miRge3

Release 0.0.1

Arun H. Patil and Marc K. Halushka

Jul 17, 2023

CONTENTS

An update to Python package to perform comprehensive analysis of small RNA sequencing data, including miRNA annotation, A-to-I editing, novel miRNA detection, isomiR analysis, visualization through IGV, processing Unique Molecular Identifieres (UMI), tRF detection and producing interactive graphical output.

miRge3.0 is developed in python v3.8 and is a recent update of our previous version miRge2.0. This build includes command line interface (CLI) and cross-platform Graphical User Interface (GUI). For more details refer to documentation link below.

CHAPTER

ONE

LINKS

- Documentation
- miRge3.0 Libraries
- Source code
- Frequently asked questions
- Report an issue
- Project page on PyPI

CHAPTER

TWO

CITATION

Arun H Patil, Marc K Halushka. **miRge3.0: a comprehensive microRNA and tRF sequencing analysis pipeline**. NAR Genomics and Bioinformatics. 2021.

CHAPTER

THREE

TABLE OF CONTENTS

3.1 Installation

3.1.1 Docker - biocontainers

For users who prefer docker, can obtain a docker image at Biocontainers

3.1.2 Linux OS

Welcome to installation protocol for Linux OS

Install python3.8 and R

This installation protocol is based on Ubuntu, please use the commands that suit your Linux distribution. For example, `apt` should be replaced with `yum` in Fedora/CentOS.

- Search and start the terminal
- Follow the commands to update Ubuntu and install python 3.8 A password will be prompted when you type `sudo`, use the one you have set during Ubuntu (or your distro) installation.

```
sudo apt update
sudo apt install software-properties-common
sudo add-apt-repository ppa:deadsnakes/ppa
sudo apt install python3.8
sudo apt install python3-setuptools
sudo apt install python3-pip
sudo apt install r-base
```

Linux (Ubuntu 18.04) comes with python2.7 installed by default. To use python3.8, creating an alias in .bashrc would do the trick.

Use vim editor if you are familiar using this editor `vi .bashrc` or open the .bashrc using text editor by `gedit .bashrc` and add the following line at the bottom of the text. `alias python=python3.8`

Save and exit. After that type `bash` on the command line -Or- simply, close the terminal.

Installing miRge3.0 with conda

```
conda install -c bioconda mirge3
```

If you want to use your own environment, please follow the instruction here.

Updating miRge3.0 with conda

```
conda update mirge3
```

Installing miRge3.0 with PyPi

First install miRge dependenceis

- Search and start the terminal, execute the command below:

```
python3.8 -m pip install --user cutadapt reportlab==3.5.42 biopython==1.78  scikit-
↪learn==0.23.1  hypothesis==5.15.1 pytest==5.4.2  scipy==1.4.1  matplotlib==3.2.1 ␣
↪joblib==0.15.1  pandas==1.0.3 future==0.18.2
```

If you encounter a WARNING, like below:

```
WARNING: The script cutadapt is installed in '/home/arun/.local/bin' which is not on␣
↪PATH.
  Consider adding this directory to PATH or, if you prefer to suppress this warning, use␣
↪--no-warn-script-location.
```

Then, open a new terminal window or type cd to get to home directory. Add bin folder PATH to the .bashrc, as shown below: Example: `export PATH=$PATH:"/home/arun/.local/bin"` Remeber to add your path /PATH_TO_USERS/bin.

Install miRge3.0 by this simple command

```
python3.8 -m pip install --user mirge3
```

To upgrade miRge3.0

```
python3.8 -m pip install --user --upgrade  mirge3
```

Install additional C-libraries based tools

Install Bowtie

- Search and start the terminal
- Download bowtie

```
wget -O bowtie-1.3.0-linux-x86_64.zip https://sourceforge.net/projects/bowtie-bio/files/
↪bowtie/1.3.0/bowtie-1.3.0-linux-x86_64.zip/download
```

- `unzip bowtie-1.3.0-linux-x86_64.zip`
- `cd bowtie-1.3.0-linux-x86_64`
- `pwd`
  - `/home/arun/software/bowtie-1.3.0-linux-x86_64`
- Add these bowtie binaries to .bashrc as shown below:

```
export PATH=$PATH:"/home/arun/software/bowtie-1.3.0-linux-x86_64"
```

- After that type `bash` on the command line -Or- simply, close the terminal.

Install Samtools

- Search and start the terminal, execute the below command: `sudo apt install samtools`

Install RNA Fold

- Search and start the terminal, execute the following commands:
- wget "https://www.tbi.univie.ac.at/RNA/download/sourcecode/2_4_x/ViennaRNA-2.4.16.tar.gz"
- cd ViennaRNA-2.4.16

```
sudo ./configure
sudo make
sudo make install
```

GUI requirements

Providing system wide access to miRge3.0, cutadapt, bowtie and bowtie-build, please type or (copy and paste) and submit each of the following commands on the terminal: **NOTE:** Make sure to change your path to python bin folder; Replace `/home/arun/.local/` with `/Path on your computer/`.

- Search and start the terminal, execute the following commands:

```
sudo ln -s /home/arun/.local/bin/miRge3.0 /usr/local/bin/miRge3.0
sudo ln -s /home/arun/.local/bin/cutadapt /usr/local/bin/cutadapt
sudo ln -s /home/arun/software/bowtie-1.3.0-linux-x86_64/bowtie /usr/local/bin/bowtie
sudo ln -s /home/arun/software/bowtie-1.3.0-linux-x86_64/bowtie-build /usr/local/bin/
↪bowtie-build
```

(continues on next page)

(continued from previous page)

```
sudo ln -s /home/arun/software/bowtie-1.3.0-linux-x86_64/bowtie-inspect /usr/local/bin/
↪bowtie-inspect
```

Downloading FASTQ files from NCBI:

- Search and start the terminal, follow the commands below:
- wget -c https://ftp-trace.ncbi.nlm.nih.gov/sra/sdk/2.10.8/sratoolkit.2.10.8-ubuntu64.tar.gz
- tar -xvzf sratoolkit.2.10.8-ubuntu64.tar.gz
- cd sratoolkit.2.10.8-ubuntu64/bin
- `pwd`
  - `/home/arun/software/sratoolkit.2.10.8-ubuntu64/bin`
- Add to .bashrc
  - `cd`
  - `vi .bashrc` or `gedit .bashrc` and add the following line at the bottom of the page
  - `export PATH=$PATH:"/home/arun/software/sratoolkit.2.10.8-ubuntu64/bin"`

Save and exit. After that type `bash` on the command line -Or- simply, close the terminal.

vdb-config Please follow these instructions for `vdb-config` here

Downloading FASTQ files, please type the following: `fastq-dump [options] < accession >` Example: fastq-dump SRR772403 SRR772404

Obtaining and installing GUI application

- Download GUI for Linux

Uninstalling miRge3.0

To uninstall open the terminal and type:

```
python3.8 -m pip uninstall mirge3
```

Conda uninstall:

```
conda remove mirge3
```

For more details on conda uninstallation process, click here

3.1.3 macOS

Welcome to installation protocol for Mac OS

System prerequisites

- Search and start the terminal, execute the following commands
- `ruby -e "$(curl -fsSL https://raw.githubusercontent.com/Homebrew/install/master/install)"`
- `brew update`
- `brew install wget`

Install python3.7

Please note, any version other than py3.7 causes error in Mac with multiprocessing, issues-1, issues-2. Download python 3.7.5 from python.org

- Search and start the terminal, execute the following commands

```
wget https://www.python.org/ftp/python/3.7.5/python-3.7.5-macosx10.9.pkg
sudo installer -pkg python-3.7.5-macosx10.9.pkg -target /
```

Mac comes with python2.7 installed by default. To use python3.7, creating an alias in .bash_profile would do the trick Open a new terminal window. Use vim editor if you are familiar using this editor `vi .bash_profile` or open the .bash_profile using text editor by `open -e .bash_profile` and add the following line at the bottom of the text.

```
alias python=python3.7
```

Save and exit. After that type `source ~/.bash_profile` on the command line -Or- simply, close the terminal.

Install R

- Search and start the terminal, execute the following command

```
brew install r
```

Installing miRge3.0 with conda

```
conda install -c bioconda mirge3
```

If you want to use your own environment, please follow the instruction here.

Error: Type `samtools --version` and make sure you don't encounter any `libcrypto.so` errors. If you do encounter, simply reinstall samtools with conda as shown below: `conda install samtools`. If the error still persists, please let us know.

Updating miRge3.0 with conda

```
conda update mirge3
```

Installing miRge3.0 with PyPi

First install miRge dependenceis

- Search and start the terminal, execute the following command

```
python3.7 -m pip install --user cutadapt reportlab==3.5.42 biopython==1.78  scikit-
↪learn==0.23.1  hypothesis==5.15.1 pytest==5.4.2  scipy==1.4.1  matplotlib==3.2.1 ␣
↪joblib==0.15.1  pandas==1.0.3 future==0.18.2
```

If you encounter a `WARNING`, like below:

```
WARNING: The script cutadapt is installed in '/Users/loaneruser/Library/Python/3.7/bin'␣
↪which is not on PATH.
  Consider adding this directory to PATH or, if you prefer to suppress this warning, use␣
↪--no-warn-script-location.
```

Then, open a new terminal window or type `cd` to get to home directory. Add `bin` folder PATH to the .bash_profile, as shown below: Example: `export PATH=$PATH:"/Users/loaneruser/Library/Python/3.7/bin/"` Remeber to add your path `/PATH_TO_USERS/Python/3.7/bin`.

Install miRge3.0 by this simple command

```
python3.7 -m pip install --user  mirge3
```

To upgrade miRge3.0

```
python3.7 -m pip install --user --upgrade  mirge3
```

Install additional C-libraries based tools

Install Bowtie

- Search and start the terminal, execute the following command
- Download bowtie

```
wget -O bowtie-1.3.0-macos-x86_64.zip  https://sourceforge.net/projects/bowtie-bio/files/
↪bowtie/1.3.0/bowtie-1.3.0-macos-x86_64.zip/download
```

- `unzip bowtie-1.3.0-macos-x86_64.zip`
- `cd bowtie-1.3.0-macos-x86_64`
- `pwd`

  - /Users/loaneruser/Software/bowtie-1.3.0-macos-x86_64

- Add these bowtie binaries to .bash_profile as shown below:

```
export PATH=$PATH:"/Users/loaneruser/Software/bowtie-1.3.0-macos-x86_64/"
```

- After that type `source ~/.bash_profile` on the command line -Or- simply, close the terminal.

Install Samtools

- Search and start the terminal, execute the following command `brew install samtools`

Install RNA Fold

- wget https://www.tbi.univie.ac.at/RNA/download/sourcecode/2_4_x/ViennaRNA-2.4.16.tar.gz
- tar -xvzf ViennaRNA-2.4.16.tar.gz
- cd ViennaRNA-2.4.16

```
sudo ./configure
sudo make
sudo make install
```

Downloading FASTQ files from NCBI:

- Search and start the terminal, execute the following command
- wget -c https://ftp-trace.ncbi.nlm.nih.gov/sra/sdk/2.10.8/sratoolkit.2.10.8-mac64.tar.gz
- tar -xvzf sratoolkit.2.10.8-mac64.tar.gz
- cd sratoolkit.2.10.8-mac64/bin
- `pwd`
  - /Users/loaneruser/Software/sratoolkit.2.10.8-mac64/bin
- Add to .bash_profile
  - `cd`
  - `vi .bash_profile` or `open -e .bash_profile` and add the following line at the bottom of the page
  - `export PATH=$PATH:"/Users/loaneruser/Software/sratoolkit.2.10.8-mac64/bin"`

Save and exit. After that type `source ~/.bash_profile` on the command line -Or- simply, close the terminal.

vdb-config Please follow these instructions for `vdb-config` here

Downloading FASTQ files, please type the following: `fastq-dump [options] < accession >` Example: fastq-dump SRR772403 SRR772404

GUI requirements

Providing system wide access to miRge3.0, cutadapt, bowtie and bowtie-build, please type or (copy and paste) and submit each of the following commands on the terminal: **NOTE:** Make sure to change your path to python bin folder; Replace `/Users/loaneruser/Library/` with `/Path on your computer/`.

- Search and start the terminal, execute the following command

```
sudo ln -s /Users/loaneruser/Library/Python/3.7/bin/miRge3.0 /usr/local/bin/miRge3.0
sudo ln -s /Users/loaneruser/Library/Python/3.7/bin/cutadapt /usr/local/bin/cutadapt
sudo ln -s /Users/loaneruser/Software/bowtie-1.3.0-macos-x86_64/bowtie /usr/local/bin/
↪bowtie
sudo ln -s /Users/loaneruser/Software/bowtie-1.3.0-macos-x86_64/bowtie-build /usr/local/
↪bin/bowtie-build
sudo ln -s /Users/loaneruser/Software/bowtie-1.3.0-macos-x86_64/bowtie-inspect /usr/
↪local/bin/bowtie-inspect
```

Obtaining and installing GUI application

- Download GUI for OSX

Uninstalling miRge3.0

To uninstall open the terminal and type:

```
python3.8 -m pip uninstall mirge3
```

Conda uninstall:

```
conda remove mirge3
```

For more details on conda uninstallation process, click here

3.1.4 Windows OS

Welcome to installation protocol for Windows OS

System prerequisites

- Require Windows 10
- Require WSL and Ubuntu 18

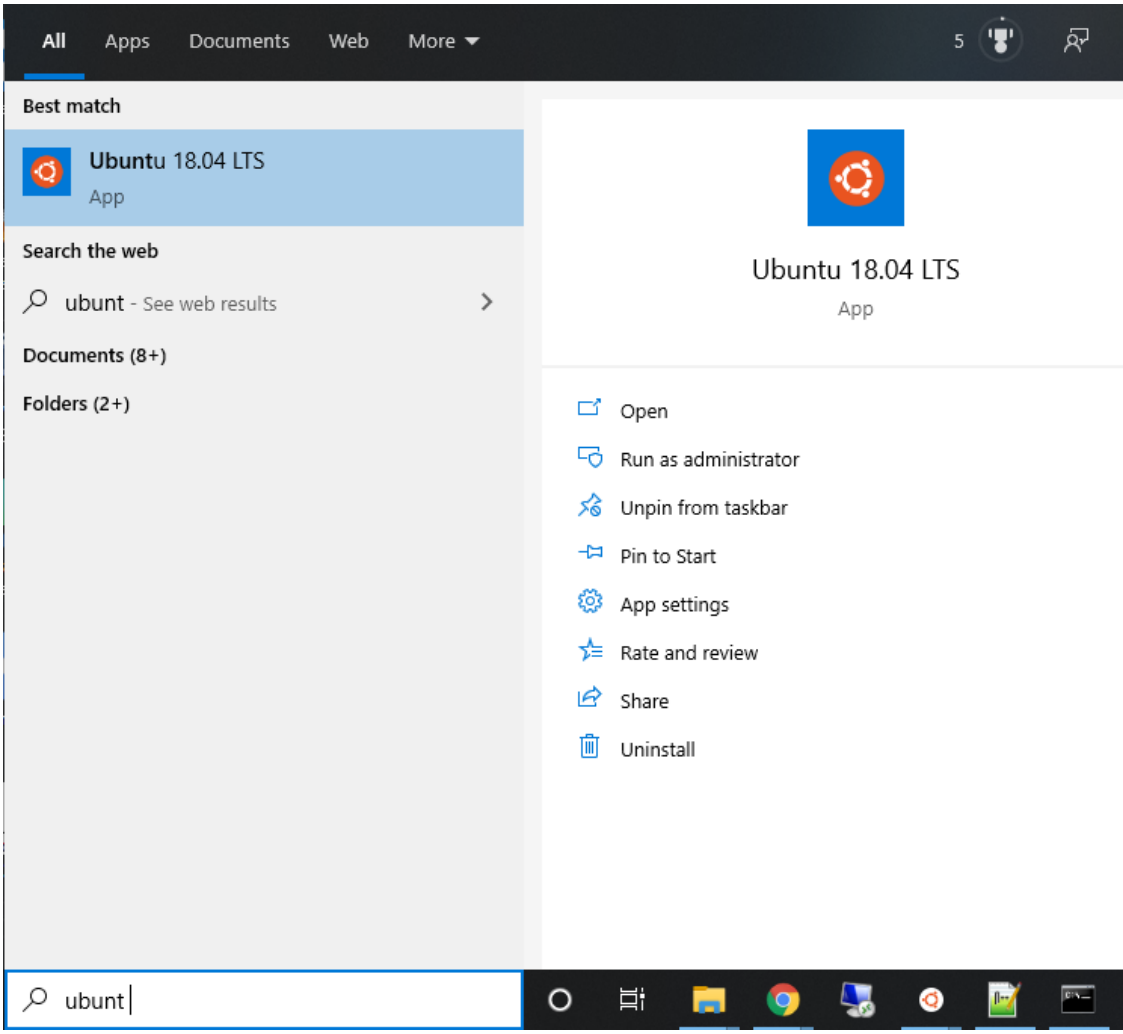
Install WSL

Please follow one of the following guidlines for installing WSL and Ubuntu 18.04 (recommended Ubuntu distribution)

- Quick and easy way
  - TopTechSkills: Watch the first 1:30 seconds, more info.
  - Patreon: Watch the first 4:04 seconds.
- Official windows page.
- `Please remember the password prompted during ubuntu installation` and use when prompted.

Install python3.8 and R

- Search and start Ubuntu



- Follow the commands to update ubuntu and install python 3.8 A password will be prompted when you type sudo, use the one you have set during Ubuntu installation.

```
sudo apt update
sudo apt install software-properties-common
```

(continues on next page)

(continued from previous page)

```
sudo add-apt-repository ppa:deadsnakes/ppa
sudo apt install python3.8
sudo apt install python3-setuptools
sudo apt install python3-pip
sudo apt install r-base
```

Linux (Ubuntu 18.04) comes with python2.7 installed by default. To use python3.8, creating an alias in .bashrc would do the trick

Use vim editor if you are familiar using this editor `vi .bashrc` or open the .bashrc using text editor by `gedit .bashrc` and add the following line at the bottom of the text. `alias python=python3.8`

Save and exit. After that type `bash` on the command line -Or- simply, close the terminal.

Installing miRge3.0 with conda

```
conda install -c bioconda mirge3
```

If you want to use your own environment, please follow the instruction here.

Updating miRge3.0 with conda

```
conda update mirge3
```

Installing miRge3.0 with PyPi

First install miRge dependenceis

- Search and start Ubuntu, execute the following command

```
python3.8 -m pip install --user cutadapt reportlab==3.5.42 biopython==1.78  scikit-
↪learn==0.23.1  hypothesis==5.15.1 pytest==5.4.2  scipy==1.4.1  matplotlib==3.2.1 ␣
↪joblib==0.15.1  pandas==1.0.3 future==0.18.2
```

If you encounter a `WARNING`, like below:

```
WARNING: The script cutadapt is installed in '/home/arun/.local/bin' which is not on␣
↪PATH.
  Consider adding this directory to PATH or, if you prefer to suppress this warning, use␣
↪--no-warn-script-location.
```

Then, open a new terminal window or type `cd` to get to home directory. Add `bin` folder PATH to the .bashrc, as shown below: Example: `export PATH=$PATH:"/home/arun/.local/bin"` Remeber to add your path `/PATH_TO_USERS/bin`.

Install miRge3.0 by this simple command

```
python3.8 -m pip install --user mirge3
```

To upgrade miRge3.0

```
python3.8 -m pip install --user --upgrade  mirge3
```

Install additional C-libraries based tools

Install Bowtie

- Search and start Ubuntu, execute the following command
- Download bowtie

```
wget -O bowtie-1.3.0-linux-x86_64.zip https://sourceforge.net/projects/bowtie-bio/files/
↪bowtie/1.3.0/bowtie-1.3.0-linux-x86_64.zip/download
```

- `unzip bowtie-1.3.0-linux-x86_64.zip`
- `cd bowtie-1.3.0-linux-x86_64.zip`
- `pwd`
  - `/home/arun/software/bowtie-1.3.0-linux-x86_64`
- Add these bowtie binaries to .bashrc as shown below:

```
export PATH=$PATH:"/home/arun/software/bowtie-1.3.0-linux-x86_64"
```

- After that type `bash` on the command line -Or- simply, close the terminal.

Install Samtools

- Search and start Ubuntu, execute the following command `sudo apt install samtools`

Install RNA Fold

- wget "https://www.tbi.univie.ac.at/RNA/download/sourcecode/2_4_x/ViennaRNA-2.4.16.tar.gz"
- cd ViennaRNA-2.4.16

```
sudo ./configure
sudo make
sudo make install
```

GUI requirements

Providing system wide access to miRge3.0, cutadapt, bowtie and bowtie-build, please type or (copy and paste) and submit each of the following commands on the terminal: **NOTE:** Make sure to change your path to python bin folder; Replace `/home/arun/.local/` with `/Path on your computer/`.

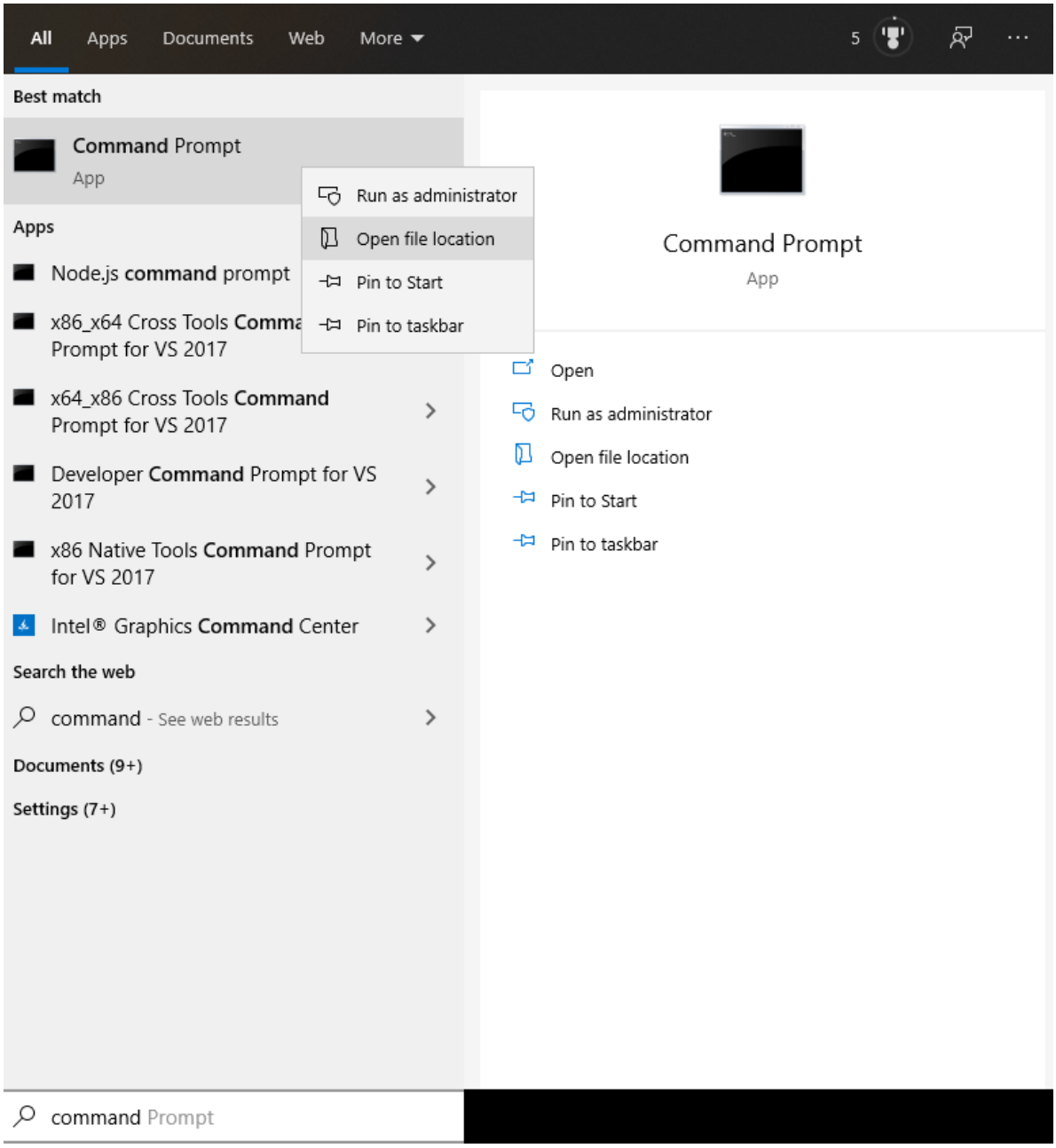
- Search and start Ubuntu, execute the following command

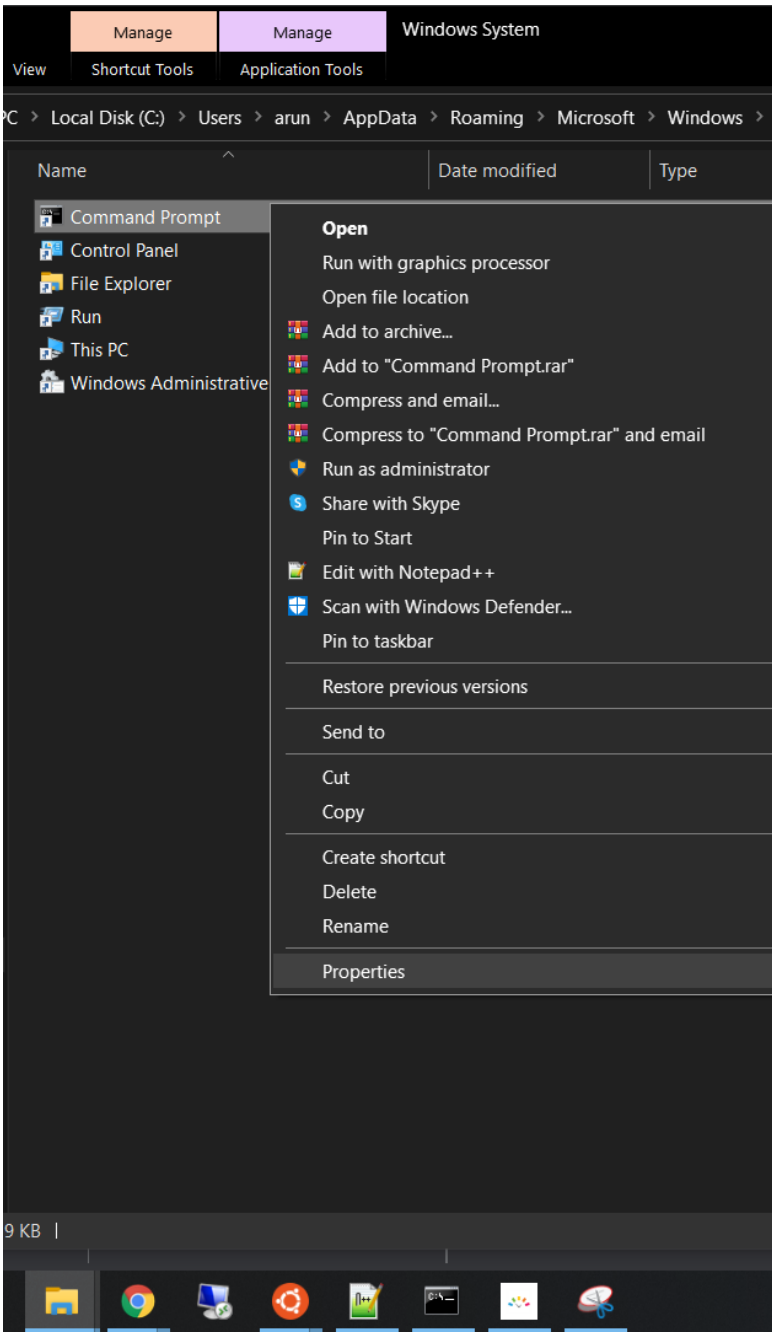
```
sudo ln -s /home/arun/.local/bin/miRge3.0 /usr/local/bin/miRge3.0
sudo ln -s /home/arun/.local/bin/cutadapt /usr/local/bin/cutadapt
sudo ln -s /home/arun/software/bowtie-1.3.0-linux-x86_64/bowtie /usr/local/bin/bowtie
sudo ln -s /home/arun/software/bowtie-1.3.0-linux-x86_64/bowtie-build /usr/local/bin/
↪bowtie-build
sudo ln -s /home/arun/software/bowtie-1.3.0-linux-x86_64/bowtie-inspect /usr/local/bin/
↪bowtie-inspect
```

Change Command Prompt Properties

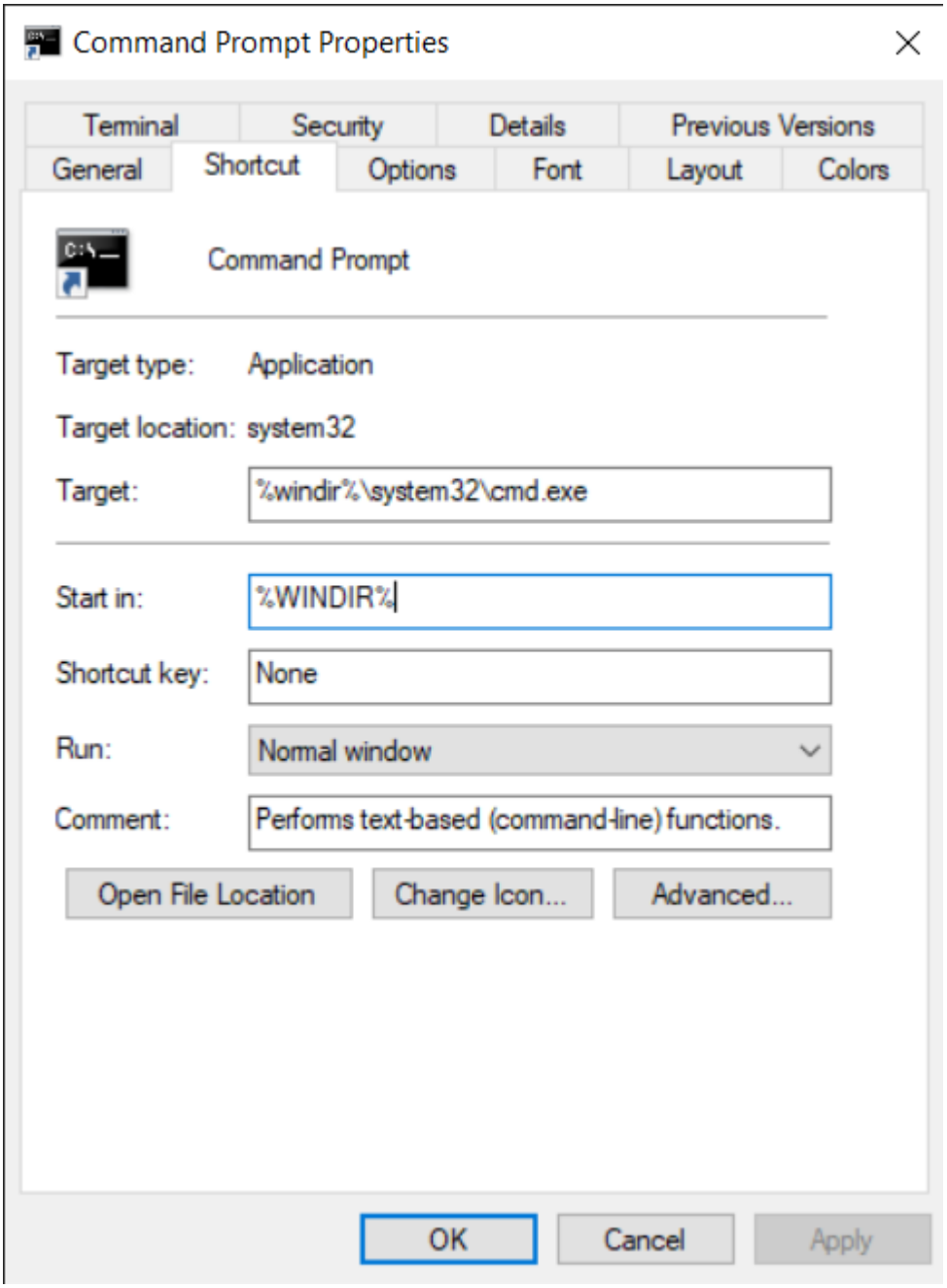
One last thing to avoid an error `The directory name is invalid`:

- Type cmd in Windows search box, right-click on Command Prompt and select Open file location.


All
Apps
Documents
Web
More
5
Best match
Command Prompt
App
Apps
Node.js command prompt
x86_x64 Cross Tools Comma
Prompt for VS 2017
x64_x86 Cross Tools Command
Prompt for VS 2017
Developer Command Prompt for VS
2017
x86 Native Tools Command Prompt
for VS 2017
Intel® Graphics Command Center
Search the web
command - See web results
Documents (9+)
Settings (7+)
Run as administrator
Open file location
Pin to Start
Pin to taskbar
Command Prompt
App
Open
Run as administrator
Open file location
Pin to Start
Pin to taskbar
command Prompt



- Right-click on Command Prompt and click on Properties.
- Under the Shortcut tab, replace `Start in` option by changing the value `%HOMEDRIVE%%HOMEPATH%` to

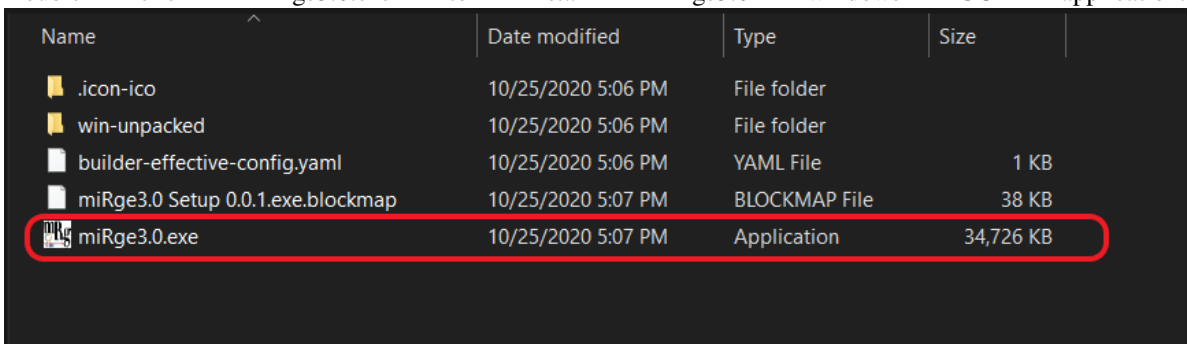


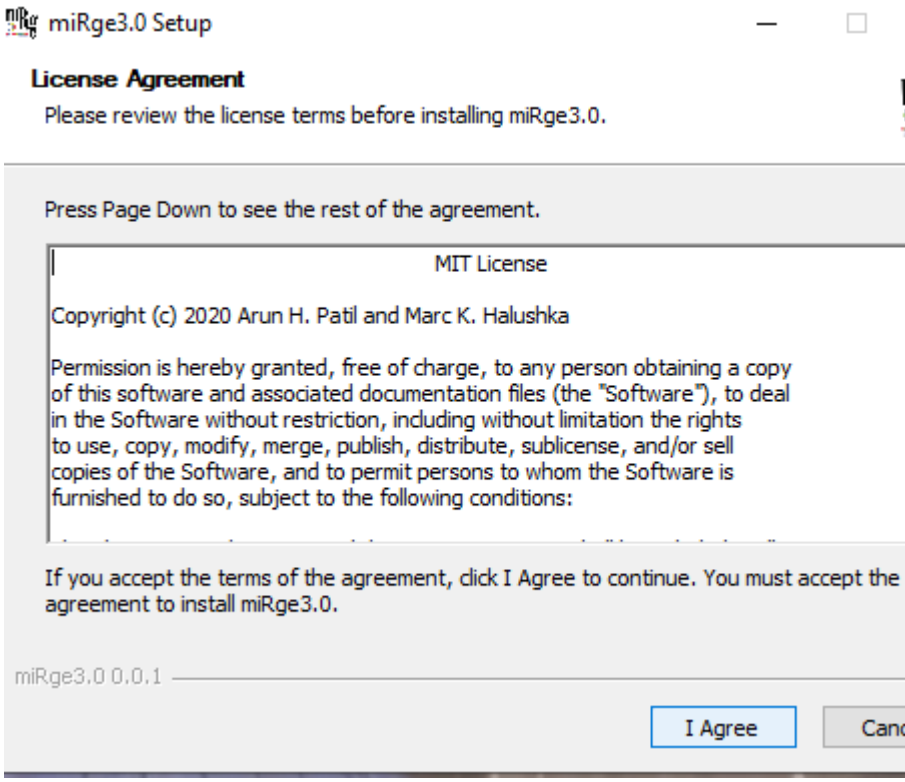
`%WINDIR%`. Click OK.

– Reference 1. Stellarinfo 2. Microsoft

Obtaining and installing GUI application

- Download GUI for Windows 10
- Double click miRge3.0.exe to install miRge3.0 windows GUI application.





- Click Next to complete miRge3.0 installation

Uninstalling miRge3.0

- Step 1: To uninstall open the terminal and type:

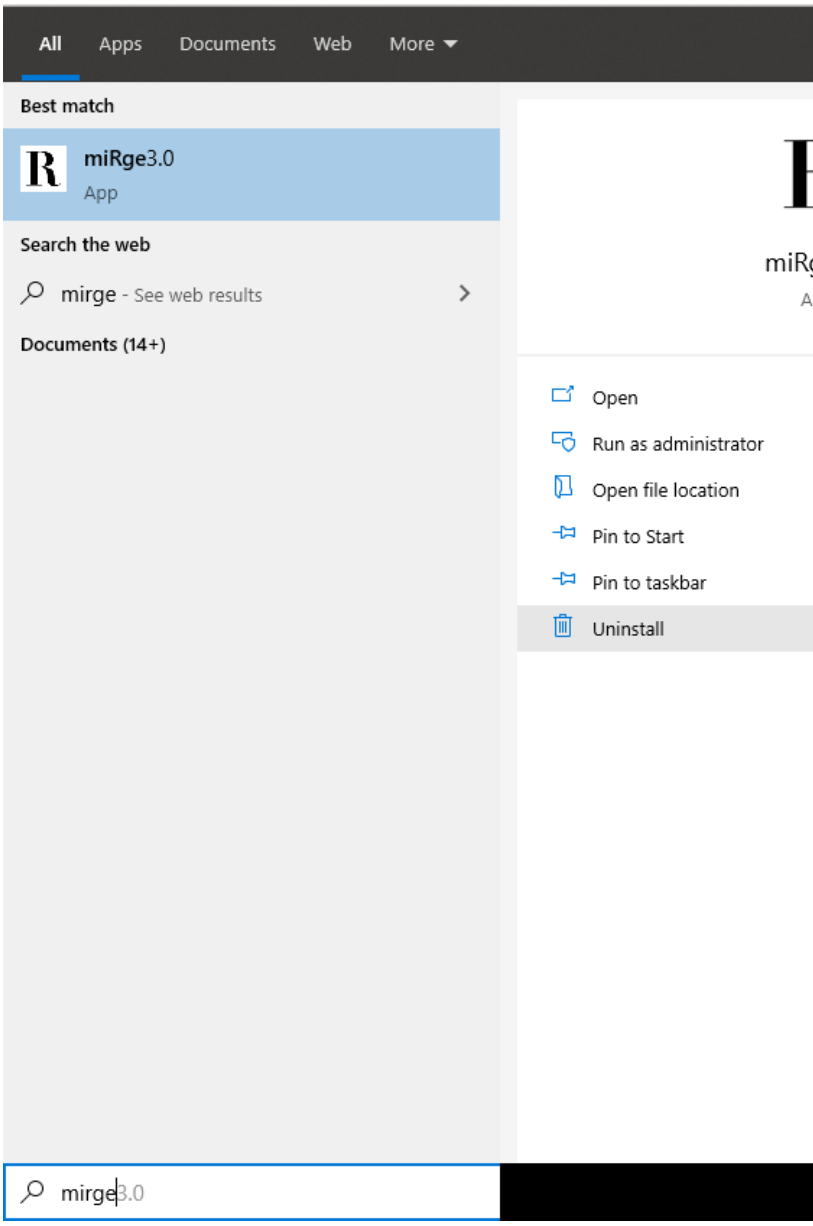
```
python3.8 -m pip uninstall mirge3
```

Conda uninstall:

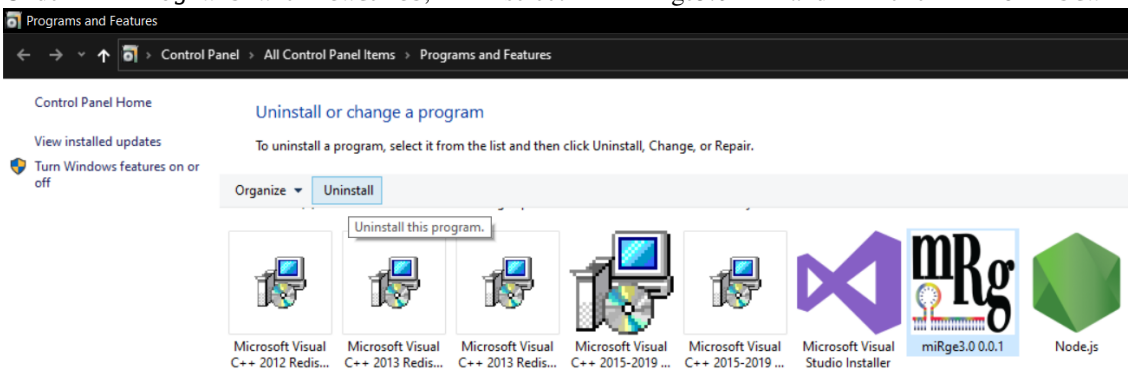
```
conda remove mirge3
```

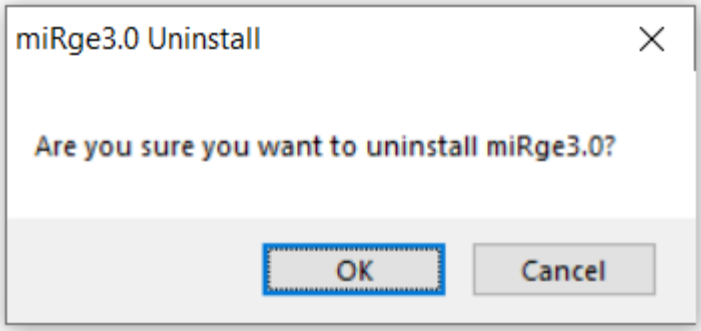
For more details on conda uninstallation process, click here

- Step 2:



- Search miRge3.0, right click and select uninstall
- Under `Programs and Features`, select miRge3.0 and click `Uninstall`





– Then select `Uninstall` by clicking `Ok`. Done.

3.2 User guide

3.2.1 Parameters

To view command-line parameters type `miRge3.0 -h`:

```
usage: miRge3.0 [options]

miRge3.0 (Comprehensive analysis of small RNA sequencing Data)

optional arguments:
  -h, --help  show this help message and exit
  --version   show program's version number and exit

Options:
  -s,   --samples            list of one or more samples separated by comma or a file␣
↪with list of samples separated by new line (accepts *.fastq, *.fastq.gz)
  -db,  --mir-DB             the reference database of miRNA. Options: miRBase and␣
↪miRGeneDB (Default: miRBase)
  -lib, --libraries-path     the path to miRge libraries
  -on,  --organism-name      the organism name can be human, mouse, fruitfly, nematode,␣
↪rat or zebrafish
  -ex,  --crThreshold        the threshold of the proportion of canonical reads for the␣
↪miRNAs to retain. Range for ex (0 - 0.5), (Default: 0.1)
  -phr, --phred64            phred64 format (Default: 33)
  -spk, --spikeIn            switch to annotate spike-ins if spike-in bowtie index␣
↪files are located at the path of bowtie's index files (Default: off)
  -ie,  --isoform-entropy    switch to calculate isomir entropy (default: off)
  -cpu, --threads            the number of processors to use for trimming, qc, and␣
↪alignment (Default: 1)
  -ai,  --AtoI               switch to calculate A to I editing (Default: off)
  -tcf  --tcf-out            switch to write trimmed and collapsed fasta file (Default:␣
↪off)
  -gff  --gff-out            switch to output isomiR results in gff format (Default:␣
↪off)
  -bam  --bam-out            switch to output results in bam format (Default: off)
  -trf  --tRNA-frag          switch to analyze tRNA fragment and halves (Default: off)
  -o    --outDir             the directory of the outputs (Default: current directory)
```

(continues on next page)

(continued from previous page)

```
  -dex   --diffex             perform differential expression with DESeq2 (Default: off)
  -mdt   --metadata           the path to metadata file (Default: off, require '.csv'␣
↪file format if -dex is opted)
  -cms   --chunkmbs           chunk memory in megabytes per thread to use during bowtie␣
↪alignment (Default: 256)
  -spl   --save-pkl           save collapsed reads in binary format for later runs␣
↪(Default: off)
  -rr    --resume             resume from collapsed reads (Default: off)
  -shh   --quiet              enable quiet/silent mode, only show warnings and errors␣
↪(Default: off)

Data pre-processing:
  -a,    --adapter            Sequence of a 3' adapter. The adapter and subsequent bases␣
↪are trimmed
  -g,    --front              Sequence of a 5' adapter. The adapter and any preceding␣
↪bases are trimmed
  -u,    --cut                Remove bases from each read. If LENGTH is positive, remove␣
↪bases from the beginning. If LENGTH is negative, remove bases from the end
  -nxt,  --nextseq-trim       NextSeq-specific quality trimming (each read). Trims also␣
↪dark cycles appearing as high-quality G bases
  -q,    --quality-cutoff     Trim low-quality bases from 5' and/or 3' ends of each read␣
↪before adapter removal. If one value is given, only the 3' end is trimmed
                              If two comma-separated cutoffs are given, the 5' end is␣
↪trimmed with the first cutoff, the 3' end with the second
  -l,    --length             Shorten reads to LENGTH. Positive values remove bases at␣
↪the end while negative ones remove bases at the beginning. This and the following
                              modifications are applied after adapter trimming
  -NX,   --trim-n             Trim N's on ends of reads
  -m,    --minimum-length     Discard reads shorter than LEN. (Default: 16)
  -umi,  --uniq-mol-ids       Trim nucleotides of specific length at 5' and 3' ends of␣
↪the read, after adapter trimming. eg: 4,4 or 0,4. (Use -udd to remove PCR duplicates)
  -udd,  --umiDedup           Specifies argument to removes PCR duplicates (Default:␣
↪False); if TRUE it will remove UMI and remove PCR duplicates otherwise it only remove␣
↪UMI and keep the raw counts (Require -umi option)
  -qumi, --qiagenumi          Removes PCR duplicates of reads obtained from Qiagen␣
↪platform (Default: Illumina; "-umi x,y " Required)


miRNA Error Correction:
  microRNA correction method for single base substitutions due to sequencing errors␣
↪(Note: Refines reads at the expense of time)
  -mEC,  --miREC              Enable miRNA error correction (miREC)
  -kh,   --threshold          the value for frequency threshold  (Default kh = 5)
  -ks,   --kmer-start         kmer range start value (k_1, default 15)
  -ke,   --kmer-end           kmer range end value (k_end, default 20)

Predicting novel miRNAs:
  The predictive model for novel miRNA detection is trained on human and mouse!
  -nmir, --novel-miRNA        include prediction of novel miRNAs
  -minl, --minLength          the minimum length of the retained reads for novel miRNA␣
↪detection (default: 16)
  -maxl, --maxLength          the maximum length of the retained reads for novel miRNA␣
```

(continues on next page)

(continued from previous page)

```
↪detection (default: 25)
  -c,    --minReadCounts      the minimum read counts supporting novel miRNA detection␣
↪(default: 2)
  -mloc, --maxMappingLoci     the maximum number of mapping loci for the retained reads␣
↪for novel miRNA detection (default: 3)
  -sl,   --seedLength         the seed length when invoking Bowtie for novel miRNA␣
↪detection (default: 25)
  -olc,  --overlapLenCutoff   the length of overlapped seqence when joining reads into␣
↪longer sequences based on the coordinate
                              on the genome for novel miRNA detection (default: 14)
  -clc,  --clusterLength      the maximum length of the clustered sequences for novel␣
↪miRNA detection (default: 30)

Optional PATH arguments:
  -pbwt, --bowtie-path        the path to system's directory containing bowtie binary
  -psam, --samtools-path      the path to system's directory containing samtools binary
  -prf,  --RNAfold-path       the path to system's directory containing RNAfold binary
```

3.2.2 miRge3.0 libraries

miRge3.0 pipeline aligns the raw reads against a set of small-RNA annotation libraries. The libraries specific to the organism of interest can be obtained from SourceForge. Downloading the libraries on terminal:

Command-line Interface (CLI)

We recommend to create a directory `miRge3_Lib` and download using wget as shown below,

```
mkdir miRge3_Lib
cd miRge3_Lib
wget -O human.tar.gz "https://sourceforge.net/projects/mirge3/files/miRge3_Lib/human.tar.
↪gz/download"
wget -O mouse.tar.gz "https://sourceforge.net/projects/mirge3/files/miRge3_Lib/mouse.tar.
↪gz/download"
wget -O rat.tar.gz "https://sourceforge.net/projects/mirge3/files/miRge3_Lib/rat.tar.gz/
↪download"
wget -O nematode.tar.gz "https://sourceforge.net/projects/mirge3/files/miRge3_Lib/
↪nematode.tar.gz/download"
wget -O fruitfly.tar.gz "https://sourceforge.net/projects/mirge3/files/miRge3_Lib/
↪fruitfly.tar.gz/download"
wget -O zebrafish.tar.gz "https://sourceforge.net/projects/mirge3/files/miRge3_Lib/
↪zebrafish.tar.gz/download"
wget -O hamster.tar.gz "https://sourceforge.net/projects/mirge3/files/miRge3_Lib/hamster.
↪tar.gz/download"
```

Users can download only what is necessary. Unzip the files once downloaded by the following command:

```
tar -xzf human.tar.gz
```

Replace `human` with the organism of interest. If you want to extract all the files at once, you could use `tar -xzf *.tar.gz` instead.

Direct download

If you are having trouble downloading files through SourceForge, please use the direct link to download the library by clicking on links: Human, Mouse, Rat, Zebrafish, Nematode, Fruitfly, Golden Hamster and md5sum.

Graphical User Interface (GUI)

We recommend to create a folder `miRge3_Lib` and download the libraries directly from SourceForge. Once downloaded, extract/unzip the compressed files.

Building new libraries

If you are interested in creating specific library for an organism that is not part of this set then please refer to miRge3_build.

3.2.3 CLI - Example usage

Example command usage:

```
miRge3.0 -s SRR772403.fastq,SRR772404.fastq,SRR772405.fastq,SRR772406.fastq -lib miRge3_
↪Lib -on human -db mirgenedb -o output_dir -gff -nmir -trf -ai -cpu 12 -a illumina
```

Output command line:

```
bowtie version: 1.2.3
Samtools version: 1.7
RNAfold version: 2.4.14
Collecting and validating input files...

miRge3.0 will process 4 out of 4 input file(s).

Cutadapt finished for file SRR772403 in 2.5358 second(s)
Collapsing finished for file SRR772403 in 0.0126 second(s)

Cutadapt finished for file SRR772404 in 7.3542 second(s)
Collapsing finished for file SRR772404 in 0.2786 second(s)

Cutadapt finished for file SRR772405 in 11.0667 second(s)
Collapsing finished for file SRR772405 in 0.8585 second(s)

Cutadapt finished for file SRR772406 in 3.5771 second(s)
Collapsing finished for file SRR772406 in 0.8677 second(s)

Matrix creation finished in 0.3838 second(s)

Data pre-processing completed in 27.2443 second(s)

Alignment in progress ...
Alignment completed in 15.8305 second(s)

Summarizing and tabulating results...
```

(continues on next page)

(continued from previous page)

```
The number of A-to-I editing sites for is less than 10 so that no heatmap is drawn.
Summary completed in 71.4691 second(s)

Predicting novel miRNAs


Performing prediction of novel miRNAs...
Start to predict
Prediction of novel miRNAs Completed (104.83 sec)

The analysis completed in 222.2487 second(s)
```

Test

The test case illustrates the usage of miRge3.0 with a sample dataset, mapping to human reference libraries.

- First download human miRge libraries as shown below:

```
mkdir miRge3_Lib
cd miRge3_Lib
wget -O human.tar.gz "https://sourceforge.net/projects/mirge3/files/miRge3_Lib/human.tar.
↪gz/download"
tar -xzf human.tar.gz
cd ..
```

- Download the sample file from Source Forge, SRR772403

```
You can download to your working directory as shown below:
wget -O SRR772403.fastq.gz "https://sourceforge.net/projects/mirge3/files/test/SRR772403.
↪fastq.gz/download"
```

- Run basic miRge3.0 command to annotate and report isomiRs

```
miRge3.0 -s SRR772403.fastq.gz -lib /mnt/d/Halushka_lab/Arun/miRge3_Lib -a illumina -on␣
↪human -db mirbase -o output_dir -gff -cpu 8

bowtie version: 1.3.0
cutadapt version: 3.1
Samtools version: 1.11
Collecting and validating input files...

miRge3.0 will process 1 out of 1 input file(s).

Cutadapt finished for file SRR772403 in 3.4343 second(s)
Collapsing finished for file SRR772403 in 0.0216 second(s)

Matrix creation finished in 0.0263 second(s)

Data pre-processing completed in 3.5111 second(s)

Alignment in progress ...
Alignment completed in 8.1488 second(s)
```

(continues on next page)

(continued from previous page)

```
Summarizing and tabulating results...
Summary completed in 2.27 second(s)


The analysis completed in 15.2276 second(s)
```

- Output folder, sample output can be accessed here

```
miRge creates a subfolder inside the folder "output_dir" and all the files will be␣
↪stored there. The test output can be accessed at the following link:
https://sourceforge.net/projects/mirge3/files/test/output_dir/miRge.2021-06-25_15-16-58/
```

Trimming both 5' and 3' adapters - Linked adapters

If the data contains adapters at both 5' and 3' ends of the reads and both the adapters need to be removed then you should perform linked adapter trimming. This is part of Cutadapt and more about linked adapters can be found here.

Example:

```
miRge3.0 -s DRR013811.fastq -lib /mnt/d/Halushka_lab/Arun/GTF_Repeats_miRge2to3/miRge3_
↪Lib/revised_hsa  -on human -db mirbase -o output_dir -g "TTAGGC...
↪TGGAATTCTCGGGTGCCAAGGAACTCCAGT"
```

Description of adapter: `"TTAGGC...TGGAATTCTCGGGTGCCAAGGAACTCCAGT"`, where `TTAGGC` is the 5' adapter and `TGGAATTCTCGGGTGCCAAGGAACTCCAGT` is the 3' adapter sequence.

Note: Complete adapter sequence must be provided (mandatory) i.e., simply specifying illumina will not be decoded to its actual adapter sequence. This will `NOT WORK`: -g `"TTAGGC...illumina"` This will `WORK`: -g `"TTAGGC... TGGAATTCTCGGGTGCCAAGGAACTCCAGT"`

Save and resume functions

Saving collapsed reads and accessory files in binary (pickle) format

For researchers interested in trying different parameters without redoing the entire run, the post-collapsed reads datafile can be saved. The parameter `-spl/--save-pkl` (save pickle) should be specified to save the pickle files. By default the internal variables such as the Pandas dataframe containing collapsed reads before alignment, read summary and sample information is saved as two different pickle files namely `collapsed.pkl` for collapsed read counts and `collapsed_accessories.pkl` for accessory files (read summary, sample information etc). An example usage is described below:

```
miRge3.0 -s SRR772403.fastq,SRR772404.fastq -a illumina -lib miRge3_Lib -on human -db␣
↪mirbase -o output_dir -spl
bowtie version: 1.3.0
cutadapt version: 4.1
Samtools version: 1.3.1
Collecting and validating input files...

miRge3.0 will process 2 out of 2 input file(s).
```

(continues on next page)

(continued from previous page)

```
Cutadapt finished for file SRR772403 in 3.8598 second(s)
Collapsing finished for file SRR772403 in 0.0259 second(s)

Cutadapt finished for file SRR772404 in 13.5832 second(s)
Collapsing finished for file SRR772404 in 0.3531 second(s)

Matrix creation finished in 0.1652 second(s)

Data pre-processing completed in 18.113 second(s)

Alignment in progress ...
Alignment completed in 20.1637 second(s)

Summarizing and tabulating results...
Summary completed in 1.9921 second(s)


The path to output directory: /mnt/d/Halushka_lab/Arun/datasets/output_dir/miRge.2022-07-
↪07_13-59-51

The analysis completed in 43.278 second(s)
```

Resuming from collapsed reads and try out different miRge3.0 parameters

The sample execution previously run with `-spl` option can only be used to resume miRge3.0 with different parameters. The sample parameter `-s` takes the path to the previous output folder (specified earlier as `-o`). Include the `-rr`/`--resume` (re-run or resume) parameter to indicate that you want to re-run miRge3.0 with different parameters. An example usage is described below:

```
miRge3.0 -s /mnt/d/Halushka_lab/Arun/datasets/output_dir/miRge.2022-07-07_13-59-51 -lib␣
↪miRge3_Lib -on human -db mirbase -o output_dir -rr -gff
bowtie version: 1.3.0
cutadapt version: 4.1
Samtools version: 1.3.1
Collecting and validating input files...

miRge3.0 will process 2 saved run(s) from binary pickle file.

Alignment in progress ...
Alignment completed in 19.9428 second(s)

Summarizing and tabulating results...
Summary completed in 7.6734 second(s)


The path to output directory: /mnt/d/Halushka_lab/Arun/datasets/output_dir/miRge.2022-07-
↪07_14-12-03

The analysis completed in 30.6275 second(s)
```

Running samples with UMI

Qiagen - based UMI

Testing sample data run on UMI obtained from Qiagen platform. Important parameters are (`-umi`, `--qiagenumi` and `-udd`)

```
miRge3.0 -s SRR13077007.fastq -db miRBase -lib miRge3_Lib -on human -a␣
↪AACTGTAGGCACCATCAAT --qiagenumi -umi 0,12 -o output_dir -cpu 10 -udd
```

Please note: As of July, 2021, the standard internal 3' adapter was AACTGTAGGCACCATCAAT ligated to 12 nucleotide UMI sequence followed by external 3' adapter sequence. If you have different internal adapter other than AACTGTAGGCACCATCAAT, then please provide that.

Example of reads, UMI and adapters for hsa-let-7a (sequence left to right in the order mentioned below with-in angular brackets):

<hsa-let-7a-5p: TGAGGTAGTAGGTTGTATAGTT><Internal 3' adapter:AACTGTAGGCACCATCAAT><12 nt UMI><external 3' adapter AGATCGGAAGAGCACACGTCT>

```
TGAGGTAGTAGGTTGTATAGTTAACTGTAGGCACCATCAATGTTAGACCTGCAAGATCGGAAGAGCACACGTCTG
TGAGGTAGTAGGTTGTATAGTTAACTGTAGGCACCATCAATCAATGACGATTTAGATCGGAAGAGCACACGTCTG
TGAGGTAGTAGGTTGTATAGTTAACTGTAGGCACCATCAATAAACAAAGATCCAGATCGGAAGAGCACACGTCTG
TGAGGTAGTAGGTTGTATAGTTAACTGTAGGCACCATCAATCGCATCGCCGACAGATCGGAAGAGCACACGTCTG
TGAGGTAGTAGGTTGTATAGTTAACTGTAGGCACCATCAATTTTGCCATTACTAGATCGGAAGAGCACACGTCTG
```

Illumina - based UMI/4N method

Testing sample data run on UMI/4N obtained from Illumina or similar platform. Important parameters are (`-umi` and `-udd`)

```
miRge3.0 -s SRR6379839.fastq -db miRBase -lib miRge3_Lib -on human -a illumina -umi 4,4 -
↪o output_dir -cpu 10 -udd
```

<04 nt UMI><hsa-let-7a-5p: TGAGGTAGTAGGTTGTATAGTT><04 nt UMI><3' adapter:TGGAATTCTCGGGTGCCAAGGAACTCCAGTCACCGGAATATCTCG>

```
TACATGAGGTAGTAGGTTGTATAGTTCCTCTGGAATTCTCGGGTGCCAAGGAACTCCAGTCACCGGAATATCTCG
TACCTGAGGTAGTAGGTTGTATAGTTACTATGGAATTCTCGGGTGCCAAGGAACTCCAGTCACCGGAATATCTCG
CAGGTGAGGTAGTAGGTTGTATAGTTGGTATGGAATTCTCGGGTGCCAAGGAACTCCAGTCACCGGAATATCTCG
AGAATGAGGTAGTAGGTTGTATAGTTACTATGGAATTCTCGGGTGACAAGGAACTCCAGTCACCGGAATATCTCG
AGGTTGAGGTAGTAGGTTGTATAGTTACTATGGAATTCTCGGGTGCCAAGGAACTCCAGTCACCGGAATATCTCG
```

Performing differential expression analysis

1. Download example datasets from NCBI SRA (Note: Tutorial on how to download SRA files is below).
2. Prepare metadata information in CSV format as shown below. For this tutorial, download the file from here.

```
id,group
SRR8497647,Control
SRR8497648,Control
SRR8497649,Control
```

(continues on next page)

(continued from previous page)

```
SRR8497650,Control
SRR8497651,treated
SRR8497652,treated
SRR8497653,treated
SRR8497654,treated
```

1. Execute the following command:

```
miRge3.0 -s SRR8497647.fastq,SRR8497648.fastq,SRR8497649.fastq,SRR8497650.fastq,
↪SRR8497651.fastq,SRR8497652.fastq,SRR8497653.fastq,SRR8497654.fastq -lib miRge3_Lib -
↪on human -db miRGeneDB -o differential_Exp -a TGGAATTCTCGG -cpu 12 -dex -mdt␣
↪DESmetadata.csv
```

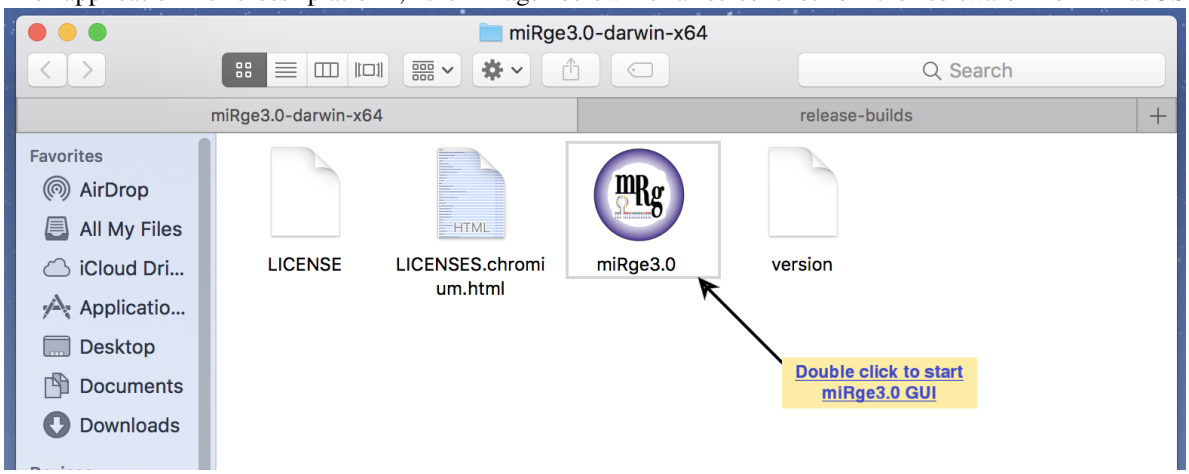
The result files for the above miRge3.0 run can be found at SourceForge

Tutorial on how to download SRA files: This turorial is only brief introduction and doesn't cover all the details of downloading NCBI SRA files. You could find YouTube tutorials on how to download SRA files.

1. Download and install NCBI SRA toolkit: You could refer to NCBI SRA Handbook or GitHub
2. Download command: One could use `fasterq-dump -t temp -e 10 SRR8497647` or simply `fastq-dump SRR8497647`. The only difference being that the fasterq-dump is faster. Similarly, download all other Runs (i.e., SRR8497648, SRR8497649 etc.)

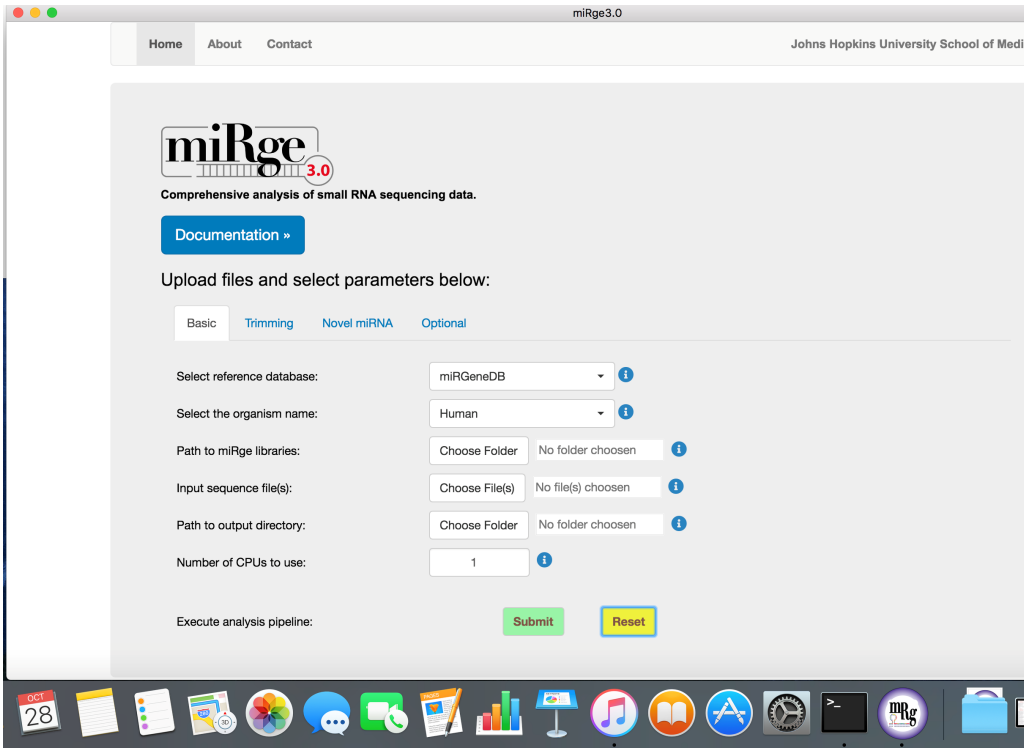
3.2.4 miRge3.0 GUI

- The application is cross platform, the image below is a screenshot of the software from MacOS

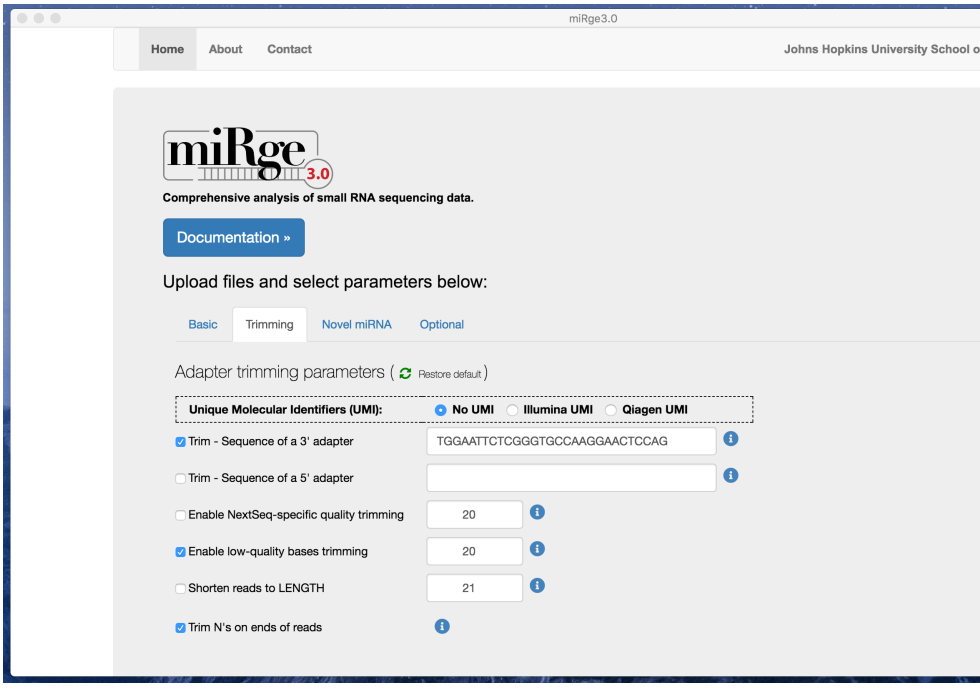


- The software is easy to use with default parameters. The parameters are tabulated into four groups such as basic, trimming parameters, novel miRNA prediction and other optional parameters.

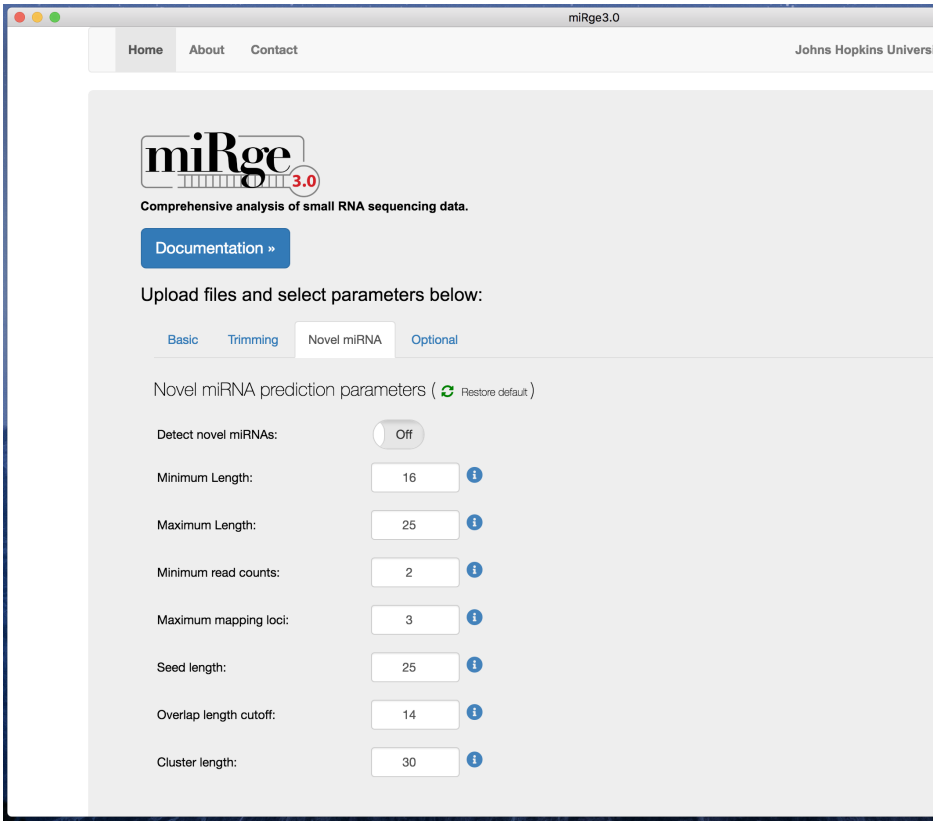
- Screenshot with basic parameters



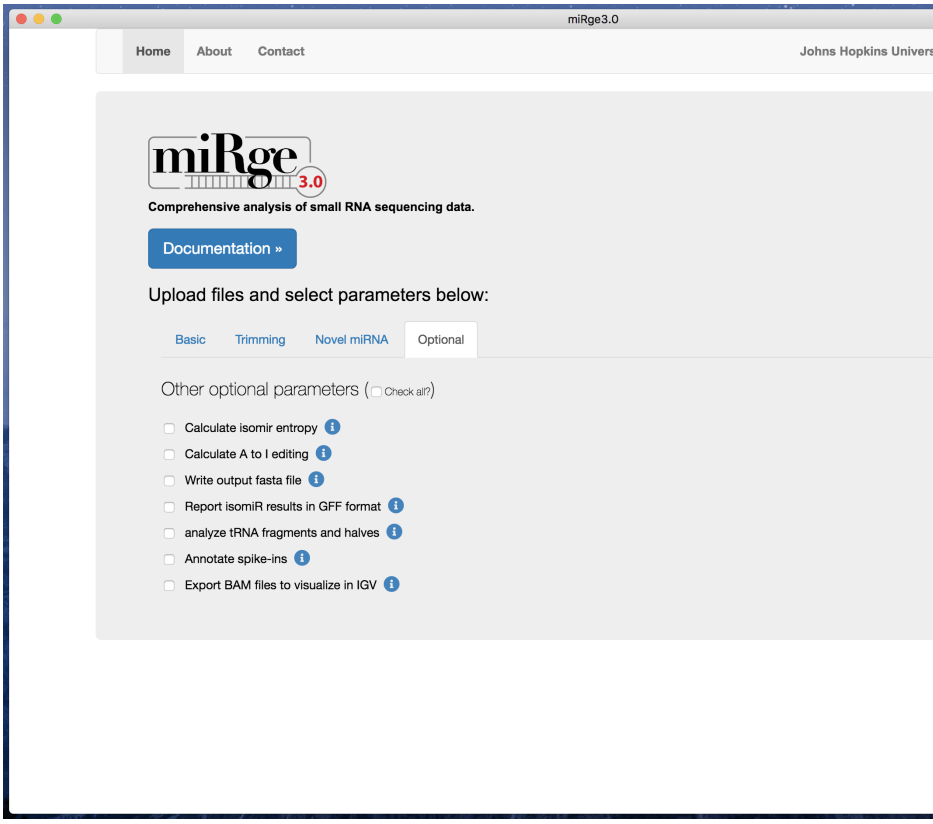
- Screenshot with trimming parameters



- Screenshot with novel miRNA predictions



- Screenshot with other optional parameters



3.2.5 Resources

- Lu, Y., et al., **miRge 2.0 for comprehensive analysis of microRNA sequencing data**. 2018. *BMC Bioinformatics*. PMID.
- Baras, S. A., et al., **miRge - A Multiplexed Method of Processing Small RNA-Seq Data to Determine MicroRNA Entropy**. 2015. *PLoS One*. PMID.

3.3 miRge3.0 output

3.3.1 Command and sample run with UMI datasets

```
miRge3.0 -s SRR8557389.fastq,SRR8557396.fastq,SRR8557398.fastq,SRR8557399.fastq -lib␣
↪miRge3_Lib -on human -db miRGeneDB \
         -o temp -a AACTGTAGGCACCATCAAT -udd --qiagenumi -umi 0,12 -cpu 12 -q 20 -NX -
↪nmir -minl 16 -maxl 25 -c 2 \
         -mloc 3 -sl 25 -olc 14 -clc 30 -gff

bowtie version: 1.2.3
cutadapt version: 2.7
Samtools version: 1.7
RNAfold version: 2.4.14
Collecting and validating input files...

miRge3.0 will process 4 out of 4 input file(s).

Cutadapt finished for file SRR8557389 in 21.0854 second(s)
Collapsing finished for file SRR8557389 in 0.0699 second(s)
Cutadapt finished for file SRR8557396 in 10.305 second(s)
Collapsing finished for file SRR8557396 in 0.6016 second(s)
Cutadapt finished for file SRR8557398 in 10.891 second(s)
Collapsing finished for file SRR8557398 in 0.911 second(s)
Cutadapt finished for file SRR8557399 in 14.2126 second(s)
Collapsing finished for file SRR8557399 in 1.1292 second(s)
Matrix creation finished in 0.4788 second(s)

Data pre-processing completed in 62.762 second(s)

Alignment in progress ...
Alignment completed in 16.9863 second(s)

Summarizing and tabulating results...
Summary completed in 7.8131 second(s)
Predicting novel miRNAs

Performing prediction of novel miRNAs...Start to predictPrediction of novel miRNAs␣
↪Completed (220.35 sec)
The analysis completed in 310.7281 second(s)
```

3.3.2 Output tree structure

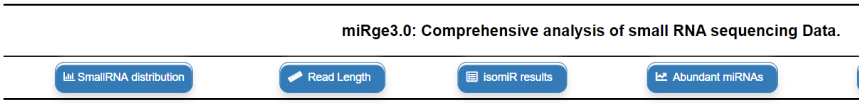
An output directory is created for each run such as `miRge.2020-10-9_1-35-53`, where the name is followed by date time format `miRge.yy-dd-mm-hr-mm-ss`.

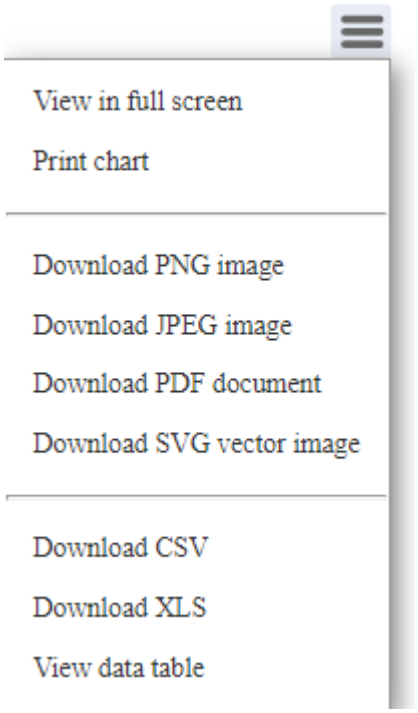
```
The following output is in general, however, the resultant output files are based on the␣
↪options selected during miRge3.0 execution.

miRge.2020-10-9_1-35-53
├── run.log (Gives the detailed log of miRge3.0 execution)
├── unmapped.log (Gives the detailed log of novel miRNA prediction)
├── mapped.csv (CSV file with read counts across each smallRNA library)
├── unmapped.csv (CSV file with unaligned/mapped reads)
├── annotation.report.csv (Basic annotation report with small RNA distribution in CSV␣
↪format)
├── annotation.report.html (Basic annotation report with small RNA distribution in HTML␣
↪format)
├── sample_miRge3.gff (GFF file with reads with isomiRs across one or more samples, if -
↪gff option selected)
├── miR.Counts.csv (miRNA raw read counts across samples)
├── miR.RPM.csv (miRNA Read Per Million - RPM counts across samples)
├── *_umiCounts.csv (Counts for each unique UMI for each sample)
├── index_data.js (Javascript file with data generated for visualization)
├── miRge3_visualization.html (HTML for data visualization)
├── FOLDER_novel_miRNAs
│   ├── *.pdf (novel miRNA structure in PDF format for each miRNA)
│   └── sample_novel_miRNAs_report.csv (Contains list of identified novel miRNAs in CSV␣
↪format)
├── a2IEditing.detail.txt
├── a2IEditing.report.csv
├── a2IEditing.report.newform.csv
├── tRFs.potential.report.tsv
├── tRF.Counts.csv
├── tRF.RP100K.csv
├── tRFs.potential.report.tsv
├── discarded.reads.summary.assigningtRFs.csv
└── tRFs.samples.tmp
    └── *.tRFs.* (Detailed summary of tRFs from each sample)
```

3.3.3 miRge - interactive visualization

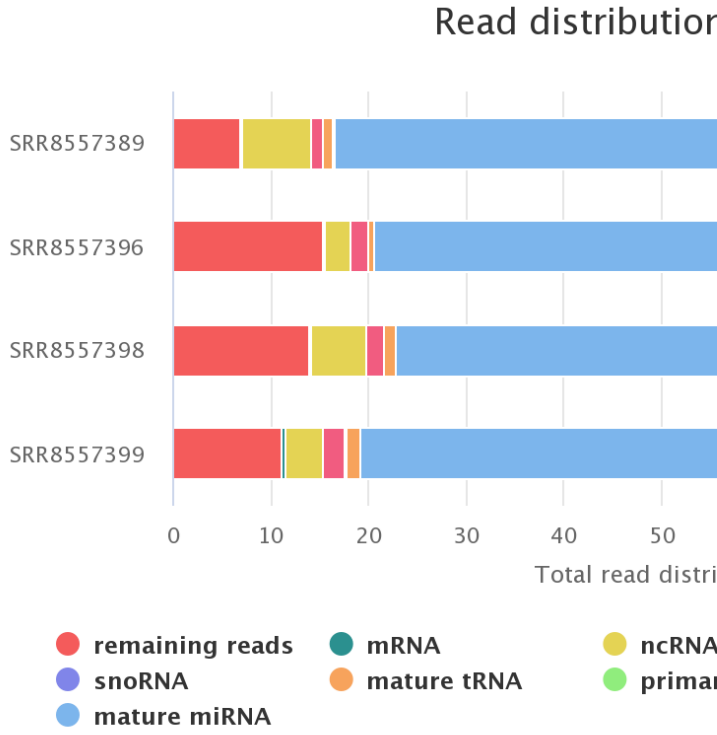
miRge3.0 produces several interactive visualization graphics as follows

- Screenshot of the miRge visualization HTML tab

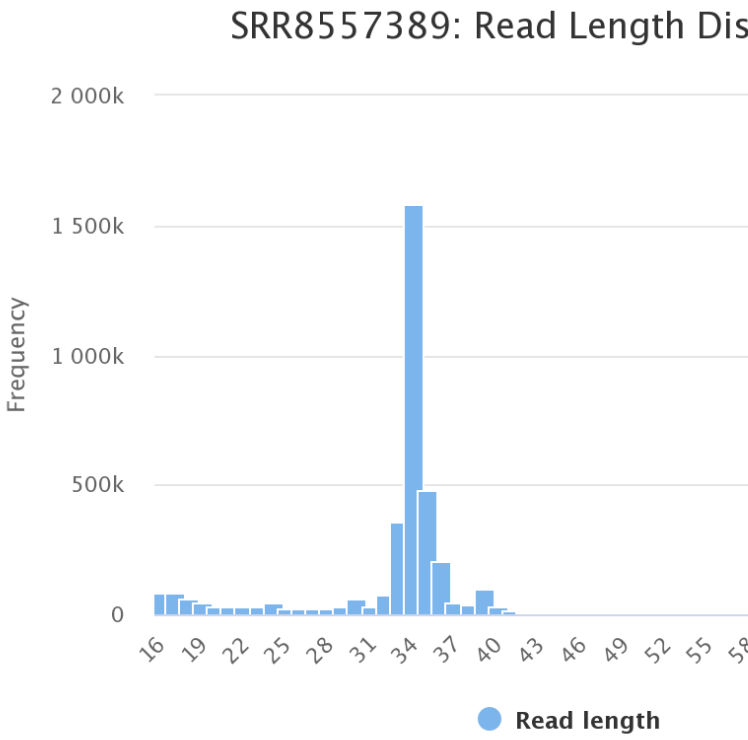




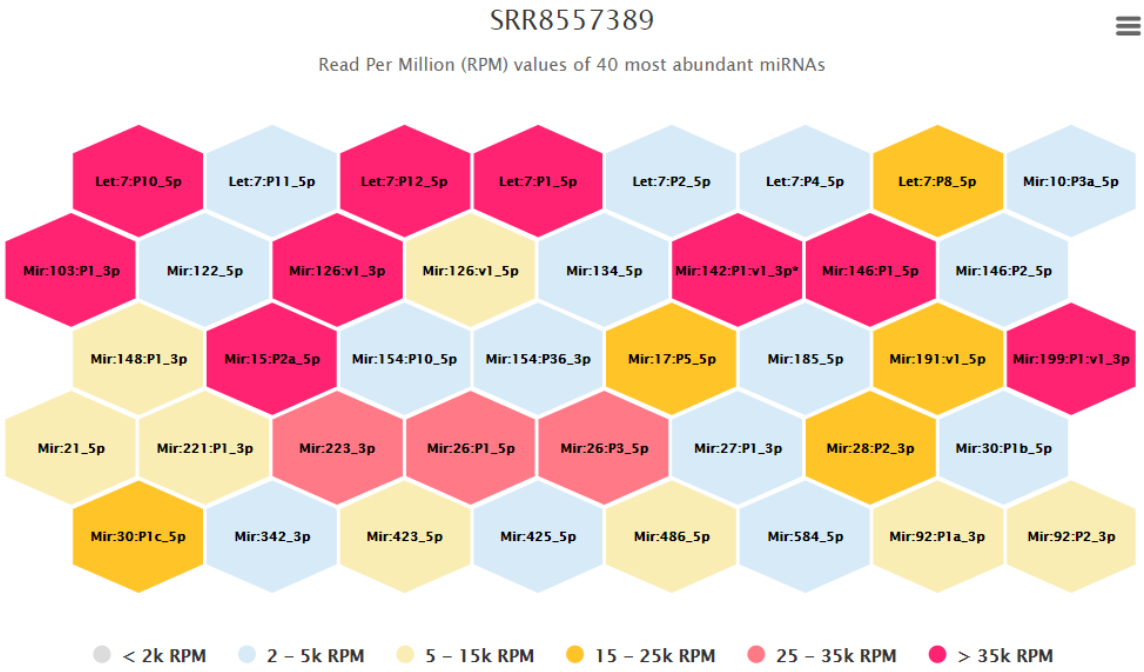
- Chart view and download options



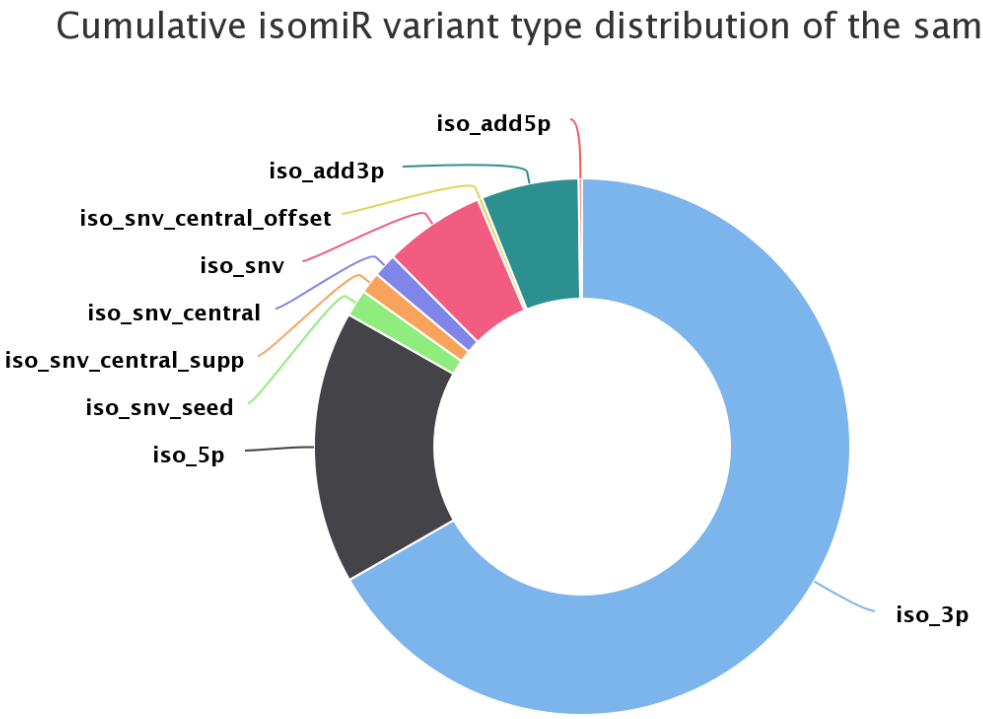
- Screenshot of the smallRNA read distribution for each sample



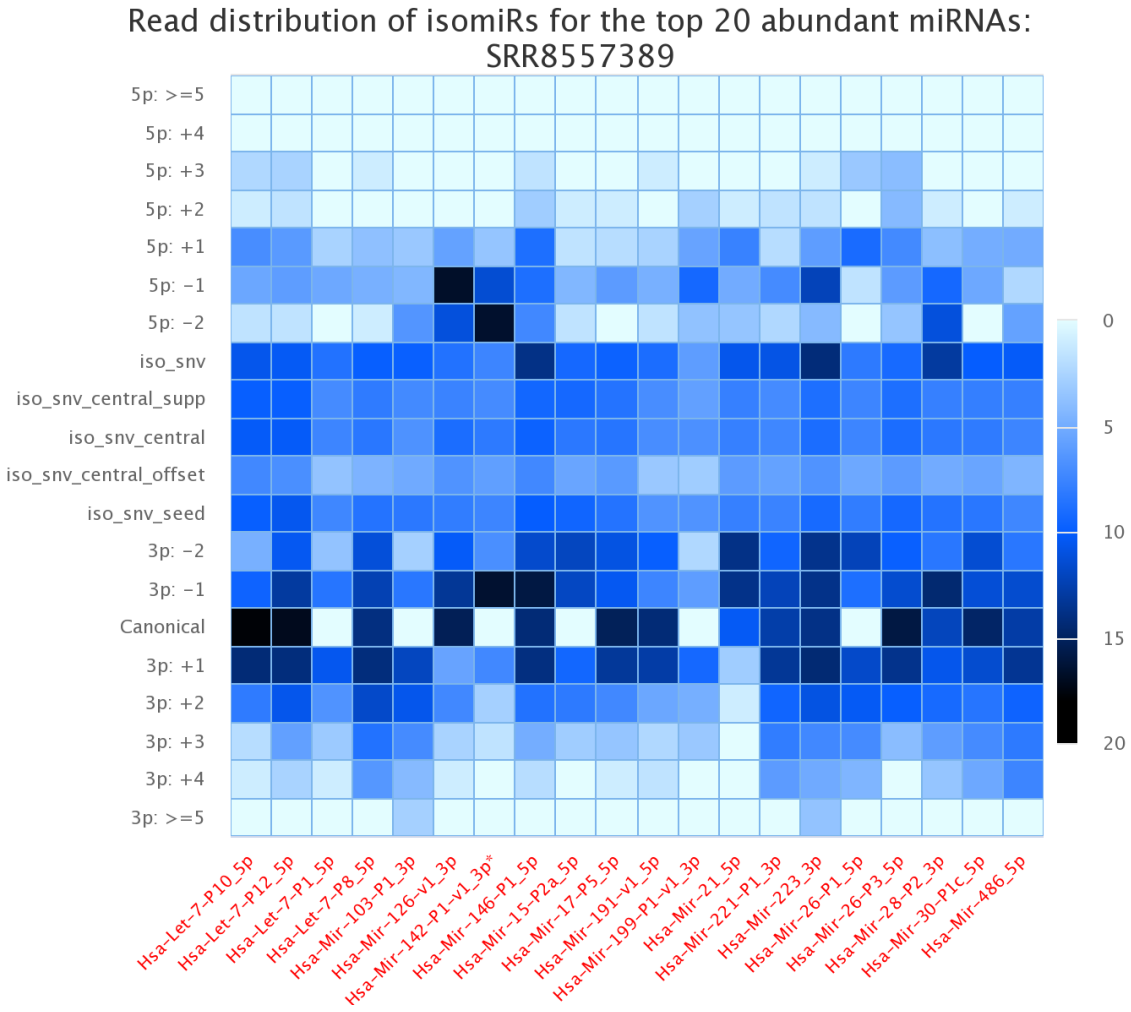
- Screenshot of the read length distribution for each sample
- Screenshot of the tile map representing top 40 high abundant miRNAs for each study



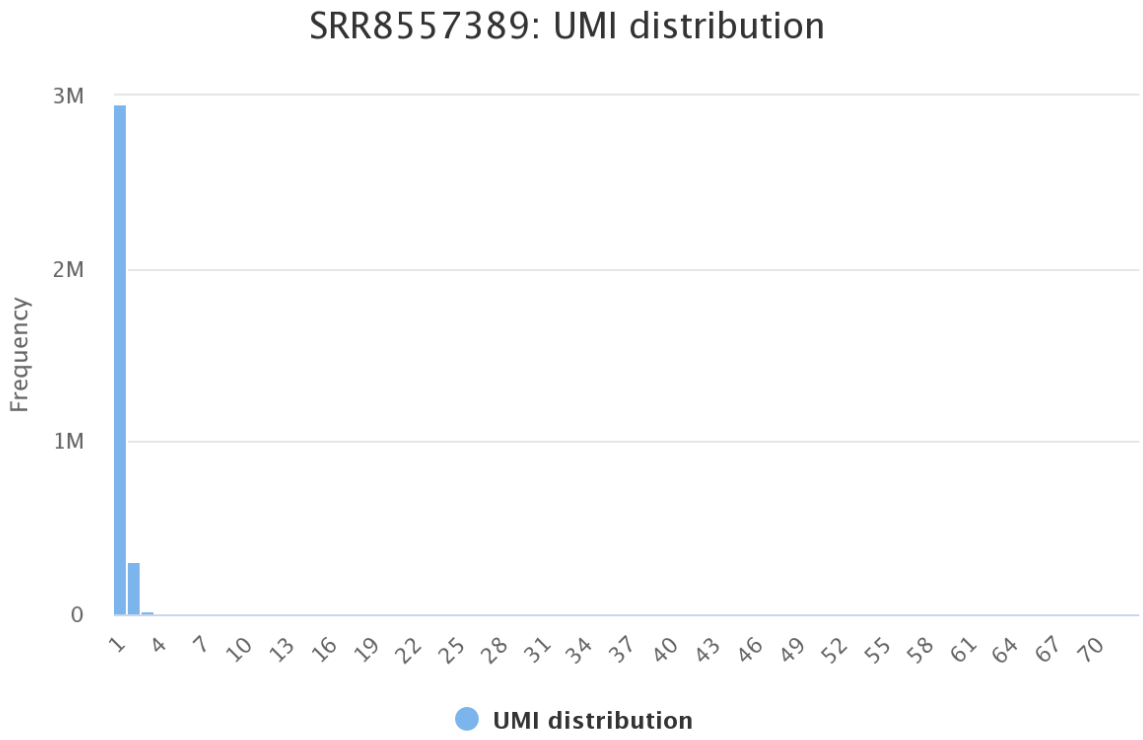
- Screenshot of the variant distribution for all samples combined (isomiRs)



- Screenshot of the heatmap representing variants for each sample for the top 20 high abundant miRNAs (isomiRs)



- Screenshot of the histogram representing UMI counts across each sample



Show 10 entries Search:

id	Name	Probability	Chr	Start pos.	End Pos.
1	SRR8557389_novel_miRNA_1	0.9999987516742271	chr2	134127130	134127
2	SRR8557389_novel_miRNA_2	0.9999915571402658	chr1	172138858	172138
3	SRR8557389_novel_miRNA_3	0.999982270847492	chr3	113594918	113594
4	SRR8557389_novel_miRNA_4	0.9963371234941661	chr8	27433413	274334
5	SRR8557389_novel_miRNA_5	0.9938801463496213	chr17	81125886	811259
6	SRR8557389_novel_miRNA_6	0.9908373862619209	chr17	74748663	747486
7	SRR8557389_novel_miRNA_7	0.9905492996218249	chr12	69584745	695847
8	SRR8557389_novel_miRNA_8	0.9781204295092701	chr22	20086072	200860
9	SRR8557389_novel_miRNA_9	0.9723790135555249	chr5	141849784	141849
10	SRR8557389_novel_miRNA_10	0.9655899863443722	chr19	13836292	138363

Showing 1 to 10 of 26 entries Previous 1

- Screenshot of a list of novel miRNAs identified across samples

3.3.4 Resources:

The graphics for miRge3.0 visualization is enabled with javascripts and CSS obtained from the following:

- Interactive charts from HighCharts
- Icons from Font Awesome
- Interactive HTML table

3.4 Frequently asked questions (FAQ)

We are very greatful and also thankful to all the users of miRge3.0 who rasied GitHub issues in the past that helped us solve few technical problems and improve miRge3.0 functionality further. We expect continued support towards this project. Here we have gathered a few frequently asked questions over the period regarding technical as well as biological/scientific questions. I hope this documentation will be useful as a ready response/solution for your queries.

Before getting started please note; if you don't find a solution to your query in this page then create a new issue and we will get back to you at the earliest. Describe the `Title` *to include the error you are facing e.g.,* `numbpy type error` *and in the* `Comment` *section, it would be best if you could put the command line used, followed by the whole error. (You can delete your file names if you prefer).*

3.4.1 *How to create an issue?*

Click create new issue and in Title: "Please describe the error you think is obvious and will be general for the scientific community to recognize", and Comment: "Give us the maximum information possible regarding the error that you can see on the standard output/terminal"

3.4.2 Frequent questions raised on GitHub:

1. How to use Unique Molecular Identifiers (UMIs)?
2. TypeError: Cannot interpret <attribute 'dtype' of 'numpy.generic' objects> as a data type
3. UnsatisfiableError: bowtie=1.3.0 -> libgcc-ng[version='>=9.3.0'] -> __glibc[version='>=2.17']
4. Is there any way to skip the adaptor trimming process? and how to determine adapter sequence of a Run?
5. How to use and tweak data with Spike-in expirements?
6. How to use -dex DESeq2 analysis?
7. What is the threshold of the proportion of canonical reads (-ex, –crThreshold)?
8. How to input paired-end sequencing data?

How to use Unique Molecular Identifiers (UMIs)?

A detailed documentation for UMI test run is available here. miRge3.0 is designed to process UMIs for Illumina and Qiagen. The parameters to trim UMIs and removing PCR duplicates are different, and also, selecting Qiagen UMI needs an additional parameter.

These following issues were raised:

- #32 (comment)
- #46 (comment)
- #28 (comment)

TypeError: Cannot interpret <attribute 'dtype' of 'numpy.generic' objects> as a data type

I suspect there is a conflict with pandas and numpy in your local machine, I want you to upgrade pandas and try the command again. You can upgrade it as shown (python3.7 if you are using py37) in the following issues:

#20 (comment) #47 (comment)

UnsatisfiableError: bowtie=1.3.0 -> libgcc-ng[version='>=9.3.0'] -> __glibc[version='>=2.17']

The discussion on this issue is available in the following GitHub issue. Thank you @asucrer, for providing solution.

#31 (comment)

```
Solution suggested by the user @asucrer, please follow the steps:

conda create -n mirge     # IMPORTANT to not specify the python version in this step
source activate mirge
conda install -c bioconda mirge3   # Every dependency (including python) is installed
conda install -c bioconda tbb=2020.2    # Solves issue associated to Bowtie installation
conda install -c bioconda openssl=1.0   # Solves issue associated to Samtools␣
↪installation
```

Is there any way to skip the adaptor trimming process? and how to determine adapter sequence of a Run?

miRge3.0 allows users to skip the adapter trimming step, and there are several options on how to provide adapter sequences and the following issue provide a list of adapter sequences for various platforms. [Curation date: January 2020].

- #20 (comment)
- #20 (looped - comment)

Please NOTE: To trim adapter sequences at both ends please follow the documentation Linked-adapters

How to use and tweak data with Spike-in expirements?

An example usage of spike-in libraries and how to add/append spike-in reads of interest to the existing libraries and interpretation is described in the following issues:

- #27 (comment)
- #48 (comment)

How to use -dex DESeq2 analysis?

The documentation for DESeq2 based differentiall expression analysis is available here

The following GitHub issues were raised:

- #41 (comment)
- #33 (comment)

What is the threshold of the proportion of canonical reads (-ex, –crThreshold)?

This was answered to an issue on why default value of 0.1 was chosen for –crThreshold in the following issue.

- #23 (comment)
- #34 (comment)

How to input paired-end sequencing data?

miRge3.0 doesn't annotate paired-end data.

- #7 (comment)

3.5 MIT License