##
####

(c) 2015-2017 Merly Escalona
merlyescalona@uvigo.es

Phylogenomics Lab. University of Vigo.

Description:

==========

Pipelines for data analysis

Running @triploid.uvigo.es

##
####
#!/bin/bash -l
##
####

Folder paths

##
####
source $HOME/src/vc-benchmark-cesga/src/vcs.variables.sh

```
################################################################################
####
simphyReplicateID=1
pipelinesName="ssp"
replicatesNumDigits=5
replicateID="$(printf "%0${replicatesNumDigits}g" $simphyReplicateID)"
ngsphyReplicatePath="$HOME/data/NGSphy_${pipelinesName}.${replicateID}"
simphyReplicatePath="$HOME/data/${pipelinesName}.${replicateID}"
referencesReplicatePath="$HOME/data/references/references.${pipelinesName}.${replicateID}*"
################################################################################
####
```

1. REFERENCE INDEXING WITH BWA

```
################################################################################
####
for fastaFile in $(find $referencesReplicatePath -name *.fasta); do
qsub $HOME/src/vc-benchmark-cesga/jobs/analysis/ssp.analysis.1.sh $fastaFile
done
################################################################################
####
```

2. GENERATION OF BWA COMMAND LINESs

```
################################################################################
####
```

qsub -t $simphyReplicateID $HOME/src/vc-benchmark-cesga/jobs/analysis/ssp.analysis.2.sh PE150OWN HiSeq2500

```
qsub -t $simphyReplicateID $HOME/src/vc-benchmark-cesga/jobs/analysis/ssp.analysis.2.sh
```

PE150DFLT HiSeq2500
qsub -t $simphyReplicateID $HOME/src/vc-benchmark-cesga/jobs/analysis/ssp.analysis.2.sh SE150DFLT HiSeq2500
qsub -t $simphyReplicateID $HOME/src/vc-benchmark-cesga/jobs/analysis/ssp.analysis.2.sh PE250DFLT MiSeqV3
qsub -t $simphyReplicateID $HOME/src/vc-benchmark-cesga/jobs/analysis/ssp.analysis.2.sh SE250DFLT MiSeqV3
##

3. MAPPINGS

##
profiles=("PE250DFLT" "SE250DFLT") # ("SE150DFLT") # ("PE150DFLT") # ("PE150OWN") #
for profileFOLDER in ${profiles[*]};do numJobs=$(find "$HOME/data/mappings/${pipelinesName}.${replicateID}/scripts/" -name "${pipelinesName}.${replicateID}.${profileFOLDER}.bwa.commands." -type f | wc -l);
echo $numJobs
qsub -t 1-$numJobs $HOME/src/vc-benchmark-cesga/jobs/analysis/ssp.analysis.3.sh "$HOME/data/mappings/${pipelinesName}.${replicateID}/scripts/${pipelinesName}.${replicateID}.${profileFOLDER}.bwa.commands"
done
##

4. Generating BAMMING SORTING commands

##
qsub -t $simphyReplicateID $HOME/src/vc-benchmark-cesga/jobs/analysis/ssp.analysis.4.sh PE150DOWN HiSeq2500
qsub -t $simphyReplicateID $HOME/src/vc-benchmark-cesga/jobs/analysis/ssp.analysis.4.sh PE150DFLT HiSeq2500
qsub -t $simphyReplicateID $HOME/src/vc-benchmark-cesga/jobs/analysis/ssp.analysis.4.sh SE150DFLT HiSeq2500
qsub -t $simphyReplicateID $HOME/src/vc-benchmark-cesga/jobs/analysis/ssp.analysis.4.sh

PE250DFLT MiSeqV3
qsub -t $simphyReplicateID $HOME/src/vc-benchmark-cesga/jobs/analysis/ssp.analysis.4.sh
SE250DFLT MiSeqV3

##
####

5. BAMMING SORTING

##
####
profiles=("PE250DFLT" "SE250DFLT") # ("SE150DFLT") # ("PE150DFLT") # ("PE150OWN") #
for profileFOLDER in ${profiles[*]};do numJobs=$(find
"$HOME/data/mappings/${pipelinesName}.${replicateID}/scripts/" -name
"${pipelinesName}.${replicateID}.${profileFOLDER}.samtools.commands." -type f | wc -l);
echo $numJobs
qsub -t 1-$numJobs $HOME/src/vc-benchmark-cesga/jobs/analysis/ssp.analysis.5.sh
"$HOME/data/mappings/${pipelinesName}.${replicateID}/scripts/${pipelinesName}.${replicateID}
.${profileFOLDER}.samtools.commands"
done

##
####

6. INFORMATION ON THE MAPPING

##
####

To ask the view command to report solely “proper pairs” we use the -f
option and ask for alignments where the second bit is true (proper pair is true).

samtools view -f 0x2 sample.sorted.bam
How many properly paired alignments are there?

samtools view -f 0x2 sample.sorted.bam | wc -l
Now, let’s ask for alignments that are NOT properly paired. To do this,
we use the -F option (note the capitalization to denote “opposite”).

samtools view -F 0x2 sample.sorted.bam
How many improperly paired alignments are there?

samtools view -F 0x2 sample.sorted.bam | wc -l