COMPUTATIONAL TOOL FOR PREDICTING DRUG-INDUCED LIVER TOXICITY (DILT)

Simone Brogi[1], Valentina Citi[1], Vincenzo Calderone[1]
[1]Department of Pharmacy, University of Pisa, Pisa - Italy

Introduction: Drug-induced liver toxicity (DILT) is a common cause of liver injury and it represents a critical issue in drug development, causing failures of drugs in clinical trials and the withdrawal of approved drugs from the market. For these reasons assessing in advance the DILT risk of drug candidates is considered an effective strategy to decrease the rate of attrition in drug discovery. We present herein the development and experimental validation of a novel and comprehensive in silico approach for evaluating the intrinsic DILT using a quantitative structure-toxicity relationship (QSTR) method.

Materials and methods: A large set of compounds (>500) was used for developing and validating the computational model. Compounds were selected considering their hepatotoxicity in HepG2cells spanning five orders of magnitude. Phase software was used for deriving the QSTR model, employing appropriate descriptors along with toxicity features as alignment rule. We set a IC_{50} threshold for the selection of hepatotoxic and non-hepatotoxic compounds ($IC_{50} \geq 50\ \mu M$ non-hepatotoxic, $5 \leq IC_{50} \leq 50\ \mu M$ moderate hepatotoxic, $IC_{50} \leq 5 \mu M$ potent hepatotoxic). Atom-based QSTR models were generated (compounds were randomly divided: 60% in the training and 40% in the test set) with a grid spacing of 0.5Å. QSTR models containing 1to 7Partial Least Squares (PLS) factors were produced and cross-validated by leave-n-out (LnO) technique. Extensive validation was performed using an external test set (1/5of selected molecules) not used for generating and cross-validating the model. A further in silico validation, by applying a decoys set, evaluating the Güner-Henry score (GH) and the Enrichment Factor (EF), and by using the Receiver Operating Characteristic (ROC) curve analysis has been performed. Finally, a virtual screening (VS), using over 1,000,000 of molecules from ZINC database, was performed selecting 20 top-ranked compounds that were tested for their hepatotoxicity in HepG2cells.

Results: The model containing 7PLS factors showed high correlation and cross-validated correlation coefficients ($r^2$and Q^2, respectively) and high Pearson R-value (R-Pearson>0.9), suggesting a close correspondence between estimated and experimental IC_{50} values, indicating a strong predictive power and significance of the model. The in silico validation using an external test set provided further indication that the predictivity of the model was not accidental (r^2ext_ts>0.7). The validation using a decoys set allowed to obtain an EF>40 (40 times more probable to find active molecules from chemical databases with respect to the chance), and a GH_score>0.6, demonstrating a great consistency of the model. The output of ROC curve provided an Area Under the Curve (AUC)_score>0.9, indicating that the QSTR model can be used to select molecules with reduced hepatotoxicity. Moreover, a prospective experimental model validation was performed by a VS procedure. The retrieved potential hepatotoxic compounds were tested in HepG2cells (common model for assessing the hepatotoxicity). Gratifyingly, the in vitro tests confirmed the validity of the chosen approach.

Discussion and conclusion: DILT is one of the major causes for drug attrition, causing discontinuation of preclinical and clinical studies, and withdrawals of marketed drugs. Computational models based on chemical structures provide an advantage in saving time and in reducing costs of the research for non-hepatotoxic leads, and they can help in prioritizing them in lab tests, preclinical and clinical studies. To this end, a high-quality QSTR model for DILT was generated, combining appropriate molecular descriptors with the experimental hepatotoxicity, and validated by in silico and in vitro procedures, demonstrating the validity of the computational approach. This effort allowed us to obtain an in-house DILT predictive tool useful for selecting compounds with reduced hepatotoxicity at the early steps of the drug discovery trajectory.