Video Article

Describing a Transcription Factor Dependent Regulation of the MicroRNA Transcriptome

Uri Rozovski[1], Inbal Hazan-Halevy[2], George Calin[3], David Harris[4], Ping Li[4], Zhiming Liu[4], Michael J. Keating[4], Zeev Estrov[4]

[1]Division of Hematology, Davidoff Cancer Center, Rabin Medical Center

[2]The Center of Nanoscience and Nanotechnology, Tel Aviv University

[3]Experimental Therapeutics, The University of Texas MD Anderson Cancer Center

[4]Department of Leukemia, The University of Texas MD Anderson Cancer Center

Correspondence to: Zeev Estrov at zestrov@mdanderson.org

URL: https://www.jove.com/video/53300
DOI: doi:10.3791/53300

Keywords: Bioengineering, Issue 112, microRNA (miR), miR transcriptome, transcription factor, miR promoter, data mining, transcription regulation

Date Published: 6/15/2016

Citation: Rozovski, U., Hazan-Halevy, I., Calin, G., Harris, D., Li, P., Liu, Z., Keating, M.J., Estrov, Z. Describing a Transcription Factor Dependent Regulation of the MicroRNA Transcriptome. *J. Vis. Exp.* (112), e53300, doi:10.3791/53300 (2016).

Abstract

While the transcription regulation of protein coding genes was extensively studied, little is known on how transcription factors are involved in transcription of non-coding RNAs, specifically of microRNAs. Here, we propose a strategy to study the potential role of transcription factor in regulating transcription of microRNAs using publically available data, computational resources and high throughput data. We use the H3K4me3 epigenetic signature to identify microRNA promoters and chromatin immunoprecipitation (ChIP)-sequencing data from the ENCODE project to identify microRNA promoters that are enriched with transcription factor binding sites. By transfecting cells of interest with shRNA targeting a transcription factor of interest and subjecting the cells to microRNA array, we study the effect of this transcription factor on the microRNA transcriptome. As an illustrative example we use our study on the effect of STAT3 on the microRNA transcriptome of chronic lymphocytic leukemia (CLL) cells.

Video Link

The video component of this article can be found at https://www.jove.com/video/53300/

Introduction

MicroRNAs are endogenous small non coding regulatory RNAs that typically function as negative regulators of mRNA expression at the posttranscriptional level. Approximately 1,000 non-coding 20 to 25 nucleotide long microRNAs are found in the human genome [1,2]. MicroRNAs regulate gene expression through canonical base pairing between the seed sequence of the microRNA and its complementary seed match sequence, which is commonly located at the 3' untranslated region (UTR) of the target mRNAs. Collectively, microRNAs regulate more than 30% of protein coding genes [3], but only little is known about the transcription from DNA of microRNAs. It has been suggested that the regulation of microRNA transcription is similar to that of mRNA [4,5]. In particular, similar to its activity in promoting transcription of protein coding genes, transcription factors are thought to activate transcription of microRNAs [6]. Transcription factor-microRNA interplay has been reported as a modulator factor of gene expression [7], and may also form feed-back and feed-forward loops. For example, Yamakuchi *et al.* reported a feedback loop in which p53 induces the expression of microRNA34a, which in turn inhibits translation of the p53 repressor SIRT and thereby increasing p53 activity [8].

Whereas specific examples of transcription factor dependent expression of microRNAs have been reported, an accepted method which provides information on how a transcription factor of interest regulates the expression of the microRNA-transcriptome is lacking. The purpose of the protocol suggested herein is to provide an in-depth description of transcription factor-dependent regulation of the microRNA-transcriptome. By combining publically available data, bioinformatics tools and using microarray technology, researchers who follow this algorithm would be able to capture on a genomic scale how any transcription factor in any cell type of interest regulates the expression of the microRNA-transcriptome and to explore a putative contribution of the transcription factor-mRNA in regulating microRNA expression.

Protocol

1. Identify Transcription Factor Binding Sites in the Promoter of MicroRNA Genes Using Data Mining Approach

1. Use the University of California Santa Cruz (UCSC) genome browser to extract chromatin immunoprecipitation (ChIP) sequencing data generated as part of the Encyclopedia of DNA element (ENCODE) project.

1. Open the table browser in the UCSC genome browser.
2. Use the following specifications to extract the table: Clade: (Mammals), genome: (Human), Assembly: (Feb2009(GRch37/hg18)), group: (regulation), track (TxnFactorChIP), table: (weEncoderesTFbsCo7steredv3), region: (Genome), output format (all genes from selected table).
3. Save the output from 1.1.2 as a .txt file and into a spreadsheet.
4. Sort and filter for the transcription factor of interest (*e.g.*, STAT2).

2. Use the list of microRNA promoters based on H3K4me3 epigenetics signature available at Baeer *et al.*[9]. Copy this list into a .txt file.
3. To match according to coordinates on the human genome, map the data from 1.1 and 1.2 (*e.g.*, STAT3 binding on putative microRNA promoters) and determine the median binding affinity using the code written in C sharp as outlined in supplemental coding file 1.

2. Use shRNA to Down-regulate the Expression of a Transcription Factor of Interest

1. Plate 1.5 x 10^6 cells from 293 human embryonic kidney cell line in 10 cm plate at approximately 50% confluence (DMEM with 10% FBS).
2. Transfect cells from 293 human embryonic kidney cell line with 5 µg of green fluorescence protein (GFP) lentivirus containing shRNA directed to the transcription factor of interest and with 5 µg of packaging vectors using transfection reagent for adherent cells according to manufacturer's protocol.
3. As a control, transfect the cells from 293 human embryonic kidney cell line with scrambled shRNA and the packaging vectors according to manufacturer's protocol.
4. Keep transfection mix on cells for 16 hr (at 37 °C, CO_2 incubator), then change media to 10 ml fresh 10% DMEM media with 10% FBS.
5. Wait 48 hr post transfection, centrifuge the cell culture (300 x g, 5 min) and collect infectious supernatant. Filter the supernatant through a 0.45 µm syringe filter (25-mm surfactant free cellulose acetate membrane) to remove any floating cells.
6. Concentrate the supernatant and collect the lentivirus using an ultracentrifugal filter device with threshold of 100 kDa. Spin at 950 x g for 30 - 60 min until the volume has been concentrate to less than 250 µl. Store the concentrated virus at -80 °C.
7. Transfect the cells with the lentivirus. Remove frozen lentivirus from -80 °C freezer and thaw to room temperature. Transfer 100 µl of viral supernatant to a fresh 1.5 ml microfuge tube.
    1. Bring up the volume in the tube to 1 ml with reduced serum medium. Add hexadimethrine bromide to 1 ml virus suspension for final concentration of 10 ng/ml, mix gently and let the mixture stand for 5 min.
8. Centrifuge 5 x 10^6 cells for each transduction and gently resuspend the cell pellet in 0.5 ml of media containing virus. Let the cells stay in the incubator for 4 - 24 hr, then add 0.5 ml of medium with 20% FBS to a final concentration of 10% FBS.
9. Wait 48 to 72 hr and stain the cells with propidium iodide (PI) and green florescent protein (GFP) according to the manufacturer's instructions. Protect the cells from light and use a FACS sorter to measure the rates of GFP^+ / PI^- cells. Since PI stains only dead cell, this rate is an estimate of transfection efficiency in living cells.
10. Sort the positive cell population to GFP expression (GFP+) by FACS sorter as previously described [10].
11. Use Western immune blotting as previously described [10] to determine the levels of a transcription factor of interest before and after infecting the cells of interest (for example, CLL cells) with designated shRNA.

3. Determine the Expression Level of MicroRNA Transcriptome in Cells Transfected with Transcription Factor-shRNA

1. Isolate RNA using a commercial kit according to manufacturer's protocol.
2. Label the RNA and hybridize it to microRNA microarray[11].
3. Determine the differentially expressed microRNA in cells transfected with transcription factor-shRNA or with empty vector controls[5].
4. Validate the microarray results for the most differentially expressed microRNAs using real-time PCR [5].

4. Determine the Overlap between the Bioinformatics and the shRNA Approach in Describing the Transcription Factor Dependent Transcriptome

1. To determine the **expected** and **observed** ratios of microRNA genes that harbor transcription factor binding sites in their promoters and were downregulated in transcription factor-shRNA transfected cells, do the following:
    1. Obtain the ratio of genes that harbor the transcription factor of interest in their promoter / total microRNA from the list generated in 1.3. This list is the **expected** ratio (*e.g.*, microRNA genes with STAT3 binding sites / total microRNA genes tested = 0.25).
    2. Determine the number of genes that were downregulated in transcription factor-shRNA transfected cells from the list generated in 3.3 **AND** has transcription factor of interest binding sites from the list generated in 1.3. This number/Total number of downregulated genes is the **observed** ratio (*e.g.*, microRNA genes with STAT3 binding sites/total number of downregulated microRNAs = 0.6).

2. Use $\chi 2$ statistics to compare the observed and expected ratios that were generated above and determine whether the list of genes that were downregulated in transcription factor-shRNA transfected cells are enriched with transcription factor binding sites in their promoter.

Representative Results

STAT3 is a transcription factor which typically induces the transcription of genes that have anti apoptotic and proliferative effects [12] . Whether STAT3 also affect the non-coding RNA transcriptome is currently unknown. In all CLL cells STAT3 is constitutively phosphorylated on serine 707 residues [10,13]. Phosphoserine STAT3 shuttles to the nucleus, binds to DNA, and activates genes known to be activated by tyrosine pSTAT3 in other cell types [10]. Because CLL is characterized by global deregulation of the microRNA network [14], we hypothesized that the presence of serine pSTAT3 affects the expression of microRNAs in CLL cells.

To test this hypothesis promoters of microRNAs that harbor STAT3 binding sites had to be identified. By crossing data generated by Baer *et al.* [9] of regions with the H2K4me3 histone modification which characterize promoter sites, with STAT3 binding sites identified by ChIP-seq experiment [15], putative STAT3 binding sites were identified. Using this approach 160 putative promoters were detected in nearly 25% of the microRNA genes examined (N=200) with binding scores ranging from 100 (the lowest score) to 1,000 (the highest score) (**Table 1**).

Subsequently CLL cells were transfected with STAT3-shRNA or with an empty vector and using microRNA array and identified 63 microRNAs that were down regulated following the transfection (**Figure 1**) suggesting that STAT3 promotes the transcription of these microRNAs. For 60% of the 63 downregulated microRNA genes (n = 38) ChIP-seq data confirmed STAT3 binding in a putative promoter upstream of the gene location, significantly more than expected by chance (p<0.0001). Nine microRNAs that were down-regulated after the transfection, suggesting that STAT3 negatively regulate its levels of transcription.

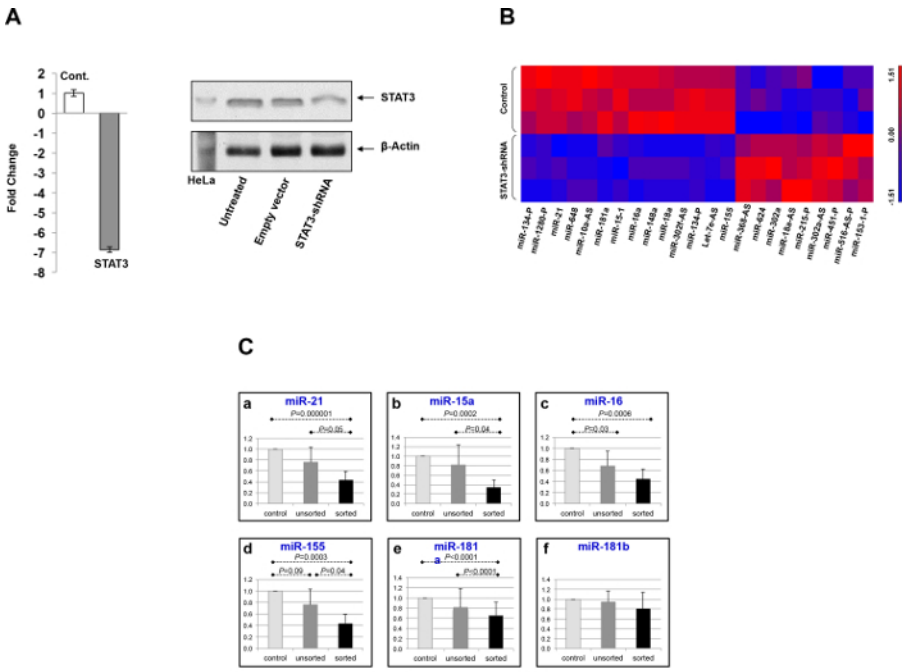

Figure 1. Transfection of CLL cells with STAT3-shRNA reduces the expression levels of STAT3 protein and STAT3 mRNA. A: After transfection of CLL cells with STAT3-shRNA levels of STAT3 mRNA (left panel, detected by Quantitative RT-PCR) and levels of the STAT3 protein (right panel, detected by western immunoblotting) significantly decreased. **B**: microRNA array of CLL cells depicted 23 microRNAs whose expression differed significantly between CLL cells transfected with STAT3-shRNA and CLL cells transfected with an empty vector. P value of less than 0.01 was considered statistically significant. **C**: The mean expression quantified by RT-PCR of 7 microRNAs who had differential expression after STAT3-shRNA treatment using the microRNA array. Bars represent the standard error of the mean. Please click here to view a larger version of this figure.

Supplemental Coding File 1. Please click here to download this file.

Micro RNA gene	Chromosome	Promoter start coordinates	Promoter end coordinates	Median (range) STAT3 binding score*
miR-1205, miR-1206,miR-1207	8 q24.21	128961454	128962791	1000 (1000-1000)
miR-1537	1 q42.3	236045425	236047415	1000 (1000-1000)
miR-21	17 q23.1	57901872	57921277	1000 (112-1000)
miR-3124	1 q44	249115404	249123965	1000 (1000-1000)
miR-451	17q11.2	27222251	27224114	1000 (1000-1000)
miR-92b	1 q22	155162340	155168439	1000 (1000-1000)
miR-3197	21 q22.2	42537544	42543023	943 (943-943)
miR-646	20 q13.33	58712550	58715320	789 (789-789)
miR-629	15 q23	70383751	70394586	773 (661-885)
miR-30e, miR-30c-1	1 p34.2	41173077	41177703	759 (759-759)
miR-3125	2 p24.3	12855381	12862915	756 (756-756)
miR-3145	6 q23.3	138776942	138779365	743 (487-1000)
miR-645	20 q13.13	49199911	49201187	743 (743-743)
miR-1256	1 p36.12	21346830	21350211	725 (725-725)
miR-619	12 q24.11	109248263	109253306	719 (719-719)
miR-181a-2, miR-181b-2	9 q33.3	127418928	127426139	710 (710-710)
miR-29a, miR-29b-1	7 q32.3	130583383	130597803	697 (482-1000)
miR-202	10 q26.3	135069499	135077337	696 (393-1000)
miR-3142, miR-146a	5 q34	159890882	159899475	671 (671-671)
miR-548c	12 q14.2	65000968	65011503	660 (660-660)
miR-630	15 q24.1	72764289	72769197	627 (255-1000)
miR-135b	1 q32.1	205416952	205452990	622 (245-1000)
miR-29c, miR-29b-2	1 q32.2	207991044	208002382	608 (608-608)
miR-1825	20 q11.21	30791020	30798310	604 (209-1000)
miR-548h-1	14 q23.2	64578834	64581657	587 (174-1000)
miR-612	11 q13.1	65183633	65198528	581 (157-1000)
miR-148b	12 q13.13	54717640	54721204	578 (578-578)
miR-3174	15 q26.1	90543381	90549092	576 (152-1000)
let7a-3, let7b	22	46480680	46481826	573 (146-1000)
miR-1255a	4 q24	102263848	102272541	557 (557-557)

Table 1. Putative microRNA's promoters with STAT3 binding sites. The protocol provides a method to identify transcription factor binding on putative promoters of microRNAs using data mining of published data. As an example, we present here a table depicting the binding of STAT3 to putative microRNA promoters. The binding score is given at a 0 to 1,000 scale from whole genome ChIP seq data published as part of the ENCODE project [15]. The promoters are identified by the presence of the H3K4me3 epigenetic signature [9].

Discussion

The mechanism underlying the RNA polymerase II- dependent transcription of protein coding genes has been extensively studied. While these elements make up only 1% - 2% of the human genome, evidence from the ENCODE project suggest that over 80% of the human genome may undergo transcription [17] and what regulates the transcription of the non-coding DNA elements remains largely unknown [6].

Several studies, which indicated that Pol II is also responsible for the transcription of some non-protein-coding genes including microRNAs [6], led us to develop a strategy that combine data from publically available resources, computational algorithms and *in vitro* studies to decipher the potential function of a transcription factor of interest in regulating transcription of microRNAs. The strategy suggested herein includes 2 critical steps. First we use publically available data to identify transcription factor binding in microRNA promoters. To that end we use Chip-seq data published by the ENCODE consortium to identify transcription factor binding and epigenetic signature that typifies promoter as an indirect marker of microRNA-promoters. Crossing the genetic coordinates from these datasets provides a gross estimation of how frequent a transcription factor

of interest binds to microRNA-promoter. Second by using shRNA technology to silence the expression of a transcription factor and subjecting the cells to microRNA-array, it is possible to explore the functional significance of a transcription factor on the microRNA-transcriptome.

The variability in microRNA expression is only partially explained by transcription factor dependent regulation. Stoichiometric variability and other cellular or extracellular factors play an important role and are not simulated in the proposed algorithm. Other limitations include the following: The promoter analysis based on the H3K4me3 signature was done on 939 annotated microRNAs. Since then the genomic locations of many more microRNAs have been identified. However to the best of our knowledge a more comprehensive list that is based on an updated database has not been published yet. Of the 939 microRNA genes, the H3K4me3 which typifies the promoter region was identified in 781 microRNA genes (83%). Hence, while this analysis is clearly based on an incomplete dataset it captures a significant fraction of the microRNA-transcriptome.

Moreover, epigenetics signature is in part imprinted and in part cell-specific. Therefore, the generability of putative promoters that were defined by epigenetics markers may be questioned. Because H3K4me3 persists independent of transcription [18] it is generally considered a marker of imprinted promoters. The analytic algorithm we propose herein may therefore miss cell-specific microRNA promoters if these promoters were identified at a different cell-type. Finally, any conclusion should be tested and confirmed empirically. Most notably, the association between down regulation of a transcription factor (using shRNA approach) and microRNA expression (identified by microRNA array) can only suggest a direct transcriptional role that should be confirmed by acceptable assays such as chromatin immunoprecipitation (ChIP) or electrophoretic mobility shift assay (EMSA). Modifications to the method suggested herein include different ways of identifying microRNA promoter and different ways of knocking down the expression of a transcription factor, for example, small interfering RNA instead of shRNA. The transcription regulation of microRNAs may be substantially different for microRNAs that reside within protein coding genes (intragenic microRNAs) and those that do not (intergenic microRNAs). Because intragenic microRNAs are commonly transcribed in conjunction with their host genes [19] the promoter is usually found immediately upstream the transcription start site. However for many intergenic microRNAs the transcription start site is poorly annotated, and prediction tools that are commonly used for protein coding genes perform poorly [20].

Disclosures

The authors declare no competing financial interests.

Acknowledgements

This study was supported by a grant from the CLL Global Research Foundation. The University of Texas MD Anderson Cancer Center is supported in part by the National Institutes of Health through a Cancer Center Support Grant (P30CA16672).

References

1. Bentwich, I. *et al.* Identification of hundreds of conserved and nonconserved human microRNAs. *Nat Genet* **37**, 766-770 (2005).
2. Hata, A. Functions of microRNAs in cardiovascular biology and disease. Annu Rev Physiol. **75**, 69-93 (2013).
3. Bartel, D. P. MicroRNAs: target recognition and regulatory functions. *Cell.* **136**, 215-233 (2009).
4. Piriyapongsa, J., Jordan, I. K., Conley, A. B., Ronan, T., & Smalheiser, N. R. Transcription factor binding sites are highly enriched within microRNA precursor sequences. *Biol Direct.* ***6***, 61 (2011).
5. Rozovski, U. *et al.* Signal transducer and activator of transcription (STAT)-3 regulates microRNA gene expression in chronic lymphocytic leukemia cells. *Mol Cancer.* ***12***, 50 (2013).
6. Turner, M. J., & Slack, F. J. Transcriptional control of microRNA expression in C. elegans: promoting better understanding. *RNA Biol.* **6**, 49-53 (2009).
7. Cui, Q., Yu, Z., Pan, Y., Purisima, E. O., & Wang, E. MicroRNAs preferentially target the genes with high transcriptional regulation complexity. *Biochem Biophys Res Commun.* **352**, 733-738 (2007).
8. Yamakuchi, M., & Lowenstein, C. J. MiR-34, SIRT1 and p53: the feedback loop. *Cell Cycle.* **8**, 712-715 (2009).
9. Baer, C. *et al.* Extensive promoter DNA hypermethylation and hypomethylation is associated with aberrant microRNA expression in chronic lymphocytic leukemia. *Cancer Res* **72**, 3775-3785 (2012).
10. Hazan-Halevy, I. *et al.* STAT3 is constitutively phosphorylated on serine 727 residues, binds DNA, and activates transcription in CLL cells. *Blood* **115**, 2852-2863 (2010).
11. Melo, S. A. *et al.* A TARBP2 mutation in human cancer impairs microRNA processing and DICER1 function. *Nat Genet* **41**, 365-370 (2009).
12. Akira, S. Functional roles of STAT family proteins: lessons from knockout mice. *Stem Cells.* **17**, 138-146 (1999).
13. Frank, D. A., Mahajan, S., & Ritz, J. B lymphocytes from patients with chronic lymphocytic leukemia contain signal transducer and activator of transcription (STAT) 1 and STAT3 constitutively phosphorylated on serine residues. *J Clin Invest.* **100**, 3140-3148 (1997).
14. Calin, G. A. *et al.* MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias. *Proc Natl Acad Sci U S A* **101**, 11755-11760 (2004).
15. Consortium, E. P. A user's guide to the encyclopedia of DNA elements (ENCODE). *PLoS biology.* **9**, e1001046 (2011).
16. Miranda, K. C. *et al.* A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. *Cell* **126**, 1203-1217 (2006).
17. Consortium, E. P. An integrated encyclopedia of DNA elements in the human genome. *Nature.* **489**, 57-74 (2012).
18. Lau, J. C., Hanel, M. L., & Wevrick, R. Tissue-specific and imprinted epigenetic modifications of the human NDN gene. *Nucl Acids Res.* **32**, 3376-3382 (2004).
19. Corcoran, D. L. *et al.* Features of mammalian microRNA promoters emerge from polymerase II chromatin immunoprecipitation data. *PLoS One.* **4**, e5279 (2009).
20. Bhattacharyya, M., Feuerbach, L., Bhadra, T., Lengauer, T., & Bandyopadhyay, S. MicroRNA transcription start site prediction with multi-objective feature selection. *Stat Appl Genet Mol Biol.* **11**, Article 6 (2012).

