



**Propuesta de Trabajo Fin de Máster**

Año académico 2026-2027

**MÁSTER EN CIENCIA DE DATOS PARA CIENCIAS EXPERIMENTALES**

<b>Proyecto Nº 4</b>
<b>Título: Análisis de microproteínas para identificar nuevas terapias antitumorales</b>
<b>Departamento/ Laboratorio:</b> DNA and RNA therapy. Lab 406. CIMA/UNAV
<b>Director:</b> Puri Fortes <b>Correo electrónico:</b> pfortes@unav.es
<b>Resumen:</b> <p>La mayor parte de las terapias que se han trasladado desde el laboratorio hasta la clínica se basan en el estudio de proteínas conservadas, cuyo número en humanos es muy limitado, casi 30.000. Sin embargo, durante los últimos años ha quedado claro que existe un gran número de RNAs no codificantes funcionales, tanto de tamaño corto (micro RNAs o miRNAs) como largo (RNAs largos no codificantes), que tienen papeles fundamentales en el funcionamiento celular y en el desarrollo de enfermedades. Por ello, se están estudiando en la actualidad como biomarcadores y como dianas terapéuticas. Recientemente se ha descubierto que gran parte del RNA que se consideraba “no codificante” produce microproteínas. Con ello, ha comenzado una nueva era de investigación en la que es necesario identificar y caracterizar estas microproteínas y afianzar su potencial terapéutico. Hasta la fecha se sabe que las microproteínas son factores esenciales en la frontera de la evolución: están poco conservadas, pero son muy numerosas y, aquellas que se han estudiado en detalle, son esenciales para la viabilidad celular. Su estudio representa un campo de un atractivo sin precedentes para el análisis bioinformático y la aplicación de la inteligencia artificial. Con este objetivo, se realizó un trabajo de Máster el curso pasado que continúa ahora en una tesis doctoral y que sirvió para identificar características clave de las microproteínas que se producen en cáncer de mama triple negativo y en hepatocarcinoma. Disponemos ahora en el laboratorio de más datos masivos de microproteínas obtenidos por secuenciación masiva (55.290 microproteínas) o por análisis proteómicos (102.306 microproteínas) y datos de pacientes con información clínica asociada. Estos datos serán esenciales para buscar características especiales en la secuencia del gen o de las microproteínas que permita su traducción, estabilización y funcionalidad. Buscaremos también asociaciones con pacientes que nos permitan predecir su pronóstico o la respuesta a un tratamiento, sobre todo, la inmunoterapia. Buscaremos aquellas microproteínas que sólo se producen en el tumor y que puedan ser excelentes candidatos para desarrollar vacunas contra el cáncer. Nuestro objetivo es que todos los resultados tengan seguimiento en el laboratorio y/o en la clínica y que impacten en los pacientes con hepatocarcinoma y cáncer de mama triple negativo.</p>

**OPTATIVAS RECOMENDADAS**

1. Análisis e interpretación de datos de alto rendimiento
2. Análisis de datos en biología
3. Advance topics in Machine learning
4. Análisis de secuencias y bioinformática estructural