Protocol of miRdeep2

By: Yang Frank Yang
28-01-2016

Introduction:

We wish to analyse a deep sequencing data of C. elegans and mapping to a region on *C. elegans* chromosome I for known and novel miRNA genes.

Preliminary files:

1. reads.fa

> *The fasta file with the deep sequence reads.*
> Here we choose a small RNA deep sequencing data of *C. elegans* (GSM1585529) from GEO. The supplementary file is **raw** and **processed** so we don't need to clip the adapter by miRDeep2.

2. mature.fa

> *The fasta file with the reference miRBase mature miRNAs for the species.*
> Here we choose the all the mature miRNAs on chrI of *C. elegans* from miRBase Release 21.

3. precursors.fa

> *The fasta file with the reference miRBase precursor miRNAs for the species.*
> Here we choose the all the precursor miRNAs on chrI of *C. elegans* from miRBase Release 21.

4. genome.fa

> *The fasta file with the reference genome.*
> Here we choose the chromosome I region of *C.elegans* genome version WS220. The genome file can be downloaded from wormBase FTP server.

Analysis:

Before the analysis you should put all 4 files in one folder and open the directory by terminal in Linux operating system.

> **Tools version:**
> Bowtie version 1.1.2
> RNAfold(ViennaRNA) version 2.1.9
> squid version 1.9g
> randfold version 2.0

STEP 1:
Build an index of the genome with the command line:
`bowtie-build genome.fa genome`

STEP 2:
Process reads and map them to the genome with the command line:
`mapper.pl reads.fa -c -j -l 18 -m -p genome -s reads_processed.fa -t reads_mapped.arf -v`

The options of mapper.pl were described in miRDeep2 documentation.

STEP 3:
Identification of known and novel miRNAs in the deep sequencing data with the command line:
`miRDeep2.pl reads_processed.fa genome.fa reads_mapped.arf mature.fa none precursors.fa -t C.elegans 2> report.log`

The options of miRDeep2.pl were described in miRDeep2 documentation.

STEP 4:
Browse the results. The result is named as "result_time.HTML" and can be opened by an internet browser.

Quote:

- The quantifier.pl is not necessary for identification.
- Due to the different version of miRDeep2 tools (Bowtie, RNAfold, etc...). It may show different results of analysis with very little differences.