



Propuesta de Trabajo Fin de Máster

Año académico 2026-2027

MÁSTER EN CIENCIA DE DATOS PARA CIENCIAS EXPERIMENTALES

Proyecto Nº 14

Título: Análisis Multimodal e Integración de Datos en Transcriptómica Espacial de Alta Resolución (Visium HD)

Departamento/ Laboratorio: Laboratorio de Terapia celular adoptiva. Programa de Inmunología e Inmunoterapia. División Cáncer. CIMA

Director: Sandra Hervás Stubbs

Correo electrónico: mshervas@unav.es

Codirector: Ángel M. Martínez Montes

Correo electrónico: amartinezmon@unav.es

Resumen:

Introducción: ¿Qué es el TCR y por qué es importante?

Para entender este proyecto, es clave conocer el **TCR (T-Cell Receptor)**. Los receptores de células T son moléculas en la superficie de los linfocitos T que les permiten reconocer y atacar amenazas, como células tumorales. Cada "clon" de linfocito tiene un TCR específico (un **clonotipo**). En este estudio, mapear el TCR no solo nos dice *qué* células inmunes hay, sino *exactamente dónde* están dentro del tumor y cómo interactúan con su entorno.

Resumen del Proyecto

La tecnología **Visium HD (10x Genomics)** representa la última frontera en transcriptómica espacial, permitiendo explorar la expresión génica en cortes de tejido con una resolución de 2 μm , cercana a la escala unicelular. A pesar de su potencial, este aumento de resolución genera retos computacionales críticos: la segmentación precisa de células individuales y la integración de datos masivos. Nuestro laboratorio ha optimizado, además, una técnica pionera para la secuenciación de TCR directamente desde Visium HD. Esto permite vincular la expansión clonal del sistema inmunitario con localizaciones espaciales específicas y perfiles de expresión génica local. En este TFM, el estudiante aplicará herramientas bioinformáticas avanzadas y modelos de **Machine Learning** para resolver la segmentación celular, analizar el repertorio de TCR e integrar estos hallazgos con datos de secuenciación de **Single-Cell**. El objetivo fundamental es localizar con precisión los TCR relevantes, identificándolos mediante el perfil molecular de las células T que los expresan, su ubicación estratégica en el tejido y sus interacciones con las células vecinas. Para ello, se trabajará con datos reales de pacientes con tumores ginecológicos generados íntegramente en nuestro centro.

Objetivos Específicos

1. **Segmentación Celular Avanzada:** Revisar y comparar algoritmos de segmentación (basados en imágenes H&E y señales transcriptómicas) para definir bordes celulares en datos de Visium HD.
2. **Análisis de Repertorio Inmune:** Analizar y procesar las librerías de expresión génica (GEX) y TCR-seq espacial para identificar clonotipos de células T.
3. **Integración Multimodal:** Integrar los resultados espaciales con datos de *single-cell* mediante técnicas de desvolución o transferencia de etiquetas para asignar tipos celulares con alta precisión.



4. **Desarrollo de Pipeline:** Crear un flujo de trabajo (*pipeline*) reproducible y escalable en **Python/R** que sea aplicable a futuros proyectos de transcriptómica espacial.

Perfil del Alumno

Estudiantes del **Máster en Ciencias de Datos** con interés en la medicina de precisión, el análisis de imágenes y la integración de datos ómicos.

Requisitos valorados:

- Dominio intermedio de **Python o R**.
- Conocimientos básicos de **Machine Learning** (especialmente en clustering y visión por computador).
- Interés por la biología computacional.

Datos Disponibles

El alumno contará con un dataset exclusivo del laboratorio que incluye datos de pacientes con tumores ginecológicos

- Datos de GEX y TCR-seq de **Visium HD**.
- Datos complementarios de GEX y TCR-seq de **secuenciación single cell** los mismos pacientes

OPTATIVAS RECOMENDADAS

1. Advance topics in machine learning.
2. Procesamiento de imágenes