# piFETCH

*version 1.0*

## - Documentation -

### 1. Scope

`piFETCH` is a tiny fetching tool to download data from piRNA cluster database without using the web interface. The latest `piFETCH` version can be found at http://www.smallRNAgroup-mainz.de/software. `piFETCH` allows to download complete proTRAC results for available NCBI SRA datasets or specified information (piRNA cluster sequence, reads mapped to a cluster, proTRAC image file) from selected piRNA clusters for a desired SRA dataset. You can also download clipped and filtered reads from any available SRA sequence set as well as sequence reads from the specified SRA dataset(s) that matched miRNA- or miRNA precursor sequences, respectively

### 2. Getting started

Running `piFETCH` on your local machine requires the installation of a Perl interpreter. Perl is pre-installed on common Linux and Mac systems. For Windows you can download and install either StrawberryPerl (www.strawberryperl.com) or ActivePerl (www.activestate.com/activeperl/downloads). Make sure that you are connceted to the internet and start `piFETCH` from the command line or terminal. Use the following command to download proTRAC results for a specified SRA dataset:

```
perl piFETCH.pl -id [SRA accession, e.g. SRR1755255]
```

You can use multiple SRA accessions at once like this:

```
perl piFETCH.pl -id SRR1755255 SRR1654828 SRR1184429
```

If you do not want to download the complete results but only genomic sequence of a specified piRNA cluster use the option `-s` (or `-cl_sequence`):

```
perl piFETCH.pl -id SRR1755255 -s 123
```

You can use the options `-r` (or `-cl_reads`) for mapped reads and `-i` (or `-cl_image`) for the proTRAC cluster image the same way. You can use all three options together or combine options and use e.g. `-ri`, `-sr` and `-sri`. You can specify an arbitrary number of clusters at once:

```
perl piFETCH.pl -id SRR1755255 -sri 10 20 100
```

Besides information on piRNA clusters you can download sequence datasets derived from SRA datasets. You can download clipped and filtered reads from any SRA sequence set with the option `-clipped_filtered`. You can also use the options `-miR` and `-miRprec` to download reads from the SRA dataset that matched miRNA- or miRNA precursor sequences, respectively:

```
perl piFETCH.pl -id SRR1755255 -clipped_filtered -miR -miRprec
```

You can change the download target directory by using the option `-d` (or `-directory`):

```
perl piFETCH.pl -id SRR1755255 -d /usr/john/downloads
```

### 3. Command line options

| OPTION | VALUE | DESCRIPTION |
|---|---|---|
| -h<br>-? | | Will print this information |

```
-help

-id             SRA accession  Specifies the desired SRA dataset(s). You can
-sra_accession  SRA accession  use multiple SRA accessions at once. When
                               providing more than one SRA accession piFETCH
                               will download the complete proTRAC results.
                               The options -s -r and -i will be ignored.

-s              integer        Specifies the desired cluster id (or number)
-cl_sequence    integer        and will download the genomic cluster sequence
                               Applies only if exactly one SRA accession was
                               provided. You can use multiple cluster id's
                               as desired.

-r              integer        Specifies the desired cluster id (or number)
-cl_reads       integer        and will download a FASTA file comprising
                               the reads that were mapped to the cluster
                               Applies only if exactly one SRA accession was
                               provided. You can use multiple cluster id's
                               as desired.

-i              integer        Specifies the desired cluster id (or number)
-cl_image       integer        and will download the proTRAC image file.
                               Applies only if exactly one SRA accession was
                               provided. You can use multiple cluster id's
                               as desired.

-sr             integer        A combination of the options -s and -r.
-si             integer        A combination of the options -s and -i.
-ri             integer        A combination of the options -r and -i.
-sri            integer        A combination of the options -s, -r and -i.

-clipped_filtered              Will download the clipped and filtered
                               sequence data for the specified SRA datasets.

-miR                           Will download sequence reads that produced
                               perfect full-length matches to mature miRNA
                               sequences deposited at miRBase.

-miRprec                       Will download sequence reads that produced
                               perfect full-length matches to miRNA hairpin
                               sequences deposited at miRBase.

-d                             Specifies the target directory for downloads.
-directory                     By default piFETCH will save the files in the
                               current working directory.
```

## 4. Contact

If you have any questions or comments or found any bugs in the software please do not hesitate to contact:

*David Rosenkranz*
*Institute of Anthropology, small RNA group*
*Johannes Gutenberg University Mainz, Germany*
*Email: rosenkranz@uni-mainz.de*
*Web: http://www.smallRNAgroup-mainz.de*