

# SPORTS1.1

Small non-coding RNA annotation Pipeline Optimized for rRNA- and tRNA-Derived Small RNAs

If you use SPORTS1.1 in your work, please cite this paper:

SPORTS1.0: a tool for annotating and profiling non-coding RNAs optimized for rRNA-and tRNA-derived small RNAs

Junchao Shi, Eun-A Ko, Kenton M. Sanders, Qi Chen, Tong Zhou. SPORTS1.0: a tool for annotating and profiling non-coding RNAs optimized for rRNA-and tRNA-derived small RNAs. Genomics, Proteomics & Bioinformatics (2018) doi.org/10.1016/j.gpb.2018.04.004

Requirements

Installation

Script description

- Example usage

Pre-compiled annotation databases instruction

Instruction for compiling annotation database by user

Copyright and licensing information

Update history

Disclaimer

Contact information

## Requirements

Linux system, enough disk space and Ram depending on the size of RNA deep sequencing data. (Tested system: ubuntu 12.04 LTS, ubuntu 16.04 LTS)

## Installation

1. Download SPORTS1.1 pipeline package.

```
wget https://github.com/junchaoshi/SPORTS1.1/archive/master.zip
```

2. Download necessary software, packages and reference databases as listed below:

1. Perl 5 (<https://www.perl.org>) (Tested version: v5.14.2, v5.22.1); Perl 5 might be already installed in the linux system.

2. Bowtie [1] (<http://bowtie-bio.sourceforge.net/index.shtml>) (Tested version: 1.1.2, 1.2.1.1)
  3. SRA Toolkit (<https://ncbi.github.io/sra-tools/>) (Tested version: 2.8.2)
  4. cutadapt [2] (<http://cutadapt.readthedocs.io/en/stable/index.html>) (Tested version: 1.11)
  5. R (<https://www.r-project.org/>) (Tested version: 3.2.3, 3.2.5)
  6. Reference database (See lists and download link of all pre-compiled species' databases in Pre-compiled Databases Instruction)
3. Installation tutorial for software and packages.

1. Install SPORTS1.1

1. Unpack SPORTS1.1 package.

```
unzip SPORTS1.1-master.zip
```

2. Attach the SPORTS directory to your PATH:

```
echo 'export PATH=$PATH:your_path_to_SPORTS1.1-master/source'
>> ~/.bashrc
```

```
chmod 755 your_path_to_SPORTS1.1-master/source/sports.pl
```

2. Install Bowtie

1. Unpack bowtie-1.x.x-linux-x86\_64.zip.

```
unzip bowtie-1.x.x-linux-x86_64.zip
```

2. Attach the bowtie directory to your PATH:

```
echo 'export PATH=$PATH:your_path_to_bowtie' >>
~/.bashrc
```

If you are administrator user, type the following command and password to easily install

```
sudo apt-get install bowtie
```

3. Install SRA Toolkit

1. Unpack SRA toolkit files.

2. Attach the SRA Toolkit executable path to your PATH:

```
echo 'export PATH=$PATH:your_path_to_sra-toolkit/bin'
>> ~/.bashrc
```

4. Install cutadapt

1. Use pip on the command line to install latest version of cutadapt:

```
pip install --user --upgrade cutadapt
```

2. Attach the cutadapt directory to your PATH:

```
echo 'export PATH=$PATH:$HOME/.local/bin' >> ~/.bashrc
```

5. Install R and R package

1. Unpack R-x.y.z.tar.gz with:

```
tar -xf R-x.y.z.tar.gz
```

2. Enter into the R-x.y.z directory:

```
cd R-x.y.z
```

3. Type following command in terminal:

```
./configure
```

```
make
```

```
make check
```

```
make install
```

4. Install R packages by typing following command in terminal:

```
R
```

```
install.packages('ggplot2', dependencies=TRUE, repos='http://cran.rstudio.com/')
```

```
install.packages('data.table', dependencies=TRUE, repos='http://cran.rstudio.com/')
```

```
install.packages('stringr', dependencies=TRUE, repos='http://cran.rstudio.com/')
```

```
q()
```

```
n
```

4. Start a new shell session to apply changes to environment variables:

```
source ~/.bashrc
```

5. Test if everything is installed properly:

```
perl -v
```

```
sports.pl -h
```

```
bowtie
```

```
fastq-dump
```

```
cutadapt -h
```

R --version

If you get any error messages you should install the software once again.

## Script description

### sports.pl

1. Input query format:

1. .sra files.
2. .fastq/.fq, .fasta/.fa files of deep sequencing reads.

Attention: compressed files need to be unpacked before input!

2. Options:

-Input:

-i <file> Input could be:

a .sra, .fastq/.fq or .fasta/.fa file;

a directory (will run all qualified files in the directory recursively);

a text document with absolute path information for each file/folder (when processing)

-Output:

-o <str> output address of annotation results (default: input address)

-k keep all the intermediate files generated during the running progress

-Alignment:

-l <int> the minimal length of the output sequences (default = 15)

-L <int> the maximal length of the output sequences (default = 45)

-M <int> the total number of mismatches in the entire alignment (default = 0)

-a Remove 5' / 3' adapters

-x <str> (if -a applied) 5' adapter sequence. Default = "GTTCAAGTTCTACAGTCCGACGATC"

-y <str> (if -a applied) 3' adapter sequence. Default = "TGGAATTCTCGGGTGCCAAGG"

-Others:

`-v` print version information

`-h` print this usage message

### 3. Example

- Example use 1:

The user wants to map a single fasta file against rat reference genome to get the mapping genome annotation only. (No output figures)

Type following command in terminal:

```
sports.pl -i reads.fa -g /foo/bar/Rattus_norvegicus/UCSC/rn6/Sequence/BowtieIndex/genome
```

- Example use 2:

The user wants to map several already trimmed human sequencing files to human reference genome, miRNA database, tRNA database, rRNA database and piRNA database by using 4 CPU threads, then to output the result to the address: `‘/foo/bar/output/’`.

Write all the fastq files’ addresses into a text document, e.g.:

```
seq_address.txt
-----
/foo/bar/fold_1/seq_1.fastq
/foo/bar/fold_2/seq_2.fq
/foo/bar/fold_2/seq_3.fq
/foo/bar/fold_3/seq_4.fasta
/foo/bar/fold_4/seq_5.fa
-----
```

Type following command in terminal:

```
sports.pl -i seq_address.txt -p 4 -g /foo/bar/Homo_sapiens/genome/hg38/genome
-m /Homo_sapiens/miRBase/21/miRBase_21-has -r /foo/bar/Homo_sapiens/rRNadb/human_rRNA
-t /foo/bar/Homo_sapiens/GtRNadb/hg19/hg19-tRNAs -w /foo/bar/Homo_sapiens/piRBase/piR_h
-o /foo/bar/output/
```

- Example use 3:

The user wants to map several untrimmed mouse sequencing files downloaded from NCBI or somewhere else to mouse reference genome, miRNA database, tRNA database, rRNA database, piRNA database, ensembl non-coding RNA database and Rfam database by using 4 CPU threads, then to output the result to the address: `‘/foo/bar/output/’` and keep all the intermediate files generated during the running progress.

Put all the sequencing files into a folder, e.g.:

folder structure:

```

-----
download_seq

    fold_1

        seq_1.sra

        seq_2.sra

    fold_2

        fold_3

            seq_3.fastq

            seq_4.fq

            seq_5.fasta

        seq_6.fa
-----

```

Type following command in terminal:

```

sports.pl -i /foo/bar/download_seq/ -p 4 -a -x GTTCAGAGTTCTACAGTCCGACGATC
-y TGGAAATTCTCGGGTGCCAAGG -g /foo/bar/Mus_musculus/genome/mm10/genome
-m /foo/bar/Mus_musculus/miRBase/21/miRbase_21-mmu -r /foo/bar/Mus_musculus/rRNadb/mous
-t /foo/bar/Mus_musculus/GtRNadb/mm10/mm10-tRNAs -w /foo/bar/Mus_musculus/piRBase/piR_m
-e /foo/bar/Mus_musculus/Ensembl/Mus_musculus.GRCm38.ncrna
-f /foo/bar/Mus_musculus/Rfam/12.3/Rfam-12.3-mouse -o /foo/bar/output/
-k

```

#### 4. Example output file structure for 1 query file input (e.g. SeqFile):

Output folder structure

1\_SeqFile

SeqFile\_fa (if -k applied)

SeqFile.fa ---unique seqs with reads number

SeqFile\_disgarded\_reads.fa ---seqs that cannot pass adapter removing

SeqFile\_too\_short\_reads.fa ---seqs that are lower than required mini

SeqFile\_too\_long\_reads.fa ---seqs that are higher than required max

SeqFile_match_genome.fa	---seqs that can match to reference genome
SeqFile_unmatch_genome.fa	---seqs that cannot match to reference genome
SeqFile_match_<X>_match_genome.fa	---seqs that can match to both <X> database and reference genome
SeqFile_match_<X>_unmatch_genome.fa	---seqs that can match to <X> database but not reference genome
SeqFile_unmatch_<X>_match_genome.fa	---seqs that cannot match to <X> database but can match to reference genome
SeqFile_unmatch_<X>_unmatch_genome.fa	---seqs that match to <X> rfam database but not reference genome
SeqFile_processed (if -k applied)	
SeqFile_output_match_genome	---seqs that match to reference genome in output
SeqFile_output_<X>_match_genome	---seqs that match to both miRNA database and reference genome in output
SeqFile_output_<X>_unmatch_genome	---seqs that match to miRNA database but not reference genome in output
SeqFile_result	
SeqFile_output.txt	---6 column table file including annotation
SeqFile_summary.txt	---3 column table file including reads number
SeqFile_length_distribution.txt	---3 column table file including reads number
SeqFile_mismatch_summary.txt	---11 column table file including mismatch
SeqFile_sncRNA_distribution.pdf	---figure of length distribution of sncRNA
SeqFile_rRNA_distribution.pdf	---figure of length distribution of rRNA
SeqFile_rRNA_mapping.pdf	---figure of rRNAs mapping against different databases
SeqFile_tRNA_mapping.pdf	---figure of length distribution of different tRNAs
processing_report (if -k applied)	
1_SeqFile.txt	---processing log file
sh_file (if -k applied)	
1__SeqFile.sh	---processing script file

Some output folders only exist when '-k' parameter is applied in sports.pl;

SeqFile\_mismatch\_summary.txt only exist when '-M' parameter is larger than 0 in sports.pl

Some output files might not exist if the file size is zero.

### **fastq2fasta.pl**

Extracted from miRDeep2 [3] (<https://github.com/rajewsky-lab/mirdeep2>)

1. Description:

Parses fastq format files into fasta format.

2. Input:

A fastq file.

3. Output:

A fasta file, one sequence per line (the sequences are expanded).

4. options:

-

5. Example usage:

```
fastq2fasta.pl reads.fq > reads.fa
```

### **fastaparse.pl**

Extracted from miRDeep2 [3] (<https://github.com/rajewsky-lab/mirdeep2>)

1. Description:

Performs simple filtering of entries in a fasta file.

2. Input:

A fasta file

3. Output:

A filtered fasta file

4. Options:

-a only output entries where the sequence is minimum int nts long

-b remove all entries that have a sequence that contains letters other than  
a,c,g,t,u,n,A,C,G,T,U,N.

-s output progress



5. Example usage:

```
fastaparse.pl reads.fa -a 15 -s > reads_no_short.fa 2>
reads_discarded.fa
```

### **combine\_reads.pl**

1. Description:

Combine reads in the fasta file to get unique sequence and its read number.

2. Input:

A fasta file

3. Output:

A filtered fasta file.

```
-----
>t000000001 1234567
TCCCTGGTGGTCTAGTGGTTAGGATTGGCGC
-----
```

't000000001' is the unique ID of the sequence, representing the abundance ranking among all sequences.  
'1234567' represents the reads number of sequence 'TCCCTGGTGGTCTAGTGGTTAGGATTGGCGC'

4. Options:

-

5. Example usage:

```
combine_reads.pl reads.fa > combined_reads.fa
```

### **tRNA\_db\_processing.pl**

1. Description:

This script adds CCA end to original genomic tRNA 3' end and add G to Histidine tRNA 5' end.

2. Input:

The reference tRNA database file in .fa format

A fasta file

3. Options: -

4. Example usage:

```
tRNA_db_processing.pl input_tRNA_file.fa
```

## annotation.pl

### 1. Description:

Combine the annotation information generated from sports.pl

### 2. Input:

sports.pl output folder address:

### 3. Output:

1. \_\_output.txt: A 6 column table file including annotation information for every unique sequence.

ID	Sequence	Length	Reads	Match_Genome
t00000001	TCCCTGGTGGTCTAGTGGTTAGGATTCGGCGC	32	1234567	Yes

-ID: t00000001 --Represents the unique ID of the sequence.

-Sequence: TCCCTGGTGGTCTAGTGGTTAGGATTCGGCGC --Represents the sequence.

-Length: 32 --Length of the sequence.

-Reads: 1234567 --Reads number of the sequence.

-Match\_Genome: Yes --If the sequence can match the reference genome.

-Annotation: tRNA-Glu-CTC\_5\_end --The annotation of the sequence. The annotation is the same as the one in the sports.pl output.

2. <seq\_fold>\_summary.txt: A 3 column table file including reads number of each major- and sub- classes.

Class	Sub_Class	Reads
tRNAdb-tRNA_5_end_Match_Genome	-	123456
-	tRNA-Glu-CTC_5_end	78910

-Class: tRNAdb-tRNA\_5\_end\_Match\_Genome --The major class name.

-Sub\_Class: tRNA-Glu-CTC\_5\_end --The sub class name.

-Reads: 123456 --The reads number of the class.

3. <seq\_fold>\_length\_distribution.txt: A 3 column table file including reads number of each length distribution of each major class.

Class	Length	Reads
-------	--------	-------

tRNAdb-tRNA_5_end_Match_Genome	30	1234
tRNAdb-tRNA_5_end_Match_Genome	31	23456
tRNAdb-tRNA_5_end_Match_Genome	32	34567
tRNAdb-tRNA_5_end_Match_Genome	33	4567
.....		

```

-----
-Class: tRNAdb-tRNA_5_end_Match_Genome    --The major class name.
-Length: 30                                --Length of the sequence.
-Reads: 1234                               --The reads number of the class.

```

#### 4. Options:

-

#### 5. Example usage:

```
annotation.pl <SPORTS_output_fold_address>
```

### **tRNA\_mapping.pl**

#### 1. Description:

Perl script calculates expression level of genomic loci by bowtie 1 output information.

#### 2. Input:

bowtie 1 output file

#### 3. Output:

expression summary file

#### 4. Options:

-

#### 5. Example usage:

```
tRNA_mapping.pl bowtie_output_file summary_file
```

### **mismatch\_summary.pl**

#### 1. Description:

Perl script summarizes mismatch information generated by Bowtie 1 if parameter '-M' >0 applied in sports.pl

#### 2. Input:

bowtie 1 output file and threshold of reads number.

3. Output:

mismatch summary file

4. Options:

-

5. Example usage:

```
mismatch_summary.pl bowtie_output_file 10 > mismatch_summary.txt
```

### **overall\_RNA\_length\_distribution.R**

1. Description:

Generate figure of length distribution of miRNA, rsRNA, tsRNA, piRNA and other RNAs, if sequence matches exists.

2. Input:

Files generated by annotation.pl

3. Output:

<seq\_fold>\_sncRNA\_distribution.pdf Generate figure of length distribution of different types of rRNAs, if sequence matches exists.

4. Options:

-

5. Example usage:

```
Rscript --vanilla overall_RNA_length_distribution.R <SPORTS_output_fold_address>  
<dataset_name>
```

### **rRNA\_length\_distribution.R**

1. Description:

Generate figure of length distribution of different types of rRNAs, if sequence matches exists. (e.g. 4.5S, 5S, 5.3S, 5.8S, 12S, 16S, 18S, 28S, 45S ...)

2. Input:

Files generated by annotation.pl

3. Output:

<seq\_fold>\_rRNA\_distribution.pdf

4. Options:

-

5. Example usage:

```
Rscript --vanilla r_RNA_length_distribution.R <SPORTS_output_fold_address>  
<dataset_name>
```

### **rRNA\_mapping.R**

1. Description:

Generate figure of rRNAs mapping against different types of rRNAs if sequence matches existed. (e.g. 4.5S, 5S, 5.3S, 5.8S, 12S, 16S, 18S, 28S, 45S ...)

2. Input:

Files generated by annotation.pl

3. Output:

<seq\_fold>\_rRNA\_mapping.pdf

4. Options:

-

5. Example usage:

```
Rscript --vanilla rRNA_mapping.R <SPORTS_output_fold_address>  
<dataset_name>
```

### **tRNA\_mapping.R**

1. Description:

Generate figure of rRNAs mapping against different types of tRNAs if sequence matches existed.

2. Input:

Files generated by tRNA\_mapping.pl

3. Output:

<seq\_fold>\_tRNA\_mapping.pdf

4. Options:

-

5. Example usage:

```
Rscript --vanilla tRNA_mapping.R <SPORTS_output_fold_address>_tRNA_mapping.txt  
<dataset_name>_tRNA_mapping.pdf
```

## **mismatch\_stat.R**

1. Description:

Generate statistics result for mismatch information generated from mismatch\_summary.pl

2. Input:

File generated by mismatch\_summary.pl and base-calling error

3. Output:

<seq\_fold>\_mismatch\_summary.txt

4. Options:

-

5. Example usage:

```
Rscript --vanilla mismatch_stat.R <SPORTS_output_fold_address>_mismatch_summary.txt  
0.01
```

## **Pre-compiled annotation databases instruction**

Available species lists of bowtie-index based reference database

To build bowtie-index of your own:

```
cd /foo/bar/your_reference_database/
```

```
bowtie-build <your_reference_database_name>.fa <your_reference_database_name>
```

The built bowtie index will be: /foo/bar/your\_reference\_database/your\_reference\_database

Unpack reference genome (e.g. human genome):

```
tar zxvf Homo_sapiens_UCSC_hg38.tar.gz
```

Main database source:

-mirbase database [4] (Original source: <http://www.mirbase.org/index.shtml>)

-rRNA database (Original source: <https://www.ncbi.nlm.nih.gov/nuccore/>)

-GtRNadb database [5] (Original source: <http://gtrnadb.ucsc.edu/>)

-mitotRNadb database [6] (Original source: <http://mttrna.bioinf.uni-leipzig.de/mtDataOutput/>)

-piRBase database [7] (Original source: <http://www.regulatoryrna.org/database/piRNA/>)

-piRNABank [8] (Original source: <http://pirnabank.ibab.ac.in/index.shtml>)

-ensembl ncRNA database [9] (Original source: <http://www.ensembl.org/index.html>)

-rfam database [10] (Original source: <http://rfam.xfam.org/>)

To download annotation databases of multiple species please use this link:

[https://ncrnainfo-my.sharepoint.com/:f/g/personal/sports\\_ncrna\\_info/EvZP50IfGFRJlNJwZRujI00BM-CP\\_w7gITsv8iU45m1rxQ?e=Pk89b9](https://ncrnainfo-my.sharepoint.com/:f/g/personal/sports_ncrna_info/EvZP50IfGFRJlNJwZRujI00BM-CP_w7gITsv8iU45m1rxQ?e=Pk89b9)

1. Homo sapiens (Human)

1. annotation databases: (We provide a download link for all databases

listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest8CDc](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest8CDc))

2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Homo_sapiens/genome/hg38/genome
-m /<your_defined_address>/Homo_sapiens/miRBase/21/miRBase_21-hsa
-r /<your_defined_address>/Homo_sapiens/rRNadb/human_rRNA
-t /<your_defined_address>/Homo_sapiens/GtRNadb/hg19/hg19-tRNAs
-w /<your_defined_address>/Homo_sapiens/piRBase/piR_human
-e /<your_defined_address>/Homo_sapiens/Ensembl/release-89/Homo_sapiens.GRCh38.ncrna
-f /<your_defined_address>/Homo_sapiens/Rfam/12.3/Rfam-12.3-human
```

2. Gorilla gorilla gorilla (Gorilla)

1. annotation database: (We provide a download link for all databases

listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guestIjvagrWepHYNF3k](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guestIjvagrWepHYNF3k))

2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Gorilla_gorilla/UCSC/gorGor5/Sequence/BowtieIndex/genome
-m /<your_defined_address>/Gorilla_gorilla/miRBase_21/miRBase_21-ggo
-t /<your_defined_address>/Gorilla_gorilla/GtRNadb/gorGor3-tRNAs
-f /<your_defined_address>/Gorilla_gorilla/Rfam_12.3/Rfam-12.3-gorilla
```

3. Pan paniscus (Bonobo)

1. annotation database: (We provide a download link for all databases

listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest)

2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Pan_paniscus/UCSC/panPan1/Sequence/BowtieIndex/genome
-m /<your_defined_address>/Pan_paniscus/miRBase_21/miRBase_21-ppa
-f /<your_defined_address>/Pan_paniscus/Rfam_12.3/Rfam-12.3-Bonobo
```

#### 4. Pan troglodytes (Chimp)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guestSbiP2C6Qjpc](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guestSbiP2C6Qjpc))

2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Pan_troglodytes/UCSC/panTro4/Sequence/BowtieIndex/genome
-m /<your_defined_address>/Pan_troglodytes/miRBase_21/miRBase_21-ptr
-t /<your_defined_address>/Pan_troglodytes/GtRNadb/panTro4-tRNAs
-e /<your_defined_address>/Pan_troglodytes/Ensembl/Pan_troglodytes.CHIMP2.1.4.ncrna
-f /<your_defined_address>/Pan_troglodytes/Rfam_12.3/Rfam-12.3-chimp
```

#### 5. Pongo abelii (Orangutan)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))

2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Pongo_abelii/UCSC/ponAbe2/Sequence/BowtieIndex/genome
-m /<your_defined_address>/Pongo_abelii/miRBase_21/miRBase_21-ppy
-t /<your_defined_address>/Pongo_abelii/GtRNadb/ponAbe2-tRNAs
-e /<your_defined_address>/Pongo_abelii/Ensembl/Pongo_abelii.PPYG2.ncrna
-f /<your_defined_address>/Pongo_abelii/Rfam_12.3/Rfam-12.3-orangutan
```

#### 6. Nomascus leucogenys (Gibbon)

- 6.1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guestacce](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guestacce))

- 6.2 SPORTS1.1 related parameters if you download recommended reference database:

...

```
-g /<your_defined_address>/Nomascus_leucogenys/UCSC/nomLeu3/Sequence/BowtieIndex/genome
-t /<your_defined_address>/Nomascus_leucogenys/GtRNadb/nomLeu3-tRNAs
-f /<your_defined_address>/Nomascus_leucogenys/Rfam_12.3/Rfam-12.3-gibbon
...
```

#### 7. Macaca mulatta (Rhesus)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guestCESs](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guestCESs))



2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Macaca_mulatta/UCSC/rheMac8/Sequence/BowtieIndex/genome
-m /<your_defined_address>/Macaca_mulatta/miRBase_21/miRBase_21-mm1
-r /<your_defined_address>/Macaca_mulatta/rRNadb/rhesus_rRNA
-t /<your_defined_address>/Macaca_mulatta/GtRNadb/rheMac3-tRNAs
-f /<your_defined_address>/Macaca_mulatta/Rfam_12.3/Rfam-12.3-rhesus
```

#### 8. Papio anubis (Baboon)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest9BLoH\\_Vmw](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest9BLoH_Vmw))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Papio_anubis/UCSC/papAnu2/Sequence/BowtieIndex/genome
-t /<your_defined_address>/Papio_anubis/GtRNadb/papAnu2-tRNAs
-f /<your_defined_address>/Papio_anubis/Rfam_12.3/Rfam-12.3-baboon
```

#### 9. Callithrix jacchus (Marmoset)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guestCWgupu6IA](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guestCWgupu6IA))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Callithrix_jacchus/UCSC/calJac3/Sequence/BowtieIndex/genome
-m /<your_defined_address>/Callithrix_jacchus/rRNadb/marmoset_rRNA
-t /<your_defined_address>/Callithrix_jacchus/GtRNadb/calJac3-tRNAs
-f /<your_defined_address>/Callithrix_jacchus/Rfam_12.3/Rfam-12.3-marmoset
```

#### 10. Carlito syrichta (Tarsier)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guesthXdASOIimxNL4I](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guesthXdASOIimxNL4I))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Carlito_syrichta/UCSC/tarSyr2/Sequence/BowtieIndex/genome
-t /<your_defined_address>/Carlito_syrichta/GtRNadb/tarSyr2-tRNAs
-f /<your_defined_address>/Carlito_syrichta/Rfam_12.3/Rfam-12.3-tarsier
```

#### 11. Rattus norvegicus (rat)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guestQMqBJoQ](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guestQMqBJoQ))

2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Rattus_norvegicus/UCSC/rn6/Sequence/BowtieIndex/genome
-m /<your_defined_address>/Rattus_norvegicus/miRBase_21/miRBase_21-rno
-r /<your_defined_address>/Rattus_norvegicus/rRNadb/mouse_rRNA
-t /<your_defined_address>/Rattus_norvegicus/GtRNadb/rn5-tRNAs
-w /<your_defined_address>/Rattus_norvegicus/piRBase/piR_rat
-e /<your_defined_address>/Rattus_norvegicus/Ensembl/Rattus_norvegicus.Rnor_6.0.ncrna
-f /<your_defined_address>/Rattus_norvegicus/Rfam_12.3/Rfam-12.3-rat
```

#### 12. Mus musculus (mouse)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guestQMqBJoQ](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guestQMqBJoQ))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Mus_musculus/genome/mm10/genome
-m /<your_defined_address>/Mus_musculus/miRBase/21/miRBase_21-mmu
-r /<your_defined_address>/Mus_musculus/rRNadb/mouse_rRNA
-t /<your_defined_address>/Mus_musculus/GtRNadb/mm10/mm10-tRNAs
-w /<your_defined_address>/Mus_musculus/piRBase/piR_mouse
-e /<your_defined_address>/Mus_musculus/Ensembl/release-89/Mus_musculus.GRCm38.ncrna
-f /<your_defined_address>/Mus_musculus/Rfam/12.3/Rfam-12.3-mouse
```

#### 13. Cricetulus griseus (Hamster)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guestQMqBJoQ](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guestQMqBJoQ))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Cricetulus_griseus/UCSC/criGri1/Sequence/BowtieIndex/genome
-m /<your_defined_address>/Cricetulus_griseus/miRBase_21/miRBase_21-cgr
-r /<your_defined_address>/Cricetulus_griseus/rRNA_db/hamster_rRNA
-t /<your_defined_address>/Cricetulus_griseus/GtRNadb/criGri1-tRNAs
-f /<your_defined_address>/Cricetulus_griseus/Rfam_12.3/Rfam-12.3-hamster
```

#### 14. Cavia porcellus (Guinea pig)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guestQMqBJoQ](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guestQMqBJoQ))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Cavia_porcellus/UCSC/cavPor3/Sequence/BowtieIndex/genome
-r /<your_defined_address>/Cavia_porcellus/rRNA_db/guinea_rRNA
-t /<your_defined_address>/Cavia_porcellus/GtRNAdb/cavPor3-tRNAs
-f /<your_defined_address>/Cavia_porcellus/Rfam_12.3/Rfam-12.3-guinea
```

15. *Heterocephalus glaber* (Naked mole-rat)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Heterocephalus_glaber/UCSC/hetGla2/Sequence/BowtieIndex/genome
-r /<your_defined_address>/Heterocephalus_glaber/rRNA_db/mole_rRNA
-t /<your_defined_address>/Heterocephalus_glaber/GtRNAdb/hetGla2-tRNAs
-f /<your_defined_address>/Heterocephalus_glaber/Rfam_12.3/Rfam-12.3-mole
```

16. *Ictidomys tridecemlineatus* (Squirrel)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Ictidomys_tridecemlineatus/UCSC/speTri2/Sequence/BowtieIndex/genome
-t /<your_defined_address>/Ictidomys_tridecemlineatus/GtRNAdb/speTri2-tRNAs
-f /<your_defined_address>/Ictidomys_tridecemlineatus/Rfam_12.3/Rfam-12.3-squirrel
```

17. *Ochotona princeps* (Pika)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest) DcRVQ)
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Ochotona_princeps/UCSC/ochPri3/Sequence/BowtieIndex/genome
-t /<your_defined_address>/Ochotona_princeps/GtRNAdb/ochPri3-tRNAs
-f /<your_defined_address>/Ochotona_princeps/Rfam_12.3/Rfam-12.3-pika
```

18. *Oryctolagus cuniculus* (Rabbit)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest) WJX10nDdXwvNygc)
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Oryctolagus_cuniculus/UCSC/oryCun2/Sequence/BowtieIndex/genome
-m /<your_defined_address>/Oryctolagus_cuniculus/miRBase_21/miRBase_21-ocu
```

```
-r /<your_defined_address>/Oryctolagus_cuniculus/rRNA_db/rabbit_rRNA
-t /<your_defined_address>/Oryctolagus_cuniculus/GtRNAdb/oryCun2-tRNAs
-f /<your_defined_address>/Oryctolagus_cuniculus/Rfam_12.3/Rfam-12.3-rabbit
```

#### 19. *Ovis aries* (Sheep)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Ovis_aries/UCSC/oviAri3/Sequence/BowtieIndex/genome
-m /<your_defined_address>/Ovis_aries/miRBase_21/miRBase_21-oar
-t /<your_defined_address>/Ovis_aries/GtRNAdb/oviAri1-tRNAs
-f /<your_defined_address>/Ovis_aries/Rfam_12.3/Rfam-12.3-sheep
```

#### 20. *Bos taurus* (Cow)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Bos_taurus/UCSC/bosTau8/Sequence/BowtieIndex/genome
-m /<your_defined_address>/Bos_taurus/miRBase_21/miRBase_21-bta
-r /<your_defined_address>/Bos_taurus/rRNA_db/cow_rRNA
-t /<your_defined_address>/Bos_taurus/GtRNAdb/bosTau8-tRNAs
-e /<your_defined_address>/Bos_taurus/Ensembl/Bos_taurus.UMD3.1.ncrna
-f /<your_defined_address>/Bos_taurus/Rfam_12.3/Rfam-12.3-cow
```

#### 21. *Sus scrofa* (Pig)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Sus_scrofa/UCSC/susScr3/Sequence/BowtieIndex/genome
-m /<your_defined_address>/Sus_scrofa/miRBase_21/miRBase_21-ssc
-r /<your_defined_address>/Sus_scrofa/rRNA_db/pig_rRNA
-t /<your_defined_address>/Sus_scrofa/GtRNAdb/susScr3-tRNAs
-f /<your_defined_address>/Sus_scrofa/Rfam_12.3/Rfam-12.3-pig
```

#### 22. *Tursiops truncatus* (Dolphin)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Tursiops_truncatus/UCSC/turTru2/Sequence/BowtieIndex/genome
-t /<your_defined_address>/Tursiops_truncatus/GtRNAdb/turTru2-tRNAs
-f /<your_defined_address>/Tursiops_truncatus/Rfam_12.3/Rfam-12.3-dolphin
```

23. Balaenoptera acutorostrata (Minke whale)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest-q0lzVEEBIth8V8](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest-q0lzVEEBIth8V8))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Balaenoptera_acutorostrata/UCSC/balAcu1/Sequence/BowtieIndex/genome
-t /<your_defined_address>/Balaenoptera_acutorostrata/GtRNAdb/balAcu1-tRNAs
-f /<your_defined_address>/Balaenoptera_acutorostrata/Rfam_12.3/Rfam-12.3-whale
```

24. Erinaceus europaeus (Hedgehog)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest-q0lzVEEBIth8V8](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest-q0lzVEEBIth8V8))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Erinaceus_europaeus/UCSC/eriEur2/Sequence/BowtieIndex/genome
-t /<your_defined_address>/Erinaceus_europaeus/GtRNAdb/eriEur2-tRNAs
-f /<your_defined_address>/Erinaceus_europaeus/Rfam_12.3/Rfam-12.3-hedgehog
```

25. Sorex araneus (Shrew)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest-q0lzVEEBIth8V8](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest-q0lzVEEBIth8V8))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Sorex_araneus/UCSC/sorAra2/Sequence/BowtieIndex/genome
-t /<your_defined_address>/Sorex_araneus/GtRNAdb/sorAra2-tRNAs
-f /<your_defined_address>/Sorex_araneus/Rfam_12.3/Rfam-12.3-shrew
```

26. Canis familiaris (Dog)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest-q0lzVEEBIth8V8](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest-q0lzVEEBIth8V8))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Canis_familiaris/UCSC/canFam3/Sequence/BowtieIndex/genome
-m /<your_defined_address>/Canis_familiaris/miRBase_21/miRBase_21-cfa
-t /<your_defined_address>/Canis_familiaris/GtRNAdb/canFam3-tRNAs
-f /<your_defined_address>/Canis_familiaris/Rfam_12.3/Rfam-12.3-dog
```

27. *Mustela putorius furo* (Ferret)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Mustela_furo/UCSC/musFur1/Sequence/BowtieIndex/genome  
-t /<your_defined_address>/Mustela_furo/GtRNAdb/musFur1-tRNAs  
-f /<your_defined_address>/Mustela_furo/Rfam_12.3-ferret
```

28. *Ailuropoda melanoleuca* (Panda)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Ailuropoda_melanoleuca/UCSC/ailMel1/Sequence/BowtieIndex/genome  
-t /<your_defined_address>/Ailuropoda_melanoleuca/GtRNAdb/ailMel1-tRNAs  
-f /<your_defined_address>/Ailuropoda_melanoleuca/Rfam-12.3-panda
```

29. *Felis catus* (Cat)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Felis_catus/UCSC/felCat8/Sequence/BowtieIndex/genome  
-r /<your_defined_address>/Felis_catus/rRNA_db/cat_rRNA  
-t /<your_defined_address>/Felis_catus/GtRNAdb/felCat5-tRNAs  
-f /<your_defined_address>/Felis_catus/Rfam_12.3/Rfam-12.3-cat
```

30. *Equus caballus* (Horse)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Equus_caballus/UCSC/equCab2/Sequence/BowtieIndex/genome  
-m /<your_defined_address>/Equus_caballus/miRBase_21/miRBase_21-eca  
-r /<your_defined_address>/Equus_caballus/rRNA_db/horse_rRNA  
-t /<your_defined_address>/Equus_caballus/GtRNAdb/equCab2-tRNAs  
-f /<your_defined_address>/Equus_caballus/Rfam_12.3/Rfam-12.3-horse
```

31. *Ceratotherium simum* (White rhinoceros)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
  2. SPORTS1.1 related parameters if you download recommended reference database:
 

```
-g /<your_defined_address>/Ceratotherium_simum/UCSC/cerSim1/Sequence/BowtieIndex/genome
-t /<your_defined_address>/Ceratotherium_simum/GtRNADB/cerSim1-tRNAs
-f /<your_defined_address>/Ceratotherium_simum/Rfam_12.3/Rfam-12.3-rhinoceros
```
32. *Myotis lucifugus* (Microbat)
1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest) 2sRUrTGn7k)
  2. SPORTS1.1 related parameters if you download recommended reference database:
 

```
-g /<your_defined_address>/Myotis_lucifugus/UCSC/myoLuc2/Sequence/BowtieIndex/genome
-t /<your_defined_address>/Myotis_lucifugus/GtRNADB/myoLuc2-tRNAs
-f /<your_defined_address>/Myotis_lucifugus/Rfam_12.3/Rfam-12.3-bat
```
33. *Trichechus manatus* (Manatee)
1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
  2. SPORTS1.1 related parameters if you download recommended reference database:
 

```
-g /<your_defined_address>/Trichechus_manatus/UCSC/triMan1/Sequence/BowtieIndex/genome
-t /<your_defined_address>/Trichechus_manatus/GtRNADB/triMan1-tRNAs
-f /<your_defined_address>/Trichechus_manatus/Rfam_12.3/Rfam-12.3-manatee
```
34. *Loxodonta africana* (Elephant)
1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest) 7s)
  2. SPORTS1.1 related parameters if you download recommended reference database:
 

```
-g /<your_defined_address>/Loxodonta_africana/UCSC/loxAfr3/Sequence/BowtieIndex/genome
-t /<your_defined_address>/Loxodonta_africana/GtRNADB/loxAfr3-tRNAs
-f /<your_defined_address>/Loxodonta_africana/Rfam_12.3/Rfam-12.3-elephant
```
35. *Dasypus novemcinctus* (Armadillo)
1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))

2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Dasypus_novemcinctus/UCSC/dasNov3/Sequence/BowtieIndex/g
-t /<your_defined_address>/Dasypus_novemcinctus/GtRNAdb/dasNov3-tRNAs
-f /<your_defined_address>/Dasypus_novemcinctus/Rfam_12.3/Rfam-12.3-armadillo
```

#### 36. *Notamacropus eugenii* (Wallaby)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guestdF1Mm7C0QH9K00WHH4](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guestdF1Mm7C0QH9K00WHH4))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Notamacropus_eugenii/UCSC/macEug2/Sequence/BowtieIndex/g
-m /<your_defined_address>/Notamacropus_eugenii/miRBase_21/miRBase_21-meu
-t /<your_defined_address>/Notamacropus_eugenii/GtRNAdb/macEug2-tRNAs
-f /<your_defined_address>/Notamacropus_eugenii/Rfam_12.3/Rfam-12.3-wallaby
```

#### 37. *Sarcophilus harrisii* (Tasmanian devil)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guestD6mL0](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guestD6mL0))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Sarcophilus_harrisii/UCSC/sarHar1/Sequence/BowtieIndex/g
-m /<your_defined_address>/Sarcophilus_harrisii/miRBase_21/miRBase_21-sha
-t /<your_defined_address>/Sarcophilus_harrisii/GtRNAdb/sarHar1-tRNAs
-f /<your_defined_address>/Sarcophilus_harrisii/Rfam_12.3/Rfam-12.3-tasmanian
```

#### 38. *Monodelphis domestica* (Opossum)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Monodelphis_domestica/UCSC/monDom5/Sequence/BowtieIndex/
-m /<your_defined_address>/Monodelphis_domestica/miRBase_21/miRBase_21-mdo
-t /<your_defined_address>/Monodelphis_domestica/GtRNAdb/monDom5-tRNAs
-e /<your_defined_address>/Monodelphis_domestica/Ensembl/Monodelphis_domestica.BROA
-f /<your_defined_address>/Monodelphis_domestica/Rfam_12.3/Rfam-12.3-opossum
```

#### 39. *Ornithorhynchus anatinus* (Platypus)



1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guestwfpJ81M](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guestwfpJ81M))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Ornithorhynchus_anatinus/UCSC/ornAna2/Sequence/BowtieIndex/
-m /<your_defined_address>/Ornithorhynchus_anatinus/miRBase_21/miRBase_21-oan
-t /<your_defined_address>/Ornithorhynchus_anatinus/GtRNAdb/ornAna1-tRNAs
-w /<your_defined_address>/Ornithorhynchus_anatinus/piRBase/piR_platypus
-e /<your_defined_address>/Ornithorhynchus_anatinus/Ensembl/Ornithorhynchus_anatinu
-f /<your_defined_address>/Ornithorhynchus_anatinus/Rfam_12.3/Rfam-12.3-platypus
```

#### 40. Taeniopygia guttata (Zebra finch)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guestwfpJ81M](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guestwfpJ81M))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Taeniopygia_guttata/UCSC/taeGut2/Sequence/BowtieIndex/g
-m /<your_defined_address>/Taeniopygia_guttata/miRBase_21/miRBase_21-tgu
-t /<your_defined_address>/Taeniopygia_guttata/GtRNAdb/taeGut2-tRNAs
-f /<your_defined_address>/Taeniopygia_guttata/Rfam_12.3/Rfam-12.3-finch
```

#### 41. Melopsittacus undulatus (Budgerigar)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guestwfpJ81M](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guestwfpJ81M))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Melopsittacus_undulatus/UCSC/melUnd1/Sequence/BowtieIndex/
-t /<your_defined_address>/Melopsittacus_undulatus/GtRNAdb/melUnd1-tRNAs
-f /<your_defined_address>/Melopsittacus_undulatus/Rfam_12.3/Rfam-12.3-budgerigar
```

#### 42. Gallus gallus (Chicken)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guestwfpJ81M](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guestwfpJ81M))
2. PORTS1.0 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Gallus_gallus/UCSC/galGal5/Sequence/BowtieIndex/genome
-m /<your_defined_address>/Gallus_gallus/miRBase_21/miRBase_21-gga
-r /<your_defined_address>/Gallus_gallus/rRNA_db/chicken_rRNA
-t /<your_defined_address>/Gallus_gallus/GtRNAdb/galGal4-tRNAs
-w /<your_defined_address>/Gallus_gallus/piRBase/piR_gga_v1.0
-e /<your_defined_address>/Gallus_gallus/Ensembl/Gallus_gallus.Gallus_gallus-5.0.no
```

- f /<your\_defined\_address>/Gallus\_gallus/Rfam\_12.3/Rfam-12.3-chicken
43. Meleagris gallopavo (Turkey)
1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
  2. SPORTS1.1 related parameters if you download recommended reference database:
 

```
-g /<your_defined_address>/Meleagris_gallopavo/UCSC/melGal1/Sequence/BowtieIndex/genome
-t /<your_defined_address>/Meleagris_gallopavo/GtRNadb/melGal1-tRNAs
-f /<your_defined_address>/Meleagris_gallopavo/Rfam_12.3/Rfam-12.3-turkey
```
44. Chrysemys picta (Painted Turtle)
1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
  2. SPORTS1.1 related parameters if you download recommended reference database:
 

```
-g /<your_defined_address>/Chrysemys_picta/UCSC/chrPic1/Sequence/BowtieIndex/genome
-t /<your_defined_address>/Chrysemys_picta/GtRNadb/chrPic1-tRNAs
-f /<your_defined_address>/Chrysemys_picta/Rfam_12.3/Rfam-12.3-turtle
```
45. Anolis carolinensis (Lizard)
1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest) i58-1fjJeqcDHZUpLLo)
  2. SPORTS1.1 related parameters if you download recommended reference database:
 

```
-g /<your_defined_address>/Anolis_carolinensis/UCSC/anoCar2/Sequence/BowtieIndex/genome
-m /<your_defined_address>/Anolis_carolinensis/miRBase_21/miRBase_21-aca
-t /<your_defined_address>/Anolis_carolinensis/GtRNadb/anoCar2-tRNAs
-e /<your_defined_address>/Anolis_carolinensis/Ensembl/Anolis_carolinensis.AnoCar2
-f /<your_defined_address>/Anolis_carolinensis/Rfam_12.3/Rfam-12.3-lizard
```
46. Xenopus laevis (Frog)
1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
  2. SPORTS1.1 related parameters if you download recommended reference database:
 

```
-g /<your_defined_address>/Xenopus_laevis/UCSC/xenTro7/Sequence/BowtieIndex/genome
-m /<your_defined_address>/Xenopus_laevis/miRBase_21/miRBase_21-xtr
-r /<your_defined_address>/Xenopus_laevis/rRNA_db/frog_rRNA
-t /<your_defined_address>/Xenopus_laevis/GtRNadb/xenTro3-tRNAs
-w /<your_defined_address>/Xenopus_laevis/piRBase/piR_xtr_v1.0
```

```
-e /<your_defined_address>/Xenopus_laevis/Ensembl/Xenopus_tropicalis.JGI_4.2.ncrna
-f /<your_defined_address>/Xenopus_laevis/Rfam_12.3/Rfam-12.3-frog
```

47. *Latimeria chalumnae* (Coelacanth)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guestiphiQFRDzuKoJtA](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guestiphiQFRDzuKoJtA))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Latimeria_chalumnae/UCSC/latCha1/Sequence/BowtieIndex/genome
-r /<your_defined_address>/Latimeria_chalumnae/rRNA_db/coelacanth_rRNA
-t /<your_defined_address>/Latimeria_chalumnae/GtRNadb/latCha1-tRNAs
-f /<your_defined_address>/Latimeria_chalumnae/Rfam_12.3/Rfam-12.3-coelacanth
```

48. *Tetraodon nigroviridis* (Tetraodon)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guestiphiQFRDzuKoJtA](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guestiphiQFRDzuKoJtA))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Tetraodon_nigroviridis/UCSC/tetNig2/Sequence/BowtieIndex/genome
-m /<your_defined_address>/Tetraodon_nigroviridis/miRBase_21/miRBase_21-tni
-t /<your_defined_address>/Tetraodon_nigroviridis/GtRNadb/tetNig2-tRNAs
-f /<your_defined_address>/Tetraodon_nigroviridis/Rfam_12.3/Rfam-12.3-tetraodon
```

49. *Takifugu rubripes* (Fugu)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest1dy5wnCfdQ](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest1dy5wnCfdQ))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Takifugu_rubripes/UCSC/fr3/Sequence/BowtieIndex/genome
-t /<your_defined_address>/Takifugu_rubripes/GtRNadb/fr3-tRNAs
-f /<your_defined_address>/Takifugu_rubripes/Rfam_12.3/Rfam-12.3-fugu
```

50. *Gasterosteus aculeatus* (Stickleback)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest1dy5wnCfdQ](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest1dy5wnCfdQ))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Gasterosteus_aculeatus/UCSC/gasAcu1/Sequence/BowtieIndex/genome
-t /<your_defined_address>/Gasterosteus_aculeatus/GtRNadb/gasAcu1-tRNAs
-e /<your_defined_address>/Gasterosteus_aculeatus/Ensembl/Gasterosteus_aculeatus.BF
```

- f /<your\_defined\_address>/Gasterosteus\_aculeatus/Rfam\_12.3/Rfam-12.3-stickleback
51. *Oryzias latipes* (Medaka)
1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
  2. SPORTS1.1 related parameters if you download recommended reference database:
 

```
-g /<your_defined_address>/Oryzias_latipes/UCSC/oryLat2/Sequence/BowtieIndex/genome
-m /<your_defined_address>/Oryzias_latipes/miRBase_21/miRBase_21-ola
-t /<your_defined_address>/Oryzias_latipes/GtRNAdb/oryLat2-tRNAs
-f /<your_defined_address>/Oryzias_latipes/Rfam_12.3/Rfam-12.3-medaka
```
52. *Oreochromis niloticus* (Nile tilapia)
1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
  2. SPORTS1.1 related parameters if you download recommended reference database:
 

```
-g /<your_defined_address>/Oreochromis_niloticus/UCSC/oreNil2/Sequence/BowtieIndex/
-t /<your_defined_address>/Oreochromis_niloticus/GtRNAdb/oreNil2-tRNAs
-f /<your_defined_address>/Oreochromis_niloticus/Rfam_12.3/Rfam-12.3-tilapia
```
53. *Gadus morhua* (Atlantic cod)
1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
  2. SPORTS1.1 related parameters if you download recommended reference database:
 

```
-g /<your_defined_address>/Gadus_morhua/UCSC/gadMor1/Sequence/BowtieIndex/genome
-r /<your_defined_address>/Gadus_morhua/rRNA_db/cod_rRNA
-t /<your_defined_address>/Gadus_morhua/GtRNAdb/gadMor1-tRNAs
-f /<your_defined_address>/Gadus_morhua/Rfam_12.3/Rfam-12.3-cod
```
54. *Danio rerio* (Zebrafish)
1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest) VjshX-GKLZI7limHFlo)
  2. SPORTS1.1 related parameters if you download recommended reference database:
 

```
-g /<your_defined_address>/Danio_rerio/UCSC/danRer10/Sequence/BowtieIndex/genome
-m /<your_defined_address>/Danio_rerio/miRBase_21/miRBase_21-dre
-r /<your_defined_address>/Danio_rerio/rRNA_db/zebrafish_rRNA
-t /<your_defined_address>/Danio_rerio/GtRNAdb/danRer6-tRNAs
-w /<your_defined_address>/Danio_rerio/piRBase/piR_dre_v1.0
```

```
-e /<your_defined_address>/Danio_rerio/Ensembl/Danio_rerio.GRCz10.ncrna
-f /<your_defined_address>/Danio_rerio/Rfam_12.3/Rfam-12.3-zebrafish
```

55. *Callorhinchus milii* (Elephant shark)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Callorhinchus_milii/UCSC/calMil1/Sequence/BowtieIndex/g
-t /<your_defined_address>/Callorhinchus_milii/GtRNAdb/calMil1-tRNAs
-f /<your_defined_address>/Callorhinchus_milii/Rfam_12.3/Rfam-12.3-shark
```

56. *Petromyzon marinus* (Lamprey)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Petromyzon_marinus/UCSC/petMar2/Sequence/BowtieIndex/gen
-r /<your_defined_address>/Petromyzon_marinus/rRNA_db/lamprey_rRNA
-t /<your_defined_address>/Petromyzon_marinus/GtRNAdb/petMar2-tRNAs
-f /<your_defined_address>/Petromyzon_marinus/Rfam_12.3/Rfam-12.3-lamprey
```

57. *Strongylocentrotus purpuratus* (Sea urchin)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Strongylocentrotus_purpuratus/UCSC/strPur2/Sequence/Bowt
-m /<your_defined_address>/Strongylocentrotus_purpuratus/miRBase_21/miRBase_21-spu
-r /<your_defined_address>/Strongylocentrotus_purpuratus/rRNA_db/urchin_rRNA
-t /<your_defined_address>/Strongylocentrotus_purpuratus/GtRNAdb/Spurp-tRNAs
-e /<your_defined_address>/Strongylocentrotus_purpuratus/Ensembl/Strongylocentrotus
-f /<your_defined_address>/Strongylocentrotus_purpuratus/Rfam_12.3/Rfam-12.3-urchin
```

58. *Drosophila melanogaster* (Drosophila)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Drosophila_melanogaster/UCSC/dm6/Sequence/BowtieIndex/g
-m /<your_defined_address>/Drosophila_melanogaster/miRBase_21/miRBase_21-dme
-r /<your_defined_address>/Drosophila_melanogaster/rRNA_db/drosophila_rRNA
```

```
-t /<your_defined_address>/Drosophila_melanogaster/GtRNAdb/dm6-tRNAs
-w /<your_defined_address>/Drosophila_melanogaster/piRBase/piR_dme
-e /<your_defined_address>/Drosophila_melanogaster/Ensembl/Drosophila_melanogaster
-f /<your_defined_address>/Drosophila_melanogaster/Rfam_12.3/Rfam-12.3-drosophila
```

59. *Anopheles gambiae* (Mosquito)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Anopheles_gambiae/UCSC/anoGam1/Sequence/BowtieIndex/gen
-m /<your_defined_address>/Anopheles_gambiae/miRBase_21/miRBase_21-aga
-t /<your_defined_address>/Anopheles_gambiae/GtRNAdb/Agamb-tRNAs
-e /<your_defined_address>/Anopheles_gambiae/Ensembl/Anopheles_gambiae.AgamP4.ncrna
-f /<your_defined_address>/Anopheles_gambiae/Rfam_12.3/Rfam-12.3-mosquito
```

60. *Pristionchus pacificus* (Roundworm)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Pristionchus_pacificus/UCSC/priPac1/Sequence/BowtieIndex
-m /<your_defined_address>/Pristionchus_pacificus/miRBase_21/miRBase_21-ppc
-r /<your_defined_address>/Pristionchus_pacificus/rRNA_db/roundworm_rRNA
-t /<your_defined_address>/Pristionchus_pacificus/GtRNAdb/priPac1-tRNAs
-f /<your_defined_address>/Pristionchus_pacificus/Rfam_12.3/Rfam-12.3-roundworm
```

61. *Caenorhabditis elegans* (Nematode):

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest) pCI)
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Caenorhabditis_elegans/UCSC/ce10/Sequence/BowtieIndex/g
-m /<your_defined_address>/Caenorhabditis_elegans/miRBase_21/miRBase_21-cel
-r /<your_defined_address>/Caenorhabditis_elegans/rRNA_db/cel_rRNA
-t /<your_defined_address>/Caenorhabditis_elegans/GtRNAdb/ce10-tRNAs
-w /<your_defined_address>/Caenorhabditis_elegans/piRBase/piR_cel_v1.0
-e /<your_defined_address>/Caenorhabditis_elegans/Ensembl/Caenorhabditis_elegans.W
-f /<your_defined_address>/Caenorhabditis_elegans/Rfam_12.3/Rfam-12.3-nematode
```

62. *Saccharomyces cerevisiae* (Yeast)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
2. SPORTS1.1 related parameters if you download recommended reference database:
 

```
-g /<your_defined_address>/Saccharomyces_cerevisiae/UCSC/sacCer3/Sequence/BowtieIndex
-r /<your_defined_address>/Saccharomyces_cerevisiae/rRNA_db/yeast_rRNA
-t /<your_defined_address>/Saccharomyces_cerevisiae/GtRNADB/sacCer3-tRNAs
-e /<your_defined_address>/Saccharomyces_cerevisiae/Ensembl/Saccharomyces_cerevisiae
-f /<your_defined_address>/Saccharomyces_cerevisiae/Rfam_12.3/Rfam-12.3-yeast
```

#### 63. Zea mays (Corn)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
2. SPORTS1.1 related parameters if you download recommended reference database:
 

```
-g /<your_defined_address>/Zea_mays/Ensembl/AGPv4/Sequence/BowtieIndex/genome
-m /<your_defined_address>/Zea_mays/miRBase_21/miRBase_21-zma
-r /<your_defined_address>/Zea_mays/rRNA_db/corn_rRNA
-t /<your_defined_address>/Zea_mays/GtRNADB/zeaMay5-tRNAs
-e /<your_defined_address>/Zea_mays/Ensembl/Zea_mays.AGPv4.ncrna
-f /<your_defined_address>/Zea_mays/Rfam_12.3/Rfam-12.3-corn
```

#### 64. Sorghum bicolor (Sorghum)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
2. SPORTS1.1 related parameters if you download recommended reference database:
 

```
-g /<your_defined_address>/Sorghum_bicolor/Ensembl/Sbi1/Sequence/BowtieIndex/genome
-m /<your_defined_address>/Sorghum_bicolor/miRBase_21/miRBase_21-sbi
-r /<your_defined_address>/Sorghum_bicolor/rRNA_db/sorghum_rRNA
-t /<your_defined_address>/Sorghum_bicolor/GtRNADB/Sbico-tRNAs
-e /<your_defined_address>/Sorghum_bicolor/Ensembl/Sorghum_bicolor.Sorghum_bicolor
-f /<your_defined_address>/Sorghum_bicolor/Rfam_12.3/Rfam-12.3-sorghum
```

#### 65. Oryza sativa (Rice)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
2. SPORTS1.1 related parameters if you download recommended reference database:
 

```
-g /<your_defined_address>/Oryza_sativa/Ensembl/IRGSP-1.0/Sequence/BowtieIndex/genome
-m /<your_defined_address>/Oryza_sativa/miRBase_21/miRBase_21-osa
-r /<your_defined_address>/Oryza_sativa/rRNA_db/rice_rRNA
```

```
-t /<your_defined_address>/Oryza_sativa/GtRNAb/Osati-tRNAs
-e /<your_defined_address>/Oryza_sativa/Ensembl/Oryza_sativa.IRGSP-1.0.ncrna
-f /<your_defined_address>/Oryza_sativa/Rfam_12.3/Rfam-12.3-rice
```

#### 66. Arabidopsis thaliana (Arabidopsis)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Arabidopsis_thaliana/Ensembl/TAIR10/Sequence/BowtieIndex
-m /<your_defined_address>/Arabidopsis_thaliana/miRBase_21/miRBase_21-ath
-r /<your_defined_address>/Arabidopsis_thaliana/rRNA_db/Arabidopsis_rRNA
-t /<your_defined_address>/Arabidopsis_thaliana/GtRNAb/araTha1-tRNAs
-e /<your_defined_address>/Arabidopsis_thaliana/Ensembl/Arabidopsis_thaliana.TAIR10
-f /<your_defined_address>/Arabidopsis_thaliana/Rfam_12.3/Rfam-12.3-arabidopsis
```

#### 67. Glycine max (Soybean)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest))
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Glycine_max/Ensembl/Gm01/Sequence/BowtieIndex/genome
-m /<your_defined_address>/Glycine_max/miRBase_21/miRBase_21-gma
-r /<your_defined_address>/Glycine_max/rRNA_db/soybean_rRNA
-t /<your_defined_address>/Glycine_max/GtRNAb/glyMax2-tRNAs
-f /<your_defined_address>/Glycine_max/Rfam_12.3/Rfam-12.3-soybean
```

#### 68. Escherichia coli (E.coli)

1. annotation database: (We provide a download link for all databases listed below: [https://ncrnainfo-my.sharepoint.com/personal/sports\\_ncrna\\_info/\\_layouts/15/guest](https://ncrnainfo-my.sharepoint.com/personal/sports_ncrna_info/_layouts/15/guest) SpOIXaUik)
2. SPORTS1.1 related parameters if you download recommended reference database:

```
-g /<your_defined_address>/Escherichia_coli/Ensembl/EB1/Sequence/BowtieIndex/genome
-r /<your_defined_address>/Escherichia_coli/rRNA_db/e_coli_rRNA
-t /<your_defined_address>/Escherichia_coli/GtRNAb/eschColi-tRNAs
-e /<your_defined_address>/Escherichia_coli/Ensembl/Escherichia_coli.HUSEC2011CHR1
-f /<your_defined_address>/Escherichia_coli/Rfam_12.3/Rfam-12.3-e_coli
```

### Instruction for compiling annotation database by user

1. Prepare the reference databases for your species X



1. Download the reference genome in .fa format into the defined folder address: <your\_defined\_address>; (necessary)
2. Download and extract the miRNA sequences belong to the species in .fa format from miRBase (<http://www.mirbase.org/>), convert RNA sequences to cDNA sequences (U -> T), then put the file X\_mirbase.fa into the defined folder address: <your\_defined\_address>; (optional)
3. Download the rRNA sequences in .fa format from NCBI or anywhere else, convert RNA sequences to cDNA sequences (U -> T), name and list each rRNA sequences separately, as exemplified below, and put the files such X\_rRNA\_5S.fa, X\_rRNA\_5.8S.fa into the defined folder address: <your\_defined\_address>; (optional)

human\_rRNA\_5S.fa

>Human 5S ribosomal RNA

GTCTACGGCCATACCACCCTGAACGCGCCCGATCTCGTCTGATCTCGGAAGCTAAGCAGGGTCGGGCCTG  
GTTAGTACTTGGATGGGAGACCGCCTGGGAATACCGGGTGCTGTAGGCTTT

human\_rRNA\_5.8S.fa

>Human 5.8S ribosomal RNA

CGACTCTTAGCGGTGGATCACTCGGCTCGTGCCTCGATGAAGAACGCGCTAGCTGCGAGAATTAATG  
TGAATTGCAGGACACATTGATCATCGACACTTCGAACGCACTTGCGGCCCCGGGTTCTCCCGGGGCTAC  
GCCTGTCTGAGCGTCGCTT

4. Download the high confidence tRNA sequences (X-tRNAs.fa) and high confidence mature tRNA sequences (X-mature-tRNAs.fa) in .fa format from GtRNAdb (<http://gtrnadb.ucsc.edu/>) and put the files X-mature-tRNAs.fa and X-tRNAs.fa into the defined folder address: <your\_defined\_address>; (optional)
5. Download mitochondrial tRNA sequences from mitotRNAdb (<http://mttrna.bioinf.uni-leipzig.de/mtDataOutput/>) in .fa format and reformat the file X\_mito-tRNA.fa by using the scripts provided below and put the generated files X-mt\_tRNA.fa and X-mt\_tRNA\_CCA.fa into the defined folder address: <your\_defined\_address>; (optional)

mt\_tRNA\_transfer.pl X\_mito-tRNA.fa X-mt\_tRNA

6. Download the noncoding RNA sequences in .fa format from Ensembl database (<http://ensemblgenomes.org/>) and put the file X\_ensembl\_ncrna.fa into the defined folder address: <your\_defined\_address>; (optional)
7. Download and extract the noncoding RNA sequences in .fa format from Rfam database (<http://ensemblgenomes.org/>) and

put the file X\_rfam\_ncrna.fa into the defined folder address:  
<your\_defined\_address>; (optional)

8. Download and extract the noncoding RNA sequences belong to the species in .fa format from Rfam database (<https://rfam.xfam.org/>) and put the file X\_rfam.fa into the defined folder address:  
<your\_defined\_address>; (optional)

9. Download piRNA sequences in .fa format from piRNA databases and put the file X\_piRNA.fa into the defined folder address:  
<your\_defined\_address>. (optional)

2. Type following command in terminal:

```
sports.pl -i /foo/bar/download_seq/ -p 4 -g <your_defined_address>/X  
-m <your_defined_address>/X_mirbase -r <your_defined_address>/X_rRNA  
-t <your_defined_address>/X-tRNAs -e <your_defined_address>/X_ensembl_ncrna  
-f <your_defined_address>/X_rfam_ncrna -w <your_defined_address>/X_piRNA  
-o /foo/bar/output/
```

## References

1. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 2009;10(3):R25. doi: 10.1186/gb-2009-10-3-r25.
2. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal.* 2011;17(1):3. doi: <http://dx.doi.org/10.14806/ej.17.1.200>.
3. Friedlander MR, Chen W, Adamidi C, Maaskola J, Einspanier R, Kneipel S, et al. Discovering microRNAs from deep sequencing data using miRDeep. *Nat Biotechnol.* 2008;26(4):407-15. doi: 10.1038/nbt1394.
4. Kozomara A, Griffiths-Jones S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic acids research.* 2014;42(Database issue):D68-73. doi: 10.1093/nar/gkt1181.
5. Chan PP, Lowe TM. GtRNAdb 2.0: an expanded database of transfer RNA genes identified in complete and draft genomes. *Nucleic acids research.* 2016;44(D1):D184-9. doi: 10.1093/nar/gkv1309.
6. Jühling F, Mörl M, Hartmann RK, Sprinzl M, Stadler PF, Pütz J. tRNAdb 2009: compilation of tRNA sequences and tRNA genes. *Nucleic acids research.* 2009 Jan;37(Database issue):D159-62. doi: 10.1093/nar/gkn772.
7. Zhang P, Si X, Skogerbo G, Wang J, Cui D, Li Y, et al. piRBase: a web resource assisting piRNA functional study. *Database (Oxford).* 2014;2014:bau110. doi: 10.1093/database/bau110.
8. Sai Lakshmi S, Agrawal S. piRNABank: a web resource on classified and clustered Piwi-interacting RNAs. *Nucleic acids research.* 2008;36(Database issue):D173-7. doi: 10.1093/nar/gkm696.
9. Yates A, Akanni W, Amode MR, Barrell D, Billis K, Carvalho-Silva D, et al. Ensembl 2016. *Nucleic acids research.* 2016;44(D1):D710-6. doi:

10.1093/nar/gkv1157.

10. Nawrocki EP, Burge SW, Bateman A, Daub J, Eberhardt RY, Eddy SR, et al. Rfam 12.0: updates to the RNA families database. *Nucleic acids research*. 2015;43(Database issue):D130-7. doi: 10.1093/nar/gku1063.

## Update history

1.1.0 - 02/20/2020 1. The tsRNA annotation information has been refined to pre-tsRNA and mature-tsRNA based on updated tRNA database; 2. Annotation for Mitochondrial tRNAs (obtained from mitotRNAdb <http://mttrna.bioinf.uni-leipzig.de/mtDataOutput/>) and yRNAs (obtained from <https://www.ncbi.nlm.nih.gov/>) are now available. Pre-compiled databases for mouse and human have been updated (databases for other species will updated soon if applicable), please re-download the databases for better results; 3. The miRNA annotation process has been optimized to get more accurate results. 4. Bowtie index building process has been imbedded in sports.pl to reduce the size of pre-compiled database (the bowtie index will build when the user using SPORTS for the first time); 5. New parameter '-z' is added to make mismatch information statistics optional ('-M' > 0 is needed for the statistics). 6. Instruction for compiling annotation database by user is added.

1.0.5 - 09/25/2018 1. Changed annotation strategy to fit more species tRNA annotation in GtRNAdb. 2. Fixed some bugs in the previous version. Thanks Catrin Lloyd to figure one out!

1.0.4 - 03/25/2018 1. New function added: if '-M' > 0 applied in sports.pl, statistics of mismatch information will generated to predict potential nucleic acid modification loci; 2. The figure of tRNA mapping information will generated if tRNA database is defined; 3. A sample output with the parameter '-M 1' of dataset GSM2304822 (mouse sperm) is uploaded here.

1.0.3 - 01/27/2018 1. Changed the mapping order to make the mapping results more accurate according to the paper " Juan Pablo Tosar, Carlos Rovira, Alfonso Cayota. Non-coding RNA fragments account for the majority of annotated piRNAs expressed in somatic non-gonadal tissues. *Communications Biology*. 2018;1, Article number: 2. doi: 10.1038/s42003-017-0001-7 ".

1.0.2 - 01/22/2018 1. Fixed bugs to generate more accurate mature tRNA database.

1.0.1 - 01/13/2018 1. Fixed bugs that caused unexpected exit. 2. Optimized script to speed up tRNA annotation process.

## **Copyright and licensing information**

SPORTS1.1 is available under the GNU General Public License version 3 (GPLv3).

## **Disclaimer**

The SPORTS1.1 package is provided as is without any guarantees or warranty for correctness. The authors are not responsible for any damage or loss of any kind caused by the use or misuse of the scripts included in the software package. The authors are not under obligation to provide support, service, corrections, or upgrades to the package.

## **Contact information**

Contact author: Junchao Shi

E-mail: [sports.rna@gmail.com](mailto:sports.rna@gmail.com); [junchao.shi@yahoo.com](mailto:junchao.shi@yahoo.com)