

DEBATE 2 INTRODUCCIÓN A LA ULTRASECUENCIACIÓN

En comparación con las técnicas de secuenciación de nueva generación, la secuenciación Sanger tiene la ventaja de proporcionar secuencias más precisas y confiables en secuencias cortas. Sin embargo, su capacidad de secuenciación masiva es limitada, ya que solo puede secuenciar una molécula a la vez, lo que la hace menos eficiente y más costosa en términos de tiempo y recursos. Por otro lado, las técnicas de secuenciación de nueva generación, como la secuenciación masiva paralela, son capaces de secuenciar miles o incluso millones de moléculas de ADN simultáneamente, lo que las hace más rápidas y eficientes en términos de costos. Aunque estas técnicas son menos precisas que la secuenciación Sanger, pueden compensar esta desventaja mediante la producción de grandes cantidades de datos, lo que las hace más adecuadas para proyectos de secuenciación a gran escala.

Es mi primera vez utilizando Galaxy y fastqc. He decidido utilizar la misma fuente de la que provienen los datos del tutorial: <https://zenodo.org/record/3736457#.ZBmB5HbMKUk>.

He utilizado los datos del siguiente archivo fastq: 3_OHara_S2_18S_2019_minq7.fastq. Se trata de la secuencia número 2 del gen ribosomal 18s comúnmente utilizado para la identificación y clasificación de especies. Para llevar a cabo la secuenciación se utilizó Sanger/Illumina 1.9. Observamos una calidad de los datos bastante baja. En el mismo nombre del archivo se indica la baja calidad de los datos. Minq7 es el valor de calidad mínimo utilizado para filtrar las secuencias de baja calidad en el proceso de secuenciación de ADN. Los motivos por los que la calidad de la secuencia es baja pueden ser: errores en la preparación de la muestra, fragmentación del ADN, errores de secuenciación, contaminación de la muestra o variabilidad natural del ADN (si es una región repetitiva puede ser difícil de secuenciar con precisión).

