



1. DATOS BÁSICOS DEL TFG:

Título: Análisis de célula única en cáncer: una propuesta FAIR para el estudio de datos abiertos

Descripción general (resumen y metodología):

El avance revolucionario de las tecnologías de secuenciación de ARN a nivel de célula única (scRNA-seq) ha transformado nuestra comprensión de la heterogeneidad tumoral, revelando la extraordinaria diversidad celular que caracteriza el microambiente del cáncer.

Este Trabajo de Fin de Grado aborda la intersección entre la biología del cáncer y la ciencia de datos abiertos, proponiendo un análisis exhaustivo de datos de célula única en muestras oncológicas bajo el riguroso prisma de los principios FAIR (Encontrables, Accesibles, Interoperables y Reutilizables). Mediante la aplicación de herramientas bioinformáticas de vanguardia como Seurat y Scanpy a conjuntos de datos públicos cuidadosamente seleccionados, se busca no solo identificar firmas moleculares específicas y poblaciones celulares relevantes en el desarrollo tumoral, sino también establecer un flujo de trabajo transparente, documentado y completamente reproducible que sirva como modelo metodológico para futuras investigaciones en este campo emergente, contribuyendo así al avance de la medicina de precisión y la ciencia abierta en oncología.

Tipología: Estudio de casos, teóricos o prácticos, relacionados con la temática del Grado.

Objetivos planteados:

Seleccionar y recopilar conjuntos de datos públicos de scRNA-seq en cáncer provenientes de repositorios como GEO o ArrayExpress, asegurando que dispongan de adecuada anotación clínica y experimental.

Aplicar metodologías de análisis bioinformático utilizando herramientas como Seurat (R) o Scanpy (Python) para:

- Identificar los distintos tipos celulares presentes en el microambiente tumoral
- Determinar genes marcadores específicos de cada población celular
- Caracterizar patrones de expresión génica relevantes en el contexto oncológico

Implementar los principios FAIR en todo el proceso analítico, garantizando que tanto los datos de entrada como los resultados sean:

- Encontrables (Findable): mediante metadatos adecuados y identificadores únicos
- Accesibles (Accessible): a través de protocolos estandarizados y abiertos
- Interoperables (Interoperable): usando formatos y vocabularios comunes
- Reutilizables (Reusable): con licencias claras y documentación detallada

Documentar exhaustivamente el flujo de trabajo mediante cuadernos interactivos (Jupyter o RMarkdown), asegurando la reproducibilidad completa del análisis.

Compartir públicamente el código, los datos procesados y los resultados en plataformas abiertas como GitHub y Zenodo, contribuyendo así a la comunidad científica y promoviendo las buenas prácticas en ciencia abierta y bioinformática reproducible.

Bibliografía básica:

Hao, Y., Stuart, T., Kowalski, M.H. et al. Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat Biotechnol* **42**, 293–304 (2024). <https://doi.org/10.1038/s41587-023-01767->

y

Camacho, J; Sorchan Armstrong; García-Mátinez, L; Díaz, Gómez-Llorente, C; Single-cell spatial (scs) omics: Recent developments in data analysis. <https://doi.org/10.1016/j.trac.2024.118109.183-118109>

Hrovatin, K., Sikkema, L., Shitov, V.A. et al. Considerations for building and using integrated single-cell atlases. *Nat Methods* **22**, 41-57 (2025). <https://doi.org/10.1038/s41592-024-02532-y>

Recomendaciones y orientaciones para el estudiante:

- El trabajo se podrá realizar de manera telemática realizando las reuniones con el tutor mediante videollamadas

Plazas: 1

2. DATOS DEL TUTOR/A:

Nombre y apellidos: MARÍA CORAL DEL VAL MUÑOZ

Ámbito de conocimiento/Departamento: CIENCIAS DE LA COMPUTACIÓN E INTELIGENCIA ARTIFICIAL

Correo electrónico: delval@ugr.es

3. COTUTOR/A DE LA UGR (en su caso):

Nombre y apellidos:

Ámbito de conocimiento/Departamento:

Correo electrónico:

4. COTUTOR/A EXTERNO/A (en su caso):

Nombre y apellidos: Laura Terrón Camero

Correo electrónico: laura.terron@csic.es

Nombre de la empresa o institución: INSTITUTO DE PARASITOLOGÍA Y BIOMEDICINA LÓPEZ-NEYRA (IPBLN-CSIC)

Dirección postal: Parque Tecnológico de Ciencias de la Salud Avenida del Conocimiento, 17 18016 Armilla (Granada)

Puesto del tutor en la empresa o institución: Investigadora

Centro de convenio Externo: INSTITUTO DE PARASITOLOGÍA Y BIOMEDICINA LÓPEZ-NEYRA (IPBLN-CSIC)

5. DATOS DEL ESTUDIANTE:

Nombre y apellidos: ALVARO EXTREMERA PALOMINO

Correo electrónico: alvaroextrmera@correo.ugr.es