1. miRMOD (Version: 0.3)

miRMOD is a **miR**NA **mod**ification prediction tool. It identifies modified miRNAs (5' and 3' non-templated nucleotide addition as well as trimming) using small RNA (sRNA) sequencing data and their corresponding targets. The graphical display of output in various formats helps in the analyzing the modification pattern at local as well as global level. Further it calculates variations (matches, mis-matches and minimum free-energy of binding) caused by the modifications w.r.t. target binding.

2. Prerequisite

Operating system: Windows XP, vista, 7 or Windows 8.x/10 (Recommended).

Environment: .NET Framework 4.0 (pre-installed in windows 8.x/10)

Third party software: RNAhybrid.

3. Installation

Download compressed miRMOD installation package. Unzip compressed miRMOD installation package and run miRMOD.exe.

miRMOD package (GUI) contains two executables – "*prepare_input*" and "*miRMOD*".

4. Input files

4.1 miRNA modification

Before running miRMOD, we highly recommend users to pre-process the sequencing data to contain only high quality reads to minimize the probability of getting false-positive results.

The very first step before executing miRMOD is preparation of input files. It requires three input files (sample files provided):

1*. miRNAs*: List of different mature miRNAs for which modification has to be predicted (fasta format). Such fasta file containing mature miRNA sequences can be downloaded from several databases like miRBase. Example file containing all mature miRNA sequences of Homo sapiens (2578 sequences; source: miRBase version 20) is available in miRMOD package.

2*. sRNA NGS reads*: Processed non-redundant set of small RNA sequencing reads (fasta format). It is required that headers of all reads should have desired syntax which includes

> ###_xRC

Where '###' is unique identifier for each read and 'RC' is the read count of that read. For example:

> read2_x1200

This unique identifier should not have colon (:) or white spaces or '_x' other than as delimiter of read count. **You should use "*prepare_input*" program** (included in miRMOD installation package) to generate such read file before aligning it to reference genome or premiRNAs using bowtie.

3. *Alignment file*: Output file generated by bowtie after aligning sRNA reads to its reference genome or pre-miRNAs. User can set any set of desired parameters for bowtie to generate output file. Example alignment file is provided with miRMOD package.

4.2 Target Variation Analysis (TVA)

Several studies reported the alteration of targets due to terminal modification of mature miRNA. Therefore, miRMOD is equipped with a unique feature whereby user can explore if modification is playing any role in altering the miRNA-target interactions.
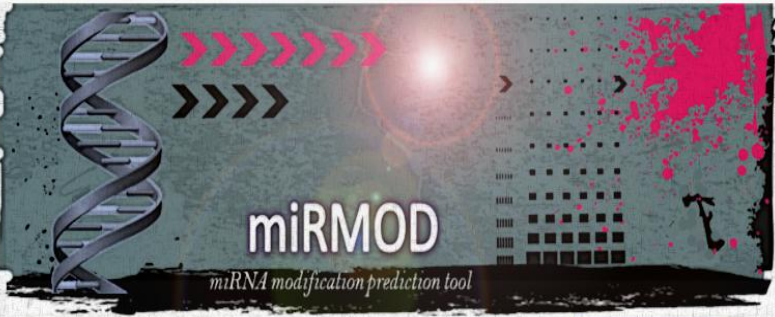
User can submit most probable targets of miRNAs (eg. 3'UTR sequences) in fasta format. miRMOD then compares the binding energy and binding site of miRNA-target and modified miRNA-target interactions. It calculates variations (matches, mis-matches and minimum free-energy of binding) caused by the modifications w.r.t. target binding. It, thus, may also help in predicting altered or novel targets of modified miRNAs.

RNAhybrid is used to predict miRNA target from the list of sequences. If the name of your organism, for RNAhybrid, is not listed then you can create your own database file for RNAhybrid containing <xi> and <theta> values and select option 'Other'. Name of this file must be "Other" (case sensitive, without quotes) and must be located in miRMOD directory (having miRMOD executable).

5. Screen shot manual:

Step 1: Upload sRNA read file.

Step 2: Upload miRNA sequence(s) for which modification(s) has to be predicted.

Step 3: Upload Alignment file generated after aligning sRNS reads to reference genome or pre-miRNA. You can also select read count threshold to filter off reads with low read count.

Step 4 and 5: Choose options, your favorite output directory and submit job.

Immediately after job submission miRMOD will start working in background and you can monitor the progress by clicking miRMOD icon in system tray.

Usually, alignment file generated by aligning reads to whole genome has high file size. Such file requires large processing time, depending upon file size. Once this alignment file is processed a new window appears where user can visualize number of modifications currently found and how much processing is left.

Once processing is over first result page with six tabs appear. The first tab i.e. **'summary'** summarizes all results in single window. This has three major sections:

1.) Basic result summary to find out miRNA with highest modifications, length distribution of modifications etc.

2.) Modification composition: To find out which modification type is most common e.g. 3' Addition in current example and percentage of reads modified in the dataset.

3.). Total number of modifications per miRNA. The graph helps in finding those miRNA with comparatively more modifications than other.

Summary of results
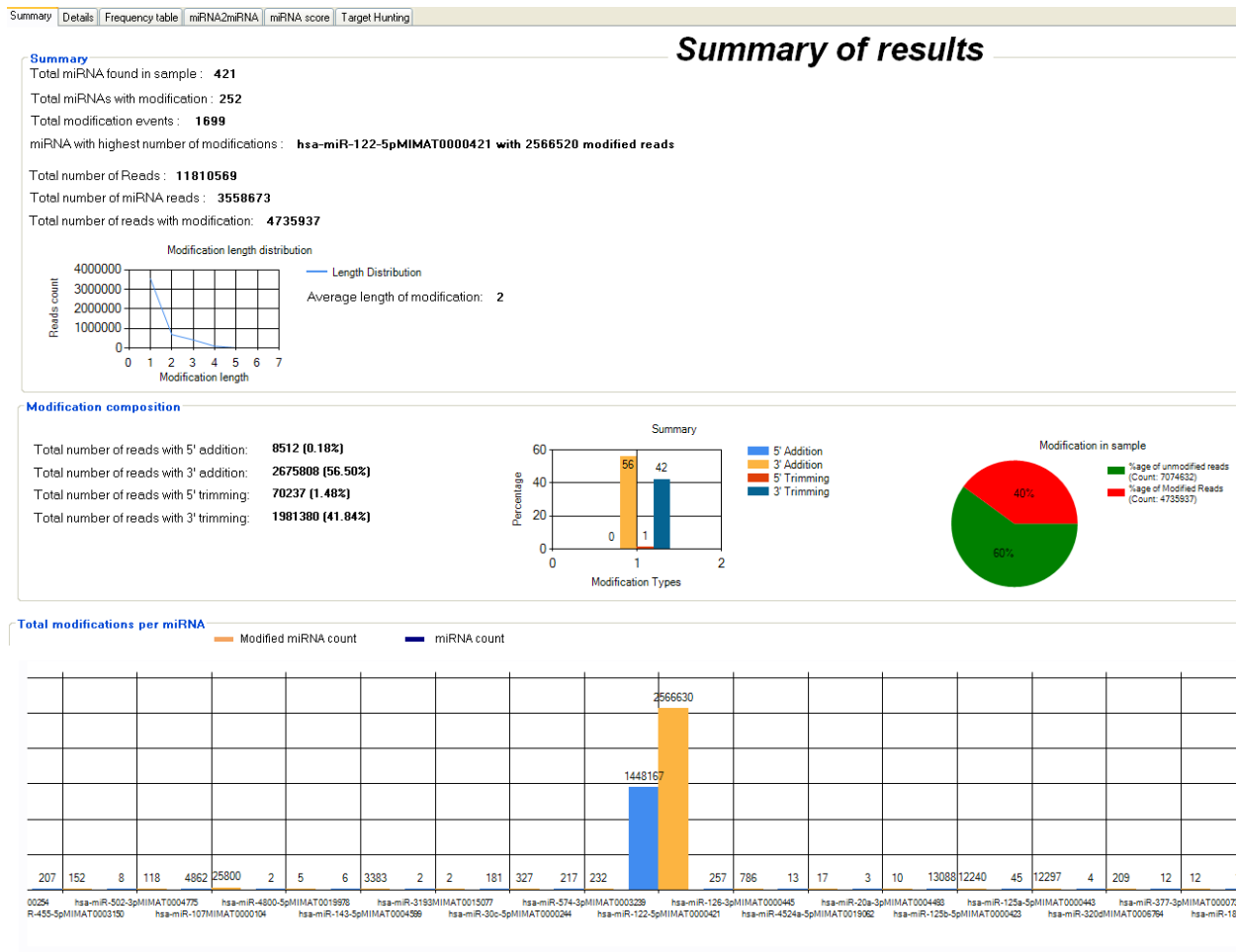
Each tab contains miRNA modifications in different format to facilitate analysis. Next tab **'Details'** gives the detailed information of the analysis.
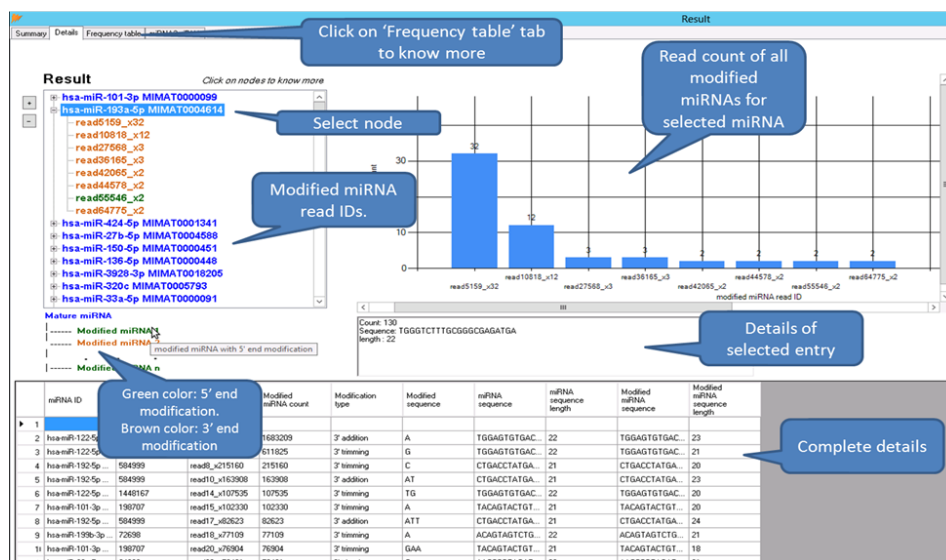
'Frequency table' contains the percentage of occurrence of a modified sequence in the dataset.

It was also found that some mature miRNA sequences (uploaded by user) get trim to generate new mature miRNA sequence (uploaded by user). List of such **'miRNA to miRNA'** conversion is provided in this tab.

'miRNA score' tab calculates Z-score for each miRNA which measures its relative tendency to get modified under given options (details are given in the manuscript). User can selection miRNAs according to their Z-score for the next analysis i.e., target hunting.

Z-score

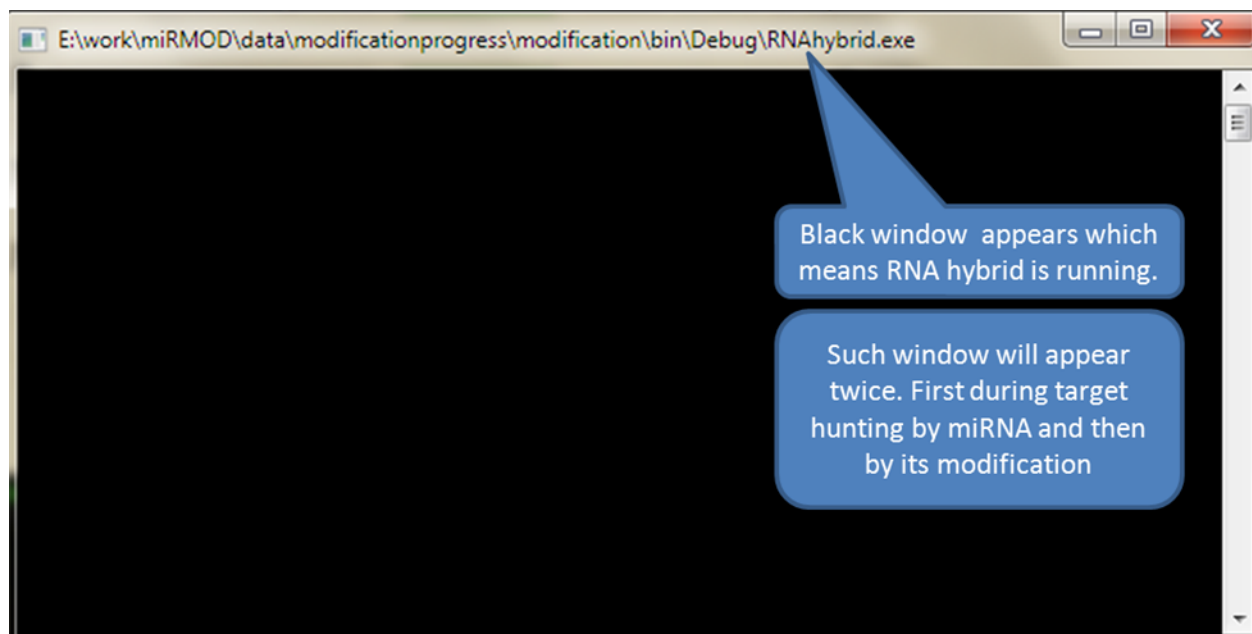|	miRNA ID	miRNA Sequence	Number of modifica	Mature miRNA count	Total modifications count	3' modification(s)	5' modification(s)	Score
2	hsa-miR-122-5pMIMAT000...	TGGAGTGTGACAATGGTGT...	81	1448167	2566520	2540631	25889	15.69245736126...
3	hsa-miR-192-5pMIMAT000...	CTGACCTATGAATTGACAGCC	128	584999	655168	637512	17656	1.492089171356...
8	hsa-let-7a-5pMIMAT0000062	TGAGGTAGTAGGTTGTATA...	27	239751	118733	116237	2496	-0.63942322041...
4	hsa-miR-101-3pMIMAT000...	TACAGTACTGTGATAACTGAA	59	198707	260988	258539	2449	0.568590088949...
1	hsa-let-7c-5pMIMAT0000064	TGAGGTAGTAGGTTGTATG...	28	117030	46971	45237	1734	-0.41219042270...
1	hsa-miR-29a-3pMIMAT000...	TAGCACCATCTGAAATCGGT...	14	160538	28848	27487	1361	-0.39165821703...
1	hsa-let-7d-5pMIMAT0000065	AGAGGTAGTAGGTTGCATA...	23	99220	23475	23130	345	-0.37399149232...
9	hsa-miR-21-5pMIMAT0000...	TAGCTTATCAGACTGATGTT...	29	18366	81617	81262	355	0.248352970278...
1	hsa-miR-148a-3pMIMAT00...	TCAGTGCACTACAGAACTTT...	18	60253	27550	26385	1165	-0.17327017956...
6	hsa-miR-99a-5pMIMAT000...	AACCCGTAGATCCGATCTTG...	20	64290	130999	130863	136	0.161328972664...
7	hsa-let-7f-5pMIMAT0000067	TGAGGTAGTAGATTGTATA...	25	159410	145174	142662	2512	-0.13276783194...
1	hsa-let-7b-5pMIMAT0000063	TGAGGTAGTAGGTTGTGTG...	27	17851	61147	60917	230	0.132584023928...
5	hsa-miR-199b-3pMIMAT00...	ACAGTAGTCTGCACATTGGT...	19	72698	128881	115302	13579	0.114919039359...
5	hsa-miR-29c-3pMIMAT000...	TAGCACCATTTGAAATCGGT...	9	15178	3031	2931	100	-0.08986009515...
1	hsa-miR-378a-3pMIMAT00...	ACTGGACTTGGAGTCAGAA...	29	22542	53829	52857	972	0.087050074786...
2	hsa-let-7g-5pMIMAT0000414	TGAGGTAGTAGTTTGTACA...	12	14236	9310	8998	312	-0.08112270127...
4	hsa-miR-26a-5pMIMAT000...	TTCAAGTAATCCAGGATAGG...	19	13301	11350	10717	633	-0.07728692079...
4	hsa-miR-185-5pMIMAT000...	TGGAGAGAAAGGCAGTTCC...	14	7189	5049	4979	70	-0.07604985853...
3	hsa-miR-320aMIMAT00005...	AAAAGCTGGGTTGAGAGGG...	28	9228	8250	7908	342	-0.07560159965...
6	hsa-miR-128-3pMIMAT000...	TCACAGTGAACCGGTCTCTTT	12	2940	1648	1636	12	-0.07353431879...
1	hsa-miR-1MIMAT0000416	TGGAATGTAAAGAAGTATG...	8	2166	351	298	53	-0.07336308947...
2	hsa-miR-125b-5pMIMAT00...	TCCCTGAGACCCTAACTTGT...	15	13088	12174	12054	120	-0.07322213850...
7	hsa-miR-146b-5pMIMAT00...	TGAGAACTGAATTCCATAGG...	16	2373	1613	1579	34	-0.07295241751...
9	hsa-miR-451aMIMAT00016...	AAACCGTTACCATTACTGAG...	6	2227	615	604	11	-0.07251947182...

And the last step is **Target Hunting:** Before using this feature, user have to select the miRNAs and corresponding modification type to be analyzed for target hunting as described below:

a) Select miRNA(s) for TVA: User have different options for selecting miRNAs for target hunting like

- Selecting miRNAs based on Z-score.
- Based on the read count of miRNA.
- Using miRNA IDs uploaded as list.
- All miRNAs.

b) Modification type: User can also specify the type of modifications to be searched for selected miRNA. If a given miRNA do not have selected modification type then it will be ignored. User can also select only highest modified read (for each selected modification type) for the target analysis.
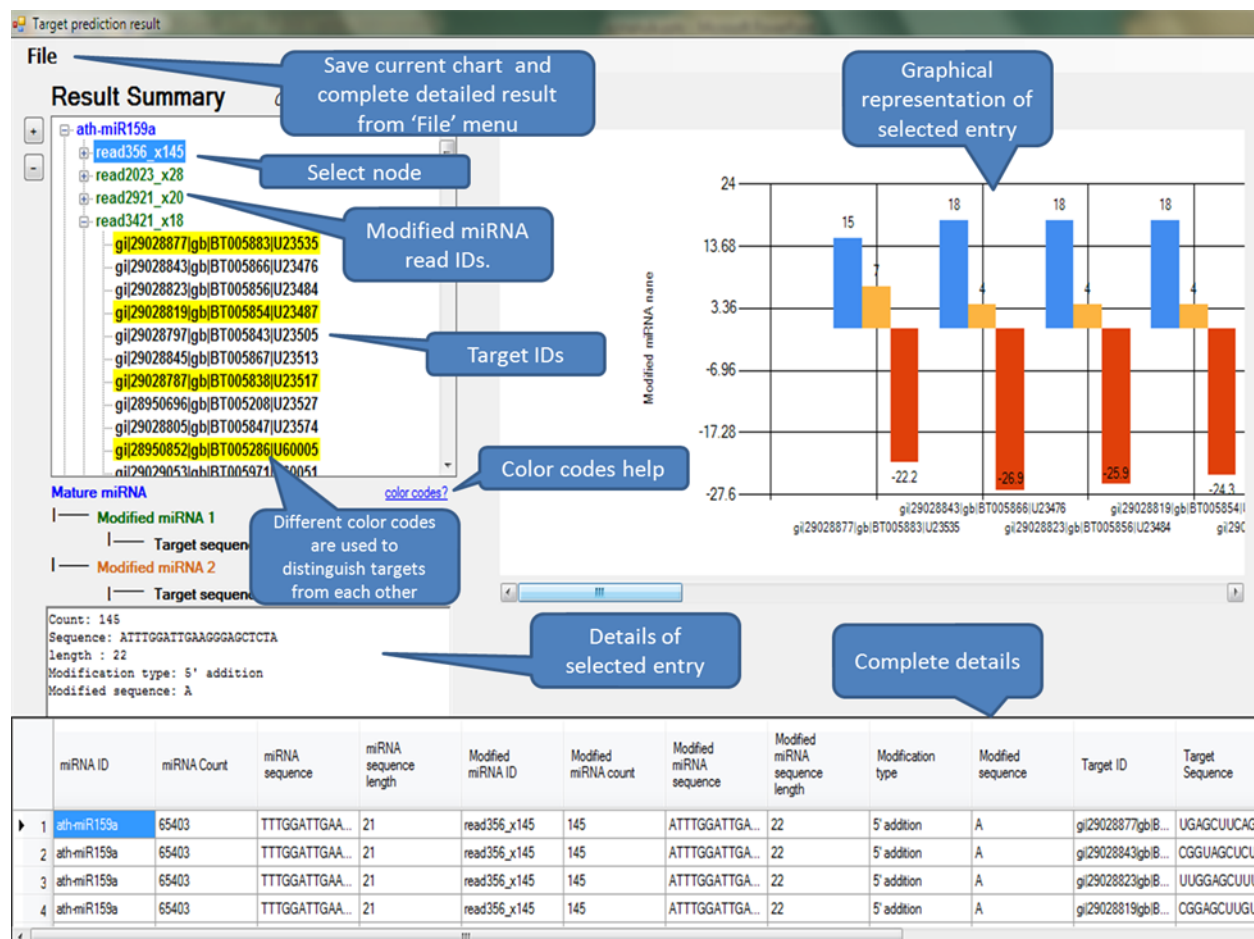
After submitting job command prompt window appears

RNAhybrid may take some time, depending upon number of miRNAs, modifications and target sequences. Immediately after the execution of RNAhybrid a new output window will appear containing the target prediction result.

6. miRMOD (Command line version)

Pre-complied miRMOD executables can also be executed via command line (command prompt or terminal) without GUI support. The command line version executables are cross-platform and can be executed in any OS (Windows, Linux and Mac) with same arguments. Three executables are available in command line package of miRMOD:

1. **prepare_inputC.exe**: To convert input sequences in fasta/fastq/TSV file into fasta format required by miRMOD.
2. **mirmodC.exe**: To execute miRMOD algorithm.
3. **TVA.exe**: To perform Target Variation Analysis.

6.1 Dependencies

Windows users do not require any special package to execute miRMOD via command prompt. Linux and Mac users, however, require mono compiler to execute miRMOD executable via terminal. Mono complier is part of *'monodevelop'* and is freely available at www.monodevelop.com/.

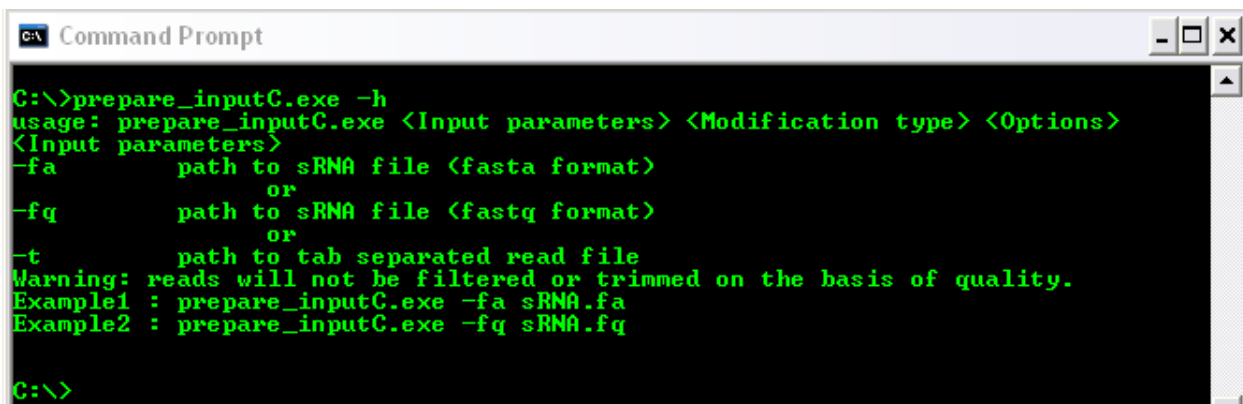Debian users can install *monodevelop* using following command:

sudo apt-get install monodevelop

Other OS users can download source files available at www.monodevelop.com/download/.

6.2 Running miRMOD via command line

Open terminal in Linux/MAC or command prompt using **'cmd'** command in the run menu. In the command prompt locate the directory having miRMOD executables using 'cd' command and execute *'prepare_inputC.exe'* in following way.

Windows
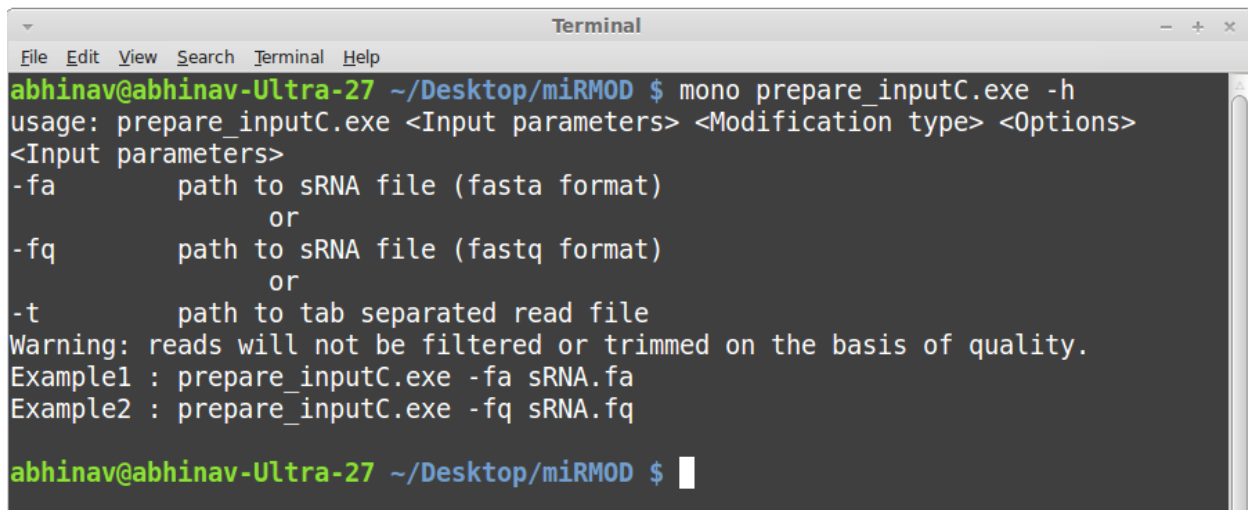
Linux

```
abhinav@abhinav-Ultra-27 ~/Desktop/miRMOD $ mono prepare_inputC.exe -h
usage: prepare_inputC.exe <Input parameters> <Modification type> <Options>
<Input parameters>
-fa        path to sRNA file (fasta format)
              or
-fq        path to sRNA file (fastq format)
              or
-t         path to tab separated read file
Warning: reads will not be filtered or trimmed on the basis of quality.
Example1 : prepare_inputC.exe -fa sRNA.fa
Example2 : prepare_inputC.exe -fq sRNA.fq

abhinav@abhinav-Ultra-27 ~/Desktop/miRMOD $ █
```
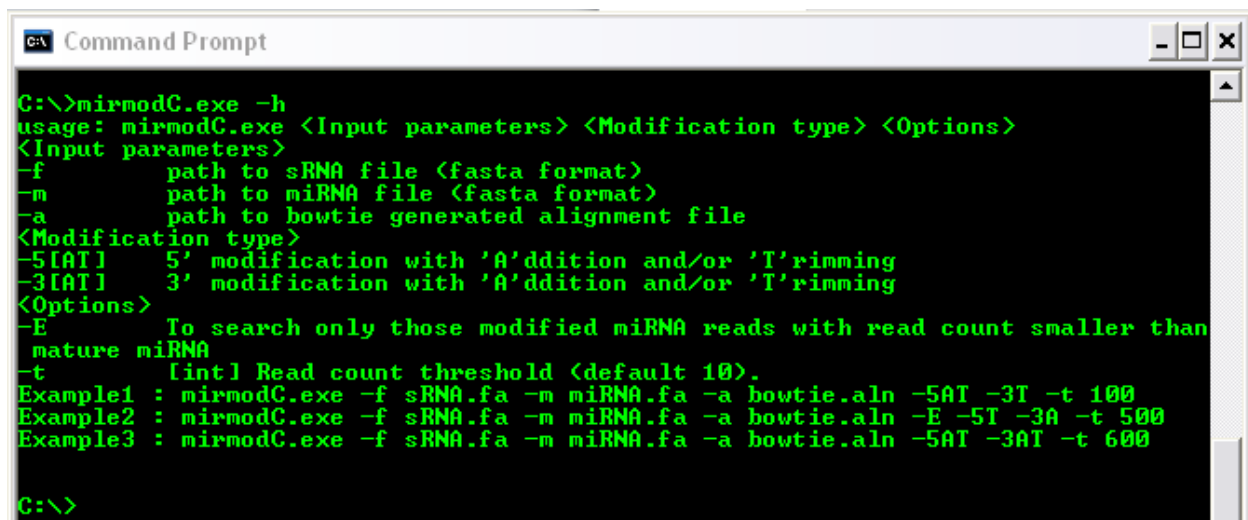
To prepare miRMOD input file user can submit the processed input file (fasta, fastq or TSV) to prepare_inputC.exe script. The script requires labels to the input file. For example, if input file is in fastq format then –fq is mandatory before filename. The output of this script is a fasta file in which header of each unique sequence include read count of that given sequence.
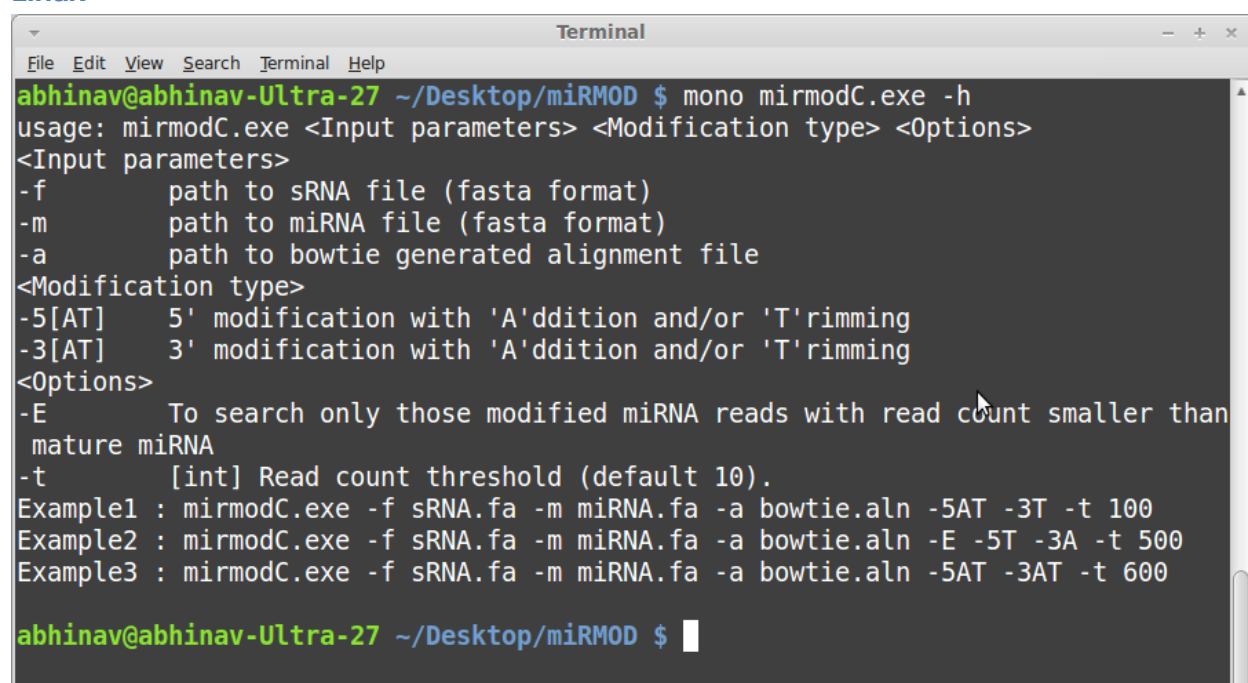
miRMODC

Once the input file is generated, user must align the fasta file to reference genome using bowtie. For more information about the input files for miRMOD please refer section 4. The resulting files can be executed via command line in following way:

Windows

```
C:\>mirmodC.exe -h
usage: mirmodC.exe <Input parameters> <Modification type> <Options>
<Input parameters>
-f         path to sRNA file (fasta format)
-m         path to miRNA file (fasta format)
-a         path to bowtie generated alignment file
<Modification type>
-5[AT]     5' modification with 'A'ddition and/or 'T'rimming
-3[AT]     3' modification with 'A'ddition and/or 'T'rimming
<Options>
-E         To search only those modified miRNA reads with read count smaller than
  mature miRNA
-t         [int] Read count threshold (default 10).
Example1 : mirmodC.exe -f sRNA.fa -m miRNA.fa -a bowtie.aln -5AT -3T -t 100
Example2 : mirmodC.exe -f sRNA.fa -m miRNA.fa -a bowtie.aln -E -5T -3A -t 500
Example3 : mirmodC.exe -f sRNA.fa -m miRNA.fa -a bowtie.aln -5AT -3AT -t 600


C:\>
```

Linux

```
                                    Terminal                        − + ×

File  Edit  View  Search  Terminal  Help

abhinav@abhinav-Ultra-27 ~/Desktop/miRMOD $ mono mirmodC.exe -h
usage: mirmodC.exe <Input parameters> <Modification type> <Options>
<Input parameters>
-f          path to sRNA file (fasta format)
-m          path to miRNA file (fasta format)
-a          path to bowtie generated alignment file
<Modification type>
-5[AT]      5' modification with 'A'ddition and/or 'T'rimming
-3[AT]      3' modification with 'A'ddition and/or 'T'rimming
<Options>
-E          To search only those modified miRNA reads with read count smaller than
 mature miRNA
-t          [int] Read count threshold (default 10).
Example1 : mirmodC.exe -f sRNA.fa -m miRNA.fa -a bowtie.aln -5AT -3T -t 100
Example2 : mirmodC.exe -f sRNA.fa -m miRNA.fa -a bowtie.aln -E -5T -3A -t 500
Example3 : mirmodC.exe -f sRNA.fa -m miRNA.fa -a bowtie.aln -5AT -3AT -t 600

abhinav@abhinav-Ultra-27 ~/Desktop/miRMOD $ █
```

Executing miRMOD via command line is very easy. Three files are required as input parameters which are discussed in section **4.** For different files different labels are required before corresponding filenames as command line arguments e.g. **–f , -m** and **–a** . Users are also requested to define what type of modifications to be searched. For example, if user have to search 5' modifications with additions and trimming then -5AT (not -5TA) should be given as command line argument. Likewise if -5T and -3AT is given as command line argument then all trimmings at 5' end will be searched, while all additions and trimming at 3' end will be included in final analysis. For more options –h argument can be used. The output of miRMOD is multiple files each representing each section as discussed in section 5. Moreover, a session file named as 'session1.mod' will also be generated by miRMOD in working directory. This session file is required by TVA.exe and also can be loaded in GUI version of miRMOD for graphical display of results.

Target Variation Analysis (TVA)

The script performs RNAhybrid analysis for the selected miRNAs and its modifications. It computes variation in the minimum free binding energy change (Kcal/mol) between miRNA-target binding and modified miRNA-target binding. TVA can be executed via command line using following way:

Windows

```
C:\>TVA.exe -h
usage: TVA.exe <Input parameters> <Options>
<Input parameters>
-i          path to miRMOD session file (.mod format)
-t          path to target sequence (fasta format)<Options>
-Z          [float] select only those miRNA with higher Z score
                or
-rc         [int] Read count threshold for miRNA (default 100).
-O          Human or Fly or Worm or Plant or Other (case sensitive)
-5[AT]      Select only 5' modification with 'A'ddition and/or 'T'rimming for sele
cted miRNA(s)
-3[AT]      Select only 3' modification with 'A'ddition and/or 'T'rimming for sele
cted miRNA(s)
-H          [T/F] Select only modified read with highest read count (default:T)
Note: Add path to RNAhybrid in file rnahybrid.txt.PLEASE DONOT REMOVE rnahybrid.
txt
=============================================================
Example1 : TVA.exe -i session1.mod -Z 0.6   -5AT -3T -H F
Example2 : TVA.exe -i session1.mod -rc 1000 -E -5T -3A -H F
Example3 : TVA.exe -i session1.mod -Z 2.0 -5AT -3AT -H T


C:\>
```

Linux

```
abhinav@abhinav-Ultra-27 ~/Desktop/miRMOD $ mono TVA.exe -h
usage: TVA.exe <Input parameters> <Options>
<Input parameters>
-i          path to miRMOD session file (.mod format)
-t          path to target sequence (fasta format)<Options>
-Z          [float] select only those miRNA with higher Z score
                or
-rc         [int] Read count threshold for miRNA (default 100).
-O          Human or Fly or Worm or Plant or Other (case sensitive)
-5[AT]      Select only 5' modification with 'A'ddition and/or 'T'rimming for sele
cted miRNA(s)
-3[AT]      Select only 3' modification with 'A'ddition and/or 'T'rimming for sele
cted miRNA(s)
-H          [T/F] Select only modified read with highest read count (default:T)
Note: Add path to RNAhybrid in file rnahybrid.txt.PLEASE DONOT REMOVE rnahybrid.
txt
=============================================================
Example1 : TVA.exe -i session1.mod -Z 0.6   -5AT -3T -H F
Example2 : TVA.exe -i session1.mod -rc 1000 -E -5T -3A -H F
Example3 : TVA.exe -i session1.mod -Z 2.0 -5AT -3AT -H T

abhinav@abhinav-Ultra-27 ~/Desktop/miRMOD $ ▯
```

TVA.exe requires two files as input:

1. Session.mod: Session file created by miRMOD (command line version or GUI version).

2. Target sequences: 3'UTR or any other miRNA target sequences.

Apart from these two files TVA.exe needs information about the miRNAs to be selected for target variation analysis. This can be specified using –Z or –rc options. Also user needs to

specify what types of modifications to be searched for the selected miRNAs using -5[AT] and/or -3[AT] options. The output of the script is session2.mod (session file) that can be loaded in GUI version of miRMOD for graphical analysis.

7. Technical details

- Programming language: C#
- Built on windows XP (SP3) having .NET Framework 4.

Please report any bug or suggestion at dinesh@icgeb.res.in or abhinav@icgeb.res.in