JOURNAL OF INTEGRATED OMICS
A METHODOLOGICAL JOURNAL
HTTP://WWW.JIOMICS.COM

ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v1i1.21

A new method for measuring functional similarity of microRNAs

Guangchuang Yu[1,a], Chuan-Le Xiao[1,a], Xiaochen Bo[2], Chun-Hua Lu[1,3], Yide Qin[4], Sheng Zhan[1], Qing-Yu He*[1].

[1]Institutes of Life and Health Engineering, Jinan University, 601 Huangpu Ave. W. Guangzhou, 510632, China; [2]Beijing Institute of Radiation Medicine, Beijing, 100850, China; [3]College of Life Science and Technology, Guangxi University, Nanning, 530004, China; [4]Department of Biochemistry and Molecular Biology, Anhui Medical University, Hefei, 230032, China. [a] These authors contributed equally in this work.

Received: 28 July 2010 **Accepted:** 18 August 2010 **Available Online**: 10 September 2010

ABSTRACT

MicroRNAs (miRNAs) are a group of small RNAs with regulatory roles at post-transcriptional level. Although they have been clustered based on their sequence or structure similarities, there is still no effective method to determine their functional similarities due to the lack of miRNAs functional annotation. To address this critical need, we presented here a novel method for systematic study of functional similarities among human miRNAs by using their target genes GO semantic similarities. The functional similarities were validated by comparing with miRNA expression similarities. To extract the highly significant clusters, we used multi-scale bootstrap re-sampling in clustering miRNAs functional similarities. The clustering of human miRNAs based on target genes molecular function annotation led to 44 significant clusters. The clustering results were coherent with biological knowledge. Our analysis suggests that systematic clustering based on target genes GO semantic similarities can aid to reveal the functional diversity of miRNA families. Additionally, this method can be extended to other species and used to predict novel miRNA functions.

Keywords: MicroRNA; Gene Ontology; Semantic Similarity; Clustering.

1. Introduction

MicroRNAs (miRNAs), approximately 22 nucleotides in length, are non-coding RNAs that play crucial roles in post-transcriptional regulation. MiRNAs are evolutionarily conserved, and generally transcribed by RNA polymerase II. MiRNAs perform their functions by RNA induced silencing complex (RISC), leading their target mRNAs to direct destructive cleavage or translational repression. MiRNAs are considered to represent one of the most important components of the cell. They involve in many critical biological processes, including cell development and differentiation [1,2], proliferation [3], apoptosis [4], development [5], immune system regulation [6,2], cancer progression [7], and virus-host interaction [8] and therefore represent potential targets for therapeutic applications. According to miRBase [9], the number of registered miRNA genes continues to grow rapidly. However, hundreds of recently identified miRNAs have unknown functions due to the lack of experimental strategy for systematic identification of their regulating targets.

In order to better understand miRNAs, it is increasingly necessary to measure their functional similarity and thus to infer novel potential functions for miRNAs. Human miRNAs have been grouped into 46 families on the basis of hairpin sequences conservation by Rfam [10], and 60 families according to pre-miRNAs sequence and secondary structures by using FOLDALIGN [11]. Many human clusters containing miRNAs without sequence homology was found [12] which indicating the current strategies for measuring miRNA similarity have some flaws. Since miRNA-mRNA duplex allows mismatch, and target recognition only matches the 6-8nt long seed region, miRNAs with similar sequences and pre-miRNAs with similar structures may have distinctive functions. Therefore, a new method for measuring miRNAs functional similarity is necessitated.

Gene Ontology (GO) is the *de facto* standard for annotation of gene products. The relationship of different genes was

***Corresponding author:** Qing-Yu He, Institutes of Life and Health Engineering, Jinan University, 601 Huangpu Ave. W. Guangzhou, 510632, China. Email address: tqyhe@jnu.edu.cn

organized as a directed acyclic graph (DAG), making it feasible for quantitative semantic comparisons. Measuring gene functional similarity based on GO has been widely used in novel GO annotations [13,14], gene function prediction [15] and similarity measurement [16,17], microarray analysis [18], cluster analysis [19,20], pathway analysis [21,22], and localization prediction [23]. The powerfulness of GO semantic similarity measurement has been verified in terms of the correlations with sequence similarities [24], gene expression profiles [25], and protein-protein interactions [26]. However, it is impractical to measure miRNAs functional similarity directly due to the lack of GO terms annotation of most miRNAs and the lack of functional annotation database [27] since most miRNAs functions remain unknown. Fortunately, most of the genes miRNAs regulated are well annotated, making it possible for inferring functional similarity of miRNAs based on their targets. The functions of miRNAs can be inferred by GO enrichment analysis of their predicted targets were reported [28,29], which also indicating that the functional similarity of miRNAs can be measured by quantitatively calculating the similarity of their target genes.

Here, we proposed a new method for systematic study of miRNAs functional similarity. The fundamental concept of our method is based on the functions of miRNAs targets GO semantic similarities. We validated our method by comparing it with miRNA expression similarity and showed that clustered miRNAs have functional relatedness through co-function of targeting genes. The similarities obtained by our method are consistent with biological knowledge of miRNA functional relationship.

2. Material and methods

2.1 *Human microRNA target prediction*

Experimental identification of target mRNAs is difficult, and TarBase [30] currently lists only 1093 verified target mRNAs for 110 human miRNAs. Due to the lack of experimental targets of miRNAs, we used computational method for miRNA genome-wide target prediction in this study, where target genes were predicted by the algorithm of Probability of Interaction by Target Accessibility (PITA) [31]. PITA uses a thermodynamic model for miRNA-mRNA interaction that was scored by an energy score, ΔΔG, which equals to the difference between the energy expended on opening the target site structure, ΔGopen, and the energy gained by forming the duplex, ΔGduplex. ΔΔG correlates well with the experimentally measured degree of mRNA suppression were reported [31].

Human miRNAs were downloaded from miRBase, version 12 [9] and human genome was downloaded from UCSC, version 18 [32] which corresponds to the human genome build 36.1 assembled by NCBI. We extracted 3' untranslated region (UTR) sequences in a single FASTA format file. For genes missing 3'-UTR annotation, 800bp downstream annotated end of the coding sequence were used as the putative UTR. Since miRNA-target interaction requires unpairing of bases flanking the targets, we used a flank of 3 upstream and 15 downstream nucleotides when performing prediction. To reduce false positive, the prediction results were narrowed down using the criteria of 7-8 bases seed length, no G:U wobble or loops, no mismatch and conservation score of 0.9 or higher. To assign an overall miRNA-target score, we computed the statistical weight to sum all the ΔΔG generated by different sites bounded to the same miRNA as defined in formula 1. Finally we screened out the results by ΔΔG below -10 kcal/mol. ΔΔG is an energetic score, and the lower its value, the stronger of miRNA-target binding. ΔΔG < -10 kcal/mol is expected to be functional in endogenous miRNA expression levels.

$$Target\_Score = -\log(\sum(\exp(-\Delta\Delta G))) \qquad (1)$$

2.2 *Human microRNA functional similarities measurement*

For measuring functional similarities among miRNAs, we used GO annotation of their target genes to define the similarity index.

Here, we developed an R package called *GOSemSim* [33], which implemented five methods proposed by Resnik [34], Lin [35], Jiang [36], Schlicker [37], and Wang [38] respectively, and was extended to support 19 species, including Anopheles, Arabidopsis, Bovine, Canine, Chicken, Chimp, Coelicolor, E coli strain K12 and Sakai, Fly, Human, Malaria, Mouse, Pig, Rhesus, Rat, Worm, Xenopus, Yeast, and Zebrafish. The program is freely distributed under GPL2 and can directly be installed from Bioconductor. The manual and source code are available at http://bioconductor.org/packages/release/bioc/html/GOSemSim.html. Since GOSemSim package only supports Entrez Gene identifier for measuring similarities among human genes, we used Bioconductor package biomaRt [39] to query BioMart [40] databases for mapping RefSeq identifiers to Entrez Gene identifiers. Molecular function ontology was used to annotate target genes, and Wang method was used to calculate similarity. Wang method was based on GO graph structure, and outperformed other algorithms based on information content and thus being more consistent with human perspectives [38]. In order to give a single distance between two miRNAs, we combined similarity scores of multiple target genes as defined in formula 2. Similarities between two gene sets associated with two miRNAs form a matrix. The similarity of these two miRNAs is the average of maximum row scores and column scores. We used this strategy and finally obtained pairwise semantic similarities among human miRNAs.

$$Sim = \frac{\sum_{1\le i\le m} \max(Score_{[i,]}) + \sum_{1\le j\le n} \max(Score_{[,j]})}{m+n} \qquad (2)$$

2.3 *Hierarchical clustering uncertainty assessment*

The resulting similarity scores were then clustered by R package *Pvclust* [41]. When performing clustering, *pvclust* used multi-scale bootstrap re-sampling to estimate the uncertainty of cluster analysis which has been popular in phylogenetic analysis. The agglomerative method, average linkage, was used, and 10,000 bootstrap replications were run, with relative sample size from 0.5 to 1.4, incrementing in steps of 0.1, for testing *p*-values. For a cluster with approximately unbiased (AU) *p*-value > 0.95, the hypothesis with "the cluster does not exist" is rejected with significance level 0.05. Roughly speaking, these clusters not only "seem to exist" attributed to sampling error, but be stably observed when we increase the number of observation [41]. All clusters in this study were extracted with *p*-value > 0.95.

3. Results and Discussion

We used PITA algorithm to predict human miRNA targets, followed by a strict criteria to control the FDR, and then measured miRNA functional similarity by *GOSemSim* package. As a result, we obtained the pairwise functional similarity of 533 miRNAs which was provided as Supplemental File 1. The pairwise functional similarity of human miRNA repertoire was shown in Figure 1. The full size of Figure 1 was provided as Supplemental File 2.

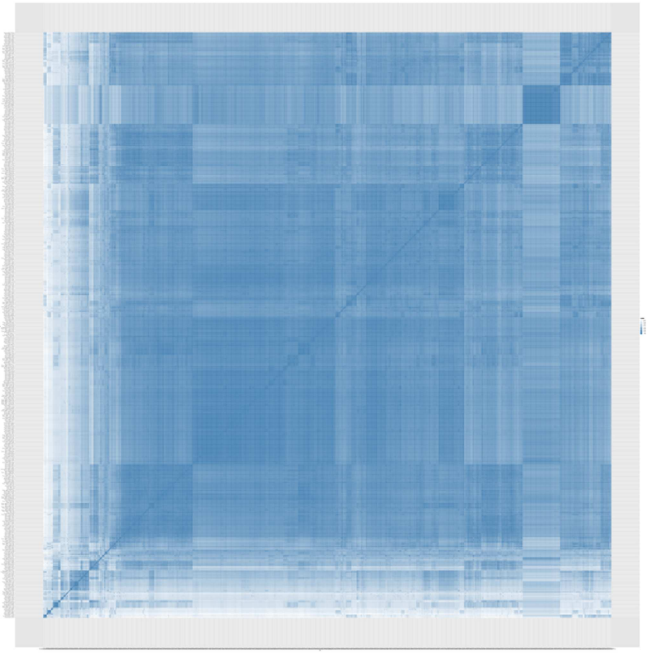

Figure 1. Human miRNAs functional similarities.

MiRNAs with high functional similarity may tend to have similar expression profiles. Taking miRNA functional similarity in pairs against expression similarity should show a positive relationship. For further evaluating the quality of our result, we study the relationship of miRNAs functional similarity and expression similarity. We used Pearson's correlation coefficients to calculate miRNA expression similarity by using miRNA expression profiles of 345 miRNAs in 40 normal human tissues obtained from Liang's investigation [42]. We classified miRNA pairs into separate groups according to functional similarity values by a step of 0.1, and calculated the average of functional similarity and expression similarity of each group, and then measured the correlation of functional similarity with expression similarity (see Supplemental File 4). As expected, functional similarity obtained by our method showed positive correlation with expression similarity (r = 0.6055), in which the functions of miRNAs can partially be explained by their expression level (r^2 = 0.3666).

The pairwise miRNAs functional similarity matrix was then clustered by *Pvclust* package. We obtained 44 *Pvclust* clusters with AU *p*-value > 0.95, containing 401 miRNAs. The result of cluster analysis with high *p*-value highlighted was provided as Supplemental File 3.

Many miRNAs were reported to be associated with diseases. It has been reported that miRNAs implicated in similar disease often have similar functions [27]. In our results, miRNAs associated to similar diseases were more likely to be grouped in the same cluster. For instance, in cluster 23, hsa-miR-215 [43], hsa-miR-221 [44], hsa-miR-194 [45], hsa-miR-193b [46], and hsa-miR-429 [47] were all reported to be associated with adenocarcinoma. In cluster 44, hsa-miR-453 and hsa-miR-219 [48], hsa-miR-93 [49] were reported to relate with breast cancer; while in cluster 29, hsa-miR-30a, hsa-miR-150, hsa-miR-223, and hsa-miR-600 were implicated to relate with Lupus Vulgaris [50].

MiRNAs conserved in evolution may regulate the cardinal biological process cooperatively. We identified that these miRNAs are tend to cluster together. For example, in cluster 24, hsa-miR-300, hsa-miR-495, hsa-miR-154, and hsa-miR-496 were reported to be conserved in genomes of human, chimp, mouse, rat, dog, and cow [51]. In cluster 33, hsa-miR-410, hsa-miR-377, hsa-miR-668, and hsa-miR-381 are conserved in many mammalian genomes and believed to act cooperatively [51]. Transposable elements (TEs) contribute to the evolution dynamics of miRNAs. We found that hsa-miR-325 and hsa-miR-545 derived from TE L2 [52] were grouped to cluster 30.

Many mammalian viruses have been shown to modulate the expression of host cellular miRNAs [53]. MiRNA expression profiles altered by viruses form a novel regulatory layer, and these miRNAs can be grouped to partially reveal the cross-talk between host and virus. In our clustering result, hsa-miR-181a and hsa-miR-15a that were altered in stable hepatitis B virus expressing cell line [54] were identified in cluster 33. Hsa-miR-24 and hsa-miR-638 that were found to have expression changes during in vitro acute hepatitis C virus infection [55] were grouped in cluster 25.

The clustering results obtained by our method were consistent with many other investigations, suggesting that our method is reliable to calculate functional similarities and sensible to cluster miRNAs. The clustering results are useful to reveal functional diversity of miRNA families.

As described previously, the target genes of miRNA were predicted by PITA algorithm, and our current results were consistent with biological data. Therefore, we can predict novel miRNA functions by miRNA pairs with high similarities or by the GO enrichment analysis of the corresponding miRNA cluster. For instance, we predicted novel miRNA functions with similarities above 0.8. Many of them were supported by newly published literature. For example, hsa-miR-107 and hsa-miR-103 regulate lipid metabolism [56], hsa-miR-449 and hsa-miR-203 regulate pRb-E2F1 activity [57, 58], hsa-miR-200b and hsa-miR-429 regulate EGF-driven invasion [59], etc. We also used the cluster information to globally assign the predicted functions to novel miRNAs. GO enrichment analysis were performed in all miRNA clusters, and all the enriched GO terms represent each miRNA cluster with their corresponding *p*-values and other information were provided in Supplemental File 5. Here, we took cluster 23 as an example. After the hypergeometric test, we select the over-represented GO terms of target genes by p-value < 0.001. Consequently, we can identify miRNAs grouped in cluster 23 to have functions of repressing binding, especially the protein binding, metal ion binding and cation binding. Overall, the method presented in this study can be used to predict potential functions of newly identify miRNA and to discover potential miRNAs involved in important pathways.

4. Concluding remarks

In summary, we proposed a novel method for inferring the functional similarities of human miRNAs by integrating the information provided by miRNA target prediction algorithm PITA with Gene Ontology annotation data. The significant miRNA families were also analyzed. This method can be extended to measure miRNA functional similarities of other species. The current method relies on the prediction of target sites that may contain false positives as well as false negatives and thus may bring bias to some extent. However, target prediction algorithms are necessary when predicting functions of newly identify miRNA. In the future, a more robust and reliable miRNA clustering may be obtained when comprehensive experimental miRNA-targets data are available. We believe that with the rapid increase of experimental miRNA-mRNA deposited in TarBase and the improvement of target prediction algorithms, our method will provide high quality miRNA similarity measurement with high sensitivity and specificity.

5. Supplementary material

Supplementary material regarding this manuscript is online available in the web page of JIOMICS.

http://www.jiomics.com/index.php/jio/rt/suppFiles/21/0

Acknowledgements

This work was partially supported by the 2007 Chang-Jiang Scholars Program, Jinan University "211" Projects (Biotechnology & Bioengineering Medicine and Biomaterial & Tissue Engineering).

References

1. Liu, S., Fu, R., Yu, H., Li, K., et al., MicroRNAs Regulation Modulated Self-Renewal and Lineage Differentiation of Stem Cells. Cell Transplantation 2009, 18, 1039-45.
2. Liang, T., Qin, C., The emerging role of microRNAs in immune cell development and differentiation. APMIS: Acta Pathologica, Microbiologica, Et Immunologica Scandinavica 2009, 117, 635-643.
3. Manni, I., Artuso, S., Careccia, S., Rizzo, M. G., et al., The microRNA miR-92 increases proliferation of myeloid cells and by targeting p63 modulates the abundance of its isoforms. The FASEB Journal 2009, 23, 3957-3966.
4. Wang, Y., Lee, C. G. L., MicroRNA and cancer--focus on apoptosis. Journal of Cellular and Molecular Medicine 2009, 13, 12-23.
5. Wienholds, E., Koudijs, M. J., van Eeden, F. J. M., Cuppen, E., Plasterk, R. H. A., The microRNA-producing enzyme Dicer1 is essential for zebrafish development. Nature Genetics 2003, 35, 217-218.
6. Lu, L., Liston, A., MicroRNA in the immune system, microRNA as an immune system. Immunology 2009, 127, 291-298.
7. Cho, W., OncomiRs: the discovery and progress of microRNAs in cancers. Molecular Cancer 2007, 6, 60.
8. Mcclure, L., Kaposi's Sarcoma Herpes Virus Taps into a Host MicroRNA Regulatory Network. Cell Host & Microbe 2008, 3, 1-3.
9. Griffiths-Jones, S., Saini, H. K., van Dongen, S., Enright, A. J., miRBase: tools for microRNA genomics. Nucl. Acids Res. 2008, 36, D154-158.
10. Griffiths-Jones, S., Bateman, A., Marshall, M., Khanna, A., Eddy, S. R., Rfam: an RNA family database. Nucleic Acids Research 2003, 31, 439-441.
11. Kaczkowski, B., Torarinsson, E., Reiche, K., Havgaard, J. H., et al., Structural profiles of human miRNA families from pairwise clustering. Bioinformatics (Oxford, England) 2009, 25, 291-294.
12. Yu, J., Wang, F., Yang, G. H., Wang, F. L., et al., Human microRNA clusters: genomic organization and expression profile in leukemia cell lines. Biochem Biophys Res Commun 2006, 349, 59-68.
13. Fontana, P., Cestaro, A., Velasco, R., Formentin, E., Toppo, S., Rapid Annotation of Anonymous Sequences from Genome Projects Using Semantic Similarities and a Weighting Scheme in Gene Ontology. PLoS ONE 2009, 4, e4619.
14. Done, B., Khatri, P., Done, A., Drăghici, S., Predicting novel human gene ontology annotations using semantic analysis. IEEE/ACM Transactions on Computational Biology and Bioinformatics / IEEE, ACM 2010, 7, 91-99.
15. Tao, Y., Sam, L., Li, J., Friedman, C., Lussier, Y. A., Information theory applied to the sparse gene ontology annotation network to predict novel gene function. Bioinformatics 2007, 23, i529-538.
16. Mistry, M., Pavlidis, P., Gene Ontology term overlap as a measure of gene functional similarity. BMC Bioinformatics 2008, 9, 327-327.
17. Pesquita, C., Faria, D., Bastos, H., Ferreira, A. E., et al., Metrics for GO based protein semantic similarity: a systematic evaluation. BMC Bioinformatics 2008, 9, S4-S4.
18. Cho, Y., Zhang, A., Xu, X., Semantic similarity based feature extraction from microarray expression data. International Journal of Data Mining and Bioinformatics 2009, 3, 333-345.

19. Bolshakova, N., Azuaje, F., Cunningham, P., A knowledge-driven approach to cluster validity assessment. Bioinformatics 2005, 21, 2546-2547.
20. Wolting, C., McGlade, C. J., Tritchler, D., Cluster analysis of protein array results via similarity of Gene Ontology annotation. BMC Bioinformatics 2006, 7, 338-338.
21. Guo, X., Shriver, C. D., Hu, H., Liebman, M. N., Analysis of Metabolic and Regulatory Pathways through Gene Ontology-Derived Semantic Similarity Measures. 2005, 2005, 972-972.
22. Gamalielsson, J., Olsson, B., Gene Ontology-based semantic alignment of biological pathways by evolutionary search. Journal of Bioinformatics and Computational Biology 2008, 6, 825-842.
23. Lei, Z., Dai, Y., Assessing protein similarity with Gene Ontology and its use in subnuclear localization prediction. BMC Bioinformatics 2006, 7, 491-491.
24. Lord, P. W., Stevens, R. D., Brass, A., Goble, C. A., Investigating semantic similarity measures across the Gene Ontology: the relationship between sequence and annotation. Bioinformatics (Oxford, England) 2003, 19, 1275-1283.
25. Sevilla, J. L., Segura, V., Podhorski, A., Guruceaga, E., et al., Correlation between gene expression and GO semantic similarity. IEEE/ACM Transactions on Computational Biology and Bioinformatics / IEEE, ACM 2005, 2, 330-338.
26. Xu, T., Du, L., Zhou, Y., Evaluation of GO-based functional similarity measures using S. cerevisiae protein interaction and expression profile data. BMC Bioinformatics 2008, 9, 472-472.
27. Wang, D., Wang, J., Lu, M., Song, F., Cui, Q., Inferring the human microRNA functional similarity and functional network based on microRNA-associated diseases. Bioinformatics 2010, 26, 1644-50.
28. Ren, J., Jin, P., Wang, E., Marincola, F., Stroncek, D., MicroRNA and gene expression patterns in the differentiation of human embryonic stem cells. Journal of Translational Medicine 2009, 7, 20.
29. Roubelakis, M., Zotos, P., Papachristoudis, G., Michalopoulos, I., et al., Human microRNA target analysis and gene ontology clustering by GOmir, a novel stand-alone application. BMC Bioinformatics 2009, 10, S20.
30. Papadopoulos, G. L., Reczko, M., Simossis, V. A., Sethupathy, P., Hatzigeorgiou, A. G., The database of experimentally supported targets: a functional update of TarBase. Nucl. Acids Res. 2009, 37, D155-158.
31. Kertesz, M., Iovino, N., Unnerstall, U., Gaul, U., Segal, E., The role of site accessibility in microRNA target recognition. Nat Genet 2007, 39, 1278-1284.
32. Pheasant, M., Meyer, L., Hsu, F., Hinrichs, A. S., et al., The UCSC Genome Browser Database: update 2009. Nucleic Acids Research 2009, 37, D755-761.
33. Yu, G., Li, F., Qin, Y., Bo, X., et al., GOSemSim: an R package for measuring semantic similarity among GO terms and gene products. Bioinformatics 2010, 26, 976-978.
34. Resnik, P., Semantic Similarity in a Taxonomy: An Information-Based Measure and its Application to Problems of Ambiguity in Natural Language. Journal of Artifical Intelligence Research 1998, 11, 130, 95.
35. Lin, D., , Proceedings of the Fifteenth International Conference on Machine Learning, In Proceedings of the 15th International Conference on Machine Learning 1998, pp. 296-304.
36. Jiang, J., Conrath, D., , International Conference Research on Computational Linguistics (ROCLING X) 1997, p. 9008.
37. Schlicker, A., Domingues, F. S., Rahnenführer, J., Lengauer, T., A new measure for functional similarity of gene products based on Gene Ontology. BMC Bioinformatics 2006, 7, 302.
38. Wang, J. Z., Du, Z., Payattakool, R., Yu, P. S., Chen, C., A new method to measure the semantic similarity of GO terms. Bioinformatics (Oxford, England) 2007, 23, 1274-1281.
39. Durinck, S., Spellman, P. T., Birney, E., Huber, W., Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nature Protocols 2009, 4, 1184-1191.
40. Haider, S., Ballester, B., Smedley, D., Zhang, J., et al., BioMart Central Portal--unified access to biological data. Nucleic Acids Research 2009, 37, W23-27.
41. Suzuki, R., Shimodaira, H., Pvclust: an R package for assessing the uncertainty in hierarchical clustering. Bioinformatics 2006, 22, 1542, 1540.
42. Liang, Y., Ridzon, D., Wong, L., Chen, C., Characterization of microRNA expression profiles in normal human tissues. BMC Genomics 2007, 8, 166.
43. Wijnhoven, B. P. L., Hussey, D. J., Watson, D. I., Tsykin, A., et al., MicroRNA profiling of Barrett's oesophagus and oesophageal adenocarcinoma. British Journal of Surgery 2010, 97, 853-861.
44. Park, J. K., Lee, E. J., Esau, C., Schmittgen, T. D., Antisense inhibition of microRNA-21 or -221 arrests cell cycle, induces apoptosis, and sensitizes the effects of gemcitabine in pancreatic adenocarcinoma. Pancreas 2009, 38, e190-e199.
45. Chung, T. K., Cheung, T., Huen, N., Wong, K. W., et al., Dysregulated microRNAs and their predicted targets associated with endometrioid endometrial adenocarcinoma in Hong Kong women. International Journal of Cancer 2009, 124, 1358-1365.
46. Wu, W., Lin, Z., Zhuang, Z., Liang, X., Expression profile of mammalian microRNAs in endometrioid adenocarcinoma. Eur J Cancer Prev 2009, 18, 50-55.
47. Mees, S. T., Mardin, W. A., Wendel, C., Baeumer, N., et al., EP300 - A miRNA-regulated metastasis suppressor gene in ductal adenocarcinomas of the pancreas. International Journal of Cancer 2010, 126, 114-124.
48. Tchatchou, S., Jung, A., Hemminki, K., Sutter, C., et al., A variant affecting a putative miRNA target site in estrogen receptor (ESR) 1 is associated with breast cancer risk in premenopausal women. Carcinogenesis 2009, 30, 59-64.
49. Zhang, L., Huang, J., Yang, N., Greshock, J., et al., microRNAs exhibit high frequency genomic alterations in human cancer. Proceedings of the National Academy of Sciences 2006, 103, 9136-9141.
50. Dai, Y., Sui, W., Lan, H., Yan, Q., et al., Comprehensive analysis of microRNA expression patterns in renal biopsies of lupus nephritis patients. Rheumatol Int 2009, 29, 749-754.
51. Glazov, E. A., McWilliam, S., Barris, W. C., Dalrymple, B. P., Origin, Evolution, and Biological Role of miRNA Cluster in DLK-DIO3 Genomic Region in Placental Mammals. Mol Biol Evol 2008, 25, 939-948.
52. Piriyapongsa, J., Marino-Ramirez, L., Jordan, I. K., Origin and Evolution of Human microRNAs From Transposable Elements. Genetics 2007, 176, 1323-1337.
53. Scaria, V., Hariharan, M., Pillai, B., Maiti, S., Brahmachari, S. K., Host-virus genome interactions: macro roles for microRNAs. Cellular Microbiology 2007, 9, 2784-2794.
54. Liu, Y., Zhao, J. J., Wang, C. M., Li, M. Y., et al., Altered expression profiles of microRNAs in a stable hepatitis B virus-expressing cell line. Chin Med J (Engl) 2009, 122, 10-14.
55. Liu, X., Wang, T., Wakita, T., Yang, W., Systematic identification of microRNA and messenger RNA profiles in hepatitis C virus-infected human hepatoma cells. Virology 2010, 398, 57-67.
56. Wilfred, B. R., Wang, W., Nelson, P. T., Energizing miRNA research: a review of the role of miRNAs in lipid metabolism, with a prediction that miR-103/107 regulates human metabolic pathways. Molecular genetics and metabolism 2007, 91, 209-

217.

57. Yang, X., Feng, M., Jiang, X., Wu, Z., et al., miR-449a and miR-449b are direct transcriptional targets of E2F1 and negatively regulate pRb–E2F1 activity through a feedback loop by targeting CDK6 and CDC25A. Genes & Development 2009, 23, 2388-2393.
58. Shi, Y., Zhu, J., Salomoni, P., Tucci, P., Lu, X., A Sino-British frontier workshop of cancer biology. Cell Death Differ 2008, 16, 648-650.
59. Uhlmann, S., Zhang, J. D., Schwäger, A., Mannsperger, H., et al., miR-200bc/429 cluster targets PLCgamma1 and differentially regulates proliferation and EGF-driven invasion than miR-200a/141 in breast cancer. Oncogene 2010, 29, 4297-4306.