Full Length Research Article

© 2024, Centre for Advanced Research in Agricultural Sciences
Research Journal of Agricultural Sciences
Volume 15; Issue 04 (Jul–Aug 2024); pp 1014–1022

ISSN: 0976-1675 (P)
ISSN: 2249-4538 (E)

# Prediction of miRNA Target Genes in Various Organisms through Computer-Based Bioinformatics Tools

Navneet Sharma*[1], Deepali Aggarwal[2] and Sheelendra Mangal Bhatt[3]

[1-3] School of Life Science and Technology, IIMT University, Meerut - 250 001, Uttar Pradesh, India

Received: 05 Jun 2024; Revised accepted: 01 Aug 2024

## Abstract

The current article discusses miRNAs called 'killer RNAs' and the method of their gene regulation, which are highly specific to target genes. There are various bioinformatics tools that can be used to find out the miRNA. Furthermore, the discovery of miRNAs has opened up new possibilities for therapeutic interventions in diseases such as cancer, where dysregulation of miRNAs has been implicated in tumor development and progression. By manipulating the expression of specific miRNAs, researchers hope to develop novel treatments that can target cancer cells with precision and minimal side effects. Overall, the study of miRNAs and their regulatory functions represents a promising area of research with far-reaching implications for understanding gene expression, disease mechanisms, and potential therapeutic strategies. As technology continues to advance, the field of miRNA research is likely to uncover even more insights into the complex world of gene regulation.

***Key words*:** miRNA, MIRGENE, siRNAs, miRNA target, Complementary, Animal miRNAs

miRNA are small RNAs (21-24 nucleotides) long that regulate gene expression in animals. Like small interfering RNAs (siRNAs), miRNAs get into the post-transcriptional gene silencing pathway, leading either to degradation of the target mRNA or to translational repression. miRNA is responsible for controlling protein expression by binding with mRNA [1]. There are many weak binding sites present in mRNA [2]. The complementary target region for miRNA binding is contained within the gene. miRNAs are also termed 'killer RNAs' that have emerged as an answer for deciphering hitherto unknown ways of gene regulation, which are highly specific to target genes [3]. The miRNAs are encoded by a distinct class of genes, called MIRGENE; the transcripts of which do not encode proteins. The transcribed RNA is partially self-complementary and folds into a characteristic structure that includes imperfect double-stranded regions from which the miRNAs originate [4]. The excision of the mature miRNA is a multi-step reaction that involves trimming the initial precursor several times. However, plant miRNAs only show homology to other plant miRNAs and animal miRNAs to animal miRNAs indicating that this mechanism of gene regulation has evolved separately in animals [5].

*In silico* identification of miRNA targets is possible because animal miRNAs generally exhibit a high degree of complementarity to their target sites [6]. Given the small size of miRNAs, computer-aided target prediction is complicated [7]. To refine the search for miRNA targets, it is important to test and update the known rules for miRNA and target interaction. Since the first discovery of miRNAs*lin-4* and *let-7* by genetic analysis of *C. elegans* developmental timing [8], numerous computational models and tools have been developed to complement biological experiments to design to study the diverse regulatory roles of miRNAs. Computational as well as experimental analysis shows that most of the human protein-coding genes are regulated by miRNAs [9]. The bioinformatics approaches have been useful in understanding the complexity of miRNA genes and their targets [10]. Most computational methods used for miRNA studies can be classified into two broad categories.

Methods do vary depending on whether a method is for animals, bacteria, or viruses since the biology of miRNAs is somewhat different in each case [11]. A miRNA target detection algorithm can be trained using the properties of known miRNA-mRNA duplexes which can then be used for finding new miRNA-mRNA duplexes [12]. In animals, miRNAs are evolutionarily conserved; many miRNAs, such as miR7 and miR18, are conserved from worm to human [13]. It has been shown that miRNAs in *Arabidopsis* evolved from random sequences by inverted duplication events [14].

These miRNAs can be grouped into 42 miRNA families according to their nucleotide sequence similarity. A majority of these miRNA families have also been found in other species, such as rice, maize, and sorghum [15]. Based on conservation, miRNAs are classified into 4 groups: highly, moderately, lowly, and non-conserved. For some miRNAs, for example, miR163 and miR158, the recent discovery of a much larger set of weakly expressed, less conserved miRNAs in animals which intern provided insights into the origin of miRNAs and hypotheses explaining the birth, selection, and death of miRNAs in land animals [16]. Searching of similar miRNA-miRNA relationships by homology-based computer searching is more useful in plants than animals [17].

***Correspondence to:** Navneet Sharma, E-mail: navneet.sh15@gmail.com; Tel: +91 9412207971

**Citation:** Sharma N, Aggarwal D, Bhatt SM. 2024. Prediction of miRNA target genes in various organisms through computer-based bioinformatics tools. *Res. Jr. Agril. Sci.* 15(4): 1014-1022.

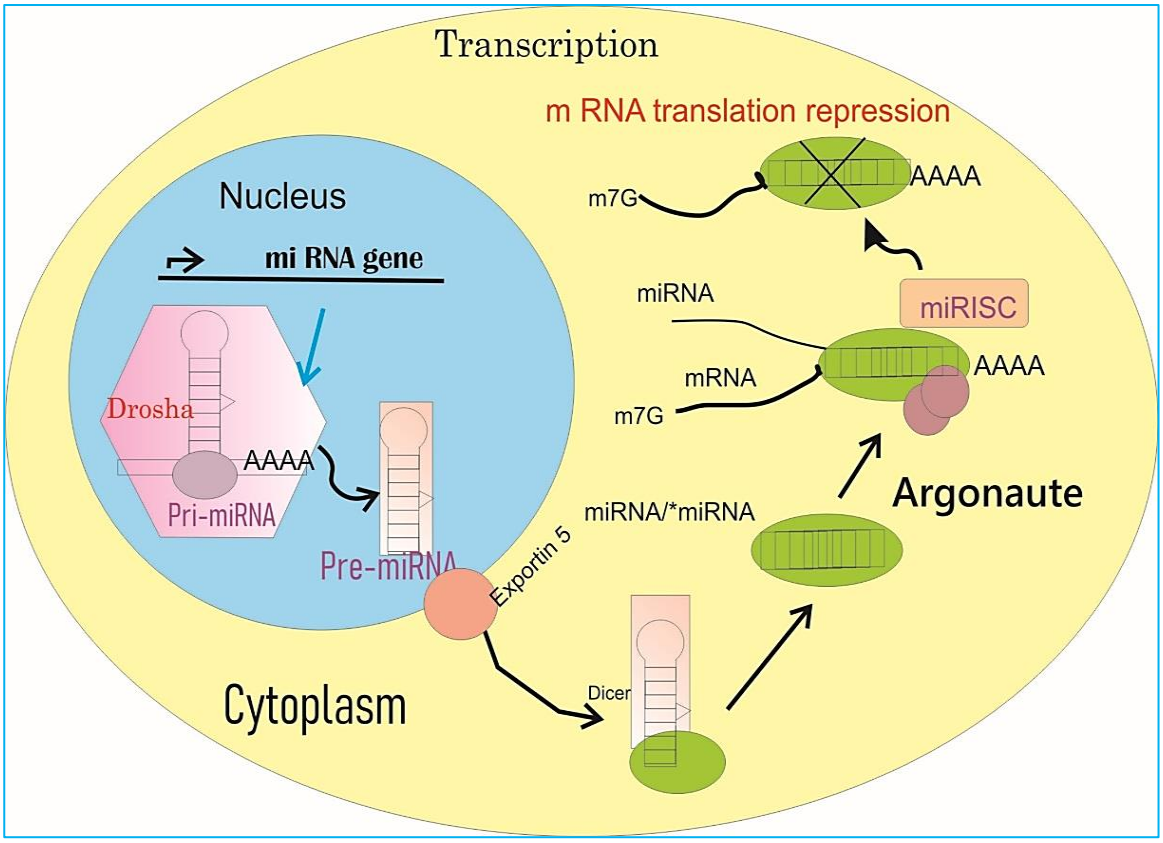



Fig 1 Pre-processing of miRNA (A diagram illustrating the process of miRNA formation and function in animals)

The biogenesis of miRNA initiates in the nucleus, where a large capped and polyadenylated transcript called primary miRNA (pri-miRNA) is transcribed by RNA-polymerase II (RNAPII). The pri-miRNA is then processed by the RNase III endonuclease, Drosha, along with its cofactor, Dgcr8, resulting in the formation of smaller stem-loop structures known as precursor miRNAs (pre-miRNA) [18]. These pre-miRNAs are transported from the nucleus to the cytosol by Exportin 5. In the cytosol, a second RNase III enzyme called Dicer further processes the pre-miRNA, leading to the production of mature miRNA. The mature miRNA then associates with the miRNA-induced silencing complex (miRISC) [19]. Within the miRISC, the seed sequence of the mature miRNA forms Watson-Crick base pairs with complementary sequences primarily located within the 3'-UTRs of mRNA molecules. This base-pairing interaction ultimately results in post-transcriptional gene silencing [20].

MicroRNAs are produced from either their own genes or from introns. It is expected that some non-conserved or lowly conserved miRNAs may be reclassified as more miRNAs are discovered. Conservation is an indicator of miRNA function; conserved miRNAs play an important role in conserved gene regulation, such as leaf and flower morphology or signal transduction [21].

miRNA targets include a large number of transcription factors, suggesting their role in the control of regulatory networks, cellular growth, and development [22]. Conserved miRNAs do not necessarily exhibit the same levels of expression, patterns of the stage of expression in different species. Therefore, sequence and expression divergence in miRNAs between species may affect miRNA accumulation and target regulation leading to developmental changes and phenotypic variation. The conservation of miRNAs and other small RNAs between the conifers and the angiosperms indicates important RNA silencing processes were highly developed in the earliest spermatophytes [23].

A total of 1642 animal miRNAs have been deposited in the miRBase, a database of miRNAs. MicroRNAs (miRNAs) are observed as important l post-transcriptional gene regulators, having diverse impacts on diverse pathological and physiological pathways [24]. Experimental identification of miRNA targets is a difficult and time-consuming process. So, there is still a great demand to understand more clearly the total number of existing miRNAs and their full impact on regulatory pathways [25]. As a consequence, several computational prediction methods have been devised to predict targets for follow-up experimental validation such as the Stacking Binding Matrix (SBM) that uses both the information about the miRNAs as well as experimentally validated target sequences in the search for candidate target sequences [26]. Recently, 79 miRNA candidates were discovered by using wheat EST sequences. Similarly, deep sequencing has proved to be an effective strategy for discovering miRNAs of low-abundance and candidate miRNAs and to recover some of the hits missed by existing methods 454 sequencing can be a valuable tool to assign probability-based scores to them [27].

## MATERIALS AND METHODS

*Identification of miRNA*

*Downloading EST data sets*

Nucleotide sequences (ffn) of *Mycobacterium leprae*, *Mus musculus*, *H. sapiens*, *Drosophila melanogaster*, *C. brigasse*, and *C. elegans* were downloaded from the National Centre for Biotechnology Information (NCBI) (www.ncbi.nlm.nih.gov) in FASTA format.

*miRNA data source*

The miRNAs were taken from miRBase (*http://microrna.sanger.ac.uk*), a growing repository of all published miRNAs of viruses, animals, and plants [28]. The

miRBase Sequence Database is a searchable database of published miRNA sequences and annotations [29]. The miRBase Registry continues to provide gene hunters with unique names for novel miRNA genes prior to the publication of results. Every entry in the miRBase gives a predicted hairpin portion of a miRNA transcript (termed mir in the database), with information on the location and sequence of the mature miRNA sequence (called miR) [30]. Both hairpin and mature sequences are available for BLAST, and entries can also be retrieved by name, keyword, references, and annotation. Starting from 120-200 bases stem-loop precursor to mature 20-21 bases long sequences are available in FASTA format for download [31]. Along with miRNA ID, other information like accession number, chromosomal location, sense or antisense strand, start and end point, etc. are available. These miRNA sequences were downloaded in FASTA format and saved in a text file to find their potential targets [32]. In animals, these target sites are usually present at the 3′ untranslated region (3′UTR) of the mRNA [33]. Genomic sequences of six species were used in this work, including *C. briggsae, C. elegans, D. melanogaster, Mus musculus,* and *Homo sapiens.* Most of these miRNAs were identified or verified by experiments, and others were computationally predicted as their close homologs [34]. In this comparative approach, we tested the algorithms by using a publicly available miRNA knockdown [35]. Prediction of miRNA targets get affected by improper sequence specificity, easy availability of target site, and mRNA structure thermodynamic [36].

*Construction of local database for standalone BLAST*

To identify miRNA targets, features just as seed matching, site conservation, free energy, and site accessibility are checked [37] (Wen et al., 2018). The FASTA formatted text file of Mycobacterium leprae was uploaded to create a local nucleotide database using the "Accessory application" available in BioEdit software [38]. The database was constructed separately for Mycobacterium leprae as a "local nucleotide database" which was used for target identification by standalone blasts [39]. Then we did-

*BLAST Local BLAST*

*Prediction of miRNA targets*

Despite continued progress, bioinformatics prediction of microRNA targets remains a challenge since many software lack accuracy and sensitivity [40]. Several preprocessing steps are taken to draw miRNA candidate sequences from the genome [41]. The reverse complementary miRNAs were used one by one as a query sequence and a standalone blast was performed against the respective "local nucleotide database" of *Mycobacterium leprae* [42].

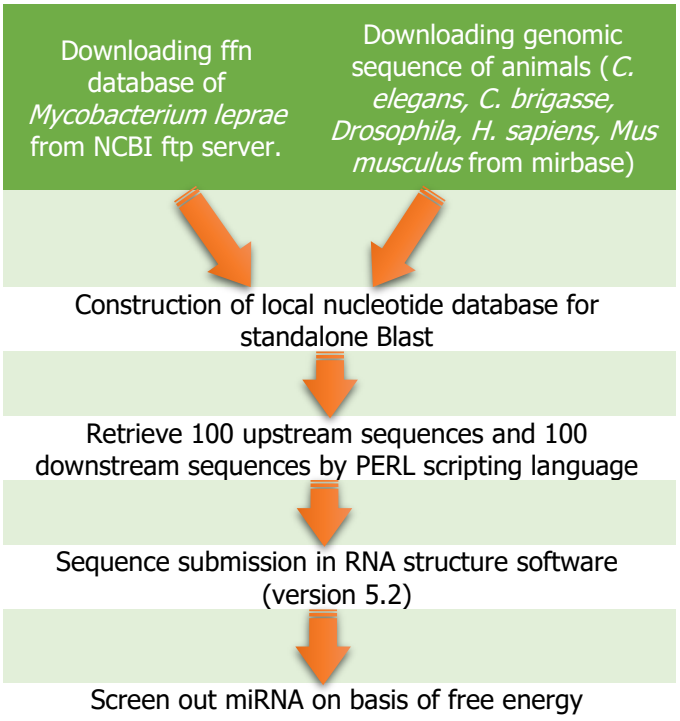
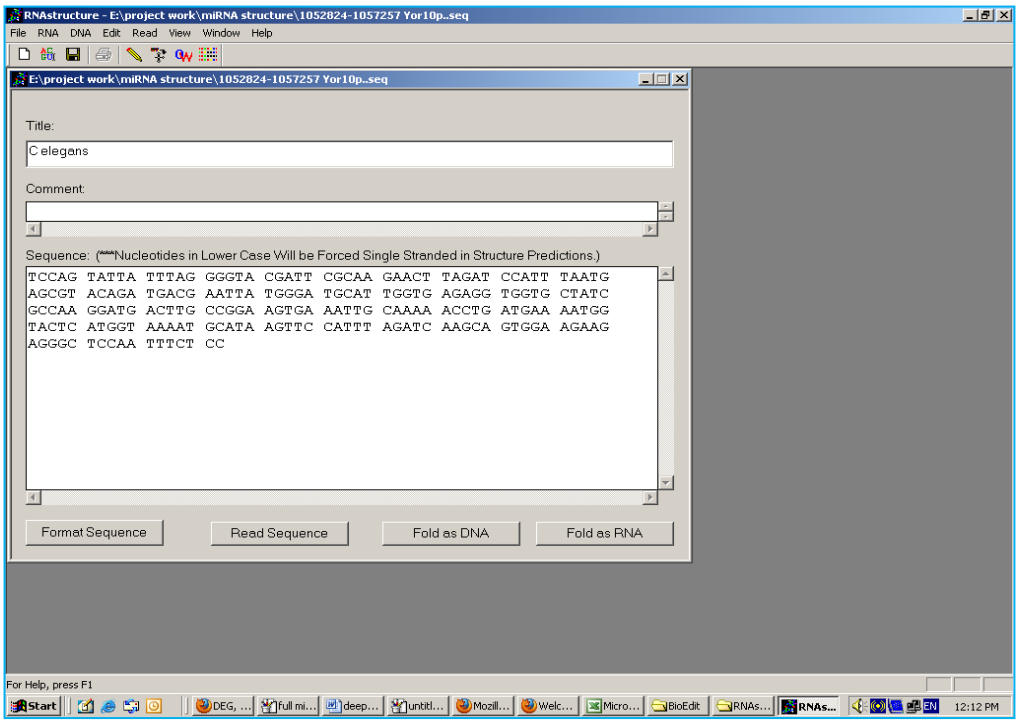



Fig 2 Flow chart depicting screening of miRNA on the basis of free energy




https://rna.urmc.rochester.edu/RNAstructure.html

Fig 3 Sequence submission in RNA structure software

## RESULTS AND DISCUSSION

Number of miRNA target genes predicted in *C. elegans* = 9
Number of miRNA target genes predicted in *C. brigasse* = 5
Number of miRNA target genes predicted in *Drosophila melanogaster* = 4
Number of miRNA target genes predicted in *Homo sapiens* = 19
Number of miRNA target genes predicted in *Mus musculus* = 78
Number of miRNA target genes predicted in *Mycobacterium leprae* = 14

The (Table 1) presents data on predicted microRNA (miRNA) target genes across different species, including *Caenorhabditis elegans*, *C. brigasse*, *Drosophila melanogaster*, *Homo sapiens*, and *Mus musculus*.

Table 1 Prediction of miRNA target genes in various organisms

| Gene id | Name of protein | mi-RNA ID | Length | Bit score | Identity (%) | E-value | Contig sequence |
|---|---|---|---|---|---|---|---|
| | | *Caenorhabditis elegans* | | | | | |
| NC_002677.1\|:c2514415-2513870 | Hypothetical protein ml2113 | cel-miR-2 | 23 | 28.2 | 100 | 0.06 | CAGCTTTGATGTGC |
| NC_002677.1\|:2145165-2146505 | Sugar transport periplasmicbinding protein | cel-miR-60* | 22 | 32.2 | 100 | 0 | ACTGGAAGAGTGCCAT |
| NC_002677.1\|:c2883434-2882646 | Putative cytochrome c-typebiogenesis protein | cel-miR-70 | 23 | 28.2 | 100 | 0.06 | GTCGTTGGTGTTTC |
| NC_002677.1\|:c65112-63319 | Hypothetical protein ml00052 | cel-miR-358* | 23 | 28.2 | 100 | 0.06 | CCTGGCCAGGCATT |
| NC_002677.1\|:218458-219639 | Succinyl-coasynthetase subunit beta | cel-miR-796 | 23 | 28.2 | 100 | 0.06 | ATGTAGTTGAGGTT |
| NC_002677.1\|:659721-660932 | S-denosylmethioninesynthetase | cel-miR-1832 | 21 | 30.2 | 100 | 0.01 | GGCGGAGCGAATCGA |
| NC_002677.1\|:c2532909-2531614 | Type ii citrate synthase | cel-miR-2209c | 23 | 28.2 | 100 | 0.06 | ACCGGTTACACTAC |
| NC_002677.1\|:1339580-1340836 | Udp-n-acetylglucosamine 1-carboxyvinyltransferase | cel-miR-1832b | 23 | 28.2 | 100 | 0.06 | GGCAGAGCGATTCG |
| NC_002677.1\|:1417029-1417634 | Lipoprotein signal peptidase | cel-miR-2217 | 24 | 30.2 | 94 | 0.02 | TGGGAAGTCGGTGTCGATC |
| | | *C. brigasse* | | | | | |
| NC_002677.1\|:c1925737-1924844 | Site-specific tyrosine recombinase xerc | cbr-miR-124 | 21 | 30.2 | 100 | 0.01 | CACGCGGTGAATGCC |
| NC_002677.1\|:c707588-706461 | Transaldolase | cbr-miR-40 | 22 | 28.2 | 100 | 0.05 | GGTGTCAATCAGCT |
| NC_002677.1\|:2457996-2460191 | Malate synthase g | cbr-miR-76 | 20 | 28.2 | 94 | 0.04 | TTCGTTGTTGACGAAGTC |
| NC_002677.1\|:c2804872-2799404 | Polyketide synthase | cbr-miR-235b* | 24 | 28.2 | 100 | 0.06 | GCTCTGGTCATGTG |
| NC_002677.1\|:c488641-486263 | Sugar phosphorylase | crm-miR-786 | 23 | 28.2 | 100 | 0.06 | CTTGCTGAGATTCC |
| | | *Drosophila melanogaster* | | | | | |
| NC_002677.1\|:c3131700-3128833 | Integral membrane protein | dme-miR-306* | 22 | 30.2 | 94 | 0.01 | GGTCGCTCTGTGCCTGTGC |
| NC_002677.1\|:1073832-1075859 | Penicillin-binding protein 2 | dme-miR-985 | 22 | 28.2 | 100 | 0.05 | CCAATGGTCGGGCA |
| NC_002677.1\|:c1939991-1939452 | 16s rrna-processing protein rim | dme-miR-252 | 22 | 28.2 | 100 | 0.05 | AAGTACTAGTGCCG |
| NC_002677.1\|:303740-305263 | Lysyl-trnasynthetase | dme-miR-1012 | 23 | 28.2 | 100 | 0.06 | GTCAAAGATTTTCC |
| | | *H. sapiens* | | | | | |
| NC_002677.1\|:c2769043-2767652 | Hypothetical protein ml2336 | hsa-miR-21 | 21 | 28.2 | 100 | 0.05 | CAGTCGATGGGCTG |
| NC_002677.1\|:1443804-1444400 | Hypothetical protein ml1222 | hsa-miR-16-1* | 22 | 28.2 | 100 | 0.05 | TATTAACTGTGCTG |
| NC_002677.1\|:3090766-3091569 | Hypothetical protein ml2587 | hsa-miR-24 | 22 | 28.2 | 100 | 0.05 | TGGCTCAGTTCAGC |
| NC_002677.1\|:1699648-1700454 | Cdp-diacylglycerolpyrophosphatase | hsa-miR-33a | 22 | 30.2 | 100 | 0.01 | TTTCCACAGTGCATC |
| NC_002677.1\|:1352912-1354093 | Acetyl-coaacetyltransferase | hsa-miR-27b* | 22 | 28.2 | 100 | 0.05 | GCTGATTGGTGAAC |
| NC_002677.1\|:c1925737-1924844 | Site-specific tyrosine recombinase xerc | hsa-miR-124 | 20 | 30.2 | 100 | 0.01 | CACGCGGTGAATGCC |
| NC_002677.1\|:c798689-797292 | Acyl-coa dehydrogenase | hsa-miR-150* | 22 | 32.2 | 100 | 0 | ACAGGCCTGGGGGACA |
| NC_002677.1\|:c3131700-3128833 | Integral membrane protein | hsa-miR-185 | 22 | 28.2 | 100 | 0.05 | AAGGCAGTTCCTGA |
| NC_002677.1\|:c2189475-2188468 | Hypothetical protein ml1808 | hsa-miR-194 | 22 | 28.2 | 100 | 0.05 | CAGCAACTCCATGT |
| NC_002677.1\|:c1805474-1803621 | Hypothetical protein ml1497 | hsa-miR-106b | 21 | 28.2 | 100 | 0.05 | GTGCTGACAGTGCA |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| NC_002677.1\|:c270449-268857 | Hypothetical protein ml0201 | hsa-miR-200a* | 22 | 28.2 | 100 | 0.05 | CCGGACAGTGCTGG |
| NC_002677.1\|:c1858031-1856520 | Hypothetical protein ml1539 | hsa-miR-323-5p | 22 | 28.2 | 100 | 0.05 | GGTCCGTGGCGCGT |
| NC_002677.1\|:188551-189252 | Lipoprotein | hsa-miR-338-3p | 22 | 28.2 | 100 | 0.05 | CAGTGATTTTGTTG |
| NC_002677.1\|:c88719-88246 | Ribonuclease activity regulator | hsa-miR-425 | 23 | 28.2 | 100 | 0.06 | AATGACACGATCAC |
| NC_002677.1\|:2068541-2069857 | Transferase | hsa-miR-425* | 22 | 28.2 | 100 | 0.05 | TGTCGTGTCCGCCC |
| NC_002677.1\|:c2086347-2084623 | Cytochrome c oxidase, polypeptide | hsa-miR-486-5p | 22 | 28.2 | 100 | 0.05 | ACTGAGCTGCCCCG |
| NC_002677.1\|:2021235-2021810 | Hypothetical protein ml1677 | hsa-miR-493 | 22 | 28.2 | 94 | 0.05 | GGTCTACTGTGTACCAGG |
| NC_002677.1\|:2476953-2478869 | Multidrug resistance pump | hsa-miR-497 | 21 | 28.2 | 94 | 0.05 | GCAGCACACTGTTGTTTG |
| NC_002677.1\|:1352912-1354093 | Acetyl-coaacetyltransferas | mmu-miR-27b | 22 | 28.2 | 100 | 0.05 | GCTGATTGGTGAAC |
| | | *Mus musculus* | | | | | |
| NC_002677.1\|:c1925737-1924844 | Site-specific tyrosine recombinase xerc | mmu-miR-124 | 20 | 30.2 | 100 | 0.01 | CACGCGGTGAATGCC |
| NC_002677.1\|:c798689-797292 | Acyl-coa dehydrogenase | mmu-miR-150* | 22 | 28.2 | 100 | 0.05 | ACAGGCCTGGGGGA |
| NC_002677.1\|:c2443504-2442641 | Cell surface protein | mmu-miR-181a | 23 | 28.2 | 100 | 0.06 | CAACGCTGTCGGTG |
| NC_002677.1\|:c3131700-3128833 | Integral membrane protein | mmu-miR-185 | 22 | 28.2 | 100 | 0.05 | AAGGCAGTTCCTGA |
| NC_002677.1\|:3090766-3091569 | Hypothetical protein ml2587 | mmu-miR-24 | 22 | 28.2 | 100 | 0.05 | TGGCTCAGTTCAGC |
| NC_002677.1\|:c2561489-2561097 | Hypothetical protein ml2158 | mmu-miR-297a | 22 | 28.2 | 100 | 0.05 | GTGCATGTGCATGT |
| NC_002677.1\|:c1805474-1803621 | Hypothetical protein ml1497 | mmu-miR-106b | 21 | 28.2 | 100 | 0.05 | GTGCTGACAGTGCA |
| NC_002677.1\|:c1915424-1914867 | Ribosome recycling factor | mmu-miR-30c-2* M | 22 | 28.2 | 100 | 0.05 | GGAGAAGGCTGTTT |
| NC_002677.1\|:2925762-2926499 | Hypothetical protein ml2450 | mmu-miR-148a* | 22 | 28.2 | 100 | 0.05 | GAGACACTCCGACT |
| NC_002677.1\|:c270449-268857 | Hypothetical protein ml0201 | mmu-miR-200a* | 22 | 28.2 | 100 | 0.05 | CCGGACAGTGCTGG |
| NC_002677.1\|:c2769043-2767652 | Hypothetical protein ml2336 | mmu-miR-21* | 22 | 30.2 | 100 | 0.01 | GCAGTCGATGGGCTG |
| NC_002677.1\|:c1858031-1856520 | Hypothetical protein ml1539 | mmu-miR-323-5p | 22 | 28.2 | 100 | 0.05 | GGTCCGTGGCGCGT |
| NC_002677.1\|:c1124822-1123017 | Pyruvate phosphate dikinase | mmu-miR-328* | 22 | 30.2 | 100 | 0.01 | GGCAGGAGGGGCTCA |
| NC_002677.1\|:188551-189252 | Lipoprotein | mmu-miR-338-3p | 22 | 28.2 | 100 | 0.05 | CAGTGATTTTGTTG |
| NC_002677.1\|:1668337-1671246 | Excinucleaseabc subunit A | mmu-miR-223 | 22 | 28.2 | 100 | 0.05 | TTGTCAAATACCCC |
| NC_002677.1\|:1699648-1700454 | Cdp-diacylglycerolpyrophosphatase | mmu-miR-33 | 22 | 30.2 | 100 | 0.01 | TTTCCACAGTGCATC |
| NC_002677.1\|:411592-412068 | 2-c-methyl-d-erythritol | hsa-miR-194* | 22 | 28.2 | 100 | 0.05 | GTGGGGCTGCTGTT |
| NC_002677.1\|:3082923-3084725 | Hypothetical protein ml2582 | mmu-miR-382 | 22 | 28.2 | 100 | 0.05 | CGGACAACACTTTT |
| NC_002677.1\|:c88719-88246 | Ribonuclease activity regulator | mmu-miR-425 | 23 | 28.2 | 100 | 0.06 | AATGACACGATCAC |
| NC_002677.1\|:c2922708-2921383 | Lipoprotein | mmu-miR-431 | 21 | 30.2 | 100 | 0.01 | TGTCTTGCAGGCCGT |
| NC_002677.1\|:918368-920137 | Lipoprotein lpqb | mmu-miR-470 | 23 | 28.2 | 100 | 0.06 | TTCTTGGACTGGCA |
| NC_002677.1\|:c2086347-2084623 | Cytochrome c oxidase, polypeptide i | mmu-miR-486 | 22 | 28.2 | 100 | 0.05 | ACTGAGCTGCCCCG |
| NC_002677.1\|:582395-583444 | Holliday junction dna helicase ruvb | mmu-miR-546 | 16 | 26.3 | 100 | 0.1 | ATGGTGGCACGGA |
| NC_002677.1\|:c902902-901763 | 2-phospho-l-lactate transferase | mmu-miR-543 | 22 | 32.2 | 100 | 0 | ATTCGCGGTGCACTTC |
| NC_002677.1\|:1668337-1671246 | Excinucleaseabc subunit A | mmu-miR-539-5p | 22 | 28.2 | 100 | 0.05 | TTATCCTTGGTGTG |
| NC_002677.1\|:2383320-2387525 | Peptide synthase | mmu-miR-494 | 22 | 28.2 | 100 | 0.05 | TTGTCCGTGTTGTC |
| NC_002677.1\|:182173-188484 | Polyketide synthase | mmu-miR-592* | 22 | 34.2 | 100 | 0.06 | CATCACGTGGTGACGCA |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| NC_002677.1\|:3058964-3060328 | Acyl-carrier-protein | mmu-miR-551b* | 23 | 28.2 | 100 | 0.06 | GCTTGGGTGAGACC |
| NC_002677.1\|:2081-3280 | Dna polymerase iii subunit beta | mmu-miR-668 | 24 | 28.2 | 100 | 0.06 | TCGGCTCGGCCCAC |
| NC_002677.1\|:c1187205-1186213 | Integral membrane protein | mmu-miR-665* | 26 | 30.2 | 100 | 0.02 | TGCCTCTATCCAGGA |
| NC_002677.1\|:c1734472-1733639 | Glycosyltransferase | mmu-miR-667* | 25 | 28.2 | 100 | 0.07 | CGGTGCTGGTGGAG |
| NC_002677.1\|:c1466688-1465612 | Hypothetical protein ml1232 | mmu-miR-667 | 23 | 28.2 | 100 | 0.06 | CTGCCACCCAGCCC |
| NC_002677.1\|:c2214166-2212448 | Ppe-family protein | mmu-miR-761 | 22 | 28.2 | 100 | 0.05 | CAGCAGGGTGAAAC |
| NC_002677.1\|:1694884-1696296 | Gininosuccinatelyase | mmu-miR-344d-1* | 24 | 30.2 | 100 | 0.02 | CTGCTGGCTATACAC |
| NC_002677.1\|:194221-200571 | Mycocerosic synthase | mmu-miR-666-3p | 22 | 28.2 | 100 | 0.05 | CGTGATCGCCTGCT |
| NC_002677.1\|:627741-628964 | Chorismate synthase | mmu-miR-760-5p | 23 | 32.2 | 100 | 0 | GGCCACCAGAGCCCGG |
| NC_002677.1\|:c425727-424867 | Abc-transporter | mmu-miR-674 | 22 | 28.2 | 100 | 0.05 | TGAGATGGGAGTGG |
| NC_002677.1\|:1630435-1631208 | Pseudouridine synthase | mmu-miR-677 M | 22 | 28.2 | 100 | 0.05 | GATGATTAGCTTCT |
| NC_002677.1\|:2476953-2478869 | Multidrug resistance pump | mmu-miR-497 | 22 | 28.2 | 94 | 0.05 | GCAGCACACTGTTGTTTG |
| NC_002677.1\|:c2561489-2561097 | Hypothetical protein ml2158 | mmu-miR-669b | 22 | 28.2 | 100 | 0.05 | GTGCATGTGCATGT |
| NC_002677.1\|:118819-120768 | Hypothetical protein ml0096 | mmu-miR-501-5p | 22 | 28.2 | 100 | 0.05 | GTCCCTGGGTGAAA |
| NC_002677.1\|:1086005-1087030 | Ftsq-family protein | mmu-miR-676 | 21 | 28.2 | 100 | 0.05 | CGTCCTGAGGTTGT |
| NC_002677.1\|:c2372198-2369802 | Isocitratelyase | mmu-miR-877 | 20 | 28.2 | 100 | 0.04 | GTAGAGGAGATGGC |
| NC_002677.1\|:c147327-144115 | Arabinosyltransferase | mmu-miR-598 | 23 | 30.2 | 94 | 0.01 | CGGTGATGCCGATGATGCG |
| NC_002677.1\|:386629-388932 | Serine-threonine protein kinase | mmu-miR-598 | 22 | 30.2 | 100 | 0.01 | GTCATCGTCGTCATC |
| \|NC_002677.1\|:c2561489-2561097 | Hypothetical protein ml2158 | mmu-miR-669f-5p | 24 | 28.2 | 100 | 0.06 | GTGCATGTGCATGT |
| NC_002677.1\|:c2561489-2561097 | Hypothetical protein ml2158 | mmu-miR-466j | 23 | 28.2 | 100 | 0.06 | GTGCATGTGCATGT |
| NC_002677.1\|:c2561489-2561097 | Hypothetical protein ml2158 | mmu-miR-466j | 23 | 28.2 | 100 | 0.06 | GTGCATGTGCATGT |
| NC_002677.1\|:1548392-1549180 | Hypothetical protein ml1300 | mmu-miR-1197* | 21 | 34.2 | 100 | 0.06 | GGTTGACCATGGTGTGT |
| NC_002677.1\|:c212814-211150 | Glucose-6-phosphate isomerase | mmu-miR-1897-3p | 22 | 30.2 | 100 | 0.01 | CTCGTTCTGTCCGGT |
| NC_002677.1\|:503217-504401 | Hypothetical protein ml0405 | mmu-miR-1903 | 22 | 28.2 | 100 | 0.05 | CTTCTTCTTCTTCC |
| NC_002677.1\|:c899638-897971 | Dnamethylase | mmu-miR-1899 | 22 | 28.2 | 100 | 0.05 | GCGATGGCCGAATC |
| NC_002677.1\|:124599-126506 | Long-chain-fatty-acid--coa ligase | mmu-miR-1906 | 22 | 32.2 | 100 | 0 | GCAGCAGCCTGAGGCA |
| NC_002677.1\|:3204710-3209539 | Cation transport atpase | mmu-miR-1893 | 22 | 30.2 | 100 | 0.01 | CGGGCGCTGGACGCC |
| NC_002677.1\|:c1194348-1193875 | Hypothetical protein ml1027 | mmu-miR-1945 | 22 | 28.2 | 100 | 0.05 | GCGGGTACTGTCGG |
| NC_002677.1\|:c3264167-3263112 | Cell division protein | mmu-miR-1306 | 21 | 30.2 | 100 | 0.01 | TTGGCTCTGGTGGTG |
| NC_002677.1\|:3065268-3065441 | Hypothetical protein | mmu-miR-1946a | 27 | 30.2 | 100 | 0.02 | CCGGGCAGTGGTGGC |
| \|NC_002677.1\|:c2561489-2561097 | Hypothetical protein ml2158 | mmu-miR-669m-5p | 23 | 28.2 | 100 | 0.06 | GTGCATGTGCATGT |
| NC_002677.1\|:c2811474-2804869 | Polyketide synthase | mmu-miR-1955-3p | 21 | 28.2 | 100 | 0.05 | GAGCATTGCATGCT |
| NC_002677.1\|:c2475105-2472769 | Preproteintranslocase subunit seca | mmu-miR-1966 | 25 | 28.2 | 100 | 0.07 | GGAGCTGGCTCAGG |
| C_002677.1\|:3065268-3065441 | Hypothetical protein | mmu-miR-1946b | 26 | 30.2 | 100 | 0.02 | CCGGGCAGTGGTGGC |
| NC_002677.1\|:3065557-3069774 | Integral membrane protein | mmu-miR-2861 | 19 | 28.2 | 100 | 0.04 | GGCCTGGCGGCGGG |
| NC_002677.1\|:2617849-2618970 | Hypothetical protein ml2203 | mmu-miR-3057-5p | 24 | 28.2 | 100 | 0.06 | GCTGAGATTCTGCG |
| NC_002677.1\|:c526941-526294 | Oligoribonuclease | mmu-miR-3057-5p | 24 | 28.2 | 100 | 0.06 | GGAGCTGAGATTCT |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| NC_002677.1\|:c587947-584381 | Acyl-coasynthetase | mmu-miR-3064-5p | 22 | 28.2 | 100 | 0.05 | TGGCTGTTGTGGTG |
| NC_002677.1\|:188551-189252 | Lipoprotein | mmu-miR-3065 | 22 | 28.2 | 100 | 0.05 | |
| NC_002677.1\|:c2039801-2038215 | D-3-phosphoglycerate dehydrogenase | mmu-miR-3067* | 23 | 30.2 | 100 | 0.01 | GGCTGCCCTGGGAGA |
| NC_002677.1\|:c2861605-2861300 | Hypothetical protein ml2390 | mmu-miR-3068* | 22 | 28.2 | 100 | 0.05 | GGTGAATTGCAGTA |
| NC_002677.1\|:c3170139-3169324 | Hypothetical protein ml2649 | mmu-miR-3073-3p | 22 | 28.2 | 100 | 0.05 | TGTCCACTGTGACC |
| NC_002677.1\|:3090766-3091569 | Hypothetical protein ml2587 | mmu-miR-3074-5p | 22 | 30.2 | 100 | 0.01 | GCTGAACTGAGCCAG |
| NC_002677.1\|:2303405-2304298 | Dnaglycosylase | mmu-miR-3080-3p | 22 | 28.2 | 100 | 0.05 | CGGGCAAAGCGCTT |
| NC_002677.1\|:700214-703018 | Phosphoenolpyruvate carboxylase | mmu-miR-3084 | 21 | 28.2 | 100 | 0.05 | GCCAGTCTCCTTCA |
| NC_002677.1\|:219659-220561 | Succinyl-coasynthetase subunit | mmu-miR-3086-5p | 20 | 30.2 | 100 | 0.01 | AGATTGTAGGCCCAT |
| NC_002677.1\|:c2561489-2561097 | Hypothetical protein ml2158 | mmu-miR-466m-5p | 23 | 28.2 | 100 | 0.06 | GTGCATGTGCATGT |
| NC_002677.1\|:1877242-1878684 | Prolyl-trnasynthetase | mmu-miR-3093-5p | 22 | 28.2 | 100 | 0.05 | ACCCCGCGGAGCTC |
| NC_002677.1\|:c2229420-2228980 | 50s ribosomal protein l15 | mmu-miR-3105-5p | 22 | 28.2 | 100 | 0.05 | AAGCCCGTAAGCAG |
| NC_002677.1\|:1519218-1519658 | Hypothetical protein ml1276 | mmu-miR-3105-5p | 22 | 28.2 | 100 | 0.05 | AGCCCGTAAGCAGC |
| NC_002677.1\|:c2086347-2084623 | Cytochrome c oxidase, polypept | mmu-miR-3107 | 22 | 28.2 | 100 | 0.05 | ACTGAGCTGCCCCG |

Overall, the (Table 1) provides a comprehensive view of miRNA-target gene interactions across various species, highlighting high-confidence predictions with 100% identity and significant bit scores. The E-values indicate the statistical significance of these interactions, with lower values representing higher confidence.

Hence prediction of miRNA drug target genes involves availability of latest software's and well-versed data sets. Source of miRNA, miRNA ID, number of mismatch sequences, length of contig sequence, Bit score, E value was easily determined by above methods. In this study we have analyzed the genomic nucleotide sequence of six species *Mus musculus*, *Homo sapiens*, *Drosophila melanogaster*, *C. brigasse*, *Caenorhabditis elegans* and *Mycobacterium leprae*. Form the above study it is clear that number of miRNA predicted genes are found highest in Mus musculus than H sapiens and also proteins coded by them were also predicted. 42% proteins of *Mycobacterium leprae* are homologous with Human genome while rest are non-homologous. Non-homologous sequences are selected as target genes for miRNA. Protein encoded by these drug target gene searched by ExPaSy proteome server to find out their functions. 51% genes were involved in endoplasmic reticulum while 36% used in periplasmic process. However, 13% (34) genes are helpful in the synthesis of cell membrane and identified as target, which can be targeted in future. The sequences work as a magic bullet for anti-bacterial and antifungal drug. A total of 115 miRNA predicted drug target genes are identified by computational method.

Table 2 miRNA of M leprae NM: Number of mismatch; LM: Length of mature miRNA; LP: Length of precursor; MFEI: Minimal folding free energy index

| S. No. | New miRNA | Source | miRNA sequence | NM/nU | LM/nU | LP/nU | MFEI |
|---|---|---|---|---|---|---|---|
| 1 | micro-MIR 59 | GSS | GCAGUCGAUGGGCUG | 0 | 15 | 215 | -65.4 |
| 2 | micro-MIR 8 | GSS | GGCGGAGCGAAUCGA | 0 | 15 | 215 | -92.1 |
| 3 | micro-MIR 79 | GSS | CGGUGCUGGUGGAG | 0 | 14 | 214 | -83.5 |
| 4 | micro-MIR 39 | GSS | CCGGACAGUGCUGG | 0 | 14 | 214 | -105.1 |
| 5 | micro-MIR 58 | GSS | CCGGACAGUGCUGG | 0 | 14 | 214 | -105.1 |
| 6 | micro-MIR 11 | GSS | UGGGCAGUCGGUGUCGAUC | 6 | 18 | 218 | -77.6 |
| 7 | micro-MIR 31 | GSS | UGGCUCAGUUCAGC | 0 | 14 | 214 | -88.4 |
| 8 | micro-MIR 70 | GSS | ACUGAGCUGCCCCG | 0 | 14 | 214 | -83.4 |
| 9 | micro-MIR 23 | GSS | CCAAUGGUCGGGCA | 0 | 14 | 214 | -69.9 |
| 10 | micro-MIR 25 | GSS | GUCAAAGAUUUUCC | 0 | 14 | 214 | -62 |
| 11 | micro-MIR 82 | GSS | CUGCUGGCUAUACAC | 0 | 15 | 215 | -95 |
| 12 | micro-MIR 43 | GSS | UGUCGUGUCCGCCC | 0 | 15 | 215 | -82.6 |
| 13 | micro-MIR 22 | GSS | GGUCACUCUGUGCCUGUGC | 6 | 18 | 218 | -10.8 |
| 14 | micro-MIR 24 | GSS | AAGUACUAGUGCCG | 0 | 14 | 214 | -67.8 |

## CONCLUSION

Hence from the above study it has been concluded that prediction of miRNA target genes was easily achieved through computer-based software and web-based bioinformatics tools. For prediction of miRNA target genes, we have to identify certain features such as miRNA binding sites and target down regulation by various methods. All the web-based bioinformatics tools, databases and software provides valuable miRNA target genes in various species of rat, dog and chicken also. By use of PERL scripting language and sequence submission in RNA structure software, 115 miRNA drug target genes have been predicted in various organisms like *Mus musculus*, *Homo sapiens*, *Drosophila melanogaster*, *C. brigasse* and *Caenorhabditis elegans* and 14 miRNA targets have been identified in Mycobacterium leprae which appeared to be related to the development, growth, metabolism and other physiological processes such as stress response of *Mycobacterium leprae*. The combined usage of nucleotide databases of various organisms and their sequence submission in RNA mediated software will increases the true prediction and decreases the availability of false positive results. The findings from this study will contribute to further researches in miRNAs function and regulatory mechanisms in Mus musculus which will be very beneficial to study some essential features in Mus musculus and also studies the confirmation of interaction between predicted miRNA-miRNA duplexes of various model and novel organisms. These findings will lead to open new avenues in computational research and miRNA will act as potential therapeutic targets for human disease.

## LITERATURE CITED


1. Ferri C, Di Biase A, Bocchetti M, Zappavigna S, Wagner S, Le Vu P, Luce A, Cossu AM, Vadakekolathu J, Miles A, Boocock DJ. 2022. MiR-423-5p prevents MALAT1-mediated proliferation and metastasis in prostate cancer. *Journal of Experimental and Clinical Cancer Research* 41(1): 20.
2. Garcia DM, Baek D, Shin C, Bell GW, Grimson A, Bartel DP. 2011. Weak seed-pairing stability and high target-site abundance decrease the proficiency of lsy-6 and other microRNAs. *Nature Structural and Molecular Biology* 18(10): 1139-1146.
3. Van Meter EN, Onyango JA, Teske KA. 2020. A review of currently identified small molecule modulators of microRNA function. *European Journal of Medicinal Chemistry* 188: 112008.
4. Baek D, Villén J, Shin C, Camargo FD, Gygi SP, Bartel DP. 2008. The impact of microRNAs on protein output. *Nature* 455(7209): 64-71.
5. Broughton JP, Pasquinelli AE. 2016. A tale of two sequences: microRNA-target chimeric reads. *Genetics Selection Evolution* 48: 1-7.
6. Jones-Rhoades MW, Bartel DP. 2004. Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Molecular Cell* 14(6): 787-799.
7. Robins H, Li Y, Padgett RW. 2005. Incorporating structure to predict microRNA targets. *Proceedings of the National Academy of Sciences* 102(11): 4006-4009.
8. Reinhart BJ, Weinstein EG, Rhoades MW, Bartel B, Bartel DP. 2002. MicroRNAs in plants. *Genes and Development* 16(13): 1616-1626.
9. Fan Y, Xia J. 2018. miRNet—functional analysis and visual exploration of miRNA–target interactions in a network context. *Computational Cell Biology: Methods and Protocols*. pp 215-233.
10. Ae P. 2000. Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature* 408: 86-89.
11. Dayal S, Chaubey D, Joshi DC, Ranmale S, Pillai B. 2024. Noncoding RNAs: emerging regulators of behavioral complexity. *Wiley Interdisciplinary Reviews: RNA* 15(3): p.e1847.
12. Akhtar MM, Micolucci L, Islam MS, Olivieri F, Procopio AD. 2016. Bioinformatic tools for microRNA dissection. *Nucleic Acids Research* 44(1): 24-44.
13. Weber MJ. 2005. New human and mouse microRNA genes found by homology search. *The FEBS Journal* 272(1): 59-73.
14. De Felippes FF, Schneeberger K, Dezulian T, Huson DH, Weigel D. 2008. Evolution of Arabidopsis thaliana microRNAs from random sequences. *RNA* 14(12): 2455-2459.
15. Agarwal V, Bell GW, Nam JW, Bartel DP. 2015. Predicting effective microRNA target sites in mammalian mRNAs. *E-life* 4: e05005.
16. Ekimler S, Sahin K. 2014. Computational methods for microRNA target prediction. *Genes* 5(3): 671-683.
17. Ajila V, Colley L, Ste-Croix DT, Nissan N, Cober ER, Mimee B, Samanfar B, Green JR. 2023. Species-specific microRNA discovery and target prediction in the soybean cyst nematode. *Scientific Reports* 13(1): 17657.
18. Liu J, Zhou F, Guan Y, Meng F, Zhao Z, Su Q, Bao W, Wang X, Zhao J, Huo Z, Zhang L. 2022. The biogenesis of miRNAs and their role in the development of amyotrophic lateral sclerosis. *Cells* 11(3): 572.
19. Li M, Yu B. 2021. Recent advances in the regulation of plant miRNA biogenesis. *RNA Biology* 18(12): 2087-2096.
20. Hajarnis, S., Lakhia, R. and Patel, V., 2015. MicroRNAs and polycystic kidney disease. *Exon Publications*, pp.313-334.
21. Yu T, Xu N, Haque N, Gao C, Huang W, Huang Z. 2020. Popular computational tools used for miRNA prediction and their future development prospects. *Interdisciplinary Sciences: Computational Life Sciences* 12: 395-413.
22. Riffo-Campos ÁL, Riquelme I, Brebi-Mieville P. 2016. Tools for sequence-based miRNA target prediction: what to choose? *International Journal of Molecular Sciences* 17(12): 1987.
23. Morin RD, Aksay G, Dolgosheina E, Ebhardt HA, Magrini V, Mardis ER, Sahinalp SC, Unrau PJ. 2008. Comparative analysis of the small RNA transcriptomes of Pinus contorta and Oryza sativa. *Genome Research* 18(4): 571-584.
24. Yadav P, Tamilselvan R, Harita M, Singh KK. 2024. MicroRNA-mediated regulation of nonsense-mediated mRNA decay factors: Insights into microRNA prediction tools and profiling techniques. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*. pp 195022.
25. Luna Buitrago D, Lovering RC, Caporali A. 2023. Insights into online microRNA bioinformatics tools. *Non-coding RNA* 9(2): 18.
26. Axtell MJ, Bartel DP. 2005. Antiquity of microRNAs and their targets in land plants. *The Plant Cell* 17(6): 1658-1673.

27. Vacic V, Jin H, Zhu JK, Lonardi S. 2008. A probabilistic method for small RNA flow gram matching. *Biocomputing* 2008: 75-86.
28. Kozomara A, Griffiths-Jones S. 2014. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Research* 42(D1): D68-D73.
29. Kumar A, Wong AKL, Tizard ML, Moore RJ, Lefèvre C. 2012. miRNA targets: a database for miRNA target predictions in coding and non-coding regions of mRNAs. *Genomics* 100(6): 352-356.
30. Selbach M, Schwanhäusser B, Thierfelder N, Fang Z, Khanin R, Rajewsky N. 2008. Widespread changes in protein synthesis induced by microRNAs. *Nature* 455(7209): 58-63.
31. Karere GM, Glenn JP, Li G, Konar A, VandeBerg JL, Cox LA. 2023. Potential miRNA biomarkers and therapeutic targets for early atherosclerotic lesions. *Scientific Reports* 13(1): 3467.
32. Kern F, Aparicio-Puerta E, Li Y, Fehlmann T, Kehl T, Wagner V, Ray K, Ludwig N, Lenhof HP, Meese E, Keller A. 2021. miRTargetLink 2.0—interactive miRNA target gene and target pathway networks. *Nucleic Acids Research* 49(W1): W409-W416.
33. Videm P, Kumar A, Zharkov O, Grüning BA, Backofen R. 2021. ChiRA: an integrated framework for chimeric read analysis from RNA-RNA interactome and RNA structurome data. *GigaScience* 10(2): giaa158.
34. Dweep H, Gretz N. 2015. miRWalk2. 0: a comprehensive atlas of microRNA-target interactions. *Nature Methods* 12(8): 697-697.
35. Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano M, Jungkamp AC, Munschauer M, Ulrich A. 2010. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* 141(1): 129-141.
36. Roberts JT, Borchert GM. 2017. Computational prediction of microRNA target genes, target prediction databases, and web resources. *Bioinformatics in MicroRNA Research*. pp 109-122.
37. Wen M, Cong P, Zhang Z, Lu H, Li T. 2018. DeepMirTar: a deep-learning approach for predicting human miRNA targets. *Bioinformatics* 34(22): 3781-3787.
38. Xu A, Chen J, Peng H, Han G, Cai H. 2019. Simultaneous interrogation of canceromics to identify subtypes with significant clinical differences. *Frontiers in Genetics* 10: 436387.
39. Loganantharaj R, Randall TA. 2017. The limitations of existing approaches in improving MicroRNA target prediction accuracy. *Bioinformatics in MicroRNA Research*. pp 133-158.
40. Zheng X, Chen L, Li X, Zhang Y, Xu S, Huang X. 2020. Prediction of miRNA targets by learning from interaction sequences. *Plos One* 15(5): p.e0232578.
41. Shenoy A, Blelloch RH. 2014. Regulation of microRNA functions in somatic stem cell proliferation and differentiation. *Nature Reviews Molecular Cell Biology* 15(9): 565-576.
42. Sedaghat N, Fathy M, Modarressi MH, Shojaie A. 2017. Combining supervised and unsupervised learning for improved miRNA target prediction. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 15(5): 1594-1604.