MiRNA-ProtPred: Unlocking miRNA Secrets Against Pathogens

HELP AND GUIDELINES

CONTENTS
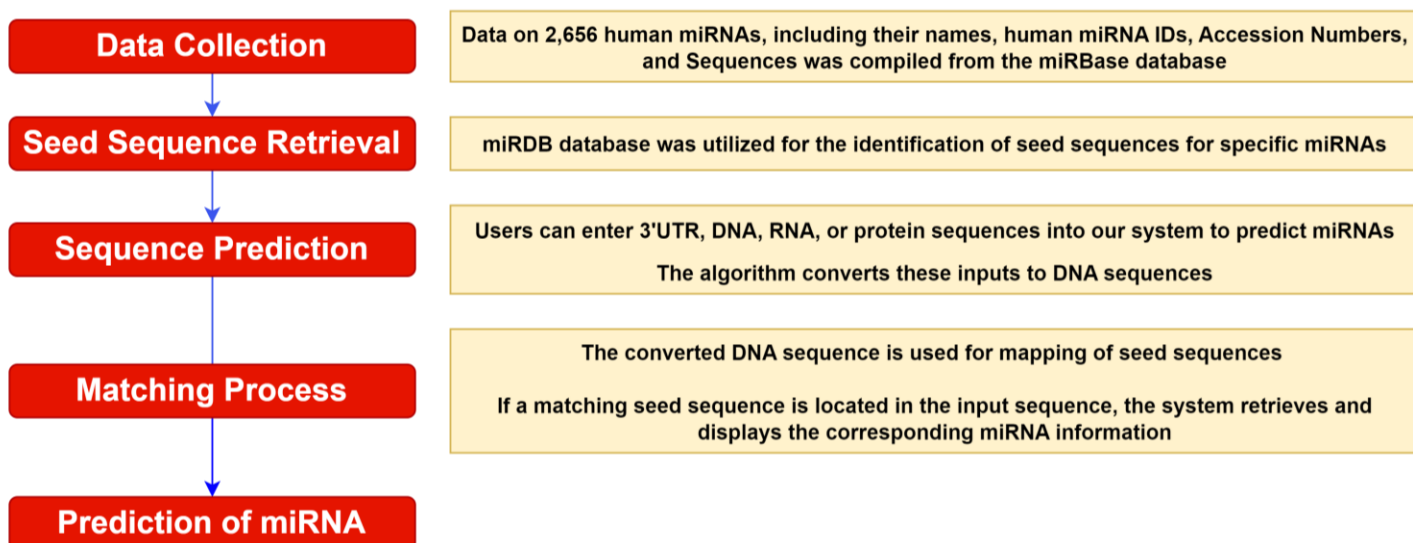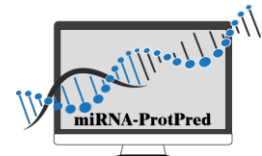
1. OVERVIEW

miRNA-ProtPred is the first online service that aids in the prediction and identification of the human miRNA with the potential to inhibit the expression of pathogen-specific genes and proteins.

2. BACKGROUND

Methodology used to train our model:

Data Collection	Data on 2,656 human miRNAs, including their names, human miRNA IDs, Accession Numbers, and Sequences was compiled from the miRBase database
Seed Sequence Retrieval	miRDB database was utilized for the identification of seed sequences for specific miRNAs
Sequence Prediction	Users can enter 3'UTR, DNA, RNA, or protein sequences into our system to predict miRNAs. The algorithm converts these inputs to DNA sequences
Matching Process	The converted DNA sequence is used for mapping of seed sequences. If a matching seed sequence is located in the input sequence, the system retrieves and displays the corresponding miRNA information
Prediction of miRNA	

Figure 1: Flow chart of development of prediction tool.

3. 3'UTRmiRNA-Pred

Users with a 3'UTR of mRNA sequence from a pathogenic source can employ this tool to forecast potential human miRNAs capable of suppressing the expression of the corresponding protein or gene. Users are encouraged to input the 3'UTR of mRNA sequence of the relevant protein.

Procedure:

3.1. Click on the **"3'UTRmiRNA-Pred"** option in the Menu Bar:

3.2. Enter a **Job Id**, although this step is not necessary, it is useful for systematic research purposes:

3.3. Enter 3'UTR sequence of the mRNA of the pathogen of interest in the text box below **"Enter Sequence (3'UTR)":**

3.4. Click the blue **"Search"** button located below the text box to obtain output:

3.5. The Output appears in the form of a table comprising of data on the following – **Description, Human miRNA ID, Accession Number, Sequence, Seed,** and **Position:**

Search

Matching Rows: Streptococcus pneumoniae

Description	Human miRNA ID	Accession Number	Sequence 5' to 3'	Seed	Position
HomosapiensmiR-494-3p	hsa-miR-494-3p	MIMAT0002816	UGAAACAUACACGGGAAACCUC	tgtttca	29
HomosapiensmiR-3133	hsa-miR-3133	MIMAT0014998	UAAAGAACUCUUAAAACCCAAU	gttcttt	37
HomosapiensmiR-10b-3p	hsa-miR-10b-3p	MIMAT0004556	ACAGAUUCGAUUCUAGGGGAAU	aatctga	49
HomosapiensmiR-34a-3p	hsa-miR-34a-3p	MIMAT0004557	CAAUCAGCAAGUAUACUGCCCU	ctgatta	52

3.5.1 Description:

Description represents information on the microRNA molecule, these include – Organism name, the type of biomolecule, identifier or name (based on order of discovery), the strand from which it originates in its precursor molecule.

For example, in the first output generated above i.e., **HomosapiensmiR-494-3p**:

> <u>Organism name</u> – *Homo sapiens*
>
> <u>Type of biomolecule</u> – miR (micro-RNA)
>
> <u>Identifier</u> - 494
>
> <u>The strand from which it originates in its precursor molecule</u> – 3p

Matching Rows: Streptococcus pneumoniae

Description	Human miRNA ID	Accession Number	Sequence 5' to 3'	Seed	Position
HomosapiensmiR-494-3p	hsa-miR-494-3p	MIMAT0002816	UGAAACAUACACGGGAAACCUC	tgtttca	29
HomosapiensmiR-3133	hsa-miR-3133	MIMAT0014998	UAAAGAACUCUUAAAACCCAAU	gttcttt	37
HomosapiensmiR-10b-3p	hsa-miR-10b-3p	MIMAT0004556	ACAGAUUCGAUUCUAGGGGAAU	aatctga	49
HomosapiensmiR-34a-3p	hsa-miR-34a-3p	MIMAT0004557	CAAUCAGCAAGUAUACUGCCCU	ctgatta	52

3.5.2 Human miRNA ID:

The miRNA IDs are frequently utilized in publications, databases, and bioinformatics tools to guarantee coherence and precision in research pertaining to miRNAs, it can be seen as a simplified version of the **Description**.

Matching Rows: Streptococcus pneumoniae

Description	Human miRNA ID	Accession Number	Sequence 5' to 3'	Seed	Position
HomosapiensmiR-494-3p	hsa-miR-494-3p	MIMAT0002816	UGAAACAUACACGGGAAACCUC	tgtttca	29
HomosapiensmiR-3133	hsa-miR-3133	MIMAT0014998	UAAAGAACUCUUAAAACCCAAU	gttcttt	37
HomosapiensmiR-10b-3p	hsa-miR-10b-3p	MIMAT0004556	ACAGAUUCGAUUCUAGGGGAAU	aatctga	49
HomosapiensmiR-34a-3p	hsa-miR-34a-3p	MIMAT0004557	CAAUCAGCAAGUAUACUGCCCU	ctgatta	52

3.5.3 Accession Number:

Accession numbers standardise miRNA sequence and information access for researchers.

miRBase is a notable miRNA database. MiRBase assigns accession numbers like "MI0000XXX" for mature miRNAs and "MI0000XXX.1" for precursor miRNAs. Each miRNA entry has a unique numerical identification "XXX".

Matching Rows: Streptococcus pneumoniae

Description	Human miRNA ID	Accession Number	Sequence 5' to 3'	Seed	Position
HomosapiensmiR-494-3p	hsa-miR-494-3p	MIMAT0002816	UGAAACAUACACGGGAAACCUC	tgtttca	29
HomosapiensmiR-3133	hsa-miR-3133	MIMAT0014998	UAAAGAACUCUUAAAACCCAAU	gttcttt	37
HomosapiensmiR-10b-3p	hsa-miR-10b-3p	MIMAT0004556	ACAGAUUCGAUUCUAGGGGAAU	aatctga	49
HomosapiensmiR-34a-3p	hsa-miR-34a-3p	MIMAT0004557	CAAUCAGCAAGUAUACUGCCCU	ctgatta	52

3.5.4 Sequence 5' to 3':

Unique nucleotide sequence of the microRNA comprising of Adenine(A), Guanine(G), Uracil(U) and Cytosine(C).

Matching Rows: Streptococcus pneumoniae

Description	Human miRNA ID	Accession Number	Sequence 5' to 3'	Seed	Position
HomosapiensmiR-494-3p	hsa-miR-494-3p	MIMAT0002816	UGAAACAUACACGGGAAACCUC	tgtttca	29
HomosapiensmiR-3133	hsa-miR-3133	MIMAT0014998	UAAAGAACUCUUAAAACCCAAU	gttcttt	37
HomosapiensmiR-10b-3p	hsa-miR-10b-3p	MIMAT0004556	ACAGAUUCGAUUCUAGGGGAAU	aatctga	49
HomosapiensmiR-34a-3p	hsa-miR-34a-3p	MIMAT0004557	CAAUCAGCAAGUAUACUGCCCU	ctgatta	52

3.5.5 Seed:

Seed Sequence refers to the heptametrical sequence located primarily at positions 2-7 from the miRNA 3' end.

Matching Rows: Streptococcus pneumoniae

Description	Human miRNA ID	Accession Number	Sequence 5' to 3'	Seed	Position
HomosapiensmiR-494-3p	hsa-miR-494-3p	MIMAT0002816	UGAAACAUACACGGGAAACCUC	tgtttca	29
HomosapiensmiR-3133	hsa-miR-3133	MIMAT0014998	UAAAGAACUCUUAAAACCCAAU	gttcttt	37
HomosapiensmiR-10b-3p	hsa-miR-10b-3p	MIMAT0004556	ACAGAUUCGAUUCUAGGGGAAU	aatctga	49
HomosapiensmiR-34a-3p	hsa-miR-34a-3p	MIMAT0004557	CAAUCAGCAAGUAUACUGCCCU	ctgatta	52

3.5.6 Position:

Position is based on DNA sequence of provided input data (3'UTR, DNA, RNA and Protein).
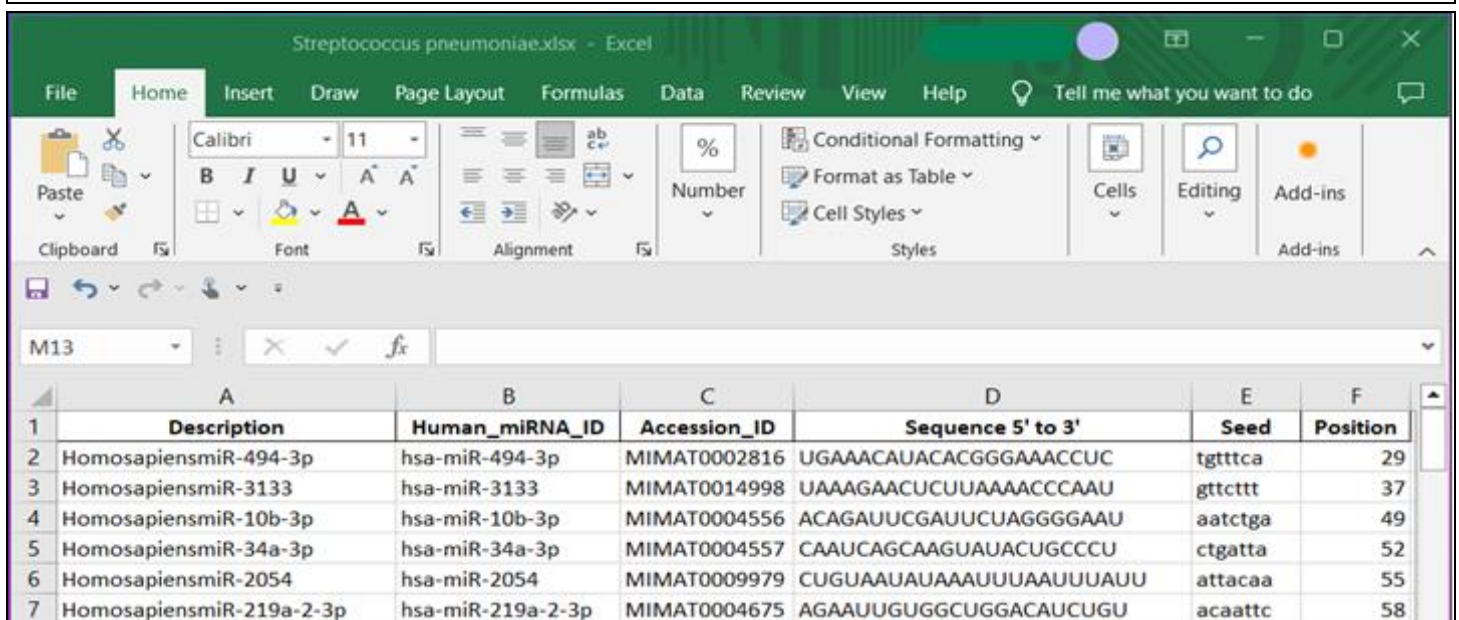
Matching Rows: Streptococcus pneumoniae

Description	Human miRNA ID	Accession Number	Sequence 5' to 3'	Seed	Position
HomosapiensmiR-494-3p	hsa-miR-494-3p	MIMAT0002816	UGAAACAUACACGGGAAACCUC	tgtttca	29
HomosapiensmiR-3133	hsa-miR-3133	MIMAT0014998	UAAAGAACUCUUAAAACCCAAU	gttcttt	37
HomosapiensmiR-10b-3p	hsa-miR-10b-3p	MIMAT0004556	ACAGAUUCGAUUCUAGGGGAAU	aatctga	49
HomosapiensmiR-34a-3p	hsa-miR-34a-3p	MIMAT0004557	CAAUCAGCAAGUAUACUGCCCU	ctgatta	52

3.5.6. You can select the "**Download Complete Data**" option below the table to obtain the output in the form of an excel sheet (.xls file format).
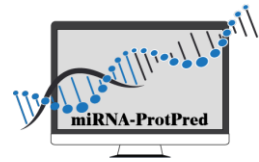
HomosapiensmiR-3185	hsa-miR-3185	MIMAT0015065	AGAAGAAGGCGGUCGGUCUGCGG	ttcttca	1281
HomosapiensmiR-4709-3p	hsa-miR-4709-3p	MIMAT0019812	UUGAAGAGGAGGUGCUCUGUAGC	tcttcaa	1282
HomosapiensmiR-4635	hsa-miR-4635	MIMAT0019692	UCUUGAAGUCAGAACCCGCAA	cttcaag	1283
HomosapiensmiR-624-5p	hsa-miR-624-5p	MIMAT0003293	UAGUACCAGUACCUUGUGUUCA	tggtact	1291

`Download Complete Data`

4. miRNA-SeqFinder

For prediction of potential human miRNA binding within coding sequence (CDS), user can use **miRNA-SeqFinder.**

The **miRNA-SeqFinder** allows users to predict human miRNA using DNA, RNA, or protein sequences as input.

Procedure:

4.1. Click on the **"miRNA-SeqFinder"** option in the Menu Bar:

4.2. Enter a **Job Id**, although this step is not necessary, it is useful for systematic research purposes:

4.3. Enter DNA, RNA or Protein sequence of the target protein of the pathogen of interest in the text box below **"Enter Sequence (DNA, RNA, or Protein)"**:

4.4. Click the blue **"Search"** button located below the text box to obtain output:

4.5. The Output appears in the form of a table comprising of data on the following – **Description, Human miRNA ID, Accession Number, Sequence, Seed,** and **Position:**

Matching Rows:

Description	Human miRNA ID	Accession Number	Sequence 5' to 3'	Seed	Position
HomosapiensmiR-3147	hsa-miR-3147	MIMAT0015019	GGUUGGGCAGUGAGGAGGGUGUGA	cccaaca	83
HomosapiensmiR-4425	hsa-miR-4425	MIMAT0018940	UGUUGGGAUUCAGCAGGACCAU	cccaaca	83
HomosapiensmiR-3192-3p	hsa-miR-3192-3p	MIMAT0027027	CUCUGAUCGCCCUCUCAGCUC	gatcagaa	56
HomosapiensmiR-3655	hsa-miR-3655	MIMAT0018075	GCUUGUCGCUGCGGUGUUGCU	cgacaaga	6
HomosapiensmiR-3940-3p	hsa-miR-3940-3p	MIMAT0018356	CAGCCCGGAUCCCAGCCCACUU	ccgggct	101
HomosapiensmiR-4778-3p	hsa-miR-4778-3p	MIMAT0019937	UCUUCUUCCUUUGCAGAGUUGA	gaagaag	40

Download Complete Data ⬅

4.6. You may select the **"Download Complete Data"** option below the table to obtain the output in the form of an excel sheet (.xlsx file format).

5. miRNA-Validator

The **miRNA-Validator** can be utilized to validate the predicted human miRNA by using experimentally reported downregulated miRNA (ERD miRNA) during pathogenic infection. Sometimes, ERD miRNA and predicted miRNA have similar seeds. Users can provide list of ERD miRNA and sequences (Protein/mRNA/DNA/3'UTR) for validation. If given ERD miRNA is positive, user can match positive miRNA with predicted miRNA and perform *in silico* validation of predicted miRNA.

Procedure:

5.1. Click on the **"miRNA-Validator"** option in the Menu Bar:

5.2. Enter either the **3'UTR sequence of the mRNA** or the **DNA, RNA or Protein sequence of the target protein of the pathogen of interest** in the text box below **"Enter Sequence (DNA, RNA, or Protein)"**:

5.3. Enter the list of human miRNAs that are downregulated during pathogenic disease in the text box below **"Enter Human miRNA IDs (comma-spaced)"**:

5. 4. Click the blue **"Search"** button located below the text box to obtain output:

5.5. The output obtained suggests whether each of the given miRNA corresponding to their respective IDs are potential binder (**Positive**) or non-binder (**Negative**) to the given sequences:

Final search: User can use the positive miRNA (predicted as positive) for further analysis. If user will search for positive miRNA in the list of predicted miRNAs, they can find some common miRNA. This will ultimately validate that the current approach is a promising strategy for identification of potential human miRNA for development of therapeutics against pathogens.

Future study:

- Users can use the finally selected miRNA for bioinformatics analysis or experimental validation.

- Bioinformatics can include miRNA-mRNA docking, MD simulation, pathway analysis, and so on.

- Wet lab experimental validation can include binding assay, pathogenic protein expression inhibition test, pathogen growth inhibition assay (GIA), and so on.

- For experimental analysis, users can order miRNA mimics from authorized vendors.