

# Computational Analysis of Human microRNA and its Disease Target Network

Vaijayanthi Raghavan<sup>1+</sup> and Manasa D.B.<sup>2\*</sup>

<sup>2</sup> Department of Biotechnology, PES Institute of Technology, Bangalore 560085, INDIA

**Abstract.** microRNAs are associated with diseases and play a vital role in the development of various cancers. Biological network analysis provides novel insights in understanding basic mechanisms controlling normal biological processes and disease pathologies. Computational prediction of disease-associated human microRNAs is one of the complementary means. In the present study we computationally determined the disease association of all human microRNA predicted targets and the leads were identified for the most prevalent disease. The complete list of human microRNAs was downloaded and the target genes of individual microRNA were retrieved based on miTG score using target prediction program. The resultant network was investigated based on the association of nodes with a relevant pathway, disease and pathological event. Structure-based virtual screening involves docking of screened compounds and the protein target and the leads were identified based on the affinity and free energy of binding values. After extensive research, breast cancer was identified as the most significant disease and the compounds such as Indolocarbozole, Camptothecin, Lucidinic Acid, Quercetin and Staurosporine were considered as best leads for ZNF439.

**Keywords:** *microRNA* (miRNA), ZNF439, Biological network, Structure-based virtual screening

## 1. Introduction

microRNAs are single-stranded RNA molecules of approximately 22 nucleotides in length and the level of translation is regulated by gene expression. microRNAs binds to the 3' *untranslated region* (3'UTR) of target mRNAs by mediating posttranscriptional regulation of protein-coding genes and their main function is to downregulate gene expression including translational repression, mRNA cleavage and deadenylation. Thus miRNAs play an important role in fine-tuning the diverse cellular functions such as development, differentiation, proliferation, apoptosis and metabolism [1].

Zinc finger proteins are encoded with zinc mediated nucleic acid binding proteins. The ZNFs are abundantly found in eukaryotic genomes and their main functions are *deoxyribonucleic acid* (DNA) recognition, protein folding and assembly, transcriptional activation, lipid binding, regulation of apoptosis and *ribonucleic acid* (RNA) packaging, the structure of ZNF439 is as shown in Fig. 1.

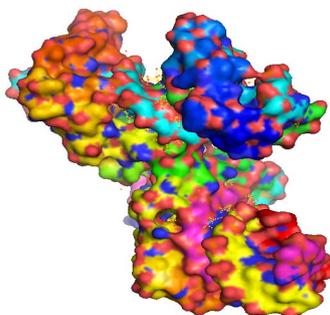


Fig. 1: Structure of ZNF439

+ Corresponding Author: Tel.: +91 09845518380; Fax: +91 80 26720886. E-mail address: vrr\_94@pes.edu

\*Presenting Author: Tel.: +91 9880605265; Fax: +91 80 26720886. E-mail address: manasadb@gmail.com

## 2. Materials and Methods

### 2.1. Network Analysis

The human miRNAs was extracted from miRBASE Release 17 [2]. The target genes of individual miRNA were explored using Diana-microT 3.0 target prediction program [3]. Highly reliable targets were selected based on the miTG score  $\geq 20$ , because the targets exhibited significantly a lower precision score with the miTG score  $< 20$ . The gene names of miRNA target genes were uploaded onto Cytoscape [4]. Reactome FIs and Nemo are the Cytoscape plugins. Reactome FIs plugin constructs human miRNA gene interaction network based on the gene set analysis and Nemo plugin analyze the biological network of miRNA and reveals the ranking based on the clustered network. The score of clustered network was consistent from rank 1 to rank 7 (i.e., Score= 49.872). Considering the gene IDs from rank1 to rank7 for finding the pathway, disease and pathological event, the biological networking of human miRNA (Fig. 2) is an undirected graph consisted of nodes and edges where each node represents one miRNA and an edge represents connection between two nodes (miRNAs).

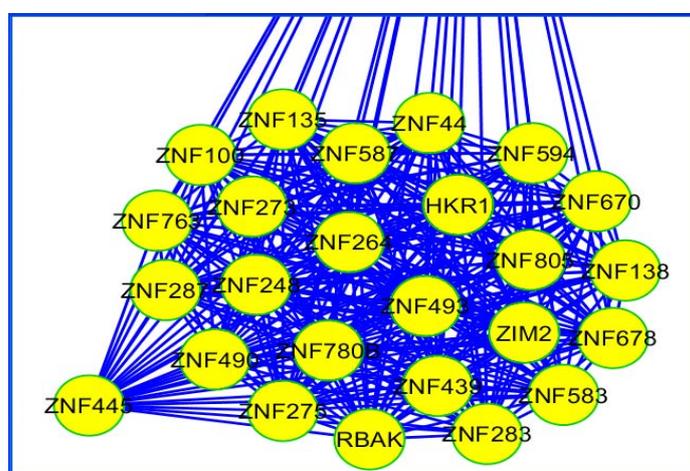


Fig. 2: Biological Networking of Human miRNA

### 2.2. Target Identification

After finding the pathway, disease and pathological event for an individual gene ID, the number of occurrences of pathway, disease and pathological event was analysed. The analysis results of pathway, disease and pathological event are as represented graphically below in the Fig. 3, 4 & 5. The results of biological networking revealed that the rank one pathway as p53 signalling pathway, the rank one disease and the rank one pathological event as breast cancer. The network analysis of breast cancer genes was performed using VisANT and the most interacting genes which is predominantly influencing for breast cancer was ZNF439 (PDBID: 2I13) [5].

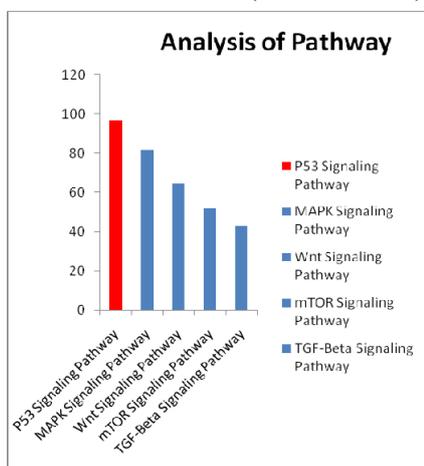


Fig. 3: Analysis of Pathway

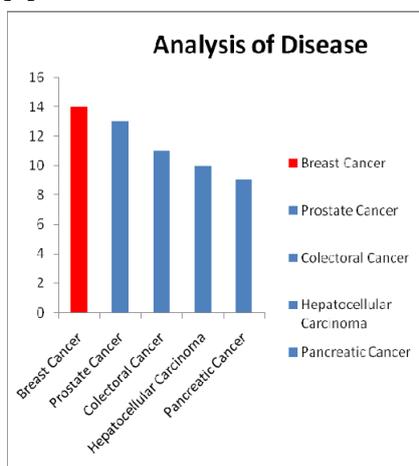


Fig. 4: Analysis of Disease

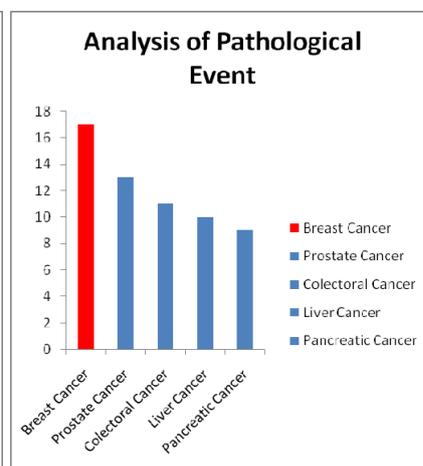


Fig. 5: Analysis of Pathological event

## 2.3. Virtual Screening

A small molecule library of anti-cancer agents for breast cancer containing 191 compounds was created. Lipinski's rule of five was chosen as compound filter for virtual screening of this library using ZNF439 as the protein target and hence 122 screened compounds were qualified. The successful compounds were subjected to energy minimization using Marvin Sketch. Structure-based virtual screening using AutoDock Vina was then performed on these minimum energy conformers [6]. The compounds were ranked based on their docking scores. Top ten compounds were redocked with Hex 6.3 and obtained free energy of binding results [7]. The Binding sites of these ten compounds were predicted using Q-SiteFinder and hydrogen bond interaction was utilized as a filter for the post processing of docking results, out of which five compounds showed Hydrogen bond interactions using PyMOL [8] as shown in the below Fig. 6, 7, 8, 9 & 10. Hence, these five compounds were proposed to be the best Lead molecules for ZNF439 with their predicted half maximal inhibitory concentration (IC<sub>50</sub>) value with EcoliTox being reported [9].

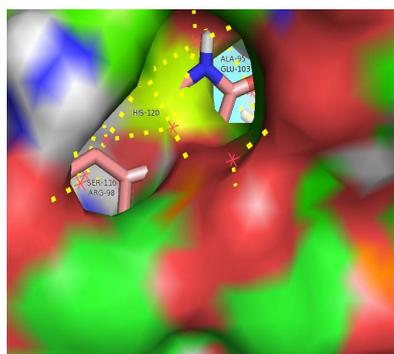


Fig. 6: Indolocarbozole and ZNF439

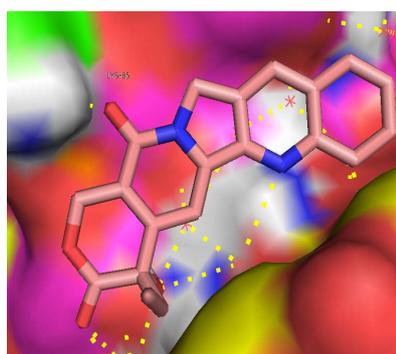


Fig. 7: Camptothecin and ZNF439

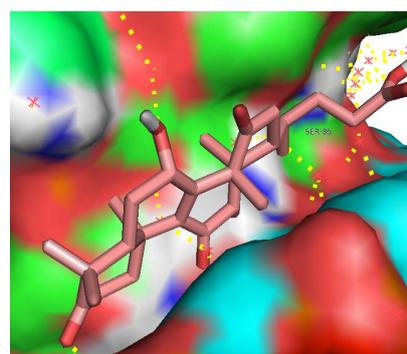


Fig. 8: Lucidenic Acid and ZNF439

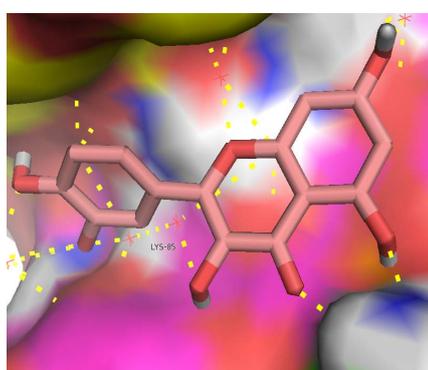


Fig. 9: Quercetin and ZNF439

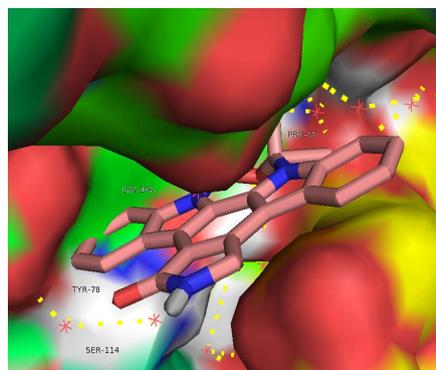


Fig. 10: Staurosporine and ZNF439

## 3. Results

Table 1: Analysis of Interaction with ZNF439, Autodock Vina<sup>a</sup>, Hex 6.3<sup>b</sup> results

Serial No.	POTENTIAL LEAD COMPOUND	DOCKING SCORE <sup>a</sup> (kcal/mol)	FREE ENERGY OF BINDING <sup>b</sup>	INTERACTING RESIDUES
1)	INDOLOCARBOZOLE	- 9.2	- 272.39	GLU -103, HIS-120, ALA-95, SER-116, ARG-98, GLN-117, THR-99
2)	CAMPTOTHECIN	- 8.8	- 255.77	LYS-85
3)	LUCIDENIC ACID	- 8.5	- 255.35	SER-86
4)	QUERCETIN	- 8.5	- 274.02	LYS-85
5)	STAUROSPORINE	- 8.5	- 268.06	PRO-77, TYR-78, LYS-104, SER-114

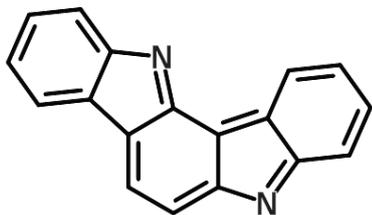


Fig. 10: Structure of Indolocarbozole

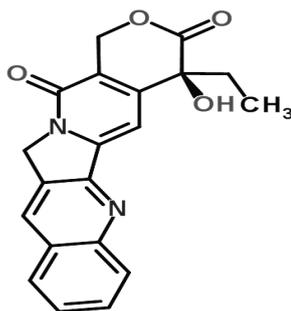


Fig. 11: Structure of Camptothecin

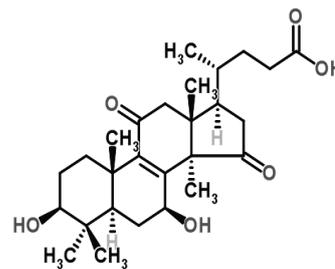


Fig. 12: Structure of Lucidenic Acid

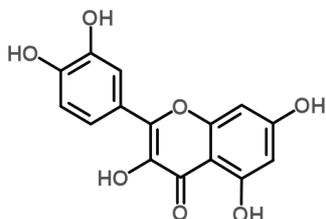


Fig. 13: Structure of Quercetin

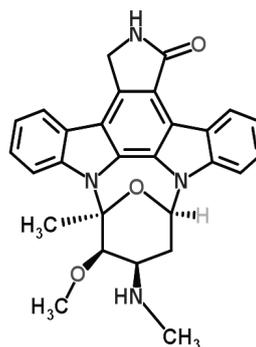


Fig. 14: Structure of Staurosporine

## 4. Discussion

In general, hundreds of target mRNAs are concurrently downregulated by a single miRNA by pairing to the 3'UTR of mRNA from either perfect or imperfect sequence complementarity. miRNAs provide a wide range of cellular functions such as differentiation, metabolism, development, apoptosis and proliferation [10]. In the present study, analysis of miRNA target network showed that the most relevant pathological event is 'breast cancer'. Furthermore, the highly relevant diseases include 'breast cancer', 'prostate cancer', and 'colorectal cancer'. The human microRNAome plays a more specialized role in regulation of tumorigenesis. Therefore, the miRNA might serve as the most effective approach to suppressing the tumorigenic potential of a wide range of cancers simultaneously miRNA-based therapy directed to targeting many cancer-associated pathways [10].

Familial breast cancers are associated with defects in RING-finger proteins. Recently, the genome-wide sequencing studies in breast cancer have identified *zinc finger protein 668* (ZNF668) as a breast tumor suppressor gene which mainly functions in regulating p53 stability [12].

The Lead compounds such as Indolocarbozole, Camptothecin, Lucidenic Acid, Quercetin and Staurosporine are the natural products and provides significant anti-cancer activity. Staurosporine and Rebeccamycin are the natural products belong to the family of indolocarbazole alkaloids with antitumor properties. Currently an intense effort exists for the creation of indolocarbazole derivatives for the treatment of cancer and neurodegenerative disorders [13].

## 5. Conclusion

The protocol used in our work can be applied to find targets for other diseases and propose leads for these targets. In the present study, the reliable predicted targets of all human miRNAs constructed biologically meaningful disease gene network of 1509 nodes and 7476 edges. In the human miRNA target network, the most relevant pathway predicted is p53 Signaling Pathway, the disease is Breast Cancer and the pathological event is Breast Cancer. The most suitable target for breast cancer has been identified as ZNF439. A small molecule library of anti-cancer agents was screened to identify best Leads. The compounds such as

Indolocarbozole, Camptothecin, Lucidenic Acid, Quercetin and Staurosporine were considered as best leads for ZNF439 which shows favourable interactions computationally. As these are predicted to be promising lead compounds, they can be prioritized for synthesis and testing, through In-vitro assays and optimization of breast cancer and other related cancers.

## 6. Acknowledgements

It is with immense gratefulness that we acknowledge Institute of Computational Biology, Bangalore for the supervision and provision of the most modern laboratory facilities. Special thanks to our faculty and management members - Department of Biotechnology, PES Institute of Technology, Bangalore for their continuous support and cooperation.

## 7. References

- [1] Bartel DP: *MicroRNAs: target recognition and regulatory functions*. Cell, 2009.
- [2] Ana Kozomara, Sam Griffiths-Jones: *miRBase: integrating microRNA annotation and deep-sequencing data*. Nucleic Acids Research, 2011.
- [3] M. Maragkakis, P. Alexiou, G. L. Papadopoulos, M. Reczko, T. Dalamagas, G. Giannopoulos, G. Goumas, E. Koukis, K. Kourtis, V. A. Simossis, P. Sethupathy, T. Vergoulis, N. Koziris, T. Sellis, P. Tsanakas, A. G. Hatzigeorgiou: *Accurate microRNA target prediction correlates with protein repression levels*. BMC Bioinformatics, 2009.
- [4] Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T: *Cytoscape 2.8: new features for data integration and network visualization*. Bioinformatics, 2011.
- [5] Hu Z, Mellor J, Wu J, DeLisi C.: *VisANT: an online visualization and analysis tool for biological interaction data*. BMC Bioinformatics, 2004.
- [6] Garrett M. Morris, Ruth Huey, William Lindstrom, Michel F. Sanner, Richard K. Belew, David S. Goodsell, Arthur J. Olson: *Autodock4 and AutoDockTools4: automated docking with selective receptor flexibility*. J. Computational Chemistry, 2009.
- [7] David W. Ritchie, Vishwesh Venkatraman: *Ultra-fast FFT protein docking on graphics processors*. Bioinformatics, 2010.
- [8] Madhavi Sastry, Jeffrey F. Lowrie, Steven L. Dixon, Woody Sherman: *Large-Scale Systematic Analysis of 2D Fingerprint Methods and Parameters to Improve Virtual Screening Enrichments*. J. Chem. Inf. Model., 2010.
- [9] Planson AG, Carbonell P, Paillard E, Pollet N, Faulon JL.: *Compound toxicity screening and structure-activity relationship modeling in Escherichia coli*. Biotechnol Bioeng, 2012.
- [10] Ming Lu, Qipeng Zhang, Min Deng, Jing Miao, Yanhong Guo, Wei Gao, Qinghua Cui: *An Analysis of Human MicroRNA and Disease Associations*. PLoS ONE, 2008.
- [11] Jun-ichi Satoh, Hiroko Tabunoki: *Comprehensive analysis of human miRNA target networks* BioData Mining, 2011.
- [12] Ruozhen Hu, Guang Peng, Hui Dai, et al: *ZNF668 Functions as a Tumor Suppressor by Regulating p53 Stability and Function in Breast Cancer*. Cancer Res, 2011.
- [13] Cesar Sanchez, Lili Zhu, Alfredo F. Brana, Aaroa P. Salas, Jurgen Rohr, Carmen Mendez, Jose A. Salas: *Combinatorial biosynthesis of antitumor indolocarbozole compounds*. PNAS, 2005.