



Propuesta de Trabajo Fin de Máster

Año académico 2026-2027

MÁSTER EN CIENCIA DE DATOS PARA CIENCIAS EXPERIMENTALES

Proyecto Nº 39

Título: Discovery of Shared Non-Canonical Tumor Antigens and Candidate T-Cell Receptors for Chronic Lymphocytic Leukemia Immunotherapy

Departamento/ Laboratorio: Machine Learning en Biomedicina, CIMA

Director: Mikel Hernáez

Correo electrónico: mhernaez@unav.es

Codirector:

Correo electrónico:

Resumen:

Peptide-vaccine strategies for **chronic lymphocytic leukemia** (CLL) currently use warehouse antigen collections. These are premanufactured peptide collections that have to be combined into personalized peptide cocktails designed for each patient's HLA type. This personalization limits their broad applicability ([Nelde et al., 2021](#)). Ribosome profiling has revealed that non-canonical open reading frames are a substantial source of **tumor specific antigens** in CLL ([Ouspenskaia et al., 2022](#)). Moreover, recent immunopeptidomic studies in bone-marrow malignancies found recurrent non-canonical antigens with different HLA types and even across different cancer types ([Kehl et al., 2026](#)). These findings suggest that existing CLL immunopeptidomic data may contain shared or “public” antigens that were missed by searches restricted to canonical proteins.

This project will reanalyse the CLL **immunopeptidomics** dataset PXD024871 using the CLL-specific **Ribo-seq**-derived sequences in nuORFdb as search space, identifying recurrent non-canonical peptides that are potentially shared across the whole cohort regardless of their HLA genotype. Then, we will use **structural diffusion models** to generate and prioritize candidate paired TCR α /TCR β sequences predicted to recognize selected peptide–HLA complexes. This would allow the interrogation of single-cell V(D)J datasets to validate that the T-Cell clonotypes binding these novel antigens are expanded and transcriptionally activated in vivo. Finally, experimental assays will confirm antigen specificity.

OPTATIVAS RECOMENDADAS

1. Deep Learning
2. Adv. Machine Learning
3. Procesamiento de imágenes
- 4.