Deep Learning Prediction and in vitro Validation of Novel Anti-Cancer Peptides From Marine Taxa Database

Koushik, Aditya (School: La Cueva High School)

Anti-cancer peptides (ACPs), a class of small peptide molecules, have gained increasing attention in cancer research due to their ability to target and kill cancer cells selectively, sparing normal cells. ACPs, unlike chemotherapy, have less toxicity and fewer side effects, are highly specific to cancer cells, are easy to synthesize and modify, and are cost-effective therapeutics. Unfortunately, the experimental identification of novel ACPs is time-consuming and expensive, therefore computational methods to identify key features of ACPs is promising. Here, a robust deep-learning model was developed that recognized molecular features of MCF-7 breast-cancer ACPs from the CancerPPD database. Next, 20,000 peptide entries from marine taxa catalogued in the J. Craig Venter Institute PhyloDB database were screened for potentially novel ACPs by the Deep learning model. After 3-fold cross-validation, the model showed a sensitivity and specificity of 95% and an accuracy of 98% in predicting known ACPs in the CancerPPD database. After learning features of previously validated MCF-7 ACPs, the model selected top 40 novel ACPs in the PhyloDB database with >90% ACP probability. MCF-7 Breast Cancer Cell culture validation of the top 4 novel ACPs shows that the ACPs exhibit a statically significant cytotoxic effect on cancer cells at concentrations above 10-100 µg/ml. In summary, for the first time, the deep learning approach described here applies learned information on ACPs to make new predictions in an unknown dataset, and would present the first AI (artificial intelligence) predicted novel ACPs to be validated in vitro.