Journal of
Data Mining in Genomics & Proteomics

Mallick B, J Data Mining Genomics & Proteomics
2016, 7:3
DOI: 10.4172/2153-0602.1000e124

Editorial Open Access

Decrypting the Treasures of Regulatory Non-coding RNAs in High-throughput Era

Bibekanand Mallick*

Department of Life Science, National Institute of Technology, Rourkela-769008, India

***Corresponding author:** Bibekanand Mallick, RNAi and Functional Genomics Laboratory, Department of Life Science, National Institute of Technology, Rourkela-769008, India, Tel: +91-661-2462685; Fax: +91-661-2472926; E-mail: vivek.iitian@gmail.com, mallickb@nitrkl.ac.in

Received date: June 23, 2016; **Accepted date:** June 25, 2016; **Published date:** June 27, 2016

Editorial

Over the past decade, there has been an increasing curiosity in the field of regulatory non-coding RNAs (rncRNAs). This is due to the discovery of a huge number of rncRNA genes and their involvement in multiple regulatory functions. Further, recent technological advances have added icing on the cake to this venture. The advent of high-throughput next-generation sequencing (NGS), especially RNA-sequencing (RNA-Seq) enabled genome-scale profiling and quantification of not only coding transcripts but also rncRNA transcripts by involving direct sequencing of complementary DNAs (cDNAs) [1,2]. This revolutionary unearthing of human transcriptome tended to shift the paradigm that the proteins (encoded by messenger RNAs) control the fate of the cells. The new classes of rncRNAs discovered from the high-throughput RNA-Seq studies demonstrated their potential regulatory roles in diverse cellular processes such as cellular differentiation, development, homeostasis and many more biological processes [3-7]. These rncRNAs include microRNA (miRNA), small interfering RNA (siRNA), piwi-interacting RNA (piRNA), long non-coding RNA (lncRNA), circular RNA (circRNA) and much more.

In RNA-Seq, a population of RNA (total or fractionated, such small rncRNAs) is converted to a library of cDNAs with adaptors attached to one or both end followed by sequencing in a high-throughput manner. The sequencing yields millions of short sequences (termed as reads) from one end (single-end sequencing) or both ends (pair-end sequencing). These reads are post-processed for quality control and adaptor trimming by using tools such as FastQC, NGSQC [8], FASTX-Toolkit, Trimmomatic [9], Cutadapt, and few more. The reads are typically of 30–400 nucleotides in size, depending on the sequencing technology/platform used. The platforms used for the RNA-Seq purpose are Illumina, SOLiD, Ion Torrent and Roche 454 Life Science systems, etc. The reads after passing QC and adaptor removal are aligned to the corresponding reference genome (say, hg19 for human genome) using NGS read aligner tools such as Bowtie, an efficient and widely used the genome-scale alignment tool [10]. The other read aligner tools are BWA [11], BLAT [12], ELAND, BFAST [13], GMAP [14], SOAP [15], SeqMap [16], SHRiMP [17], NextGenMap [18], etc. The uniquely mapped reads with not more than one mismatch to the corresponding genome are considered for their annotations and subsequent analysis. The reference sequences of different genomes are obtained from the databases such as NCBI and UCSC genome browser.

The uniquely mapped reads showing length distribution in the range of rncRNAs (either small ncRNAs or lncRNAs) are analyzed by adopting different strategies. Moreover, sequencing of small RNAs and long RNAs are performed separately to pool out all the desired RNAs efficiently as per our need. As an example, the sequencing of RNAs within the range of 16-40 nucleotides is recommended if we intend to identify small rncRNAs such as miRNAs and piRNAs. The preprocessed reads are then analyzed using tools such as iMir and mirtoools [19,20] to identify known and novel piRNAs as well as miRNAs. The annotations of genomic regions from which these small RNAs might have originated are done by aligning these to the coordinates of mRNAs (5'UTR, CDS, 3'UTR), repeats, pseudogenes, introns, small ncRNAs and lncRNAs using any publicly available tools or in-house programs. We can obtain these individual annotation track files from UCSC FTP site.

The expression level of a transcript (ei,j) in individual i obtained from the RNA-Seq experiments is computed as follows:

$$e_{i,j} = g_{i,j} \times t/a_i \times l_i$$

Where gi,j is the number of reads mapped onto transcript j (transcript, exon, etc.) in individual i, a_i is the total number of mappable reads in individual i, and l is the length of the transcript in individual i. The expression of a transcript (ei,j) is equivalent to RPKM (reads per kilobase per million mapped reads) when the scaling factor is t=109. The measure of expression is also represented in terms of TPM (transcripts per million) in RNA-Seq. 1 TPM is equal to 1 RPKM, which is approximately 1 transcript per cell when the average transcript length is 1 kb. It is recommended to have at least three replicates for each condition for computing differential expression between two samples. The tools publicly available for performing differential expression analysis are edgeR [21], DESeq [22], baySeq [23], EBSeq [24], Cufflinks, etc. Some of these tools accept raw reads, and others accept normalized reads for computing expression levels.

As mentioned previously, the RNA-Seq method is used to profile and quantify any transcripts including lncRNAs and small ncRNAs. The methods are described below.

Identification of lncRNAs

The lncRNAs are >200 nucleotide long transcripts without the protein-coding potential. These are known to play vital roles in various biological processes [25]. Many types of NGS studies [26-28] have been performed (outline below) to identify and characterize these lncRNAs in various types of organisms, cells, tissues, etc.

RNA-Seq: The sequencing of short cDNA fragments using RNA-Seq is used to reconstruct virtually an entire transcriptome to uncover functional lncRNAs. The reads generated from the RNA-Seq of the transcriptome is aligned to the corresponding reference genome followed by assembling of the transcripts. The assembled transcripts matching to the protein coding transcripts are removed. The remaining transcripts with length >200 nts are selected (Figure 1) and coding

potential is evaluated using tools such as CPC (Coding Potential Calculator) [29], CPAT (Coding-Potential Assessment Tool) [30], Coding-Non-Coding Index (CNCI) [31] and PLEK (predictor of long non-coding RNAs and messenger RNAs based on an improved k-mer scheme) [32] to annotate these as lncRNAs.

3SEQ: 3SEQ (3'-end sequencing for expression quantification) is a modified version of the traditional RNA-Seq method that is based on enrichment of 3' ends of transcripts by relying on the polyA+ selection of fragmented RNAs [28]. In 3SEQ, only the 3'-end-most polyadenylated fragment of each transcript is isolated and sequenced which substantially reduce the depth of sequencing enabling discovery and quantification of rare RNAs including lncRNAs. This method is advantageous over the traditional RNA-Seq method by employing strand-specific libraries which allow to quantify transcripts separately based on strands using directional information of each read.

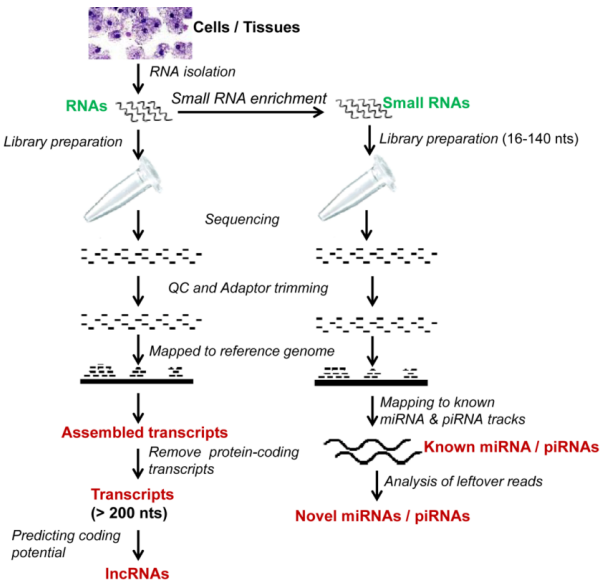

Figure 1: Flowchart of RNA-Seq for identification of non-coding RNAs (lncRNA, miRNA, piRNA).

Full-length cDNA sequencing: Full-lenghth cDNA sequencing of >21000 cDNA clones was performed by the FANTOM (Functional Annotation of The Mammalian Genome) project, which aimed to characterize the 'unclassifiable' transcripts (ncRNAs) in addition to the expected protein-coding transcripts [33]. These 'unclassifiable' transcripts accounted for ~30% of the total transcripts which included lncRNAs. Further, identification of transcriptional start sites (TSSs) using cap-analysis of gene expression (CAGE) tag sequences in FANTOM3 also predicted additional lncRNAs [33].

RNA pol II ChIP-Seq: The ChIP-Seq is a powerful method that combines chromatin immunoprecipitation (ChIP) with parallel DNA sequencing for identifying genome-wide DNA binding sites for transcription factors and other proteins [34]. It is known that ChIP-Seq profiling of H3K4me3 and H3K36me3 preferentially identifies TSSs and transcribed gene bodies respectively [35]. This observation led to the collective use of these marks to identify 'K4–K36 domains' outside of known protein-coding genes that may represent novel lncRNA loci. There are increasing evidence of the use of ChIP-seq data of H3K4me3, H3K36me3, and Pol II to detect lncRNAs in various organisms [36,37].

Identification of miRNAs and piRNAs

The RNA-seq is highly versatile and can be modified to investigate specific RNAs such as small RNAs. The identification of small RNAs can be performed by employing the method termed as small RNA-Seq where gel-based size selection is used to enrich RNAs within the range of length of desired small RNAs. The construction of library can be fractionated to select 16-40 nts RNAs using gel electrophoresis if one intends to identify and quantify both miRNAs and piRNAs (through RNA-Seq), the most promising sncRNAs discovered till date. The miRNAs are ~18–22-nt single-stranded ncRNAs primarily found in eukaryotes [38] is known to regulate mRNA transcripts post-transcriptionally as well as crosstalk with other ncRNAs such as lncRNAs through competing endogenous RNA (ceRNA) network [39,40]. While piRNAs are comparatively less investigated and young ncRNAs of about 25-33 nts in length primarily reported in germ lines to regulate the transposable elements (TEs) [41]. There are computational challenges to identify and validate real and novel mi/piRNAs from the RNA-Seq demanding extensive experimental validations to prove newly identified mi/piRNAs as real. Nevertheless, several computational tools have been developed to identify known small ncRNAs (miRNAs and piRNAs) accurately and also complement experimental approaches to identify novel small ncRNAs, of course with less sensitivity. The most popular tools used for predicting miRNAs and piRNAs from RNA-Seq data are miRDeep2 [42], miRanalyzer [43], Dario [44], SSCprofiler [45], miReader [46], mirtools2 [20], piPipes [47] and iMir [19].

Downstream Analysis of ncRNAs

The functions of ncRNAs can be deduced by analyzing their probable targets and enrichment in various biological processes. The majority of the ncRNAs such as miRNA, piRNAs, and lncRNAs primarily participate in gene regulation at epigenetic, transcriptional and post-transcriptional level adopting different mechanisms. The miRNAs act at post-transcriptional level for fine-tuning of gene expression in the cell by binding to target mRNAs at their 3'UTR, CDS or 5'UTR in a sequence-specific manner through partial or full complementarity. The miRNAs, when perfectly base-paired to their target mRNA, direct cleavage of the target mRNA and direct transcriptional repression when partially base-paired, preferably to 2-7/8 bases at 5' end of the miRNAs (designated as 'seed') [48]. Many miRNA targets have been computationally predicted, but only a limited number of these were found to be experimentally correct. There are a variety of tools for predicting miRNA targets, among which most of them relies on sequence complementarity of 7-8 bases. The most widely used tools are targetscan [49], pictar [50], mirsvr [51], starmir [52], pita [53] etc. Among these, mirsvr and starmir are based on sequence, structure and thermodynamic features of target mRNAs trained on Ago2-identified CLIP-Seq studies. However, none of these tools are 100% sensitive in predicting targets and hence finding a functional miRNA target is still a challenging problem. To overcome this challenge, many CLIP-Seq studies are booming up along with the development of experimentally verified miRNA target databases to meet the requirement of the RNA biologists. Some of the widely known databases are miRTarBase [54], miRWalk2.0 [55], StarBase [56], StarmirDB [53], and miRecords [57], etc.

The piRNAs are primarily known to execute repeat inactivation by targeting transposons, but these are currently known to regulate endogenous mRNAs as well. The principle of targeting by piRNAs is not identical to that of miRNAs as thought earlier. There are several

theories gradually evolving to decipher the mechanism of targeting by piRNAs. In a recent report, Goh et al, 2015 deduced that potential piRNAs bind to target mRNAs through perfect sequence complementarity in nts 2–11 (primary seed), with maximum of four mismatches being tolerated in nts 12–21 (secondary seed) of the piRNAs [58]. The possible regions of target mRNAs where piRNAs can bind are 3'UTR, CDS, 5'UTR similar to that of miRNA targeting. However, there are also other rules of piRNA targeting reported in other studies [59,60]. These indicate that these concepts are still in nascent stage and are likely to evolve in coming years to sharpen our understanding of mechanisms of piRNA-mediated regulations.

lncRNAs which was initially treated as transcriptional noise are now established as an important group of functional ncRNAs playing diverse roles in genetic imprinting, genome rearrangement, chromatin modification, cell cycle regulation, transcription, splicing, mRNA decay, and translation [5]. These regulate expression of genes at epigenetics, transcriptional and post-transcriptional level. The functions of lncRNAs are elucidated from studying the changes in genome-wide or individual gene expression level after knocking down or overexpressing the lncRNA. Some of the lncRNAs act as miRNA decoys, known as competing for endogenous RNAs (ceRNAs) which are reported to play key roles in diseases including cancer [61,62]. ceRNAs can sequester shared miRNAs through competition for binding to miRNAs leading to de-repression of the other target transcript [63]. Several computational methods have been developed in the past to predict ceRNAs, but none these are efficient [56,64,65]. Moreover, all the components of a ceRNA network and how they interact among themselves to modulate cellular mechanisms are yet to be understood.

Conclusion

Unveiling the entire regulatory ncRNA landscape and their functional characterization is a rapidly expanding area of research. Within the past few years, it has become clear that although the majority of the transcripts in cells do not undergo translation still, they perform important regulatory functions. Many of these new rncRNAs belong to an important, relatively unexplored class of regulatory elements. Rapid improvements in high-throughput sequencing technologies have made it possible to discover a significant amount of these uncharacterized transcripts. With such rapidity of advancement, if sequencing costs keep decreasing, we can plausibly expect this technology to become a key component of rncRNomics. Moreover, much remains to be done regarding accurate characterization and functional analysis of rncRNAs. Despite the difficulties encountered regarding characterizing novel rncRNA transcripts, we believe that new and more sophisticated technnologies and algorithms will come to the rescue. With the advent of such new technologies, we shall be able to probe into this large novel class of rncRNAs which may well define one of the turning points of modern biology.

Acknowledgement

The skillful assistance of Jyoti Roy of my laboratory is gratefully acknowledged. This work is supported by grants from Council of Scientific & Industrial Research (CSIR) and Department of Biotechnology (DBT) of Govt. of India awarded to BM.

References

1. Ozsolak F, Milos PM (2011) RNA sequencing: advances, challenges and opportunities. Nat Rev Genet 12: 87-98.
2. Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10: 57-63.
3. Ameres SL, Zamore PD (2013) Diversifying microRNA sequence and function. Nat Rev Mol Cell Biol 14: 475-488.
4. Iwasaki YW, Siomi MC, Siomi H (2015) PIWI-Interacting RNA: Its Biogenesis and Functions. Annu Rev Biochem 84: 405-433.
5. Iyer MK (2015) The landscape of long noncoding RNAs in the human transcriptome. Nat Genet 47: 199-208.
6. Chen LL (2016) The biogenesis and emerging roles of circular RNAs. Nat Rev Mol Cell Biol 17: 205-211.
7. Samantarrai D, Sahu M, Roy J, Mohanty BB, Singh G, et al. (2015) Unraveling novel TF-miRNA regulatory crosstalk in metastasis of Soft Tissue Sarcoma. Sci Rep 5: 9742.
8. Dai M, Thompson RC, Maher C, Contreras-Galindo R, Kaplan MH, et al. (2010) NGSQC: cross-platform quality analysis pipeline for deep sequencing data. BMC Genomics 11: S7.
9. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30: 2114-2120.
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. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25: 1754-1760.
12. Kent WJ (2002) BLAT-the BLAST-like alignment tool. Genome Res 12: 656-664.
13. Homer N, Merriman B, Nelson SF (2009) BFAST: an alignment tool for large scale genome resequencing. PLoS One 4: e7767.
14. Wu TD, Watanabe CK (2005) GMAP: a genomic mapping and alignment program for mRNA and EST sequences. Bioinformatics 21: 1859-1875.
15. Li R, Li Y, Kristiansen K, Wang J (2008) SOAP: short oligonucleotide alignment program. Bioinformatics 24: 713-714.
16. Jiang H, Wong WH (2008) SeqMap: mapping massive amount of oligonucleotides to the genome. Bioinformatics 24: 2395-2396.
17. Rumble SM, Lacroute P, Dalca AV, Fiume M, Sidow A, et al. (2009) SHRiMP: accurate mapping of short color-space reads. PLoS Comput Biol 5: e1000386.
18. Sedlazeck FJ, Rescheneder P, von Haeseler A (2013) NextGenMap: fast and accurate read mapping in highly polymorphic genomes. Bioinformatics 29: 2790-2791.
19. Giurato G (2013) iMir: an integrated pipeline for high-throughput analysis of small non-coding RNA data obtained by smallRNA-Seq. BMC Bioinformatics 14: 362.
20. Wu J, Liu Q, Wang X, Zheng J, Wang T, et al. (2013) mirTools 2.0 for non-coding RNA discovery, profiling, and functional annotation based on high-throughput sequencing. RNA Biol 10: 1087-1092.
21. Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26: 139-140.
22. Anders S, Huber W (2010) Differential expression analysis for sequence count data. Genome Biol 11: R106.
23. Hardcastle TJ, Kelly KA (2010) baySeq: empirical Bayesian methods for identifying differential expression in sequence count data. BMC Bioinformatics 11: 422.
24. Leng N (2015) EBSeq-HMM: a Bayesian approach for identifying gene-expression changes in ordered RNA-seq experiments. Bioinformatics 31: 2614-2622.
25. Guo X, Gao L, Wang Y, Chiu DK, Wang T, et al. (2016) Advances in long noncoding RNAs: identification, structure prediction and function annotation. Brief Funct Genomics 15: 38-46.

26. Chen B, Zhang Y, Zhang X, Jia S, Chen S, et al. (2016) Genome-wide identification and developmental expression profiling of long noncoding RNAs during Drosophila metamorphosis. Sci Rep 6: 23330.
27. Ran M, Chen B, Li Z, Wu M, Liu X, et al. (2016) Systematic Identification of Long Noncoding RNAs in Immature and Mature Porcine Testes. Biol Reprod 94: 77.
28. Beck AH (2010) 3'-end sequencing for expression quantification (3SEQ) from archival tumor samples. PLoS One 5: e8768.
29. Kong L, Zhang Y, Ye ZQ, Liu XQ, Zhao SQ, et al. (2007) CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. Nucleic Acids Res 35: W345-W349.
30. Wang L, Park HJ, Dasari S, Wang S, Kocher JP, et al. (2013) CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. Nucleic Acids Res 41: e74.
31. Sun L (2013) Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. Nucleic Acids Res 41: e166.
32. Li A, Zhang J, Zhou Z (2014) PLEK: a tool for predicting long non-coding RNAs and messenger RNAs based on an improved k-mer scheme. BMC Bioinformatics 15: 311.
33. Carninci P (2005) The transcriptional landscape of the mammalian genome. Science 309: 1559-1563.
34. Quinodoz M, Gobet C, Naef F, Gustafson KB (2014) Characteristic bimodal profiles of RNA polymerase II at thousands of active mammalian promoters. Genome Biol 15: R85.
35. Young MD, Willson TA, Wakefield MJ, Trounson E, Hilton DJ, et al. (2011) ChIP-seq analysis reveals distinct H3K27me3 profiles that correlate with transcriptional activity. Nucleic Acids Res 39: 7415-7427.
36. Chen MJ (2016) Integrating RNA-seq and ChIP-seq data to characterize long non-coding RNAs in Drosophila melanogaster. BMC Genomics 17: 220.
37. Lu L, Sun K, Chen X, Zhao Y, Wang L, et al. (2013) Genome-wide survey by ChIP-seq reveals YY1 regulation of lincRNAs in skeletal myogenesis. EMBO J 32: 2575-2588.
38. Siomi H, Siomi MC (2010) Posttranscriptional regulation of microRNA biogenesis in animals. Mol Cell 38: 323-332.
39. Krol J, Sobczak K, Wilczynska U, Drath M, Jasinska A, et al. (2004) Structural features of microRNA (miRNA) precursors and their relevance to miRNA biogenesis and small interfering RNA/short hairpin RNA design. J Biol Chem 279: 42230-42239.
40. Dai Q, Li J, Zhou K, Liang T (2015) Competing endogenous RNA: A novel posttranscriptional regulatory dimension associated with the progression of cancer. Oncol Lett 10: 2683-2690.
41. Meister G (2013) Argonaute proteins: functional insights and emerging roles. Nat Rev Genet 14: 447-459.
42. Mackowiak SD (2011) Identification of novel and known miRNAs in deep-sequencing data with miRDeep2. Curr Protoc Bioinformatics.
43. Hackenberg M, Rodriguez-Ezpeleta N, Aransay AM (2011) miRanalyzer: an update on the detection and analysis of microRNAs in high-throughput sequencing experiments. Nucleic Acids Res 39: W132-W138.
44. Fasold M, Langenberger D, Binder H, Stadler PF, Hoffmann S (2011) DARIO: a ncRNA detection and analysis tool for next-generation sequencing experiments. Nucleic Acids Res 39: W112-W117.
45. Oulas A, Poirazi P (2011) Utilization of SSCprofiler to predict a new miRNA gene. Methods Mol Biol 676: 243-252.
46. Jha A, Shankar R (2013) miReader: Discovering Novel miRNAs in Species without Sequenced Genome. PLoS One 8: e66857.
47. Han BW, Wang W, Zamore PD, Weng Z (2015) piPipes: a set of pipelines for piRNA and transposon analysis via small RNA-seq, RNA-seq, degradome- and CAGE-seq, ChIP-seq and genomic DNA sequencing. Bioinformatics 31: 593-595.
48. Liu C, Rennie WA, Mallick B, Kanoria S, Long D, et al. (2014) MicroRNA binding sites in C. elegans 3' UTRs. RNA Biol 11: 693-701.
49. Lewis BP, Burge CB, Bartel DP (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 120: 15-20.
50. Krek A (2005) Combinatorial microRNA target predictions. Nat Genet 37: 495-500.
51. Betel D, Koppal A, Agius P, Sander C, Leslie C (2010) Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites. Genome Biol 11: R90.
52. Liu C, Mallick B, Long D, Rennie WA, Wolenc A, et al. (2013) CLIP-based prediction of mammalian microRNA binding sites. Nucleic Acids Res 41: e138.
53. Rennie W (2016) STarMirDB: A database of microRNA binding sites. RNA Biol 13: 554-560.
54. Hsu HD (2011) miRTarBase: a database curates experimentally validated microRNA-target interactions. Nucleic Acids Res 39: D163-D169.
55. Dweep H, Gretz N, Sticht C (2014) miRWalk database for miRNA-target interactions. Methods Mol Biol 1182: 289-305.
56. Li JH, Liu S, Zhou H, Qu LH, Yang JH (2014) Starbase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. Nucleic Acids Res 42: D92-D97.
57. Xiao F, Zuo Z, Cai G, Kang S, Gao X, et al. (2009) miRecords: an integrated resource for microRNA-target interactions. Nucleic Acids Res 37: D105-D110.
58. Goh WS, Falciatori I, Tam OH, Burgess R, Meikar O, et al. (2015) piRNA-directed cleavage of meiotic transcripts regulates spermatogenesis. Genes Dev 29: 1032-1044.
59. Zhang P (2015) MIWI and piRNA-mediated cleavage of messenger RNAs in mouse testes. Cell Res 25: 193-207.
60. Vourekas A, Alexiou P, Vrettos N, Maragkakis M, Mourelatos Z (2016) Sequence-dependent but not sequence-specific piRNA adhesion traps mRNAs to the germ plasm. Nature 531: 390-394.
61. Peng W, Si S, Zhang Q, Li C, Zhao F, et al. (2015) Long non-coding RNA MEG3 functions as a competing endogenous RNA to regulate gastric cancer progression. J Exp Clin Cancer Res 34: 79.
62. Xiao H (2015) LncRNA MALAT1 functions as a competing endogenous RNA to regulate ZEB2 expression by sponging miR-200s in clear cell kidney carcinoma. Oncotarget 6: 38005-38015.
63. Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP (2011) A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? Cell 146: 353-358.
64. Sarver AL, Subramanian S (2012) Competing endogenous RNA database. Bioinformation 8: 731-733.
65. Xu J (2015) The mRNA related ceRNA-ceRNA landscape and significance across 20 major cancer types. Nucleic Acids Res 43: 8169-8182.