

## Archivos utilizados en el transcurso de la investigación

Nombre de los archivos utilizados en el transcurso de la presente tesis, así como su formato y descripción del contenido del archivo disponible.

Formato	Nombre	Descripción de contenido
PDB	<alelo>	9 archivos diferentes de las estructuras de HLA de clase I y II.
FASTA	proteinas	Secuencias de las 13 proteínas de <i>M. tuberculosis</i> .
FASTA	peptidos_A	Secuencias de los péptidos generados por el ventaneo de la población A (9 residuos de aminoácidos).
FASTA	peptidos_B	Secuencias de los péptidos generados por el ventaneo de la población B (15 residuos de aminoácidos).
TXT	poblacion_A	Secuencias de los péptidos filtrados con CD-HIT de la población A (9 residuos de aminoácidos).
TXT	poblacion_B	Secuencias de los péptidos filtrados con CD-HIT de la población B (15 residuos de aminoácidos).
CSV	embeddings_A	Embeddings (encajes) generados con ESM2 de la población A.
CSV	embeddings_B	Embeddings (encajes) generados con ESM2 de la población B.
CSV	preprocesado_A	Predicciones de las características de la población A.
CSV	preprocesado_B	Predicciones de las características de la población B.
TXT	poblacion_pruebaA	Secuencias de los péptidos de prueba.
CSV	poblacion_pruebaA	Predicciones de las características de la población de prueba.
FASTA	control_A	Secuencias de los péptidos control generados por el ventaneo para la población A (9 residuos de aminoácidos).
TXT	control_A	Secuencias de los péptidos control filtrados con CD-HIT para la población A (9 residuos de aminoácidos).
CSV	embeddings_PruebaA	Embeddings (encajes) generados con ESM2 de la población de prueba A.
CSV	evaluacion_controlA	Evaluación de los péptidos control para la población A (predicciones y promedio).
CSV	evaluacion_controlB	Evaluación de los péptidos control para la población B (predicciones y promedio).
PYTHON	June-A	Algoritmo evolutivo para la población A
PYTHON	June-B	Algoritmo evolutivo para la población B
PKL	A_<1,2,3,4,5>	Semillas del código para replicabilidad de resultados, de la población A.
PKL	B_<1,2,3,4,5>	Semillas del código para replicabilidad de resultados, de la población B.