README.md

RADish

Scripts aiding file processing and analysis of RADseq and GBS data

LICENSE

All code within the `RADish` v0.1.0 repository is available "AS IS" under a generous GNU license. See the LICENSE file for more information.

CITATION

If you use scripts from this repository as part of your published research, I require that you cite the repository as follows (also see DOI information below):

- Bagley, J.C. 2018. RADish v0.1.0. GitHub repository, Available at: http://github.com/justincbagley/RADish.

Alternatively, please provide the following link to this software repository in your manuscript:

- https://github.com/justincbagley/RADish

USAGE

fastq_ReadLengthChecker

Example usage code and output to screen during a `fastq_ReadLengthChecker.sh` run, in which reads were discovered to range from 78 bp to 140 bp in length:

```
$ ./fastq_readLengthChecker.sh .

#############################################################################
#                     fastq_ReadLengthChecker v0.1.0, February 2018         #
#############################################################################

INFO      | Wed Feb 21 17:58:12 EST 2018 | Found 1123 .fastq files in current working directory.
INFO      | Wed Feb 21 17:58:12 EST 2018 | Shortest read length: 78 bp
INFO      | Wed Feb 21 17:58:12 EST 2018 | Longest read length: 140 bp
INFO      | Wed Feb 21 17:58:12 EST 2018 | Results output to 'fastq_lengths_summary.txt' in current working direc
INFO      | Wed Feb 21 17:58:12 EST 2018 |
INFO      | Wed Feb 21 17:58:12 EST 2018 | ...Cleaning up workspace...
INFO      | Wed Feb 21 17:58:12 EST 2018 |
INFO      | Wed Feb 21 17:58:12 EST 2018 | Done checking fastq read lengths.
INFO      | Wed Feb 21 17:58:12 EST 2018 | Bye.
```

fastqTrimmer

Usage

```
$ fastqTrimmer -h
```

```
Usage: $(basename "$0") [Help: -h help] [Options: -i l] workingDir
 ## Help:
  -h   help text (also: -help)

 ## Options:
  -i   inputFile (def: NULL) file name, in case of single input file
  -b   startingBase (def: 1) nucleotide position to start keeping bases from
       (=also starting position for counting up to trim point)
  -l   trimLength (def: $MY_TRIM_LENGTH) desired final length of reads
  -o   output (def: NULL; e.g. 'output') basename for output .fastq file name

 OVERVIEW
 Trims one or more fastq files from their original length to the length (trimLength)
 desired by the user, by trimming off bases from the right (3') end of sequence reads
 in the fastq file(s). Useful for trimming reads from multiple lanes or assemblies so
 that they all have the same length prior to calling SNPs or merging assemblies. For
 example, the author has used this script to prepare fastqs from different lanes of
 Illumina sequencing on different or mixed sets of samples prior to de novo or
 reference-based assembly and SNP calling in pyRAD (Eaton 2014) or ipyrad (Eaton and
 Overcast 2017).

 Several options are available. If an input file is specified using the -i flag, then only
 that file will be trimmed, and if no output basename (-o) is given then the trimmed file
 will replace the original file. If no input file or output names are given, then the
 script will trim all .fastq files in the current working directory (final workingDir
 argument).

 CITATION
 Bagley, J.C. 2018. RADish v0.1.0. GitHub repository, Available at:
        <http://github.com/justincbagley/RADish>.

 REFERENCES
 Eaton DA (2014) PyRAD: assembly of de novo RADseq loci for phylogenetic analyses.
        Bioinformatics, 30, 1844-1849.
 Eaton DAR, Overcast I (2017) ipyrad: interactive assembly and analysis of RADseq data sets.
        Available at: <http://ipyrad.readthedocs.io/>.
```

Example usage code and output to screen during a `fastqTrimmer.sh` run, in which all fastq files within the working directory were trimmed down to the first 1-91 bases:

```
$ ./fastqTrimmer.sh -l 91 .

################################################################################
#                            fastqTrimmer v0.1.0, February 2018                #
################################################################################

INFO      | Thu Feb 22 09:32:59 MST 2018 | STEP #1: SETUP.
INFO      | Thu Feb 22 09:32:59 MST 2018 |          Setting working directory to: .
INFO      | Thu Feb 22 09:32:59 MST 2018 | STEP #2: TRIMMING READS.
INFO      | Thu Feb 22 09:32:59 MST 2018 |          Trimming reads in all .fastq files in current directory. Trim
INFO      | Thu Feb 22 09:32:59 MST 2018 |          ./trimmed_fastq/, and original fastq files will be moved to ./
./APA1_1.R1.fastq
./APA1_2.R1.fastq
./APA1_3.R1.fastq
./APA1_4.R1.fastq
./APA1_5.R1.fastq
./APA1_6.R1.fastq
./APA1_7.R1.fastq
./APA1_8.R1.fastq
./CAP1H_10.R1.fastq
./CAP1H_1.R1.fastq
./CAP1H_2.R1.fastq
./CAP1H_3.R1.fastq
```

```
./CAP1H_4.R1.fastq
./CAP1H_5.R1.fastq
./CAP1H_6.R1.fastq
./CAP1H_7.R1.fastq
./CAP1H_8.R1.fastq
./CAP1H_9.R1.fastq
./CAP1L_10.R1.fastq
./CAP1L_1.R1.fastq
.
.
.
```

February 22, 2018 Justin C. Bagley, Richmond, VA, USA