



UNIVERSIDAD
DE GRANADA



Facultad de Ciencias



BIOTECNOLOGÍA
UGR

Propuesta TFG_BIOTEC
Curso: 2023-24
DEPARTAMENTO: Ciencias de la
computación

CÓDIGO DEL TFG: CCIA-03

1. DATOS DEL TFG OFERTADO:

Título: Poniendo a prueba las nuevas tecnologías para generar genomas perfectos. ¿Estamos tan cerca como pensamos?

Resumen (máx 250 palabras, estructurado en Introducción, Objetivos y Plan de trabajo):

* Introducción:

Actualmente, los enormes avances relacionados con técnicas de secuenciación masiva (*NGS*) han hecho posible una resolución nunca visto anteriormente. En el campo de la genómica, se ha hecho posible obtener el genoma de un organismo, con precios y condiciones muy accesibles. Las técnicas de secuenciación de lecturas largas (*long reads*) han sido reconocidas como “Método del Año” por la revista *Nature* en 2022, y el desarrollo tecnológico es constante, con múltiples empresas y nuevos protocolos prometiendo la obtención de “genomas perfectos”. Sin embargo, las *long reads* presentan problemas metodológicos, que en última instancia afectan a la precisión e integridad de las secuencias obtenidas. Dado que el análisis bioinformático de la mayoría de técnicas *NGS*-ómicas requiere en primer lugar el uso de un genoma considerado “de referencia”, posibles errores de secuenciación pueden ser cruciales en estudios de genómica funcional. Por ello, en paralelo al desarrollo de nuevas técnicas de secuenciación y de “ensamblado” de las secuencias para obtener genomas, se han desarrollado herramientas bioinformáticas para corregir las secuencias y asegurar su exactitud. Tradicionalmente, se han basado en aproximaciones “híbridas”, que utilizan varias tecnologías de secuenciación para complementar los puntos débiles.

* Objetivo:

En este proyecto, se utilizarán datos genómicos ya publicados y se compararán diferentes soluciones para ensamblar y corregir automáticamente las secuencias.

* Plan de trabajo:

1. Búsqueda bibliográfica de herramientas y *datasets* disponibles
2. Diseño experimental para el análisis de los datos y *benchmarking* de las herramientas
3. Ejecución del flujo de trabajo y generación de resultados
4. Preparación de la memoria, exposición del TFG, y posible publicación científica

Tabla de actividades y dedicación
estimada:

Planteamiento, orientación, supervisión, y preparación de la memoria	20
Preparación de la memoria	9
Desarrollo del trabajo	120
Exposición del trabajo	1
TOTAL (6 ECTS)	150 horas

OFERTADO POR:

Profesor del Departamento

Profesor del Departamento junto con Empresa o Institución

Propuesto por alumno (*)

(*) En el caso de TFG propuesto por alumno, por favor completar la siguiente información sobre el mismo:

Apellidos:

Nombre:

e-mail institucional:

2. MODALIDAD: 5

1. Estudio de casos, teóricos o prácticos, bibliográficos relacionados con la temática del Grado
2. Elaboración de un informe o un proyecto de naturaleza profesional
3. Elaboración de un plan de empresas
4. Simulación de encargos profesionales
5. Trabajos experimentales, de toma de datos
6. Trabajos derivados de la experiencia desarrollada en prácticas externas

3. COMPETENCIAS Y RESULTADOS DEL APRENDIZAJE:

Básicas y generales: CG5, CB2-5

Transversales: CT1-5, CT8-9

Específicas: CE2, CE3, CE5

4. BIBLIOGRAFIA BÁSICA:

*Method of the Year 2022: long-read sequencing. Nat Methods 20, 1 (2023).

*Ruiz, José L., et al. "From contigs to chromosomes: automatic Improvement of Long Read Assemblies (ILRA)." bioRxiv (2021): 2021-07.

*Zhang, X. et al. Benchmarking of long-read sequencing, assemblers and polishers for yeast genome. Brief Bioinform 23, 1–13 (2022).

*Southwood, Dean, et al. "Pyro: A Comprehensive Pipeline for Eukaryotic Genome Assembly." bioRxiv (2023): 2023-04.

*Southwood, Dean, et al. "Exhaustive benchmarking of de novo assembly methods for eukaryotic genomes." bioRxiv (2023): 2023-04.

5. ACLARACIONES PARA EL ESTUDIANTE:

El proyecto tendrá un fuerte componente formativo en cuanto a su carácter bioinformático. No se espera un gran conocimiento previo, pero sí un interés genuino por las Ciencias de la Salud y por formarse en el campo de la Biología Computacional. Se proveerá la infraestructura computacional necesaria, y se priorizará en todo momento el éxito del TFG del estudiante, así como la publicación científica del trabajo si fuera posible.

3. DATOS DEL TUTOR/A UGR:

Apellidos: Del Val Muñoz
Teléfono: 958240468

Nombre: Coral
e-mail: delval@ugr.es

**En el caso de trabajos desarrollados en Empresas u otras Instituciones ajenas a la Universidad de Granada, por favor completar la siguiente información:

TUTOR/A DE LA EMPRESA O INSTITUCIÓN:

Apellidos: RUIZ RODRÍGUEZ

Empresa/Institución: INSTITUTO DE PARASITOLOGÍA Y BIOMEDICINA LÓPEZ-NEYRA (IPBLN-CSIC)

Teléfono: 958181640

Nombre: JOSÉ LUIS

e-mail: JOSELUIS.RUIZ@CSIC.ES