Universidad Regional Amazónica Ikiam

Bioinformática

Andrés Benalcázar

1.10.1

Change directory to CSB/unix/sandbox.

```
User@DESKTOP-U73I4LB MINGW64 ~ (master)
$ cd Documents/CSB-master/unix/sandbox

User@DESKTOP-U73I4LB MINGW64 ~/Documents/CSB-master/unix/sandbox (master)
$ S
```

What is the size of the file Marra2014_data.fasta?

```
User@DESKTOP-U73I4LB MINGW64 ~/Documents/CSB-master/unix/sandbox (master)
$ ls -l -h
total 556K
drwxr-xr-x 1 User 197121    0 Jan 28  2019 'Papers and reviews'/
-rw-r--r-- 1 User 197121 553K Nov 18 00:17  my_file.fasta
drwxr-xr-x 1 User 197121    0 Nov 18 00:09  pruebas/
```

Create a copy of Marra2014_data.fasta in the sandbox and name it my_file.fasta.

```
User@DESKTOP-U73I4LB MINGW64 ~/Documents/CSB-master/unix/sandbox (master)
$ cp ../data/Marra2014_data.fasta my_file.fasta

User@DESKTOP-U73I4LB MINGW64 ~/Documents/CSB-master/unix/sandbox (master)
$ ls
'Papers and reviews'/   my_file.fasta   pruebas/   sandbox
```

How many contigs are classified as isogroup00036?

```
User@DESKTOP-U73I4LB MINGW64 ~/Documents/CSB-master/unix/sandbox (master)
$ grep -c isogroup00036 my_file.fasta
16
```

Replace the original “two-spaces” delimiter with a comma.

```
User@DESKTOP-U73I4LB MINGW64 ~/Documents/CSB-master/unix/sandbox (master)
$ cat my_file.fasta | tr -s " " "," > my_file2.fasta

User@DESKTOP-U73I4LB MINGW64 ~/Documents/CSB-master/unix/sandbox (master)
$ cat my_file2.fasta | head -n 3
>contig00001,length=527,numreads=2,gene=isogroup00001,status=it_thresh
ATCCTAGCTACTCTGGAGACTGAGGATTGAAGTTCAAAGTCAGCTCAAGCAAGAGATTTG
TTTACAATTAACCCACAAAAGGCTGTTACTGAAGGTGTGGCTTAAGTGTCAGAGCAACAG
```

How many unique isogroups are in the file?

```
User@DESKTOP-U73I4LB MINGW64 ~/Documents/CSB-master/unix/sandbox (master)
$ grep ">" my_file2.fasta | cut -d "," -f 4 | sort | uniq | wc -l
43
```

Which contig has the highest number of reads (numreads)? How many reads does it have?

```
User@DESKTOP-U73I4LB MINGW64 ~/Documents/CSB-master/unix/sandbox (master)
$ grep ">" my_file2.fasta | cut -d "," -f 1,3 | sort -t "=" -k 2 -n -r | head -n 1
>contig00302,numreads=3330
```

1.10.2

Literal a y c

```
User@DESKTOP-U73I4LB MINGW64 ~/Documents/CSB-master/unix/sandbox (master)
$ cat ejericio2.sh
#1.10.2
echo "Lietral a"
cut -f 1 ../data/Gesquiere2011_data.csv | grep -c -w 3

cut -f 1 ../data/Gesquiere2011_data.csv | grep -c -w 27


echo "literal b es un archivo bash aparte"


echo "lietral 3"

myIDS=`tail -n +2 ejercicio.csv | cut -f 1 | sort -n | uniq`

for id in $myIDS
do
    mycounts=`bash count_baboons.sh ../data/Gesquiere2011_data.csv $id`
    echo "ID:" $id "counts:" $mycounts
done


tail -n +2 ejercicio.csv | cut -f 1 | sort -n | uniq
```

Literal b

```
User@DESKTOP-U73I4LB MINGW64 ~/Documents/CSB-master/unix/sandbox (master)
$ nano literal_b

User@DESKTOP-U73I4LB MINGW64 ~/Documents/CSB-master/unix/sandbox (master)
$ cat literal_b
#Literal b
cut -f 1 $1 | grep -c -w $2

User@DESKTOP-U73I4LB MINGW64 ~/Documents/CSB-master/unix/sandbox (master)
$ nano literal_b.sh

User@DESKTOP-U73I4LB MINGW64 ~/Documents/CSB-master/unix/sandbox (master)
$ cat literal_b.sh
#literalb

cut -f 1 $1 | grep -c -w $2


User@DESKTOP-U73I4LB MINGW64 ~/Documents/CSB-master/unix/sandbox (master)
$ bash literal_b.sh ejercicio.csv 3
61
```

1.10.3

```
User@DESKTOP-U73I4LB MINGW64 ~/Documents/CSB-master/unix/sandbox/Saavedra_2013 (master)
$ nano ejercicio3.sh

User@DESKTOP-U73I4LB MINGW64 ~/Documents/CSB-master/unix/sandbox/Saavedra_2013 (master)
$ cat ejercicio3.sh
echo "literal a"

 echo "El numero de columna de n7.txt es"
     head -n 1 n7.txt | grep -o " " | wc -l
  echo "el numero de filas de n7.txt es"
      wc -l n7.txt

echo "literal b"

for i in n*.txt

do
  echo "El numero de columna es"
     head -n 1 $i | grep -o " " | wc -l
  echo "el numero de filas es"
      wc -l $i

  echo "fin de este archivo"
done


User@DESKTOP-U73I4LB MINGW64 ~/Documents/CSB-master/unix/sandbox/Saavedra_2013 (master)
$ bash ejercicio3.sh
```

```
User@DESKTOP-U73I4LB MINGW64 ~/Documents/CSB-master/unix/sandbox/Saavedra_2013 (master)
$ cat ejercicio3.sh
echo "literal a"

 echo "El numero de columna de n7.txt es"
     head -n 1 n7.txt | grep -o " " | wc -l
  echo "el numero de filas de n7.txt es"
      wc -l n7.txt

echo "literal b"

for i in n*.txt

do
  echo "El numero de columna es"
     head -n 1 $i | grep -o " " | wc -l
  echo "el numero de filas es"
      wc -l $i

  echo "fin de este archivo"
done

echo "literal c"

for i in n*txt

do
  echo "mayor cant columnas"
    head -n 1 $1 | grep -o " " | wc -l | sort -r | head -n 1
  echo "mayor cant filas"
    wc -l $i | sort -r | head -n 1
done
```