Abstract No: 116 ***Life Sciences***

IN SILICO DRUG DISCOVERY USING THE KNIME ANALYTICS PLATFORM TO IDENTIFY PHYTOCHEMICALS WITH *VEGFR-2* INHIBITORY PROPERTIES

M.R.V. Perera[1], L.R.LS. Kumari[2] and W.R.P. Wijesinghe[2*]

[1]Department of Molecular Biology and Biotechnology, Faculty of Science, University of Peradeniya, Peradeniya, Sri Lanka.
[2]Department of Botany, Faculty of Science, University of Peradeniya, Peradeniya, Sri Lanka.
[*]priyangaw@sci.pdn.ac.lk

Conventional drug discovery is a challenging process. Computer-aided drug design (CADD), especially Quantitative Structure-Activity Relationship (QSAR), offers an efficient alternative by correlating chemical structure to biological activity using statistical and computational models. Lately, QSAR has been integrated with KNIME, making CADD straightforward. The Vascular Endothelial Growth Factor Receptor-2 (VEGFR-2) binds to the Vascular Endothelial Growth Factor (VEGF) and promotes angiogenesis, which is crucial for tumour progression. Thus, VEGFR-2 inhibitors suppress angiogenesis, interrupting cancer progression. This study incorporated a modified-TeachOpenCADD workflow of KNIME to identify phytochemicals with VEGFR-2 inhibitory properties. The data of 343 phytochemicals, 12 VEGFR-2 inhibitors, and 10 non-cancer drugs were acquired from the ChEMBL database, and MACCS fingerprints were generated using the RDKit toolkit. The compounds were classified based on activity using Random Forest (RF), Artificial Neural Network (ANN) and Support Vector Machine (SVM), each with 10-fold cross-validations. The predicted phytochemicals' drug likeliness was assessed using Lipinski's rule of five. The ANN, RF and SVM performed with accuracies of 96.36%, 96.64% and 97.20%, respectively, which aligns with existing CADD studies. The root mean square error (RMSE) of RF, ANN and SVM were, respectively, 0.183, 0.191 and 0.167. The multilayer feedforward ANN with one hidden layer (each with 10 neurons) identified three active phytochemicals. With the split criterion as information gain ratio, RF only returned known VEGFR-2 inhibitors. Similarly, SVM with Radial Basis Function (RBF) in the kernel only predicted VEGFR-2 inhibitors. This is because ANN, compared to RF and SVM, excels in decision-making based on training data. All three predicted phytochemicals comply with Lipinski's rule and could be potential VEGFR-2 inhibitors. Notably, two predicted phytochemicals are found in *Butea monosperma* and *Saraca indica* which are native to Sri Lanka. However, subsequent *in vitro* and *in vivo* studies are required before these can be established as anticancer agents.

Keywords: Drug discovery, KNIME, Phytochemicals, QSAR, VEGFR-2 inhibitors