International Journal of Advanced Biotechnology and Research
ISSN 0976-2612, Online ISSN 2278–599X,
Vol 4, Issue 2, 2013, pp 253-259
http://www.bipublication.com

IN SILICO PREDICTION OF MIRNA IN *Curcuma longa* AND THEIR ROLE IN HUMAN METABOLOMICS

Rashmi Rameshwari[1*], Divya Singhal[2], Rachit Narang[2], Apurvi Maheshwari[2] and T. V. Prasad[3]

[1, 2]Dept. of Biotechnology, Manav Rachna International University, Faridabad, India
[3]Computing Sciences, Visvodaya Technical Academy, Kavali, AP, India
Email: rashmi.rameshwari@gmail.com, divyacitm@yahoo.com, rachit_narang@yahoo.co.in, apurvim@yahoo.com, tvprasad2002@yahoo.com
[*]**Corresponding author:** Email: rashmi.rameshwari@gmail.com, Tel: +919810039850.

[Received-11/01/2013, Accepted-29/03/2013]

ABSTRACT:

MicroRNAs (miRNAs) are a small, non- coding, single-stranded RNA molecules directly involved in regulating gene expression at the post transcriptional level. In this study miRNAs of *Curcuma longa* are predicted along with their possible target genes. A total of 12,593 ESTs were downloaded from dbEST database and processed and trimmed through SeqClean. This contig database was now used to find the putative miRNAs by performing a local BLAST with the miRNAs of *Arabidopsis thaliana* retrieved from miRBase. The targets were scanned by hybridizing screened ESTs with the UTRs of human using miRanda software. Finally, 12 putative miRNAs were found to hybridize with the various targets of signal transduction and apoptosis that may play significant role in preventing diseases like diabetes mellitus type 2, cardiovascular disorders, alzheimer, cancer, thalassemia by gene silencing.

Keywords: microRNA, computational methods, bioinformatics database, systems biology, gene expression.

[1] INTRODUCTION

One of the most important advances in biology in recent years is the discovery of RNAs that can regulate gene expression. As one kind of such functional non coding RNAs, microRNAs (miRNAs) form a class of endogenous 19–23-nucleotide RNAs that can have important regulatory roles in animals and plants by targeting transcripts for cleavage or translational repression [1]. Since the discovery of the very first miRNAs, computational methods have been an invaluable tool that can complement experimental approaches to understand the biology of miRNAs. Most computational approaches associated with miRNA research can be classified into two broad categories, namely miRNA gene identification and miRNA target prediction [2].

The discovery of microRNAs (miRNAs), a new class of negative regulator *that repress gene expression by pairing with* their target messenger RNAs (mRNAs), has revealed a natural pathway for controlling gene expression[3]. MicroRNAs (miRNAs) are a novel growing family of endogenous, small, non- coding, single-stranded RNA molecules directly involved in regulating

gene expression at the posttranscriptional level. High conservation of miRNAs in plant provides the foundation for identification of new miRNAs in other plant species through homology alignment [4]. They have been found to play key roles in regulatory functions of gene expression. These endogenous RNA sequences are the interest of intensive research in various model organisms, ranging from plants to mammals.

It is estimated that only 1% of the genomic transcripts in mammalian cells encode miRNA, nearly one-third of the encoded genes are regulated by miRNA [5]. Various bioinformatics databases, tools, and algorithms have been developed to predict the sequences of miRNAs and their target genes. In combination with the *in silico* approaches in systems biology, experimental studies on miRNA provide a new bioengineering approach for understanding the mechanism of gene regulation [6] .

To know miRNAs action at systems level, one has to know genome for miRNAs, their targeted genes and finally tracing the pathways and pin pointing the enzymes involved. This information will be highly helpful to develop the genetic and protein networks, which may lead to understand at phenotypic level changes. This may leads complete system biology approach for disease. This approach gives holistic approach for finding disease which occurs due to dysfunction of gene-gene interaction, gene-protein or protein-protein interaction. There are many tools for predicting this type of network [7]. The silent genes can be traced by KEGG (Kyoto Encyclopedia of Genome and Genomic Research) database to know their functional annotation in metabolomics of human.

1.1. *CURCUMA LONGA* A HIGH MEDICINAL VALUE HERB

Till date more than 120 species of *Curcuma* has been known. In this work miRNA of *Curcuma longa* and its effect on human metabolomics was predicted, a series of computational tools was used. Turmeric derived from the plant *curcuma longa* is used mainly in Indian subcontinent and many other Asian countries like China, Pakistan and Bangladesh. The active compound in *Curcuma longa* is curcumin which is used as coloring agent in dye for textile industries [8]. It is not only used as dye but also in health care and for the preservation of food. It is commonly used as spices and flavoring agent in many Indian cuisines.

From very early times spices have played an important role in Ayurvedic preparations. *Curcuma longa* (Zingiberaceae) commonly known as 'Haldi' in Hindi is also an important dietary spice. It is extensively used in Ayurveda, Unani and Siddha systems of medicine as home remedy for various diseases.

Curcumin is a active component *Cucurma longa* which is non-water-soluble polyphenol that can be derived from *C. longa* by acetone extraction [9]. Curcumin is having many properties such as antioxidant, anti-inflammatory, anti-viral, antibacterial, antifungal and anticancer activities and also works against various malignant diseases such as diabetes [10], Arthrosclerosis [11], allergies [12], arthritis, Alzheimer's [13] and other chronic diseases has been reported. In a study, diabetic rats were given curcumin showed a significant reduction in renal dysfunction and oxidative stress, which may indicate that curcumin has a protective role against diabetic nephropathy. Animal studies involving rats and mice, as well as *in vitro* studies utilizing human cell lines, have demonstrated curcumin's ability to inhibit carcinogenesis at three stages: tumor promotion, angiogenesis, and tumor growth.

Curcumin are also capable of suppressing the activity of several common mutagens and carcinogens in a variety of cell types, both *in vitro* and *in vivo*. A potent anti-inflammatory, turmeric works in two main ways: It suppresses an enzyme called cyclo-oxygenase (COX) that creates pro-inflammatory signals in the body, and it inhibits a gene that enhances production of pro-inflammatory molecules [14]. It has been observed

in vitro that curcumin is having strong antispasmodic effect. Curcuma also helps in curing lever disorders [15].

Recent studies show that curcumin is also having effect on cellular enzyme. In addition to this it also enhances immunity. When researchers looked at the lining of the intestine after ingestion of curcumin, they found that CD4+, T-helper and B type immune cells were greater in number. It has been observed that mice given curcumin has increased antibodies and more immune action.

[II] METHODS

miRNA genes can be searched by motif searches combining sequence, structure and conservation information. In order to facilitate the investigation into miRNA function, numerous bioinformatics methods were developed in order to allow high throughput prediction of miRNA target genes. Most miRNA target prediction algorithms use similar general principles in the development of their algorithm. Most algorithms search for targets in the 3'-UTR region of mRNAs, where almost all miRNA-target interactions occur in vertebrates. Algorithms usually account for the possibility of multiple target sites for more than one miRNA in each mRNA 3'-UTR region, though they differ in the degree in which combinations of miRNA target sites are incorporated into the prediction algorithm. The degree of sequence conservation is another criterion commonly used to filter possible miRNA targets. Most target prediction algorithms identify orthologous 3'-UTR sequences and check whether the miRNA-target interaction is conserved between closely related species. Many algorithms depend on an initial input of a miRNA or miRNAs to be queried and a set of genes to test for targets.

Earlier work shows that miRNA regulates gene expression by targeting the 3' untranslated region of specific messenger RNAs for degradation or translational repression. For the prediction of miRNA in *Curcuma longa* and the study of its effect on human metabolomics, a series of computational tools was used.

2.1. *Cucurma longa* EST Dataset: 12,593 raw ESTs sequences of Cucurma longa were downloaded from dbEST (Database of Express Sequence Tags). ESTs database has highest numbers of impurities associated with them which has to be removed for further processing. To remove impurities it was processed through SeqClean [16] a vector removing and trimming tool. It removes non redundant sequences and poly A tails. After that 12,590 ESTs were obtained. Out of these sequences 3 were trashed and 107 sequences were trimmed and then after this ESTs were processed to RepeatMasker for masking the repeated sequence. It is a program that screens DNA sequences for interspersed repeats and low complexity DNA sequences. The output of the program is a detailed annotation of the repeats that are present in the query sequence as well as a modified version of the query sequence in which all the annotated repeats have been masked (default: replaced by Ns). Masked file was a input for the TGICL [17] a clustering tool which clustered the sequences in masked file and made contig and .ace files. Clustering of ESTs is must for gene prediction and to know about the functional annotations and to understand the important genetic information's with variations such as those which are involved in diseases.

2.2. miRNA Homology detection: One simple way to search for new miRNAs is to search for homologs to already known pre-miRNA. Although most sequence-searching algorithms only use information in the primary sequence and neglect structural information, many pre-miRNAs are sufficiently conserved also on the sequence level to allow reliable detection. This method used characteristic features of known plant miRNAs as criteria to search for miRNAs conserved between *A. thaliana* and *C. longa*. miRNA of *Arabidopsis thaliana* [18] were downloaded from miRBase [19], a database of miRNAs and these miRNA were used as a query for local BLAST[4] against

contig obtained from TGICL. 12 putative miRNAs were obtained through local BLAST performed in CLC genomics work bench.

2.3. Strategy to predict mature miRNA from clustered ESTs: The miRNAs of *Arabidopsis thaliana* were compared with the assembled ESTs of *Curcuma longa* to identify regions in the human genome where experimentally validated miRNAs of *A. Thaliana* shows its presence in the contigs of *Curcuma longa*. Both mature and precursor miRNA matches were checked out in the resultant clustered contigs and singletons. The resultant miRNAs was compared with the other miRNAs through BLAST searches to generate a consensus for the predictions of novel mature miRNAs.

miRNA have folding energies comparable to that of random RNA, minimum folding energy value of 8 best putative miRNAs are taken. The query that whether the miRNAs are blocking or repressing any of the human genes was solved by hybridizing putative miRNAs with the 3' UTRs of human through miRanda software 8 out of 10 putative miRNAs were found to hybridize with the potential targets with a probability to block the action of those genes.

Figure1 *[last page]

[III] RESULTS AND DISCUSSIONS

Human 3' UTR were downloaded from the UTR database. Putative miRNAs were hybridized with human 3' UTR using miRanda a miRNA target finding tool [20]. The targets for the putative miRNAs were then scanned in the 3' UTRs with the help of miRanda software. Eight miRNA were found to hybridized with human 3' UTR. With the help of KEGG Brite pathway the genes for various metabolic pathways in human were tracked. This might silence the gene expression of diseases like Diabetes Mellitus type 2, cardiovascular disorders, Alzheimer, cancer, thalassemia and many more disease by gene silencing, (Table 1).

The query that whether the miRNAs are blocking or repressing any of the human genes was solved by hybridizing putative miRNA with 3' UTR of human through miRanda software. Table 1 shows the comparative account of finally selected miRNAs on the basis of hybridization energy, Blast score and minimum folding energy.

- >ath-miR167d was found to hybridize with the EIF2AK2 genes responsible for blocking pathway of protein processing in endoplasmic reticulum. Malfunction in this gene leads to disease like Alzheimer, Parkinson and type II diabetes mellitus [21]. Again it was observed that >ath-miR167d can hybridize with ZFYVE16 gene. This gene is involved in regulating membrane trafficking in the endosomal pathway. Over expression induces endosome aggregation which leads to disease like Arteriosclerosis, Hyperglycemia [22].
- >ath-miR169g hybridizes with PAPOLB, PAPT, TPAP. This gene is responsible for the mRNA surveillance pathway is a quality control mechanism that detects and degrades abnormal mRNAs. These pathways include nonsense-mediated mRNA decay (NMD), nonstop mRNA decay (NSD), and no-go decay. This is also responsible for ER stress, which leads to disease like Mendelian disorders, Wolfram syndrome and complex diseases amyotrophic lateral sclerosis and diabetes [23].
- >ath-miR169g* was found to be responsible for silencing CHRM3, which is candidate gene for type 2 diabetes in pima Indians [24].
- >ath-miR167d hybridizes with CFC1, GDF1, NKX2-6, TBX1. This gene is responsible for Conotruncal heart malformations comprise a group of clinically significant congenital heart defects [25]. Gene GSTM1, GSTM5 was found to hybridize with
- >ath-miR854d which in turn is associated with early ageing, methyl mercury detoxification [26].
- >ath-mir851-3p hybridizes with HSD17B3 gene. This gene is associated with reproductive system disease [27]. >ath-mir851-3p hybridization product is ATM,

MBD5 and STAT5B. This gene is responsible for primary immunodeficiency, infectious disease, congenital disorders, mental retardation and musculoskeletal diseases [28].

[IV] CONCLUSION

Present studies shows that *C. longa* is a very important medicinal plant for human against various diseases. The miRNA predicted from it were found to target various genes on metabolic pathway.

Out of 12 only 8 miRNA were finally hybridized with 3' UTR of human and they silent the genes responsible for the signal transduction, apoptosis might be preventing the diseases like Diabetes Mellitus type 2, Cardiovascular disorders, Alzheimer, Cancer, Thalassemia and many more diseases by gene silencing.

In a study it has been observed that diabetes mellitus is usually associated with elevated serum lipid levels considered as risk factors for coronary heart diseases [29, 30]. Lowering of these elevated levels either though drug or dietary therapy seems to be associated with a decrease in the risk of vascular disease.

Beta-Amyloid (betaA)-induced oxidative stress is a well-established pathway of neuronal cell death in Alzheimer's disease. curcumin may reduce levels of amyloid and oxidized proteins and prevent cognitive deficits.

Based on animal study, oral administration of curcumin may reduce expression of several cytokines, chemokines, and proteinases known to mediate aneurismal degeneration [31]. Curcumin also inhibits the secretion of collagenase, elastase, and hyaluronidase. Inhibition of neutrophil function has been note [32], and *in vitro* research demonstrates that curcumin inhibits 5-hydroxy-eicosatetraenoic acid (5-HETE) in intact human neutrophils.

In vitro, curcumin inhibits platelet aggregation induced by ADP, epinephrine, or collagen [33]. Turmeric appears to inhibit arachidonic acid incorporation into platelet phospholipids, degradation of phospholipids, and cyclooxygenase

The study concludes that this first reporting of the scientific exploration of multiple action potential of *C. longa* miRNA with special reference in type I diabetes models is not only interesting but very encouraging too for medicinal chemists as well as diabetic patients globally. These predicted miRNA may serve as potential resource for initiating experimental validation which may provide valuable information for functional analysis of miRNA regulated pathway.

REFERENCES

[1] Bartel DP, [2004] "MicroRNAs: genomics, biogenesis, mechanism, and function". Cell, 116(2), 281-97.

[2] Morten Lindow and Jan Gorodkin [2007] "Principles and limitations of Computational microRNA Gene and Target Finding". DNA and Cell Biology,Volume 26,Number 5.

[3] Pang KC, Frith MC, Mattick JS. Rapid evolution of noncoding RNAs [2006.] "lack of conservation does not mean lack of function". Trends Genet.22, 1-5.

[4] Altschul SF et al. [1990] "Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes". Proc. Nat. Acad Sci USA. 87, 2264–2268.

[5] John B, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS.[2004]. "Human MicroRNA targets", PLoS Biol 2(11), e363.

[6] Philipp Berninger , Dimos Gaidatzis, Erik van Nimwegen, Mihaela Zavolan.[2008] "Computational analysis of small RNA cloning data". Elsevier, Methods 44, 13-21.

[7] Rameshwari R, Prasad T.V.Prasad. [2011] "Systematic and Integrative Analysis of Proteomic Data using Bioinformatics Tools". International Journal of Advanced Computer Science and Applications, Volume 2 issue .

[8] Bharat B. Aggarwal, Chitra Sundaram, Nikita Malani, and Haruyo Ichikawa.[2007] "Molecular target and therapeutics uses of Curcumin in health and diseases". Springer.

[9] Revathy.S et al. [2011] "Isolation, Purification and Identification of Curcuminoids from Turmeric (*Curcuma longa* L.) by Column Chromatography". Journal of experimental sciences, 2(7):21-25.

[10] Saran Shantikumar, Andrea Caporali, and Costanza Emanueli [December 2011] "Role of microRNAs in diabetes and its Cardiovascular complications", European society of cardiology.

[11] A. Ramirez Bosca, M. A. Carrión Gutierrez, et. al. [1997] "Effects of the antioxidant turmeric on

lipoprotein peroxides: Implications for the prevention of atherosclerosis". AGE, Volume 20, Number(3), 165-168,.

[12] Liddle M, Hull C, Liu C. Powell D. [2006] "Contact urticaria from curcumin", Dermatitis, 17:196 7

[13] Shrikant Mishra and Kalpana Palanivelu. [2008] "The effect of curcumin (turmeric) on Alzheimer's disease: An overview", Indian Acad Neurol, Jan-Mar; 11(1): 13–19.

[14] Chainani N. [2003] "Safety and anti-inflammatory activity of Curcumin component of turmeric (curcuma longa)", Compl Med. ;9:161–8.

[15] Ammon HP, Wahl MA [1991] "Pharmacology of Curcuma longa", Planta Med., Feb, 57(1):1-7.

[16] Falgueras J, et al.[2010] BMC Bioinformatics. 11: 38 [PMID: 20089148].

[17] Geo Pertea, et al.[2002] "TIGR Gene Indices clustering tools (TGICL), a software system for fast clustering of large EST datasets". Bioinformatics, Vol. 19, issue 5.

[18] A. Adail et al. [2005] "Computational prediction of miRNAs in Arabidopsis thaliana". Genome Res. 15, 78–91.

[19] S. Griffiths-Jones et al. [2006] "miRBase: microRNA sequences targets and gene nomenclature", Nucleic acids Res. 34, D140-144.

[20] Miranda KC, Huynh T, Tay Y, Ang YS, Tam WL, Thomson AM, Lim B, Rigoutsos I [2006] "A pattern based method for the identification of microRNA binding sites and their corresponding heteroduplexes". Cell 126(6), 1203–1217.

[21] Han SG, et al. [2012] "EGCG protects endothelial cells against PCB 126-induced inflammation through inhibition of AhR and induction of Nrf2-regulated gene", Toxicol Appl Pharmacol , Jun 1;261(2):181-8.

[22] Seet LF, Liu N, Hanson BJ, Hong W [2004]."Endofin recruits TOM1 to endosomes". J. Biol. Chem. 279 (6): 4670–9.

[23] Dombroski BA et al., [2010]. "Gene expression and genetic variation in response to endoplasmic reticulum stress in human cells", American Journal of Human Genetics, May14; 86(15),719-29.

[24] Yan Guo et. al. [2010]. "CHRM3 Gene Variation Is Associated With Decreased Acute Insulin Secretion and Increased Risk for Early-Onset Type 2 Diabetes in Pima Indians", American Diabetes Association, Vol. 55, no. 12, 3625-3629.

[25] Davit-Spraul A. [2008] "CFC1 gene involvement in biliary atresia with polysplenia syndrome", J Pediatr Gastroenterol Nutr., Jan;46(1):111-2.

[26] Mazzaron Barcelos GR, et al. [2012]. "Evaluation of glutathione S-transferase GSTM1 and GSTT1 polymorphisms and methyl mercury metabolism in an exposed Amazon population, J Toxicol Environ Health A.

[27] Lee YS, Kirk JM, Stanhope RG, et al. [2007]. "Phenotypic variability in 17beta-hydroxysteroid dehydrogenase-3 deficiency and diagnostic pitfalls." Clin. Endocrinol. (Oxf) 67 (1): 20–8.

[28] Talkowski ME, et al. [2011] "Assessment of 2q23.1 microdeletion syndrome implicates MBD5 as a single causal locus of intellectual disability, epilepsy, and autism spectrum disorder", Am J Hum Genet, 89(4):551-63.

[29] Hofteizer V. [1973] "Comparison of streptozotocin-induced diabetes in the rat inducing volumetric quantitation of the pancreatic islets", Diabetologia, *9*: 178-84.

[30] Goodman LS, Gilman A. [1985] "The Pharmacological Basis of Therapeutics",7th ed., Macmillan, New York; 1490-510.

[31] Joe, B. and Lokesh, B. R.[1997] "Effect of curcumin and capsaicin on arachidonic acid metabolism and lysosomal enzyme secretion by rat peritoneal macrophages". Lipids;32(11):1173-1180.

[32] Ammon, H. P., Safayhi, H., Mack, T., and Sabieraj, J.[1993] "Mechanism of antiinflammatory actions of curcumine and boswellic acids". J Ethnopharmacol;38(2-3):113-119,

[33] Srivastava, K. C., Bordia, A., and Verma, S. K. [1995] "Curcumin, a major component of food spice turmeric (Curcuma longa) inhibits aggregation and alters eicosanoid metabolism in human blood platelets". Prostaglandins Leukot Essent. Fatty Acids,;52(4):223-227.

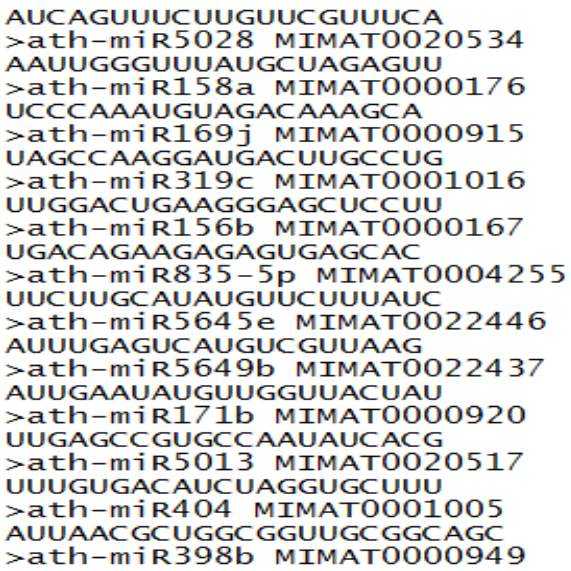

Figure 1: Arabidopsis thaliana miRNAs

[Table -1]

S. No	Identifiers	Contigs	miRNA	BLAST Score	MFE Value	GENE ID	Gene affected	Linked Disease/ Disorder
1	NM_002953	CL1853CONTIG1	>ath-miR167d	159	-25.959999	6195	EIF2AK2	Alzheimer, Parkinson and type II diabetes mellitus
2	NM_007324	CL1853CONTIG1	>ath-miR167d	164	-22.760000	9372	ZFYVE16, ENDOFIN, PPP1R69	Arteriosclerosis, Hyperglycemia
3	NM_022894	CL2599CONTIG1	>ath-miR169g	152	-21.670000	64895	PAPOLB, PAPT, TPAP	Mendelian disorders, Wolfram syndrome and complex diseases amyotrophic lateral sclerosis and diabetes
4	NM_001202519	CL210CONTIG1	>ath-miR169g*	152	-21.870001	8837	CHRM3	for type 2 diabetes
5	NM_000651	CL511CONTIG1	>ath-miR167d	152	-21.110001	1378	CFC1, GDF1, NKX2-6, TBX1	congenital heart defects
6	NM_000851	CL1925CONTIG1	>ath-miR854d	173	-24.840000	2949	GSTM1, GSTM5	early ageing
7	NM_176877	CL311CONTIG1	>ath-mir851-3p	159	-26.190001	10207	HSD17B3	reproductive system disease
8	NM_000721	CL311CONTIG1	>ath-mir851-3p	151	-21.370001	777	MBD5, ATM, LIG4	primary immunodeficiency, infectious disease, congenital disorders, mental retardation and musculoskeletal diseases

Table 1: Different miRNAs having minimum folding energy (MFE) which can silence genes responsible for different diseases.