[BIOINFO 2023]

AMP-BERT: Prediction of antimicrobial peptide function based on a BERT model

Hansol Lee[1], Songyeon Lee[1], Ingoo Lee[1], and Hojung Nam[1,2*]

[1]*School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology (GIST)*
[2]*AI Graduate School, Gwangju Institute of Science and Technology*
**Corresponding author: hjnam@gist.ac.kr*

The evolution of antimicrobial resistance in bacteria, viruses, and other harmful microbes has become a serious health concern. As possible alternatives to traditional antibiotics, antimicrobial peptides (AMPs) have been studied and developed to counter the resistance problem. AMPs are innately produced host defense peptides, which protect against unwanted microorganisms using non-specific mechanisms that hinder the microorganisms from developing resistance. We present AMP-BERT, a fine-tuned deep learning AMP classification model based on the bidirectional encoder representations from transformers (BERT) architecture that can extract structural information from peptide sequences and predict input samples as AMP or non-AMP. In comparison to other machine/deep learning-based methods for AMP classification, our model achieved the best prediction results when applied to our curated external dataset. Furthermore, we performed attention analysis on the sequence residue features from AMP-BERT that showed which residues in known AMPs contribute to the overall structure and antimicrobial function. Our findings demonstrate AMP-BERT's ability to learn the structural characteristics of peptides through model training and to classify input sequences as AMP or non-AMP. We expect AMP-BERT to play a significant role in the pre-screening process for potential AMPs, an important step in the AMP research and drug discovery field.