## Bioinformática y análisis de datos ómicos

© Ignacio Varela Egocheaga Este material se publica bajo licencia Creative Commons CC BY-NC-SA 4.0

## <u>EJERCICIO TEMA 7</u>

## <u>Alumno:</u>

Realiza las distintas actividades de abajo copiando debajo de cada actividad los comandos necesarios. Para ello descárgate la <u>carpeta para realizar los ejercicios del Tema 7</u>. Finalmente, convierte este documento a PDF y entrégalo a través del moodle.

1.- Utiliza el programa BWA para generar los índices Burrows-Wheeler del archivo Mus\_musculus.GRCm39.dna.chromosome.6.fa

2.- Alinea los archivos fastq de las muestras 1181-1 y 1181-6 frente al genoma de referencia que has indexado utilizando el programa BWA.

3.- Utiliza la herramienta SAMTOOLS para convertir los archivos a formato BAM (samtools view), arreglar las coordenadas de las secuencias pareadas (samtools fixmate), ordenar los archivos BAM (samtools sort) y generar un índice para los archivos BAM (samtools index).

4.- Utiliza la herramienta SAMTOOLS para generar archivos en formato mpileup, con la opción de especificar el genoma de referencia, con las coordenadas del genoma de las que se ha generado información de secuencia.

5.- Utiliza la herramienta VARSCAN para identificar tanto mutaciones de un solo nucleótido (SNP) como pequeñas inserciones y deleciones (indels) en cada una de las dos muestras imprimiendo las mutaciones en formato vcf

6.- Utiliza la herramienta snpEff para anotar la consecuencia funcional de las mutaciones identificadas en cada uno de los archivos. Utiliza la versión GRCm39.105 del modelo de genes.

7.- Alinea los archivos de secuencia de la muestra 788-6 frente al genoma de referencia que has indexado en el ejercicio 1. Convierte el archivo en formato BAM, arregla las coordenadas de las secuencias pareadas con la opción -m, ordena e indexa el archivo BAM resultante. Finalmente anota los duplicados de PCR utilizando la herramienta SAMTOOLS y vuelve a indexar el BAM generado.

8.- Utiliza la herramienta DELLY para identificar reorganizaciones genómicas en la nueva muestra alineada.

9.- Convierte el archivo de salida generado en formato vcf utilizando bcftools.