An Integrated Analysis Method for miRNA, lncRNA and mRNA Profiles Based on Their Functional and Positional Relationships*

Li Guo, Sheng Yang, Yang Zhao, Hui Zhang, Qian Wu, Feng Chen[#]

Department of Epidemiology and Biostatistics and Ministry of Education Key Lab for Modern Toxicology,
School of Public Health, Nanjing Medical University, Nanjing, China
Email: lguo@njmu.edu.cn, [#]fengchen@njmu.edu.cn

ABSTRACT

ncRNAs have been identified as potential regulatory molecules and have multiple biological roles. Aberrant expression of specific ncRNAs contributes to multiple biological processes and many human diseases. Herein, we simultaneously profiled miRNA, lncRNA and mRNA in human HepG2 and L02 cells applying high-throughput sequencing and microarray technologies. Abnormal miRNA, lncRNA and mRNA profiles were assessed through fold change filtering. A cross-platform integrated analysis method was developed to analyze differentially expressed miRNA, lncRNA and mRNA profiles. miRNA-mRNA interaction was analyzed according to their functional relationships. Target mRNAs of aberrantly expressed miRNAs were obtained from experimentally validated datasets or predicted using some programs. Generally, multiple target mRNAs were involved, and they have versatile roles by functional enrichment analysis. According to actual expression datasets in the study, compared to deregulated miRNAs, these theoretical target mRNAs showed various expression patterns. The consistent or inconsistent expression was mainly derived from complex, multiple, flexible and alternative regulatory relationships between miRNA and mRNA. Further, miRNA/mRNA and lncRNA were completely surveyed based on their location distributions on human chromosomes. Many miRNA-lncRNA and mRNA-lncRNA pairs always were located on the same strand or different strands in the specific genomic region. Due to the location distributions, they might have partly or completely overlapped regions or they could be reverse complementarily binding. These miRNA/mRNA-lncRNA pairs showed consistent or inconsistent expression patterns, although they might have functional relationships through reverse complementarily binding events. Moreover, we also detected and analyzed various isomiRs from a given miRNA locus, including those isomiRs with 3' additional non-template nucleotides. These isomiRs, especially for those 5' isomiRs with the new "seed sequences" through "seed shifting" events, maybe have potential biological roles as well as isomiR repertoire and their expression patterns. The integrative analysis provides potential functional relationships between miRNA, lncRNA and mRNA across different datasets. The complex and various expression patterns suggest a robust regulatory network across different regulatory molecules and their targets.

Keywords: ncRNA; mRNA; Profile; Integrative Analysis

1. Introduction

As important players of gene regulation, non-coding RNAs (ncRNAs), have attracted considerable interests in many fields. Different lines of evidence suggest that ncRNAs play an important role in multiple biological processes and contribute to occurrence and development of many human diseases, including microRNAs (miR-NAs) and long non-coding RNAs (lncRNAs). miRNAs contribute to biological processes by targeting mRNAs through conserved complementarily to the seed (nucleotides 2 - 8) of the miRNA [1]. Similarly, lncRNAs have a broad range of biological roles in regulation of expression of genes and chromosomes. As potential regulators, they contribute to various basic cellular functions, such as cell proliferation, differentiation, death and tumorigenesis [2].

Concerning the potential relationships between different molecules, especially for their contributions to complex regulation network in multiple biological processes, cross-platform analysis has been a new hot focus. How-

ever, it is difficult to unveil the potential feedback loops across different molecules in specific space and time in regulatory network. Here, based on potential positional and functional relationships between different molecules, we developed a method to integrative analyze datasets of ncRNA-mRNA from different cells applying high-through-put technologies. The sequence and potential functional correlations and expression profiles of different RNA molecules can provide more complex interaction and biological roles across molecules *in vivo*. The integrative analysis method will systematically reveal the complex relationships in regulatory network and contributions in tumorigenesis.

2. Materials and Methods

High-throughput datasets of miRNA, lncRNA and mRNA were obtained from HepG2 and L02 cells applying Solexa sequencing platform and microarray technologies, respectively. miRNAs were identified from the raw miRNA datasets using Novoalign software (v2.07.11, http://www.novocraft.com) based on the known human miRNA precursors (pre-miRNAs) in the miRBase database (Release 18.0, http://www.mirbase.org/) [3]. The only one mismatch was allowed in the mapping analysis. Specially, the multiple isomiRs [4-7] from a given miRNA locus were also comprehensively surveyed and analyzed. Those sequences that could match the pre-miRNAs in the known mature miRNA or miRNA* region ± 4 nt were defaulted as isomiRs due to alternative and imprecise cleavage of Drosha and Dicer through pre-miRNA processing. These isomiRs may be detected with 3' additional non-template nucleotide (3' addition event).

LncRNA and mRNAs were obtained from array images using agilent Feature Extraction software (version 10.7.3.1) after quantile normalization of the raw data. Fold change was calculated to obtain the differentially expressed miRNA, lncRNA and mRNA profiles. Further, isomiR profiles were also obtained based on potential various sequences.

Herein, the integrated analysis method was developed based on the potential functional and positional relationships between ncRNAs and mRNAs (**Figure 1**). The self-developed scripts were used to survey and obtain potential relationships between mRNA-miRNA, miRNA-lncRNA and mRNA-lncRNA based on location distributions of aberrantly expressed ncRNA and mRNA profiles. Expression patterns of these pairs were further analyzed in HepG2 and L02 cells, especially for their deregulation patterns in tumor cells. Functional enrichment analysis, pathway and GO analysis were applied to define the potential biological roles of these differentially expressed functional mRNA molecules.
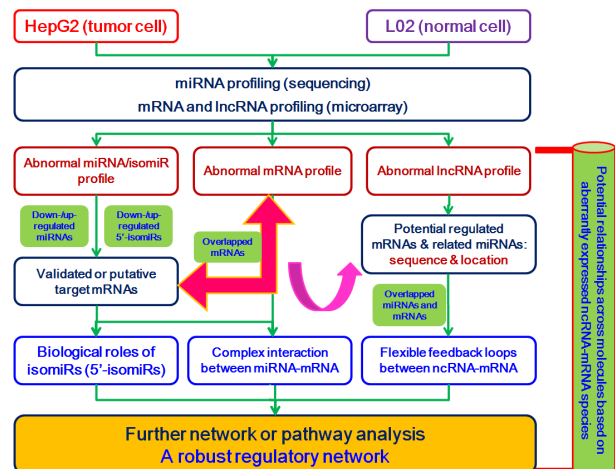

Figure 1. A flowchart indicates the integrative analysis of ncRNA-mRNA based on their potential relationships.

3. Results and Discussion

Based on the original miRNA, lncRNA and mRNA profiles, the aberrantly expressed ncRNA and mRNA profiles were obtained in HepG2 cells according to fold change filtering. Multiple isomiRs from a given miRNA locus, including those isomiRs with 3' additional non-template nucleotides, were also obtained the abnormal isomiR profiles. Firstly, according to the experimentally validated target mRNAs in the miRTarBase database [8], all the targets of aberrantly expressed miRNAs were collected. For those miRNAs that had no or less validated mRNAs, we predicted their targets using the TargetScan program [9]. IsomiRs might have various 5' and/or 3' ends due to alternative cleavage of Drosha and Dicer through pre-miRNA processing, especially 5' isomiRs will have the new "seed sequences" with "seed shifting" events [7,10]. The new seed sequences might lead to the shifting or change of their target mRNAs. Therefore, the potential novel target mRNAs of dominantly expressed 5' isomiRs were also predicted and further analyzed. Although these 5' isomiRs and isomiRs with 3' addition always were not the abundant sequences from a given miRNA locus, they might show unexpectedly higher relative expression levels. The functional enrichment analysis of their targets indicated that they have important roles in multiple biological processes, including some human diseases, such as cell cycle and Prostate cancer (**Table 1**). Moreover, we also found inconsistent isomiR repertoires and expression patterns in tumor and control cells based on specific miRNA locus (data not shown) [7]. The interesting phenomenon indicated potential biological roles of isomiR repertoires in regulatory network.

According to these abnormal miRNAs and isomiRs, a series of target mRNAs (experimentally validated or predicted target mRNAs) were obtained. Theoretically,

Table 1. An example of pathway enrichment analysis of mRNA targets of deregulated miRNAs.

Pathway	No.	P-value	Target Genes
Cell cycle	18	3.01E−30	ATM; **CCNA2**; CCND1; **CCND2**; CCNE1; CDC25A; CDK6; CDKN1A; CDKN1B; CDKN2A; E2F1; **E2F2**; E2F3; EP300; RB1; **RBL2**; TP53; WEE1
Prostate cancer	15	3.74E−26	AKT1; BCL2; CCND1; CCNE1; CDKN1A; CDKN1B; E2F1; **E2F2**; E2F3; EP300; IGF1R; NFKB1; **NRAS**; RB1; TP53
Pancreatic cancer	14	3.03E−25	ACVR1C; AKT1; CCND1; **CDC42**; CDK6; CDKN2A; E2F1; **E2F2**; E2F3; NFKB1; RAC1; RB1; TP53; VEGFA
Melanoma	13	3.21E−23	AKT1; CCND1; CDK6; CDKN1A; CDKN2A; E2F1; **E2F2**; E2F3; IGF1R; **MET**; **NRAS**; RB1; TP53

These target mRNAs are regulated by at least 2 abnormal miRNAs. mRNAs in bold type indicate up-regulated species, underlined mRNAs indicate down-regulated species, and others indicate that they are stably expressed or not detected.

these mRNAs might be targeted by these aberrant miR-NA/isomiRs and lead to abnormal expression levels by miRNA-mRNA interaction. Indeed, consistent or inconsistent expression patterns were detected from the actual aberrantly expressed mRNA profiles (**Table 1**). Although miRNAs were down- or up-regulated in tumor cells, their target mRNAs might be up- or down-regulated, or stably expressed (**Table 1**). The phenomenon is mainly derived from complex and multiple regulatory patterns between miRNAs and mRNAs. Generally, one miRNA can target a great amount of mRNAs, whereas one mRNA can be targeted by a series of miRNAs by miRNA-mRNA interaction. The complex feedback loops contribute to the whole flexible regulatory network.

Compared to miRNA, another non-coding regulatory molecule, lncRNA, still remains mysterious. Herein, we obtained the potential relationships between lncRNA-mRNA and lncRNA-miRNA based on sequences and location distributions. Some miRNA-lncRNA and mRNA-lncRNA pairs were surveyed based on their locations on human chromosomes. They might be located on different strands in the same genomic region, or located on the same strand with the completely or partly over-lapped regions. Based on the relationships of sequences or locations, miRNA/mRNA and lncRNA could have potential functional relationships. The expression analysis indicated that these pairs showed consistent or inconsistent expression patterns, even though they could be reverse complementarily binding (from the sense/antisense strands). The results were similar to miRNA-mRNA interaction. Although different RNA molecules may have functional relationships, they might show various expressions as well as consistent or inconsistent deregulated expression patterns. Among of these, consistent expression patterns between miRNA-lncRNA and mRNA-lncRNA were more popular.

The genome-wide analysis of ncRNA-mRNA based on their positional and functional relationships provides a systematical integrative method to unveil potential relationships across different RNA molecules, including ncRNAs (as regulatory RNA molecules) and mRNAs (as functional RNA molecules). Here, based on aberrantly expressed miRNA, lncRNA and mRNA profiles in tumor cells, the integrative analysis suggested a robust regulatory network in tumorigenesis.

REFERENCES

[1] D. P. Bartel, "MicroRNAs: Genomics, Biogenesis, Mechanism, and Function," *Cell*, Vol. 116, 2004, pp. 281-297. http://dx.doi.org/10.1016/S0092-8674(04)00045-5

[2] H. W. Hwang and J. T. Mendell, "MicroRNAs in Cell Proliferation, Cell Death, and Tumorigenesis," *British Journal of Cancer*, vol. 94, 2006, pp. 776-780. http://dx.doi.org/10.1038/sj.bjc.6603023

[3] A. Kozomara and S. Griffiths-Jones, "miRBase: Integrating microRNA Annotation and Deep-Sequencing Data," *Nucleic Acids Research*, Vol. 39, 2011, pp. D152-D157. http://dx.doi.org/10.1093/nar/gkq1027

[4] S. L. Fernandez-Valverde, R. J. Taft and J. S. Mattick, "Dynamic isomiR Regulation in Drosophila Development," *RNA—A Publication of the RNA Society*, Vol. 16, 2010, pp. 1881-1888. http://dx.doi.org/10.1261/rna.2379610

[5] L. W. Lee, S. Zhang, A. Etheridge, L. Ma, D. Martin, D. Galas and K. Wang, "Complexity of the microRNA Repertoire Revealed by Next Generation Sequencing," *RNA—A Publication of the RNA Society*, Vol. 16, 2010, pp. 2170-2180. http://dx.doi.org/10.1261/rna.2225110

[6] A. M. Burroughs, Y. Ando, M. J. L. de Hoon, Y. Tomaru, T. Nishibu, R. Ukekawa, T. Funakoshi, T. Kurokawa, H. Suzuki, Y. Hayashizaki and C. O. Daub, "A Comprehensive Survey of 3' Animal miRNA Modification Events and a Possible Role for 3' Adenylation in Modulating miRNA Targeting Effectiveness," *Genome Research*, Vol. 20, 2010, pp. 1398-1410. http://dx.doi.org/10.1101/gr.106054.110

[7] L. Guo, Q. Yang, J. Lu, H. Li, Q. Ge, W. Gu, Y. Bai and Z. Lu, "A Comprehensive Survey of miRNA Repertoire and 3' Addition Events in the Placentas of Patients with Pre-eclampsia from High-Throughput Sequencing," *PLoS ONE*, Vol. 6, 2011, Article ID: e21072. http://dx.doi.org/10.1371/journal.pone.0021072

[8] S. D. Hsu, F. M. Lin, W. Y. Wu, C. Liang, W. C. Huang and W. L. Chan, "miRTarBase: A Database Curates Ex-

ENG

perimentally Validated microRNA-Target Interactions," *Nucleic Acids Research*, Vol. 39, 2011, pp. D163-D169. http://dx.doi.org/10.1093/nar/gkq1107

[9] B. P. Lewis, I. H. Shih, M. W. Jones-Rhoades, D. P. Bartel and C. B. Burge, "Prediction of Mammalian microRNA Targets," *Cell*, Vol. 115, 2003, pp. 787-798. http://dx.doi.org/10.1016/S0092-8674(03)01018-3

[10] L. Guo, H. Li, J. Lu, Q. Yang, Q. Ge, W. Gu, Y. Bai and Z. Lu, "Tracking miRNA Precursor Metabolic Products and Processing Sites through Completely Analyzing High-Throughput Sequencing Data," *Molecular Biology Reports*, Vol. 39, 2012, pp. 2031-2038. http://dx.doi.org/10.1007/s11033-011-0950-8