

SuperCYPsPred—a web server for the prediction of cytochrome activity

Priyanka Banerjee¹*, Mathias Dunkel, Emanuel Kemmler and Robert Preissner

Structural Bioinformatics Group, Institute for Physiology & ECRC, Charité, University Medicine Berlin, 10115 Berlin, Germany

Received February 13, 2020; Revised February 26, 2020; Editorial Decision March 02, 2020; Accepted March 05, 2020

ABSTRACT

Cytochrome P450 enzymes (CYPs)-mediated drug metabolism influences drug pharmacokinetics and results in adverse outcomes in patients through drug–drug interactions (DDIs). Absorption, distribution, metabolism, excretion and toxicity (ADMET) issues are the leading causes for the failure of a drug in the clinical trials. As details on their metabolism are known for just half of the approved drugs, a tool for reliable prediction of CYPs specificity is needed. The SuperCYPsPred web server is currently focused on five major CYPs isoenzymes, which includes CYP1A2, CYP2C19, CYP2D6, CYP2C9 and CYP3A4 that are responsible for more than 80% of the metabolism of clinical drugs. The prediction models for classification of the CYPs inhibition are based on well-established machine learning methods. The models were validated both on cross-validation and external validation sets and achieved good performance. The web server takes a 2D chemical structure as input and reports the CYP inhibition profile of the chemical for 10 models using different molecular fingerprints, along with confidence scores, similar compounds, known CYPs information of drugs—published in literature, detailed interaction profile of individual cytochromes including a DDIs table and an overall CYPs prediction radar chart (<http://insilico-cyp.charite.de/SuperCYPsPred/>). The web server does not require log in or registration and is free to use.

INTRODUCTION

A successful drug candidate is often characterized by sticking an appropriate balance of potency, efficacy, safety and favourable pharmacokinetics (1). In other words, the subtle distinction between drug efficacy and toxicity are controlled by the interplay of pharmacokinetic, pharmacody-

namic and genetic factors. Drug metabolism represents an integral contributor to the many physiological processes that govern the pharmacokinetic fate of most therapeutic agents. In rudimentary terms, drug metabolism can be described as the biological transformation of lipophilic, nonpolar molecules to more hydrophilic, polar molecules known as metabolites, which are in turn readily eliminated from the body (2). Based on the chemical nature of biotransformation, the process of drug metabolism reactions can be divided into two broad categories: phase I (oxidative reactions) and phase II (conjugative reactions) (3). The human cytochrome P450 family (phase I enzymes) contains 57 isozymes and these isozymes metabolize approximately two-thirds of known drugs in human with 80% of this attribute to five isozymes—1A2, 3A4, 2C9, 2C19 and 2D6 (3). Most of these CYPs responsible for phase I reactions are concentrated in the liver (3). Drug safety is a major challenge in bringing new drugs to market—especially with a majority of polytreated patients (4). As details on their metabolism are known for just half of the approved drugs, a tool for reliable prediction of CYPs specificity is needed. Unexpected toxicities due to drug–drug interactions (DDIs) are a major source of adverse effects concerning the post-marketing safety of the drugs, which results in unexpected morbidity and mortality (3). The evaluation of the interactions of CYPs with small molecules (such as drugs) constitutes a fundamental step for drug design, as well as for toxicity assessment. The DDIs is one of the major risks associated with the pre-clinical screening of drugs, undesired DDIs can result in enhancement of drugs toxicity or reduction of therapeutic efficacy (3). Most DDIs are known to be mediated via CYPs, such interactions may indeed increase for genetic variants (with frequent CYPs polymorphisms ‘poor metabolizers’ or inhibitors) or increased metabolism (with ‘rapid metabolizers’ or CYPs inducers). Food and Drug Administration (FDA) guidance documents include pre-screening of drugs for CYPs inhibition and induction (5). CYPs–drug interactions are responsible for several unwanted adverse effects resulting in alteration of drug pharmacokinetics, such as changes in plasma concentration–time profiles of drugs

*To whom correspondence should be addressed. Tel: +49 30 450 528 505; Fax: +49 30 450 540 955; Email: priyanka.banerjee@charite.de
Present address: Priyanka Banerjee, Institute of Physiology, Charité–University Medicine Berlin, Philippstrasse 12, Berlin 10115, Germany.

when given concomitantly (5). Polymedication in elderly patients is a well-established source of variability in drug response. Adverse drug reactions (ADRs) effects induced by DDIs may result in early termination of drug development or even withdrawal of drugs from the market (such as in the case of mibefradil and cerivastatin) (3). As estimated by a US meta-analysis—the incidence of severe side effects is 6.7% due to ADRs and results in 100 000 deaths yearly (4). Therefore, to avoid undesirable DDIs, a screening test for new drug candidates' is regularly performed by pharmaceutical companies. Traditionally, *in vivo* and *in vitro* tests are performed to investigate drug safety, and adverse drug effects. Currently, experimental cellular assays on human hepatocytes are being used to evaluate the CYP metabolism of the drug candidates and its possible risk (5). However, there are few challenges associated with the *in vitro* testing such as it is time-consuming, lacks the ability to provide information on the structure-activity relationship data. In contrast, computational approaches are faster for evaluation of large number of compounds and can be applied at the early phase of drug design owing to their low cost (6). *In silico* approaches have the strength to predict the activities of hypothetical compounds, which are yet to be synthesized. In the past several years, a variety of computational approaches have been used to develop classification models for the prediction of chemical activities (2,7–8).

Several studies have been reported in the literature for the screening of CYP450 inhibition based on quantitative structure-activity relationship (9), molecular dynamics simulation and recently machine learning (ML) methods (9,10). Several of them vNN-ADMET (11), AdmetSAR (12), WhichCyp (13) are designed for the prediction of cytochrome inhibition.

We present SuperCYPsPred web server which includes ML models based on the RF algorithm (14), and different types of data sampling methods (7). The models presented here discriminate the inhibitors from the non-inhibitors for five major CYP450 isoforms. The statistical significance of the predictive models was assessed by the 10-fold fragment-based CLUSTER cross-validation for the training sets, models were evaluated on different criteria like prediction accuracy, sensitivity, specificity, area under the curve (AUC) of a receiver operating characteristic (ROC) and F1 measure. The applicability domain of the models was assessed using fragment-based and structural similarity-based approaches. Apart from predicting a particular compound as active (inhibitor) or inactive (non-inhibitor) for a defined CYPs isoform, the web server also shows most similar compounds in the training set, including if there are known cytochrome interactions available for the input compound in the literature. Additionally, each model prediction performance was validated by the external test set. The SuperCYPsPred web server enables user to check DDIs to understand the likelihood of DDIs when given a combination of medications.

MATERIALS AND METHODS

Software implementation

SuperCYPsPred data is stored in a relational MySQL database. RDKit package (<http://www.rdkit.org>) is used

for handling the chemical information in the database. ChemDoodle Web components (<https://web.chemdoodle.com/>), an open source, JavaScript library for chemistry interface was used in the server. The website back-end is built using PHP and Python; web access is enabled via the Apache HTTP Server. Redis is employed for queuing and assessing the API requests (8). The server has been tested on the recent version of Mozilla Firefox, Google Chrome and Apple Safari.

Input and output

The user interface of the SuperCYPsPred is easy-to-use and self-informative. The web server offers the user four ways to submit small molecules. The user can upload a standard molecule file, name of the molecule, draw a molecule or enter SMILES (Simplified Molecular-Input Line-Entry System) string of the compound. Optionally, the user may select different models or all models for prediction, including different molecular fingerprints. The results are displayed in tabular format and radar plot in the browser and includes the molecular structure with the physicochemical properties and the three most similar molecules from the training set that gave the most important contribution to the decision of the CYPs prediction models. The user can access the result in the results section, in case the prediction results cannot be shown immediately. These prediction results are also displayed as a radar plot comparing the average confidence score of the active compounds in the training set of each model, to that of the input compound (see Figure 1). Via the DDIs checker available on the web server user can understand the likelihood of DDIs when given a combination of medications as well get information on alternative drugs that can be used with the same therapeutic effects.

Datasets

The training set for the models was collected from literature and two different databases. A total of 17 143 substances tested using quantitative high-throughput screening with *in vitro* bioluminescent assay against five major isoforms of CYPs was obtained from the PubChem Bioassay database, AID: 1815 (15). Much of the information on the cytochrome P450 enzymes were extracted from the literature using our in-house data mining platform and stored in the SuperCYP database (16). This database currently contains information on 1170 drugs with more than 3800 known interactions. Inorganic compounds, salts and mixtures as well as entries classified as inconclusive were removed from the final dataset. For each of the five cytochromes, compounds were divided into training and external validation set, keeping the ratio of the actives (inhibitors) and inactives (non-inhibitors) constant (See Supplementary Data S1).

Evaluation of the SuperCYPsPred models

Each model was validated using 10-fold CLUSTER cross-validation. The data was divided using different sampling methods and keeping the ratio of active and inactive instances in all the folds constant as published in our previous

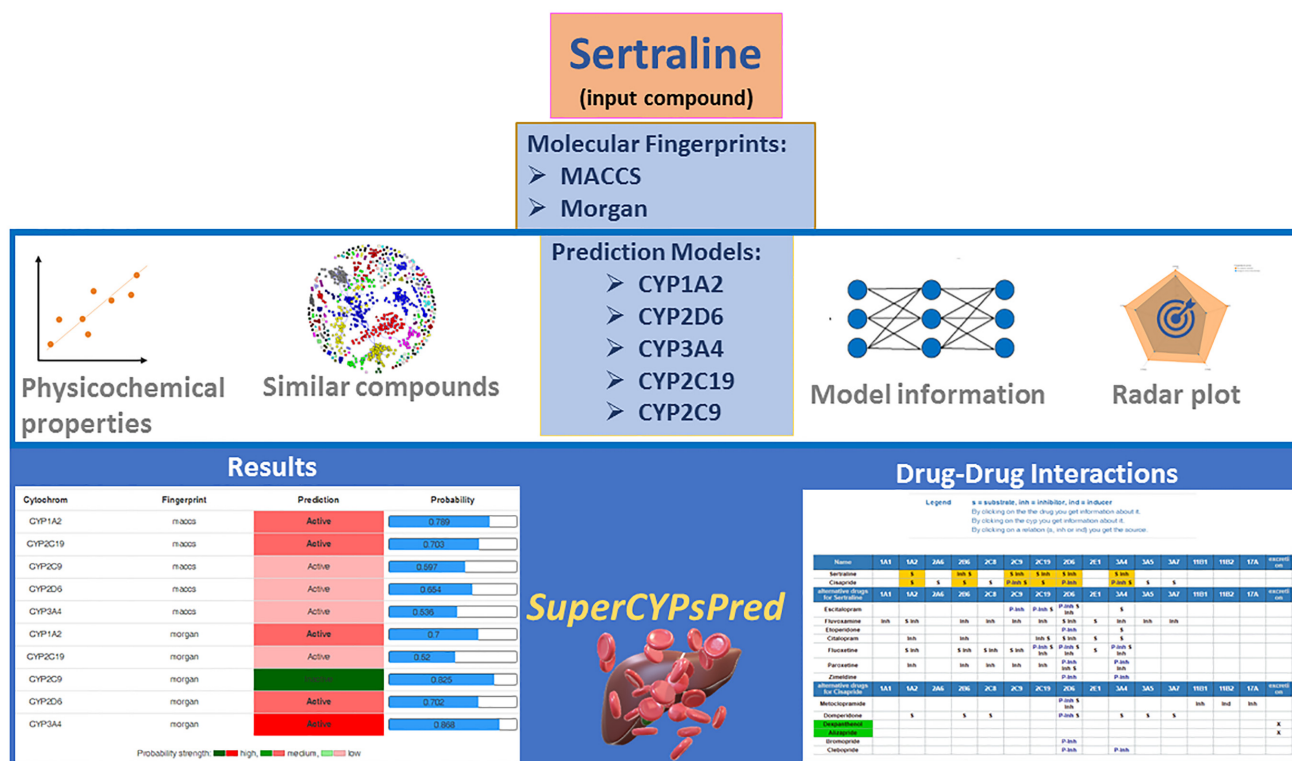


Figure 1. Illustration of an example compound (sertraline) used as an application case. Sertraline is the input compound; the user can choose either MACCS or Morgan fingerprints or both for the prediction. In this case both were selected. The results displayed shows the cytochrome inhibition profile for the five major isoforms. The result page also includes information on similar compounds, with known CYPs reported in literature and overall radar plot. The DDI matrix shows the cytochrome–drug interactions of sertraline, when given with combination with drugs (such as cisapride or ibuprofen) can result in major interactions with side effects (please check the DDI example on the server for more details).

work (7). Additionally, an external set was used for the evaluation of the predictive capacity of each model. Each model was evaluated by the following parameters:

- **Prediction Accuracy** is defined as the ability of a model to differentiate actives and inactives cases correctly.
- **Sensitivity** is the model's ability to correctly identify the positive instances.
- **Specificity** is defined as the model's ability to correctly identify the negative instances.
- **The AUC of a ROC curve** plots the true positive rate against the false positive rate at different thresholds. The AUC-ROC has been used as an effective measure for binary classifiers (17).
- **F1 measure** is a measure of a test's accuracy and is defined as the weighted harmonic mean of the precision and recall of the test.

All cross-validations and external set validation results for the best performing CYPs isoform models are summarized with respect to the different performance measures in Table 1. Most for the models achieved a prediction accuracy of 90% and higher, except for 2D6 for cross-validation and 2D6, 3A4 for external validation. The specificity and sensitivity of all the models are balanced, and scored above 85% and above, exception 2D6, 2C9 on external sets. The

least ROC-AUC value of the models is 85% and the highest is 99% (Table 1).

Prediction models

The SuperCYPsPred prediction models were developed by ML approaches. These models enable a data-driven approach to cytochrome activity predictions that can identify chemical patterns that otherwise would be overlooked. These models are based on traditional ML-like Random Forest (RF) which can produce interpretable models with low complexity (14). Ensemble learning algorithms such as RF are less susceptible to class bias and overfitting. The optimal parameters of the RF models were determined using the grid search method, which is implemented by the Scikit-learn package (version 0.20) in Python (version 3.6.6), and the 10-fold cross-validation was used for the model optimization. Two different chemical based fingerprints are used: MACCS molecular fingerprints-166 bits and Morgan circular fingerprints-2048 bits (<http://www.rdkit.org/>). These two fingerprints have shown an optimal performance for the prediction of chemical activity for several prediction methods (7,18). A detailed information on the construction of the models and evaluation can be found at the web server FAQ section as well as in the published work (7). Data curation and standardization of the SMILES strings in the dataset were processed by KNIME (19). More details on

Table 1. Statistics for the models applied to fragment-based CLUSTER cross-validation and external validation sets

Cytochromes isoforms		1A2	2C9	2C19	2D6	3A4
Data sampling method		SMOTETC	SMOTETC	RandOS	kMedoids1	AugRandUS
Chemical fingerprints		Morgan	Morgan	MACCS	MACCS	Morgan
Cross-validation	Prediction	0.95	0.97	0.97	0.84	0.92
	Accuracy					
	Sensitivity	0.97	0.97	0.99	0.88	0.93
	Specificity	0.97	0.96	0.94	0.86	0.92
	ROC-AUC	0.99	0.98	0.97	0.92	0.96
External validation	F-measure	0.93	0.94	0.99	0.85	0.92
	Prediction	0.90	0.90	0.95	0.80	0.86
	Accuracy					
	Sensitivity	0.84	0.61	0.95	0.80	0.82
	Specificity	0.92	0.94	0.95	0.79	0.87
	ROC-AUC	0.95	0.97	0.87	0.85	0.93
	F-measure	0.76	0.58	0.98	0.60	0.74

individual models and the feature responsible for class predictions are provided in the 'Model Information' section of the web server.

Model CYP3A4. The most abundant hepatic CYP human isoform is CYP3A4, and the metabolism of almost 50% clinically approved drugs is mediated by this isoform (5). The undesired CYP3A4 inhibition by co-administered drugs can result in clinically adverse DDIs. The CYP3A4 model is based on RF algorithm and Augmented Random Under Sampling (AugRandUS) data sampling method (7). This model has a prediction accuracy of 92% on cross-validation and 86% on external validation. The AUC-ROC values of cross-validation and external validation are 0.96 and 0.93, respectively (Table 1).

Model CYP1A2. CYP1A2 is an important metabolizing enzyme in the human liver, which accounts for 13% of the total hepatic CYPs and is responsible for the metabolism of 4% marketed drugs (5). The CYP1A2 model is based on RF algorithm and Synthetic Minority Over-Sampling Technique using Tanimoto Coefficient (SMOTETC) data sampling method (7). The CYP1A2 model has achieved prediction accuracy of 95% on cross-validation and 90% on external validation. The AUC-ROC values of cross-validation and external validation are 0.99 and 0.95, respectively (Table 1).

Model CYP2C9. Approximately 20% of the total hepatic CYPs include CYP2C9 isoform. This isoform is involved in many clinically relevant drug interactions and is responsible for the metabolism of 15% of the clinically approved drugs (5). The CYP2C9 model is based on RF algorithm and SMOTETC data sampling method (7). The CYP2C9 model has achieved prediction accuracy of 97% on cross-validation and 90% on external validation. The AUC-ROC values of cross-validation and external validation are 0.98 and 0.97, respectively (Table 1).

Model CYP2C19. CYP2C19 contributes to about 16% of total hepatic content. This enzyme genetic polymorphism may affect several clinically important drugs with narrow therapeutic index (5). The CYP2C19 model is based on the RF algorithm and Random Over Sampling (RandOS) data sampling method (7). The CYP2C19 model has achieved

prediction accuracy of 97% on cross-validation and 95% on external validation. The AUC-ROC values of cross-validation and external validation are 0.97 and 0.87, respectively (Table 1).

Model CYP2D6. CYP2D6 only accounts for 2–4% of the total hepatic CYPs; however it is responsible for ~30% of all marketed drugs (5). The CYP2D6 model is based on the RF algorithm and *K*-Medoids Under Sampling (*k*Medoids1) data sampling method (7). The CYP2D6 model has achieved prediction accuracy of 84% on cross-validation and 80% on external validation. The AUC-ROC values of cross-validation and external validation are 0.92 and 0.85, respectively (Table 1).

Drug–Drug Interactions (DDIs)

DDIs are a major concern in the clinical field owing to the aging population and increase of prescription of multiple medications. Numerous adverse drug interactions have been identified and reported in literature and electronic health records (ECRs) (4). If two or multiple drugs are prescribed at the same time, DDIs have shown to delay, decrease or enhance the absorption of one or more of the drugs present in the combination. This can cause severe side effects either by decreasing or increasing the action of the drugs (such as in the case of fluoxetine and phenelzine, digoxin and quinidine, sildenafil and isosorbide mononitrate) (20). Therefore, the initial screening of possible drug–CYP inhibitors interaction is required to highlight the potential clinically relevant complications associated with such interactions. On the other hand, drugs can be intentionally combined to take advantage of CYP450 inhibition. Such in the case of the drug ritonavir, a protease inhibitor and potent CYP3A4 inhibitor. This drug is added to lopinavir, another protease inhibitor, to boost serum levels in patients with human immunodeficiency virus (3). An updated version of the SuperCYP database was used to predict the likelihood of DDIs when given a combination of medications (16). The DDIs matrix in the SuperCYPsPred web server is based on manually curated dataset from literature (known interactions with reference) as well as prediction based on computational models (unknown interactions). Additionally, the interpretation of results from different models can be connected to make better elucidation and

understanding of the mechanisms leading to DDIs. With the help of DDIs checker assessable via the SuperCYPsPred web server, user can check if the metabolisms of the multiple drug combinations (also known as drug cocktail) interact with each other. User can provide a list of drugs name such as enalapril (angiotensin converting enzyme (ACE)-inhibitor) and tasosartan (angiotensin II (AT2)-receptor blocker) and submit the names through 'get interactions' button. The resulting table will then contain information on the metabolic enzymes for each drug and the mode of activity (substrate, inhibitor or inducer). In this case both are substrates of CYP3A4; the table additionally provides alternatives drug with the same therapeutic effect that do not interact with each other (e.g. enalapril and candesartan (CYP2C9) (16).

Application case

As an application case, Sertraline is been chosen in order to show the functionality of the web server and to discuss the results in detail. Sertraline is an antidepressant belongs to the class of drugs known as selective reuptake inhibitors (SSRIs). Sertraline is widely used to treat depression, panic, anxiety or obsessive-compulsive symptoms. Sertraline interactions with drugs like aspirin, cimetidine, ibuprofen and Monoamine Oxidase (MAO) inhibitors can change the course of action of sertraline and may cause a serious (sometimes fatal) drug interaction (20). In the Figure 1, sertraline is used as the input compound. The user has the possibility to choose the different models or all models, including the choice of molecular fingerprints. The resulting page will then show the physicochemical properties and similar compounds from the training set, along with the resulting table which contains the CYPs isoforms and prediction confidence respectively. The table in this case shows sertraline is active for five major isoforms with variable (weak to strong) confidence scores. Additionally, the radar plot enables the user to understand the prediction strength of the input compound as compared to average prediction strength achieved by the respective model on the training set compounds. The DDI results shows drugs (sertraline, cisapride) taken in combination can result in major drug interaction. Sertraline is an inhibitor/substrate for five major CYPs (shown in yellow). On the other hand, cisapride, a gastrokinetic agent also shown to interact with five major CYPs isoforms (shown in yellow). This drug combination is avoided by the medical practitioners as this can increase the risk of an irregular heart rhythm and may result in potential life-threatening condition. An alternative drug as a replacement for cisapride (shown in green) is highlighted. The alternative drug is suggested when the drug is not known or predicted to bind with any CYPs isoforms. User can click on the drug or cytochrome name to get detail information. For each cytochrome, synonyms, Uniprot ID as well as known drug interactions along with the type of interaction and literature source are provided. For the drugs, CID, known CYPs interaction with the type of interactions including the Phase2 interactions profile and half-life are shown.

CONCLUSION AND FUTURE UPDATES

The adverse effect of DDIs contributes largely to drug toxicity. Though it is often interpreted the toxicity of drug metabolites can be only be determined empirically; however, learning from previously reported ADRs data can demand intensive drug testing for potentially toxic drugs (21). Here, we present SuperCYPsPred web server which implements the state-of-art ML methods to build predictive models for five major cytochromes involved in the metabolism of most clinically available drugs. The computational models are focussed on the first step of the safety assessment. That is, if a particular compound is active (inhibitor) or inactive (non-inhibitor) for a defined CYPs isoform. The outcome of these predictions helps us to compute probability of a compound to be highly active or slightly active for the CYPs, rather than its exact activity value. When compared with other standard published predictive models for CYP inhibition (11,12), all the models of SuperCYPsPred performed from the range of comparatively good to better in some cases. However, a fair comparison using performance measures like accuracy, sensitivity, specificity and AUC-ROC has been provided as Supplementary Data S2 and S3.

One of the major challenges of computing any prediction models is the availability of quality data required for such predictions. Currently, SuperCYPsPred web server only contains models for CYPs inhibition. In future, when sufficient standard data are obtained, models for substrate and inducer prediction will be made available via the web server updates.

We hope that the understanding of the DDIs enabled via SuperCYPsPred web server will help to approximately adjust or reinvent research and development strategies. This is important in order to overcome the attrition during the clinical trial phases of drug discovery and post-market withdrawal of approved drugs. The application of SuperCYPsPred web server will include the identification of useful CYPs inhibitors as well as assessment of new drug candidates for its clinically relevant DDIs potential.

As an evolutionary step, SuperCYPsPred will focus on method development to foster better characterization of clinically relevant adverse effects associated with undesired DDIs, considering genetic polymorphisms of individual CYPs. Furthermore, to maintain the high standard of the SuperCYPsPred web server, regular updates will be executed, including addition of new models for the prediction of substrates and inducers.

SUPPLEMENTARY DATA

[Supplementary Data](#) are available at NAR Online.

ACKNOWLEDGEMENTS

We thank the students of the Structural Bioinformatics Group at Charité for testing the SuperCYPsPred web server. Furthermore, the authors wish to thank Andreas-Oliver Eckert for his input regarding the SuperCYPsPred web server API script.

Author contributions: *Project idea:* PB and RP. PB designed the workflow and curated the dataset. PB developed the models, validated and implemented the models. *Implementation of web server:* PB, EK and MD. *Writing of manuscript:* PB. *Project coordination:* PB. All authors proofread the manuscript.

FUNDING

Berlin-Brandenburg Research Platform BB3R; Federal Ministry of Education and Research (BMBF), Germany [031A262C]; DKTK. Funding for open access charge: Charité—University Medicine Berlin.

Conflict of interest statement. None declared.

REFERENCES

- Walker, D.K. (2004) The use of pharmacokinetic and pharmacodynamic data in the assessment of drug safety in early drug development. *Br. J. Clin. Pharmacol.*, **58**, 601–608.
- Corsini, A. and Bortolini, M. (2013) Drug-induced liver injury: the role of drug metabolism and transport. *J. Clin. Pharmacol.*, **53**, 463–474.
- Lynch, T.O.M., Price, A.M.Y. and Virginia, E. (2007) The effect of cytochrome P450 metabolism on drug response, interactions, and adverse effects. *Am. Fam. Physician*, **76**, 2–6.
- Lazarou, J., Pomeranz, B.H. and Corey, P.N. (1998) Incidence of adverse drug reactions in hospitalized patients a meta-analysis of prospective studies. *JAMA*, **279**, 1200–1215.
- Kato, H. (2019) Computational prediction of cytochrome P450 inhibition and induction. *Drug Metab. Pharmacokinet.*, **35**, 30–44.
- Basile, A.O., Yahi, A. and Tatonetti, N.P. (2019) Artificial intelligence for drug toxicity and safety. *Trends Pharmacol. Sci.*, **40**, 624–635.
- Banerjee, P., Dehnbostel, F.O. and Preissner, R. (2018) Prediction is a balancing Act: importance of sampling methods to balance sensitivity and specificity of predictive models based on imbalanced chemical data sets. *Front. Chem.*, **6**, 1–11.
- Banerjee, P., Eckert, A.O., Schrey, A.K. and Preissner, R. (2018) ProTox-II: a webserver for the prediction of toxicity of chemicals. *Nucleic Acids Res.*, **46**, 257–263.
- Guo, Y., Lucksiri, A., Dickinson, G.L., Vuppalachchi, R.K., Hilligoss, J.K. and Hall, S.D. (2020) Quantitative prediction of CYP3A4- and CYP3A5-Mediated drug interactions. *Clin. Pharmacol. Ther.*, **107**, 246–256.
- Lee, J.H., Basith, S., Cui, M., Kim, B. and Choi, S. (2017) In silico prediction of multiple-category classification model for cytochrome P450 inhibitors and non-inhibitors using machine-learning method. *SAR QSAR Environ. Res.*, **28**, 863–874.
- Schyman, P., Liu, R., Desai, V. and Wallqvist, A. (2017) vNN web server for ADMET predictions. *Front. Pharmacol.*, **8**, 1–14.
- Yang, H., Lou, C., Sun, L., Li, J., Cai, Y., Wang, Z., Li, W., Liu, G. and Tang, Y. (2019) admetSAR 2.0: web-service for prediction and optimization of chemical ADMET properties. *Bioinformatics*, **35**, 1067–1069.
- Rostkowski, M., Spjuth, O. and Rydberg, P. (2013) WhichCyp: prediction of cytochromes P450 inhibition. *Bioinformatics*, **29**, 2051–2052.
- Banerjee, P., Siramshetty, V.B., Drwal, M.N. and Preissner, R. (2016) Computational methods for prediction of in vitro effects of new chemical structures. *J. Cheminform.*, **8**, 1–11.
- Veith, H., Southall, N., Huang, R., James, T., Fayne, D., Artemenko, N., Shen, M., Inglese, J., Austin, C.P., Lloyd, D.G. *et al.* (2009) Comprehensive characterization of cytochrome P450 isozyme selectivity across chemical libraries. *Nat. Biotechnol.*, **27**, 1050–1055.
- Preissner, S., Kroll, K., Dunkel, M., Senger, C., Goldsobel, G., Kuzman, D., Guenther, S., Winnenburg, R., Schroeder, M. and Preissner, R. (2010) SuperCYP: a comprehensive database on Cytochrome P450 enzymes including a tool for analysis of CYP-drug interactions. *Nucleic Acids Res.*, **38**, D237–D243.
- van Erkel, A.R. and Pattynama, P.M. (1998) Receiver operating characteristic (ROC) analysis: basic principles and applications in radiology. *Eur. J. Radiol.*, **27**, 88–94.
- Li, X., Xu, Y., Lai, L. and Pei, J. (2018) Prediction of human cytochrome P450 inhibition using a multitask deep autoencoder neural network. *Mol. Pharm.*, **15**, 4336–4345.
- Fillbrunn, A., Dietz, C., Pfeuffer, J., Rahn, R., Landrum, G.A. and Berthold, M.R. (2017) KNIME for reproducible cross-domain analysis of life science data. *J. Biotechnol.*, **261**, 149–156.
- Muijsers, R.B.R., Plosker, G.L., Noble, S., Kurzthaler, I. and Preskorn, S.H. (2002) Sertraline a review of its use in the management of major depressive disorder in elderly patients. *Drugs Aging*, **19**, 377–392.
- Hua, L., Chiang, C.-W., Cong, W., Li, J., Wang, X., Cheng, L., Feng, W., Quinney, S.K., Wang, L. and Li, L. (2019) The cancer drug fraction of metabolism database. *CPT: Pharmacometrics Syst. Pharmacol.*, **8**, 511–519.