



Propuesta de Trabajo Fin de Máster

Año académico 2024-2025

MÁSTER EN MÉTODOS COMPUTACIONALES EN CIENCIAS

Proyecto Nº 07

Título: Optimización de una pipeline de análisis metataxonómico a partir de long reads de ONT

Departamento/ Laboratorio: BIOQUÍMICA Y GENÉTICA

Director: FRANCISCO JAVIER NOVO VILLAVERDE

Correo electrónico: fnovo@unav.es

Codirector: SERGIO ROA GÓMEZ

Correo electrónico: sroa@unav.es

Resumen:

La utilización de long-read sequencing tiene ciertas ventajas sobre las short reads a la hora de clasificar bacterias usando la secuencia del ADN ribosomal 16S. En el grupo tenemos la capacidad de secuenciar en un secuenciador Mk1C de Oxford Nanopore Technologies (ONT) y hemos adquirido cierta experiencia en varios proyectos y TFGs previos. ONT ha actualizado recientemente sus workflows de análisis de EPI2ME, que ahora se ejecutan en entorno Linux usando nextflow y docker, permitiendo ajustar más parámetros de análisis y alinear las secuencias frente a distintas bases de datos de 16S.

En este proyecto se plantea optimizar una pipeline completa de análisis, partiendo de los archivos fastq de long-reads generadas por el Mk1C y utilizando el nuevo workflow 16S de EPI2ME para realizar el alineamiento y generar la tabla de ASVs y la tabla de taxonomía. A continuación, se generará un objeto phyloseq en R/Bioconductor para los análisis de diversidad y abundancia diferencial, y se integrará con MicrobiomeAnalyst para los análisis funcionales.

La optimización se llevará a cabo con un dataset de benchmarking generado por nosotros, en el que se estudiará la influencia de variar los distintos parámetros de análisis en los resultados finales. Finalmente, la pipeline desarrollada se aplicará al estudio de un nuevo dataset que se generará en el laboratorio a lo largo del proyecto.

Presupuesto solicitado (máximo 1500€) y breve justificación

Se estima un gasto de 1.500 euros que cubra la compra de una flowcell MINION FLO-MIN114 (€665.00), suficiente para secuenciar 12 muestras con suficiente profundidad) y un 16S Barcoding Kit 24 V14 (€855.00) para generar las librerías.

OPTATIVAS RECOMENDADAS

1. Análisis e interpretación de datos de alto rendimiento
2. Programación avanzada
3. Machine Learning I
4. Machine Learning II