



Characterization of microRNA expression profiles by deep sequencing of small RNA libraries in leukemia patients from Naxi ethnic

Zhenglei Shen^{1#}, Xuezhong Gu^{2#}, Wenwen Mao³, Honghua Cao¹, Rui Zhang¹, Yeying Zhou¹, Kunmei Liu¹, Lilan Wang¹, Zhe Zhang⁴, Liefen Yin⁴

¹Department of Hematology, The Third Affiliated Hospital of Kunming Medical University, Kunming 650000, China; ²Department of Hematology, The First People's Hospital of Yunnan Province, Kunming 650000, China; ³Department of Geriatrics, The Second Hospital of Kunming, Kunming 650000, China; ⁴Department of Hematology, The Second Affiliated Hospital of Kunming Medical University, Kunming 650031, China

Contributions: (I) Conception and design: Z Shen, X Gu, L Yin; (II) Administrative support: Z Shen, X Gu, L Yin; (III) Provision of study materials or patients: H Cao, R Zhang, Z Zhang; (IV) Collection and assembly of data: Y Zhou, K Liu, L Wang; (V) Data analysis and interpretation: X Gu, W Mao; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Liefen Yin, Department of Hematology, The second Affiliated Hospital of Kunming Medical University, Kunming 650031, China. Email: ylfynkm@126.com.

Background: Leukemia is a hematological malignancy characterized by the proliferation of early lymphoid precursors that replaces normal hematopoietic cells of the bone marrow. Naxi (Naxi) ethnic minorities considered to be an area of low incidence. MicroRNAs (miRNAs) are a class of small noncoding RNAs that regulate the expression of other genes in various biological processes. The purpose of this work is to study the molecular mechanism of miRNAs in the leukemia from Naxi.

Methods: Six leukemia patients (case 2 to case 7) and one healthy person (case 1) from Naxi (Naxi) ethnic minorities were recruited. Total RNA was extracted from these samples and small RNA deep sequencing was performed.

Results: A list of miRNAs (1,392 known and candidate 125 novels) expressed in leukocytes were identified, and many differentially regulated targets involved in several cellular pathways, such as cancer, Rap1 signaling pathway, Ras signaling pathway, and endocytosis. Additionally, quantitative real time-polymerase chain reaction (qRT-PCR) results show that hsa-miR-181b-5p, hsa-miR-181a-3p, hsa-miR-181a-5p, and hsa-miR-342-3p has different expression patterns in different cancer cells, hsa-miR-450a-5p, and hsa-miR-1255a were dysregulated in all leukemia cells.

Conclusions: Several abnormal expressed miRNAs in leukemia patients were identified, the correlation of miRNAs dysregulation and leukemia biology demonstrates that specific miRNA can be potential therapeutic target.

Keywords: MicroRNA (miRNA); RNA sequencing; expression profile; leukemia; pathway; Naxi; novel; leukemia cells

Submitted Aug 09, 2018. Accepted for publication Jan 07, 2019.

doi: 10.21037/tcr.2019.01.18

View this article at: <http://dx.doi.org/10.21037/tcr.2019.01.18>

Introduction

MicroRNAs (miRNAs), a class of small noncoding RNAs, regulate the expression of other genes by targeting protein-coding transcripts in the post-transcriptional regulation (1). Approximately, 1,500 miRNAs were annotated in the miRNA database (miRBase) (2) and up to 60% of protein-coding genes were estimated to be regulated by miRNAs (3). miRNAs interact with 3'UTR of messenger RNAs (mRNAs) through base pairing and lead to their degradation, destabilization, or repression of translation through RNA-induced silencing complex (RISC), a complex of multiple proteins and miRNA-mRNA adduct (4,5). Occasionally, these miRNAs can also up-regulate gene expression (6).

It has been reported that miRNAs are essential for normal mammalian development and are involved in many biological processes, such as cell proliferation, differentiation, apoptosis, and metabolism, and their involvement in cancer has sparked increased interest in miRNA biology (7-9). Convincing evidence has indicated that miRNAs are key regulators of hematopoiesis (10,11). Leukemia is a group of hematopoietic neoplasms featuring impaired hematopoiesis and bone marrow failure caused by clonal expansion of undifferentiated myeloid precursors (12). The disease is frequently found to have recurrent chromosomal aberrations and gene mutations, and mostly carry driver mutations relevant to patient clinical outcomes (13). It is grouped by how quickly the disease develops (acute or chronic) as well as by the type of blood cell that is affected (lymphocytes or myelocytes). There are four main types of leukemia, including acute lymphocytic leukemia (ALL), chronic lymphocytic leukemia (CLL), acute myelocytic leukemia (AML), and chronic myelocytic leukemia (CML).

Recently, most researchers use next-generation sequencing (NGS) to profile miRNAs and then identify novel miRNAs associated with a disease of interest (14,15). However, most cohort studies were focused on the major populations such as Asia, America, Europe people etc. Little were known about the molecular profile changes (DNA mutations, RNA transcript and miRNAs, etc.) in the Chinese minorities. The Naxi people are one of the unique ethnic groups in Yunnan and live in a high concentration. In the present study, we analyzed differential expression profiles of miRNAs by quantitative small RNA NGS in six leukemia patients and one normal control. All these people are from Naxi ethnic minorities located in the northwest of Yunnan province, and considered to be an area of low incidence. This approach facilitated in identifying several

miRNAs relevant to different leukemia pathogenesis. Additionally, we also verified six miRNAs by quantitative real time-polymerase chain reaction (qRT-PCR) in human leukemic cell lines.

Methods

Patient cohort and sample collection

With informed consent and ethics approval, six leukemia patients (case 2 and case 3 have AML; case 4, case 5, and case 7 have ALL; case 6 has CML) and one healthy person (case 1) were recruited from Tumor Hospital in Yunnan Province. All seven person are from Naxi (Naxi) ethnic minorities, which considered to be an area of low incidence. The samples were collected in 1.5 mL microcentrifuge tubes, immediately snap-frozen in liquid nitrogen and stored at -80°C until further processing.

RNA isolation and deep sequencing

Total RNA isolation was carried out from peripheral blood or cell lines using TRIzol[®] Reagent (Invitrogen) according to the manufacturer's instructions. The RNA concentration was measured by Nanodrop ND-100 (Thermo Scientific, Waltham, USA), and RNA quality was assessed by the Agilent 2100 Bioanalyzer (Agilent, Santa Clara, USA). Small RNA sequencing libraries were constructed following IlluminaHTruSeq[™] Small RNA Sample Preparation protocol. In brief, 3' and 5' adapters were ligated to sRNA population and ligated RNAs were purified by running on urea-PAGE. Modified sRNAs were reverse transcribed to generate cDNA libraries and PCR amplified to add unique index sequence to each library.

Sequencing data analysis

The sequences were converted into FASTQ format and de-multiplexing were performed using Illumina bcl2fastq2 software version 2.17 (<http://support.illumina.com/downloads/>). Adaptor trimming was performed using the FASTQ Toolkit App of Illumina BaseSpace (<http://basespace.illumina.com/apps/>). Quality of the sequenced reads before and after adapter trimming was evaluated using FastQC software (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Cleaned sequences were aligned to the most recent mirBASE database release 21 (<http://mirBase.org/>) using the Small RNA App of Illumina

BaseSpace. Quantification of miRNA expression was performed by counting aligned reads to miRNA genes.

Novel miRNA prediction

The novel miRNA prediction strategy is based on first removing all known RNAs including those derived from exonic regions and then identifying those that are derived from intronic and intergenic regions. Identifying novel miRNAs from sequencing data commonly apply a combination of (I) evaluating miRNA secondary structures; and (II) ranking miRNA candidates by utilizing existing annotation or evolutionary conservation of the mature miRNA sequence. Structure threshold parameters of most algorithms are often optimized based on animal miRNA precursor structures.

Prediction of miRNA targets

The most probable targets of the differentially regulated miRNAs were predicted by two criteria. (I) Prediction by established target prediction software, which facilitate in identifying all the target genes using intersection between miRanda and miTarget; (II) inverse correlation in expression pattern between miRNA and coding genes. This approach compares the expression of putative genes with the different regulated miRNAs, and only genes that show inverse correlation to the miRNA levels were considered as most putative targets of the select miRNA.

qRT-PCR

qRT-PCR was performed to determine expression of six miRNAs: hsa-miR-181b-5p, hsa-miR-181a-3p, hsa-miR-181a-5p, hsa-miR-342-3p, hsa-miR-450a-5p, hsa-miR-1255a in three different human leukemic cell lines (Jurkat, HL-60, K-562). cDNA synthesis was performed using M-MLV Reverse Transcriptase (Promega) according to the manufacturer's protocol. All samples were run in three replicates. Ct values for miRNAs were normalized against U6 and the relative expression was calculated using $2^{-\Delta Ct}$ (Gene-U6) method. All the RT-qPCR primers are provided by RIBOBIO (<http://www.ribobio.com/>), the item number was listed in *Table S1*.

Statistical analysis

All bioinformatics-associated statistical analyses were

performed using the R package for statistical computing. One-way analysis of variance (ANOVA) was calculated and analyzed using GraphPad Prism. $P < 0.05$ is considered as statistically significant.

Results

Small RNAs sequencing and annotation

sRNAs from six leukemia patients (case 2 to case 7) and 1 normal control (case 1) were used and size selected by gel electrophoresis, then sequenced using a Solexa platform (Illumina, San Diego, USA). First, we removed the adaptor sequences from the sequence reads, and only those reads that were greater than 10 nucleotides were considered for further analysis. On average, ~13 million reads mapping to the human genome were obtained per sample (*Figure 1A*). The length of the detected sequences varied between 18 and 32 nucleotides in all these samples (*Figure S1*), similar to some other studies in animals (18–35 nucleotides) and plants (18–30 nucleotides). Then, we clustered these sequences on the basis of sequence similarity and performed similarity searches using specific databases (miRNAs, rRNA, tRNA, snRNA, snoRNAs). The frequency of reads mapping to miRNAs ranged from 59% to 71% and had a median of 65% (*Figure 1B*). In total, 1,392 known miRNAs (miRBase 21) were detected in at least one of seven samples sequenced (*Table 1*). To detect the proportion of mRNA derived by degradation, we checked if there are some sequences derived from intronic and exonic regions, and found that the median read frequency is 7.572% (*Figure 1C*). Additionally, a small fraction of reads mapped to rRNA, tRNA, snRNA, snoRNA, together comprising a frequency of ~0.8% (*Figure 1C*).

Identification of candidate novel miRNA genes

The representative hairpin structure of miRNA precursor is normally used to identify putative novel miRNAs. In this study, we integrated two miRNA prediction software, miRNAEvo (16) and mirdeep2 (17) for the new miRNA prediction. As shown in *Table 2*. Ninety-nine mature miRNAs, 26-star miRNAs, and 102 miRNA precursors were detected in at least one of the seven samples sequenced. Almost the same number of candidate novel miRNAs was obtained from each sample. During miRNA processing, Dicer enzyme plays an important role from precursor miRNAs to mature miRNAs. Due to the specificity of the

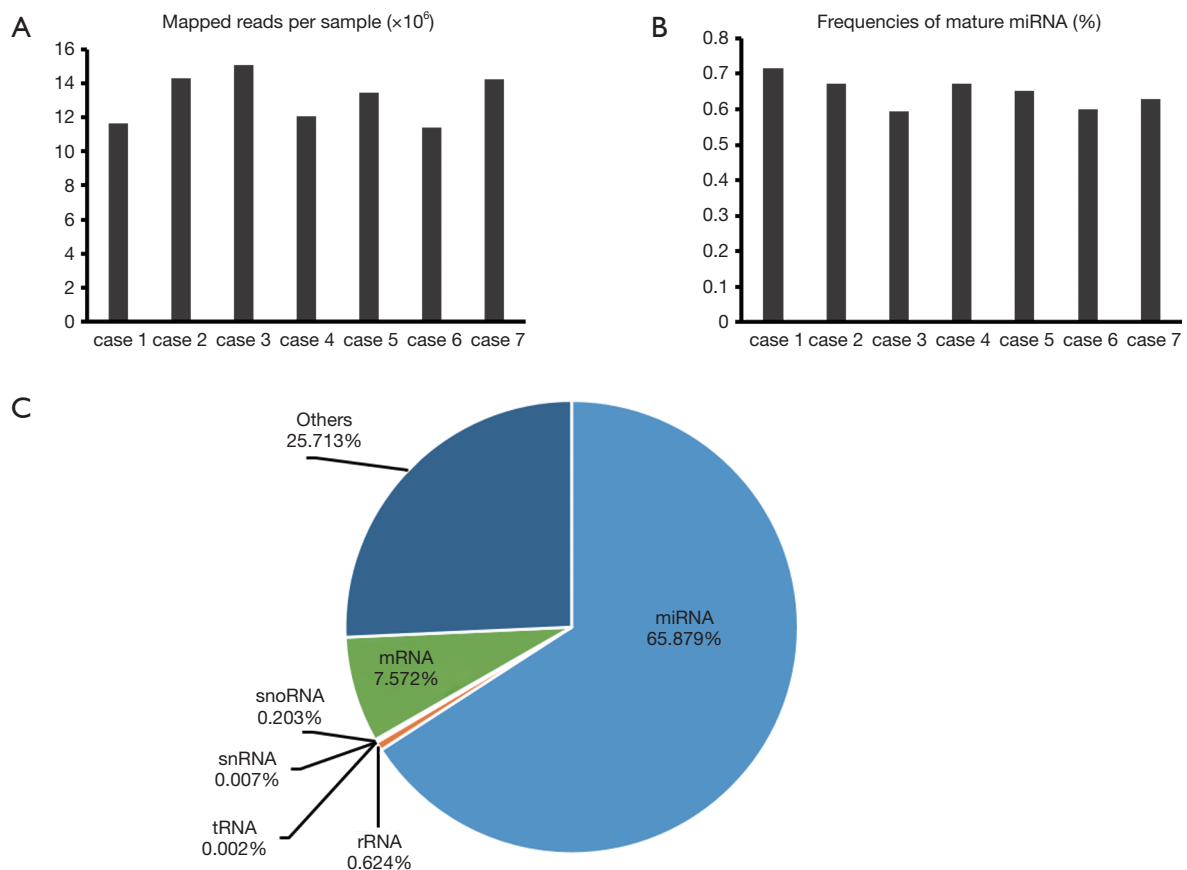


Figure 1 Overview of mapped reads, mature miRNAs and frequencies of RNA classes. (A) A number of reads ($\times 10^6$) mapped to the human genome for all samples; (B) the frequency of reads mapped to annotated mature miRNAs for all samples using the miRNA database (miRBase release 21); (C) the frequency of different classes of RNA species found in the dataset. miRNA, microRNA.

enzyme cleavage sites, the first base of miRNAs might display bias. So, we then analyzed the first base distribution of miRNAs in different length, and the base distribution of all miRNAs. As observed, the nucleotide distribution of 18–30 nucleotides novel miRNAs at the first nucleotide position had a strong bias (Figures 2,S2). Additionally, the base distributions of all miRNAs display that there are bias at each position (Figures 2,S3).

miRNA expression pattern

In order to identify miRNAs relevant to different types of leukemia, we divided our samples into four groups, acute leukemia (jxzl), chronic leukemia (mxzl), myeloid leukemia (sxzl) and lymphocytic leukemia (lxzl). To compare the expression levels between different samples, it is necessary to normalize against the read count by calculating a normalization factor. Here, the absolute sequence reads

were transformed into transcript abundance by normalizing the data in “transcripts per million (TPM)” for each library first. The expression levels ranged from less than 10 to more than 100,000 counts (Figure S4). Thus, the sequencing data revealed a wide range of expression levels spanning five orders of magnitude.

Using an arbitrary P value <0.05 and two-fold change cut-off, we found that in comparison to control, 21 and 125 miRNAs were up-regulated in jxzl and mxzl respectively, while 5 and 53 miRNAs were down-regulated in jxzl and mxzl respectively (Figure 3; Tables S2-S7). Venn diagram displayed that 18 miRNAs specifically regulated in jxzl and 170 miRNAs in mxzl (Figure S5). We then compared the differently expressed miRNAs between jxzl and mxzl, and found that 120 miRNAs were up-regulated and 203 miRNAs were down-regulated in jxzl. The up-regulated genes include six novel miRNAs (novel_581, novel_133, novel_692, novel_224, novel_401, novel_94) and several

Table 1 The number of known miRNAs in each sample

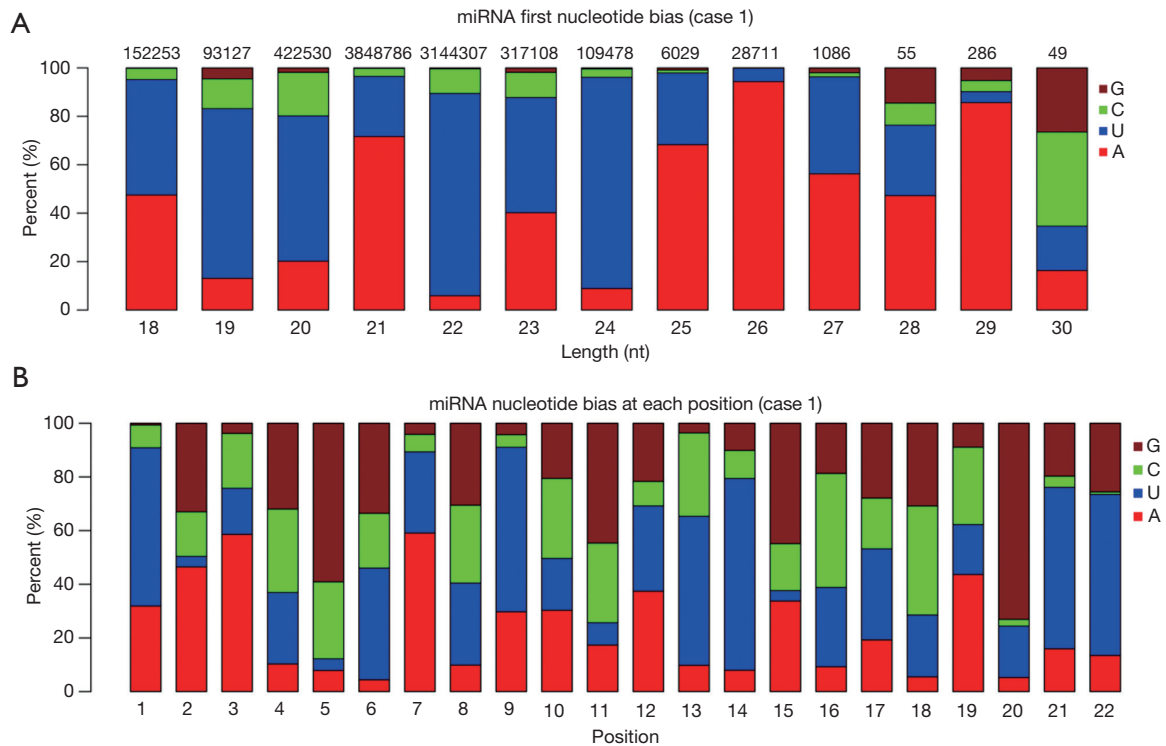
Types	Total	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7
Mapped mature miRNA	1,392	942	894	1,021	909	981	938	918
Mapped pre-miRNA	1,156	828	793	895	830	857	836	820

miRNA, microRNA.

Table 2 The number of novel miRNAs in each sample

Types	Total	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7
Mapped mature miRNA	99	53	53	63	62	45	63	57
Mapped star miRNA	26	10	8	14	9	10	8	8
Mapped pre-miRNA	102	57	57	71	64	49	65	59

miRNA, microRNA.

**Figure 2** The nucleotide distribution of novel miRNAs. (A) The nucleotide distribution of 18–30 nucleotide novel miRNAs at the first nucleotide position; (B) the nucleotide distribution of all novel miRNAs at each nucleotide. Different colors represent different nucleotide. G is brown; C is green; U is blue; A is red. miRNA, microRNA.

known miRNAs, such as hsa-miR-103a-2-5p, hsa-miR-320b, hsa-miR-320c, hsa-miR-1180-3p, hsa-let-7a-5p, hsa-let-7b-5p, hsa-let-7c-5p, and hsa-let-7d-5p. And the downregulated genes include novel_498, novel_377,

novel_580, hsa-miR-582-5p, hsa-miR-582-3p, hsa-miR-450b-5p, and hsa-miR-450a-5p (Tables S2-S7). These differentially expressed miRNAs might play crucial role in regulating the rate of the disease development.

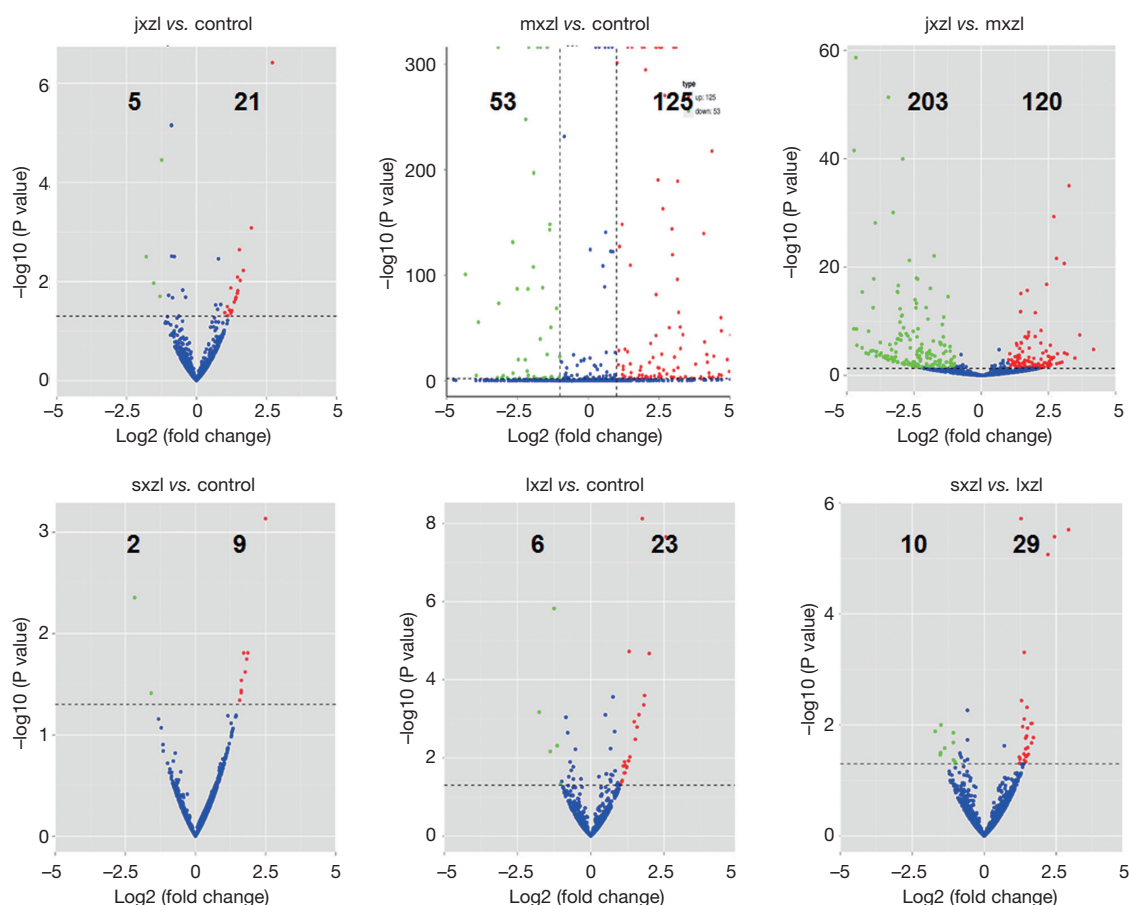


Figure 3 The identity of miRNAs with altered expression levels in the indicated comparisons. Volcano plots of expression changes were generated. Log₂ of fold change is shown on the horizontal axis, and $-\log_{10}$ of the Q value (adjusted P value) is shown on the vertical axis. Red and green dots represent significantly up- and downregulated miRNAs (FDR <0.01). Blue dots are genes with a nonsignificant change. miRNA, microRNA.

Next, we compared the differentially expressed miRNAs between the sxzl or lxzl with control. As shown, 9 and 23 miRNAs were up-regulated in sxzl and lxzl, 2 and 6 miRNAs were down-regulated in sxzl and lxzl (*Figure 3*; *Tables S2-S7*). When compared the differentially expressed miRNAs in sxzl and lxzl, the results showed that 29 known miRNAs were up-regulated and 10 known miRNAs were downregulated, including hsa-miR-181-5p, hsa-miR-181a-3p, hsa-miR-181b-5p, hsa-miR-342-3p, hsa-let-7f-1-3p, and hsa-miR-30a-5p (*Tables S2-S7*). These miRNAs might play an important role in the specific type of blood cell development.

Predicted targets of differentially regulated miRNAs

In order to understand the biological role of the

differentially regulated miRNAs, it is necessary to figure out the putative targets of the detected miRNAs. In this study, two approaches were employed for target prediction. First, predict target using two popular software tools, miRanda, and miTarget. Then, compare the expression levels between miRNA and putative target genes. Those genes that coexist in two databases, and show inversion correlation to its miRNAs were considered to be the targets. KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analysis was performed for predicted targets. We analyzed the pathways of differentially expressed genes between jxzl and control, and found that the most significant gene-enrichment included “pathways in cancer”, “Rap1 signaling pathways”, “Ras signaling pathways”, and “endocytosis”. However, when compared between mxzl and control, the most significant gene-enrichment was metabolic

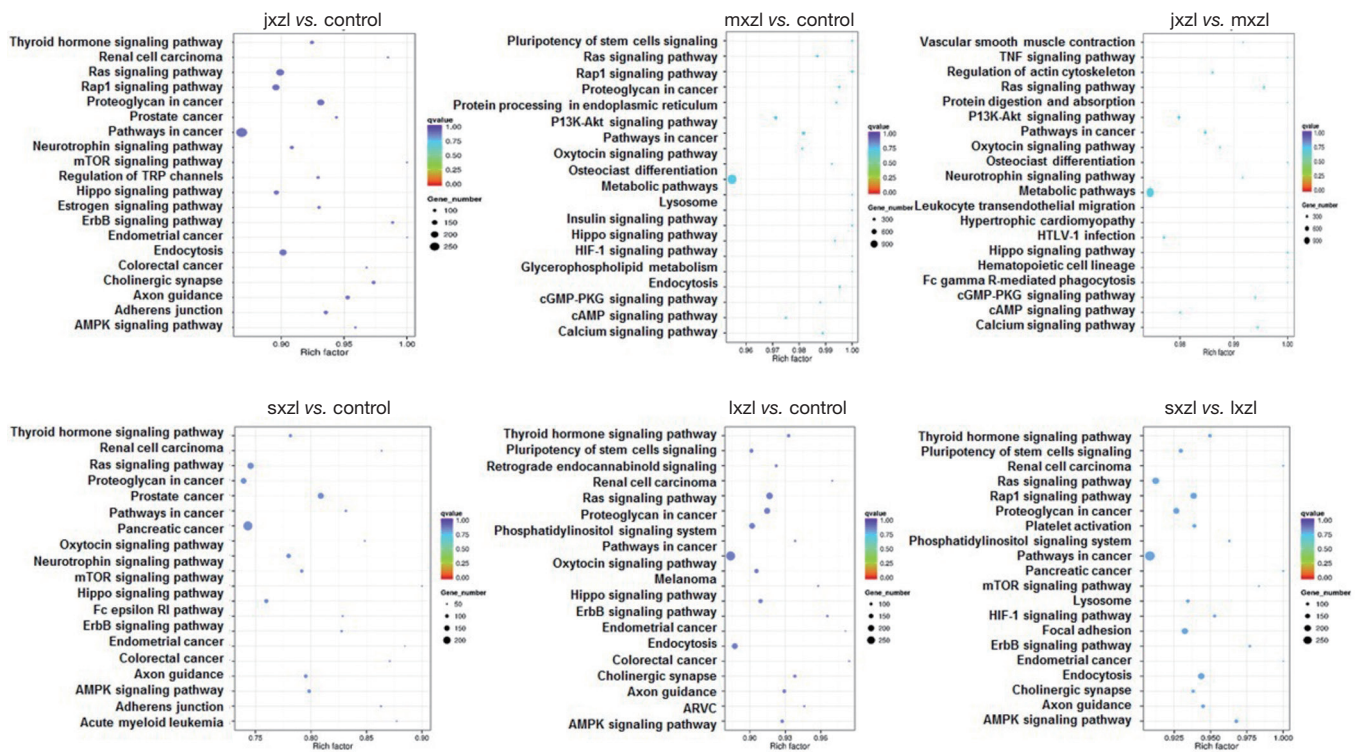


Figure 4 Pathway analysis for the predicted targets of different expression miRNAs in the indicated comparisons. The 20 most significant pathways are shown with the number of genes present in their respective Q value (adjusted P value). miRNA, microRNA.

pathways. Similar results were found in jxzl and mxzl, which suggests that the metabolic pathway is responsible for the development of leukemia. Then we compared the pathways of differently expressed genes between sxzl and control, between lxzl and control (Figure 4). The results display that cancer development is the most significant pathway, which indicates that the miRNAs play an important role in determining the type of cancer and its development.

Validation of candidate miRNAs by qRT-PCR

To validate these findings, we examined the expression of six miRNAs, hsa-miR-181b-5p, hsa-miR-181a-3p, hsa-miR-181a-5p, hsa-miR-342-3p, hsa-miR-450a-5p, and hsa-miR-1255a in three different human leukemic cell lines (Jurkat, HL-60, K-562) by RT-qPCR (Figure 5). As known, Jurkat is acute lymphoid leukemia, HL-60 is AML, and K-562 is CML and consistent with the small RNAseq results, we found that hsa-miR-181a-5p and hsa-miR-181b-5p displayed higher expression in Jurkat and HL-60 than that in K-562. However, the expression level of hsa-miR-342a-3p in Jurkat was higher than other two (HL-60 and

K-562). And hsa-miR-181a-3p showed different expression in three cell lines (Jurkat had highest and K-562 had lowest). Additionally, both hsa-miR-450a-5p and hsa-miR-1255a displayed low expression in all three cell lines. These results suggest that hsa-miR-181b-5p, hsa-miR-181a-3p, hsa-miR-181a-5p, and hsa-miR-342-3p showed different expression patterns in different types of cancer cells, and hsa-miR-450a-5p, hsa-miR-1255a were dysregulated in all the leukemia cells.

Discussion

The study of miRNA expression has a certain guiding role in clinical practice as the abnormal expression of miRNA may be used as a prognostic indicator for leukemia patients (18,19). However, this is based on a large number of races and there was very little research on ethnic minorities. To provide accurate treatment for minority people, the differences between Han people in genetic characteristics and molecular maps need to be analyzed. In this study, 6 patients from the Naxi ethnic minority were selected for miRNA analysis, the Naxi ethnic group lives in the western

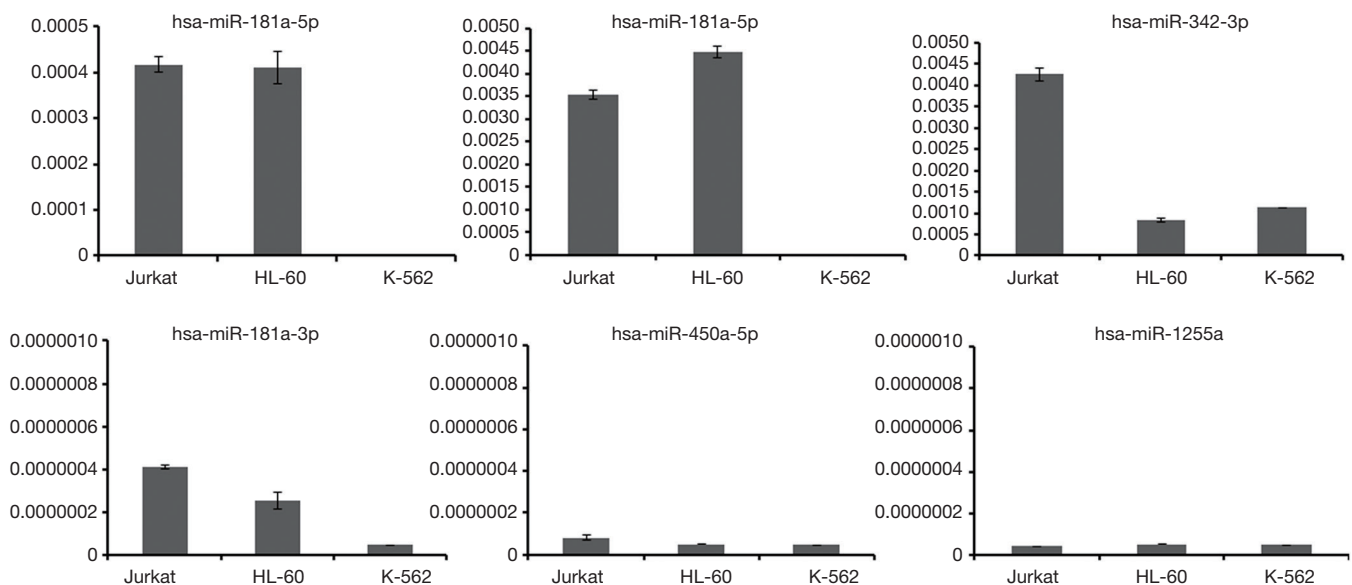


Figure 5 The relative expression levels of six known miRNAs. RT-PCR was performed in three cell lines. U6 was used as an internal control and standard deviations were calculated from three technical repeats. RT-PCR, real time-polymerase chain reaction; miRNA, microRNA.

Yunnan Plateau in the northwest of Yunnan Province, with unique national culture and regional characteristics. This is the first report of the Naxi ethnic miRNA expression. In spite of the small number of examples, it has laid a certain foundation for the follow-up study.

In this study, six patients from the Naxi ethnic minority were selected for miRNA analysis, the Naxi ethnic group located in the western Yunnan plateau, the northwest of Yunnan province, with a unique national culture and regional characteristics. Our paper, for the first time, studied the miRNA expression in Naxi people. Although the number of samples is limited, it lays the foundation for future research.

Certain investigations indicated that some miRNAs can affect the directional differentiation of hematopoietic cells, and associated with the pathogenesis of leukemia and lymphoma. Understanding and clarifying the relationship between miRNAs and leukemia cells can help to reveal the molecular mechanism of leukemia, and miRNAs and specific target genes can also be used to achieve targeted therapeutics. Till now, several miRNAs expressed tags, as well as miRNAs, responsible for the AML and CLL have been reported (18,20,21). For example, it was found that frequent deletions and down-regulation of miRNA15 and miRNA16 were happened in ALL (22); Garzon *et al.* found that the expression of miR-191 and miR-199a was up-regulated in AML, and its high expression was associated with

poor prognosis in AML patients (18). Dixon-McIver *et al.* also confirmed that t(15;17) chromosome heterotopia is related to the upregulation of miR-127, miR-154, miR-299, miR-323, miR-368 and miR-370 (19).

In the present study, deep sequencing technology was used to quantify miRNA expression in six leukemia patients from Naxi ethnic. From these samples, we observed several interesting findings, including the discovery of novel miRNAs and a valuable list of differentially expressed miRNAs in chronic and acute leukemia, as well as lymphocytic and myelocytic leukemia. Our major findings are a list of expressed miRNAs (1,392 known and 125 candidate novels) in leukocytes and the discovery that the differential expression of miRNAs is most abundant in the pathway of cancer. These 125 candidate miRNAs should be further studied to exclude the false positive results and find the specific miRNAs after more samples would be collected. Our qRT-PCR results show convincingly that such miRNAs are likely to be real and not a result of an artifact of sequencing.

The known miRNAs found to be enriched in acute leukemia were members of the let-7 family, miR185 and miR320, which were also reported in peripheral blood mononuclear cells using “Taqman miRNA assay” (23). Therefore, our results are in agreement with previous studies based on a different methodology. There are multiple mechanisms that are likely to regulate miRNA

levels similar to that of mRNAs.

Additionally, we found for the first time that novel-94, 692, 581, 580, 498, 401, 377, 224 and 133 showed abnormal expression in leukemia patients. We also validate our RNAseq data by RT-qPCR in three leukemia cell lines, the expression of hsa-miR-181b-5p, hsa-miR-181a-3p, hsa-miR-181a-5p, hsa-miR-342-3p, hsa-miR-450a-5p, and hsa-miR-1255a were consistent with our sequence data. Most interestingly, WT1 gene was predicted to be the target of miR-1255a by TargetScan v7.1. As known, WT1 was highly expressed in peripheral blood or narrow bone of most AML patients, and supposed to be the marker of detecting minimal residual disease (MRD). Our results indicated that the expression of miR-1255a was significantly down-regulated in AML patients, which suggested that the WT1 gene might be regulated by miR-1255a. This new finding provides a new direction for the further research.

Acknowledgments

Funding: This research was supported by grants from National Natural Science Foundation of China (81360089) and Scientific and Technological Commission of Yunnan Province (2015FB072).

Footnote

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <http://dx.doi.org/10.21037/tcr.2019.01.18>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). Informed consent was taken from all patients. The study was approved by the Tumor Hospital in Yunnan Province (2015FB072).

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the

formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Kusenda B, Mraz M, Mayer J, et al. MicroRNA biogenesis, functionality and cancer relevance. *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub* 2006;150:205-15.
2. Griffiths-Jones S, Grocock RJ, van Dongen S, et al. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res* 2006;34:D140-4.
3. Friedman RC, Farh KK, Burge CB, et al. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* 2009;19:92-105.
4. Rana TM. Illuminating the silence: understanding the structure and function of small RNAs. *Nat Rev Mol Cell Biol* 2007;8:23-36.
5. Hammell CM. The microRNA-argonaute complex: a platform for mRNA modulation. *RNA Biol* 2008;5:123-7.
6. Ørom UA, Nielsen FC, Lund AH. MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. *Mol Cell* 2008;30:460-71.
7. Tsuchiya S, Okuno Y, Tsujimoto G. MicroRNA: biogenetic and functional mechanisms and involvements in cell differentiation and cancer. *J Pharmacol Sci* 2006;101:267-70.
8. Tang JT, Fang JY. MicroRNA regulatory network in human colorectal cancer. *Mini Rev Med Chem* 2009;9:921-6.
9. Sassen S, Miska EA, Caldas C. MicroRNA: implications for cancer. *Virchows Arch* 2008;452:1-10.
10. Ling H, Fabbri M, Calin GA. MicroRNAs and other non-coding RNAs as targets for anticancer drug development. *Nat Rev Drug Discov* 2013;12:847-65.
11. Larsen MT, Häger M, Glenthøj A, et al. miRNA-130a regulates C/EBP-ε expression during granulopoiesis. *Blood* 2014;123:1079-89.
12. Lee YG, Kim I, Oh S, et al. Small RNA sequencing profiles of mir-181 and mir-221, the most relevant microRNAs in acute myeloid leukemia. *Korean J Intern Med* 2019;34:178-83.
13. Papaemmanuil E, Gerstung M, Bullinger L, et al. Genomic Classification and Prognosis in Acute Myeloid Leukemia. *N Engl J Med* 2016;374:2209-21.
14. Schee K, Lorenz S, Worren MM, et al. Deep Sequencing the MicroRNA Transcriptome in Colorectal Cancer. *PLoS One* 2013;8:e66165.
15. Vaz C, Ahmad HM, Sharma P, et al. Analysis of microRNA

- transcriptome by deep sequencing of small RNA libraries of peripheral blood. *BMC Genomics* 2010;11:288.
16. Wen M, Shen Y, Shi S, et al. miREvo: an integrative microRNA evolutionary analysis platform for next-generation sequencing experiments. *BMC Bioinformatics*. 2012;13:140.
 17. Friedländer MR, Mackowiak SD, Li N, et al. miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res* 2012;40:37-52.
 18. Garzon R, Volinia S, Liu CG, et al. MicroRNA signatures associated with cytogenetics and prognosis in acute myeloid leukemia. *Blood* 2008;111:3183-9.
 19. Dixon-McIver A, East P, Mein CA, et al. Distinctive patterns of microRNA expression associated with karyotype in acute myeloid leukaemia. *PLoS One* 2008;3:e2141.
 20. Luo HC, Zhang ZZ, Zhang X, et al. MicroRNA expression signature in gastric cancer. *Chin J Cancer Res* 2009;21:74-80.
 21. Navarro A, Gaya A, Martinez A, et al. MicroRNA expression profiling in classic Hodgkin lymphoma. *Blood* 2008;111:2825-32.
 22. Calin GA, Dumitru CD, Shimizu M, et al. Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc Natl Acad Sci U S A* 2002;99:15524-9.
 23. Liang Y, Ridzon D, Wong L, et al. Characterization of microRNA expression profiles in normal human tissues. *BMC Genomics* 2007;8:166.

Cite this article as: Shen Z, Gu X, Mao W, Cao H, Zhang R, Zhou Y, Liu K, Wang L, Zhang Z, Yin L. Characterization of microRNA expression profiles by deep sequencing of small RNA libraries in leukemia patients from Naxi ethnic. *Transl Cancer Res* 2019;8(1):160-169. doi: 10.21037/tcr.2019.01.18

Table S1 The item number of RT-qPCR primers provided by RIBOBIO

Species	Gene name	Item number
Human	hsa-miR-181b-5p	MIMAT0000257
Human	hsa-miR-181a-3p	MIMAT0000270
Human	hsa-miR-181a-5p	MIMAT0000256
Human	hsa-miR-342-3p	MIMAT0000753
Human	hsa-miR-450a-5p	MIMAT0001545
Human	hsa-miR-1255a	MIMAT0005906

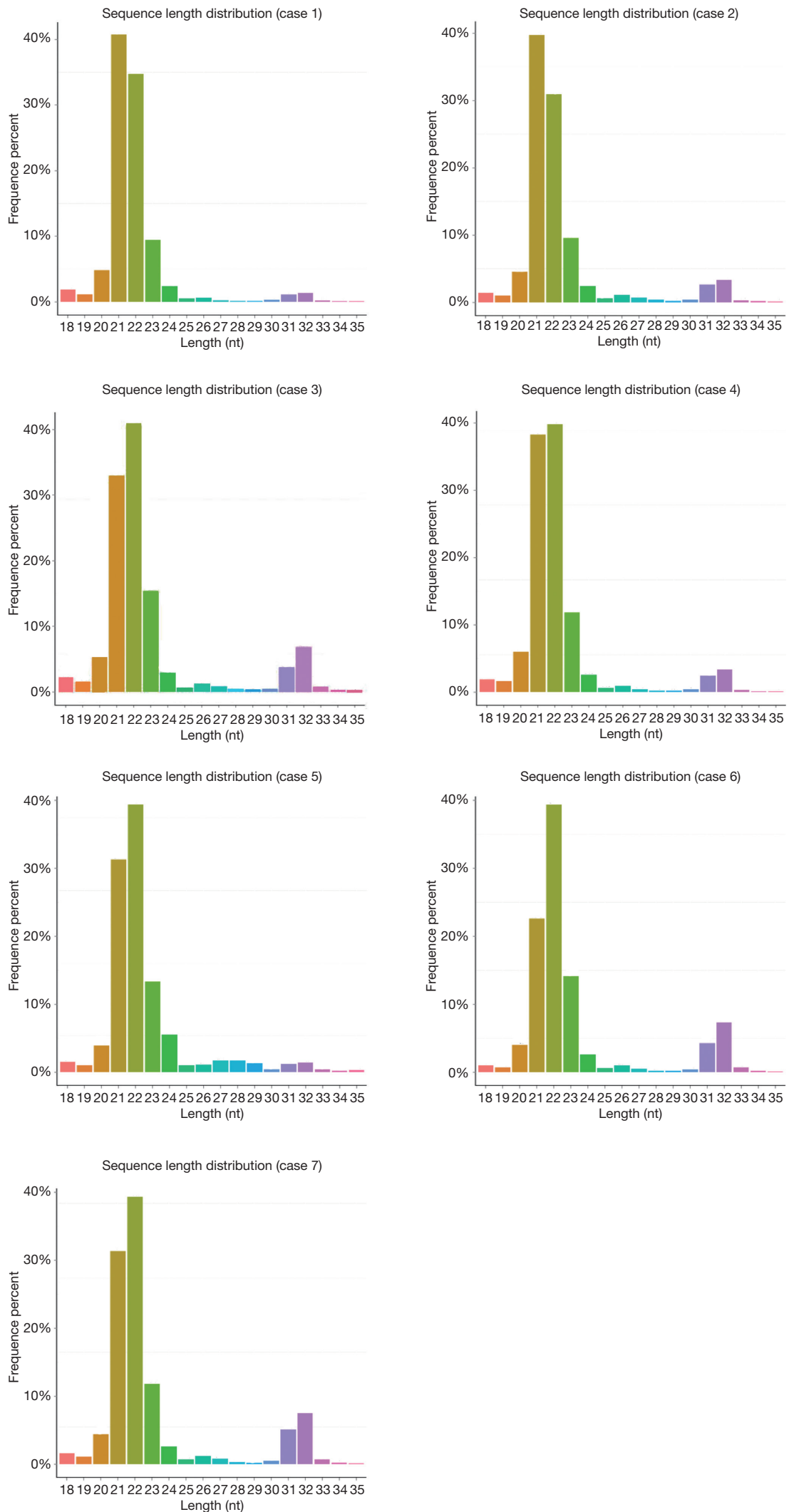


Figure S1 Small RNA length distribution and the frequency percent for all samples.

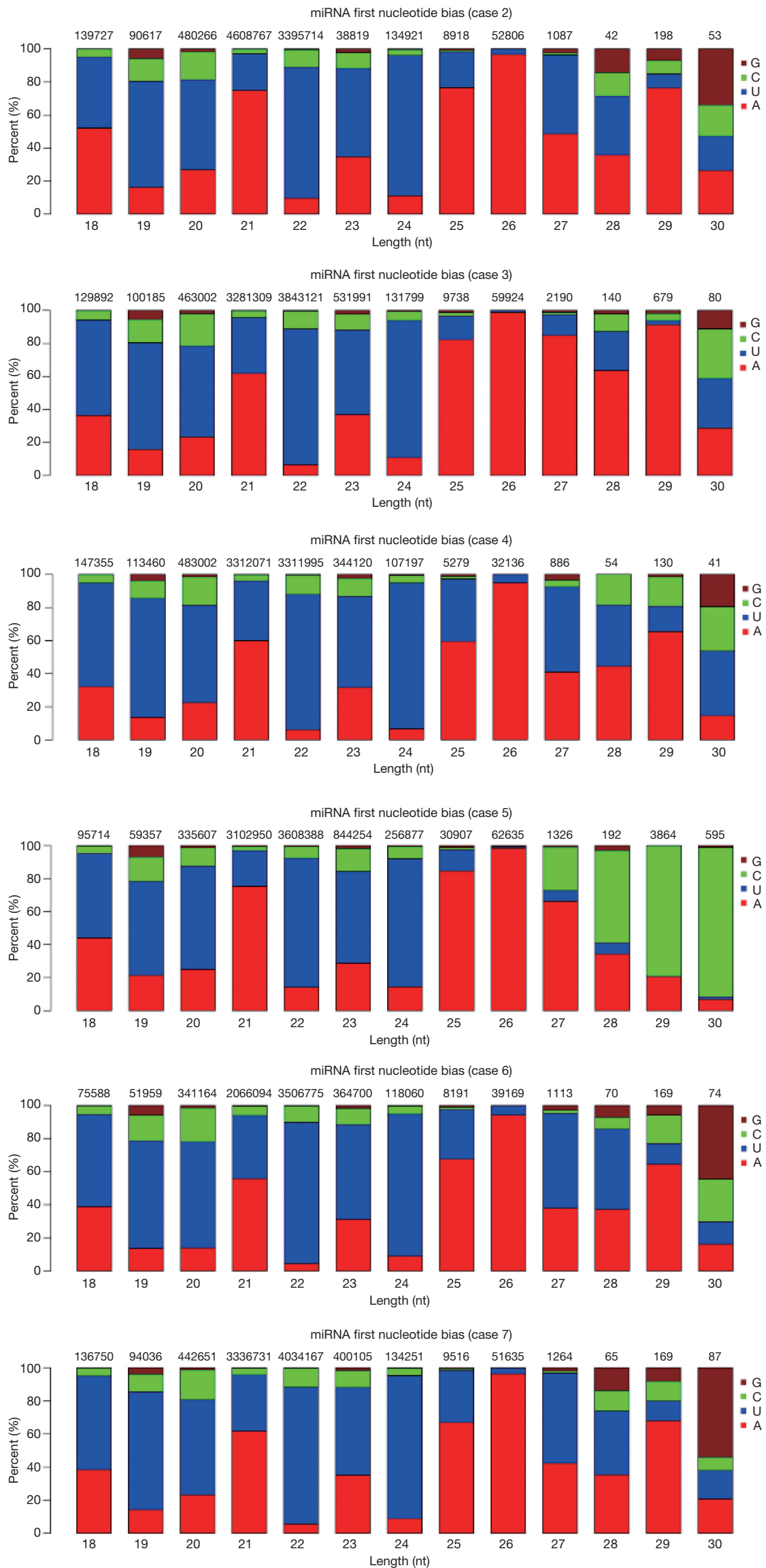


Figure S2 The nucleotide distribution of 18nt- 30nt novel miRNAs at the first nucleotide position. Different colors represent different nucleotide. G is brown; C is green; U is blue; A is red.

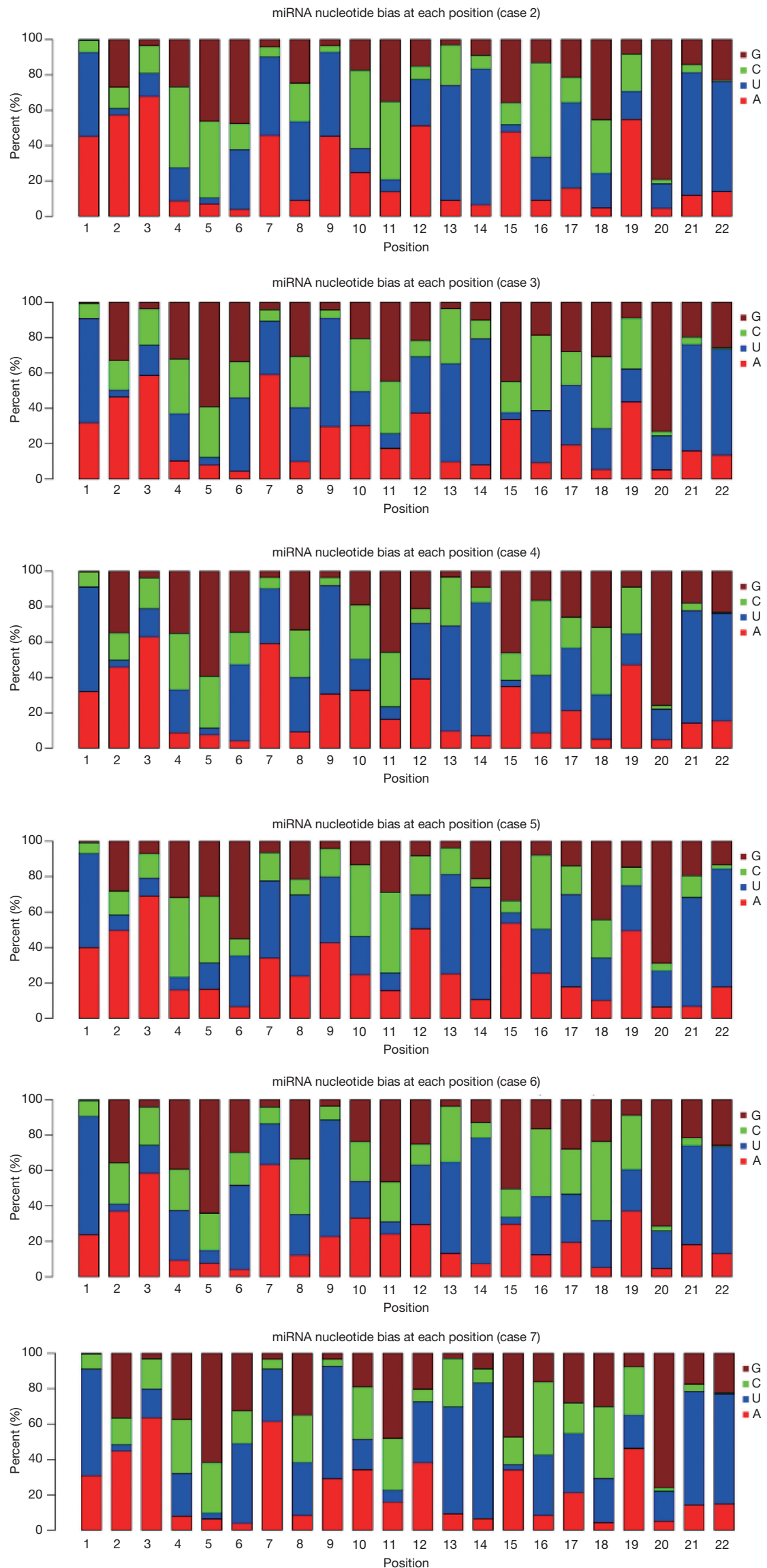


Figure S3 The nucleotide distribution of all novel miRNAs at each nucleotide. Different colors represent different nucleotide. G is brown; C is green; U is blue; A is red.

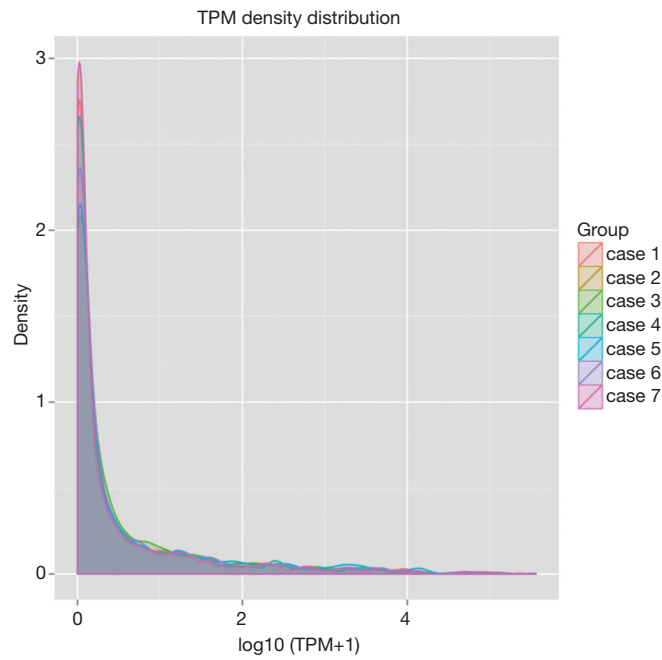


Figure S4 TPM density distribution of known and novel miRNAs for each sample. The overall level of expression of miRNAs with respect to a number of miRNAs is shown, Numbers of sequence reads are taken as miRNA levels and the values are represented in the form of a range of values. The expression levels of the miRNAs span up to five orders of magnitude.

Table S2 *jxzl vs. control*

sRNA	jxzl readcount	Control readcount	Log2 fold change	P value	Q value
Up-regulated					
hsa-miR-1255a	55.92721169	3.347829955	2.7099	3.84E-07	0.000219
hsa-miR-99b-3p	27.37110049	2.231886637	1.9569	0.000831	0.094769
hsa-miR-1270	26.99996407	6.69565991	1.5303	0.002276	0.21624
hsa-miR-34a-5p	177.0151767	11.15943318	1.6723	0.005981	0.34091
hsa-miR-335-3p	24.84632076	5.579716592	1.4625	0.00816	0.42281
hsa-miR-139-5p	226.9695179	26.78263964	1.5594	0.009461	0.44938
hsa-miR-1226-3p	29.98544425	10.04348987	1.2247	0.013456	0.5463
hsa-miR-99b-5p	298.7417542	31.24641292	1.4743	0.015335	0.5463
hsa-miR-149-5p	10.96456363	0	1.4556	0.017281	0.56172
hsa-miR-8485	11.14732578	0	1.4116	0.020799	0.56172
hsa-miR-3117-3p	15.64362549	0	1.4008	0.021681	0.56172
hsa-miR-7854-3p	6.27480693	0	1.3922	0.022724	0.56315
hsa-miR-3656	7.953821644	1.115943318	1.3472	0.02591	0.61537
hsa-miR-652-5p	18.29332169	6.69565991	1.101	0.032129	0.67828
hsa-miR-23b-5p	11.61892026	3.347829955	1.1754	0.037711	0.72538
hsa-miR-4433a-3p	188.5453394	20.08697973	1.2659	0.03874	0.72538
hsa-miR-4433b-5p	188.154882	20.08697973	1.2614	0.03945	0.72538
hsa-miR-1271-5p	21.83756653	8.927546547	1.0088	0.042338	0.75415
hsa-miR-6786-3p	5.197110923	0	1.2277	0.043122	NA
hsa-miR-495-3p	9.17037747	0	1.2207	0.04373	0.75534
hsa-let-7e-5p	245.6468219	77.00008897	1.1127	0.048766	0.81754
Down-regulated					
hsa-miR-450a-5p	147.230475	369.3772384	-1.2423	3.55E-05	0.005056
hsa-miR-642a-5p	1.521240321	12.2753765	-1.795	0.003171	NA
hsa-miR-642b-3p	1.521240321	12.2753765	-1.795	0.003171	NA
hsa-miR-3690	3.130164922	30.1304696	-1.5273	0.010778	0.47255
hsa-miR-4435	5.723576538	20.08697973	-1.3048	0.019909	0.56172

Table S4 | jxzl vs. mxzl

Table with 7 columns: sRNA, jxzl readcount, mxzl readcount, Log2 fold change, P value, Q value. The table lists up-regulated and down-regulated sRNAs with their respective read counts, fold changes, and significance values.

Table S5 *szxl vs. control*

sRNA	szxl readcount	Control readcount	Log2 fold change	P value	Q value
Up-regulated					
hsa-miR-1255a	73.93453	3.670145	2.4982	0.000734	0.98318
hsa-miR-99b-3p	34.05871	2.446764	1.8701	0.015512	0.98318
hsa-miR-342-3p	1,564.196	274.0375	1.7191	0.015543	0.98318
hsa-miR-34a-5p	144.9126	12.23382	1.8276	0.017862	0.98318
hsa-miR-181a-3p	1,231.747	58.72232	1.7736	0.023952	0.98318
hsa-miR-181b-5p	1,733.788	261.8037	1.6401	0.028975	0.98318
hsa-miR-125a-5p	565.7181	50.15865	1.6352	0.036442	0.98318
hsa-miR-8485	12.10476	0	1.6302	0.038292	0.98318
hsa-miR-99b-5p	552.4379	34.25469	1.575	0.045443	0.98318
Down-regulated					
hsa-miR-30a-5p	554.4653	5,402.454	-2.1709	0.004414	0.98318
hsa-miR-30a-3p	41.46146	225.1022	-1.5811	0.038759	0.98318

Table S6 *lxzl vs. control*

sRNA	lxzl readcount	Control readcount	Log2 fold change	P value	Q value
Up-regulated					
hsa-miR-125a-5p	164.4959262	42.54988	1.7568	7.52E-09	3.92E-06
hsa-miR-1255a	47.68217947	3.113406	2.566	2.19E-08	5.71E-06
hsa-let-7e-5p	196.18214	71.60834	1.309	1.89E-05	0.00222
hsa-miR-99b-5p	228.78698	29.05846	1.992	2.13E-05	0.00222
hsa-miR-99b-3p	23.73255619	2.075604	1.8339	0.000253	0.020466
hsa-miR-34a-5p	226.0002671	10.37802	1.8112	0.000441	0.028716
hsa-miR-127-3p	140.8358884	21.79384	1.6454	0.000785	0.036749
hsa-miR-1270	26.93738349	6.226812	1.4834	0.001185	0.044086
hsa-miR-139-5p	175.0936703	24.90725	1.5811	0.001626	0.056467
hsa-miR-493-3p	20.24950193	1.037802	1.5177	0.00335	0.096951
hsa-miR-493-5p	36.66853096	5.18901	1.3349	0.009469	0.22425
hsa-miR-143-5p	7.137243546	0	1.2882	0.011921	NA
hsa-miR-1226-3p	28.83038399	9.340218	1.1487	0.012641	0.27767
hsa-miR-335-3p	18.89945023	5.18901	1.1775	0.015629	0.32382
hsa-miR-1303	108.0305569	36.32307	1.1114	0.01616	0.32382
hsa-miR-7854-3p	6.459986627	0	1.2208	0.016689	NA
hsa-miR-379-5p	15.65690359	1.037802	1.2222	0.017663	0.32865
hsa-miR-3656	8.119232076	1.037802	1.1662	0.0242	NA
hsa-miR-6786-3p	6.151685313	0	1.14	0.024322	NA
hsa-miR-381-3p	97.77120798	20.75604	1.0721	0.037705	0.49031
hsa-miR-182-3p	10.42158481	2.075604	1.0584	0.040467	0.49031
hsa-miR-4433b-5p	174.713051	18.68044	1.0067	0.048513	0.52705
hsa-miR-4433a-3p	175.0375348	18.68044	1.0055	0.048736	0.52705
Down-regulated					
hsa-miR-450a-5p	135.4842658	343.5125	-1.2519	1.49E-06	0.000259
hsa-miR-3690	2.625944012	28.02066	-1.7583	0.000675	0.036749
hsa-miR-3135a	22.32185019	59.15472	-1.1441	0.004872	0.13359
hsa-miR-642a-5p	1.308257088	11.41582	-1.3814	0.00686	NA
hsa-miR-642b-3p	1.308257088	11.41582	-1.3814	0.00686	NA
hsa-miR-4435	5.142299703	18.68044	-1.0286	0.046412	0.52705

Table S7 *sxzl vs. lxzl*

sRNA	sxzl readcount	lxzl readcount	Log2 fold change	P value	Q value
Up-regulated					
hsa-miR-181a-5p	39,056.69	15,625.02	1.2746	1.92E-06	0.000148
hsa-miR-181a-3p	1,215.375	77.99342	2.9279	3.03E-06	0.000148
hsa-miR-181b-5p	1,693.275	224.783	2.4416	4.07E-06	0.000148
hsa-miR-342-3p	1,523.522	257.5529	2.2109	8.51E-06	0.000232
hsa-miR-146b-3p	401.1346	140.4814	1.3875	0.000492	NA
hsa-miR-15b-3p	2,750.452	1,008.29	1.2939	0.003621	0.078929
hsa-miR-16-2-3p	1,296.128	381.1239	1.4855	0.004811	0.084379
hsa-miR-1260b	76.25221	24.49384	1.3869	0.007796	NA
hsa-miR-1-3p	90.20319	18.62308	1.6595	0.009336	NA
hsa-miR-1268a	11.95946	2.560318	1.63	0.009398	NA
hsa-miR-150-5p	2,148.384	716.1593	1.3371	0.010691	0.12348
hsa-miR-197-3p	2,577.615	680.3299	1.5043	0.011328	0.12348
hsa-let-7f-1-3p	19.77389	5.190162	1.4708	0.015771	NA
hsa-miR-200a-3p	4.417496	0	1.7056	0.016799	NA
hsa-miR-1307-5p	44.33103	12.17739	1.4361	0.017191	NA
hsa-miR-504-5p	31.89215	4.118537	1.6451	0.021032	NA
hsa-miR-4707-3p	19.32852	3.991825	1.5099	0.025443	NA
hsa-miR-92a-1-5p	15.18709	3.689128	1.4387	0.026903	NA
hsa-miR-27b-5p	16.97949	4.421235	1.3786	0.033286	NA
hsa-miR-3154	4.762506	0.373006	1.5324	0.033713	NA
hsa-miR-6770-5p	2.865023	0	1.4812	0.035719	NA
hsa-miR-1275	5.691972	0	1.4454	0.037007	NA
hsa-miR-23a-5p	3.877873	0	1.4563	0.037146	NA
hsa-miR-125a-5p	543.5221	187.4418	1.2164	0.037688	NA
hsa-miR-499a-5p	126.6113	41.07934	1.2463	0.040416	NA
hsa-miR-499b-3p	125.6438	41.07934	1.2387	0.041507	NA
hsa-miR-125b-2-3p	8.459143	1.865813	1.3897	0.042802	NA
hsa-miR-125a-3p	12.62621	1.487911	1.4403	0.045977	NA
hsa-miR-32-3p	101.2775	32.47671	1.2287	0.048942	NA
Down-regulated					
hsa-miR-30a-5p	529.2648	1,956.122	-1.4991	0.009967	0.12348
hsa-miR-6753-3p	0.869689	5.90425	-1.7068	0.013047	NA
hsa-miR-151a-3p	10,040.77	22,962.51	-1.0725	0.013772	0.13647
hsa-miR-3133	15.02586	34.93526	-1.0734	0.020626	NA
hsa-miR-1288-3p	2.661609	9.564785	-1.3787	0.026051	NA
hsa-miR-4753-5p	0	2.201217	-1.5165	0.031505	NA
hsa-miR-30c-2-3p	0.269812	3.661317	-1.5266	0.034302	NA
hsa-miR-449c-5p	5.607974	13.67431	-1.0701	0.042598	NA
hsa-miR-6815-5p	6.328357	14.35981	-1.0136	0.046006	NA
hsa-miR-1303	53.95024	123.3373	-1.0132	0.048477	NA

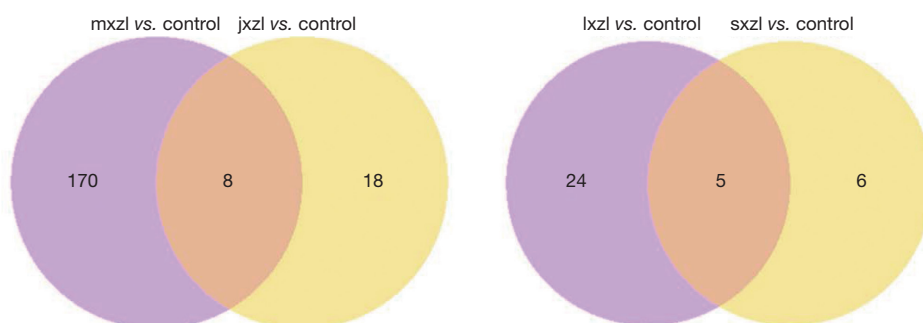


Figure S5 The venny plot between different comparison was shown.