

Análisis de datos ómicos - Primera prueba de evaluación continua

Aràntzazu Alonso Carrasco

12 de abril, 2023

1. Descarga del archivo

Primero, debemos descargar el archivo *fastq* con el que trabajaremos a lo largo de la Prueba de Evaluación Continua (PEC). Para hacerlo, debemos ir al *Google drive* de la asignatura y buscar el archivo que corresponda. Dado que el último dígito numérico de mi DNI es un 4, el archivo que debo descargar corresponde a la muestra “S04_Bac03_read1.fastq”.

2. Descripción de los datos

Una vez descargados los datos, podemos proceder a describirlos. El formato del archivo es *.fastq*, que es un formato ampliamente utilizado en bioinformática para representar datos de secuenciación de ADN o ARN en bruto. Concretamente, los archivos en formato *.fastq* contienen secuencias de nucleótidos junto con información de calidad asociada a cada nucleótido de la secuencia. Los archivos *.fastq* son muy utilizados en la secuenciación de nueva generación (NGS) para representar datos de secuenciación de alta calidad generados por tecnologías como Illumina.

Particularmente, el archivo *S04_Bac03_read1.fastq* contiene 250000 secuencias.

FALTA COMPLETAR

3. Control de calidad con Galaxy