**miRPlant manual**

2014-01-10

Jiyuan An, John Lai, Atul Sajjanhar, Melanie Lehman, Colleen Nelson: Australian Prostate Cancer Research Centre – Queensland (APCRC-Q), Institute of Health and Biomedical Innovation (IHBI), Queensland University of Technology (QUT), Princess Alexandra Hospital (PAH), Translational Research Institute (TRI), Brisbane, Australia
j.an@qut.edu.au
+61 7 3443 7280

Web page:
http://www.australianprostatecentre.org/research/software/mirplant
https://sourceforge.net/projects/mirplant/

# 1. Execution with Demo-data

## 1.1 Install JDK

a. Download JDK 7 (version >= 1.7) to run miRPlant from http://www.oracle.com/technetwork/java/javase/downloads/index.html
b. 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.

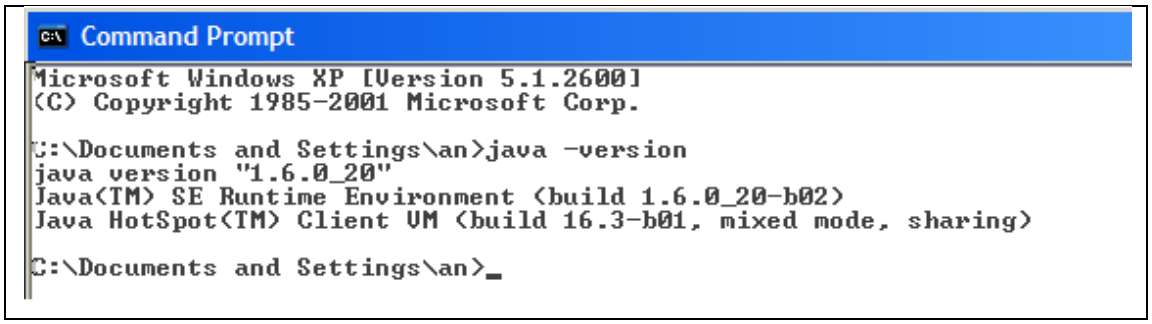



**Figure 1 - Check whether JDK has been installed in windows.**

## 1.2 Download files

a. Download jar and demo data files from http://www.australianprostatecentre.org/research/software/mirplant.
b. There are "miRPlant.sh" and "miRPlant.bat" files in the unzipped directory. Windows uses the "miRPlant.bat" file to run miRPlant, while "miRPlant.sh" is used in Linux.
c. In Linux, to execute miRPlant, type "miRPlant.sh" in command line:
>miRPlant.sh
d. In Windows, double click "miRPlant.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: "demo.fastq".
g. miRPlant 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.

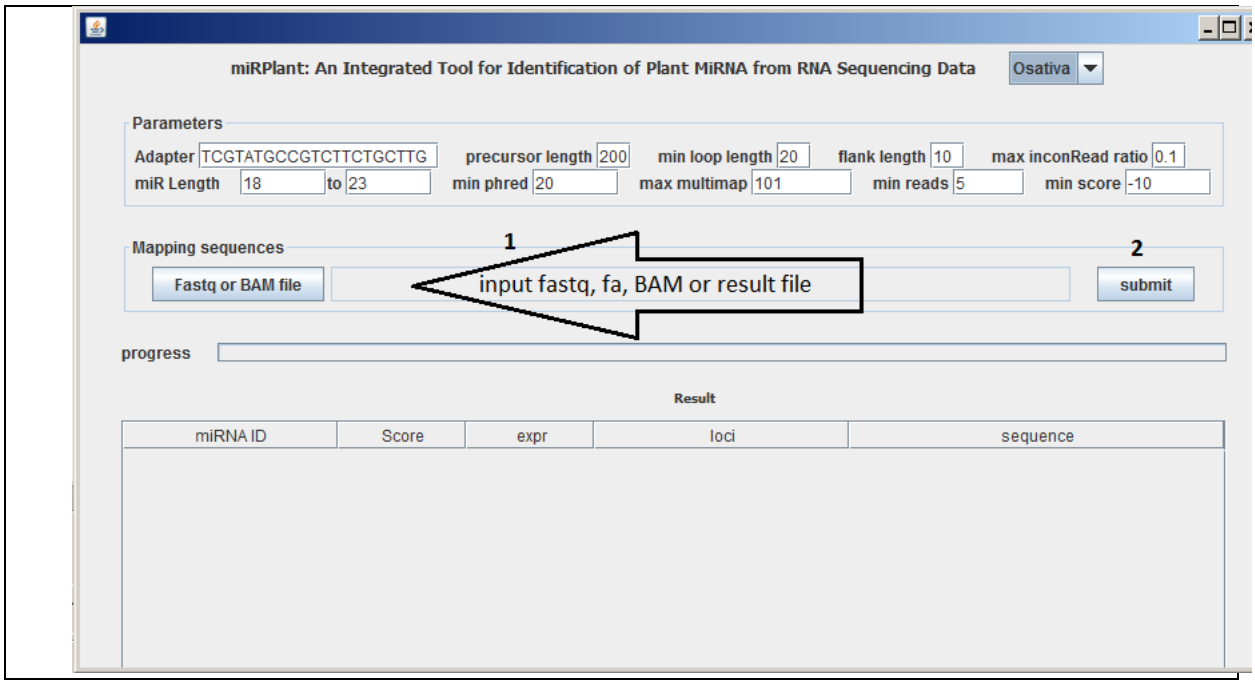
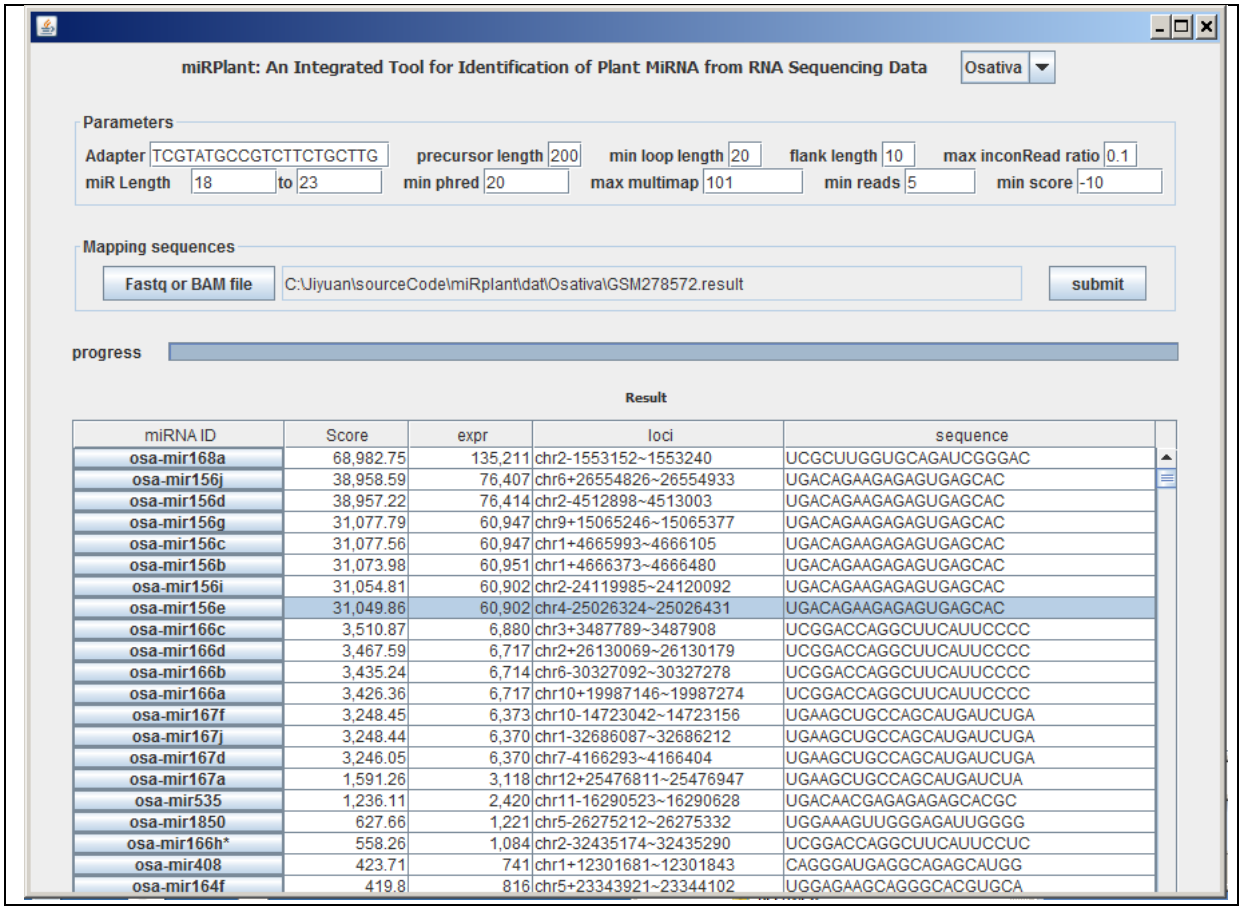



**Figure 2 screen shot of miRPlant**

## 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.




| miRNA ID | Score | expr | loci | sequence |
|---|---|---|---|---|
| osa-mir168a | 68,982.75 | 135,211 | chr2-1553152~1553240 | UCGCUUGGUGCAGAUCGGGAC |
| osa-mir156j | 38,958.59 | 76,407 | chr6+26554826~26554933 | UGACAGAAGAGAGUGAGCAC |
| osa-mir156d | 38,957.22 | 76,414 | chr2-4512898~4513003 | UGACAGAAGAGAGUGAGCAC |
| osa-mir156g | 31,077.79 | 60,947 | chr9+15065246~15065377 | UGACAGAAGAGAGUGAGCAC |
| osa-mir156c | 31,077.56 | 60,947 | chr1+4665993~4666105 | UGACAGAAGAGAGUGAGCAC |
| osa-mir156b | 31,073.98 | 60,951 | chr1+4666373~4666480 | UGACAGAAGAGAGUGAGCAC |
| osa-mir156i | 31,054.81 | 60,902 | chr2-24119985~24120092 | UGACAGAAGAGAGUGAGCAC |
| osa-mir156e | 31,049.86 | 60,902 | chr4-25026324~25026431 | UGACAGAAGAGAGUGAGCAC |
| osa-mir166c | 3,510.87 | 6,880 | chr3+3487789~3487908 | UCGGACCAGGCUUCAUUCCCC |
| osa-mir166d | 3,467.59 | 6,717 | chr2+26130069~26130179 | UCGGACCAGGCUUCAUUCCCC |
| osa-mir166b | 3,435.24 | 6,714 | chr6-30327092~30327278 | UCGGACCAGGCUUCAUUCCCC |
| osa-mir166a | 3,426.36 | 6,717 | chr10+19987146~19987274 | UCGGACCAGGCUUCAUUCCCC |
| osa-mir167f | 3,248.45 | 6,373 | chr10-14723042~14723156 | UGAAGCUGCCAGCAUGAUCUGA |
| osa-mir167j | 3,248.44 | 6,370 | chr1-32686087~32686212 | UGAAGCUGCCAGCAUGAUCUGA |
| osa-mir167d | 3,246.05 | 6,370 | chr7-4166293~4166404 | UGAAGCUGCCAGCAUGAUCUGA |
| osa-mir167a | 1,591.26 | 3,118 | chr12+25476811~25476947 | UGAAGCUGCCAGCAUGAUCUA |
| osa-mir535 | 1,236.11 | 2,420 | chr11-16290523~16290628 | UGACAACGAGAGAGAGCACGC |
| osa-mir1850 | 627.66 | 1,221 | chr5-26275212~26275332 | UGGAAAGUUGGGAGAUUGGGG |
| osa-mir166h* | 558.26 | 1,084 | chr2-32435174~32435290 | UCGGACCAGGCUUCAUUCCUC |
| osa-mir408 | 423.71 | 741 | chr1+12301681~12301843 | CAGGGAUGAGGCAGAGCAUGG |
| osa-mir164f | 419.8 | 816 | chr5+23343921~23344102 | UGGAGAAGCAGGGCACGUGCA |

**Figure 3 result screen of miRPlant**

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

a. 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.

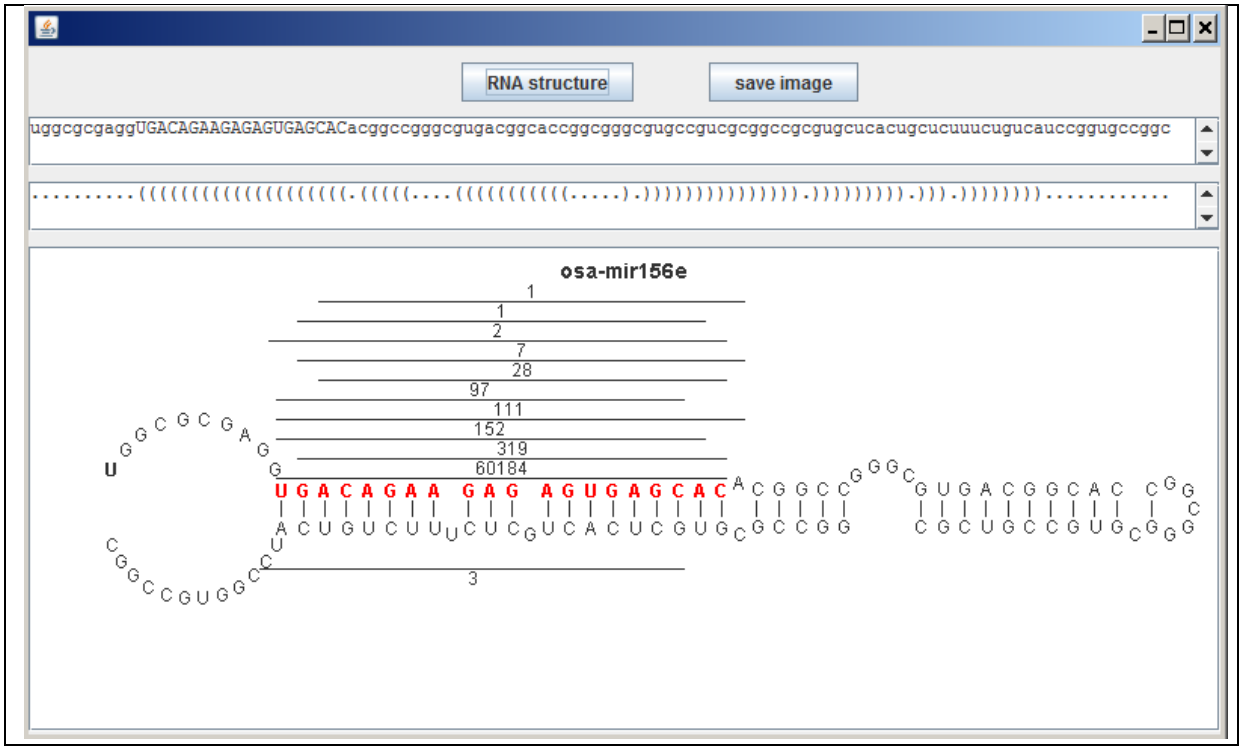



**Figure 4 HairPin structure**

## 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
ATGGGTTTGCAGTCCTCGGTTTAAAAAAAAGACGTC
+
SNSNISPSNENSSPINNESNSNIIIOIIONAAAAA<
```

### 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

miRPlant 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

miRPlant saves the result file as "XXX.result". You can load this file in miRPlant to show the structure.

## 3. miRPlant used in other species

### 3.1 Change genome assembly Osativa, or TAIR9 (Osativa has already become default assembly in miRPlant_v1)

### 3.2 Other species

Download and unzip build_bwt_idx_v31.zip from http://sourceforge.net/mirplant. follow the readme to create index files. Copy created index file folder into miRPlant/genome. the new species genome will be shown in genome combobox. Please notice: you need to replace knownMiR.gff3 in ../miRPlant/genome/xxx(assembly)/miRBase/ with corresponding genome's know miRNA gff file. The known miRNA gff file can be downloaded from http://www.mirbase.org/ftp.shtml. Its file name HAS TO BE RENAMED as knownMiR.gff3. for example

>mv ath.gff3 knownMiR.gff3

.