

# Deep Sequencing Identification of Novel Glucocorticoid-Responsive miRNAs in Apoptotic Primary Lymphocytes

Lindsay K. Smith<sup>1</sup>✉, Arpit Tandon<sup>2</sup>✉, Ruchir R. Shah<sup>2</sup>✉, Deepak Mav<sup>2</sup>, Alyson B. Scoltock<sup>1</sup>, John A. Cidlowski<sup>1\*</sup>

**1** Molecular Endocrinology Group, Laboratory of Signal Transduction, NIEHS, NIH, Department of Health and Human Services, Research Triangle Park, North Carolina, United States of America, **2** SRA International, Durham, North Carolina, United States of America

## Abstract

Apoptosis of lymphocytes governs the response of the immune system to environmental stress and toxic insult. Signaling through the ubiquitously expressed glucocorticoid receptor, stress-induced glucocorticoid hormones induce apoptosis via mechanisms requiring altered gene expression. Several reports have detailed the changes in gene expression mediating glucocorticoid-induced apoptosis of lymphocytes. However, few studies have examined the role of non-coding miRNAs in this essential physiological process. Previously, using hybridization-based gene expression analysis and deep sequencing of small RNAs, we described the prevalent post-transcriptional repression of annotated miRNAs during glucocorticoid-induced apoptosis of lymphocytes. Here, we describe the development of a customized bioinformatics pipeline that facilitates the deep sequencing-mediated discovery of novel glucocorticoid-responsive miRNAs in apoptotic primary lymphocytes. This analysis identifies the potential presence of over 200 novel glucocorticoid-responsive miRNAs. We have validated the expression of two novel glucocorticoid-responsive miRNAs using small RNA-specific qPCR. Furthermore, through the use of Ingenuity Pathways Analysis (IPA) we determined that the putative targets of these novel validated miRNAs are predicted to regulate cell death processes. These findings identify two and predict the presence of additional novel glucocorticoid-responsive miRNAs in the rat transcriptome, suggesting a potential role for both annotated and novel miRNAs in glucocorticoid-induced apoptosis of lymphocytes.

**Citation:** Smith LK, Tandon A, Shah RR, Mav D, Scoltock AB, et al. (2013) Deep Sequencing Identification of Novel Glucocorticoid-Responsive miRNAs in Apoptotic Primary Lymphocytes. *PLoS ONE* 8(10): e78316. doi:10.1371/journal.pone.0078316

**Editor:** Zhengqi Wang, Emory University, United States of America

**Received:** May 28, 2013; **Accepted:** September 11, 2013; **Published:** October 24, 2013

This is an open-access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the Creative Commons CC0 public domain dedication.

**Funding:** This research was supported by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences. No additional external funding was received for this study. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have the following interests. Arpit Tandon, Ruchir R. Shah and Deepak Mav are employed by SRA International. The work was done in collaboration (under a government contract) with researchers from SRA International. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials.

\* E-mail: cidlows1@niehs.nih.gov

✉ These authors contributed equally to this work.

## Introduction

Apoptosis of lymphocytes is critical for the homeostatic balance of the immune system. The escape of lymphocytes from apoptotic constraint results in dire consequences including the development of hematolymphoid malignancy and autoimmune disorders. Glucocorticoid hormones are potent inducers of lymphocyte apoptosis [1]. Endogenous glucocorticoids regulate immune development through the elimination of unwanted immature thymocytes during the T-cell selection process [2]. Furthermore, given their aggressive pro-apoptotic properties, synthetic glucocorticoids are a mainstay of hematolymphoid chemotherapeutic regimens.

Glucocorticoids are a class of essential stress-induced steroid hormones regulating cardiovascular, metabolic,

homeostatic and immunologic functions. Endogenous glucocorticoids are synthesized and secreted under the control of the hypothalamic-pituitary-adrenal axis in response to stressors, including environmental stress, nociception, and emotion [3]. The pleiotropic effects of glucocorticoids are mediated by the ubiquitously expressed glucocorticoid receptor (GR), which serves as a sensor of environmental stress, mediating the response of the immune system to environmental stress and toxic insult. Glucocorticoid-induced apoptosis of lymphocytes is a multifaceted process, requiring signaling through the GR and the altered expression of apoptotic effector genes [4-6]. Several laboratories have performed genome-wide microarray analysis to delineate the changes in gene expression that modulate glucocorticoid-induced apoptosis. Most notably, the expression of the pro-

apoptotic BH3-only Bcl-2 family member Bim is induced by glucocorticoid-treatment in murine lymphoma cell lines, human leukemic cell lines, mouse primary thymocytes, as well as human primary chronic lymphoblastic leukemia and acute lymphoblastic leukemia samples [7-9]. While not the only mechanism involved in this complex process, the upregulation of Bim is likely an important mediator of glucocorticoid-induced apoptosis, as both in-vivo and in-vitro depletion of Bim expression in lymphocytes decreases sensitivity to glucocorticoid-induced apoptosis [10-12]. However, until recently, gene expression analysis of lymphocytes undergoing glucocorticoid-induced apoptosis has largely ignored the examination of non-coding RNAs, or miRNAs.

MiRNAs are non-coding, ~21mer, single-stranded post-transcriptional regulators of gene expression [13,14]. First discovered in *C. elegans* fifteen years ago, highly conserved miRNAs have now been identified and cloned in plants, *D. melanogaster*, rodents, humans and numerous other species [15-18]. The interaction of a miRNA with mRNA (via imperfect “seed sequence” binding) hinders target mRNA translation while increasing evidence demonstrates that miRNAs can also promote the deadenylation and subsequent degradation of their mRNA targets [19].

To date, miRNAs have been assigned regulatory roles in fundamental biological processes, including differentiation, proliferation, embryonic development, and cell death [20]. Accordingly, the dysregulation of miRNA expression and function is a common observation in numerous and diverse human diseases [21]. Currently, there are over 2000 annotated mature human miRNAs, each with the capacity to regulate hundreds of target mRNAs (or approximately 30% of coding genes), establishing miRNAs as a substantial class of gene regulatory elements [22]. Importantly, miRNAs also regulate lymphocyte function and survival through both the induction and antagonism of apoptosis [23].

Previously, using both microarray and deep sequencing analysis, we described the prevalent repression of annotated miRNA expression during glucocorticoid-induced apoptosis of primary lymphocytes [1]. Further functional studies demonstrated for the first time a regulatory role for specific miRNAs and miRNA processors in the execution of glucocorticoid-induced apoptosis. Interestingly, this analysis also indicated the potential presence of numerous novel glucocorticoid-responsive miRNAs.

Here, we have developed a customized bioinformatics pipeline that facilitates the deep sequencing-mediated discovery of novel miRNAs. Using this approach, we describe the identification of hundreds of potentially novel glucocorticoid-responsive miRNAs in the transcriptome of apoptotic primary lymphocytes. Furthermore, we validated the glucocorticoid-dependent repression of two candidate novel miRNAs and Ingenuity Pathways Analysis (Ingenuity® Systems, [www.ingenuity.com](http://www.ingenuity.com)) predicted that these novel glucocorticoid-responsive miRNAs may contribute to glucocorticoid-induced apoptosis. In summary, these computational findings describe the discovery of novel glucocorticoid-responsive miRNAs and further suggest a potential role for both annotated and novel miRNAs in the glucocorticoid-induced apoptosis program.

## Results

### Discovery of novel miRNAs from deep sequencing data: Generation of test and training sets

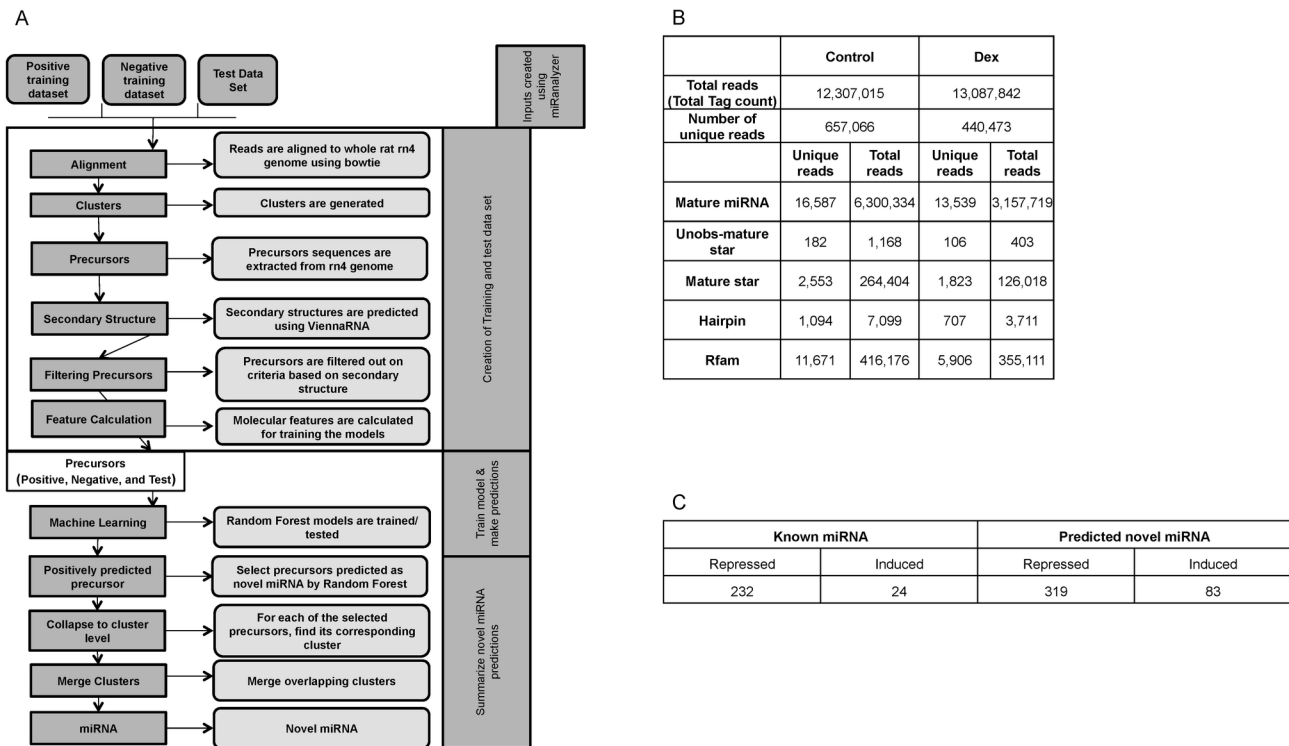
To identify glucocorticoid-responsive novel miRNAs from deep sequencing data we employed a customized bioinformatics pipeline. This pipeline is based on miRanalyzer, a previously published methodology (also available via web-server) [24]; however, we implemented several significant modifications to the original miRanalyzer approach (see methods). The basis of this computational analysis was to first align miR-analyzer-generated reads to the genome and use ‘machine learning’ to learn from the signal profile of known miRNAs and known non-miRNAs (*training*). Once the models are trained and able to accurately classify known miRNAs from non-miRNAs, we then use the models to predict novel miRNAs from signals at unannotated regions of the genome (*testing*) (Figure 1A).

This analysis employed reads previously generated by miRNA-seq analysis of annotated miRNAs during glucocorticoid-induced apoptosis [1]. Reads were generated by next generation sequencing on the Illumina platform using total RNA extracted from dexamethasone (Dex) treated and untreated (Control) primary thymocytes (see [1] for detailed description of apoptosis analysis). We obtained approximately 12-13 million reads for each sample and performed quality control analysis using FastQC (<http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc>). We then trimmed all reads at the 3' end to remove adapter sequences. Trimmed reads were subjected to a step-wise alignment protocol adopted from miRanalyzer [24] which first attempts to align reads to known miRNA sequences, and the remaining unaligned reads are then sequentially aligned to mature, mature-star\*, unobserved mature-star\*, hairpin, Refseq, and Rfam transcripts, sequentially (Figure S1). As a final alignment step, the remaining reads are aligned to the whole rat genome (Rn4). As expected, a large number of the total ~12-13 million reads obtained from deep sequencing of each sample aligned to known miRNAs when compared to the aforementioned RNA subtypes (Figure 1B). Reads that aligned to known miRNAs were used to generate the “Positive” training set while reads that aligned to other RNA subtypes were used to generate the “Negative” training set. Reads that did not align to any annotated RNA species but did align to the genome were used as “Test” data (Figure 1A).

To generate sequences belonging to the “Training” and “Test” datasets, reads with overlapping genomic coordinates were grouped together to form ‘clusters’ (totaling a length of 20-27 nucleotides) and several ‘precursor’ sequences were generated from each cluster. Precursors encompassed a genomic window centered at the cluster and extending on both the 5' and 3' ends of the cluster. We obtained 284 Control and 236 Dex clusters in the true “Positive” and 5,499 Control and 3,179 Dex clusters in the true “Negative” training data (Table S1A). Generated precursor sequences were then subjected to secondary structure selection criteria (Figure 1A).

The secondary structure of each precursor sequence was generated using Vienna RNA [25] and the precursor sequence

Figure 1



**Figure 1. Development of a customized bioinformatics pipeline for the discovery of novel miRNAs from deep sequencing data.** (A) This bioinformatics analysis workflow describes the novel miRNA discovery process adapted from miRanalyzer. The analysis pipeline uses next generation sequencing (miRNA-seq) data from untreated (control) or dexamethasone-treated rat primary thymocytes as input. This pipeline divides reads into three files: reads that align to an annotated mature miRNA (“Positive” training set), reads that align to other RNA subtypes (“Negative” training set), or reads that align at unannotated regions (“Test” set). Reads from each of these files are then aligned and alignment results are methodically processed to generate clusters, precursors and predicted secondary structures. Random forest machine learning is then employed to train the models for the prediction of novel miRNAs in the “Test” dataset. The output provides the genomic coordinates of predicted putative novel miRNAs. (B) Table describes total number of reads generated by miRNA-seq of control and dexamethasone treated primary thymocytes analyzed using the novel bioinformatics workflow described above. As expected, the majority of these reads align to known miRNAs when compared to other RNA subtypes. (C) Table summarizes the total number of known and predicted novel miRNAs identified by the bioinformatics workflow as induced or repressed in control and dexamethasone treated rat primary thymocytes. Both known and predicted novel miRNAs exhibit a trend of repressed expression during glucocorticoid-induced apoptosis.

doi: 10.1371/journal.pone.0078316.g001

was discarded from further consideration if the secondary structure did not meet stringent criteria. Pre-miRNAs are characterized by a canonical stem loop structure, hence the selection criteria was designed to discard all precursors whose secondary structure did not exhibit the desired number of base pairing and a stable hairpin structure. The filtered precursor sequences that met these criteria were used to generate molecular features that describe the unique sequence and/or secondary structure attributes of the precursor candidate in question. We chose a set of ten molecular features that best characterize attributes distinguishing a miRNA from other RNA subtypes (Figure 1A). These include features that characterize the degree of conservation of the miRNA sequence, the signal intensity at each putative miRNA location, and characteristics

of the predicted secondary structure including the minimum free energy (Figure S2).

### Training of random forest models

We generated molecular features for all filtered precursor sequences within the “Positive” and “Negative” training sets and constructed random forest models for both the Control and Dex datasets (Figure 1A). While our modeling technique used information from all molecular features for classification purposes, our analysis indicated that certain molecular features had more influence on the classification. The sequence conservation score was the most informative feature whereas the number of bases in the overhang of the secondary

structure was among the least informative (Figure S2). The training classification error for random forest models ranged from 93.3% to 99.8% denoting a high degree of accuracy (Table S2A and S2B).

### Prediction of novel miRNAs using trained models

The “Test” dataset was processed in a manner identical to the “Training” dataset in terms of preparation of clusters and precursor sequences; however, we eliminated clusters from further analysis if miRNA expression signal was below 11 raw read count to focus on only those miRNAs displaying moderate to high expression levels. We obtained 15,332 Control and 9,876 Dex clusters resulting in 52,354 Control and 33,646 Dex precursors in the “Test” dataset (Table S1B). These precursors were subjected to classification using our modeling technique.

The precursor sequences that were predicted as novel miRNAs were further filtered based on two criteria: (i) the minimum free energy of their predicted secondary structures, and (ii) signal intensity. This yielded 515 and 346 novel miRNAs predicted for Control and Dex samples, respectively, with 220 common between the two sample types. Previously, we reported that the majority of known miRNAs are repressed during glucocorticoid-induced apoptosis of lymphocytes [1]. Interestingly, this trend extends to our analysis of miRNA-seq-derived novel miRNAs. Here, approximately 80% of predicted novel miRNAs were repressed in response to dexamethasone treatment (Figure 1C).

### Validation of novel glucocorticoid-responsive miRNAs

To verify the glucocorticoid-induced repression of miRNAs, a combination of both annotated and novel miRNA candidates were selected for qPCR validation. Two novel miRNA candidates, candidate 44 and candidate 166, were chosen for validation on the basis of their predicted secondary structure. Both candidates demonstrate a canonical stem-loop structure and a putative mature miRNA sequence (Figure 2A). Furthermore, the expression of each novel miRNA candidate (as visualized in the UCSC Genome Browser [26]) is repressed in response to dexamethasone treatment (Figure 2B). This observation parallels the trend of prevalent repression of annotated miRNAs during glucocorticoid-induced apoptosis of lymphocytes, suggesting that these novel candidates are biologically similar to annotated miRNAs. Interestingly, candidate 166 also exhibits detectable signal at the proximal mature miRNA *no-miR-6324*, a recently annotated mature miRNA arising from the same precursor as candidate 166 [27]. While the basal expression of *no-miR-3624* is lower than candidate 166, it is also repressed in response to dexamethasone treatment (Figure 2B).

A total of five candidates, three annotated miRNAs (*miR-1949*, *miR-3559-5p*, and *miR-362\**) and the two predicted novel miRNAs (candidates 44 and 166) were subjected to small-RNA qPCR analysis. Each validation candidate exhibited sufficient basal signal for qPCR analysis and a degree of glucocorticoid-responsiveness as determined by the percent control value generated from computationally derived expression signals (Figure 2C and 2D). Custom Taqman Small RNA assays were designed to the mature 5'-3' sequence of

each candidate miRNA and used for the targeted quantitation of novel glucocorticoid-responsive miRNAs. These assays employ a sequence-specific stem-loop 3' reverse transcription primer, thereby assuring the definitive analysis of small RNAs [28]. This analysis confirmed the significant repression of both the annotated positive controls as well as the novel candidate miRNAs during glucocorticoid-induced apoptosis of primary lymphocytes (Figure 2E). Interestingly, the percent of control values generated by qPCR analysis closely mirror those derived from the miRNA-seq data (Figure 2D). These findings confirm the presence of two and predict the existence of numerous additional novel glucocorticoid-responsive miRNAs in the rat transcriptome (Figure 1C). To explore the potential functional roles of these novel glucocorticoid-responsive novel miRNAs, we performed further computational analysis to identify the predicted gene targets for each of the two qPCR-validated novel miRNAs.

### Pathways analysis predicts novel miRNA targets may contribute to glucocorticoid-induced apoptosis

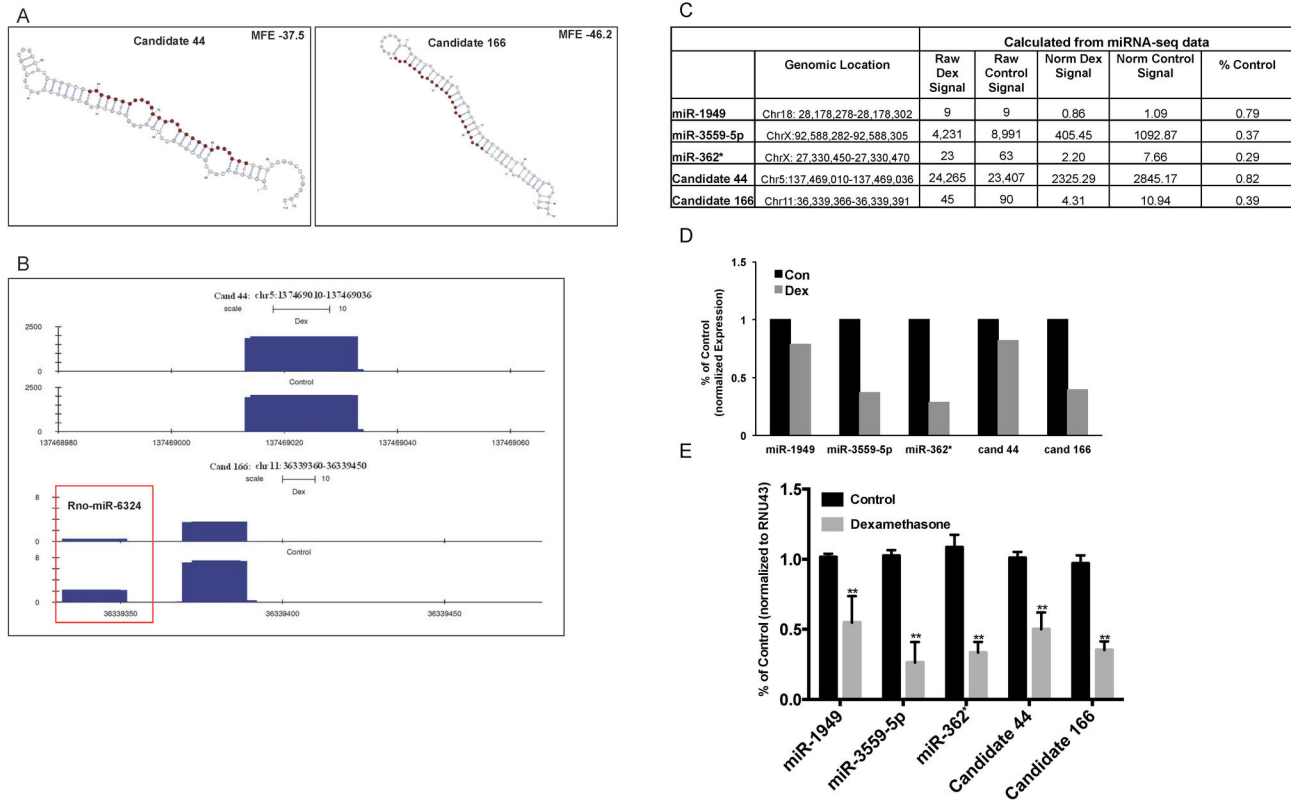
Using the mature sequence of novel miRNA candidates 44 and 166, gene target predictions were made against the 3' untranslated regions of RefSeq transcripts via the miRanda miRNA target prediction algorithm [29]. Numerous gene targets were predicted for both candidate novel miRNAs (Figure 3A). To assess the potential role of these predicted targets in the glucocorticoid-induced apoptosis program, whole genome gene expression microarray was performed on both untreated and dexamethasone treated primary thymocytes (3 biological replicates each). Ingenuity Pathways Analysis (IPA) of genes deemed differentially expressed ( $p$ -value < 0.01 and absolute fold change > 1.2) suggests that they govern molecular and cellular functions involving cell proliferation, cell division, and cell death (Figure 3B). Interestingly, IPA of the predicted novel miRNA targets suggests that these miRNAs may contribute to many of the same molecular and cellular functions identified by the whole genome microarray analysis. Specifically, cell death and cell survival is a top IPA-generated molecular and cellular function for the miRanda predicted targets of both candidates 44 and 166 (Figure 3B).

Further Venn diagram analysis identified specific mRNA targets of candidates 44 and 166 differentially expressed during glucocorticoid-induced apoptosis (Figure 3A). IPA of this combined gene list identified cell death and survival as a top predicted molecular and cellular function of these differentially expressed potential targets, as well as other functions critical to the induction and execution of glucocorticoid-induced apoptosis, including changes in cell morphology, cell cycle and cell signaling (Figure 3C). These computational findings suggest that these novel glucocorticoid-responsive miRNAs may contribute to glucocorticoid-induced apoptosis.

### Discussion

Previously, using both microarray and deep sequencing analysis, we described the prevalent repression of annotated miRNAs during glucocorticoid-induced apoptosis of primary rat thymocytes [1]. Additional studies have demonstrated the

Figure 2



**Figure 2. Validation of novel glucocorticoid responsive miRNAs.** (A) Secondary structure of two novel miRNA generated by ViennaRNA. The predicted ‘mature’ sequence is highlighted in red; the remaining hairpin contains the putative stem loop and mature-star\* sequence, the minimum free energy (MFE) of each structure is indicated. The VARNA visualization applet was used to draw the RNA secondary structure [66].

(B) The expression of the candidate novel miRNAs, candidates 44 and 166 visualized in UCSC genome browser (Dex treated is top bar, Control is bottom bar). Both of the predicted novel miRNAs are repressed during glucocorticoid-induced apoptosis of primary lymphocytes. Visualization of novel miRNA candidate 166 also detects a glucocorticoid-responsive signal at the proximal newly annotated mature miRNA rno-miR-6324, which is antisense to candidate 166 (indicated in the red box).

(C) Percent control values of the five miRNAs (3 known and 2 predicted novel candidates) selected for qPCR validation. Percent control was calculated as (Dex/Control) using computationally derived signal values for control and dexamethasone-treated rat primary thymocytes. Signal values were generated using stringent sequence alignment criteria of miRNA-seq data.

(D) Graphic representation of percent control values for control and dexamethasone-treated samples generated using computationally derived expression signals from the miRNA-seq data. Raw read counts at each miRNA were normalized to the total number of aligned reads in the respective sample to generate normalized signal.

(E) Rat primary thymocytes were untreated (control) or treated with 100nM dexamethasone for 6 hours (apoptosis was monitored as previously described [1]). The expression of annotated positive controls and individual mature candidates was evaluated via quantitative PCR using custom TaqMan Small RNA Assays. The expression of RNU43 small nuclear RNA served as an endogenous control. Results are reported as mean percent control values +/- SEM values for 3 biological replicates (\*\*p<.01).

doi: 10.1371/journal.pone.0078316.g002

glucocorticoid-mediated regulation of specific miRNAs in lymphoid cells, and further delineated a functional role for these miRNAs in the execution of glucocorticoid-induced apoptosis [30-33]. For example, studies by both Harada et al. and Molitoris et al. report the glucocorticoid-mediated repression of the miR-17 family, resulting in increased Bim expression, and, consequently, increased sensitivity to glucocorticoid-induced

apoptosis [31,32]. Alternatively, several studies have reported that specific miRNAs regulate glucocorticoid sensitivity and contribute to glucocorticoid-resistance in lymphoid malignancies [34-37].

In our present study, we propose the existence of novel, unannotated, glucocorticoid-responsive miRNAs with expression profiles similar to those we previously described for

Figure 3

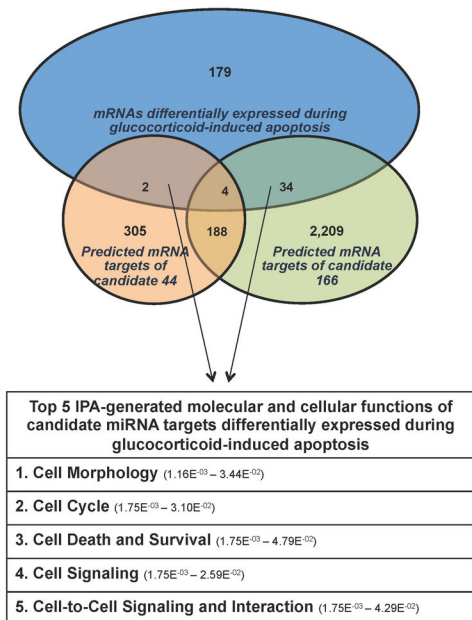
A

MicroRNA	Number of predicted gene targets	Targets differentially expressed
Candidate 44	499	6
Candidate 166	2,435	38

B

	Significantly Regulated Genes (microarray)	Candidate 44 targets (499)	Candidate 166 targets (2,435)
Cell Death and Survival	1 (3.73E <sup>-06</sup> - 2.27E <sup>-02</sup> )	5 (1.30E <sup>-09</sup> - 4.36E <sup>-03</sup> )	1 (1.39E <sup>-09</sup> - 1.20E <sup>-02</sup> )
Cell-to-Cell Signaling and Interaction	2 (3.00E <sup>-04</sup> - 2.27E <sup>-02</sup> )	1 (5.43E <sup>-09</sup> - 4.28E <sup>-03</sup> )	
Cell Cycle	3 (3.87E <sup>-04</sup> - 2.27E <sup>-02</sup> )		
Cellular Development	4 (3.87E <sup>-04</sup> - 2.27E <sup>-02</sup> )	3 (7.02E <sup>-06</sup> - 4.37E <sup>-03</sup> )	
Cellular Growth and Proliferation	5 (3.87E <sup>-04</sup> - 2.27E <sup>-02</sup> )	4 (1.03E <sup>-09</sup> - 4.37E <sup>-02</sup> )	
Cellular Function and Maintenance		2 (2.70E <sup>-06</sup> - 4.23E <sup>-03</sup> )	5 (5.20E <sup>-09</sup> - 1.20E <sup>-02</sup> )
Cellular Morphology			2 (1.55E <sup>-09</sup> - 1.20E <sup>-02</sup> )
Gene Expression			3 (3.52E <sup>-09</sup> - 1.15E <sup>-02</sup> )
Cellular Assembly and Organization			4 (5.20E <sup>-09</sup> - 1.20E <sup>-02</sup> )

C



**Figure 3. Pathways analysis predicts novel miRNA targets may contribute to glucocorticoid-induced apoptosis.** (A) miRNA target predictions for novel miRNA candidates 44 and 166 were performed using the miRanda miRNA target prediction algorithm. The number of target mRNAs differentially expressed during glucocorticoid-induced apoptosis ( $p < 0.01$ ; fold change  $> 1.2$ ) is indicated for each candidate.

(B) IPA-generated ranking of the top five molecular and cellular functions of genes differentially expressed during glucocorticoid-induced apoptosis ( $p < 0.01$ ; fold change  $> 1.2$ ), as well as the predicted targets of both candidates 44 and 166 ( $p$ -values for top functions are indicated beneath each ranking). Genes differentially expressed during glucocorticoid-induced apoptosis were identified by whole genome microarray analysis of untreated and 100nM dexamethasone-treated thymocytes (6 hours, 3 biological replicates).

(C) Venn diagram analysis identified specific novel candidate predicted targets differentially expressed during glucocorticoid-induced apoptosis ( $p < 0.01$ ) and the application IPA to this combined gene list (40 genes) generated a top 5 ranking of molecular and cellular functions regulated by these predicted targets ( $p$ -values for top functions are indicated beneath each ranking).

doi: 10.1371/journal.pone.0078316.g003

annotated miRNAs (dexamethasone-induced repression). Given that deep sequencing technology provides a powerful, unbiased platform to measure the expression of miRNAs we sought to further explore and catalogue the presence of novel glucocorticoid-responsive miRNAs in the rat transcriptome. To this end, we developed a bioinformatics pipeline combining elements of miRanalyzer [24], a peer-reviewed publically available miRNA discovery approach, and a customized machine learning technique to facilitate the identification of novel miRNAs from deep sequencing data.

The discovery of novel miRNAs from deep sequencing data is a rapidly expanding area of bioinformatics research. To date, numerous studies have reported the deep sequencing-mediated discovery of novel miRNAs in diverse systems including viruses [38,39], plants [40-43], insects [44], lower vertebrates [45,46], mammals [47,48], cell culture [49,50], and human patient samples [51-55]. Interestingly, several of these studies report the altered expression profile of these newly

identified miRNAs during pathophysiological conditions including aging, Sjogren's Syndrome, psoriasis, b-cell malignancy, and lung cancer [48,51-54]. Our study extends these findings to non-transformed, mammalian primary lymphocytes and, to our knowledge, is the first to report the hormonal-regulation of novel miRNA expression. Importantly, the recent, independent discovery of rno-miR-6324 (a mature miRNA in the anti-sense orientation to candidate 166) strengthens the evidence that candidate 166 is a novel, glucocorticoid-responsive miRNA and that our approach to the identification of novel miRNAs from deep-sequencing data is both accurate and reproducible [27].

We next employed IPA to characterize the potential cellular and molecular functions of the newly validated glucocorticoid-responsive miRNAs. This analysis indicated that the putative targets of these novel miRNAs are predicted to influence cell death. Pathways analysis of specific novel miRNA candidate targets differentially regulated during glucocorticoid-induced

apoptosis identified cell death and survival as a top-regulated predicted cellular and molecular function as well as other cellular processes essential for the glucocorticoid-induced cell death program, including changes in cellular morphology, cell cycle, and cellular signaling [56]. Presently, further functional analyses of these novel miRNAs in this model system are not possible, since rat primary thymocytes are not amenable to genetic manipulation in-vitro. However, these preliminary IPA-derived functional predictions provide a promising basis for the future validation and functional analysis of both novel miRNAs in an alternative, adaptable model system.

In summary, these studies employ a customized bioinformatic pipeline that enables the discovery of novel miRNAs from deep sequencing data and further describes the repression of two novel miRNAs (candidates 44 and 166) during glucocorticoid-induced apoptosis of primary thymocytes. Computational analysis predicts that miRNA candidates 44 and 166 may contribute to the glucocorticoid-induced apoptosis program through the regulation of target mRNAs involved in cell death and survival functions. These findings are the first to identify the presence of novel, glucocorticoid-responsive miRNAs in the rat transcriptome.

## Materials and Methods

### Ethics Statement

All animal experiments were approved by the National Institute of Environmental Health Sciences Institutional Animal Care and Use Committee and complied with USDA Column C classification (minimal, transient, or no pain or distress). Experimental animals were routinely monitored by NIEHS veterinary staff and investigators for pain or distress.

### Rat primary thymocyte isolation

Rat primary thymocytes were isolated from adrenalectomized (60-75g) male Sprague-Dawley rats (Charles River Laboratories, Wilmington, MA) approximately 1-2 weeks after surgery. Following decapitation, the thymi of three animals were removed and pooled in RPMI 1640 medium containing 10% heat-inactivated fetal bovine serum, 4 mM glutamine, 75 units/ml streptomycin, and 100 units/ml penicillin. Thymi were gently sheared with surgical scissors at room temperature. Sheared cells were filtered through 200-micron nylon mesh twice and centrifuged at 3K for 5 minutes at room temperature. The cell pellet was then resuspended in fresh media and filtered into a sterile conical tube. Cells were cultured at a final concentration of  $2 \times 10^6$  cells/mL and incubated at 37°C, 5% CO<sub>2</sub> atmosphere.

### miRNA deep sequencing

Rat primary thymocytes were isolated and cultured in the presence or absence of 100nM dexamethasone for 6 hours. Following treatment, total RNA was isolated using the Ambion mirVana miRNA isolation kit (Austin, TX) from untreated control and dexamethasone-treated samples and subjected to miRNA Deep Sequencing. Small RNA cDNA libraries were prepared according to manufacturer's protocol (Small RNA Sample Prep

Kit Oligo Only, protocol 71003, Illumina, Inc., San Diego, CA). Small RNA cDNA libraries were then sequenced according to manufacturer's instructions on the Illumina Genome Analyzer II (Illumina, Inc., San Diego, CA). The data discussed in this publication have been deposited in NCBI's Sequence Read Archive [57] and are accessible through SRA accession number SRP019941.

### Bioinformatic analysis of miRNA deep sequencing data

Deep sequencing data for one lane each of Dex and Control samples were received in the fasta format. Read lengths of Dex samples was 35 nucleotides whereas for control it was 25 nucleotides. However approximate length of a mature miRNA is around 18-22 nucleotides therefore it is likely that the 3' end of the read sequence may contain adapter sequences. To remove possible adapter sequences we trimmed the reads at 3' end such that resulting reads were 20 nucleotides in length. Next, the sequence reads were collapsed into a fasta formatted file where only unique sequences remain and duplicated sequences were counted and recorded in the header information for each sequence. Out of 13,087,842 and 12,307,015 reads in Dex and Control respectively, the data was compressed to 440,473 and 657,066 unique reads respectively. The resulting files were used for further analysis, which included discovery of novel miRNAs and calculation of differential expression in Dex vs. Control for novel and existing miRNAs.

### Computational prediction of novel miRNAs

To discover novel miRNAs from deep sequencing data, we designed a bioinformatics pipeline based on the miRanalyzer methodology. miRanalyzer is a web server that uses input short sequence reads of lengths up to 25nts and outputs predicted novel miRNAs [24]. It is also available in a stand alone version [58]. Moreover, our experimentation with the software determined that the implementation of the random forest prediction approach within miRanalyzer was not robust enough to yield reproducible results. To overcome these limitations we implemented a number of new ideas within the novel miRNA discovery paradigm. To this end, we designed a data analysis workflow that uses the general framework and certain components from miRanalyzer and combines it with our novel machine learning approach.

First, we implemented a sequence alignment strategy as described in miRanalyzer. The fasta files from Dex and Control samples were used as input. Alignments were performed in a sequential manner. First, reads were aligned to known miRNAs, followed by alignment to mature-star\*, mature-star\* unobserved and hairpin precursors. Next, the remaining reads were aligned to known mRNAs and RNA families as defined by RFAM. The sequences that map to any of the above RNA subtypes are then removed. Remaining sequences are aligned to the Rn4 genome (Figure S1).

Reads aligning to mature miRNAs were used to build the true "Positive" training dataset and reads aligned to other RNA types such as RFAM was used to build "Negative" training dataset. Reads that did not map to any known RNAs but map to unannotated locations in the genome were used to build the

“Test” dataset. The alignments for training/test dataset were generated using bowtie (0.12.7) [59] with `-best` and `-strata` options. We allowed up to 2 mismatches in the seed length of 17 and up to 6 alignments were allowed per read. Only the longest alignments that maintained the number of observed mismatches within the seed were kept for further analysis.

Following the miRanalyzer approach, all overlapping aligned reads were grouped together and ‘clusters’ were formed. ‘Precursor’ sequences were then generated from each cluster [58]. We predicted the secondary structure of each precursor sequence using the ViennaRNA (version 2.0.6) [60] tool and removed precursors if any of the following were true:

1. It doesn’t have single stem hairpin structure
2. If it has less than 19 bindings to the candidate precursor sequence
3. If it has less than 11 bindings to the region occupied by the read cluster
4. If candidate precursor genomic location doesn’t overlap with a known miRNA (only in case of true positive data set)

For the remaining precursor sequences, we calculated the molecular attributes that best describe the sequence and secondary structure characteristics of the precursor sequences. These characteristics are then used as input to the machine learning methods to train the models. The molecular features used include:

1. Total number of bindings within the read cluster
2. Total number of bindings in whole candidate precursor secondary structure
3. The length of the read cluster
4. The expression of mature-star\* sequence
5. Total tag counts in the read cluster
6. The minimum free energy (MFE)
7. Normalized Energy (MFE/candidate precursor length)
8. The difference in the number of nucleotides that don’t bind between the arms
9. The expression of overlapping conserved region
10. The number of unbinding nucleotides in overhang region

Using these features calculated for each precursor sequence in the positive and negative training dataset, we built two random forest models, one each for the Control and Dex data using “randomForest” R-Package [61]. The random forest model consisted of 1000 binary decision trees, each constructed from 66% of randomly selected training precursors and 3 randomly selected training features. For each training sample, aggregated classification votes were computed from all the trees in which the sample under consideration was excluded. Next, the out of bag training error/accuracy rates were computed, using above classification vote counts. The importance of each of the training feature is assessed using change in out of bag training accuracy, after permuting the values of feature of interest (Figure S2 shows the ranking of features). Ranking is calculated by mean decrease in accuracy associated with each feature.

Our training models displayed significantly high classification accuracy (i.e. low class error) as described by the confusion matrix (Table S2). We employed 1000 trees for modeling, a

significantly large number compared to the miRanalyzer, to ensure that training and testing results are consistent and reproducible.

We used reads from the “Test” dataset and generated clusters and precursor sequences as described earlier. Here, we discarded clusters with raw read counts (expression value) lower than 11 prior to precursor generation step to avoid regions with low expression. The resulting precursor sequences from the test data were used for feature generation and as input for classification using the two random forest models trained from the Control and Dex data as described earlier.

Precursor sequences that were predicted to be novel miRNAs were identified, and the parent ‘cluster’ sequence for each of those precursors was used as the novel ‘mature’ miRNA sequence. These novel miRNAs were further discarded if they met either of the following criteria:

1. If the predicted novel miRNA localizes to chrUn or any of the ‘random’ chromosomes.
2. If the MFE of predicted novel miRNA is greater than -25.

In cases where the chromosomal coordinates of the novel miRNAs overlapped each other, they were merged to form one novel miRNA.

#### miRNA signal and differential expression calculation

To determine computationally derived expression values at each annotated and predicted novel miRNA, we counted the total number of reads aligned at a genomic locus normalized by the total aligned reads for a given sample (Reads Per Million). For this calculation, we only included those reads that met very stringent sequence alignment criteria (reads may have a maximum of 3 alignments and only one mismatched position within the 17nt seed length). To determine whether a given miRNA is induced or repressed in response to dexamethasone treatment, we calculated the ratio of signal in Dex divided by Control. If the ratio is above 1, we consider the miRNA induced by Dex, if the ratio is below 1, we consider the miRNA repressed by Dex.

#### Novel miRNA qPCR

Total RNAs were isolated from control and dexamethasone-treated (100nM, 6 hours) thymocytes using the Ambion mirVana miRNA isolation kit (Austin, TX). For annotated and novel miRNA validations, total RNAs were reverse transcribed using the Taqman miRNA Reverse Transcription kit (Applied Biosystems, CA, USA) and analyzed using custom-designed Taqman Small RNA Assays (Applied Biosystems, CA, USA) per manufacturer instructions. Single-tube primer/probes for each candidate were designed using the Custom TaqMan Small RNA Assay Design Tool using the predicted (or annotated) mature miRNA sequence as the design template. Prior to submission, template sequences were evaluated for specificity via the Basic Local Alignment Search Tool (BLAST) [62]. Primer template sequences for each candidate novel miRNA were:

Candidate 44: CGCGGATGATGACACCTGGGTAT  
Candidate 166: GCTCTGCTGACTGCCTATGGGCT



Each customized small RNA assay was evaluated for signal in both reverse transcriptase minus and the cDNA minus non-template controls, indicating the detection of small-RNA-specific signal. Each primer/probe was normalized to the expression of the small-nucleolar RNA RNU43.

### Whole genome microarray

Rat primary thymocytes were isolated and cultured in the presence or absence of 100nM dexamethasone for 6 hours. Following treatment, total RNA was isolated from three biological replicates using the Ambion mirVana miRNA isolation kit (Austin, TX) and subjected to whole genome microarray analysis. Gene expression analysis was conducted using Agilent Whole Rat Genome 4x44 multiplex format oligo arrays (014879) (Agilent Technologies) following the Agilent 1-color microarray-based gene expression analysis protocol. Starting with 500ng of total RNA, Cy3 labeled cRNA was produced according to manufacturer's protocol. For each sample, 1.65ug of Cy3 labeled cRNAs were fragmented and hybridized for 17 hours in a rotating hybridization oven. Slides were washed and then scanned with an Agilent Scanner. Data was obtained using the Agilent Feature Extraction software (v9.5), using the 1-color defaults for all parameters. The Agilent Feature Extraction Software performed error modeling, adjusting for additive and multiplicative noise. The resulting data were processed using the Rosetta Resolver® system (version 7.2) (Rosetta Biosoftware, Kirkland, WA). The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus [63] and are accessible through GEO Series accession number GSE45560.

### Analysis of whole genome microarray data

The feature extractor processed raw signal was  $\log_2$ -transformed, quantile normalized and summarized for each probe using median polish algorithm. Next, we identified differentially expressed genes in Dex treated compared to Control samples using signal to noise statistic defined as the ratio of average signal difference and sum of between replicate standard deviations. The adjusted and unadjusted p-values for this signal to noise statistic was computed using left/right tail of empirical distribution generated by 10,000 sample/probe permutations (similar to [64]). We used a nominal p-value threshold of 0.01 (nominal p-value  $\leq$  0.01) and absolute fold change threshold of 1.2 (absolute fold  $\geq$  1.2) to identify differentially expressed probes. We used available probe annotation to map probe IDs to corresponding RefSeq genes. We identified 219 genes with statistically significant differential expression.

### Prediction novel miRNA targets

Prediction of gene targets for a given miRNA was conducted using miRanda software [65]. We used the mature miRNA sequence as an input to the program and the software generated predicted gene targets by comparing complementarity in the seed region of the miRNA sequence to the 3' UTR sequence of all known mRNAs in the genome. The list of gene targets for each of the two candidate novel miRNA

was further analyzed for enrichment of biological pathways using IPA.

### Pathway analysis using IPA

We employed Ingenuity Pathway Analysis software to identify enriched biological pathways and molecular functions within a given gene list. We performed IPA on a list of gene targets for each of the two candidate novel miRNAs (candidate 44 and 166), and also performed IPA on the list of differentially expressed genes identified by the microarray analysis. IPA was also performed on the subset of gene targets, as identified by the Venn Diagram analysis, to be differentially expressed in the microarray analysis (Figure 3).

### Supporting Information

**Figure S1. Alignment workflow based on original miRanalyzer.** Work-flow diagram of sequence alignment as implemented in miRanalyzer. The figure was adapted from the miRanalyzer manuscript [58]. (TIFF)

**Figure S2. Accuracy of molecular features used in computational prediction of miRNAs.** Figure displays ranking of molecular features used in computational prediction of miRNAs. The x-axis reports the mean decrease in accuracy of the model for each of the molecular features in question. Conservation is the most informative feature in this analysis. (TIFF)

**Table S1. Summary of training and test data sets.** (A) Table describes number of unique clusters and resulting precursor sequences for Dex and Control samples in positive and negative training sets. (B) Table describes number of unique clusters and resulting precursor sequences for Dex and Control samples in test set. (TIFF)

**Table S2. Confusion matrix for training data for control and dexamethasone-treated thymocytes.** (A and B) Confusion matrix displaying predicted and actual number of microRNAs and non-microRNAs as identified by our computational analysis. The classification accuracies of predicting microRNAs are listed for both control and dexamethasone-treated samples. (TIFF)

### Acknowledgements

We acknowledge the generous contributions of the University of North Carolina High Throughput Sequencing, NIEHS Microarray, and NIEHS Flow Cytometry Core Facilities. We thank Dr. Carl Bortner and Dr. David Miller for their insights and critical review of the manuscript.

## Author Contributions

Conceived and designed the experiments: LKS JAC RRS.  
Performed the experiments: LKS ABS. Analyzed the data: LKS

AT RRS DM. Contributed reagents/materials/analysis tools:  
LKS AT RRS DM JAC. Wrote the manuscript: LKS RRS.

## References

- Smith LK, Shah RR, Cidlowski JA (2010) Glucocorticoids modulate microRNA expression and processing during lymphocyte apoptosis. *J Biol Chem* 285: 36698-36708. doi:10.1074/jbc.M110.162123. PubMed: 20847043.
- Ashwell JD, Lu FW, Vacchio MS (2000) Glucocorticoids in T cell development and function\*. *Annu Rev Immunol* 18: 309-345. doi: 10.1146/annurev.immunol.18.1.309. PubMed: 10837061.
- Rhen T, Cidlowski JA (2005) Antiinflammatory action of glucocorticoids--new mechanisms for old drugs. *N Engl J Med* 353: 1711-1723. doi:10.1056/NEJMra050541. PubMed: 16236742.
- Cifone MG, Migliorati G, Parroni R, Marchetti C, Millimaggi D et al. (1999) Dexamethasone-induced thymocyte apoptosis: apoptotic signal involves the sequential activation of phosphoinositide-specific phospholipase C, acidic sphingomyelinase, and caspases. *Blood* 93: 2282-2296. PubMed: 10090938.
- Mann CL, Hughes FM Jr., Cidlowski JA (2000) Delineation of the signaling pathways involved in glucocorticoid-induced and spontaneous apoptosis of rat thymocytes. *Endocrinology* 141: 528-538. doi: 10.1210/en.141.2.528. PubMed: 10650932.
- Wang D, Müller N, McPherson KG, Reichardt HM (2006) Glucocorticoids engage different signal transduction pathways to induce apoptosis in thymocytes and mature T cells. *J Immunol* 176: 1695-1702. PubMed: 16424199.
- Distelhorst CW (2002) Recent insights into the mechanism of glucocorticosteroid-induced apoptosis. *Cell Death Differ* 9: 6-19. doi: 10.1038/sj.cdd.4400969. PubMed: 11803370.
- Iglesias-Serret D, de Frias M, Santidrián AF, Coll-Mulet L, Cosiáls AM et al. (2007) Regulation of the proapoptotic BH3-only protein BIM by glucocorticoids, survival signals and proteasome in chronic lymphocytic leukemia cells. *Leukemia* 21: 281-287. doi:10.1038/sj.leu.2404483. PubMed: 17151701.
- Wang Z, Malone MH, He H, McColl KS, Distelhorst CW (2003) Microarray analysis uncovers the induction of the proapoptotic BH3-only protein Bim in multiple models of glucocorticoid-induced apoptosis. *J Biol Chem* 278: 23861-23867. doi:10.1074/jbc.M301843200. PubMed: 12676946.
- Abrams MT, Robertson NM, Yoon K, Wickstrom E (2004) Inhibition of glucocorticoid-induced apoptosis by targeting the major splice variants of BIM mRNA with small interfering RNA and short hairpin RNA. *J Biol Chem* 279: 55809-55817. doi:10.1074/jbc.M411767200. PubMed: 15509554.
- Bouillet P, Metcalf D, Huang DC, Tarlinton DM, Kay TW et al. (1999) Proapoptotic Bcl-2 relative Bim required for certain apoptotic responses, leukocyte homeostasis, and to preclude autoimmunity. *Science* 286: 1735-1738. doi:10.1126/science.286.5445.1735. PubMed: 10576740.
- Lu J, Quearry B, Harada H (2006) p38-MAP kinase activation followed by BIM induction is essential for glucocorticoid-induced apoptosis in lymphoblastic leukemia cells. *FEBS Lett* 580: 3539-3544. doi:10.1016/j.febslet.2006.05.031. PubMed: 16730715.
- Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116: 281-297. doi:10.1016/S0092-8674(04)00045-5. PubMed: 14744438.
- Esquela-Kerschner A, Slack FJ (2006) Oncomirs - microRNAs with a role in cancer. *Nat Rev Cancer* 6: 259-269. doi:10.1038/nrc1840. PubMed: 16557279.
- Lagos-Quintana M, Rauhut R, Yalcin A, Meyer J, Lendeckel W et al. (2002) Identification of tissue-specific microRNAs from mouse. *Curr Biol* 12: 735-739. doi:10.1016/S0960-9822(02)00809-6. PubMed: 12007417.
- Lee RC, Ambros V (2001) An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* 294: 862-864. doi:10.1126/science.1065329. PubMed: 11679672.
- Lee RC, Feinbaum RL, Ambros V (1993) The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 75: 843-854. doi:10.1016/0092-8674(93)90529-Y. PubMed: 8252621.
- Lim LP, Glasner ME, Yekta S, Burge CB, Bartel DP (2003) Vertebrate microRNA genes. *Science* 299: 1540. doi:10.1126/science.1080372. PubMed: 12624257.
- Fabian MR, Sonenberg N, Filipowicz W (2010) Regulation of mRNA translation and stability by microRNAs. *Annu Rev Biochem* 79: 351-379. doi:10.1146/annurev-biochem-060308-103103. PubMed: 20533884.
- Bushati N, Cohen SM (2007) microRNA functions. *Annu Rev Cell Dev Biol* 23: 175-205. doi:10.1146/annurev.cellbio.23.090506.123406. PubMed: 17506695.
- Jiang Q, Wang Y, Hao Y, Juan L, Teng M et al. (2009) miR2Disease: a manually curated database for microRNA deregulation in human disease. *Nucleic Acids Res* 37: D98-104. doi:10.1093/nar/gkn714. PubMed: 18927107.
- Filipowicz W, Bhattacharyya SN, Sonenberg N (2008) Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat Rev Genet* 9: 102-114. PubMed: 18197166.
- Subramanian S, Steer CJ (2010) MicroRNAs as gatekeepers of apoptosis. *J Cell Physiol* 223: 289-298. PubMed: 20112282.
- Hackenberg M, Sturm M, Langenberger D, Falcón-Pérez JM, Aransay AM (2009) miRanalyzer: a microRNA detection and analysis tool for next-generation sequencing experiments. *Nucleic Acids Res* 37: W68-W76. doi:10.1093/nar/gkp221. PubMed: 19433510.
- Hofacker IL (2003) Vienna RNA secondary structure server. *Nucleic Acids Res* 31: 3429-3431. doi:10.1093/nar/gkg599. PubMed: 12824340.
- Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH et al. (2002) The human genome browser at UCSC. *Genome Res* 12: 996-1006. doi:10.1101/gr.229102. Article published online before print in May 2002 PubMed: 12045153
- Clokier SJ, Lau P, Kim HH, Coon SL, Klein DC (2012) MicroRNAs in the pineal gland: miR-483 regulates melatonin synthesis by targeting arylalkylamine N-acetyltransferase. *J Biol Chem* 287: 25312-25324. doi:10.1074/jbc.M112.356733. PubMed: 22908386.
- Chen C, Ridzon DA, Broomer AJ, Zhou Z, Lee DH et al. (2005) Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res* 33: e179. doi:10.1093/nar/gni178. PubMed: 16314309.
- Betel D, Wilson M, Gabow A, Marks DS, Sander C (2008) The microRNA.org resource: targets and expression. *Nucleic Acids Res* 36: D149-D153. doi:10.1093/nar/gkn715. PubMed: 18158296.
- Davis TE, Kis-Toth K, Szanto A, Tsokos GC (2013) Glucocorticoids Suppress T Cell Function by Up-Regulating MicroRNA-98. *Arthritis Rheum* 65: 1882-1890. doi:10.1002/art.37966. PubMed: 23575983.
- Harada M, Pokrovskaja-Tamm K, Söderhäll S, Heyman M, Grandér D et al. (2012) Involvement of miR17 pathway in glucocorticoid-induced cell death in pediatric acute lymphoblastic leukemia. *Leuk Lymphoma* 53: 2041-2050. doi:10.3109/10428194.2012.678004. PubMed: 22475310.
- Molitoris JK, McColl KS, Distelhorst CW (2011) Glucocorticoid-mediated repression of the oncogenic microRNA cluster miR-17~92 contributes to the induction of Bim and initiation of apoptosis. *Mol Endocrinol* 25: 409-420. doi:10.1210/me.2010-0402. PubMed: 21239610.
- Sionov RV (2013) MicroRNAs and Glucocorticoid-Induced Apoptosis in Lymphoid Malignancies. *ISRN Hematol*, 2013: 348212. PubMed: 23431463
- Kotani A, Ha D, Hsieh J, Rao PK, Schotte D et al. (2009) miR-128b is a potent glucocorticoid sensitizer in MLL-AF4 acute lymphocytic leukemia cells and exerts cooperative effects with miR-221. *Blood* 114: 4169-4178. doi:10.1182/blood-2008-12-191619. PubMed: 19749093.
- Kotani A, Ha D, Schotte D, den Boer ML, Armstrong SA et al. (2010) A novel mutation in the miR-128b gene reduces miRNA processing and leads to glucocorticoid resistance of MLL-AF4 acute lymphocytic leukemia cells. *Cell Cycle* 9: 1037-1042. doi:10.4161/cc.9.6.11011. PubMed: 20237425.
- Tessel MA, Benham AL, Krett NL, Rosen ST, Gunaratne PH (2011) Role for microRNAs in regulating glucocorticoid response and resistance in multiple myeloma. *Horm Cancer* 2: 182-189. doi:10.1007/s12672-011-0072-8. PubMed: 21761344.
- Yang A, Ma J, Wu M, Qin W, Zhao B et al. (2012) Aberrant microRNA-182 expression is associated with glucocorticoid resistance in lymphoblastic malignancies. *Leuk Lymphoma* 53: 2465-2473. doi: 10.3109/10428194.2012.693178. PubMed: 22582938.

38. Yao Y, Smith LP, Petherbridge L, Watson M, Nair V. (2012) Novel microRNAs encoded by duck enteritis virus. *J Gen Virol* 93: 1530-1536. doi:10.1099/vir.0.040634-0. PubMed: 22492913.
39. Zhu JY, Strehle M, Frohn A, Kremmer E, Höfig KP et al. (2010) Identification and analysis of expression of novel microRNAs of murine gammaherpesvirus 68. *J Virol* 84: 10266-10275. doi:10.1128/JVI.01119-10. PubMed: 20668074.
40. Song C, Wang C, Zhang C, Korir NK, Yu H et al. (2010) Deep sequencing discovery of novel and conserved microRNAs in trifoliolate orange (*Citrus trifoliata*). *BMC Genomics* 11: 431. doi:10.1186/1471-2164-11-431. PubMed: 20626894.
41. Gao ZH, Wei JH, Yang Y, Zhang Z, Xiong HY et al. (2012) Identification of conserved and novel microRNAs in *Aquilaria sinensis* based on small RNA sequencing and transcriptome sequence data. *Gene* 505: 167-175. doi:10.1016/j.gene.2012.03.072. PubMed: 22521867.
42. Gébelin V, Argout X, Engchuan W, Pitollat B, Duan C et al. (2012) Identification of novel microRNAs in *Hevea brasiliensis* and computational prediction of their targets. *BMC Plant Biol* 12: 18. doi:10.1186/1471-2229-12-18. PubMed: 22330773.
43. Lelandais-Brière C, Naya L, Sallet E, Calenge F, Frugier F et al. (2009) Genome-wide *Medicago truncatula* small RNA analysis revealed novel microRNAs and isoforms differentially regulated in roots and nodules. *Plant Cell* 21: 2780-2796. doi:10.1105/tpc.109.068130. PubMed: 19767456.
44. Jagadeeswaran G, Zheng Y, Sumathipala N, Jiang H, Arrese EL et al. (2010) Deep sequencing of small RNA libraries reveals dynamic regulation of conserved and novel microRNAs and microRNA-stars during silkworm development. *BMC Genomics* 11: 52. doi:10.1186/1471-2164-11-52. PubMed: 20089182.
45. Ambady S, Wu Z, Dominko T (2012) Identification of novel microRNAs in *Xenopus laevis* metaphase II arrested eggs. *Genesis* 50: 286-299. doi:10.1002/dvg.22010. PubMed: 22223599.
46. Glazov EA, Cottee PA, Barris WC, Moore RJ, Dalrymple BP et al. (2008) A microRNA catalog of the developing chicken embryo identified by a deep sequencing approach. *Genome Res* 18: 957-964. doi:10.1101/gr.074740.107. PubMed: 18469162.
47. Chen C, Deng B, Qiao M, Zheng R, Chai J et al. (2012) Solexa sequencing identification of conserved and novel microRNAs in backfat of Large White and Chinese Meishan pigs. *PLOS ONE* 7: e31426. doi:10.1371/journal.pone.0031426. PubMed: 22355364.
48. Inukai S, de Lencastre A, Turner M, Slack F (2012) Novel microRNAs differentially expressed during aging in the mouse brain. *PLOS ONE* 7: e40028. doi:10.1371/journal.pone.0040028. PubMed: 22844398.
49. Dhahbi JM, Atamna H, Boffelli D, Magis W, Spindler SR et al. (2011) Deep sequencing reveals novel microRNAs and regulation of microRNA expression during cell senescence. *PLOS ONE* 6: e20509. doi:10.1371/journal.pone.0020509. PubMed: 21637828.
50. Ryu S, Joshi N, McDonnell K, Woo J, Choi H et al. (2011) Discovery of novel human breast cancer microRNAs from deep sequencing data by analysis of pri-microRNA secondary structures. *PLOS ONE* 6: e16403. doi:10.1371/journal.pone.0016403. PubMed: 21346806.
51. Tandon M, Gallo A, Jang SI, Illei GG, Alevizos I (2012) Deep sequencing of short RNAs reveals novel microRNAs in minor salivary glands of patients with Sjögren's syndrome. *Oral Dis* 18: 127-131. doi:10.1111/j.1601-0825.2011.01849.x. PubMed: 21895886.
52. Jima DD, Zhang J, Jacobs C, Richards KL, Dunphy CH et al. (2010) Deep sequencing of the small RNA transcriptome of normal and malignant human B cells identifies hundreds of novel microRNAs. *Blood* 116: e118-e127. doi:10.1182/blood-2010-05-285403. PubMed: 20733160.
53. Joyce CE, Zhou X, Xia J, Ryan C, Thrash B et al. (2011) Deep sequencing of small RNAs from human skin reveals major alterations in the psoriasis miRNAome. *Hum Mol Genet* 20: 4025-4040. doi:10.1093/hmg/ddr331. PubMed: 21807764.
54. Keller A, Backes C, Leidinger P, Kefer N, Boisguerin V et al. (2011) Next-generation sequencing identifies novel microRNAs in peripheral blood of lung cancer patients. *Mol Biosyst* 7: 3187-3199. doi:10.1039/c1mb05353a. PubMed: 22027949.
55. Creighton CJ, Benham AL, Zhu H, Khan MF, Reid JG et al. (2010) Discovery of novel microRNAs in female reproductive tract using next generation sequencing. *PLOS ONE* 5: e9637. doi:10.1371/journal.pone.0009637. PubMed: 20224791.
56. Smith LK, Cidlowski JA (2010) Glucocorticoid-induced apoptosis of healthy and malignant lymphocytes. *Prog Brain Res* 182: 1-30. doi:10.1016/S0079-6123(10)82001-1. PubMed: 20541659.
57. Leinonen R, Sugawara H, Shumway M (2011) The sequence read archive. *Nucleic Acids Res* 39: D19-D21. doi:10.1093/nar/gkq768. PubMed: 21062823.
58. Hackenberg M, Rodríguez-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. doi:10.1093/nar/gkq738. PubMed: 21515631.
59. 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. doi:10.1186/gb-2009-10-3-r25. PubMed: 19261174.
60. Lorenz R, Bernhart SH, Höner Zu Siederdisen C, Tafer H, Flamm C et al. (2011) ViennaRNA Package 2.0. *Algorithms Mol Biol* 6: 26. doi:10.1186/1748-7188-6-26. PubMed: 22115189.
61. Liaw AW (2002) Classification and Regression by randomForest. *R NEWS* 2: 18-22.
62. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215: 403-410. doi:10.1016/S0022-2836(05)80360-2. PubMed: 2231712.
63. Edgar R, Domrachev M, Lash AE (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 30: 207-210. doi:10.1093/nar/30.1.207. PubMed: 11752295.
64. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL et al. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102: 15545-15550. doi:10.1073/pnas.0506580102. PubMed: 16199517.
65. Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ (2008) miRBase: tools for microRNA genomics. *Nucleic Acids Res* 36: D154-D158. doi:10.1093/nar/gkn221. PubMed: 17991681.
66. Darty K, Denise A, Ponty Y (2009) VARNA: Interactive drawing and editing of the RNA secondary structure. *Bioinformatics* 25: 1974-1975. doi:10.1093/bioinformatics/btp250. PubMed: 19398448.