Bash scripts: Test data

RGI

Bash file to run RGI (only forward reads): `bashScripts/RGI_test.sh`

```
#!/usr/bin/env bash
if [ $# -ne 1 ]
    then
    echo "usage: card_RGI.sh forward"
    echo "forward format: fastq or fastq.gz"
    echo "Downloads CARD database and aligns forward and reverse FASTQ
reads using Bowtie2"
    exit
fi

forward=$1
filenamef=${forward##*/}
sample=${filenamef%_*}

if [ ! -f $forward ]
then
    echo "File '$forward' not found! Cannot continue"
    exit 1
fi

mkdir RGI_analysis
cd RGI_analysis

set -e

wget https://card.mcmaster.ca/latest/data ##generates data document
tar -xvf data ./card.json ##generates card.json in the current folder
rgi load --card_json card.json --local ##creates a folder called localdb

version=$(rgi database --version --local) ##obtain the card version we
just downloaded

##This commands will generate: card_annotation.log and
card_database_xx.fasta objects
rgi card_annotation -i card.json > card_annotation.log 2>&1
rgi load -i card.json --card_annotation card_database_v$version.fasta --
local
```

```
echo "Downloaded card database version v$version"
echo "Aligning forward and reverse FASTQ reads using Bowtie2 against
v$version CARD database"

rgi bwt --read_one  $forward\
 --aligner bowtie2\
 --output_file ${sample}_RGI_output\
 --threads 8 --local
```

```
conda activate rgi
script=/home/erubio/Documentos/UOCMaster/bashScripts/RGI_test.sh
test=/home/erubio/Documentos/UOCMaster/data_testing/argannot-150bp-
10000-reads.fastq
bash $script $test
```

Groot

Bash file to run Groot (only forward reads): `bashScripts/RGI_test.sh`

```
#!/usr/bin/env bash
MYDIR=/home/erubio/Documentos/UOCMaster/bashScripts

if [ $# -ne 2 ]
    then
    echo "usage: card_groot.sh forward reverse seqlen"
    echo "forward format: fastq"
    echo "short reads length from input sequences (numeric)"
    echo "Downloads CARD database and aligns forward and reverse FASTQ
using GROOT pipeline"
    exit
fi
set -e

forward=$1
filenamef=${forward##*/}
sample=${filenamef%.*}
seqlen=$2 ##short reads length


if [ ! -f $forward ]
then
    echo "File '$forward' not found! Cannot continue"
    exit 1
fi

mkdir groot_analysis
cd groot_analysis
```

```
groot get -d card
##generates a folder called card.90 in working directory with clustered
card database

groot index -m card.90 -i grootIndex$seqlen -w $seqlen -p 8
##Convert a set of clustered reference sequences to variation graphs and
then index them

groot align -i grootIndex$seqlen -f $forward -p 8 -g $sample-groot-
graphs >$sample.bam
##generates bam file from both sequences
##Multiple FASTQ files can be specified as input
##however all are treated as the same sample and paired-end info isn't
used

samtools view -F 256 -h $sample.bam > $sample.sam
##Transform BAM file to SAM file (to execute python function).
##We have removed sequences with flag "not primary alignment"

#########################################################################
##Execute python function
python $MYDIR/../pythonScripts/groot_uniqseq.py $sample.sam
#########################################################################

##Continue with bash:

grep "\S" $sample-uniqseq.txt > $sample-uniqseq2.txt

mv -f $sample-uniqseq2.txt $sample-uniqseq.txt

cp $sample-uniqseq.txt $sample-uniqseq.sam
samtools view -S -b $sample-uniqseq.sam > $sample-uniqseq.bam

groot report -c 0 --bamFile $sample-uniqseq.bam >$sample-uniqseq-0report
groot report --bamFile $sample-uniqseq.bam >$sample-uniqseq-0.97report
groot report --bamFile $sample-uniqseq.bam --lowCov>$sample-uniqseq-
lowCov-report
echo "Report: This will report gene, read count, gene length, coverage
cigar"
```

```
conda activate Groot
script=/home/erubio/Documentos/UOCMaster/bashScripts/groot_test.sh
bash $script $test 150
```

ARIBA

ARIBA requires forward and reverse sequences: we split the fastq file in two using fastqsplitter:

```
fastqsplitter -i argannot-150bp-10000-reads.fastq -o argannot-150bp-
10000-reads_split.1.fastq -o argannot-150bp-10000-reads_split2.fastq
```

```
conda activate ariba2
script=/home/erubio/Documentos/UOCMaster/bashScripts/card_ariba.sh
test1=/home/erubio/Documentos/UOCMaster/data_testing/argannot-150bp-
10000-reads_split.1.fastq
test2=/home/erubio/Documentos/UOCMaster/data_testing/argannot-150bp-
10000-reads_split2.fastq
```

```
bash $script $test1 $test2
```