

Computational prediction of miRNA in *Gmelina arborea* and their role in human metabolomics

Alka Dubey^{1,*}, Swinderjeet Singh Kalra², Nandita Trivedi³

¹Department of Zoology, Bioinformatics infrastructure facility centre (DBT), Ministry of Science and Technology (Govt. of India), D.G (PG) College civil lines Kanpur, INDIA

²Department of Chemistry, D.A-V, College Kanpur INDIA

³Department of Biotechnology, CSJM University, Kanpur INDIA

Email address:

Alkabioinfo964@gmail.com (A. Dubey)

To cite this article:

Alka Dubey, Swinderjeet Singh Kalra, Nandita Trivedi. Computational Prediction of miRNA in *Gmelina Arborea* and their Role in Human Metabolomics. *American Journal of Bioscience and Bioengineering*. Vol. 1, No. 5, 2013, pp. 62-74.

doi: 10.11648/j.bio.20130105.12

Abstract: Micro RNA are the noncoding class of RNA which plays major role post transcriptional regulation of genes and 19-22 nt in length and endogenous in nature. In this study miRNA *Gmelina arborea* are predicted along with possible target genes, A total of 44 ESTs are taken from dbEST database and proceed and trimmed through EGASSEMBLER. This contig sequence was now used to find out the putative miRNA by performing a local alignment with miRNA of *Arabidopsis thaliana* retrieved from miRBase. The targets were scanned by hybridizing screened ESTs with the UTRs of human using DIANA tar software. Finally 6 putative miRNAs were found to hybridize with the various targets of signal transduction and apoptosis that play significant role in preventing disease like cancer, blood borne disease, and other urinary infections and various study indicated that it possesses antioxidants, diuretic, anthelmintic activity, cardio protective, anti diabetic, immunomodulatory, antipyretic and analgesic, antimicrobial in the treatment of jaundice, piles bases on their targets

Keywords: *Gmelina Arborea*, miRNA, Dbest, DIANA, Local Alignment

1. Introduction

1.1 About miRNA

MicroRNA is a family of small non-coding RNAs that regulate gene expression in a sequence-specific manner. miRNAs are a class of post-transcriptional regulators. miRNAs are a family of 19 to 25 small [1] miRNAs function via base-pairing with complementary sequences within mRNA molecules, usually resulting in gene silencing via translational repression or target degradation.[2] in plant miRNA biogenesis in plants differs from animal biogenesis mainly in the steps of nuclear processing and export. Instead of being cleaved by two different enzymes, once inside and once outside the nucleus, both cleavages of the

Plant miRNA is performed by a Dicer homolog, called Dicer-like1 (DL1). DL1 is only expressed in the nucleus of plant cells, which indicates that both reactions take place inside the nucleus. Before plant miRNA:miRNA* duplexes are transported out of the nucleus, its 3' overhangs

are methylated by a RNA methyl transferase protein called Hua-Enhancer1 (HEN1). The duplex is then transported out of the nucleus to the cytoplasm by a protein called Hasty (HST), an Exportin 5 homolog, where they disassemble and the mature miRNA is incorporated into the RISC.[3]

1.2. *Gmelina Arborea*. a High Medicinal Value Herb

The root and bark of *Gmelina arborea* are claimed to be stomachic, galactagogue laxative and anthelmintic; improve appetite, useful in hallucination, piles, abdominal pains, burning sensations, fevers, 'tridosha' and urinary discharge. Leaf paste is applied to relieve headache and juice is used as wash for ulcers. Flowers are sweet, cooling, bitter, acrid and astringent. They are useful in leprosy and blood diseases. In Ayurveda, it has been observed that Gamhar fruit is acrid, sour, bitter, sweet, cooling, diuretic tonic, aphrodisiac, alternative astringent to the bowels, promote growth of hairs, useful in 'vata', thirst, anaemia, leprosy, ulcers and vaginal discharge. The plant is recommended in combination with other drugs for the treatment of snakebite

and scorpion sting. In snakebite a decoction of the root and bark is given internally. [4]

2. Materials and Methods

The EST sequences of *Gmelina arborea* were retrieved from dbEST (NCBI), for clean the EST via EG assembler (seq cleaner) were used to find out the contig files and the precursors and secondary energy (MFE) was found with the help of RNA fold then we used to find out the putative miRNA from *Gmelina arborea* via the local alignment with *Arabidopsis thaliana*. This miRNA predicted the target site, and the target site prediction and UTRs hybridization and chromosomes position analysis with involves gene metabolic activity via the KEGG path way analysis completed with the help of DIANA web server. 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 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. 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 *Gmelina arborea* and the study of its effect on human metabolomics, a series of computational tools was used. [5]

2.1. *Gmelina Arborea* EST Dataset

44 raw ESTs sequences of *Gmelina arborea* 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 [6] a vector removing and trimming tool. It removes non redundant sequences and polyA tails. After that 44 ESTs were obtained. Out of these sequences 40 were trashed and 1 sequence was trimmed and then after this ESTs were processed to Repeat Masker for masking the repeated sequence. It is a program that screens DNA sequences for interspersed repeats and low complexity

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 an input for the TGICL [7] 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.

3. Results

Computational prediction of miRNA in *Gmelina arborea* and their role in human metabolomics analysis depends Upon different computational prediction analysis.

3.1. Analysis of Retrieved ESTs for *Gmelina Arborea*

For this analysis EST retrieves from dbEST via NCBI according raw data GA(*Gmelina arborea*) carrying a 44 sequences. An expressed sequence tag or EST is a short sub-sequence of a cDNA sequence. They may be used to identify gene transcripts, and are instrumental in gene discovery and gene sequence determination. [8] The identification of ESTs has proceeded rapidly, with approximately 74.2 million ESTs now available in public databases.

3.2. Analysis of Sequence Cleaner

Because 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 Seq Cleaner (EG assembler) then analyze the basic five outputs like sequence cleaning, repeat masking, vector masking, organelle masking, sequence was assembling.

3.3. Strategy to Predict Mature miRNA and their Targeted Genes

The miRNAs of *Arabidopsis thaliana* were compared with the assembled ESTs of *Gmelina arborea* to identify regions in the human genome where experimentally validated miRNAs of *A. Thaliana* shows its presence in the contigs of *Gmelina arborea*. 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 clustal omega searches to generate a consensus for the predictions of novel mature miRNAs basis on conserved regions and alignments between *Arabidopsis thaliana* miRNAs and *Gmelina arborea* contigs precursors the putative miRNA given below.

Table 1: miRNA sequence obtained through contig file and local alignments

Serial numbers	Annotated name of miRNA	miRNA seq in nucleotide
1.	miRNA1_GA_CONTIG1	UGGUAGGGCCCUUCUGAA
2.	miRNA2_GA_CONTIG1	GGCACCGCUGCCAGCGGUAU
3.	miRNA3_GA_CONTIG1	AUGCUGUUUUGGUGGUAGGGCCC
4.	miRNA4_GA_CONTIG1	GAAUUGCAGAAUUCAG
5.	miRNA5_GA_CONTIG1	GCAUGCCUGUCCGAGC
6.	miRNA6_GA_CONTIG1	ACUCAAGACUCUUUUGUUCU

Table 1 explains : first coulam explains the serial number and second columns explains the annotated name of miRNA where miRNA as micro RNA and GA as *Gmelina arborea* and CONTIG means are given RNAs Obtained from contig 1 and Colum thirds explains the nucleotide sequence as miRNA

Table: 2. Name of target genes via putative miRNA from *Gmelina arborea*

>miRNA1_GA_CONTIG1 (UGGUAGGGCCCUUCUGAA)						
S.No	Names Of Genes	Position On Chromosome	3'UTR	Target Sites Involved In Conserving Information		
1.	SH3PXD2A (SH3 and PX domain-containing protein 2A (SH3 multiple domains protein 1) (Five SH3 domain-containing protein) (Adaptor protein TKS5).	10: 105351051 - 105351079	483 - 511	Not Conserved		
		10: 105349750 - 105349778	1784 - 1812	Not Conserved		
		10: 105349745 - 105349773	1789 - 1817	Not Conserved		
		10: 105347560 - 105347588	3974 - 4002	oryCun1		
		10: 105347307 - 105347335	4227 - 4255	Not Conserved		
		10: 105346886 - 105346914	4648 - 4676	canFam2, loxAfr1		
		10: 105345560 - 105345588	5974 - 6002	Not Conserved		
2.	ZNF275 (Zinc finger protein 275.)	10: 105344507 - 105344535	7027 - 7055	Not Conserved		
		X: 152270243 - 152270271	3817 - 3845	monDom4		
		X: 152270400 - 152270428	3974 - 4002	Not Conserved		
		X: 152270484 - 152270512	4058 - 4086	Not Conserved		
		X: 152271494 - 152271522	5068 - 5096	rn4, mm8, oryCun1, canFam2		
		>miRNA2_GG_CONTIG1 (GGCACCGCUGCCAGCGGUAU)				
		1.	MAPK1 Mitogen-activated protein kinase 1 (EC 2.7.11.24) (Extracellular signal-regulated kinase 2) (ERK-2) (Mitogen-activated protein kinase 2) (MAP kinase 2) (MAPK 2) (p42-MAPK) (ERT1).	22: 20444943 - 20444971	3567 - 3595	oryCun1, bosTau2, canFam2, dasNov1, loxAfr1, echTel1, galGal2, xenTro1
				22: 20447232 - 20447260	1278 - 1306	mm8, oryCun1, canFam2, xenTro1
22: 20444382 - 20444410	4128 - 4156			Not Conserved		
>miRNA3_GA_CONTIG1 (AUGCUGUUUUGGUGGUAGGGCCC)						
1	DST (Bullous pemphigoid antigen 1)	6: 56435863 - 56435891	14487 - 14515	rn4, mm8, oryCun1, bosTau2, canFam2, dasNov1, loxAfr1, echTel1, galGal2, xenTro1, tetNig1		
		6:56509694;56524941 56509697;56524966	8200 - 8228	oryCun1, canFam2, loxAfr1		
		6: 56571310 - 56571338	3455 - 3483	canFam2, galGal2		
		6: 56505225 - 56505253	8570 - 8598	canFam2, loxAfr1, echTel1		
		6: 56481326 - 56481354	10774 - 10802	oryCun1, galGal2		

		66474445;56476754	11143 - 11171	bosTau2, echTel1
		56474450;56476777		
		6: 56474298 - 56474326	11290 - 11318	canFam2, dasNov1, loxAfr1, echTel1, monDom4, tetNig1
2.	ZNF 621 (Zinc finger protein 621)	3: 40551577 - 40551605	1991 - 2019	Not Conserved
		3: 40553291 - 40553319	3705 - 3733	Not Conserved
		3: 40554081 - 40554109	4495 - 4523	Not Conserved
		3: 40554636 - 40554664	5050 - 5078	Not Conserved
		3: 40555613 - 40555641	6027 - 6055	Not Conserved
		3: 40555683 - 40555711	6097 - 6125	Not Conserved
3.	APKAP9 (A-kinase anchor protein 9 (Protein kinase A-anchoring protein 9) (PRKA9) (A-kinase anchor protein 450 kDa) (AKAP 450) (A-kinase anchor protein 350 kDa) (AKAP 350) (hgAKAP 350) (AKAP 120-like protein) (Protein hyperion) (Protein yotiao) (Centrosome- and Go	7: 91565375 - 91565403	5706 - 5734	rn4, oryCun1, bosTau2, canFam2, dasNov1, loxAfr1, echTel1
		7: 91509396 - 91509424	124 - 152	bosTau2, canFam2, loxAfr1, galGal2
		7: 91512300 - 91512328	309 - 337	oryCun1, bosTau2, loxAfr1
		7: 91529641 - 91529669	983 - 1011	bosTau2, canFam2, dasNov1
		7: 91533725 - 91533753	1358 - 1386	canFam2, loxAfr1, echTel1, monDom4
		7: 91546571 - 91546599	2282 - 2310	bosTau2, canFam2, dasNov1
		7: 91560530 - 91560558	4628 - 4656	dasNov1, loxAfr1
4	NTRK3 (NT-3 growth factor receptor precursor (EC 2.7.10.1) (Neurotrophic tyrosine kinase receptor type 3) (TrkC tyrosine kinase) (GP145-TrkC) (Trk-C).	15: 86216544 - 86216572	4597 - 4625	bosTau2, canFam2, dasNov1
		15: 86220424 - 86220452	717 - 745	Not Conserved
		15: 86213289 - 86213317	7852 - 7880	Not Conserved
		15: 86208464 - 86208492	12677 - 12705	Not Conserved
		15: 86205238 - 86205266	15903 - 15931	Not Conserved
5	SMAD2 (Mothers against decapentaplegic homolog 2 (SMAD 2) (Mothers against DPP homolog 2) (Mad-related protein 2) (hMAD-2) (JV18-1) (hSMAD2).	18: 43622080 - 43622108	87 - 115	rn4, mm8, bosTau2, canFam2, dasNov1, loxAfr1, echTel1, monDom4, galGal2, xenTro1
		18: 43621421 - 43621449	746 - 774	Not Conserved
		18: 43621117 - 43621145	1050 - 1078	Not Conserved
		18: 43615408 - 43615436	6759 - 6787	Not Conserved
6	TRPM3 (Transient receptor potential cation channel subfamily M member 3 (Long transient receptor potential channel 3) (LTpC3) (Melastatin-2) (MLSN2).	9: 72631703 - 72631731	843 - 871	dasNov1
		9: 72630398 - 72630426	2148 - 2176	Not Conserved
		9: 72629666 - 72629694	2880 - 2908	Not Conserved
		9: 72629626 - 72629654	2920 - 2948	Not Conserved
		9: 72629518 - 72629546	3028 - 3056	canFam2
7	UBR4 (Zinc finger UBR1-type protein 1 (Retinoblastoma-associated factor of 600 kDa) (600 kDa retinoblastoma protein-associated factor) (RBAF600) (p600)	1: 19286981 - 19287009	4322 - 4350	dasNov1
		1: 19285828 - 19285856	4480 - 4508	rn4, canFam2, dasNov1, loxAfr1, echTel1
		1: 19313037 - 19313065	1202 - 1230	Not Conserved
		1: 19293167 - 19293195	3666 - 3694	rn4, mm8, bosTau2, dasNov1, echTel1, galGal2, xenTro1
		1: 19292494 - 19292522	3876 - 3904	rn4, mm8, oryCun1, bosTau2, dasNov1, loxAfr1, echTel1, galGal2, xenTro1, tetNig1
8	TAP2 (Antigen peptide transporter 2 (APT2) (ATP-binding cassette sub-family B	6: 32902929 - 32902957	1703 - 1731	rn4, mm8
		6: 32902004 - 32902032	2628 - 2656	Not Conserved
		6: 32901236 - 32901264	3396 - 3424	Not Conserved

	member 3) (Peptide transporter TAP2) (Peptide transporter PSF2) (Peptide supply factor 2) (PSF-2) (Peptide transporter involved in antigen processing 2)			
9	RBM33 (Proline-rich protein 8)	7: 155261061 - 155261089	564 - 592	Not Conserved
		7: 155263770 - 155263798	3273 - 3301	Not Conserved
		7: 155265407 - 155265435	4910 - 4938	Not Conserved
10	FCER2 (Low affinity immunoglobulin epsilon Fc receptor (Lymphocyte IgE receptor) (Fc-epsilon-RII) (BLAST-2) (Immunoglobulin E-binding factor) (CD23 antigen) [Contains: Low affinity immunoglobulin epsilon Fc receptor membrane-bound form; Low affinity immunoglobul	19: 7659905 - 7659933	145 - 173	Not Conserved
	19: 7659874 - 7659902	176 - 204	Not Conserved	
	19: 7659842 - 7659870	208 - 236	Not Conserved	
	19: 7659810 - 7659838	240 - 268	Not Conserved	
	19: 7659778 - 7659806	272 - 300	Not Conserved	
11	CPNE3 (Copine-3 (Copine III).	8: 87640290 - 87640318	535 - 563	Not Conserved
		8: 87640856 - 87640884	1101 - 1129	Not Conserved
		8: 87641578 - 87641606	1823 - 1851	canFam2
12	NBPF9 (NBPF8 isoform 6.)	1: 143533394 - 143533422	760 - 788	Not Conserved
		1: 143540031 - 143540059	1717 - 1745	Not Conserved
		1: 143540738 - 143540766	2424 - 2452	Not Conserved
		1: 143540816 - 143540844	2502 - 2530	Not Conserved
13	C20orf177 (Uncharacterized protein C20orf177)	20: 57954912 - 57954940	1366 - 1394	Not Conserved
		20: 57955263 - 57955291	1717 - 1745	Not Conserved
		20: 57957086 - 57957114	3540 - 3568	Not Conserved
14	CDH23 (Cadherin-23 precursor (Otocadherin).	10: 73142541 - 73142569	43 - 71	rn4, mm8, oryCun1, bosTau2, canFam2, echTel1
		10: 73209174 - 73209202	2029 - 2057	bosTau2, echTel1, monDom4, galGal2, tetNig1
		10:73220160;73220895 73220176;73220907	2725 - 2753	Not Conserved
		10: 73220990 - 73221018	2836 - 2864	galGal2, xenTro1
		10: 73237705 - 73237733	5335 - 5363	oryCun1, canFam2, dasNov1, echTel1, galGal2
15	KIAA529 (KIAA1529)	9: 99172151 - 99172179	2133 - 2161	Not Conserved
		9: 99172841 - 99172869	2264 - 2292	Not Conserved
		9: 99179146 - 99179174	2935 - 2963	Not Conserved
		9: 99179293 - 99179321	3082 - 3110	Not Conserved
16	MLL (Zinc finger protein HRX (ALL-1) (Trithorax-like protein).	11: 117870287 - 117870315	3792 - 3820	rn4, mm8, oryCun1, bosTau2, canFam2, dasNov1, loxAfr1, echTel1, tetNig1
		11: 117877013 - 117877041	4791 - 4819	mm8, oryCun1, bosTau2, dasNov1, loxAfr1, echTel1, xenTro1
		11: 117899388 - 117899416	11730 - 11758	bosTau2, canFam2
>miRNA4_GA_CONTIG1 (GAAUUGCAGAAUUCAG				

1	ESSRG (Estrogen-related receptor gamma (Estrogen receptor-related protein 3) (ERR gamma-2).	1: 214745325 - 214745353	1550 - 1578	rn4, mm8, bosTau2, loxAfr1, echTel1
		1: 214745751 - 214745779	1124 - 1152	rn4, mm8, oryCun1, bosTau2, canFam2, loxAfr1, echTel1, monDom4, galGal2, xenTro1
		1: 214744387 - 214744415	2488 - 2516	rn4, mm8, oryCun1, bosTau2, canFam2, dasNov1, loxAfr1, echTel1, monDom4, galGal2
		1: 214745953 - 214745981	922 - 950	oryCun1, bosTau2, loxAfr1
2	NFIB (Nuclear factor 1 B-type (Nuclear factor 1/B) (NF1-B) (NFI-B) (NF-1/B) (CCAAT-box-binding transcription factor) (CTF) (TGGCA-binding protein).	9: 14073314 - 14073342	4965 - 4993	bosTau2, canFam2, dasNov1, echTel1, galGal2, xenTro1
		9: 14077424 - 14077452	855 - 883	rn4, mm8, oryCun1, bosTau2, canFam2, dasNov1, loxAfr1, monDom4, galGal2
		9: 14073041 - 14073069	5238 - 5266	rn4, mm8, oryCun1, bosTau2, canFam2, echTel1, monDom4, galGal2, xenTro1
		9: 14073133 - 14073161	5146 - 5174	mm8, bosTau2, canFam2, monDom4, xenTro1
3	VCPIP1 (Deubiquitinating protein VCIP135 (EC 3.4.22.-) (Valosin-containing protein p97/p47 complex-interacting protein p135) (Valosin-containing protein p97/p47 complex-interacting protein 1).	8: 67707963 - 67707991	1298 - 1326	rn4
		8: 67706404 - 67706432	2857 - 2885	bosTau2, canFam2
		8: 67706812 - 67706840	2449 - 2477	Not Conserved
		8: 67706374 - 67706402	2887 - 2915	Not Conserved
4	SATB1 (DNA-binding protein SATB1 (Special AT-rich sequence-binding protein 1).	3: 18364785 - 18364813	852 - 880	rn4, mm8, oryCun1, bosTau2, canFam2, dasNov1, loxAfr1, echTel1, monDom4, galGal2, xenTro1
		3: 18364748 - 18364776	889 - 917	rn4, mm8, oryCun1, bosTau2, canFam2, loxAfr1, echTel1, monDom4, galGal2, xenTro1, tetNig1
		3: 18364682 - 18364710	955 - 983	rn4, mm8, oryCun1, bosTau2, canFam2, dasNov1, loxAfr1, echTel1, monDom4, galGal2, xenTro1
5	SGK269 (Tyrosine-protein kinase SgK269 (EC 2.7.10.2) (Sugen kinase 269).	15: 75192237 - 75192265	1287 - 1315	rn4, mm8, bosTau2, dasNov1, galGal2
		15: 75192695 - 75192723	829 - 857	mm8, loxAfr1
		15: 75189373 - 75189401	4151 - 4179	Not Conserved
		15: 75190325 - 75190353	3199 - 3227	Not Conserved
6	G3BP2 (Ras GTPase-activating protein-binding protein 2 (G3BP-2) (GAP SH3 domain-binding protein 2)	4: 76789250 - 76789278	359 - 387	rn4, mm8, oryCun1, bosTau2, canFam2, dasNov1, loxAfr1, echTel1, monDom4, galGal2, tetNig1
		4: 76789241 - 76789269	368 - 396	rn4, mm8, bosTau2, canFam2, loxAfr1, echTel1, monDom4
		4: 76788319 - 76788347	1290 - 1318	rn4, mm8, oryCun1, bosTau2, canFam2, monDom4, galGal2
7	ONECUT2 (One cut domain family member 2 (Transcription factor ONECUT-2) (OC-2).	18: 53302147 - 53302175	7193 - 7221	rn4, oryCun1, bosTau2, canFam2, echTel1
		18: 53304701 - 53304729	9747 - 9775	rn4, mm8, bosTau2, dasNov1, loxAfr1
		18: 53309471 - 53309499	14517 - 14545	oryCun1, bosTau2, canFam2, loxAfr1, xenTro1
		18: 53307520 - 53307548	12566 - 12594	Not Conserved
8	BCAT1 (Branched-chain-amino-acid aminotransferase)	12: 24859551 - 24859579	2629 - 2657	mm8, oryCun1, bosTau2, canFam2, loxAfr1, echTel1
		12: 24859509 - 24859537	2671 - 2699	bosTau2, dasNov1, loxAfr1, echTel1
		12: 24856005 - 24856033	6175 - 6203	loxAfr1, galGal2
		12: 24861843 - 24861871	337 - 365	Not Conserved

9	PDS5B	13: 32248010 - 32248038	819 - 847	rn4, mm8, oryCun1, bosTau2, canFam2, dasNov1, loxAfr1
	(Androgen-induced proliferation inhibitor (Androgen-induced prostate proliferative shutoff-associated protein AS3) (Sister chromatid cohesion protein PDS5 homolog B).	13: 32249400 - 32249428	2209 - 2237	rn4, mm8
		13: 32247595 - 32247623	404 - 432	rn4, mm8, oryCun1, canFam2, dasNov1
		13: 32248890 - 32248918	1699 - 1727	Not Conserved
10	TNRC6B	22: 39053089 - 39053117	3897 - 3925	bosTau2, canFam2, dasNov1, loxAfr1
	(Trinucleotide repeat-containing 6B protein.)	22: 39051494 - 39051522	2302 - 2330	Not Conserved
		22: 39057884 - 39057912	8692 - 872	Not Conserved
		22: 39058466 - 39058494	9274 - 9302	Not Conserved
11	GRK7	3: 143018910 - 143018938	327 - 355	oryCun1, canFam2, dasNov1, echTel1, monDom4
	(G protein-coupled receptor kinase 7 precursor (EC 2.7.11.16) (G protein-coupled receptor kinase GRK7).	3: 143019117 - 143019145	534 - 562	Not Conserved
		3: 143019734 - 143019762	1151 - 1179	Not Conserved
12	EEA1	12: 91690948 - 91690976	2941 - 2969	canFam2, dasNov1, loxAfr1, echTel1
	(Early endosome antigen 1)	12: 91690576 - 91690604	3313 - 3341	oryCun1
	(Endosome-associated protein p162)	12: 91693577 - 91693605	312 - 340	Not Conserved
	(Zinc finger FYVE domain-containing protein 2).	12: 91690736 - 91690764	3153 - 3181	Not Conserved
13	SOCS5	2: 46842834 - 46842862	2049 - 2077	rn4, mm8, bosTau2, canFam2, dasNov1, loxAfr1, echTel1, monDom4
	(Suppressor of cytokine signaling 5 (SOCS-5) (Cytokine-inducible SH2-containing protein 5)	2: 46841257 - 46841285	472 - 500	mm8, oryCun1, canFam2, dasNov1, loxAfr1
	(Cytokine-inducible SH2 protein 6) (CIS-6).	2: 46841767 - 46841795	982 - 1010	echTel1
		2: 46843207 - 46843235	2422 - 2450	echTel1
>miRNA5_GA_CONTIG1 (GCAUGCCUGUCCGAGC)				
1	ZNF677 (Zinc finger protein 677.)	19: 58431395 - 58431423	615 - 643	Not Conserved
		19: 58431140 - 58431168	870 - 898	Not Conserved
		19: 58430990 - 58431018	1020 - 1048	Not Conserved
		19: 58429002 - 58429030	2602 - 2630	Not Conserved
		19: 58427732 - 58427760	3872 - 3900	Not Conserved
		19: 58426257 - 58426285	5347 - 5375	Not Conserved
		19: 58423998;58425445 58424011;58425460	6172 - 6200	Not Conserved
2	RAB11FIP4 (Rab11 family-interacting protein 4 (Rab11-FIP4).	17: 26884596 - 26884624	1725 - 1753	dasNov1
		17: 26885526 - 26885554	2655 - 2683	Not Conserved
		17: 26885656 - 26885684	2785 - 2813	Not Conserved
		17: 26885770 - 26885798	2899 - 2927	Not Conserved
		17: 26886259 - 26886287	3388 - 3416	Not Conserved
3	SPN (Leukosialin precursor (Leukocyte	16: 29585767 - 29585795	2013 - 2041	Not Conserved
		16: 29586434 - 29586462	2680 - 2708	Not Conserved

	sialoglycoprotein) (Sialophorin)	16: 29587132 - 29587160	3378 - 3406	Not Conserved
	(Galactoglycoprotein) (GALGP)	16: 29588144 - 29588172	4390 - 4418	Not Conserved
	(CD43 antigen).	16: 29589017 - 29589045	5263 - 5291	Not Conserved
		16: 29589155 - 29589183	5401 - 5429	Not Conserved
4	ZNF529 (Zinc finger protein 529.)	19: 41728138 - 41728166	1441 - 1469	Not Conserved
		19: 41728004 - 41728032	1575 - 1603	Not Conserved
		19: 41727831 - 41727859	1748 - 1776	Not Conserved
		19: 41727702 - 41727730	1877 - 1905	Not Conserved
5	ZC3H6 (Zinc finger CCCH domain-containing protein 6.	2: 112810108 - 112810136	3571 - 3599	dasNov1
		2: 112811625 - 112811653	5088 - 5116	Not Conserved
		2: 112811925 - 112811953	5388 - 5416	Not Conserved
		2: 112812117 - 112812145	5580 - 5608	Not Conserved
		2: 112812253 - 112812281	5716 - 5744	Not Conserved
		2: 112813197 - 112813225	6660 - 6688	Not Conserved
6	ZNF587 (Zinc finger protein 587.)	19: 63063703 - 63063731	382 - 410	Not Conserved
		19: 63063768 - 63063796	447 - 475	Not Conserved
		19: 63064631 - 63064659	1310 - 1338	Not Conserved
		19: 63065070 - 63065098	1749 - 1777	Not Conserved
		19: 63066551 - 63066579	3230 - 3258	Not Conserved
		19: 63067634 - 63067662	4313 - 4341	Not Conserved
7	ANGEL2 (Protein angel homolog 2.)	1: 211234115 - 211234143	862 - 890	Not Conserved
		1: 211233364 - 211233392	1613 - 1641	oryCun1, canFam2
		1: 211233143 - 211233171	1834 - 1862	Not Conserved
		1: 211233006 - 211233034	1971 - 1999	Not Conserved
8	NF2 (Merlin (Moesin-ezrin-radixin-like protein) (Neurofibromin-2) (Schwannomin) (Schwannomerlin)	22: 28421224 - 28421252	781 - 809	Not Conserved
		22: 28422186 - 28422214	1743 - 1771	Not Conserved
		22: 28422286 - 28422314	1843 - 1871	Not Conserved
		22: 28422929 - 28422957	2486 - 2514	Not Conserved
		22: 28423813 - 28423841	3370 - 3398	Not Conserved
9	KIAA1217 (Sickle tail protein homolog)	10: 24874195 - 24874223	141 - 169	rn4, mm8, bosTau2, canFam2, dasNov1, echTel1, monDom4, galGal2, xenTro1
		10: 24876216 - 24876244	2162 - 2190	rn4, mm8, bosTau2, canFam2, loxAfr1, echTel1, monDom4, galGal2, tetNig1
		10: 24876389 - 24876417	2335 - 2363	rn4, mm8, bosTau2, canFam2, loxAfr1, echTel1, monDom4, galGal2
		10: 24874144 - 24874172	90 - 118	rn4, mm8, bosTau2, canFam2, echTel1, monDom4, galGal2
		10: 24875849 - 24875877	1795 - 1823	Not Conserved
10	SLC7A2 (Low-affinity cationic amino acid transporter 2 (CAT-2) (CAT2)	8: 17468612 - 17468640	1682 - 1710	Not Conserved
		8: 17468959 - 17468987	2029 - 2057	Not Conserved
		8: 17470445 - 17470473	3515 - 3543	Not Conserved
		8: 17470592 - 17470620	3662 - 3690	Not Conserved

11	ZFP14 (Zinc finger protein 14 homolog (Zfp-14) (Zinc finger protein 531).	19: 41522534 - 41522562	403 - 431	Not Conserved
		19: 41522286 - 41522314	651 - 67	Not Conserved
		19: 41521162 - 41521190	1775 - 1803	Not Conserved
		19: 41520886 - 41520914	2051 - 2079	Not Conserved
		19: 41519135 - 41519163	3802 - 3830	Not Conserved
12	LONRF2 (LON peptidase N-terminal domain and RING finger protein 2 (RING finger protein 192) (Neuroblastoma apoptosis-related protease).	2: 100267134 - 100267162	29 - 57	Not Conserved
		2: 100264591 - 100264619	2572 - 2600	Not Conserved
		2: 100263404 - 10026343	3759 - 3787	Not Conserved
		2: 100260564 - 100260592	6599 - 6627	Not Conserved
		2: 100257882 - 100257910	9281 - 9309	Not Conserved
13	C20orf142 (Uncharacterized protein C20orf142 precursor.)	20: 42368233 - 42368261	417 - 445	Not Conserved
		20: 42367745 - 42367773	905 - 933	Not Conserved
		20: 42367611 - 42367639	1039 - 1067	Not Conserved
		20: 42365958 - 42365986	2692 - 2720	Not Conserved
14	ZNF284 (Zinc finger protein 284.)	19: 49284018 - 49284046	764 - 792	Not Conserved
		19: 49284161 - 49284189	907 - 935	Not Conserved
		19: 49284840 - 49284868	1586 - 1614	Not Conserved
		19: 49284933 - 49284961	1679 - 1707	Not Conserved
		19: 49285240 - 49285268	1986 - 2014	Not Conserved
15	MDN1 (Midasin (MIDAS-containing protein).	6: 90459275 - 90459303	6443 - 6471	rn4, mm8, oryCun1, bosTau2, dasNov1, loxAfr1, echTel1, monDom4, galGal2, xenTro1, tetNig1
		6: 90510050 - 90510078	564 - 592	rn4, mm8, oryCun1, bosTau2, canFam2, dasNov1, echTel1, monDom4, xenTro1
		6: 90499093 - 90499121	1123 - 1151	rn4, mm8, oryCun1, canFam2, dasNov1, loxAfr1, echTel1, monDom4
		6: 90511937 - 90511965	317 - 345	bosTau2, loxAfr1, echTel1
		6: 90442601 - 90442629	8820 - 8848	canFam2
16	ZFTB8 (zinc finger and BTB domain containing 8)	1: 32838716 - 32838744	108 - 136	rn4, mm8, oryCun1, bosTau2, canFam2, dasNov1, loxAfr1, echTel1, monDom4
		1: 32840994 - 32841022	2386 - 2414	rn4, mm8, oryCun1, bosTau2, canFam2, dasNov1, loxAfr1, echTel1, monDom4, galGal2, xenTro1
		1: 32841438 - 32841466	2830 - 2858	Not Conserved
		1: 32841571 - 32841599	2963 - 2991	Not Conserved
		1: 32843593 - 32843621	4985 - 5013	Not Conserved
		1: 32843729 - 32843757	5121 - 5149	Not Conserved
17	SGCB (Beta-sarcoglycan (Beta-SG) (43 kDa dystrophin-associated glycoprotein) (43DAG) (A3b).	4: 52584823 - 52584851	28 - 56	oryCun1, bosTau2, dasNov1, loxAfr1, echTel1, monDom4, galGal2
		4: 52584017 - 52584045	834 - 862	Not Conserved
		4: 52583591 - 52583619	1260 - 1288	Not Conserved
		4: 52582449 - 52582477	2402 - 2430	Not Conserved
18	CHST6 (Carbohydrate sulfotransferase 6 (EC	16: 74068297 - 74068325	1714 - 1742	Not Conserved
		16: 74067685 - 74067713	2326 - 2354	Not Conserved

	2.8.2.-) (N-acetylglucosamine 6-O-sulfotransferase 5) (GlcNAc6ST-5) (Corneal N-acetylglucosamine-6-O-sulfotransferase) (C-GlcNAc6ST) (hCGn6ST) (Galactose/N-acetylglucosamine/N-acetylglucosamine 6-O-sulfotransferase 4)	16: 74066483 - 74066511	3528 - 3556	Not Conserved
		16: 74065305 - 74065333	4706 - 4734	Not Conserved
19	MTF1 (Metal-regulatory transcription factor 1) (Transcription factor MTF-1) (MRE-binding transcription factor).	1: 38051264 - 38051292	2102 - 2130	mm4, mm8, oryCun1, bosTau2, canFam2, dasNov1, loxAfr1, echTel1, monDom4, galGal2, xenTro1
		1: 38053344 - 38053372	22 - 50	Not Conserved
		1: 38051950 - 38051978	1416 - 1444	Not Conserved
		1: 38051408 - 38051436	1958 - 1986	Not Conserved
		1: 38049191 - 38049219	4175 - 4203	Not Conserved
20	AZU1 (Azurocidin precursor (Cationic antimicrobial protein CAP37) (Heparin-binding protein) (HBP).	19: 782201 - 782229	214 - 242	Not Conserved
		19: 782333 - 782361	346 - 374	Not Conserved
		19: 782421 - 782449	434 - 462	Not Conserved
		19: 782452 - 782480	465 - 493	Not Conserved
21	ZNF264 (Zinc finger protein 264.)	19: 62417788 - 62417816	1626 - 1654	Not Conserved
		19: 62419368 - 62419396	3206 - 3234	Not Conserved
		19: 62423944 - 62423972	7782 - 7810	Not Conserved
		19: 62424079 - 62424107	7917 - 7945	Not Conserved
		19: 62424484 - 62424512	8322 - 8350	Not Conserved
22	PHACTR4 (Phosphatase and actin regulator 4.)	1: 28696509 - 28696537	860 - 888	Not Conserved
		1: 28697883 - 28697911	2234 - 2262	Not Conserved
		1: 28698355 - 28698383	2706 - 2734	Not Conserved
		1: 28698489 - 28698517	2840 - 2868	Not Conserved
23	GFOD1 (Glucose-fructose oxidoreductase domain-containing protein 1 precursor (EC 1.-.-.-)).	6: 13466492 - 13466520	6433 - 6461	mm4, mm8, oryCun1, bosTau2, canFam2, loxAfr1, echTel1
		6: 13470408 - 13470436	2517 - 2545	Not Conserved
		6: 13470020 - 13470048	2905 - 2933	Not Conserved
		6: 13467894 - 13467922	5031 - 5059	Not Conserved
		6: 13467424 - 13467452	5501 - 5529	Not Conserved
24	C1orf69 (Uncharacterized protein C1orf69)	1: 226431477 - 226431505	1639 - 1667	Not Conserved
		1: 226432018 - 226432046	2180 - 2208	Not Conserved
		1: 226432741 - 226432769	2903 - 2931	Not Conserved
		1: 226434564 - 226434592	4726 - 4754	Not Conserved
25	ZNF555 (Zinc finger protein 555.)	19: 2806385 - 2806413	1434 - 1462	Not Conserved
		19: 2807323 - 2807351	2372 - 2400	Not Conserved
		19: 2807458 - 2807486	2507 - 2535	Not Conserved
26	RIPK5 (Receptor-interacting serine/threonine-protein kinase 5 (EC 2.7.11.1) (Dusty protein kinase) (Dusty PK) (RIP-homologous kinase) (Sugen	1: 203382610 - 203382638	670 - 698	Not Conserved
		1: 203381433 - 203381461	1847 - 1875	Not Conserved
		1: 203379297 - 203379325	3983 - 4011	Not Conserved
		1: 203379051 - 203379079	4229 - 4257	Not Conserved

	kinase 496).			
27	FLJ45803 (CDNA FLJ45803 fis)	11: 110671723 - 110671751	287 - 315	Not Conserved
		11: 110671509 - 110671537	501 - 529	Not Conserved
		11: 110671376 - 110671404	634 - 662	Not Conserved
		11: 110670401 - 110670429	1609 - 163	Not Conserved
28	TNRC6B (Trinucleotide repeat-containing 6B protein.)	22: 39049792 - 39049820	600 - 628	rn4, mm8, oryCun1, canFam2, dasNov1, loxAfr1, echTel1, monDom4
		22: 39051365 - 39051393	2173 - 2201	Not Conserved
		22: 39058550 - 39058578	9358 - 9386	mm8, oryCun1, loxAfr1
29	C1orf216 (CDNA FLJ39147 fis)	1: 35953768 - 35953796	23 - 51	oryCun1, canFam2
		1: 35953714 - 35953742	77 - 105	Not Conserved
		1: 35953672 - 35953700	119 - 147	Not Conserved
		1: 35953359 - 35953387	432 - 460	dasNov1
30	(CDNA FLJ45156 fis)	1: 159484654 - 159484682	606 - 634	Not Conserved
		1: 159485346 - 159485374	1298 - 1326	Not Conserved
		1: 159485948 - 159485976	1900 - 1928	Not Conserved
		1: 159486082 - 159486110	2034 - 2062	Not Conserved
31	GJA7 (Gap junction alpha-7 protein (Connexin-45) (Cx45)	17: 40235766 - 40235794	1726 - 1754	Not Conserved
		17: 40233290 - 40233318	4202 - 4230	Not Conserved
		17: 40232208 - 40232236	5284 - 5312	Not Conserved
		17: 40231872 - 40231900	5620 - 5648	Not Conserved
32	RSL1D1 (Ribosomal L1 domain-containing protein 1 (Cellular senescence-inhibited gene protein) (Protein PBK1) (CATX-11).	16: 11838477 - 11838505	639 - 667	Not Conserved
		16: 11836693 - 11836721	2423 - 2451	Not Conserved
		16: 11836336 - 11836364	2780 - 2808	Not Conserved
33	ENDOGL1 (Endonuclease G-like 1 (EC 3.1.30.-) (Endo G-like).	3: 38540876 - 38540904	18 - 46	Not Conserved
		3: 38541861 - 38541889	1003 - 1031	Not Conserved
		3: 38542139 - 38542167	1281 - 1309	Not Conserved
34	NTSDC3 (5'-nucleotidase domain containing 3 isoform 1)	12: 102695687 - 102695715	21 - 49	Not Conserved
		12: 102694332 - 102694360	1376 - 1404	Not Conserved
		12: 102692324 - 102692352	3384 - 3412	Not Conserved
35	GMEMB1 (Glucocorticoid modulatory element-binding protein 1 (GMEB-1) (Parvovirus initiation factor p96) (PIF p96) (DNA-binding protein p96PIF)	1: 28914174 - 28914202	301 - 329	Not Conserved
		1: 28914360 - 28914388	487 - 515	oryCun1
		1: 28916544 - 28916572	2671 - 2699	Not Conserved
		1: 28917401 - 28917429	3528 - 3556	Not Conserved
>miRNA6_GA_CONTIG1 (ACUCAAGACUCUUUGUUCU)				
1	NTRK3 (NT-3 growth factor receptor precursor (EC 2.7.10.1) (Neurotrophic tyrosine kinase receptor type 3) (TrkC tyrosine	15: 86216493 - 86216521	4648 - 4676	bosTau2, canFam2
		15: 86212170 - 86212198	8971 - 8999	bosTau2
		15: 86211855 - 86211883	9286 - 9314	Not Conserved
		15: 86211054 - 86211082	10087 - 10115	dasNov1, loxAfr1

	kinase) (GP145-TrkC) (Trk-C).	15: 86206441 - 86206469	14700 - 14728	Not Conserved
2	NUAK family SNF1-like kinase 1 (EC 2.7.11.1) (AMPK-related protein kinase 5).	12: 104981720 - 104981748	2961 - 2989	mm4, mm8, oryCun1, bosTau2, canFam2, dasNov1, loxAfr1
		12: 104983004 - 104983032	1677 - 1705	bosTau2
		12: 104984225 - 104984253	456 - 484	Not Conserved
		12: 104983836 - 104983864	845 - 873	Not Conserved
3	ZNF578 (zinc finger protein 765)	19: 58604982 - 58605010	789 - 817	Not Conserved
		19: 58604353 - 58604381	160 - 188	Not Conserved
		19: 58606114 - 58606142	1921 - 1949	Not Conserved
		19: 58606244 - 58606272	2051 - 2079	Not Conserved
4	PAG1 (Phosphoprotein associated with glycosphingolipid-enriched microdomains 1 (Transmembrane adapter protein PAG) (Csk-binding protein) (Transmembrane phosphoprotein Cbp).	8: 82048478 - 82048506	2827 - 2855	loxAfr1
		8: 82047549 - 82047577	3756 - 3784	oryCun1, loxAfr1
		8: 82043898 - 82043926	7407 - 7435	Not Conserved
		8: 82043625 - 82043653	7680 - 7708	Not Conserved
5	PRDM2 (PR domain zinc finger protein 2 (PR domain-containing protein 2) (Retinoblastoma protein-interacting zinc finger protein) (Zinc finger protein RIZ) (MTE-binding protein) (MTB-ZF) (GATA-3-binding protein G3B).	1: 14022570 - 14022598	351 - 379	Not Conserved
		1: 14022812 - 14022840	593 - 621	bosTau2
		1: 14022855 - 14022883	636 - 664	loxAfr1
		1: 14023828 - 14023856	1609 - 1637	bosTau2, canFam2, loxAfr1
6	SETD7 (Histone-lysine N-methyltransferase)	4: 140651343 - 140651371	895 - 923	Not Conserved
		4: 140649691 - 140649719	2547 - 2575	Not Conserved
		4: 140648972 - 140649000	3266 - 3294	Not Conserved
		4: 140647315 - 140647343	4923 - 4951	mm8, bosTau2, dasNov1

Table 2 explains: name of genes and position on chromosomes with UTRs and gene target information given by conserved information and this prediction analysis mainly completed by DIANA tar with KEGG analysis for metabolomic activity .yellow color highlighter shows the miRNA with nucleotide sequences.

4. Discussion and Conclusion

Human 3UTRs were found via the DIANA tar total 6 miRNA were found to hybridizes with human 3UTR.with the help of KEGG pathway in a genes for various metabolic pathways in human were tracked .this might silence the gene expression of disease like cancer, blood borne disease, , diuretic, anthelmintic activity, cardio protective, anti diabetic, immunomodulatory, and many more disease by gene silencing the query whether the miRNAs are blocking or repressing any of the human genes was solved by hybridizing putative miRNAs with 3UTR of human through DIANA tar web server. Present studies shows that *Gmelina arborea* 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 *the study concludes that this first reporting of the scientific exploration of multiple action potential of Gmelina arborea miRNA with special reference in cancer and blood brone disease models is not only interesting but very encouraging too for medicinal chemists as well as cancer 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.

Acknowledgements

It is my proud privilege to express deepest sense of reverence and heart full thanks DBT, for providing me scientific computational platform

References

- [1] VL Saxena and A Dwivedi ,insilico identification of miRNA and their target prediction from Japanese encephalitis. *Journal of Bioinformatics and Sequence Analysis*, 2013, Vol. 5(2), pp. 25-33.
- [2] Bartel DP , "MicroRNAs: target recognition and regulatory functions, 2009, *Cell* 136(2): 215–33.
- [3] Lelandais-Brière C, Sorin C, Declerck M, Benslimane A, Crespi M, Hartmann C ,Small RNA diversity in plants and its impact in development ,2010,*Current Genomics* 11 (1): 14–23.
- [4] A.S.R. Anjaneyulu, A.Madhusudhana rao, V.Kameswara Rao and L.Ramachandra, Novel hydroxy lignans from the heartwood of *gmelina arborea* Row, *Tetrahedron*, 1977, Volume 33, Issue 1, Pages 133–143,
- [5] Rashmi Rameshwari I, Divya Singhal, Rachit Narang, Apurvi Maheshwari, and T. V. Prasad, IN SILICO PREDICTION OF MIRNA IN *Curcuma longa* AND THEIR ROLE IN HUMAN METABOLOMICS. *International Journal of Advanced Biotechnology and Research*, 2013 Vol 4, Issue 2, pp 253-259.
- [6] Falgueras J, SeqTrim, a high-throughput pipeline for pre-processing any type of sequence read., 2010,*BMC Bioinformatics*,11:38
- [7] Geo Perte, TIGR Gene Indices clustering tools (TGICL), a software system for fast clustering of large EST datasets. *Bioinformatics*. 2002, Vol. 19, issue 5.
- [8] Adams MD, Kelley JM, Gocayne JD, Complementary DNA sequencing: expressed sequence tags and human genome project ,1991,*Science* 252 (2013): 1651–6.