**LABORATIO I: GABRIELA LÓPEZ-ALONSO Y LAURA FABREGA**

Ejercicio 1: (4 ptos)

Usa el fichero con formato .vcf file.vcf. Explora el archivo usando el comando head. Aquí hay una explicación del tipo de archivo, aunque se comentará en clase.

Selecciona aquellos SNPs que han pasado todos los filtros (Tienen el texto PASS). Selecciona también aquellos SNPs que no han pasado el filtro de calidad (Tienen el texto q10). Indica que comandos debes usar si quieres generar un archivo de texto con esa información.

grep “PAS” file.vcf

grep “q10” file.vcf

grep “PAS” file.vcf > reultado.txt

grep “q19” file.vcf >> reultado.txt

nano resultado.txt (para ver el archivo de texto creado)

Tenemos un nuevo fichero file2.vcf que ha perdido la información de la cabecera. Sabemos que la cabecera del fichero file.vcf nos vale, añadir la cabecera de file.vcf al fichero file2.vcf creando un nuevo fichero. Indica los comandos usados.

head -n 1 file.vcf >> cabeza.txt

cat cabeza.txt file2.vcf >> file3.vcf

nano file3.vcf (para ver lo creado)

Ejercicio 2: (5 ptos)

Preparar un directorio de trabajo, llamado ejercicio2

mkdir “ejercicio2”

cd “ejercicio2”

En el directorio ejercicio2, descargar (mediante la línea de comandos) este link: https://ftp.ncbi.nlm.nih.gov/genomes/GENOME\_REPORTS/prokaryotes.txt con el archivo prokaryots.txt que contiene todos los genomas secuenciados de procariotas hasta la fecha. Para ello usa el comando:

curl https://ftp.ncbi.nlm.nih.gov/genomes/GENOME\_REPORTS/prokaryotes.txt -o prokaryotes.txt

Observa la información que hay en el archivo descargado.

En el mismo directorio crea un nuevo archivo de texto llamado akkermansia.txt donde se seleccionen los siguientes aspectos (Indica los comandos usados)

La cabecera del archivo original prokaryots.txt

Todos los genomas de “Akkermansia”

head -n 1 prokaryotes.txt > akkermansia.txt

grep “akkermansia” prokaryotes.txt >> akkermansia.txt

Sobre el archivo akkermansia.txt conteste indicando el comando usado:

¿Cómo pueden verse las primeras 10 líneas del archivo?

head -n 10 akkermansia.txt

¿Cuántos genomas de Akkermansia muciniphila hay?

grep “akkermansia muciniphila” prokaryotes.txt

El número de genomas de akkermansia muciniphila es de 1205

¿Cuántos de Akkermasia biwaensis?

grep “akkermansia biwaensis” prokaryotes.txt

El número de genomas de akkermansia muciniphila es de 1

Por último elimina el archivo prokaryots.txt

rm “prokaryotes.txt”