JJBS

Volume 12, Number 4,September 2019
ISSN 1995-6673
Pages 477 - 485

Jordan Journal of Biological Sciences

Next Generation Approach of Biological Data Analysis through *in silico* Prediction and Analysis of Small ncRNA-RNA Interactions in Human Tissues

Imon Chakraborty[1] and Shohini Chakraborty[2,*]

[1] *School of Bioscience and Engineering, Jadavpur University, Kolkata, WB, 700032;* [2] *Botany Department, University of Gour Banga, Malda, 732103, West Bengal, India*

Received November 25, 2018; Revised January 1, 2019; Accepted January 15, 2019

Abstract

This study investigates the analysis of small non-coding RNAs based on the identification of micro RNAs (miRNAs) and other small non-coding RNAs (sncRNAs) to study tissue-specific interactions which lead to specific diseases. The RNA-RNA interaction revealed the tissue specificity to understand the impact on human health. miRNAs were identified using miRDeep2, and other sncRNAs were identified using the Database of Small Human non-coding RNAs (DASHR) which gave the information on sncRNAs and a high throughput validation. Moreover, the RNA Association Interaction Database (RAID) gave the profiled interactions data of sncRNAs for six different healthy human tissues. The results show the identification of a total of 777 miRNAs and their level of expression in each tissue. In the DASHR analysis, 1560 small ncRNAs were found in all of the tissues. miRNAs and other small ncRNAs expressions were visualized using the R software-based Heatmap. RAID detailed 502 interactions with 229 tissue-specific interactions in the liver from 85 miRNAs and 102 small ncRNAs. The *In silico* analysis was suggested to describe the non-coding RNAs tissue-specific interactions during healthy conditions which might be associated with disease conditions.

Keywords: Biological data analysis, *In-silico*, miRNA, non-coding RNA, RNA-RNA interaction.

1. Introduction

Ribonucleic acid (RNA) is one of the three major biological macromolecules that are fundamental for all known forms of life (along with DNA and proteins) (Baker *et al.*, 2003). RNA is a significant macromolecule for cells because it relays information encoded in DNA within the cell as a coding and non-coding segment. The coding part of RNA molecule is translated into proteins according to the RNA's instructions followed by a translation mechanism (Hutvágner *et al.*, 2001).

A non-coding RNA is an RNA molecule that is transcribed from DNA, but is not translated into proteins (Anastasiadou *et al.*, 2018). In general, non- coding RNAs regulate gene expression at the transcriptional and post-transcriptional levels. These non-coding RNAs can be divided into two main groups: the small ncRNAs (small ncRNAs) and the long non-coding RNAs. The major classes of small ncRNAs are micro RNAs (miRNAs), small nucleolar RNAs (snoRNAs), short interfering RNAs (siRNAs) and PIWI interacting RNAs (piRNAs) (Esteller, 2011).

Small ncRNAs play a role in heterochromatin formation, histone modification, DNA methylation targeting, and gene silencing. Long non-coding RNAs are responsible for chromatin remodeling, transcriptional regulation, post-transcriptional regulation, and as precursors for siRNAs. miRNAs are exigent regulators for many basic cellular processes (Pritchard *et al.*, 2012), hence, a little alteration or deviation in the amount, timing or location of miRNA expression can have larger effects on cell and organisms' growth. The analyses of primary, precursor, and mature miRNA levels as well as the identification and characterization of miRNA targets are crucial for determining the biogenesis or functional changes of miRNAs due to a particular mutant or disease (Van Wynsberghe *et al.*, 2011).

Small nucleolar RNAs (snoRNAs) primarily guide chemical modifications of other RNAs. They are also associated with small nucleolar ribonucleoprotein particles (snoRNPs) where they serve as guide molecules in the post-transcriptional modification of ribosomal RNA (rRNA) and small nuclear RNA (snRNA). siRNAs function in an analogous way to miRNAs to mediate post-transcriptional gene-silencing as a result of mRNA degradation. piRNAs are involved in chromatin regulation and the suppression of transposon activity in germline and somatic cells. Many long non-coding RNAs can form a complex with chromatin-modifying proteins, and recruit their catalytic activity to specific sites in the genome by altering chromatin states and influencing gene expression (Martens-Uzunova *et al.*, 2013).

In this study, the *in-silico* (an expression via computer simulation) approach for the identification of miRNAs from a large pool of sequenced transcripts from deep

* Corresponding author e-mail: shohini.chakraborty@gmail.com.

sequencing experiments has been carried out. The study has used miRDeep2 (Friedländer *et al*., 2008), which implements a probabilistic model of miRNA biogenesis to score compatibility of the position and frequency of sequenced RNA with the secondary structure of the miRNA precursor. miRDeep2 can accurately identify known and novel miRNAs, distinguish miRNAs from other small RNAs, and give the high throughput validation with little consumption of time and memory combined with a user-friendly interactive graphic output.

Another aspect of this study involves the identification of all the sncRNAs in five different human tissues for which the DASHR (Database of Small Human non-coding RNAs) (Leung *et al*., 2015) has been used. This is a database containing comprehensive information on small human non-coding RNAs and mature sncRNA products. Next, possible RNA-RNA interactions in the small RNA sequencing data using RNA Association Interaction Database (RAID) (Zhang *et al*., 2014) have been identified and profiled based on their occurrence in different tissues. Consequently, RNA-RNA interactions or differentially-expressed non-coding RNAs that may be tissue-specific (if any) have been determined.

2. Materials and Methods

The current study has taken different datasets of human tissues (Frontal Cortex, Cerebellum, Heart, Kidney, Testis and Liver) from the public data repository of Gene Expression Omnibus (GEO). The sequencing has been done by using hg19 sequencing platform. Finally the data were deposited in the GEO database by the researchers. After taking the datasets from the GEO, the researchers have identified miRNAs present in the datasets using miRDeep2 modules. Then, the identification of small ncRNAs present in the datasets using DASHR has been done. In the next stage, possible RNA-RNA interactions have been identified using RAID, and then visualization was performed. For the Heatmap visualization, raw count data have been loaded to the R session. Then, the data were normalized using DSeq2 (Love *et al*., 2014), and the variance stabilizing transformation was performed. Finally, a heatmap was generated using the complex Heatmap R package. Figure 1 shows the flowchart of the whole process, and each of the steps are explained in details with further flowcharts afterwards.

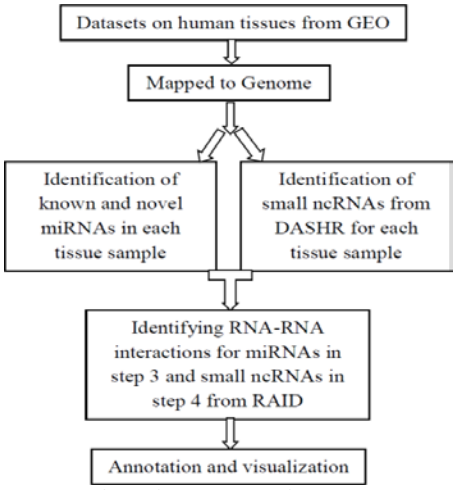

Figure 1. Workflow for the identification and analysis of sncRNAs. Datasets are taken from Geo database and mapped to genome. In the next step, the identification of miRNAs and other small ncRNAs using miRDeep2 and DASHR was performed respectively. Finally RNA-RNA interaction has been performed using DASHR. In the last step, visualization was performed using *Heatmap and venn-diagram (Heberle et al., 2015).*

2.1. Identification of miRNAs Using miRDeep2

miRNAs identification using miRDeep2 has been categorized in two modules, namely "Mapper" and "miRDeep2". The Mapper module is for preprocessing the raw illumina output; it reads it in fastq format, and maps it to the reference genome. The miRDeep2 module is run to identify novel and known miRNAs in high-throughput sequencing data analysis. Details of the miRDeep2 analysis are described in Figure 2.

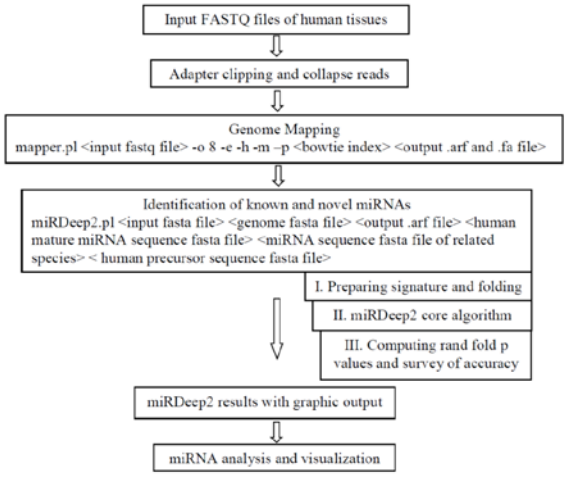

Figure 2. Workflow for the identification of miRNAs from small RNA sequence datasets. Here '–o' denotes the number of threads used for bowtie, '-e' refers to the input file in fasta format, '-h' refers to parsing the fasta format, '-m' refers to the collapsed reads and '-p' denotes the map to genome. Output contains .arf (advanced recording file) and .fa (fasta) files. In genome mapping, bowtie runs to create the index of genome of any size. The genome mapped output arf and fasta file have been used in the identification of miRNAs. miRDeep2 runs in several steps to give high throughput sequence analysis with a graphical output. Few of the steps are linked in the identification steps. The final analyzed results have been visualized using the R software-based Heatmap.

2.2. Identification of Small ncRNAs Using DASHR

The collapsed reads are mapped against the DASHR (Database of Small Human non-coding RNAs) database to identify the highly abundant sncRNAs from any specific datasets. The detailed steps for the identification of small ncRNAs are shown in Figure 3.

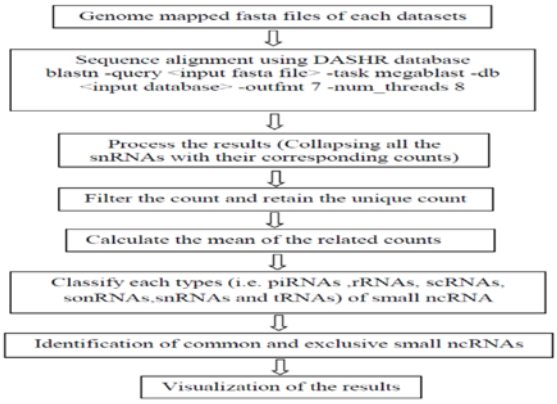

Figure 3. Workflow for other Small ncRNAs identification using RNA sequencing datasets. In this flowchart, '-task megablast' refers to very similar sequence, '–db' refers to the database, '-outfmt' denotes out format and the '-num threads' indicates the num of threads. The Balstn sequence alignment results were processed in three steps for a proper identification of common and exclusive small ncRNAs using venn diagram. The final analyzed output has been visualized using the R software-based Heatmap.

2.3. Identification of RNA-RNA Interaction Using RAID

The RNA-RNA interaction gives the information of any specific miRNAs interaction with other small ncRNAs. The process identified the small ncRNAs, which have been described below in the flowchart (Figure 4).

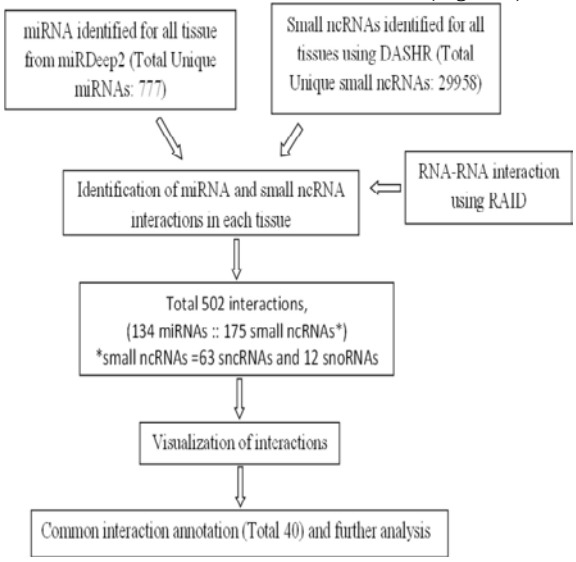

Figure 4. Workflow for RNA-RNA interaction and analysis. The identified total unique miRNAs and small ncRNAs are combined in a single file, and the RNA-RNA interaction was performed using the RAID database (database provides RAID template for comparison between the interactions). The interaction output has been visualized using Heatmap. Finally 40 interactions have been analyzed and annotated where both interactors were present in the sample. The forty common interactions were obtained from 11 miRNAs and 27 small ncRNAs.

3. Results

After trimming the adapters from the raw fastq files, the adapter trimmed file has been used for all RNAs identification and analysis as described below. The length distribution was calculated using the awk command from the fastq files of each tissue. The subsequent length distribution plot is shown in Figure 5.

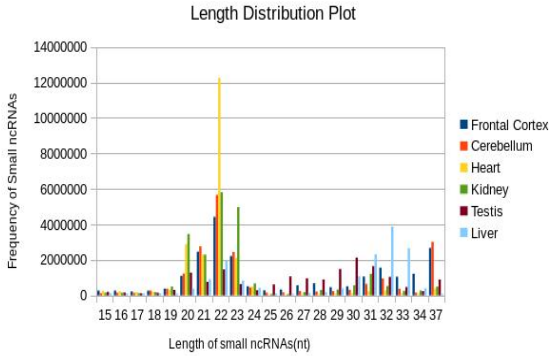

Figure 5. Shows the length-distribution plot of all datasets. The first peak shown around 22nt indicates the presence of miRNAs. Another peak shown around 32nt denotes the enriched piRNAs in the datasets.

This comparison plot shows a peak around 22nt. It is evident from the plot that the sample libraries were enriched for miRNA. There is another peak at around 30 to 33 nucleotides which denotes the presence of piRNAs as well. In comparison to other samples, heart tissue has twelve million small ncRNAs of twenty-two nucleotides (generally miRNAs). The testis dataset has a diverse high range of small ncRNAs between twenty-six and thirty nucleotides. On the basis of these outcomes, further identification and analysis must take place.

Table 1. Details of raw datasets count and mapping for each tissue.

Tissue Type with GEO ID	Raw Count	Unique Raw Count	Mapping Count	Mapping Percentage
Frontal Cortex GSM995300	2,25,44,135	32,81,425	12,63,379	38.5%
Cerebellum GSM995301	2,00,82,273	29,57,982	6,41,519	21.69%
Hear GSM995302	2,27,67,169	4,43,156	98,810	22.29%
Kidney GSM995303	2,26,16,113	5,67,035	1,61,008	28.39%
Testis GSM995304	1,68,38,653	39,94,670	25,41,858	63.63%
Liver GSE57381	1,82,27,059	8,09,611	4,05,632	50.1%

In Table 1, the first column contains the details of tissue types with their GEO accession ID corresponding to row count (i.e. the number of reads coming from fastq files), the unique raw count (i.e. number of collapsed reads), and the mapping count (i.e. number of reads mapped to the genome). The mapping percentage has been calculated using the difference between raw count and mapping to the mapping count, as shown in the last column. In testis, the mapping percentage is 63.63 %, whereas the percentage is much less in cerebellum (21.69 %). Here, four uninfected liver tissues (GSM1381481, GSM1381482, GSM1381483 and GSM1381484) have been taken and merged to calculate the read counts and the mapping percentage and to carry out further analysis.

3.1. Analysis of miRNAs

A total of 777 miRNAs have been identified using miRDeep2. The analysis of miRDeep2 output, resulted in the count of novel miRNAs, known miRNAs and unique miRNAs on the basis of precursor sequence followed by highly expressed miRNAs in each tissue, as shown in Table 2.

Table 2. Number of novel and known unique miRNA identified on the basis of precursor sequence.

Tissue Types	Novel miRNAs (Count)	Known miRNAs (Count)	Unique miRNAs on precursor sequence(count)	Highest expressed miRNA (Count)
Frontal Cortex	187	518	496	hsa-let-7b-5p(1,886,870)
Cerebellum	141	477	450	hsa-let-7b-5p(3,608,413)
Heart	68	464	441	hsa-miR-1-3p (7,019,605)
Kidney	65	459	434	hsa-miR-143-5p(3,653,125)
Testis	85	382	359	hsa-miR-143-5p(1,553,310)
Liver	156	425	325	hsa-miR-192-5p (723,267)

As for data analysis, the highest number of known miRNAs was found in the frontal cortex tissue. miRNA, hsa-miR-1-3p (count 7019605) was found in the heart which has the highest expression level in comparison to all miRNAs. Novel miRNAs refer to the unreported miRNAs in humans but reported in some other species or predicted using the characteristic features of miRNA. Known miRNAs denotes reported miRNAs annotated in miRBase (Kozomara and Griffiths-Jones, 2013) for humans. On the basis of miRDeep2 results, highly expressed miRNAs for each tissue have been listed in last column with counts. hsa-let-7b-5p is highly expressed in both brain tissues (i.e. frontal cortex and cerebellum). Also miRNA hsa-miR-143-5p has a high expression level in the kidney tissues as well as testis. On the other hand, hsa-miR-192-5p is highly expressed in liver tissues while showing a low expression level in other tissues.

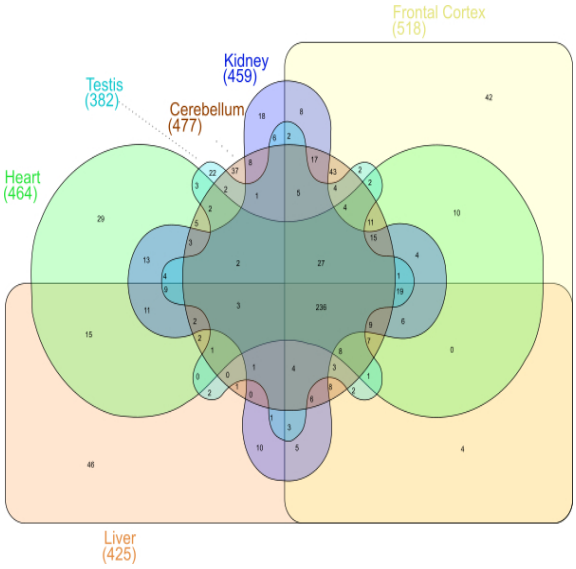

Figure 6. shows the number of common and exclusive miRNAs in each tissue. Visualization shows that half of the miRNAs are present in all of the tissues as for the count of total miRNAs present in each tissue. In the testis, the number of miRNAs is relatively less compared to other tissues. The highest number of miRNAs was found in the frontal cortex (i.e. miRNAs count 518).

A total of 777 miRNAs were sequenced, and 236 were present in all of the tissues (as shown in Figure 6). Forty-six exclusive miRNAs were found in the liver tissue. Another forty-three miRNAs were commonly found in brain tissues (i.e. frontal cortex and cerebellum) which are exclusive in terms of other tissues. Table 3 shows the exclusive miRNAs presence in each tissue.

In this study the number of exclusive miRNAs in the liver is higher (exclusive count 46) than in other tissues. On the other hand, the kidney has only eighteen exclusive miRNAs.

In this heatmap, the color scheme has been defined by 'RColorBrewer' and the palette used 'colorRampPalette' OrRd (i.e. orange- Red: 9 colors distributed in 100 colors shades for visualization) (Figure 7).

Table 3. Exclusive miRNAs shown for each tissue type

Tissue type	Exclusive miRNAs
Frontal Cortex	miR-2682-5p,-320e,-6734-5p,-448,-1226-5p,-7156-5p,-6874-5p,-4454,-5002-5p,-5680,-3663-5p,-7854-3p,-522-5p,-522-3p,-519b-5p,-3124-5p,-6864-5p,-4727-5p,-3139,-5688,-4753-5p,-3193,-6748-5p,-516a-5p,-4723-5p,-3173-5p,-6890-5p,-4785,-4773,-4721,-4506,-4501,-4450,-3181,-2278,-6846-5p,-6739-5p,-609,-520d-5p,-4795-5p,-3175,-6817-5p
Cerebellum	miR-6499-5p,-4429,-3159,-3158-3p,-1229-5p,-6762-5p,-6885-5p,-7855-5p,-6804-5p,-6808-5p,-6816-3p,-3692-5p,-3190-3p,-892b,-1234-3p,-3612,-4448,-4646-5p,-6820-5p,-892a,-3679-5p,-4665-5p,-6851-5p,-4685-3p,-4685-5p,-5191,-5581-5p,-6736-5p,-6786-5p,-3167,-4423-5p,-4470,-4644,-4757-5p,-6774-5p,-6831-5p,-6866-5p
Heart	miR-29b-3p,-208b-5p,-1285-3p,-1285-5p,-378h,-3605-5p,-208a-5p,-302c-5p,-302d-5p,-7846-3p,-4444,-4633-5p,-4786-5p,-4508,-3129-5p,-4670-3p,-5684,-3126-5p,-367-5p,-4490,-4766-5p,-548ao-5p,-6717-5p,-1245b-5p,-1265,-5003-3p,-5003-5p,-6811-5p,-6827-5p
Kidney	miR-934,-559,-1275,-4709-3p,-4709-5p,-3128,-4473,-4724-5p,-1263,-4647,-890,-4679,-4729,-6878-5p,-1257,-5701,-603,-643
Testis	miR-513b-5p,-520f-5p,-1323,-518d-5p,-517-5p,-517b-3p,-670-5p,-7162-5p,-4433b-3p,-523-5p,-525-5p,-3923,-515-5p,-518e-3p,-147b,-3660,-498,-518a-5p,-548a-5p,-520e,-5186,-4760-5p
Liver	miR-99b-5p,-484,-331-5p,-664a-5p,-5589-5p,-641,-548av-3p,-4775,-5690,-6715a-3p,-2467-5p,-211-5p,-7705,-3614-5p,-5590-5p,-4686,-1304-5p,-4492,-5696,-548y,-7974,-3074-5p,-548al,-600,-1228-5p,-4645-3p,-5010-3p,-548f-5p,-802,-1255a,-2115-5p,-3130-5p,-3161,-3664-5p,-3681-5p,-371b-5p,-4642,-4687-5p,-4699-5p,-4750-5p,-496,-548s,-6769b-5p,-6837-5p,-6852-5p,-7976

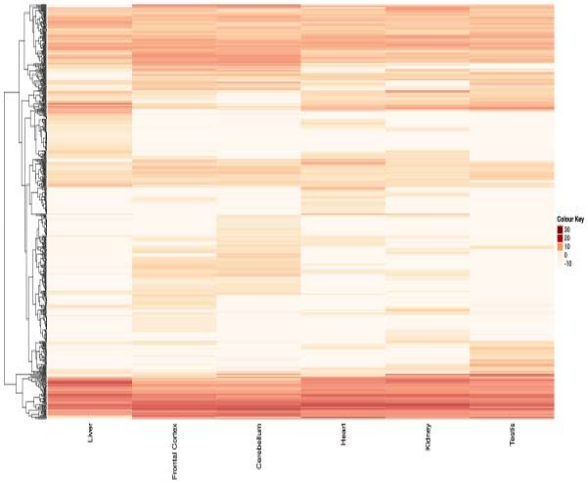

Figure 7. Visualization of miRNAs expression for different tissues using Heatmap. In this heatmap, the dark to light color shows the level of high to low expression. For example, in the bottom band of each column, a dark band is shown across all tissues except for the liver.

This visualization concludes that few miRNAs are expressed in all of the tissues, but they are not expressed in the liver i.e. hsa-miR-34b-5p is present in all of the tissues except the liver. An opposite scenario is also seen where some miRNAs (like miR-3681, -496,-600 etc.) are expressed in the liver, but not in other tissues. The miRNA expression in the two brain samples is similar which is clearly understandable in the heatmap visualization.

3.2. Analysis of Small ncRNAs

DASHR provides the most comprehensive information on small ncRNAs. DASHR analysis gives a huge number of small ncRNA-processing information output. In this study, the researchers have processed the results and excluded the miRNAs for further analysis to carry out interaction studies. The classification of the small ncRNAs obtained from the DASHR results is shown in Table 4.

Small ncRNAs	Frontal Cortex (count)	Cerebellum (Count)	Heart (Count)	Kidney (Count)	Testis (Count)	Liver (Count)
piRNAs	7,058	8,335	1,561	1,517	25,843	4,547
rRNAs	118	107	91	93	116	138
scRNAs	70	111	13	15	47	50
snoRNAs	303	277	226	290	285	383
snRNAs	803	481	435	445	920	736
tRNAs	472	481	410	433	534	520
Total RNAs	8,824	9,792	2,736	2,793	27,745	6,374

Table 4. Shows the filtered results of DASHR analysis

Here, small ncRNAs are categorized in six types (i.e. piRNAs, rRNAs, scRNAs, snoRNAs, snRNAs, and tRNAs). In all of the tissues, the piRNAs count was highest in comparison to other types of small ncRNAs. Small ncRNAs are classified and the corresponding counts for each tissue have been shown in Table 4. The results show that testis has a huge number of piRNAs among all the tissues as piRNAs are primarily involved in preserving genomic integrity in germline cells. The number of scRNAs is relatively less in all of the tissue in comparison to all other categories of small ncRNAs.

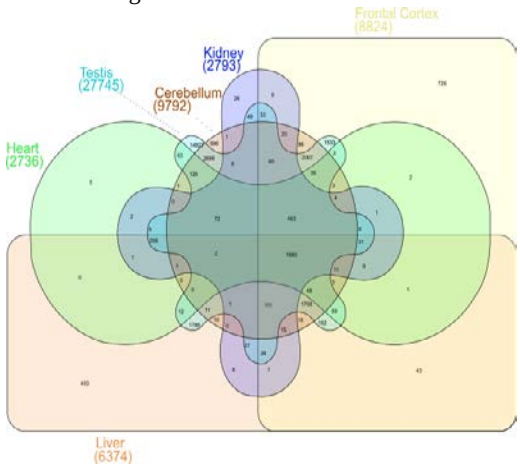

Figure 8. Shows the number of common and exclusive other small ncRNAs in each tissue. The highest number of small ncRNAs (i.e. 27,745) was found in the testis. In this analysis, only five exclusive small ncRNAs were found in the heart tissue which is very low compared with the other tissues.

In this analysis, 1,560 small ncRNAs have been found as common to all tissues. The testis tissue shows a large number of small ncRNAs (exclusive count 14,827) as the presence of huge amount of piRNAs. On the other hand, the heart and kidney tissues show a lesser number of small ncRNAs (Figure 8).

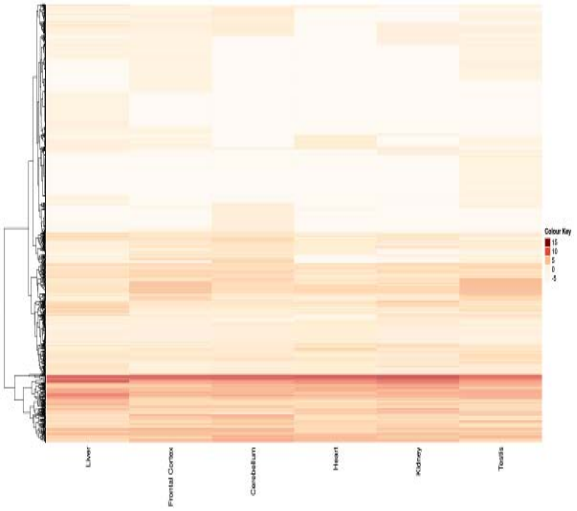

Figure 9. Visualization of other small ncRNAs expression for different tissues using Heatmap. The same line dark band shows the few 1,560 common small ncRNAs which were present in all of the tissues.

From this heatmap visualization, piRNAs have been excluded. Here, the dark band across all of the tissues means a high expression level of particular small ncRNAs among all the samples. The uppermost part of the heatmap shows a lesser expression level for each tissue (Figure 9).

3.3. Analysis of Small ncRNA-RNA Interactions

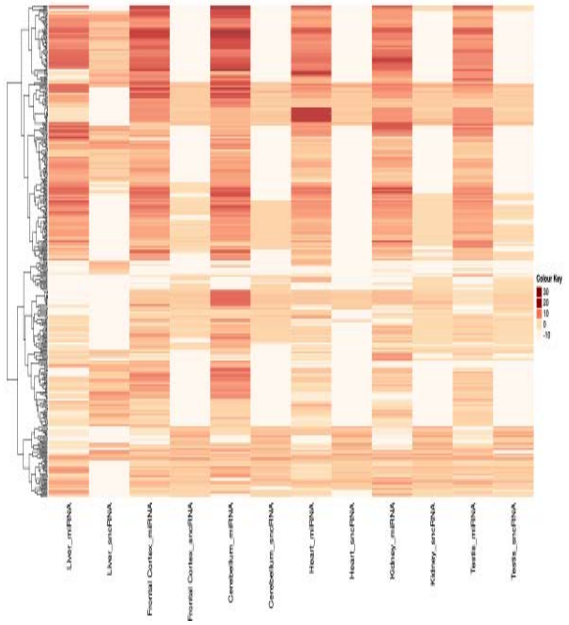

Figure 10. Shows the comparison of the expression of RNA-RNA interaction for different tissues. Here in this interaction, the brain tissue shows several common expressions.

This heatmap shows all interactions between miRNAs and small ncRNAs (Figure 10). As shown in this heatmap, the bottom portion has some interactions where both interactors are present in high numbers except for the liver tissues. For example, hsa-miR-429 interacts with SNORD116-21; this has been found in each tissue sample except for the liver. Some interactions (such as when hsa-miR383-5p interacted with RNA5SP165 and SNORD111B) have also been found only in the liver tissues, but were not captured in other tissues.

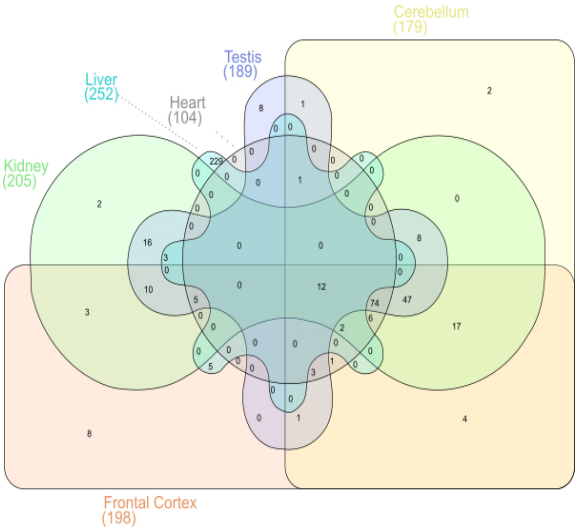

Figure 11. Common and exclusive RNA-RNA interactions identified for all of the tissues. Twelve common interactions have been found among all interactions. Another interesting fact is that 229 exclusive interactions were captured in the testis tissue (highest exclusive interaction), whereas seventy-four interactions were found in all of the tissues with the exception of the liver.

Twelve out of 502 interactions were captured in all of the tissue samples. It has been seen that miRNAs interacted with either snoRNAs or sncRNAs (Figure 11). Out of 502 interactions, 134 miRNAs interacted with 175 small ncRNAs (i.e. 63 sncRNAs and 112 snoRNAs).

Table 5. Presents a list of common miRNAs interactions with other small ncRNAs.

miRNAs	Small ncRNAs
hsa-miR133b	SNORD124, SNORD114-17, SNORD114-3, SNORD113-4, SNORD116-13, SNORD114-14,SNORD114-12
hsa-miR145-5p	SNORD127, SNORA3,SNORD65,SNORD6, SNORD124
hsa-miR183-5p	SNORA46, SNORD123, SNORA48, SCARNA1
hsa-miR190b	SNORA49, SNORD121A, SNORA24, SNORA15, SNORD124
hsa-miR224-5p	SNORA24, SNORD115-6, SNORD115-21, SNORA60, SNORD123, SNORD115-3, SNORD116-29, SNORD115-10
hsa-miR-243p	SNORD124
hsa-miR26a-5p	SNORD115-35, SNORA60, SNORD123
hsa-miR26b-5p	SNORD115-35, SNORD123, SNORA60
hsa-miR374a-5p	SNORD127
hsa-miR374b-5p	SNORD127
hsa-miR873-5p	SNORD17, SNORD123

A total of forty interactions took place using these miRNAs and other small ncRNAs. The forty interactions were obtained from twelve miRNAs and twenty-seven small ncRNAs (as shown in Table 5). hsa-miR-224-5p has interacted with eight other small ncRNAs which is the highest number of interactions.

In the liver tissue, 229 specific interactions had been found, whereas in the heart tissue, no specific interaction was found. Out of the total 502 interactions, only few were tissue-specific distributed as follows: eight in in the frontal cortex, two in the cerebellum, two in the kidney, and eight in the testis. Most interestingly, 229 specific interactions were found in the liver tissue. The 229 interactions were obtained from eighty-five miRNAs and 102 small ncRNAs (as shown in Table 6).

4. Discussions

This study presents an investigation of RNA-RNA interactions using healthy human tissues. Individual small ncRNAs such as miRNAs, snoRNAs have been identified. Furthermore, their interaction characteristics were explored for the sake of understanding their molecular level and functional impact on human health.

A total of 777 unique miRNAs have been found, out of which, 3 miRNAs, namley hsa-let-7a-5p, hsa-let-7f-5p, and hsa-miR-143-5p, were present among all tissues with a high expression level. hsa-miR-143-5p has the highest count in the kidney and testis samples as shown in Table 1.

Table 6. Presents details of tissue-specific interactions:

Tissue types	Specific Interactions
Frontal Cortex	miR-30a-5p:SNORA11C, miR-30e-5p:SNORA11C, miR-448:SNORD113-4, miR-30d-5p:SNORA11C,miR-30c-5p:SNORA11C, miR-30b-5p:SNORA11C, miR-491-5p:SNORA11C, miR-3139:SNORA26
Cerebellum	miR-3167:SNORA31, miR-3167:SNORA55
Heart	No specific interaction
Kidney	miR-551a:SCARNA23, miR-205-5p:SNORA33
Testis	miR-520e:SNORD116-19,miR-520e:SNORD116-21,miR-520e:SNORA37, miR-340-5p:SNORA22,miR-185-5p:SNORA22,miR-202-5p:SNORA25, miR-218-5p:SNORA22, miR-520e:SCARNA1
Liver	229 specific interactions (miRNAs : snoRNAs) miR-504-5p:RNA5SP301,miR-504-5p:RNA5SP497,miR-30c-5p:RNY3,miR-429:SNORA36A,miR-9-5p:RNY4P14,miR-383-5p:RNA5SP165,miR-383-5p:SNORD111B,miR-15b-5p:SNORA63,miR-149-5p:SNORA15,miR-320c:SNORA44,miR-504-5p:RNA5SP324,miR-181d-5p:RNA5SP95,miR-429:SNORA56,miR-10b-5p:RNA5SP175,miR-181d-5p:SNORD116-12,miR-181d-5p:SNORA13,miR-504-5p:RNA5SP242,miR-383-5p:RNA5SP48,miR-181c-5p:SNORA13,miR-496:SNORA15,miR-383-5p:RNA5SP366,miR-30b-5p:RNY3,miR-181c-5p:RNA5SP221,miR-181c-5p:RNA5SP203,let-7e-5p:RNA5SP221,miR-497-5p:SNORA63,miR-1-3p:SNORD116-15,miR-379-5p:RNA5SP221,miR-217:SNORD98,miR-135b-5p:SNORD12B,miR-107:SNORA19,miR-181a-5p:SNORA13,miR-320d:SNORA44,miR-383-5p:RNA5SP296,miR-30c-5p:SNORA11B,miR-125a-5p:SNORA79,miR-504-5p:RNA5SP450,let-7e-5p:SNORD11,miR-196b-5p:SNORD11,miR-98-5p:SNORD11,let-7g-5p:SNORD11,miR-383-5p:RNA5SP450,miR-148a-3p:SNORD17,miR-486-5p:SCARNA16,miR-383-5p:RNA5SP382,miR-383-5p:RNA5SP429,miR-361-5p:SNORD111B,miR-361-5p:SNORA60,miR-148a-3p:SNORA8,miR-122-5p:SNORA60,miR-504-5p:RNA5SP387,miR-181b-5p:SNORA13,let-7f-5p:RNA5SP221,miR-150-5p:SNORA36A,miR-7-5p:SNORA15,miR-193b-3p:SNORA6,miR-1-3p:SNORD116-14,miR-204-5p:SNORA15,miR-340-5p:SNORD115-4,miR-190b:SNORA15,miR-101-3p:SNORA32,miR-383-5p:RNA5SP321,miR-193b-3p:SNORD12B,miR-30a-5p:SCARNA15,miR-31-5p:SNORA62,miR-106a-5p:SNORD11,miR-504-5p:RNA5SP143,miR-30b-5p:SNORA60,miR-199b-5p:SNORA61,miR-101-3p:SNORA15,miR-103a-3p:SNORA19,miR-30a-5p:SNORA60,miR-181a-5p:RNA5SP95,miR-181b-5p:RNA5SP221,miR-1-3p:SNORA77,miR-504-5p:RNA5SP283,miR-193b-3p:SNORD12,miR-379-5p:RNA5SP203,miR-383-5p:RNA5SP352,miR-383-5p:RNA5SP276,miR-383-5p:RNA5SP86,miR-182-5p:SNORA4,miR-30a-5p:SNORA11B,miR-30b-5p:RNY3P6,miR-873-5p:SNORD17,miR-218-5p:SNORA49,miR-383-5p:RNA5SP272,miR-383-5p:SNORD17,miR-140-5p:SNORD89,miR-181a-5p:RNA5SP203,let-7i-5p:RNA5SP221,miR-504-5p:RNA5SP134,miR-361-5p:SNORD12,miR-181d-5p:RNA5SP221,miR-411-5p:RNA5SP443,miR-150-5p:SNORA38,miR-504-5p:RNA5SP52,miR-499a-5p:RNY1P5,miR-138-5p:SCARNA15,miR-20b-5p:SNORD11,miR-196a-5p:SNORD11,miR-383-5p:RNA5SP145,miR-138-5p:SNORA60,let-7c-5p:SNORD11,miR-30e-5p:RNY3P6,miR-499a-5p:SNORD114-9,miR-504-5p:RNA5SP191,miR-504-5p:RNA5SP392,miR-186-5p:SNORD111B,miR-10a-5p:RNA5SP134,let-7f-5p:SNORD11,miR-504-5p:RNA5SP174,miR-218-5p:SNORD68,miR-125b-5p:SNORA6,miR-194-5p:SNORA36A,miR-383-5p:RNA5SP283,miR-181b-5p:RNA5SP203,miR-504-5p:RNA5SP388,miR-383-5p:RNA5SP259,miR-383-5p:RNA5SP263,miR-153-3p:SNORD12,miR-145-5p:SNORD65,miR-217:SNORA32,miR-9-5p:SNORD4B,miR-133b:SNORD116-13,miR-183-5p:SNORA46,miR-375:SNORD116-11,miR-101-3p:RNA5SP121,miR-383-5p:RNA5SP53,miR-153-3p:SNORA9,miR-361-5p:SNORD12B,miR-383-5p:RNA5SP355,miR-30c-5p:RNY3P6,let-7c-5p:RNA5SP221,miR-504-5p:RNA5SP221,miR-383-5p:RNA5SP215,miR-383-5p:RNA5SP368,let-7i-5p:SNORD11,miR-383-5p:RNA5SP338,miR-320b:SNORA44,miR-10a-5p:SCARNA16,miR-379-5p:SCARNA15,miR-30d-5p:SCARNA15,miR-504-5p:RNA5SP514,miR-30d-5p:SNORA60,miR-383-5p:RNA5SP284,miR-135b-5p:SNORA49,miR-383-5p:SNORD88C,miR-224-5p:SNORD115-3,miR-30e-5p:SNORA60,miR-183-5p:SNORA48,miR-16-5p:SNORA63,miR-504-5p:RNA5SP53,miR-504-5p:RNA5SP152,miR-34c-5p:RNA5SP248,miR-98-5p:RNA5SP221,miR-504-5p:RNA5SP298,miR-320a:SNORA44,miR-122-5p:SNORA4,miR-34a-5p:SNORA60,miR-199a-5p:SNORA61,miR-383-5p:SNORD12,miR-485-5p:SCARNA15,miR-125a-5p:SNORA6,miR-10b-5p:SCARNA16,miR-30d-5p:RNY3,miR-190b:SNORA49,miR-30d-5p:SNORA11B,miR-485-5p:SNORA32,miR-30e-5p:RNY3,miR-181b-5p:SNORD116-12,miR-135a-5p:SNORD12B,miR-7-5p:SNORA60,miR-383-5p:RNA5SP493,miR-383-5p:RNA5SP191,miR-187-3p:SNORA38,let-7g-5p:RNA5SP221,miR-383-5p:RNA5SP477,miR-190a-5p:SNORA49,miR-125b-5p:SNORA79,miR-17-5p:SNORD11,miR-30b-5p:SCARNA15,miR-30c-5p:SCARNA15,miR-543:SNORA60,miR-149-5p:SNORD12B,miR-224-5p:SNORD115-6,miR-383-5p:RNA5SP202,miR-34a-5p:RNA5SP248,miR-504-5p:RNA5SP469,miR-26a-5p:SNORA60,miR-30a-5p:RNY3P6,miR-34c-5p:SNORA60,miR-383-5p:RNA5SP298,miR-224-5p:SNORD115-10,miR-190a-5p:SNORA15,miR-26b-5p:SNORA60,miR-223-3p:SNORD4B,miR-30a-5p:RNY3,miR-30d-5p:RNY3P6,miR-1-3p:SNORD116-12,miR-504-5p:RNA5SP259,miR-33b-5p:SNORA6,miR-181a-5p:SNORD116-12,let-7a-5p:RNA5SP221,miR-182-5p:SNORA56,miR-181c-5p:SNORD116-12,miR-224-5p:SNORA60,miR-30e-5p:SNORA11B,miR-379-5p:RNA5SP167,miR-15a-5p:SNORA63,miR-383-5p:RNA5SP424,let-7a-5p:SNORD11,miR-181b-5p:RNA5SP95,miR-181a-5p:RNA5SP221,miR-10b-5p:RNA5SP134,miR-135a-5p:SNORA49,miR-30c-5p:SNORA60,miR-383-5p:RNA5SP388,miR-30b-5p:SNORA11B,miR-383-5p:RNA5SP442,miR-504-5p:RNA5SP135,miR-10a-5p:RNA5SP175,miR-383-5p:RNA5SP65,miR-181c-5p:RNA5SP95,miR-181d-5p:RNA5SP203,miR-153-3p:SNORA77,miR-340-5p:SNORD116-11,miR-217:SCARNA16,miR-30e-5p:SCARNA15

All of these miRNAs have been already reported in colorectal neoplasia (Michael *et al*., 2003). hsa-let-7a and hsa-let-7f have been reported to have an association with human cervical cancer (Lui *et al*., 2007). hsa-miR-143 has also been reported to induce the apoptosis in prostate cancer (Ma *et al*., 2017). hsa-miR-1-3p has been reported to have a great impact on cardiovascular diseases or any kind of cardiac disorders. miRNA-1 (has-miR-1) has potential anti-tumorigenic properties in lung cancer cells

(Nasser *et al*., 2008). hsa-miR-1 is also reported to down-regulate non-coding RNA in urothelial cancer associated with bladder cancer. Small ncRNA, SNORD85 was found in the prefrontal cortex and was strongly down-regulated in schizophrenia (Smalheiser *et al*., 2014).

A total of 777 unique miRNAs have been found, out of which, 3 miRNAs, namley hsa-let-7a-5p, hsa-let-7f-5p, and hsa-miR-143-5p, were present among all tissues with a high expression level. hsa-miR-143-5p has the highest count in the kidney and testis samples as shown in Table 1. All of these miRNAs have been already reported in colorectal neoplasia (Michael *et al*., 2003). hsa-let-7a and hsa-let-7f have been reported to have an association with human cervical cancer (Lui *et al*., 2007). hsa-miR-143 has also been reported to induce the apoptosis in prostate cancer (Ma *et al*., 2017). hsa-miR-1-3p has been reported to have a great impact on cardiovascular diseases or any kind of cardiac disorders. miRNA-1 (has-miR-1) has potential anti-tumorigenic properties in lung cancer cells (Nasser *et al*., 2008). hsa-miR-1 is also reported to down-regulate non-coding RNA in urothelial cancer associated with bladder cancer. Small ncRNA, SNORD85 was found in the prefrontal cortex and was strongly down-regulated in schizophrenia (Smalheiser *et al*., 2014).

On the other hand, a total of 502 interactions have been found from the RAID database for the identified miRNAs and other small ncRNAs in this study as shown in Figure 4. Out of 502 interactions, forty have been identified where both interactors were sequenced in the same sample. Twelve miRNAs interacted with 27 other small ncRNAs giving rise to forty different interactions. miRNAs have been reported to interact with different small ncRNAs and showed changes in expression level for normal and diseased conditions. hsa-miR-133-b has been found as a tumor suppressor miRNA, targeting FSCN1 in esophageal squamous cell carcinoma (Smalheiser *et al*., 2014). This miRNAs has interacted with seven small ncRNAs in the current analysis (Table 5) with a low expression level in all of the tissues. It has, also, been reported in tumor suppression, and it negatively regulates TBPL1 in collateral cancer (Xiang and Li, 2014). Another miRNA, hsa-miR-145-5p, has been captured with a relatively high expression level in all of the tissues except in brain tissues. The current interaction analysis shows that hsa-miR-145-5p has interacted with SNORD -6, 65, 124, 127 and SNORA3. It has been annotated as a tumor suppressor which inhibits cancer cell growth in lung adeno-carcinoma patients with the mutation of epidermal growth factor receptor (Cho *et al*., 2009). It is also associated with non-coding RNA urothelial cancer which promotes bladder cancer cell migration and invasion through the ZERB1/2 and FSCN-1 pathway (Xue *et al*., 2016).

An interesting miRNA, hsa-miR-224-5p, has shown interaction with eight other small ncRNA (snoRNA) which targets CD40 on the progression of pancreatic ductal adeno carcinoma (Mees *et al*., 2009). These eight small ncRNAs were highly expressed in all of the tissues. SNORD123 being in one of these eight interactions, interacts with five other miRNAs among the twelve miRNAs (Table 5). It is also reported that decreased levels of miR-224 promotes colorectal tumor growth (Yuan *et al*., 2013). Another miRNA showing interactions with four snoRNAs, is the hsa-miR-183-5p, which is reported as a novel diagnostic biomarkers candidate for primary nasopharygeal carcinoma (Tang *et al*., 2014). The down-regulation of hsa-miR-183 inhibits apoptosis and enhances the invasive potential of endometrial stromal cells in endometriosis (Shi *et al*., 2014).

Furthermore, this study has found that few snoRNAs have interacted with several miRNAs such as SNORD123, SNORD124, and SNORA60. SNORD123 has been predicted to interact with five different miRNAs such as (miR-183-5p, miR-224-5p, miR-26a-5p, miR-26b-5p, and miR-873-5p). Similarly, SNORD124 has interacted with mir-133b, miR-145-5p, miR-190b, and miR-24-3p. These results suggest that the miRNA-snoRNA interactions might be important during healthy conditions, and probably, perturbations in these interactions might be associated with disease conditions such as cancer as the individual miRNAs or snoRNAs have been linked with the diseases.

5. Conclusions

This study gives a brief overview of miRNAs and small ncRNAs present in six healthy human tissues. The interaction analysis helps to explore if there are any specific RNA-RNA interactions or differentially expressed sncRNAs that may be tissue-specific. The present study confirms the significance of small ncRNAs including miRNAs and snoRNAs, and suggests that there might be some roles played by miRNA-snoRNA or other small ncRNA interactions which contribute to diseases. Such interactions require further investigations. In conclusion, the *in silico* prediction and analysis of small ncRNAs, together with the expanding data on the expression of known small ncRNAs will unravel the significance of the RNA-RNA interactions, which in turn may have therapeutic applications against various diseases.

Acknowledgement

The authors would like to express their deepest gratitude to Dr. Priyanka Pandey from the National Institute of Biomedical Genomics-India for providing support and data analysis facilities for the successful completion of this study.

References

Anastasiadou E, Jacob LS and Slack FJ. 2018. Non-coding RNA networks in cancer. *Nat Rev Cancer*, **18**:5-15.

Baker TA, Watson JD, Bell SP, Gann A, Losick MA and Levine R, 2003. **Molecular Biology of the Gene,** Benjamin-Cummings Publishing Company.

Cho WC, Chow AS and Au JS. 2009. Restoration of tumour suppressor hsa-miR-145 inhibits cancer cell growth in lung adenocarcinoma patients with epidermal growth factor receptor mutation. *Eur J Cancer*, **45**: 2197-2206.

Esteller M. 2011. Non-coding RNAs in human disease. *Nat Rev Genetics*, **12**: 861-865.

Friedländer MR, Chen W, Adamidi C, Maaskola J, Einspanier R, Knespel S and Rajewsky N. 2008. Discovering microRNAs from deep sequencing data using miRDeep2. *Nat Biotechnol*, **26**: 407-415.

Heberle H, Meirelles GV, da Silva FR, Telles GP and Minghim R. 2015. *InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. BMC Bioinformatics*, **16**: 169.

Hutvágner G, McLachlan J, Pasquinelli AE, Bálint É, Tuschl T and Zamore PD. 2001. A cellular function for the RNA-

interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science*, **293**: 834-838.

Kozomara A and Griffiths-Jones S. 2013. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res*, **42**: D68-D73.

Leung YY, Kuksa PP, Amlie-Wolf A, Valladares O, Ungar LH, Kannan S, Gregory BD and Wang LS. 2015. DASHR: database of small human noncoding RNAs. *Nucleic Acids Res*, **44**: D216-D222.

Love MI, Huber W and Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**: 550.

Lui WO, Pourmand N, Patterson BK and Fire A. 2007. Patterns of known and novel small RNAs in human cervical cancer. *Cancer Res.*, **67**: 6031-6043.

Ma Z, Luo Y and Qiu M. 2017. miR-143 Induces the Apoptosis of Prostate Cancer LNCap Cells by Suppressing Bcl-2 Expression. *Med Sci Monit*, **23**: 359.

Martens-Uzunova ES, Olvedy M and Jenster G. 2013. Beyond microRNA–novel RNAs derived from small ncRNA and their implication in cancer. *Cancer let.*, **340**: 201-211.

Mees ST, Mardin WA, Sielker S, Willscher E, Senninger N, Schleicher C, Colombo-Benkmann M and Haier J. 2009. Involvement of CD40 targeting miR-224 and miR-486 on the progression of pancreatic ductal adenocarcinomas. *Ann Surg Oncol*, **16**: 2339-2350.

Michael MZ, O'Connor SM, van Holst Pellekaan NG, Young GP and James RJ. 2003. Reduced accumulation of specific microRNAs in colorectal neoplasia. *Mol Cancer Res*, **1**: 882-891.

Nasser MW, Datta J, Nuovo G, Kutay H, Motiwala T, Majumder S, Wang B, Suster S, Jacob ST and Ghoshal K. 2008. Down-regulation of micro-RNA-1 (miR-1) in lung cancer suppression of tumorigenic property of lung cancer cells and their sensitization to doxorubicin-induced apoptosis by miR-1. *J Biol Chem*, **283**: 33394-33405.

Pritchard CC, Cheng HH and Tewari M. 2012. MicroRNA profiling: approaches and considerations. *Nat Rev Genet*, **13**: 358-369.

Shi XY, Gu L, Chen J, Guo XR and Shi YL. 2014. Downregulation of miR-183 inhibits apoptosis and enhances the invasive potential of endometrial stromal cells in endometriosis. *Int J Mol Med*, **33**: 59-67.

Smalheiser NR, Lugli G, Zhang H, Rizavi H, Cook EH and Dwivedi Y. 2014. Expression of microRNAs and other small RNAs in prefrontal cortex in schizophrenia, bipolar disorder and depressed subjects. *PLoS One*, **9**: e86469.

Tang JF, Yu ZH, Liu T, Lin ZY, Wang YH, Yang LW, He HJ, Cao J, Huang HL and Liu G. 2014. Five miRNAs as novel diagnostic biomarker candidates for primary nasopharyngeal carcinoma. *Asian Pac J Cancer Prev*, **15**: 7575-7581.

Van Wynsberghe PM, Chan SP, Slack FJ and Pasquinelli AE. 2011. Analysis of microRNA expression and function. *Methods Cell Biol*, **106**: 219.

Xiang KM and Li XR. 2014. MiR-133b acts as a tumor suppressor and negatively regulates TBPL1 in colorectal cancer cells. *Asian Pac J Cancer Prev*, 1**5**: 3767-72.

Xue M, Pang H, Li X, Li H, Pan J and Chen W. 2016. Long noncoding RNA urothelial cancer associated 1 promotes bladder cancer cell migration and invasion by way of the has miR 145–ZEB1/2–FSCN1 pathway. *Cancer Sci.*, **107**: 18-27.

Yuan K, Xie K, Fox J, Zeng H, Gao H, Huang C and Wu M. 2013. Decreased levels of miR-224 and the passenger strand of miR-221 increase MBD2, suppressing maspin and promoting colorectal tumor growth and metastasis in mice. *Gastroenterol.*, **145**: 853-864.