

miRDeep* manual

2012-08-23

Jiyuan An, John Lai, Melanie Lehman, Colleen Nelson: Australian Prostate Cancer Research Center (APCRC-Q) and Institute of Health and Biomedical Innovation (IHBI), Queensland University of Technology (QUT), Brisbane, Australia

j.an@qut.edu.au

+61 7 3176 3075

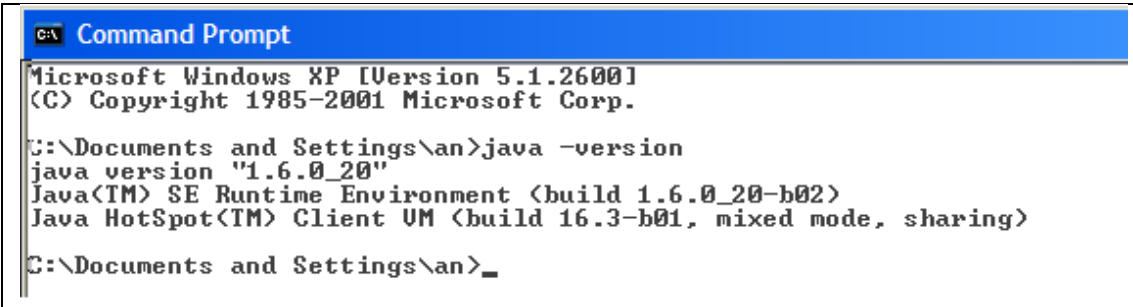
Web page:

<http://www.australianprostatecentre.org/research/software/mirdeep-star>

1. Execution with Demo-data

1.1 Install JDK

- Download JDK 6 (version $\geq 1.6.0_{20}$) to run miRDeep* from <http://jdk6.java.net/>.
- After installing JDK, the java command will run in a command prompt (Windows) as shown in Figure 1 or any other terminal (Linux). If version information is not available from the command prompt, then check whether the java path is in the environment variable.



```
C:\> Command Prompt
Microsoft Windows XP [Version 5.1.2600]
(C) Copyright 1985-2001 Microsoft Corp.

C:\Documents and Settings\an>java -version
java version "1.6.0_20"
Java(TM) SE Runtime Environment (build 1.6.0_20-b02)
Java HotSpot(TM) Client VM (build 16.3-b01, mixed mode, sharing)

C:\Documents and Settings\an>_
```

Figure 1 - Check whether JDK has been installed in windows.

1.2 Download files

- Download jar and demo data files from <http://www.australianprostatecentre.org/research/software/mirdeep-star>.
- There are "miRDeep_star.sh" and "miRDeep_star.bat" files in the unzipped directory. Windows uses the "miRDeep_star.bat" file to run miRDeep*, while "miRDeep_star.sh" is used in Linux.
- In Linux, type the following command making the file executable.
>chmod +x miRDeep.sh
To execute miRDeep*, type "miRDeep.sh" in command line:
>miRDeep.sh
- In Windows, double click "miRDeep_star.bat" to run the tool.

A window (as shown in Figure 2) will be displayed after execution of the program.

- e. To load raw sequence data in fastQ file, click the “Fastq”, “fasta” or “BAM” file button (labelled 1 in Figure 2). The file selection pop up window will appear.
- f. In the unzipped directory, there is a directory called “dat” where you will find a demo data file: “limitDemo.fastq”.
- g. miRDeep* accepts the following formats: fastq(“xxx.fastq”), fasta (“xxx.fa”), SAM/BAM(“xxx.sam” or “xxx.bam”) and result (“xxx.result”). Details are shown in the end of this README.
- h. To start the whole process of genomic mapping, and identifying miRNAs, select the file and click the submit button (labelled 2 in Figure 2) on the right hand side.

miRDeep*: An Integrated Application Tool For miRNA Identification From RNA Sequencing Data

hg18 ☐ miRDeep

Parameters

Adapter: TCGTATGCCGTCTTCTGCTTG Length: 18 to 23 min phred: 20 max multimap: 101 min reads: 5 min score: -10

Mapping sequences

Fastq or BAM file submit

1 input fastq, fasta, SAM, BAM, result file 2

progress

Result

miRNA ID	Score	expr	loci	sequence

Figure 2 screen shot of miRDeep*

1.3 Results

After execution, the identified miRNAs will be shown in a table (as shown in Figure 3). To show the hairpin structure of an identified miRNA, click the miRNA ID in the first column of the table.

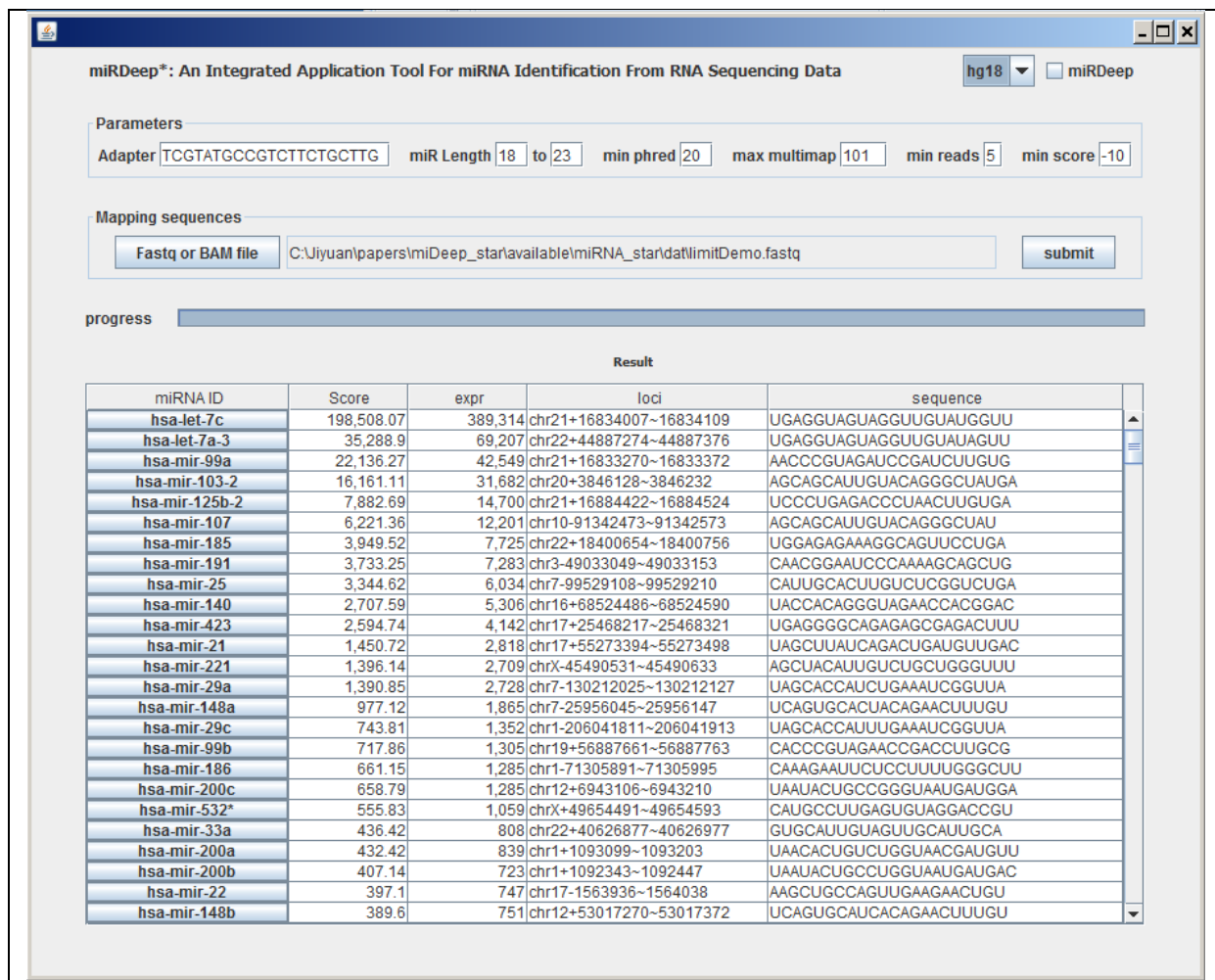


Figure 3 result screen of miRDeep*

In the hairpin structure pop up window (Figure 4), several extra functions are available:

- User can change the sequence in the top textbox. To show modified sequence: change the sequence in the first textbox, and then click the “RNA structure” button.
- Click “target genes” to show the target genes for the miRNA. This takes several minutes.

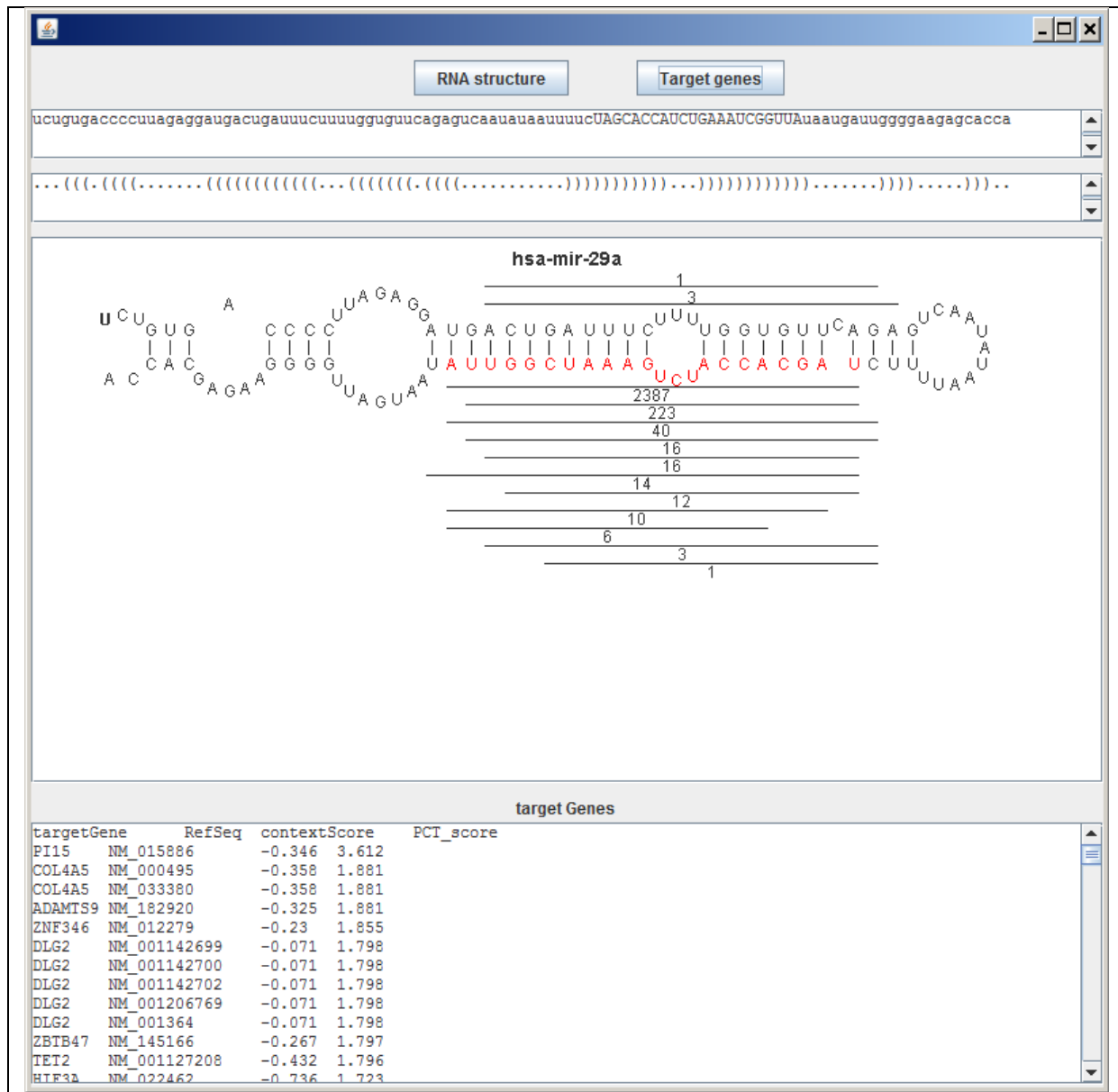


Figure 4 HairPin structure

2. Custom sequence data

2.1. Input file in fastQ format

fastQ has four lines to represent one read.

If your sequence data does not have a fourth line (quality line), simply input an "SSSS..." sequence, with the same length of read, in place of the fourth line.

@seq_0

ATGGGTTTGCAGTCCTCGGTTTAAAAAAAGACGTC

+

SNSNISPSNENSSPINNESNSNIIIOIIIONAAAAA<

2.2. Input file in fasta format

In the fasta format file, the copy number is appended in the description line within the delimitation table. For example,

```
>t1      234  
AGGCGATCACGTAGATTT
```

2.3. Input file in SAM/BAM format

miRDeep* also supports identified miRNAs from aligned sequences. There are several alignment algorithms (such as bowtie, BAW, and soap2). The aligned format should be in SAM or BAM format.

2.4. Show the results identified previously

miRDeep* saves the result file as “XXX.result”. You can load this file in miRDeep* to show the structure.