtrimmed.fq SRR12847935.cleaned.fq
```

Here, -m and -M specify the minimum and maximum length of sequences to retain after trimming, respectively. The above steps of filtering are performed on all samples, and then the filtered reads obtained from each sample are combined into a single fastq file using the cat command:

```
> cat *.cleaned.polyAtrimmed.fq > combined_reads.fq
```

The combined reads are converted into fasta format and collapsed into unique tags. For collapsing the reads, fastx_collapser from FASTX-Toolkit is used. The unique tag file is further formatted to make it compatible with miRNA prediction software to be used in downstream analysis using sed command.

```
> cat combined_reads.fq | paste - - - - | sed ’s/^@/>/g’ | cut
-f 1,2 | tr ’\t’ ’\n’ > combined_reads.fa
> fastx_collapser -i combined_reads.fa -o unique_tags.fa
> sed -i ’s/-/_x/’ unique_tags.fa
```

The final data preparation step involves the removal of read mapping to ribosomal RNA (rRNA), transfer RNA (tRNA), small nuclear (snRNA), small nucleolar RNA (snoRNA), and repeat sequences. For rRNA, tRNA, snRNA, and snoRNA removal, reads can be aligned using Bowtie to a database of r-, t-, sn-, and sno-RNA sequences (can be downloaded from the Rfam database [14]). Similarly, for eliminating read mapping to repeat regions, the Repbase database (https://www.girinst.org/repbase/) can be used. Before running the alignment step, users need to create an index file for the database using bowtie-build.

```
> bowtie-build rfam.fa rfam
> bowtie rfam -f unique_tags.fa -S unique_tags.ncRNA.sam --un
unique_tags.unaligned.ncRNA.fa
> bowtie-build repeats.fa repeats
> bowtie repeats -f unique_tags.unaligned.ncRNA.fa -S unique_-
tags.unaligned.ncRNA.repeats.sam --un unique_tags.filtered.fa
```

Here, -f and -S specify that the input file is in fasta format, and the output is in sequence alignment map (SAM) format, respectively. In the command, --un specifies the name of the file where the unaligned reads are stored.

3.2 Identification of Known miRNAs

The filtered reads (unique_tags.filtered.fa) from Subheading 3.1 are aligned against the known plant miRNAs from the miRBase database [15] using Bowtie to identify conserved or known miRNAs (*see* **Note 3**).

```
> bowtie-build mirbase.fa mirbase
> bowtie mirbase -n 2 -f unique_tags.filtered.fa -S unique_-
tags.filtered.mirbase.sam --un forNovelPrediction.fa
```

Here, the number of mismatches can be changed using the parameter "-n"; in this example, we have used two mismatches. The list of known miRNAs and their sequences can be extracted by parsing the alignment file using the command:

```
> grep -v "^@" unique_tags.filtered.mirbase.sam | awk -F"\t"
'{if($3!="*") print}' | cut -f 1,3 | sort -u > known_miRNAs.
txt
```

3.3 Identification of Novel miRNAs

The unique reads that do not align against the miRBase database are used for novel miRNA prediction. The prediction will be carried out using the miRDeep-P package in this demonstration. miRDeep-P is a freely available package that includes nine Perl scripts, which are executed sequentially to predict miRNAs based on plant-specific criteria [1]. A detailed manual and example datasets are provided with the package. The reads not aligning to known plant miRNAs are first mapped to the reference genome using Bowtie with either 0 or 1 mismatch for novel miRNA prediction.

```
> bowtie-build chickpea.genome.fa chickpea.genome
> bowtie chickpea.genome -n 0 -f forNovelPrediction.fa -S
forNovelPrediction.genome.sam
```

Next, the bowtie alignments in SAM format are converted to blast format using the script "convert_SAM_to_blast.pl"'provided with miRDeep-P package. The script requires a file with bowtie alignments, unique tags, and the reference genome (in fasta format):

```
> perl miRDeep-P/convert_SAM_to_blast.pl forNovelPrediction.
genome.sam forNovelPrediction.fa chickpea.genome.fa > genome.
bst
```

Further, the alignments are filtered to retain only those with 100% sequence identity, full-length alignment, and the number of matches that do not exceed a user-specified cutoff (−c 15 in this example and can be changed according to the species of interest) using the "filter_alignments.pl" script.

```
> perl miRDeep-P/filter_alignments.pl genome.bst -c 15 >
genome.filter15.bst
```

The reads that overlap with the known annotated features (exons, CDS, etc.) of the species under study are discarded. The corresponding annotations can be obtained from the public databases like NCBI (https://www.ncbi.nlm.nih.gov/) or Phytozome (https://phytozome.jgi.doe.gov/) depending on the species of interest. This step is executed using "overlap.pl" and "alignedselected.pl" scripts:

```
> perl miRDeep-P/overlap.pl genome.filter15.bst Chickpea.
gene.gff -b > genome.filter15.overlap_CDS
> perl miRDeep-P/alignedselected.pl genome.filter15.bst -g
genome.filter15.overlap_CDS > genome.filter15.CDS.bst
```

The following script extracts the fasta sequences of the reads filtered in the previous step.

```
> perl miRDeep-P/filter_alignments.pl genome.filter15.CDS.bst
-b forNovelPrediction.fa > genome.filter15.CDS.filtered.fa
```

Next, we extract the potential precursor sequences from the reference genome using "excise_candidate.pl". The script takes the reference genome (in fasta format) and the filtered alignments. The authors of the miRDeep-P recommend 250 bp as the optimal window size for extracting precursor sequences for both monocot and dicot plants.

```
> perl miRDeep-P/excise_candidate.pl chickpea.genome.fa gen-
ome.filter15.CDS.bst 250 > genome.filter15.CDS.precursors.fa
```

The secondary structures of the potential precursor sequences are then predicted using RNAfold utility from the ViennaRNA package [11]. The users can invoke the --noPS option to avoid the graphical output (*see* **Note 4**).

```
> cat genome.filter15.CDS.precursors.fa | RNAfold --noPS >
genome.filter15.CDS.structures
```

Now, the filtered reads are aligned to potential precursor sequences to generate miRNA signatures. The alignment file generated by mapping the reads onto the precursor sequences is converted into blast format and finally sorted to obtain signatures using the following commands:

```
> bowtie-build -f genome.filter15.CDS.precursors.fa genome.
filter15.CDS.precursors
> bowtie genome.filter15.CDS.precursors -f genome.filter15.
CDS.filtered.fa -S genome.filter15.CDS.precursors.sam
> perl miRDeep-P/convert_SAM_to_blast.pl genome.filter15.CDS.
precursors.sam genome.filter15.CDS.filtered.fa genome.fil-
ter15.CDS.precursors.fa > precursors.bst
> sort +3 -25 precursors.bst > signatures
```

Once the signatures are generated, the "miRDP.pl" script combines this information with structures obtained using RNAfold to make miRNA predictions.

```
> perl miRDeep-P/miRDP.pl signatures genome.filter15.CDS.
structures -y > predictions
```

Finally, the predicted miRNAs are filtered to remove redundant miRNAs and miRNAs that do not meet the criteria of plant miRNAs using the "rm_redundant_meet_plant.pl" script. This script requires the length of each chromosome of the reference genome, precursors, and miRNA predictions. For obtaining chromosome lengths, the faidx utility of samtools is used.

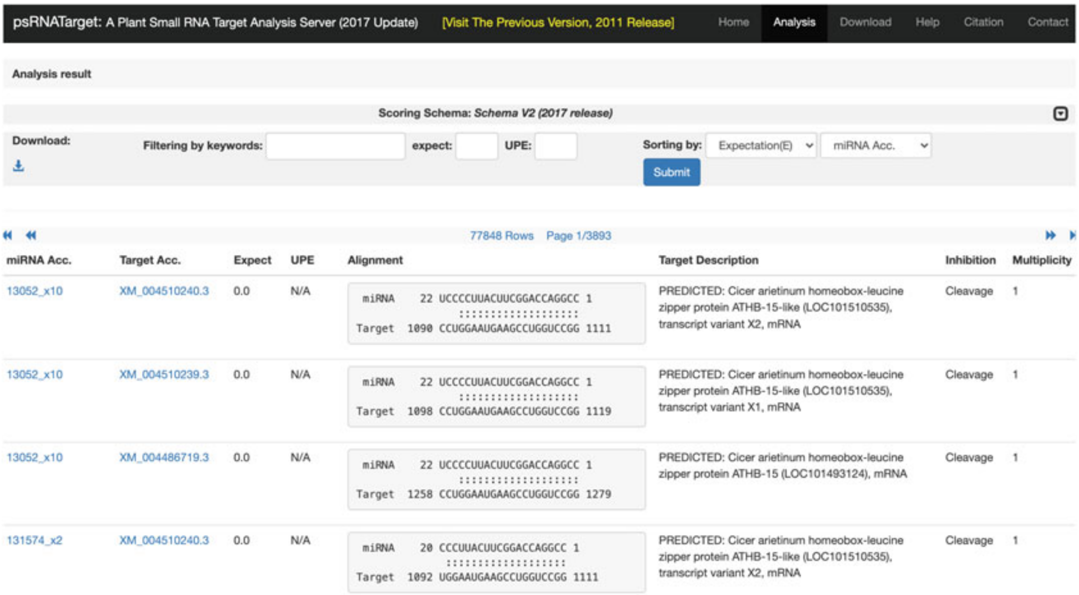
```
> samtools faidx Chickpea.genome.fa
> perl miRDeep-P/rm_redundant_meet_plant.pl Chickpea.genome.
fa.fai genome.filter15.CDS.precursors.fa predictions nr_pre-
dictions novel_plant_miRNAs
```

This step gives us novel miRNAs, which, together with known miRNAs from Subheading 3.2, constitute the final list of miRNAs (all_miRNAs).

3.4 Prediction of miRNA Targets

For understanding the function of miRNAs, it is imperative to predict their targets. In contrast to animal miRNAs, plant miRNAs are known to have perfect or near-perfect complementarity with their targets. Utilizing the complementarity attribute, several tools with varying degrees of specificity and sensitivity are available for miRNA target prediction. Some of the widely used target prediction tools include TAPIR [16], psRobot [17], comTAR [18], and psRNATarget [19]. Besides similarity-based computational tools, the miRNA targets can also be predicted using a highly sensitive and powerful approach called degradome sequencing [20]. The degradome sequencing data offers patterns of RNA degradation and can be analyzed using different computational pipelines such as CleaveLand [21], SeqTar [22], sPARTA [23], and PAREsnip2 [24].

For this demonstration, we use the psRNATarget server (http://plantgrn.noble.org/psRNATarget/analysis), which offers an easy-to-use graphical user interface. It employs an accelerated Smith–Waterman algorithm to find the best sRNA/mRNA complementarity location in the target candidate and to indicate whether the miRNA is involved in cleavage or translational inhibition. For prediction, upload the small RNA sequences into the online portal and select the corresponding species database. In case the transcript sequences are not available for a given species, the users can upload both the small RNA and the transcript sequences for prediction. An example of psRNATarget results is shown in Fig. 2.


psRNATarget: A Plant Small RNA Target Analysis Server (2017 Update)
[Visit The Previous Version, 2011 Release]
Home
Analysis
Download
Help
Citation
Contact
Submit

Analysis result

Scoring Schema: *Schema V2 (2017 release)*

Download:

Filtering by keywords:

expect:

UPE:

Sorting by: Expectation(E) miRNA Acc.

77848 Rows Page 1/3893

miRNA Acc.	Target Acc.	Expect	UPE	Alignment	Target Description	Inhibition	Multiplicity
13052_x10	XM_004510240.3	0.0	N/A	miRNA 22 UCCCCUUACUUCGGACCAGGCC 1 :::::::::::::::::::: Target 1090 CCUGGAAUGAAGCCUGGUCCGG 1111	PREDICTED: Cicer arietinum homeobox-leucine zipper protein ATHB-15-like (LOC101510535), transcript variant X2, mRNA	Cleavage	1
13052_x10	XM_004510239.3	0.0	N/A	miRNA 22 UCCCCUUACUUCGGACCAGGCC 1 :::::::::::::::::::: Target 1098 CCUGGAAUGAAGCCUGGUCCGG 1119	PREDICTED: Cicer arietinum homeobox-leucine zipper protein ATHB-15-like (LOC101510535), transcript variant X1, mRNA	Cleavage	1
13052_x10	XM_004486719.3	0.0	N/A	miRNA 22 UCCCCUUACUUCGGACCAGGCC 1 :::::::::::::::::::: Target 1258 CCUGGAAUGAAGCCUGGUCCGG 1279	PREDICTED: Cicer arietinum homeobox-leucine zipper protein ATHB-15 (LOC101493124), mRNA	Cleavage	1
131574_x2	XM_004510240.3	0.0	N/A	miRNA 20 CCCUUACUUCGGACCAGGCC 1 :::::::::::::::::::: Target 1092 UGGAAUGAAGCCUGGUCCGG 1111	PREDICTED: Cicer arietinum homeobox-leucine zipper protein ATHB-15-like (LOC101510535), transcript variant X2, mRNA	Cleavage	1

Fig. 2 Screenshot of psRNATarget results highlighting the miRNA–mRNA pair. The columns "miRNA Acc." and "Target Acc." indicate the miRNAs and their targets, respectively

3.5 Differential Expression of miRNAs

Once the miRNAs are identified, the most common downstream analysis is to study the expression patterns of these miRNAs across different samples. The expression of miRNAs for all samples is quantified by aligning the filtered reads from each sample to the final set of miRNA sequences.

```
> bowtie-build all_miRNAs.fa all_miRNAs
> bowtie all_miRNAs --best -v 2 -q SRR12847935.cleaned.poly-Atrimmed.fq -S SRR12847935.sam
> samtools view -bS SRR12847935.sam | samtools sort -o SRR12847935.bam -O bam -
> samtools index SRR12847935.bam
> samtools idxstats SRR12847935.bam | cut -f 1,3 > SRR12847935.counts.txt
```

The above commands are run on all samples to create a counts file for each sample. Next, we perform differential expression analysis using the DESeq2 package in R/Bioconductor (*see* **Note 5**). DESeq2 models raw read counts as negative binomial distribution with generalized linear models [12]. Before running DESeq2, we need to create two tab-separated text files, i.e., raw counts matrix file ("counts.txt") and samples list file ("samples.txt"). The raw counts matrix can be created by combining the individual count files. The samples list file should contain the list of all samples as the first column followed by sample details in the subsequent columns,

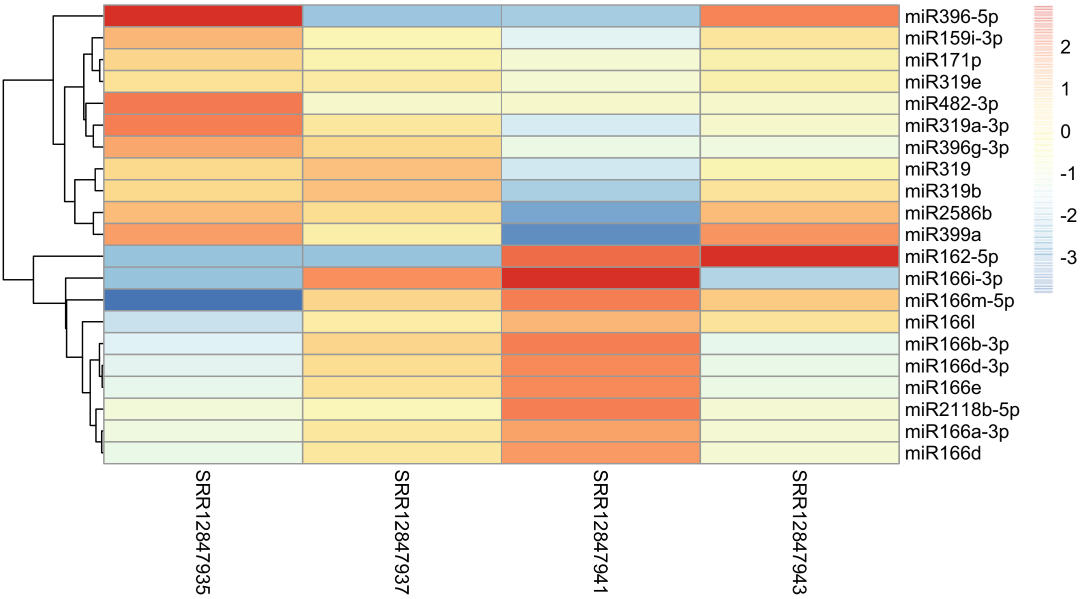


Fig. 3 Heatmap of the differentially expressed miRNAs identified from different chickpea samples. Heatmap was drawn with mean-centered log2 normalized expression values using the pheatmap R package

e.g., treatment, condition, time point. The first row in this file should specify the type of column. Once these files are ready, the users can use the command-line version of R or the GUI like RStudio to execute the commands:

```
> library(DESeq2)
> countData <- read.table("counts.txt", sep="\t", header=T,
as.is=T, row.names=1)
> colData <- read.table("samples.txt", sep="\t", header=T,
as.is=T, row.names=1)
> colData$id <- rownames(colData)
> dds <- DESeqDataSetFromMatrix(countData=countData, colDa-
ta=colData, design= ~ treatment)
> dds <- DESeq(dds, betaPrior=F)
```

The differentially expressed miRNAs between a pair of samples can be obtained using the commands:

```
> res <- results(dds, contrast=c("treatment",<sample1>,
<sample2>))
> write.table(res, file="diff_exp_miRNAs_sample1_vs_sample2.
txt", sep="\t", quote=F)
```

The differentially expressed miRNAs can be filtered using fold change ≥ 1 or ≤ -1 and a *P*-value <0.05. The expression profiles of miRNAs can be plotted as heat maps using either web-based tools such as WebMeV (http://mev.tm4.org) or R packages like pheatmap [25] and ComplexHeatmap [26]. An example heatmap showing the expression profile of different miRNAs is presented in Fig. 3.

4 Notes

1. Users should check the adapter removal statistics during the filtering step. The adapter sequences should be present in the majority of reads, ideally in over 90% of them. Lower percentages could indicate that the adapter sequence is incomplete or that the software used for adapter removal is not able to find all occurrences of the adapter. For accuracy, users can try more than one tool for this step.
2. The majority of NGS workflows include PCR duplicate removal, but in the case of sRNA-seq analysis, this step should be avoided. As sRNA libraries mostly consist of short reads with nearly identical sequences, filtering for duplicate reads will remove the highly expressed sRNAs, thus producing skewed results.
3. In the case of the identification of conserved miRNAs, miRNAs from miRBase should be selected carefully. It has been noted that many miRNAs reported in miRBase are false positives. For accurate detection, the "high confidence" miRNA list from miRBase should be selected, or a list of miRNAs conserved across all vascular plants or closely related species can be used.
4. For differential expression analysis, other R packages, edgeR [27] or limma [28], can also be used.
5. Predicting secondary structures is one of the most time-consuming steps. In case the number of candidate precursors is large, the users can divide them into small chunks and then run these chunks in parallel.

Acknowledgments

RKV thanks Bill and Melinda Gates Foundation, USA, for supporting several projects related to genomics-assisted breeding at ICRISAT and acknowledges Science and Engineering Research Board (SERB), Department of Science and Technology, Government of India for awarding the JC Bose National Fellowship.

References

1. Yang X, Li L (2011) miRDeep-P: a computational tool for analyzing the microRNA transcriptome in plants. Bioinformatics 27: 2614–2615
2. Shahid S, Axtell MJ (2014) Identification and annotation of small RNA genes using ShortStack. Methods 67:20–27
3. An J, Lai J, Sajjanhar A, Lehman ML, Nelson CC (2014) miRPlant: an integrated tool for identification of plant miRNA from RNA sequencing data. BMC Bioinformatics 15:275
4. Mathelier A, Carbone A (2010) MIReNA: finding microRNAs with high accuracy and no learning at genome scale and from deep sequencing data. Bioinformatics 26: 2226–2234
5. Guigon I, Legrand S, Berthelot J, Bini S, Lanselle D, Benmounah M et al (2019) miRkwood: a tool for the reliable identification of microRNAs in plant genomes. BMC Genomics 20:532
6. Garg V, Khan AW, Kudapa H, Kale SM, Chitikineni A, Qiwei S et al (2019) Integrated transcriptome, small RNA and degradome sequencing approaches provide insights into Ascochyta blight resistance in chickpea. Plant Biotechnol J 17:914–931
7. Varshney RK, Song C, Saxena RK, Azam S, Yu S, Sharpe AG et al (2013) Draft genome sequence of chickpea (*Cicer arietinum*) provides a resource for trait improvement. Nat Biotechnol 31:240–246
8. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120
9. Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17:10–12
10. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10:R25
11. Lorenz R, Bernhart SH, Zu Siederdissen CH, Tafer H, Flamm C, Stadler PF et al (2011) ViennaRNA Package 2.0. Algorithms Mol Biol 6:26
12. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550
13. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N et al (2009) The sequence alignment/map (SAM) format and SAMtools. Bioinformatics 25:2078–2079
14. Kalvari I, Nawrocki EP, Argasinska J, Quinones-Olvera N, Finn RD, Bateman A et al (2018) Non-coding RNA analysis using the Rfam database. Curr Protoc Bioinformatics 62:e51
15. Kozomara A, Birgaoanu M, Griffiths-Jones S (2019) miRBase: from microRNA sequences to function. Nucleic Acids Res 47:D155–D162
16. Bonnet E, He Y, Billiau K, Van de Peer Y (2010) TAPIR, a web server for the prediction of plant microRNA targets, including target mimics. Bioinformatics 26:1566–1568
17. Wu HJ, Ma Y, Chen T, Wang M, Wang X (2012) PsRobot: a web-based plant small RNA meta-analysis toolbox. Nucleic Acids Res 40:W22–W28
18. Chorostecki U, Palatnik JF (2014) comTAR: a web tool for the prediction and characterization of conserved microRNA targets in plants. Bioinformatics 30:2066–2067
19. Dai X, Zhuang Z, Zhao PX (2018) psRNATarget: a plant small RNA target analysis server (2017 release). Nucleic Acids Res 46: W49–W54
20. Addo-Quaye C, Eshoo TW, Bartel DP, Axtell MJ (2008) Endogenous siRNA and miRNA targets identified by sequencing of the *Arabidopsis* degradome. Curr Biol 18:758–762
21. Addo-Quaye C, Miller W, Axtell MJ (2009) CleaveLand: a pipeline for using degradome data to find cleaved small RNA targets. Bioinformatics 25:130–131
22. Zheng Y, Li Y, Sunkar R, Zhang W (2012) SeqTar: an effective method for identifying microRNA guided cleavage sites from degradome of polyadenylated transcripts in plants. Nucleic Acids Res 40:e28
23. Kakrana A, Hammond R, Patel P, Nakano M, Meyers BC (2014) sPARTA: a parallelized pipeline for integrated analysis of plant miRNA and cleaved mRNA data sets, including new miRNA target-identification software. Nucleic Acids Res 42:e139
24. Thody J, Folkes L, Medina-Calzada Z, Xu P, Dalmay T, Moulton V (2018) PAREsnip2: a tool for high-throughput prediction of small RNA targets from degradome sequencing data using configurable targeting rules. Nucleic Acids Res 46:8730–8739
25. Kolde R (2015) pheatmap: pretty heatmaps. R package version 1.0.8
26. Gu Z, Eils R, Schlesner M (2016) Complex heatmaps reveal patterns and correlations in

multidimensional genomic data. Bioinformatics 32:2847–2849

27. Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139–140

28. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W et al (2015) Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43:e47

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.