

METHOD, DEVICE AND STORAGE MEDIUM FOR IDENTIFYING PEPTIDES CAPABLE OF BINDING TO MHC-I/HLA-I AND BEING RECOGNIZED BY TCR

ABSTRACT

The present invention discloses a method, a device, and a storage medium for identifying peptides capable of binding to MHC-I/HLA-I and being recognized by TCR. The method includes: acquiring protein sequences, MHC-I/HLA-I sequences, and peptide sequences; training a protein sequence tokenizer using the protein sequences; training a protein representation model using the protein sequences tokenized by the protein sequence tokenizer to obtain a trained protein representation pre-trained model; fine-tuning the trained protein representation pre-trained model using the MHC-I/HLA-I sequences and the peptide sequences tokenized by the protein sequence tokenizer, to obtain a binary classification model; and identifying target wild-type/mutant peptides and classifying mutant peptides by using the binary classification model, to obtain candidate neoantigen mutant peptides that are positively correlated with TCR recognition.

[Figure 1]