Analysis of Molecular Properties of Drugs Interacting with SLC22 Transporters OAT1, OAT3, OCT1, and OCT2: A Machine-Learning Approach

Henry C. Liu, Anne Goldenberg, Yuchen Chen, Christina Lun, Wei Wu, Kevin T. Bush, Natasha Balac, Paul Rodriguez, Ruben Abagyan, Sanjay K. Nigam

From the Departments of Bioengineering (H.C.L.), Pediatrics (A.G., Y.C., C.L., K.T.B., S.K.N.), Medicine (W.W., S.K.N.), Cellular and Molecular Medicine (S.K.N.), Pharmacology (R.A.), and the San Diego Supercomputer Center (N.B., P.R.), University of California San Diego, La Jolla, CA, 92093, USA .

JPET Fast Forward. Published on August 3, 2016 as DOI: 10.1124/jpet.116.232660
This article has not been copyedited and formatted. The final version may differ from this version.

Running Title Page

Running title: *Analysis of ligand selectivity of SLC22 transporters*

To whom correspondence should be addressed: Sanjay K. Nigam, Department of Pediatrics, University of California San Diego, 9500 Gilman Drive, MC0693, La Jolla, CA 92093. USA. Phone: (858) 822-3482; Fax: (858) 822-3483; E-mail: snigam@ucsd.edu

The number of text pages: 39

Number of tables: 4

Number of figures: 8

Number of references: 44

Number of words in the Abstract: 240

Number of words in the Introduction: 1079

Number of words in the Discussion: 1304

Nonstandard abbreviations: APF, atomic property field; DDI, drug-drug interaction; DMI, drug-metabolite interaction; KNIME, Konstanz information miner; OAT, organic anion transporter; OCT, organic cation transporter; ROC, receiver operating characteristic; SLC, solute carrier; WEKA, Waikato environment for knowledge analysis.

Recommended Section Assignment: Metabolism, Transport, and Pharmacogenomics

JPET Fast Forward. Published on August 3, 2016 as DOI: 10.1124/jpet.116.232660
This article has not been copyedited and formatted. The final version may differ from this version.

ABSTRACT

Statistical analysis was performed on physicochemical descriptors of ~250 drugs known to interact with one or more SLC22 "drug" transporter (i.e., SLC22A6 or OAT1, SLC22A8 or OAT3, SLC22A1 or OCT1, and SLC22A2 or OCT2), followed by application of machine-learning methods and wet-lab testing of novel predictions. In addition to molecular charge, organic anion transporters (OATs) were found to interact with planar structures, whereas organic cation transporters (OCTs) interact with more three-dimensional structures (i.e., greater SP3 character). Moreover, compared to OAT1 ligands, OAT3 ligands possess more acyclic tetravalent bonds and have a more zwitterionic/cationic character. Multiple pharmacophore models based on the drugs were generated and, consistent with the machine-learning analyses, one unique pharmacophore created from ligands of OAT3 possessed cationic properties similar to OCT ligands; this was confirmed by quantitative atomic property field analysis (APF). Virtual screening with this pharmacophore, followed by transport assays, identified several cationic drugs that selectively interact with OAT3 but not OAT1. Although this analysis may be somewhat limited by the need to rely largely on inhibition data for modeling, wet-lab/in vitro transport studies, as well as analysis of drug/metabolite handling in *Oat* and *Oct* knockout animals support the general validity of the approach—which can also be applied to other SLC and ABC drug transporters. This may make it possible to predict the molecular properties of a drug or metabolite necessary for interaction the transporter(s), thereby enabling better predicting of drug-drug interactions (DDI) and drug-metabolite interactions (DMI).

JPET Fast Forward. Published on August 3, 2016 as DOI: 10.1124/jpet.116.232660
This article has not been copyedited and formatted. The final version may differ from this version.

INTRODUCTION

Organic anion transporter 1 (OAT1/SLC22A6), OAT3 (SLC22A8), organic cation transporter 1 (OCT1/SLC22A1), and OCT2 (SLC22A2), perhaps the best studied members of the SLC22 family of solute carriers, are responsible for the excretion of a wide variety of drugs, toxins, and metabolites in the kidney, liver and other tissues (Emami Riedmaier et al., 2012; Koepsell, 2013; Nigam, 2015; Nigam et al., 2015a; Nigam et al., 2015b). This family, originally proposed in 1997 based on three family members (Lopez-Nieto et al., 1997), now consists of over 30 members in mammals (Lopez-Nieto et al., 1997; Eraly et al., 2004; Wu et al., 2009; Zhu et al., 2015). Although sharing overall sequence and predicted structural similarities, the four transporters have distinct preferences for interaction with ligands. As their names suggest, OATs, belonging to the "organic anion" transporter subfamily, mainly interact with anions whereas OCTs, belonging to the "organic cation" transporter subfamily, mainly interact with cations (Popp et al., 2005). Nevertheless, the grouping of OATs and OCTs into two different transporter subfamilies, organic anions and organic cations, respectively, can be misleading when it comes to individual drugs, toxins and metabolites. For example, OATs have the capacity to interact with cationic drugs (Ahn et al., 2009), and both OATs and OCTs appear to interact in vitro and in vivo with zwitterionic or mildly "cationic" metabolites such as creatinine and polyamines (Ahn et al., 2011; Imamura et al., 2011; Vallon et al., 2012). However, these studies were limited to a few interacting compounds. Moreover, evolutionary analysis also indicates that the SLC22 family is likely more complex than originally thought as it appears to be comprised of at least 6 subgroups, including, apart from the Oat and Oct group, groups termed Oat-like, Oat-related, Octn and Oct-related (Zhu et al., 2015). Together these results raise certain questions about the simple conception of OATs as "organic anion" transporters and OCTs as "organic cation" transporters and demonstrate the need for deeper investigation of ligand interactions with the various SLC22 transporters.

Given that there are a large number of well-established OAT1, OAT3, OCT1 and OCT2 interacting drugs, we attempted to address this issue by performing a systematic computational and statistical analysis, as well as machine-learning analyses, based on the physicochemical descriptors of the

drugs known to interact with one or more of these four transporters. Since the crystal structures of the four transporters are unknown at this time, ligand-based computational chemistry approaches were utilized here. Among these, one commonly used method is the development of quantitative structure-activity relationship (SAR and QSAR) models, which attempt to identify the correlation between the activity, or binding affinity of ligands and transporters, and the values of the physicochemical descriptors of the ligands. Previously developed QSAR models for OAT1, OAT3, and OAT6, which were built based on inputs of approximately 10 descriptors, identified several physicochemical properties of ligands important for interaction with transporters (Kaler et al., 2007; Truong et al., 2008). In addition to QSAR models, another approach that has gained popularity is the application of machine-learning tools. Among these tools, the support vector machine method (SVM) has been used to develop models for two ABC transporters, breast cancer resistance protein (BCRP) and P-glycoprotein (P-gp) (Wang et al., 2011; Hazai et al., 2013); these models were mainly used for the in-silico prediction of new substrates. Besides SVM, other powerful machine-learning tools, such decision trees and random forests, have been used widely for different applications (Vaglio Laurin et al., 2014; Kim et al., 2015).

To investigate the functional differences between OAT1, OAT3, OCT1, and OCT2, a number of machine-learning methods were applied to understand which physicochemical properties within a set of ~250 drugs affect the interactions between individual transporters and their ligands, as well as the relative importance of these properties in contributing to selectivity for interaction with a particular transporter. In order to obtain clear results, the analysis relied heavily on inhibition (Ki) data, since this data is available for almost all of the drugs studied. Actual transport (Km) data is much more limited; in a recent review on modeling of drug transporters, it was pointed out that the limited transport data (as opposed to inhibition data) is a general issue in the field of SLC and ABC drug transporters (Matsson and Bergstrom, 2015). Likewise, while it is generally assumed that inhibition data indicates competitive inhibition of transport of a characteristic substrate (e.g., PAH, TEA), competitive versus non-competitive inhibition is rarely formally evaluated (in fact, the authors emphasize this in their “wish-list of developments needed in the field” (Matsson and Bergstrom, 2015)). Nevertheless, it is worth noting, that, for those drugs for which

transport data was available, it was generally consistent with inhibition data (Supplemental Table 1). Moreover, many of the general classes of compounds (e.g., antivirals, diuretics, antibiotics, metformin, zwitterions) have been studied in the Oat1, Oat3, Oct1 and Oct2 knockout animals or tissues derived from them, and altered handling by the kidney and other tissues has been demonstrated in many of these cases (Eraly et al., 2006; Vanwert et al., 2007; Truong et al., 2008; Vallon et al., 2008b; Vanwert et al., 2008; Nagle et al., 2011; Vallon et al., 2012; Nagle et al., 2013).

The results of the machine-learning were further supported by the generation of pharmacophore models of OAT and OCT ligands. Pharmacophore modeling studies aim to find the common features shared among the ligands in three-dimensional space and previous studies using this method have built several pharmacophore models for relatively small subsets of various ligands of renal transporters, including drugs and metabolites (Ahn et al., 2009; Kouznetsova et al., 2011; Wikoff et al., 2011; Duan et al., 2012). In order to gain a much more comprehensive understanding of binding interactions and ligand selections, it was necessary to construct pharmacophore models based on as many pharmaceutical drugs as possible (in our case, 253). This provides a more comprehensive chemical space to build complete pharmacophore models for the multi-specific transporters because each individual compound contributes some information to the broader representation of the whole chemical space of binding.

Our results indicate that, in addition to charge-related factors, OATs interact with planar structures, whereas OCTs interact with more 3-dimenional structures, indicating that in addition to charge, the topology of ligands is another important factor. In addition, subtle but important differences exist between OAT1 and OAT3; OAT3 has a propensity to bind some cations that structurally overlap with OCT ligands. This was experimentally confirmed by wet lab transport assays of ligands predicted by virtual screening.

JPET Fast Forward. Published on August 3, 2016 as DOI: 10.1124/jpet.116.232660
This article has not been copyedited and formatted. The final version may differ from this version.

MATERIALS AND METHODS

The overall computational workflow is shown in Supplemental Figure 1.

Materials—Water-soluble probenecid was purchased from Molecular Probes (Eugene, OR). The fluorescent tracers, 5-carboxyfluorescein (5CF) and 6-carboxyfluorescein (6CF), and cationic drugs (loperamide hydrochloride, nebivolol hydrochloride, darifenacin hydrobromide, paliperidone, cisapride monohydrate, and halofantrine hydrochloride) were purchased from Sigma-Aldrich (St. Louis, MO).

Selection and classification of drugs interacting with OAT1, OAT3, OCT1, and/or OCT2—A comprehensive literature and internet search was performed in order to compile a list of pharmaceutical drugs and tracers that interact with any of the four SLC22 transporters investigated in this study (Supplemental Table 1). Approximately 250 drugs were analyzed and the inhibition affinity (Ki) and/or the substrate affinity (Km) were used as a measurement of a given drug's interaction with a given transporter. This "interaction affinity" classified drugs as either high affinity (i.e., Km or Ki $\leq$ 100 μM), mid affinity (i.e., Km or Ki > 100 μM, but $\leq$ 1000 μM), low affinity (i.e., Km or Ki > 1000 μM, but $\leq$ 2000 μM), or extremely low affinity (i.e., Km or Ki > 2000 μM, but $\leq$ 12000 μM). Since these transporters share a great deal of similarity and can interact with the same compounds, albeit usually with different affinities, drugs interacting with two or more transporters with similar interaction affinities (i.e., both interact with high affinity) were excluded from the subsequent data mining analysis aimed at defining the physicochemical descriptors that separate OAT1, OAT3, OCT1, and OCT2.

The net charge states of the drugs at physiological pH (i.e., 7.4) were then determined in the computational environment of ICM software (Molsoft LLC, San Diego, CA). Drugs were considered "cationic" if their net charge was greater than zero; "anionic" if their net charge was less than zero; "neutral" if their net charge was equal to zero and they contained no charged atoms; or "zwitterionic neutral" if their net charge was zero but they contained an equal number of positively charged and negatively charged atoms. However, since it is possible for a drug to have more than one charged species co-existing at a given pH, the percentage of each charge species was determined using pH/concentration curves created using the chemicalize software (www.chemicalize.org; Chemaxon, Cambridge, MA). and

the species percentages were calculated at three different pH values, 7.2, 7.4, and 7.6. Finally, total positive species percentage, total negative species percentage, total neutral species percentage, and total zwitterionic neutral species percentage were calculated for individual charge-species bar diagrams. The results were then plotted.

Collection and preprocessing for machine-learning analyses—The pairwise comparison study employed in the machine-learning analyses was limited to non-overlapping drugs (i.e., drugs which interacted with high affinity for only one transporter). The attributes to be compared in the machine-learning models were physicochemical properties of the drugs calculated using ICM, a commercially available computational chemistry software (Molsoft, San Diego, CA), and tabulated in KNIME, an open-source workflow platform for machine-learning (Beisken et al., 2013). Using ICM, about 50 physicochemical attributes of the drugs were calculated, including molecular quantum numbers, atom counts, bond counts, polarity counts, and topology counts. KNIME includes extensions capable of collecting data from three notable open source cheminformatics toolkits, RDKIT, Indigo, and CDK. Through the KNIME platform using RDKit, Indigo and CDK, attributes were added to represent about 100 chemical features for each drug such as molecular weight, molecular volume, Log P, Log S, polar surface area, etc. In addition to these physicochemical attributes, a class variable was also added to represent with which transporter a given drug would interact.

After collecting the data, Weka, KNIME, and Excel were used to preprocess the data. Weka (cs.waikato.ac.nz/ml/weka) is an open source collection of machine-learning algorithms developed by the University of Waikato and is bundled together with tools for preprocessing data to make it more easily understood by the machine-learning algorithms. For example, the raw data extracted from KNIME and ICM contained some attributes that were overlapping, empty, or constant, and these were eliminated. The second step was to use Weka's attribute selection feature, Chi Square Evaluator, in order to rank the attributes according to their contribution to predicting the class variable. The Chi Square procedure is applied individually to each variable by first binarizing real-valued variables and then testing the expected minus observed counts with respect to the class, where the expected counts are assumed to be

independent; larger counts result in a higher Chi square statistic and suggest non-independence (Agresti and Coull, 1996).

Machine-learning analyses—After compiling and preprocessing the data, machine-learning algorithms were employed to develop models. In this case, drugs that had a "high affinity" for the transporters OAT1, OAT3, OCT1, and OCT2 were treated as "instances", and the physicochemical properties of the drugs were used as "attributes". Six pairwise comparison studies were conducted: OAT1 versus OCT1, OAT1 versus OCT2, OAT3 versus OCT1, OAT3 vs OCT2, OAT1 versus OAT3, and OCT1 versus OCT2; as described above, in each comparison study, "overlapping" drugs, or ones that displayed high-affinity interaction with both of the transporters being compared, were eliminated from the analysis.

Several Weka machine-learning models were utilized: decision trees, decision rules, support vector machine, Bayesian models, and neural networks. Classification models that were well-validated were obtained using several different techniques, but the preference was for those models which could help explain transporter binding/interaction data. For example, neural network models are a "black box" model, so, while, they are not as useful as decision rule or decision tree models for defining distinguishing properties, these are still accurate classifiers. Comparable classification success rates (Supplemental Table 2) with several different algorithms demonstrate that there is a boundary between transporter selectivity. Also, of note, within a given model, depending on the features of the model selected, different decision trees were generated likely due in part to the overlap in molecular characteristics captured by various attributes. Multiple iterations of the algorithms and parameters were explored to arrive at models with the best validation scores.

In a decision tree, each node is a variable, and each branch represents a data split that depends on the value of the variable. An instance of the data determines a path down the tree, which ultimately leads to a leaf node that represents a class prediction. The decision tree is induced by ranking how well each variable can split the data at a decision node (starting with the root), splitting the data, and repeating the process for each branch. As the data gets split more and more, eventually each node will mostly reflect

one class or the other, and the branching will stop. Typically, trees are induced and then lower levels are pruned back to improve performance in a cross-validation procedure. Compared to other techniques, a decision tree is more interpretable because the decisions are easily described.

A random forest is an ensemble of decision trees, in which each tree is trained with different bootstrap samples (1000 in our case). The ensemble is averaged together to produce an aggregate classification. The trees are made slightly de-correlated by limiting the choice of variables during tree induction so that different combinations of variables can fill out the tree branches. An additional benefit of the bootstrap is that one can estimate the detrimental effect of variable permutations on predictions for each "left out of bag" sample. That effect is averaged and normalized over all trees, leading to a measure of variable importance. Because a decision tree is nonlinear in the way it partitions the input, the variable importance is potentially a measure of both interaction and main effects (Svetnik et al., 2003).

Statistical analysis—In addition to the machine-learning approach, statistical tests were used to study the significance of the calculated differences between ligand transporter interactions. In each of the pairwise comparison studies, t-tests were performed on the physiological properties to determine if the differences in the mean values for each were statistically significant between the two groups of drugs. Then, the physiological properties were ranked according to their p-values.

Creation of pharmacophore hypotheses. Pharmacophore models were built in ICM, which performed clustering, alignment, and pharmacophore building based on the atomic property field, APF, of the drug. APF considers the 3D representation of atomic properties such as hydrogen bond donors, hydrogen bond acceptors, SP2 hybridization, lipophilicity, size of large atoms, and positive and negative charges (Totrov, 2008). High affinity drugs were chosen as "actives." Since the actives were diverse in their 3D molecular structures, hierarchical clustering of actives based on APF was first done to separate them into groups. Then, actives among the each group were aligned, and a pharmacophore model was generated from the aligned drugs. In order to be included, each group needed to be comprised of a minimum of 3 drugs with dissimilarity score less than or equal to 0.25. The dissimilarity score is an indication of how similar two compounds are in APF and ranges from 0 to 1, where 0=similarity and

1=dissimilarity. Thus, clusters containing drugs that were too dissimilar would not be considered for pharmacophore model generation. APF properties were determined for each pharmacophore model using ICM, and the vectors of each APF property across all the models were added to calculate the total for that property. More extensive descriptions of this type of approach can be found elsewhere (Khan et al., 2012).

In silico screening and uptake assays—Pharmacophore models were then used to virtually screen the *Drugbank* Database using the ICM computational software. Some top hits were selected for further testing in an in vitro transport/uptake assay for interaction with selected transporters. Uptake assays, with probenecid serving as a negative control, were performed using Chinese hamster ovary (CHO) cells constitutively expressing mouse *Oat3* or *Oat1* as previously described (Ahn et al., 2009; Wu et al., 2013; Wu et al., 2015; Zhu et al., 2015).

JPET Fast Forward. Published on August 3, 2016 as DOI: 10.1124/jpet.116.232660
This article has not been copyedited and formatted. The final version may differ from this version.

RESULTS

The overall goal was to determine whether a formal systematic analysis of the physicochemical descriptors of drugs which interact with SLC22 transporters could: 1) identify properties, other than charge, which would help in predicting whether a ligand interacts with an OAT or OCT and, 2) uncover additional molecular properties of ligands predictive for interaction with prototypical members of these subfamilies (OAT1 vs OAT3 and OCT1 vs OCT2). An extensive literature search identified a large number of pharmaceutical drugs and tracers with the ability to interact with OAT1, OAT3, OCT1, and OCT2 (i.e., 103, 105, 96, and 81, respectively) at all affinity levels (~5μM to 5mm) (Supplemental Table 1); unless otherwise specified, machine-learning analysis, statistical analysis, and pharmacophore modeling were performed using drugs interacting with the transporters in the "high affinity" range (i.e., ≤ 100 μM).

Because there is only limited direct transport data (Km) for these transporters compared to the amount of inhibition data (Ki), the analyses (Ki combined with Km) perforce is weighted toward inhibition data. The literature seems to assume competitive inhibition with a transported substrate (e.g., labeled PAH or TEA), but in nearly all cases the type of inhibition is not formally established by accepted biochemical criteria. This appears to be a general issue for most if not all SLC and ABC drug transporters (Matsson and Bergstrom, 2015). Nevertheless, we also carried out the decision tree analyses described below for those drugs with inhibition (Ki) data alone (excluding those drugs that had Km data), and generally similar results were obtained for these comparatively large datasets (Supplemental Figs. 2, 3; Supplemental Table 3). We also tried to perform the analysis on the much smaller sets of drugs for which Km data was available; while a trend similar to the "Ki plus Km analysis" and the "Ki analysis" was often seen, there did not appear to be large enough samples to achieve clear results (Supplemental Fig. 4; Supplemental Table 3) .

OAT3 has greater capacity to interact with drugs of positively-charged species and zwitteronic-neutral species. Based on the charge-species bar diagrams for individual transporters at pH 7.4 (Figs. 1A, 1B), it was noted that the charged-species that OAT1 and OAT3 mainly interacted with were negatively

charged (i.e., anionic), while OCT1 and OCT2 mainly interacted with positively charged species (i.e,. cationic). Although the next most prevalent charged species with which both OATs and OCTs interacted were the neutral species, all four transporters interacted with zwitterionic-neutral species, as well (Fig. 1). Notably, OAT3 (compared to OAT1) exhibited more ability to interact with zwitterionic-neutral species, as well as those with a charge opposite to that which is suggested by the name "organic anion transporter" (i.e. organic cations) (Fig. 1). At physiological pH (i.e., pH 7.4), OAT1 does not interact with any positively charged species with high affinity; in contrast, OAT3 was able to interact with positively charged species at all affinities (which constituted 3.55% of species with which OAT3 interacts) (Fig. 1C). Both OCT1 and OCT2 interacted with negatively charged species, and the total negatively charged species percentages were 3.80% and 3.17%, respectively (Fig. 1D). Finally, the four transporters interacted with zwitterionic-neutral species to varying degrees; the total zwitterionic-neutral species percentages for OAT1, OAT3, OCT1, and OCT2 were the following: 2.75%, 5.44%, 1.78%, and 2.15%, respectively (Fig. 1). Thus, amongst these SLC22 transporters, OAT3 had the greatest ability to interact with zwitterionic-neutral species. To determine how well individual transporters interacted with "oppositely charged" and zwitterionic-neutral species together, we explored the total percentages of "oppositely charged" species percentage plus zwitterionic-neutral species percentage for each transporter (Figs. 1C, 1D). Among the four transporters, OAT3 had a much higher total percentage than the rest of the transporters (the value for OAT3 was 8.98%, whereas the values for OAT1, OCT1, and OCT2 were 2.75%, 5.58%, and 5.33%, respectively) (Fig. 1). This began to suggest to us that, while OCT1 and OCT2 may be somewhat similar in their ligand specificities, OAT3 might be quite different than OAT1 especially with respect to the ability to interact with cations and zwitterions and may have more similarity (in terms of ligand preference) to OCTs than previously appreciated. This hypothesis was more formally explored in the studies below.

Effect of pH on the ability of transporters to interact with charged and zwitteronic neutral drugs. In addition to analyzing pH 7.4, we explored how varying the pH of the solution in-silico might change the composition of charged species with which that each of the transporters interacted. At different pH

levels, the percent composition of charged species for drugs considered to be anionic, cationic or zwitterionic at pH 7.4, would be expected to vary. In a more acidic environment, drugs would be protonated and contain more positively charged species, while in a more basic environment, drugs would be deprotonated and contain more negatively charged species. This would affect the percent composition of the charged species for a particular drug (anion/cation/zwitterion) with which individual transporters interacted (Figs. 1C, 1D). The sum of the positively charged species percentage and zwitterionic-neutral species percentage for drugs that interact with the organic anion transporters, OAT1 and OAT3, increased as pH decreased, and the sum of negatively charged and zwitterionic species of the organic cation transporters, OCT1 and OCT2, increased when pH shifted toward the basic direction. In addition, it was found that OAT3-interacting drugs (compared to drugs interacting with OAT1, OCT1, and OCT2) likely changed most dramatically throughout the pH range of 7.2 to 7.6; when pH was either lowered or increased, the sum of the total positively charged and zwitterionic-neutral species for OAT3-interacting drugs, changed from 8.31% to 9.93% as the pH was lowered from 7.6 to 7.2. In contrast, the sum of those values for OAT1, OCT1, and OCT2 changed minimally (Figs. 1C, 1D).

Ligand overlap between OAT1 and OAT3 and between OCT1 and OCT2. OAT1 and OAT3 were found to share a number of high affinity ligands with~50% of the drugs showing affinities ≤100 μM for both organic anion transporters; similarly, OCT1 and OCT2 also shared many high affinity ligands with ~35% of these drugs displaying high affinity interactions for both organic cation transporters (Fig. 2). Comparisons of OAT high affinity ligands with those of the OCTs revealed much less overlap with only ~1.8% of OAT1 and ~1.2% of OAT3 high affinity drugs also being able to interact with the organic cation transporters at affinities ≤ 100 μM (Fig. 2), which is consistent with known ligand differences between OATs and OCTs. In order to identify subtle differences in ligand specificity between transporters, the overlapping drugs (i.e., those interacting with two transporters with high affinity) were excluded from the subsequent data mining analysis aimed at defining the physicochemical descriptors that separate OAT1, OAT3, OCT1, and OCT2.

JPET Fast Forward. Published on August 3, 2016 as DOI: 10.1124/jpet.116.232660
This article has not been copyedited and formatted. The final version may differ from this version.

Machine-learning Analysis: Results of Exemplary Models for OAT1 vs OCT1. We initially applied a number of machine-learning approaches. The overall results of applying classification algorithms using decision trees, neural networks, support vector machine, decision rules, and naïve Bayes for the comparison of OAT1 drugs and OCT1 drugs are presented in Supplemental Table 2. Although comparable results were generally obtained, decision trees developed using the J48 algorithm and random forest are discussed in detail for all the pairwise comparisons. Since, in contrast to some of the other approaches which are more like "black boxes" (eg. neural networks), these models not only classified the data well, but provided a very logical way to demonstrate how physicochemical properties of the ligands affect the binding interaction between ligands and transporters. Of note, comparable classification success rates were obtained using different approaches (Supplemental Table 2) which suggests analyzable boundaries related to transporter selectivity.

Differences in substrate specificity are more likely to exist between OAT1 and OAT3 than between OCT1 and OCT2. Table 1 shows the summary of weighted average ROC areas for the six decision tree models based on high-affinity drugs when performing ten-fold cross validation. These were: OAT1/OCT1, OAT1/OCT2, OAT3/OCT1, OAT3/OCT2, OAT1/OAT3, and OCT1/OCT2. Most decision tree models were well-validated, and only two trees had ROC areas less than 0.80, which were the trees for OAT1/OAT3 and OCT1/OCT2. This was likely due to the fact that ligands for the two OATs and two OCTs were highly similar, and it is difficult to build a decision tree model to identify and predict differences. Nevertheless, the ROC areas for OAT1/OAT3 was 0.795, while that for OCT1/OCT2 was 0.639 (Table 1), indicating that the functional differences between OAT1 and OAT3 were more easily discriminated than those between OCT1 and OCT2. This is an important point for the analyses that follow.

Substrate preferences between OATs and OCTs appear to be mostly due to charge. When an OAT was compared with an OCT in decision tree analysis, it was found that the first two physicochemical attributes that separated an OAT from an OCT were the number of negative (nof_negCharge) and positive charges (nof_posCharge) (Fig. 3). This is consistent with previous experimental data across mammalian species for many OAT1 and OCT1 ligands that include not only

drugs, but also metabolites and toxins. Drugs that had the "number of negative charge greater than zero" were classified as OAT-interacting; in contrast, drugs that had the "number of positive charge greater than zero" interacted with OCTs (Fig. 3). While this is compatible with the simple view that OATs transport anions and OCTs transport cations, as we describe elsewhere, a more complex picture emerged with further analysis. For example, even with pairwise comparisons, after charge, the next determinant attribute seen in most trees was SP3 character, in which those drugs with greater SP3 values are classified as OCT drugs. This suggested that drugs with more 3 dimensional/less planar character were more likely to be OCT ligands.

Random Forest models were also used as an independent classification approach. In the variable importance plots derived from the Random Forest model for the pairwise OAT and OCT comparisons, the charge state information was also found to dominate in the ranking (Fig. 4). This supports the notion that the higher nodes in the decision tree are robustly important for classification across the bootstrap samples in the Random Forest. In addition, the variables found to be most important after the charge state were the number of acyclic double bonds (adb), acyclic oxygens (ao), followed by the "SP3 character." After 5 or 6 variables, the importance levels drop off and little is gained by considering additional variables. For the pairwise comparison between OATs, the results also confirm and justify the decision tree interpretation. However, for OCT1 versus OCT 2, the results are not aligned, which is not surprising given that the classification performance is poor (Fig. 4).

Exclusion of charge reveals potential role of physicochemical properties other than charge in substrate preference differences. The Random Forest models pointed to the potential role that other physicochemical features of the high affinity drugs might play in separating OAT-interacting drugs from OCT-interacting drugs in additional to charge. Therefore, decision trees were constructed that excluded the properties of positive and negative charge (Fig. 5). The resulting trees split on a variety of other properties; the number of acyclic double bonds ("adb"), number of acyclic oxygens ("ao"), number of acyclic nitrogens ("an"), and the "SP3 character" were dominant.

In the OAT1/OCT1 tree, the first attribute that split was "adv" (acyclic divalent nodes); drugs that had zero "adv" were classified as OCT1 drugs. The next attribute was "number of aliphatic bonds," and drugs with the greater number of aliphatic bonds were classified as OCT1 drugs. When we examined the three other OAT vs OCT trees, they followed a similar trends as the OAT1/OCT1 tree; OCT ligands generally had higher number of "an" than OAT ligands, and OAT ligands had higher numbers of "adb" and "ao" than OCT ligands. Interestingly, statistics from the accuracy of these decision tree models (which excluded charge) were not as strong as ones including charge, but were still reasonable; thus, the models correct classification of drugs was between 76% to 81% and results of ROC area were between 0.75 to 0.83 (data not shown). In addition, the attributes "ao", "adb" and "SP3 character" were confirmed as important attributes in the t-test statistical analysis (below and Table 2).

Pair-wise comparison between OAT1 and OAT3 reveals differences between the two OATs. When OAT1 and OAT3 were compared (Fig. 3E), the first attribute separating OAT1 and OAT3 ligands was the number of acyclic tetravalent nodes ("aqv"). Drugs that have the greater number of acyclic tetravalent nodes tended to be classified as interacting with OAT3. The next attribute separating the OAT ligands was the number of phosphorous atoms ("p"). Drugs that had at least one or more phosphorus atoms tended to be classified as OAT1-interacting. A third attribute that emerged from these comparisons of OAT1 and OAT3 ligands was the number of positive charges; drugs with a positive charge were associated with an OAT3 classification (Fig. 3E). (The aforementioned properties will be discussed in more detail below when we present wet lab support for the computational analysis.) In contrast to the comparison of the two OATs, the model generated for comparison of the two OCTs had poor validation performance; it appears that OCT ligands are too similar to be distinguished by the approaches we used, and hence, the results for that decision tree model will not be discussed further.

Statistical analysis confirmed the machine-learning analyses. When performing t-test analyses on individual attributes for each pairwise transporter comparison, we identified a number of attributes as statistically different between ligands interacting with each pair of transporters. The attributes that had the lowest p-values for each comparison are summarized in Table 2 and are consistent with the machine-

learning analyses. The two properties that had the lowest p-values were the "number of positive charge" and the "number of negative charge", corresponding to the results from the machine-learning analyses. After positive and negative charge, the next attributes that came out from the ranking were numbers of acyclic double bond ("adb"), acyclic oxygen ("ao"), hydrogen bond acceptor site ("hbam"), and SP3 character (Table 2). For the pair-wise comparison of the two OATs, the two properties seen in the OAT1/OAT3 decision tree (i.e., the "number of acyclic tetravalent nodes" ("aqv") and "number of positive charges") were also found to have the lowest p-values in the ranking. Again, the results from both decision trees and random forest are consistent with the statistical analysis.

Explanation of properties found to be relevant in results. As described above, based on the results of machine-learning and statistical tests, we found that ligands of the OATs (either OAT1 or OAT3) generally had higher numbers of negative charge, acyclic double bonds, acyclic oxygen, and hydrogen bond acceptor sites than an OCT ligand (either OCT1 or OCT2). These properties tend to be associated with the anionic propensity. For example, most acyclic double bonds within the structures were in the forms of carbonyl (O=C), thial (S=C), sulfoxide (S=O), and the electro-negative oxygen and sulfur within these double bonds are prominent hydrogen bond accepting sites. The "number of acyclic oxygen" is another property that expresses the anionic propensity as the acyclic oxygen also serves as a potential hydrogen bond accepting site.

Importantly, in addition to having differences in properties associated with charges and ionization, ligands of OCTs and OATs are different in geometry-related properties, particularly with respect to the SP3 character value. SP3 character is defined as the number of SP3-hybridized carbons divided by the total number of atoms; it is one measure of the degree of three-dimensionality of a compound. If a drug has a higher SP3 character value, it is more three-dimensional; likewise, a lower SP3 character value is taken to imply that the drug is more planar (Lovering et al., 2009; Over et al., 2014). In machine-learning models and statistical analyses, drugs with a stronger affinity for the OCTs had a greater SP3 character value than those with a stronger affinity for the OATs, supporting the view that the "OCT interacting drugs" are more three-dimensional than "OAT interacting drugs". As measured by SP3 character,

compared with most other drugs in the data set, amantadine, nandrolone, and atropine are three OCT drugs that have highly three-dimensional structures, each with a SP3 character value of 0.357, 0.326, and 0.227, respectively. On the other hand, OAT drugs have much lower values of SP3 character, with none of the OAT drugs having SP3 character values greater than 0.300.

Some differences are also observed among the ligands of the two OATs; OAT3 tended to interact with drugs that have more acyclic tetravalent nodes and more positive charges, whereas OAT1 tended to interact with those that have more phosphorus atoms. An acyclic tetravalent node usually is composed of a carbon-forming tetravalent bond with four elements. In the decision tree model, 11 drugs were classified as OAT3 drugs from this node; among them were verapamil, pravastatin, enalapril, and methotrexate, and along with the higher number of acyclic tetravalent nodes, these drugs have longer and more hydrophobic chains. The next attribute separating OAT1 and OAT3 ligands was the number of phosphorous atoms ("p"). Drugs that had at least one or more phosphorus atoms were classified as interacting with OAT1; the three drugs in this category were cidofovir, tenofovir, and adefovir. When looking at the chemical structures of these drugs, it was found the phosphorus atoms were in phosphate groups. Since the phosphate groups contain several oxygen atoms binding with phosphorus—some of which were deprotonated at the normal pH range—the phosphate group is highly anionic. Thus, the number of phosphorus atoms was directly correlated with the anionic propensity. In summary, even though both OAT1 and OAT3 were known to have functional overlap, there were some differences between their ligands identified in our analyses. OAT3 preferred to interact with drugs with more positive charge and long hydrophobic chains, and OAT1 ligands tended to be more anionic than OAT3.

Analysis of mid-affinity drugs supports the results of high affinity drugs. Well-described OAT ligands verified in vivo in knockouts include many compounds with an affinity greater than 100 μM (Eraly et al., 2006; Vallon et al., 2008a; Wikoff et al., 2011; Wu et al., 2013; Nigam, 2015; Nigam et al., 2015a; Nigam et al., 2015b). Thus, in addition to understanding the molecular interactions between transporters and drugs that bind with high affinity (≤100 μM), we also tried to study how OAT1, OAT3, OCT1 and OCT2 interact with drugs in the mid affinity range (100μM to 1000μM). The decision trees

based on mid affinity drugs (Fig. 6) demonstrates that major factors involved in classifying a drug as an OAT or an OCT substrate were due to charge, as in the high affinity group. But the separation was less impressive than for the high affinity (<100 μM) drugs. The decision tree comparing OAT1 and OAT3 in the mid-affinity range only had one node, which split on positive charge (Fig. 6). Drugs with a positive charge were classified as OAT3-interacting.

3D pharmacophore models showed structural similarities corresponding to the overlap in functions for OATs and for OCTs. Since it was found that OAT3 ligands also possessed some cationic characteristics based on the machine-learning analyses, pharmacophore models for OAT3, OCT1, and OCT2 interacting drugs were built to compare the functional similarities/differences between the OAT and the OCTs in three-dimensional space (Fig. 7). The models showed that OAT3 and OCTs interacted with drugs that had hydrophobic and aromatic centers. However, a slight difference in compound backbone appeared as the hydrophobic chains for OCT1 and OCT2 models would sometimes enclose cationic spheres (seen in OCT1 pharmacophore model 3, 4, 5, and 6), which is not observed in most OAT3 models. Overall, models of OAT3 interacting ligands were more anionic, and models of OCT interacting ligands were more cationic. This can also be seen from Table 3, which shows the quantitative measurements of the seven properties for individual models; as measured by the mean, the table shows that ligands of the OATs had higher “hydrogen bond acceptors” and higher “negative charges”; in contrast, ligands of the OCTs had higher “hydrogen bond donors” and higher “positive charges”.

The pharmacophore models revealed structural similarities between ligands of OAT3 and OCT1. Even though the majority of pharmacophore models for ligands of OAT3 had similar features, there was one clear exception, which was the pharmacophore model based on group 9 for OAT3 (Fig. 7). Unlike other OAT ligand models, this model contained a hydrophobic chain that tended to enclose a sphere enriched with hydrogen bond donors and positive charges, which was a pattern shared among many OCT1 and OCT2 ligand models. Thus, this model (OAT3 pharmacophore model 9) was found to be very OCT-like, and the quantitative APF measurement of this model was found to have greater values of “positive charges” and “electropositive charges.”

Interestingly, the list of drugs used to construct this model from group 9 for OAT3 was found to be highly similar to the list of drugs that was independently separated based on the first attribute or node in the OAT1/OAT3 decision tree (Fig. 3). Out of the 9 drugs used to construct the pharmacophore model, 6 of them contained more than 7 acyclic tetravalent nodes and were classified as "OAT3" drugs in the decision tree. This is important since it demonstrates that the results from the decision trees and the pharmacophore models identified the same differences found between ligands of OAT1 and OAT3, and the differences were due to the apparent capability of OAT3 to interact with OCT-like substrates.

Experimental validation of in-silico screening results identified new cationic drugs that preferentially interact with OAT3 but not OAT1. The finding that OAT3 prefers more cationic substrates than does OAT1 was thus consistent in decision tree and Random Forest analyses, and there was one (cationic) OAT3 pharmacophore model that was strikingly similar to OCT pharmacophore models. Thus, with the idea of trying to validate this experimentally, the OAT3 cationic model was used for in silico-based virtual screening. Using the pharmacophore model based on group 9 of the OAT3 substrates, a virtual screen of Drugbank Database was done to identify potential new OAT3 cationic ligands. Six top hits were selected for further wet-lab validation. These hits were then tested for their ability to interact selectively with OAT3 using wet lab transport assays in OAT1-expressing or OAT3-expressing cells. Four of the ligands were found to interact with OAT3, with strong inhibition of tracer uptake. In marked contrast, when these 6 cationic drugs were tested in the OAT1 uptake assay, it was found that only two of them inhibited OAT1 function, and, importantly, with a much lower affinity (Fig. 8). The preference of these compounds for interaction with OAT3 but not OAT1, not only supports the validity of the pharmacophore model (model 9) but it is consistent with the machine-learning analysis indicating the capability of OAT3 to interact with cationic drugs. The measured IC50 values of tested compounds against OAT1 and OAT3 are summarized in Table 4.

DISCUSSION

Recent knockout and in vitro data on a limited set of ligands suggest that the specificity of the OATs and OCTs of greatest clinical and pharmaceutical interest goes beyond whether or not the ligand is an anion or a cation (Ahn et al., 2009; Nigam, 2015; Nigam et al., 2015a; Nigam et al., 2015b). Thus, molecular properties other than ligand charge need to be carefully addressed. To systemically examine this question, an extensive literature search was first done to build a complete as possible transporter-ligand database (nearly any compound found to interact with the transporters of interest was initially curated). Within this data (Supplemental Table 1), all drugs reliably known to interact with OAT1, OAT3, OCT1, and OCT2 were selected and used to study the functional differences and similarities between the transporters by applying machine-learning tools. Among the machine-learning tools (which included neural nets, support vector machines and other methods as shown in Supplemental Table 2), decision trees and random forests were more helpful from the viewpoint of understanding this question of substrate specificity as opposed to simply fitting data (Figs. 3-6).

The results of the decision tree analyses were in agreement with the results of the random forest, and these results were further verified by conventional statistical tests (Table 2). The results indicated that, while the main difference between the ligand preferences of OATs and OCTs (with respect to physicochemical descriptors) was charge, the structure of ligands also affected the interaction with the transporters. Thus, in considering factors beyond charge, OCTs interacted with more three-dimensional structures (more SP3 character), whereas OATs interacted with planar compounds (Figs. 3-5). This may imply that the binding pockets of OCTs accommodate less planar compounds than those of OATs, which is worthy of further investigation once crystal structures of these transporters become available (Koepsell, 2013; Matsson and Bergstrom, 2015; Nigam et al., 2015a).

In addition to finding differences between OATs and OCTs, some differences among the sub-members of these families were also identified. Based on the machine-learning models and pharmacophore models, OAT1 and OAT3 were found to be different in that the latter possesses some ability to interact with cations, making it more functionally similar to OCT1 and OCT2 in this respect

(Fig. 3, 5). Among high affinity drugs (<100mm Km, Ki or IC50), OAT3 could interact with ligands with more diverse structures (per machine-learning analysis of physicochemical descriptors and pharmacophore analysis) than OAT1, again implying that OAT3 has different binding pockets than OAT1 and supporting the importance of obtaining structures for both transporters.

Based on the pharmacophore OAT3/OCT1 (Fig. 7) overlay, OAT3 binding pockets could have similarity to binding pockets of the OCTs, enabling OAT3 to bind some ligands with cationic characteristics. Our studies indicate that while OAT1, OAT3, OCT1, and OCT2 are "multispecific" (or "polyspecific"), this multispecificity (polyspecificity) is restricted, and the actual interaction of each transporter with their ligands goes beyond conventional views about charge. This is our main finding, supported by machine-learning analysis, pharmacophore modeling, and wet lab transport assays. In particular, OAT3 stands out. While OAT3 has overlapping ligands with OAT1, and, like OAT1, it has a preference for planar anionic molecules, OAT3 also accepts larger ligands and more cationic/zwitterionic ones—including those that might conventionally be viewed as OCT substrates. We support this conclusion with wet lab data using an OAT3 transport assay indicating that cationic drugs not previously reported (as far as we know) to be ligands indeed interact with OAT3. Together, the computational and wet lab analyses indicate that the boundary that separates OATs and OCTs is not as clear as the current literature suggests.

Thus, finding the differences and similarities between the transporters with respect to ligand preference can help to predict and identify new compounds that interact with the transporter (as we have done here), since the set of rules defined by decision trees can be further used for in-silico screening of new ligands/inhibitors (drugs, toxins, metabolites, signaling molecules). These rules can also be used to design new, potent, selective ligands that can target a particular transporter. These could be drugs that are aimed at targeting a particular tissue or body fluid, or alternatively, selective inhibitors of transport.

Expression of varying levels of OAT1, OAT3, OCT1, and OCT2 may thus help the cell alter the net ligand (drugs, toxins, metabolites, signaling molecules) taken up by kidney, liver, and other tissues in non-obvious ways. The potential relevance of this concept to normal physiology and pathophysiological

states has been discussed in the Remote Sensing and Signaling Hypothesis (Kaler et al., 2006; Ahn and Nigam, 2009; Wu et al., 2011; Nigam, 2015; Nigam et al., 2015a). Our results should also be useful for predicting potential drug-drug interactions (DDI) and drug-metabolite interactions (DMI).

As discussed throughout this article, the study may be somewhat limited due to paucity of direct transport data and the reliance on inhibition data. As indicated in a recent review addressing ligand-based modeling of SLC and ABC drug transporters, the limited transport data available is an issue for the whole field (Matsson and Bergstrom, 2015); even with inhibition data, competitive versus non-competitive inhibition is also generally not addressed although the former is often assumed (Matsson and Bergstrom, 2015). However, at least for the drugs studied here the limited transport data was quite consistent with binding data. In support of this notion, one can also consider in vivo studies in the *Oat* and *Oct* knockout animals. A number of general classes of organic anion, organ cation and organic zwitterion compounds analyzed here (e.g., antivirals, antibiotics, diuretics, metformin, zwitterions) have also been evaluated in the *Oat1*, *Oat3*, *Oct1* and *Oct2* knockout animals or in knockout tissues, and abnormalities in handling of these compounds consistent with inhibition affinities have been demonstrated (Eraly et al., 2006; Vanwert et al., 2007; Truong et al., 2008; Vallon et al., 2008b; Vanwert et al., 2008; Nagle et al., 2011; Vallon et al., 2012; Nagle et al., 2013). Indeed, the knockout data even seems to support the preference of Oat3 (compared to Oat1) for zwitterions such as creatinine (Vallon et al., 2012). Nevertheless, caution about relying entirely on inhibition data seems appropriate as there may be cases where high affinity binding to transporters such as Oct1 may not necessarily correspond to physiologically-relevant transport (He et al., 2016).

As discussed above, we also performed decision tree analyses on the set of drugs that had inhibition (Ki) data (not including drugs with transport data as indicated by Km values); in this analysis, similar results were obtained to the larger dataset consisting of both Ki and Km data (Supplemental Fig. 2). In addition, we attempted to obtain reliable decision trees for the considerably smaller set of compounds for which transport (Km) data had been found (Supplemental Fig.3). Although similar trends (to the Ki plus Km decision trees) were found in some cases, clear, consistent and significant results could

not be generally obtained with this limited set of compounds with Km values. This, again, highlights the need for the field to obtain transport data for all the drugs and, with respect to inhibition data, the need to distinguish competitive from non-competitive inhibition (Matsson and Bergstrom, 2015). In addition, it can be argued that ligand-based modeling for multispecific SLC drug transporters, which handle structurally diverse compounds, might be more difficult than for transporters that handle a single class of structurally similar compounds (Matsson and Bergstrom, 2015). This is one reason we believed it was reasonable to use as large a dataset as possible, despite the limitations described above—an approach that was partly experimentally validated. As more transport and other biochemical data becomes available, and as machine-learning and other data science approaches continue to improve, it may be possible to obtain an even clearer picture of the chemical features of drugs that enable transport by one or another SLC and/or ABC transporter.

Author contributions:

Conceived the hypothesis, supervised the experiments: Nigam

Conducted experiments: Liu, Goldenberg, Wu

Performed analysis: Liu, Goldenberg, Wu

Provided methodological assistance: Chen, Lun

Contributed analytical tools: Balac, Rodriguez, and Abagyan

Wrote or contributed to the writing of the manuscript: Liu, Goldenberg, Bush, and Nigam

REFERENCES

Agresti A and Coull BA (1996) Order-restricted tests for stratified comparisons of binomial proportions. *Biometrics* **52**:1103-1111.

Ahn SY, Eraly SA, Tsigelny I and Nigam SK (2009) Interaction of organic cations with organic anion transporters. *J Biol Chem* **284**:31422-31430.

Ahn SY, Jamshidi N, Mo ML, Wu W, Eraly SA, Dnyanmote A, Bush KT, Gallegos TF, Sweet DH, Palsson BO and Nigam SK (2011) Linkage of organic anion transporter-1 to metabolic pathways through integrated "omics"-driven network and functional analysis. *J Biol Chem* **286**:31522-31531.

Ahn SY and Nigam SK (2009) Toward a systems level understanding of organic anion and other multispecific drug transporters: a remote sensing and signaling hypothesis. *Mol Pharmacol* **76**:481-490.

Beisken S, Meinl T, Wiswedel B, de Figueiredo LF, Berthold M and Steinbeck C (2013) KNIME-CDK: Workflow-driven cheminformatics. *BMC Bioinformatics* **14**:257.

Duan P, Li S, Ai N, Hu L, Welsh WJ and You G (2012) Potent inhibitors of human organic anion transporters 1 and 3 from clinical drug libraries: discovery and molecular characterization. *Mol Pharm* **9**:3340-3346.

Emami Riedmaier A, Nies AT, Schaeffeler E and Schwab M (2012) Organic anion transporters and their implications in pharmacotherapy. *Pharmacol Rev* **64**:421-449.

Eraly SA, Monte JC and Nigam SK (2004) Novel slc22 transporter homologs in fly, worm, and human clarify the phylogeny of organic anion and cation transporters. *Physiol Genomics* **18**:12-24.

Eraly SA, Vallon V, Vaughn DA, Gangoiti JA, Richter K, Nagle M, Monte JC, Rieg T, Truong DM, Long JM, Barshop BA, Kaler G and Nigam SK (2006) Decreased Renal Organic Anion Secretion and Plasma Accumulation of Endogenous Organic Anions in OAT1 Knock-out Mice. *J Biol Chem* **281**:5072-5083.

Hazai E, Hazai I, Ragueneau-Majlessi I, Chung SP, Bikadi Z and Mao Q (2013) Predicting substrates of the human breast cancer resistance protein using a support vector machine method. *BMC Bioinformatics* **14**:130.

He X, Garza D, Nigam SK and Chang G (2016) Multispecific Organic Cation Transporter 1 (OCT1) from Bos taurus Has High Affinity and Slow Binding Kinetics towards Prostaglandin E2. *PLoS One* **11**:e0152969.

Imamura Y, Murayama N, Okudaira N, Kurihara A, Okazaki O, Izumi T, Inoue K, Yuasa H, Kusuhara H and Sugiyama Y (2011) Prediction of fluoroquinolone-induced elevation in serum creatinine levels: a case of drug-endogenous substance interaction involving the inhibition of renal secretion. *Clin Pharmacol Ther* **89**:81-88.

Kaler G, Truong DM, Khandelwal A, Nagle M, Eraly SA, Swaan PW and Nigam SK (2007) Structural variation governs substrate specificity for organic anion transporter (OAT) homologs. Potential remote sensing by OAT family members. *J Biol Chem* **282**:23841-23853.

Kaler G, Truong DM, Sweeney DE, Logan DW, Nagle M, Wu W, Eraly SA and Nigam SK (2006) Olfactory mucosa-expressed organic anion transporter, Oat6, manifests high affinity interactions with odorant organic anions. *Biochem Biophys Res Commun* **351**:872-876.

Khan MT, Wuxiuer Y and Sylte I (2012) Binding modes and pharmacophore modelling of thermolysin inhibitors. *Mini Rev Med Chem* **12**:515-533.

Kim SP, Gupta D, Israni AK and Kasiske BL (2015) Accept/decline decision module for the liver simulated allocation model. *Health Care Manag Sci* **18**:35-57.

Koepsell H (2013) The SLC22 family with transporters of organic cations, anions and zwitterions. *Mol Aspects Med* **34**:413-435.

Kouznetsova VL, Tsigelny IF, Nagle MA and Nigam SK (2011) Elucidation of common pharmacophores from analysis of targeted metabolites transported by the multispecific drug transporter-Organic anion transporter1 (Oat1). *Bioorg Med Chem* **19**:3320-3340.

Lopez-Nieto CE, You G, Bush KT, Barros EJ, Beier DR and Nigam SK (1997) Molecular cloning and characterization of NKT, a gene product related to the organic cation transporter family that is almost exclusively expressed in the kidney. *J Biol Chem* **272**:6471-6478.

Lovering F, Kirincich S, Wang W, Combs K, Resnick L, Sabalski JE, Butera J, Liu J, Parris K and Telliez JB (2009) Identification and SAR of squarate inhibitors of mitogen activated protein kinase-activated protein kinase 2 (MK-2). *Bioorg Med Chem* **17**:3342-3351.

Matsson P and Bergstrom CA (2015) Computational modeling to predict the functions and impact of drug transporters. *In Silico Pharmacol* **3**:8.

Nagle MA, Truong DM, Dnyanmote AV, Ahn SY, Eraly SA, Wu W and Nigam SK (2011) Analysis of three-dimensional systems for developing and mature kidneys clarifies the role of OAT1 and OAT3 in antiviral handling. *J Biol Chem* **286**:243-251.

Nagle MA, Wu W, Eraly SA and Nigam SK (2013) Organic anion transport pathways in antiviral handling in choroid plexus in Oat1 (Slc22a6) and Oat3 (Slc22a8) deficient tissue. *Neurosci Lett* **534**:133-138.

Nigam SK (2015) What do drug transporters really do? *Nat Rev Drug Discov* **14**:29-44.

Nigam SK, Bush KT, Martovetsky G, Ahn SY, Liu HC, Richard E, Bhatnagar V and Wu W (2015a) The organic anion transporter (OAT) family: a systems biology perspective. *Physiol Rev* **95**:83-123.

Nigam SK, Wu W, Bush KT, Hoenig MP, Blantz RC and Bhatnagar V (2015b) Handling of Drugs, Metabolites, and Uremic Toxins by Kidney Proximal Tubule Drug Transporters. *Clin J Am Soc Nephrol* **10**:2039-2049.

Over B, McCarren P, Artursson P, Foley M, Giordanetto F, Gronberg G, Hilgendorf C, Lee MDt, Matsson P, Muncipinto G, Pellisson M, Perry MW, Svensson R, Duvall JR and Kihlberg J (2014) Impact of stereospecific intramolecular hydrogen bonding on cell permeability and physicochemical properties. *J Med Chem* **57**:2746-2754.

Popp C, Gorboulev V, Muller TD, Gorbunov D, Shatskaya N and Koepsell H (2005) Amino acids critical for substrate affinity of rat organic cation transporter 1 line the substrate binding region in a model derived from the tertiary structure of lactose permease. *Mol Pharmacol* **67**:1600-1611.

Svetnik V, Liaw A, Tong C, Culberson JC, Sheridan RP and Feuston BP (2003) Random forest: a classification and regression tool for compound classification and QSAR modeling. *J Chem Inf Comput Sci* **43**:1947-1958.

Totrov M (2008) Atomic property fields: generalized 3D pharmacophoric potential for automated ligand superposition, pharmacophore elucidation and 3D QSAR. *Chem Biol Drug Des* **71**:15-27.

Truong DM, Kaler G, Khandelwal A, Swaan PW and Nigam SK (2008) Multi-level analysis of organic anion transporters 1, 3, and 6 reveals major differences in structural determinants of antiviral discrimination. *J Biol Chem* **283**:8654-8663.

Vaglio Laurin G, Cheung-Wai Chan J, Chen Q, Lindsell JA, Coomes DA, Guerriero L, Del Frate F, Miglietta F and Valentini R (2014) Biodiversity mapping in a tropical West African forest with airborne hyperspectral data. *PLoS One* **9**:e97910.

Vallon V, Eraly SA, Rao SR, Gerasimova M, Rose M, Nagle M, Anzai N, Smith T, Sharma K, Nigam SK and Rieg T (2012) A role for the organic anion transporter OAT3 in renal creatinine secretion in mice. *Am J Physiol Renal Physiol* **302**:F1293-1299.

Vallon V, Eraly SA, Wikoff WR, Rieg T, Kaler G, Truong DM, Ahn SY, Mahapatra NR, Mahata SK, Gangoiti JA, Wu W, Barshop BA, Siuzdak G and Nigam SK (2008a) Organic anion transporter 3 contributes to the regulation of blood pressure. *J Am Soc Nephrol* **19**:1732-1740.

Vallon V, Rieg T, Ahn SY, Wu W, Eraly SA and Nigam SK (2008b) Overlapping in vitro and in vivo specificities of the organic anion transporters OAT1 and OAT3 for loop and thiazide diuretics. *Am J Physiol Renal Physiol* **294**:F867-873.

Vanwert AL, Bailey RM and Sweet DH (2007) Organic anion transporter 3 (Oat3/Slc22a8) knockout mice exhibit altered clearance and distribution of penicillin G. *Am J Physiol Renal Physiol* **293**:F1332-1341.

Vanwert AL, Srimaroeng C and Sweet DH (2008) Organic anion transporter 3 (oat3/slc22a8) interacts with carboxyfluoroquinolones, and deletion increases systemic exposure to ciprofloxacin. *Mol Pharmacol* **74**:122-131.

Wang Z, Chen Y, Liang H, Bender A, Glen RC and Yan A (2011) P-glycoprotein substrate models using support vector machines based on a comprehensive data set. *J Chem Inf Model* **51**:1447-1456.

Wikoff WR, Nagle MA, Kouznetsova VL, Tsigelny IF and Nigam SK (2011) Untargeted metabolomics identifies enterobiome metabolites and putative uremic toxins as substrates of organic anion transporter 1 (Oat1). *J Proteome Res* **10**:2842-2851.

Wu W, Baker ME, Eraly SA, Bush KT and Nigam SK (2009) Analysis of a large cluster of SLC22 transporter genes, including novel USTs, reveals species-specific amplification of subsets of family members. *Physiol Genomics* **38**:116-124.

Wu W, Bush KT, Liu HC, Zhu C, Abagyan R and Nigam SK (2015) Shared Ligands Between Organic Anion Transporters (OAT1 and OAT6) and Odorant Receptors. *Drug Metab Dispos* **43**:1855-1863.

Wu W, Dnyanmote AV and Nigam SK (2011) Remote communication through solute carriers and ATP binding cassette drug transporter pathways: an update on the remote sensing and signaling hypothesis. *Mol Pharmacol* **79**:795-805.

Wu W, Jamshidi N, Eraly SA, Liu HC, Bush KT, Palsson BO and Nigam SK (2013) Multispecific drug transporter Slc22a8 (Oat3) regulates multiple metabolic and signaling pathways. *Drug Metab Dispos* **41**:1825-1834.

Zhu C, Nigam KB, Date RC, Bush KT, Springer SA, Saier MH, Jr., Wu W and Nigam SK (2015) Evolutionary Analysis and Classification of OATs, OCTs, OCTNs, and Other SLC22 Transporters: Structure-Function Implications and Analysis of Sequence Motifs. *PLoS One* **10**:e0140569.

FOOTNOTES

This work was supported by the National Institutes of Health Grants National Institute of General Medical Sciences [Grants R01-GM098449, RO1GM104098], and by the National Institutes of Health Grants Eunice Kennedy Shriver National Institute of Child Health and Human Development [U54-HD07160].

Financial Disclosure: The authors declare the following competing financial interest(s): Ruben Abagyan has significant financial interest in Molsoft or its products. Yuchen Chen is a Molsoft employee.

To whom correspondence should be addressed: Sanjay K. Nigam, Department of Pediatrics, University of California San Diego, 9500 Gilman Drive, MC0693, La Jolla, CA 92093. USA. E-mail: snigam@ucsd.edu

JPET Fast Forward. Published on August 3, 2016 as DOI: 10.1124/jpet.116.232660
This article has not been copyedited and formatted. The final version may differ from this version.

FIGURE LEGENDS

FIGURE 1. (A) The distribution of charge states for pharmaceuticals that interacted with each of the transporters at various binding affinity ranges. The charge states of the pharmaceuticals were defined by considering the number of positive charges and negative charges calculated in ICM at the environment of pH = 7.4. (B) The charge-species composition diagrams for the transporters. The charge species composition for individual pharmaceuticals was measured based on the pH/concentration curves found in Chemicalize.org (an online compound database supported by Chemaxon), which were then grouped according to the transporters with which the pharmaceuticals interacted. The diagrams could indicate the capability of the transporters to interact with various charge species. (C-D) Summary of the total percentage of various charge species for each transporter based on the results of charge-species composition diagrams. The potential capability of OAT3 to interact with positively charged and zwitterionic species (Arrow) was thereby clarified.

FIGURE 2. Substrate overlap among transporters. The Venn diagram demonstrates the substrate specificity and substrate multi-specificity between the transporters. Drugs found to be overlapping between the various transporters were excluded for the subsequent machine-learning analysis. Note: While cimetidine and verapamil can bind OAT1 as well (Ahn et al., 2009), the affinity is roughly 10-fold less.

FIGURE 3. Decision trees based on those drugs interacting with the transporters with high-affinity (i.e., $\leq$ 100 μM). The decision trees show that the main difference between OATs and OCTs are due to charge and charge-associated properties. Besides charge, the 3-dimensionality versus planarity of the drug, indicated by SP3 character, was found to be another important factor in separating OAT and OCT drugs. In addition, some differences were found between two OATs, specifically in number of aqv, p, and posCharge (these attributes are further explained in the text).

FIGURE 4. Results based on the Random Forest analyses. As discussed in the text, these results are highly comparable to the results from decision trees.

FIGURE 5. Decision trees excluding charge properties. The attributes of positive and negative charge were excluded in the building of the model so as to identify other important properties that potentially segregate OAT and OCT drugs. Again, it was found that some charge-associated attributes and the SP3 character were key determinants.

FIGURE 6. The decision trees based on drugs that interact with the transporters at mid-affinity range (between 100 and 1000 μM). The trees show that, in the mid-affinity range, the main differences between OAT and OCT interacting drugs were still due to charges, but to a lesser degree than drugs in the high-affinity range.

FIGURE 7. The pharmacophore models for OAT3, OCT1 and OCT2. Since the drugs interacting with each transporter were diverse in their 3D structures, the drugs were clustered into groups based on atomic property field (APF). Then, the drugs within the same clustering groups were aligned, and pharmacophore models for each group were created. In the pharmacophore models, different colors represent various APF properties: blue - hydrogen bond donor; red – hydrogen bond acceptor; white – aromaticity; yellow – hydrophobicity; light red – negative charges; light blue – positive charges. The OAT3 models and OCT1 models are found to be distinctive. With one notable exception, the OAT3 models for each group contained more characteristics of negative charges, electronegativity, and hydrogen bond acceptors, and vice versa for OCT1 models. However, the OAT3 model derived from group 9 was an exception as it contained several characteristics found largely in models from OCT groups.

FIGURE 8. Uptake inhibition assay based on the virtual screening of the OAT3 pharmacophore model from group 9 (i.e., cationic pharmacophore) against the Drugbank database. For OAT1 inhibition assay, 10 μM 6CF was used as fluorescent tracer, and for OAT3 assay, 20 μM 5CF was used. Please see methods and text for additional details.

JPET Fast Forward. Published on August 3, 2016 as DOI: 10.1124/jpet.116.232660
This article has not been copyedited and formatted. The final version may differ from this version.

TABLES

TABLE 1

[a]Weighted Average ROC Areas:

Performance validation of various decision tree analyses

Transporters Compared	High Affinity Drugs				Mid Affinity Drugs	
	Charge				Charge included as an attribute	
	Included		Excluded			
	Correctly classified	ROC area	Correctly classified	ROC area	Correctly classified	ROC area
OAT1/OCT1	86.57%	0.905	80.60%	0.823	82.50%	0.874
OAT1/OCT2	83.33%	0.932	78.95%	0.835	82.22%	0.868
OAT3/OCT1	86.33%	0.880	77.70%	0.764	80.00%	0.880
OAT3/OCT2	93.28%	0.932	72.27%	0.774	70.83%	0.779
OAT1/OAT3	69.77%	0.795	NA	NA	86.37%	0.722
[b]OCT1/OCT2	66.67%	0.639	NA	NA	45.45%	0.450

[a]The table summarizes the results using ten-fold cross validation of machine-learning decision tree models for: a) high affinity drugs (with affinity less than 100 μM), b) high affinity drugs without using charge as an attribute, and c) mid affinity drugs (with affinity between 100 to 1000 μM).

[b]Note the poor results in the OCT1/OCT2 analysis are likely due to a small data set of 6 and 5 instances.

JPET Fast Forward. Published on August 3, 2016 as DOI: 10.1124/jpet.116.232660
This article has not been copyedited and formatted. The final version may differ from this version.

TABLE 2

[a]Pairwise comparisons of Individual Attributes for the 4 SLC22 Transporters

Pair-wise comparison	Number of attributes P-value <0.005	<0.01	<0.05	Top 8 attributes ranked by p-value[a]							
OAT1 vs OCT1	27	29	37	nof_posCharge (p=5.95E-23)	nof_negCharge (p=4.92E-13)	nof_adb (p=2.4E-12)	nof_ao (p=3.17E-09)	nof_hbam (p=2.28E-08)	SP3 Character (p=1.34E-07)	Number of hydrogens (p=7.93E-07)	nof_asv (p=1.11E-06)
OAT3 Vs OCT1	17	18	23	nof_negCharge (p=1.10E-16)	nof_posCharge (p=2.44E-15)	nof_adb (p=4.98E-13)	nof_ao (p=3.07E-10)	nof_hbam (p=7.08E-09)	molPSA (p=1.38E-06)	SP3 Character (p=2.7E-06)	Topological Polar Surface Area (p=3.63E-06)
OAT1 Vs OCT2	16	21	28	nof_posCharge (p=1.45E-19)	nof_negCharge (p=1.31E-13)	nof_adb (p=1.52E-11)	nof_ao (p=1.15E-10)	nof_hbam (p=1.08E-09)	SP3 Character (p=4.59E-07)	molPSA (p=8.03E-07)	nof_adv (p=1.64E-06)
OAT3 Vs OCT2	17	20	27	nof_negCharge (p=3.52E-17)	nof_posCharge (p=8.12E-14)	nof_ao (p=1.71E-11)	nof_adb (p=3.24E-11)	nof_hbam (p=1.27E-09)	Hydrogen Bond Acceptors (p=1.03E-06)	molPSA (p=1.32983E-06)	nof_adv (p=1.66E-06)
OAT1 Vs OAT3	1	2	18	nof_aqv (p=0.00257)	nof_posCharge (p=0.00586)	nof_asv (p=0.01339)	nof_asb (p=0.01364)	nof_rbc (p=0.01432)	molVolume (p=0.01828)	Fragment Complexity (p=0.01924)	Number of hydrogens (p=0.01964)
OCT1 Vs OCT2	6	12	30	nof_s (p=0.00171)	Molecular weight (p=0.00382)	LabuteASA (p=0.00451)	molWeight (p=0.00465)	Number of heavy atoms (p=0.00495)	SMR (p=0.00497)	nof_hac (p=0.00514)	Vertex adjacency information magnitude (p=0.00527)

[a]The student's T tests calculated the p-values for each attribute for each pairwise transporter comparison, and the level of significance is indicated. The results are found to be consistent with the results from the machine-learning analyses.

JPET Fast Forward. Published on August 3, 2016 as DOI: 10.1124/jpet.116.232660
This article has not been copyedited and formatted. The final version may differ from this version.

TABLE 3

Quantitative APF Property Measurements

Transporter	Pharmacophore Model	Hydrogen Bond		Sp2 Hybridization	Lipophilic	Size (large)	Charge		Electro:		Overall space
		Donors	Acceptors				Positive	Negative	Positive	Negative	
OAT3	1	143.204	154.117	839.695	101.509	213.651	0	0	219.724	-117.089	47124
	2	68.1087	484.381	1560.94	628.674	773.567	0	-230.93	587.804	-289.52	82173
	3	7.92777	175.642	1009.03	493.385	636.479	5.5641	-193.954	529.816	-102.394	80496
	4	88.2388	189.257	1441.67	925.382	1058.15	0	-138.558	964.06	-140.411	98000
	5	35.8698	387.069	1795.93	923.052	1148.9	0	-197.94	987.326	-151.581	120744
	6	220.122	169.155	894.646	580.978	656.921	0	-38.4882	588.515	-325.512	68894
	7	197.147	244.087	881.089	588.426	752.08	57.7324	-173.197	773.25	-176.682	66924
	8	123.888	427.116	1445.62	664.117	934.116	0	-181.445	686.311	-246.445	116550
	[a]9	66.121	252.657	850.408	1164.04	1388.78	73.3208	-175.956	1331.36	-99.7383	149940
OCT1	1	48.1776	31.481	1056.59	570.64	704.827	102.635	0	823.36	-23.6594	115248
	2	90.1057	195.495	1020.17	716.752	937.245	46.1859	0	829.391	-75.0204	92752
	3	89.3848	32.1592	845.243	850.424	951.403	110.846	0	997.719	-20.1164	134160
	4	74.3945	95.9579	670.303	748.842	882.211	115.465	0	979.179	-55.4599	93240
	5	80.1732	150.189	951.337	1241.55	1355.43	115.465	0	1323.47	-66.1869	110124
	6	77.3553	68.1684	211.241	896.307	962.966	69.2789	0	1106.73	-21.3618	80360
	7	216.391	229.452	556.192	339.69	495.407	17.6656	-56.1539	503.882	-61.267	62700
	8	193.59	217.192	1639.75	1511.09	1759.88	46.1859	0	1827.44	-80.0704	164883
	9	554.411	58.0202	1109.79	377.711	472.479	346.394	0	887.07	-48.0251	105408
OCT2	1	211.184	222.252	622.403	345.852	501.321	23.0927	-69.2787	498.707	-75.3364	115248
	2	68.8115	29.6212	918.67	778.429	899.004	98.145	0	946.223	-7.10932	92752
	3	101.393	47.913	287.96	825.327	881.816	79.382	0	992.614	-23.1909	134160
	4	82.6908	86.1924	667.67	738.494	867.966	128.294	0	977.612	-49.6171	93240
	5	133.952	214.602	1024.3	1186.25	1402.03	76.9765	0	1347.41	-64.4698	110124
	6	502	39.793	1078.9	362.868	453.107	269.418	0	826.522	-24.6827	80360

[a]OAT3 pharmacophore model 9 was found to have higher value of positive charge and electro-positive charge than the rest of OAT3 pharmacophore models. The APF property values of this model were also found to be comparable with several OCT1 models, such as OCT1 model 5 and 6.

JPET Fast Forward. Published on August 3, 2016 as DOI: 10.1124/jpet.116.232660
This article has not been copyedited and formatted. The final version may differ from this version.

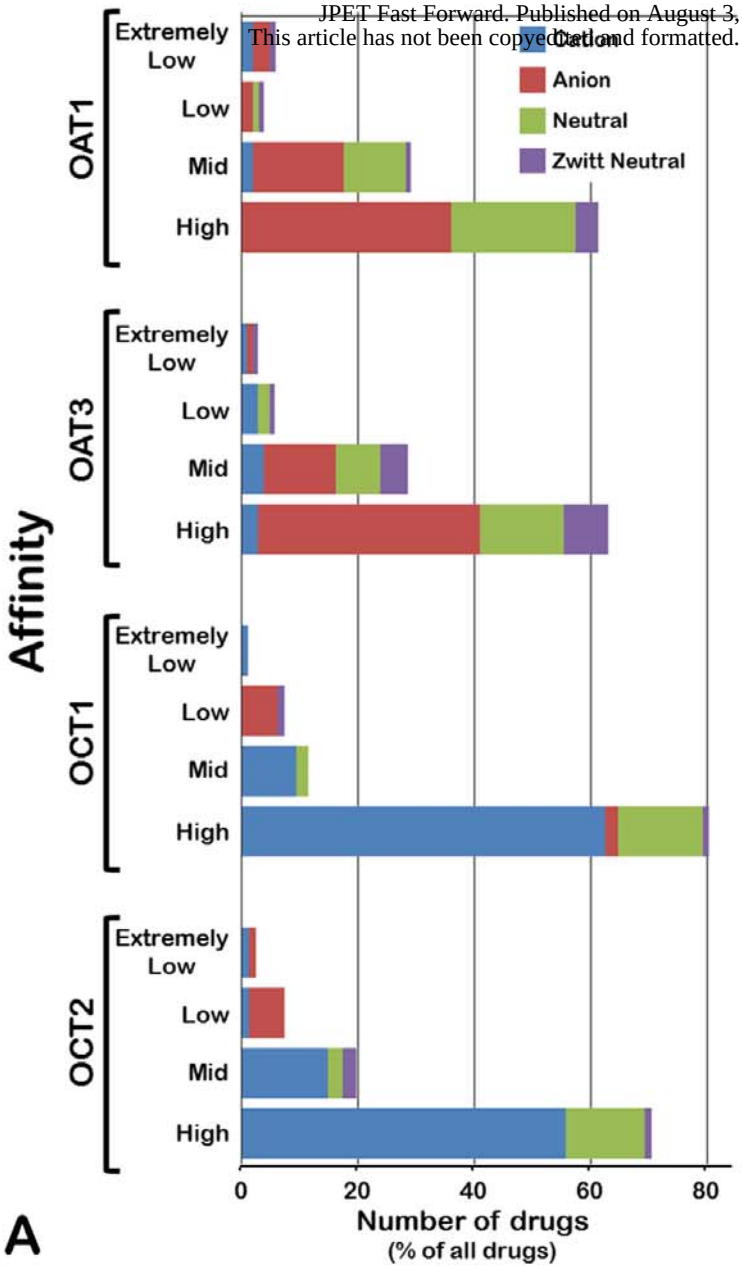
TABLE 4

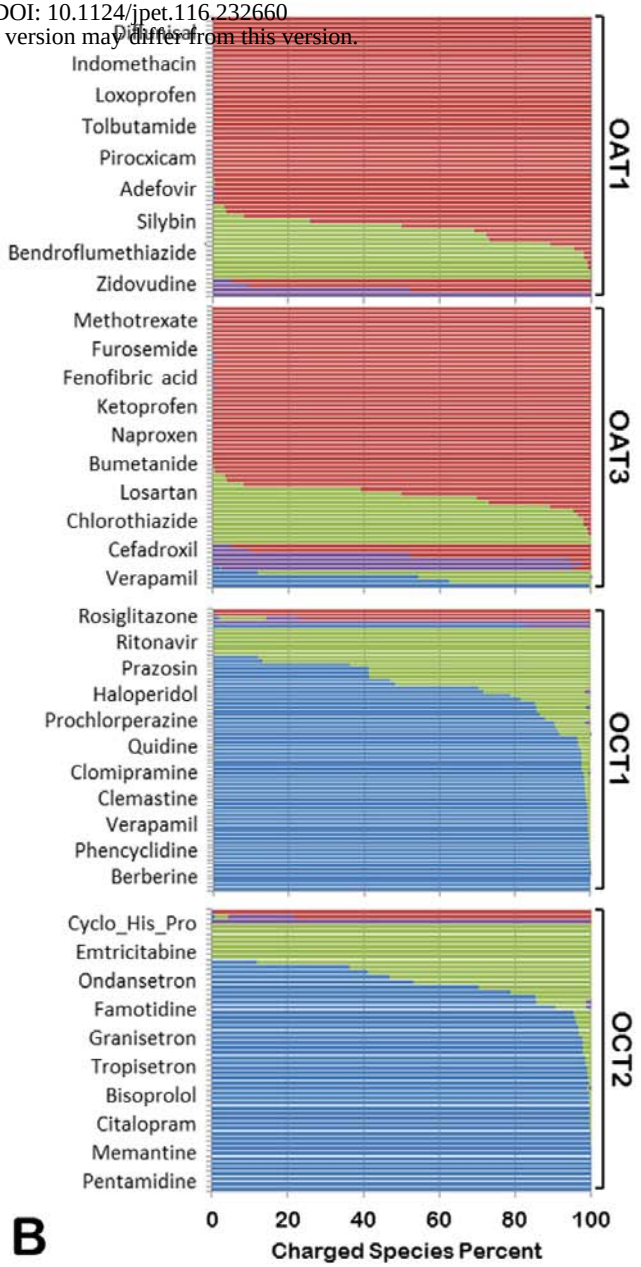
IC_{50} Values of Cationic Drugs Tested for Interaction with OAT1 and OAT3

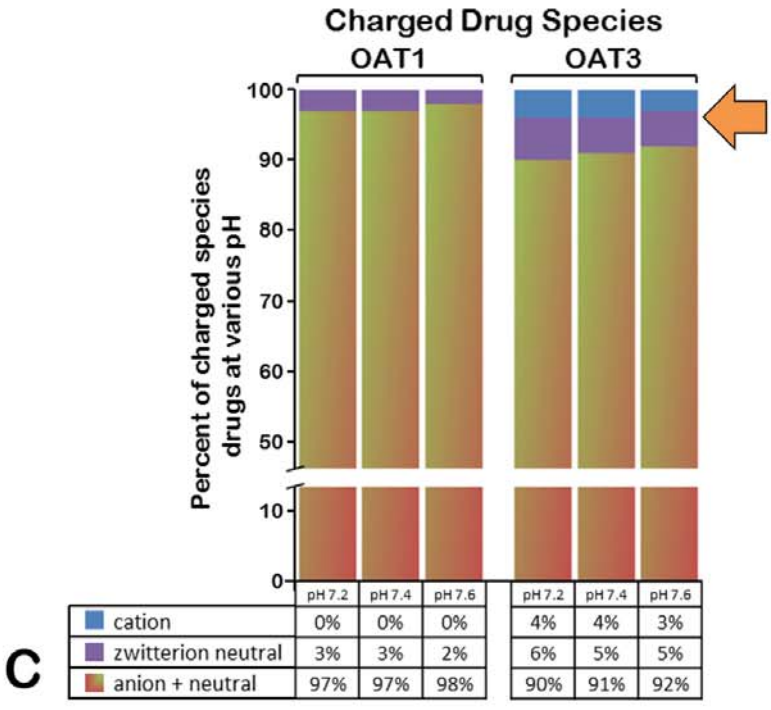
Drug Name	IC_{50} (uM)	
	OAT1	**OAT3**
[a]Probenecid	2.4	33
Darifenacin	807	198
Paliperidone	1082	260
Loperamide	no significant inhibition	95
Nebivolol	no significant inhibition	169
Halofantrine	No inhibition	No inhibition
Cisapride	No inhibition	No inhibition

[a]The data for probenecid uptake inhibition is shown as a control.

JPET Fast Forward. Published on August 3, 2016 as DOI: 10.1124/jpet.116.232660







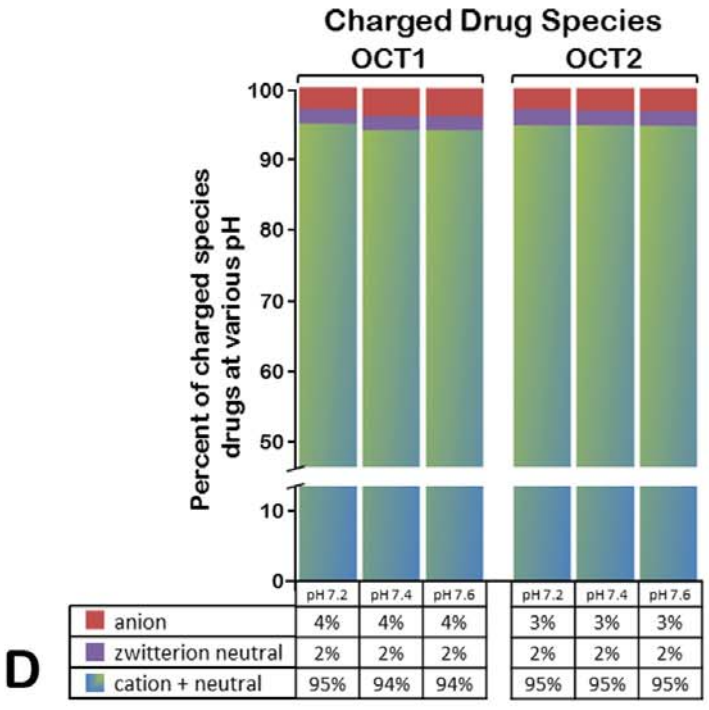


Figure 1

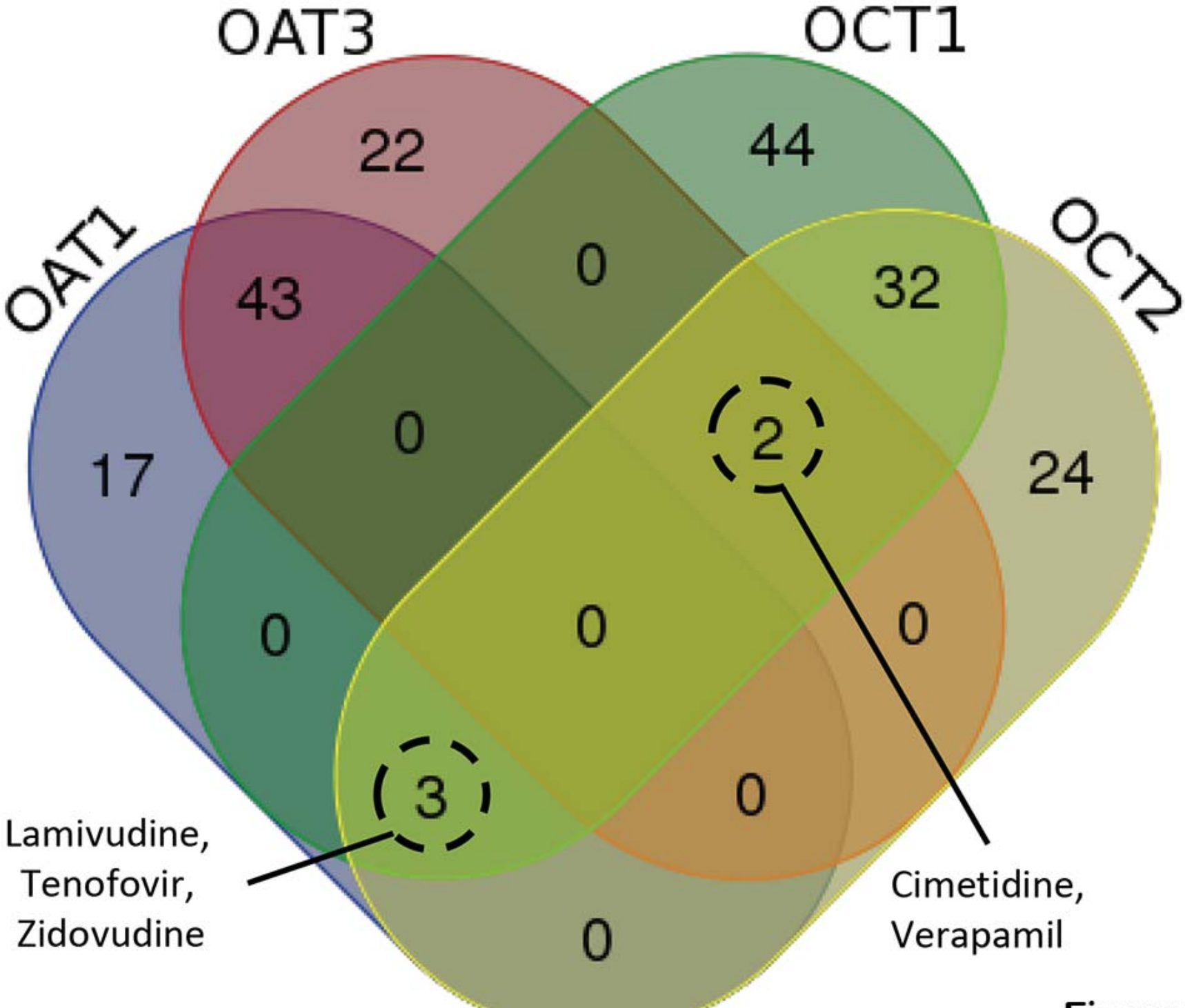


Figure 2

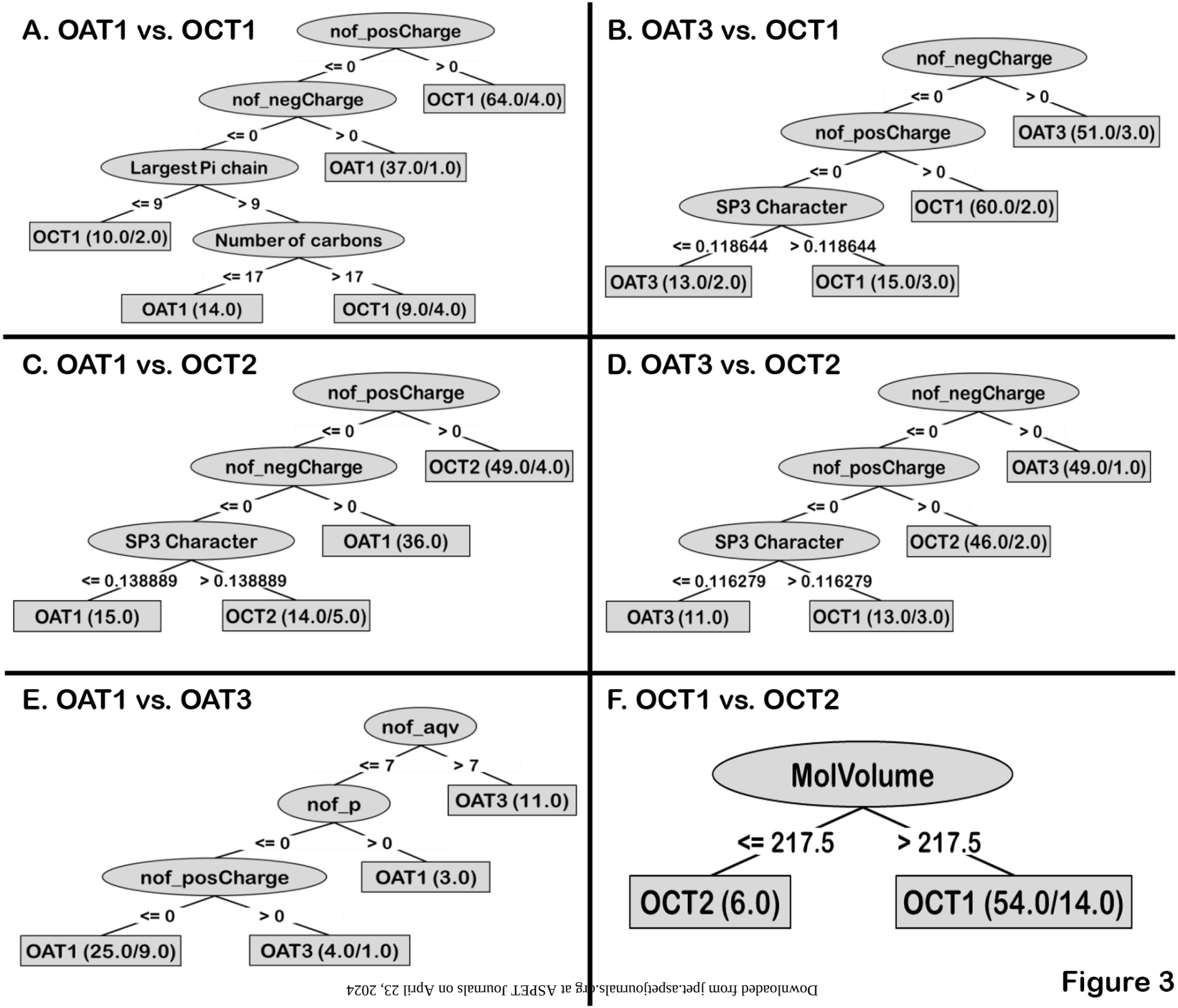


Figure 3

JPET Fast Forward. Published on August 3, 2016 as DOI: 10.1124/jpet.116.232660
This article has not been copyedited and formatted. The final version may differ from this version.

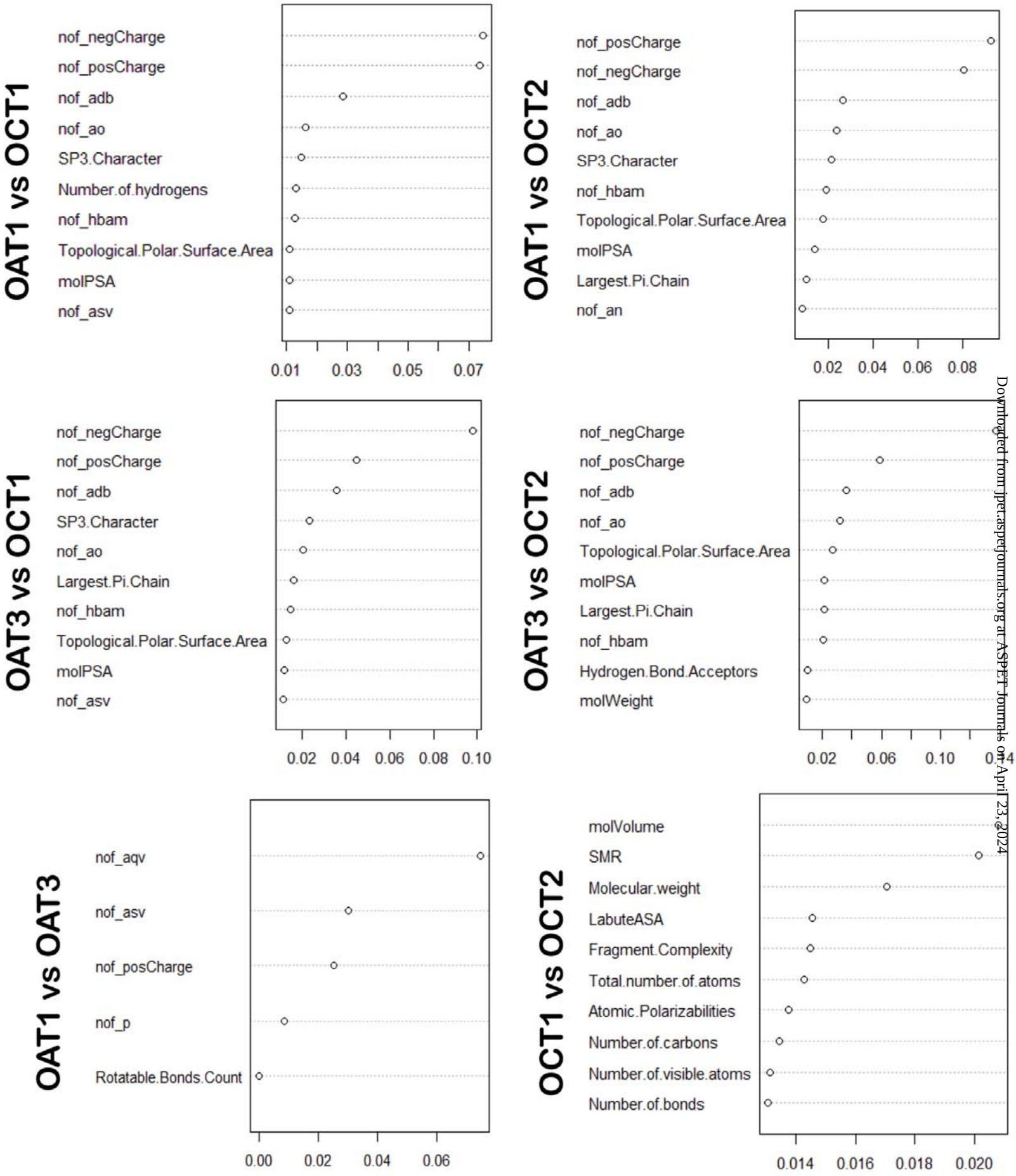
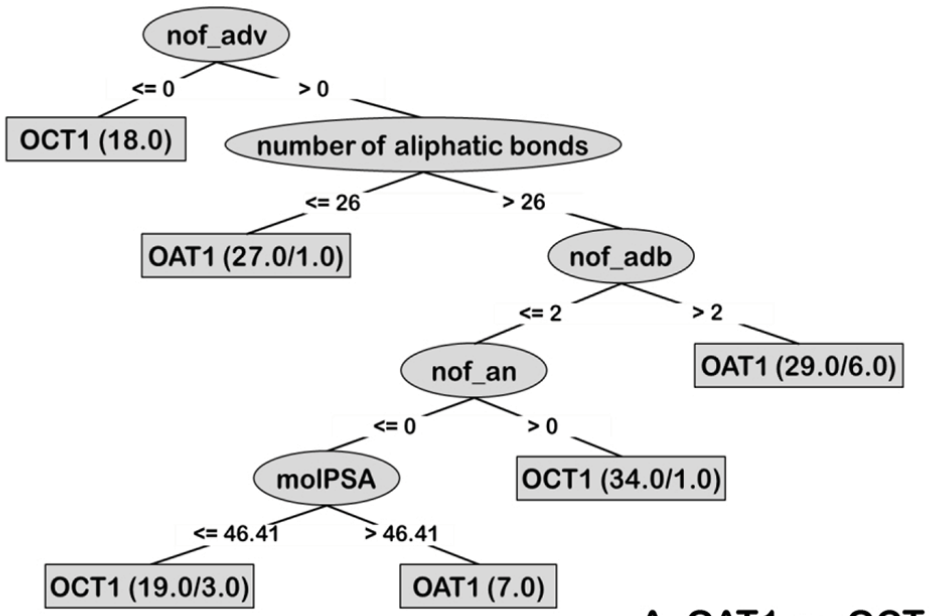
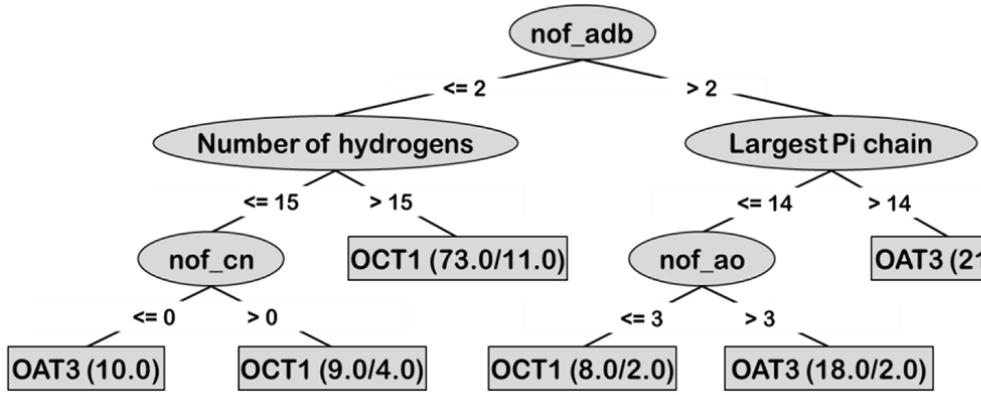


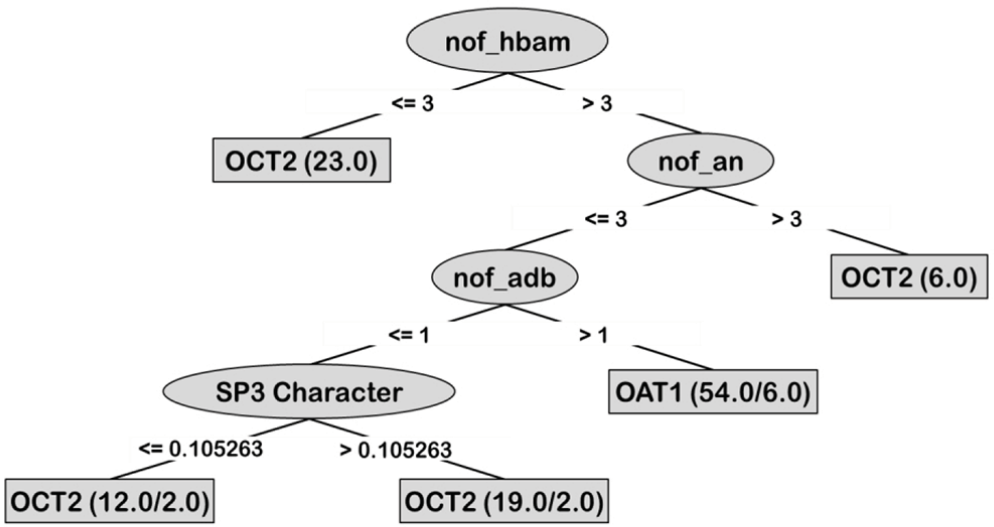
Figure 4



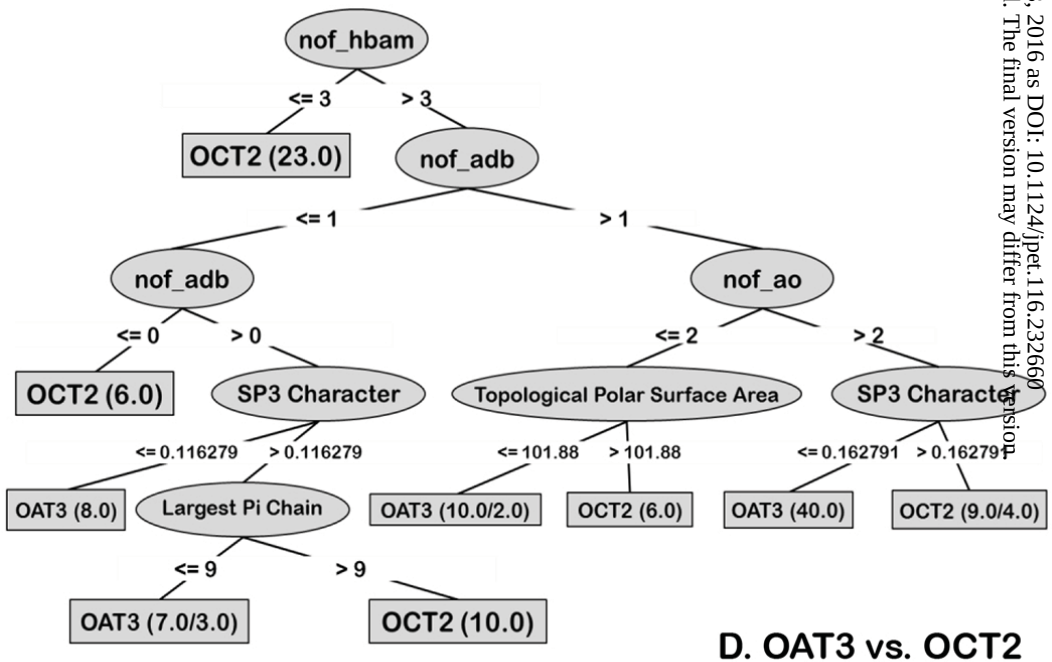
A. OAT1 vs. OCT1



B. OAT3 vs. OCT1



C. OAT1 vs. OCT2



D. OAT3 vs. OCT2

Figure 5

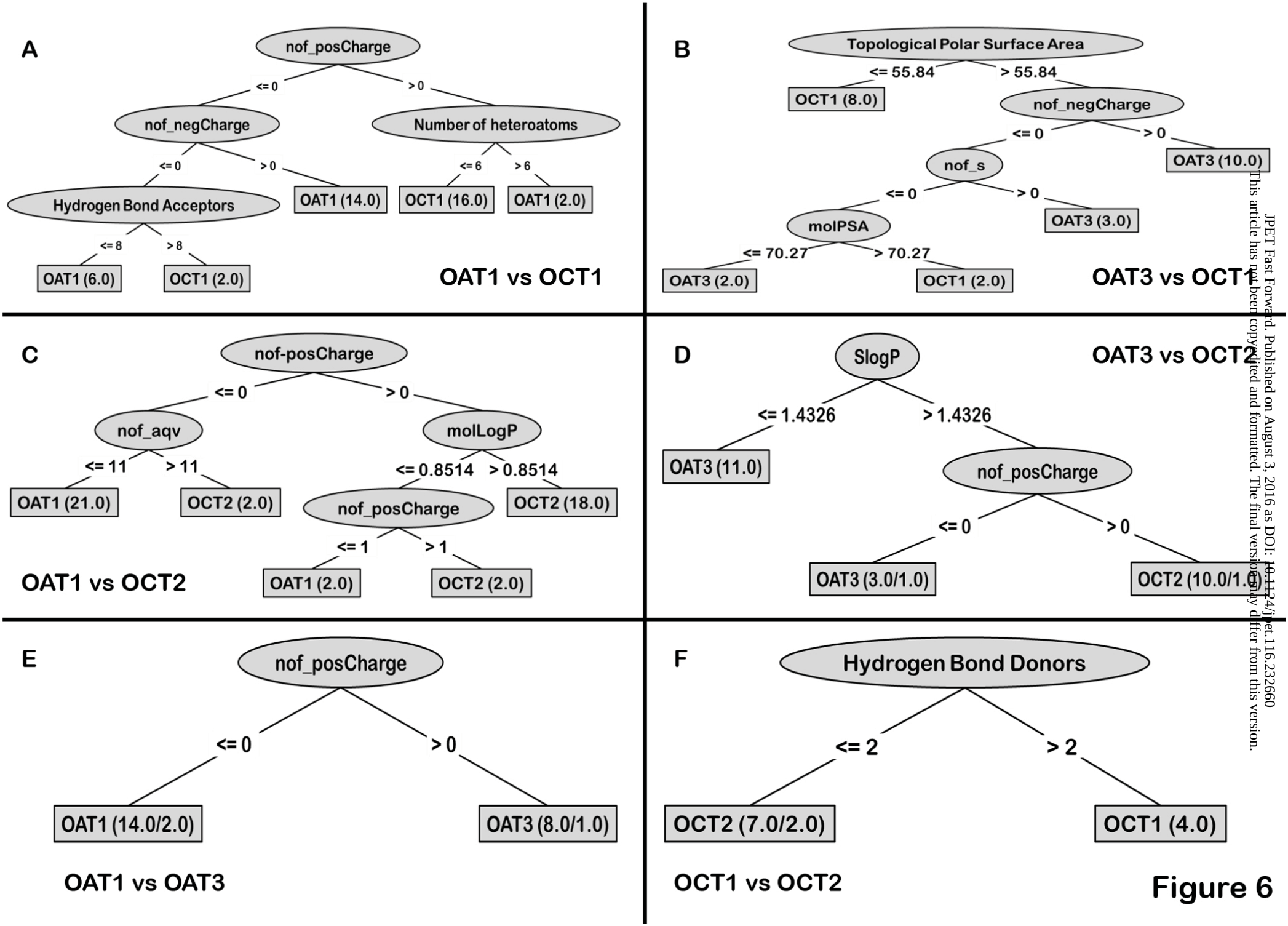


Figure 6

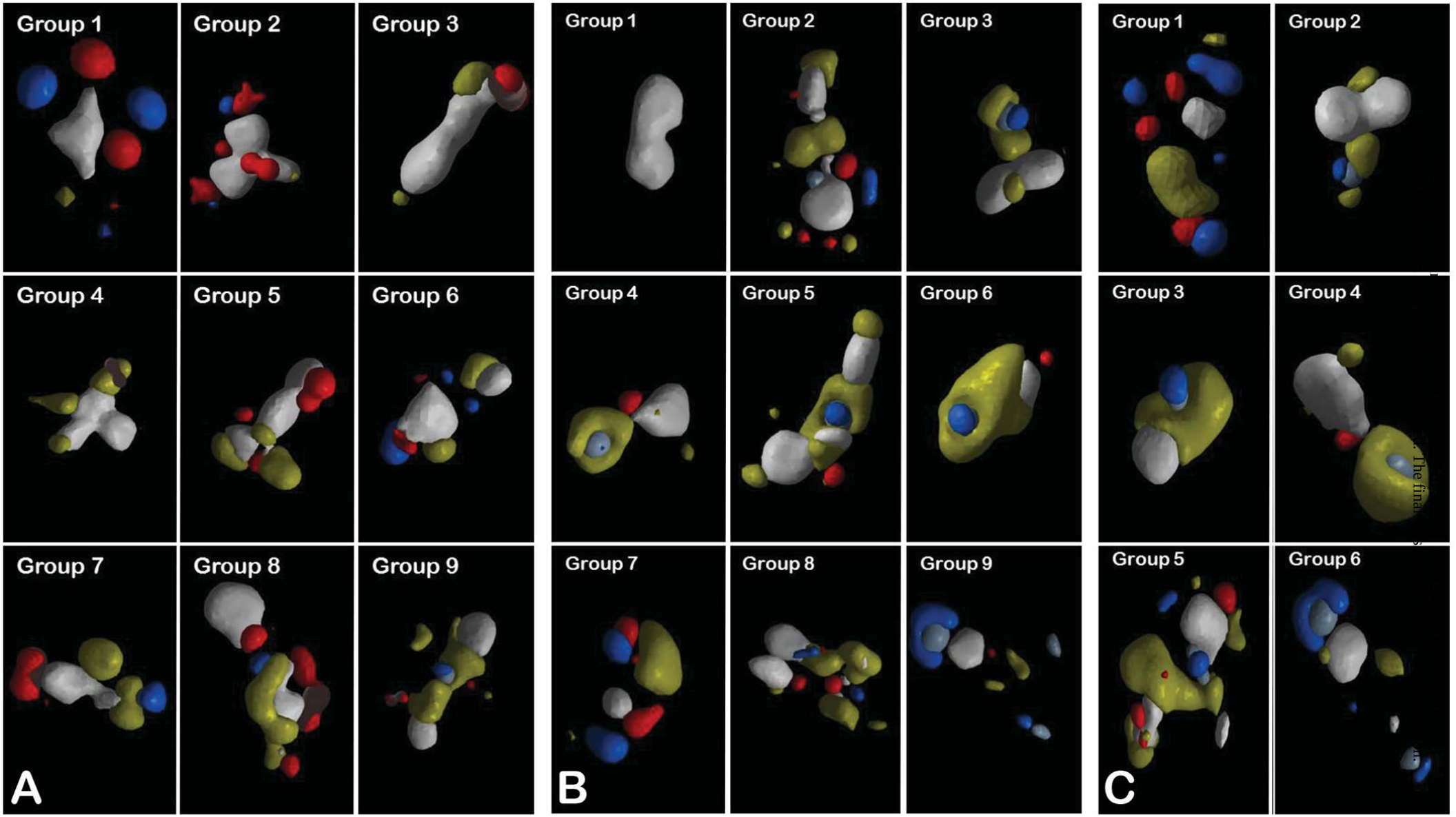


Figure 7

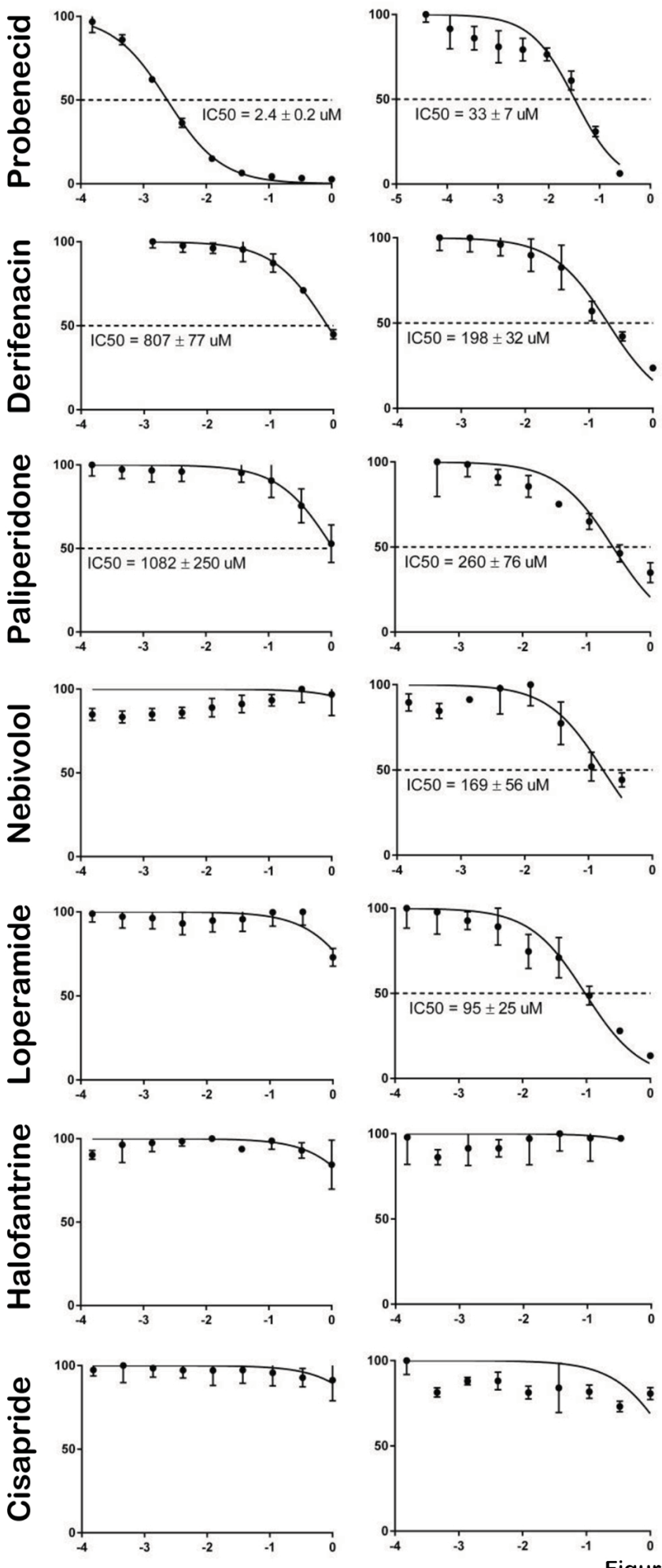
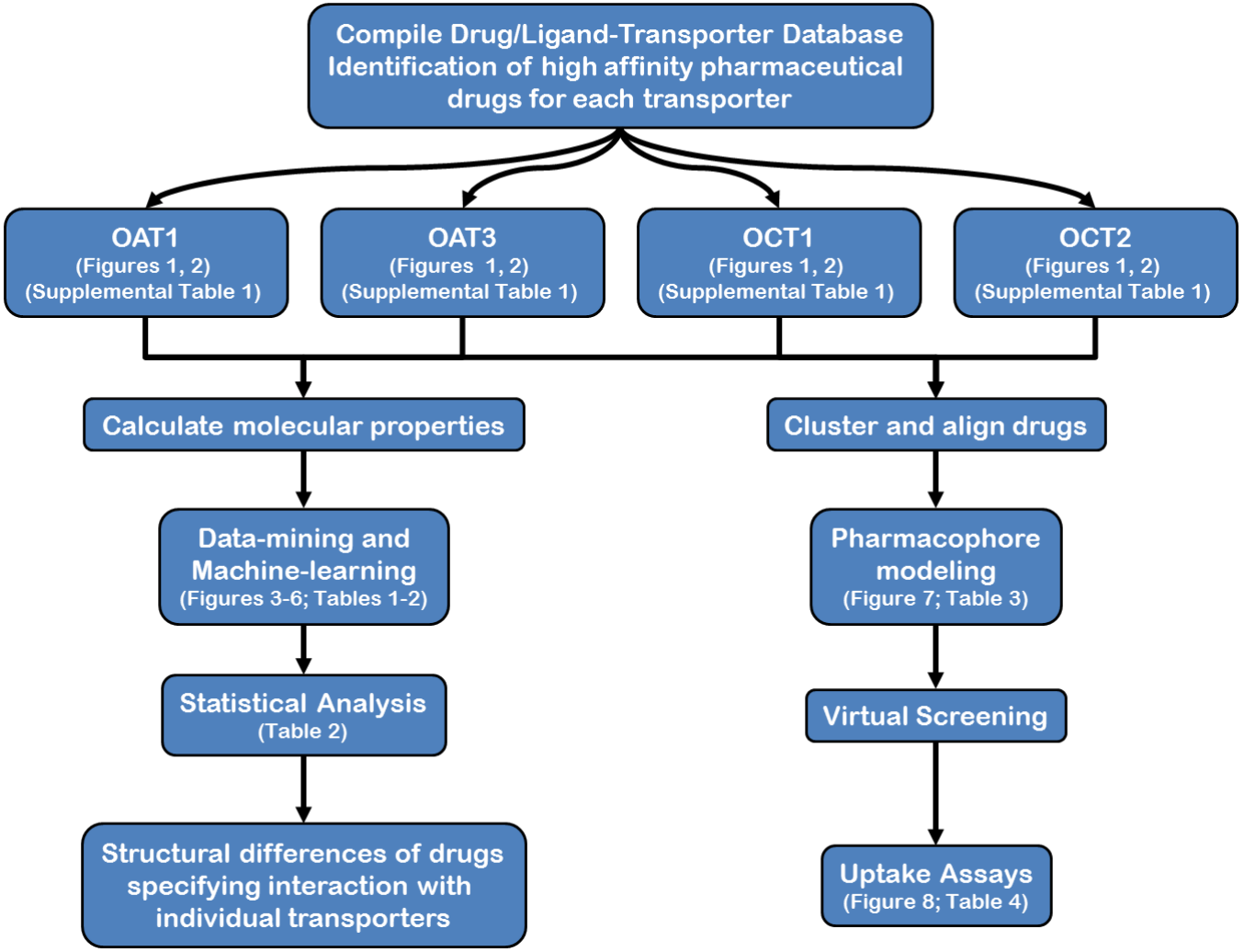


Figure 8

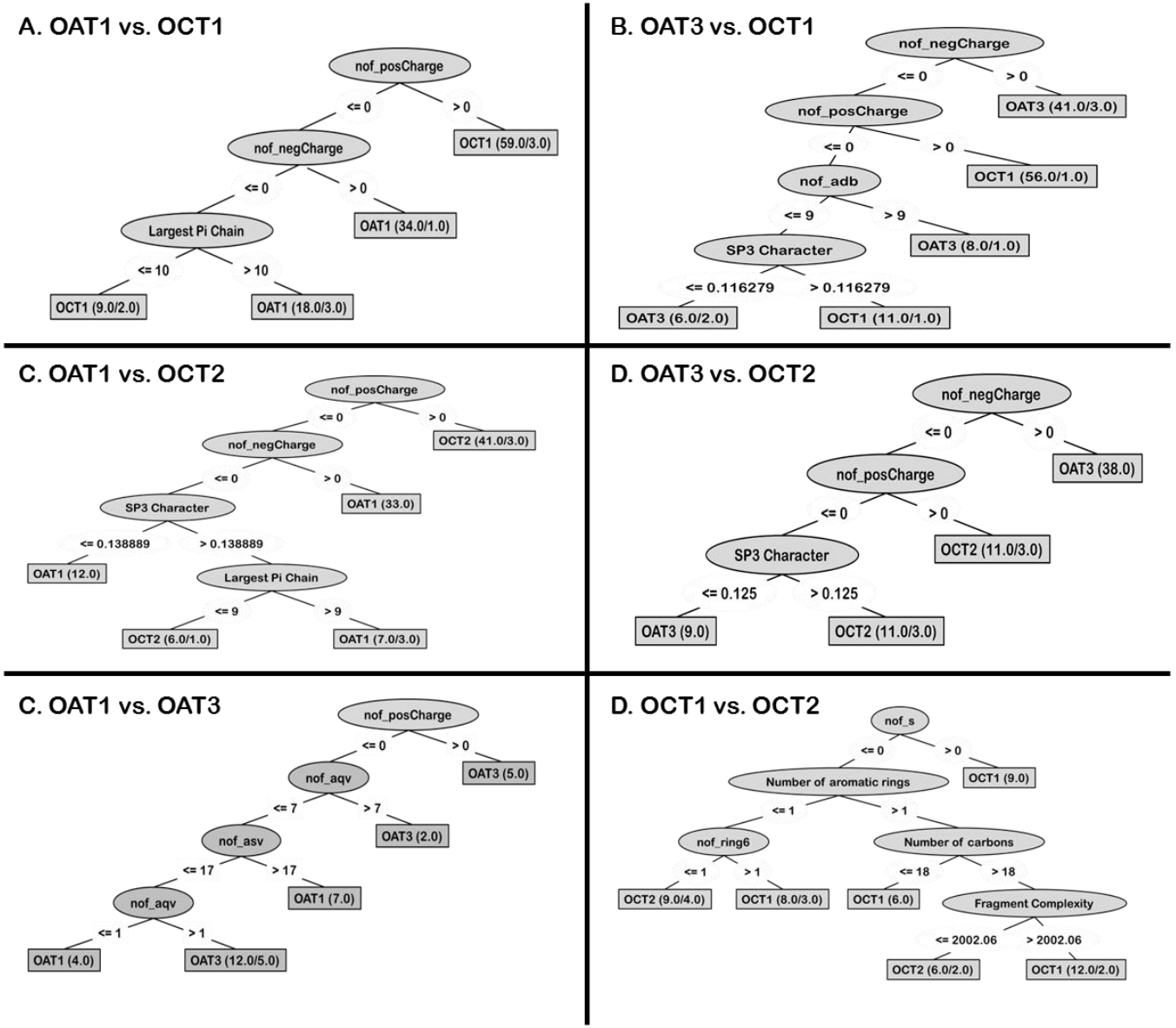
Analysis of Molecular Properties of Drugs Interacting with SLC22 Transporters OAT1, OAT3, OCT1, and OCT2: A Machine-Learning Approach

Henry C. Liu, Anne Goldenberg, Yuchen Chen, Christina Lun, Wei Wu, Kevin T. Bush, Natasha Balac, Paul Rodriguez, Ruben Abagyan, Sanjay K. Nigam

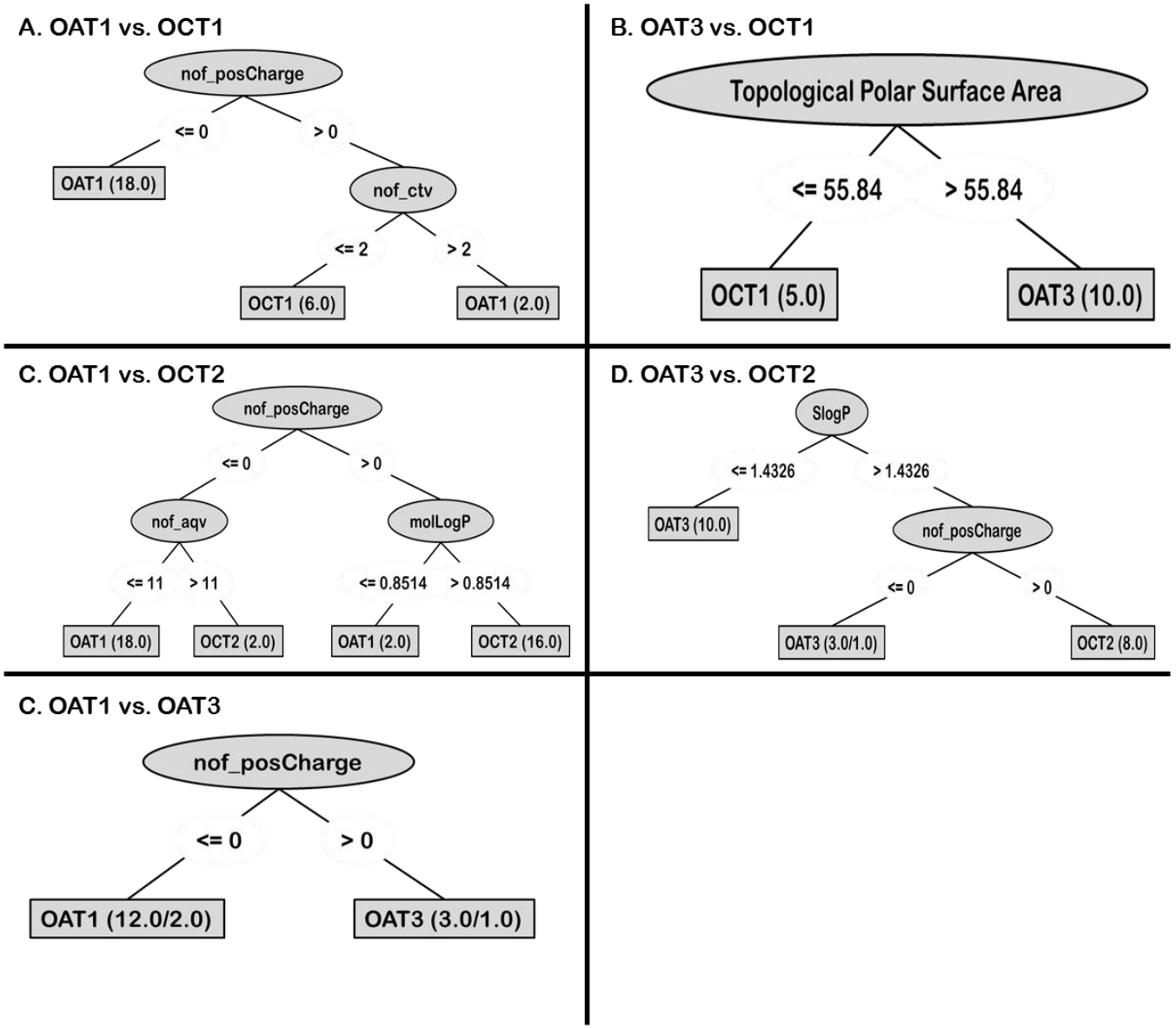
The Journal of Pharmacology and Experimental Therapeutics



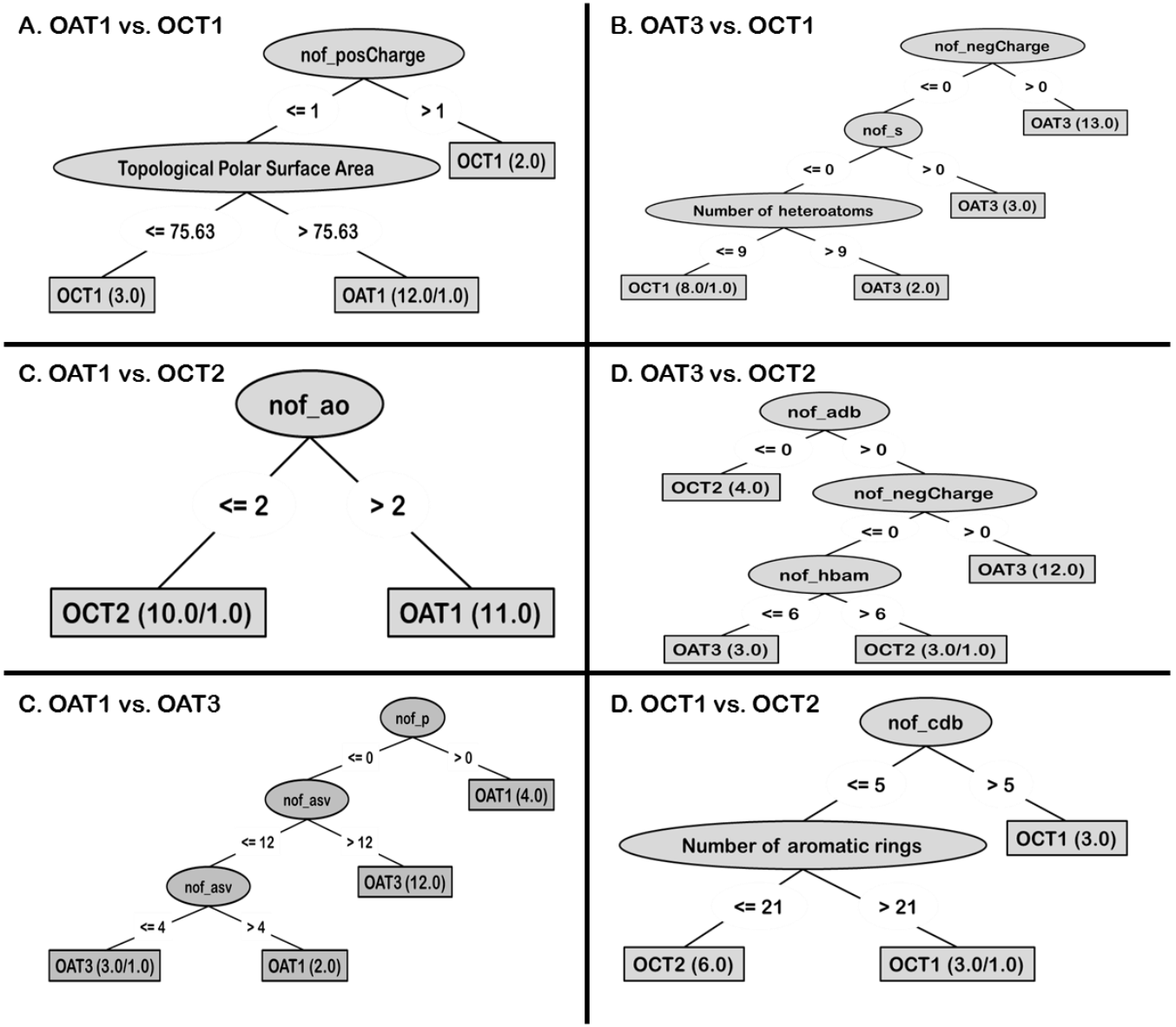
Supplemental Figure 1. A schematic of the overall strategy applied in this study. A comprehensive literature search was done to identify all the pharmaceutical drugs and tracers known to interact with any of the four transporters. Among these pharmaceutical entities, ones that have high-affinity interaction (<100 μM) were selected for further data-mining analysis and pharmacophore modeling.



Supplemental Figure 2. Decision trees based on drugs for which only a Ki value (high affinity; ≤ 100 μM) is available. The decision trees are similar to those obtained using all high affinity drugs without considering Km or Ki (Fig. 3). The trees show that besides charge, the 3-dimensionality of the drug, indicated by SP3 character, is an important factor. In addition, some differences were found between two OATs, specifically in number of aqv, p, and posCharge.



Supplemental Figure 3. Decision trees based on those drugs for which only a Ki value (medium-affinity; >100 μM and ≤ 1000 μM) is available. In this case, probably due to the reduced number of drugs used in the analysis, the main consideration distinguishing the drugs appears to be charge. Also please note that a decision tree was not generated for OCT1 vs OCT2 due to the limited number of drugs available for analysis.



Supplemental Figure 4. Decision trees based on those drugs for which Km values are available. In this case, the numbers of drugs does not appear to large enough to achieve clear results.

SUPPLEMENTAL TABLE 1
Kinetic Data for Drugs and Tracers Interacting with SLC22 Drug Transporters

	Drug Name	Km	Ki	PMID	Affinity
OAT1	6-Carboxyfluorescein	3.93		10929807	High Affinity
	6-Mercaptopurine	98		15287899	
	8-(Noradamantan-3-yl)-1,3-dipropylxanthine		7.82	11426832	
	Acyclovir	242		10945832	
	Adefovir	30		10462545	
	Bendroflumethiazide		0.9	18508962	
	Benzbromarone		2.8	12472777	
	Betamipron		23.6	11426832	
	Bumetanide		5.5	14610216	
	Candesartan		17	17674156	
	Cefadroxil		6.14	11909604	
	Cefamandole		0.03	11909604	
	Cefotaxime		3.13	11909604	
	Ceftriaxone		0.23	11909604	
	Chlorothiazide		3	14610216	
	Chlorpropamide		39.5	10854830	
	cHPMPC	309		10703662	
	Cidofovir	58		10703662	
	Cyclothiazide		67.4	14610216	
	Dibromosulfophthalein		2.74	11861777	
	Diclofenac		4.56	14722319	
	Didanosine		98.5	18174163	
	Diflunisal		0.64	10991954	
	Ellagic acid		0.138	15870380	
	Ethacrynate		23.7	14610216	
	Etodolac		38	10991954	
	Fluorescein		3	17553798	
	Flurbiprofen		1.13	10991954	
	Fluvastatin		21	14729100	
	Furosemide		9.5	10991988	
	Ganciclovir	895.5		11861798	
	Glibenclamide		1.6	10854830	
	Ibuprofen		4.33	14722319	
	Indomethacin		4.2	11099697	
	Ketoprofen		6.11	14722319	
	Lamivudine		17.1	18174163	
	Losartan		12	17674156	
	Loxoprofen		27.1	15548848	
	Mefenamic acid		0.66	12388633	
	Methotrexate	553.8		12130730	
	Mycophenolic acid		8.56	17462604	
	Naproxen		5.54	14722319	
	Nateglinide		9.2	10854830	
	Novobiocin		14.87	19282394	
	Olmesartan		0.28	17674156	
	Oxyphenbutazone		32	10220563	
	Para-aminohippurate	20.4		15846473	
	Phenylbutazone		47.9	14722319	
	Piroxicam		52	10220563	
	Pratosartan		1.5	17674156	
	Probenecid		6.4	17553798	
	Resorufin		4.7	18174163	
	Silybin		24	17920288	
	Simvastatin		58.9	14729100	
	Sulindac		99.9	14722319	
	Telmisartan		0.46	17674156	
	Temocaprilat	0.556		12660303	
	Tenofovir	33.8		11563082	
	Theophylline		53	16038872	
	Tolbutamide		55.5	10854830	
	Tolmetin		15.4	14722319	
	Torasemide		55.2	17978306	
	Trichloromethiazide		15.4	14610216	
	Valsartan		16	17674156	
	YM90K		5.8	15377641	

	Zidovudine	68		10945832	
	Zonampanel monohydrate	13.4		18443035	
	5-Carboxyfluorescein		523.4	18174163	Mid Affinity
	Acetaminophen		511	12388633	
	Acetylsalicylate		428	10220563	
	Benzoate		247	17094945	
	Buspirone		494	19737926	
	Caffeine		610	16038872	
	Carbenicillin		500	10411577	
	Cefaclor		876	16098483	
	Cefazolin		450	10411577	
	Cefdinir		553.4	16098483	
	Cefoperazone		132	15618660	
	Cefotiam		319	15618660	
	Ceftibuten		450	16098483	
	Cephalothin		530	12005172	
	Cephradine		793	10929807	
	Cimetidine		394	14978359	
	Digoxin		330	11408557	
	Hydrochlorothiazide		150	10991988	
	Melatonin		728	14737013	
	Methazolamide		350	14610216	
	Penicillin G		328	17553798	
	Phenacetin		488	10220563	
	Salicylate		145	17553798	
	Stavudine		103.1	18174163	
	Ticarcillin		530	17553798	
	Verapamil		447	19737926	
	Zalcitabine		242.9	18174163	
	Acetazolamide		1100	10991988	Low Affinity
	Cephaloridine		1320	12005172	
	Cilastatin		1470	11426832	
	N-acetyl-L-cysteine		2000	12237339	
	Pravastatin		1150	11861777	
	Cefoselis		2080	16098483	Very Low Affinity
	Ceftizoxime		2878	16098483	
	Cephalexin		2310	10411577	
	Histamine		3214	19737926	
	Nicotine		2365	19737926	
	Quinidine		4251	19737926	
	Taurocholate		2770	11408557	
OAT3	5-Fluorouracil	0.0539		15100168	High Affinity
	6-Carboxyfluorescein		10.7	18174163	
	6-Mercaptopurine	4.01		15100168	
	8-(Noradamantan-3-yl)-1,3-dipropylxanthine		3.7	11426832	
	Atorvastatin		6.8	17585018	
	Azathiopurine		15.7	15287899	
	Bendroflumethiazide		18.6	18508962	
	Betamipron		48.3	11426832	
	Bumetanide	1.01		16256982	
	Buspirone		34	19737926	
	Candesartan		0.3	17674156	
	Cefadroxil		8.62	11909604	
	Cefamandole		0.046	11909604	
	Ceftriaxone		4.39	11909604	
	Cephalothin		48	12005172	
	Chlorothiazide		64.7	14610216	
	Cimetidine	90.7		16006492	
	Ciprofloxacin	70		18381565	
	Cyclothiazide		27.6	14610216	
	Dibromosulfophthalein		3.09	11861777	
	Diclofenac		3.17	14722319	
	Enalapril		42.5	17314201	
	Ethacrynate		0.57	14610216	
	Etodolac		10	14722319	
	Famotidine	345	A	16006492	

	Fenofibric acid		2.2	17314201	
	Fexofenadine	70.2		16455804	
	Fluorescein		35	18174163	
	Fluvastatin		5.73	14729100	
	Furosemide		2.5	18508962	
	Gemfibrozil		3.4	17455113	
	Indapamide		10.9	17314201	
	Indomethacin		1.29	14722319	
	Ketoprofen		4.31	14722319	
	Losartan		1.6	17674156	
	Mefenamic acid		36	18789319	
	Meloxicam		81	18789319	
	Methazolamide		96.6	14610216	
	Methotrexate	60.6		17660957	
	Mycophenolic acid		1.5	17462604	
	Naproxen		19.1	14722319	
	Novobiocin		4.77	19282394	
	Olmesartan		0.027	17674156	
	Para-aminohippurate	64.7		10224140	
	Penicillin G	85.1		12660303	
	Phenacetin		19.2	12388633	
	Phenylbutazone		8.48	14722319	
	Piroxicam		4.19	14722319	
	Pranoprofen		88	18789319	
	Pratosartan		0.1	17674156	
	Pravastatin	13.4		11861777	
	Probenecid		4.43	11961115	
	Quinapril		6.1	17314201	
	Ranitidine	121		15319347	
	Resorufin		8.3	18174163	
	Rosuvastatin	7.4		17585018	
	Simvastatin		32	14729100	
	Sitagliptin	162		17314201	
	Sulindac		7.72	14722319	
	Telmisartan		1.6	17674156	
	Temocaprilat	1.41		12660303	
	Tetracycline	566.2		11855680	
	Tolmetin		8.78	14722319	
	Torasemide		89.9	17978306	
	Trichloromethiazide		70.5	14610216	
	Valacyclovir	57.9		11861798	
	Valsartan		0.2	17674156	
	Verapamil		31	19737926	
	YM90K		3.4	15377641	
	Zidovudine	145.1		11861798	
	Zonampanel monohydrate	53.6		18443035	
	5-Carboxyfluorescein		373	18174163	Mid Affinity
	6-Thioguanine		172	15287899	
	Acetaminophen		720	18789319	
	Acetazolamide		808	14610216	
	Adefovir		470	18174163	
	Cefaclor		119.7	16098483	
	Cefdinir		270.4	16098483	
	Cefoperazone		670	12005172	
	Cefotaxime		793	12005172	
	Cefotiam		760	18480179	
	Ceftibuten		246.4	16098483	
	Ceftizoxime		953.1	16098483	
	Cephalexin		624	12005172	
	Cilastatin		231	11426832	
	Didanosine		107	18174163	
	Gatifloxacin		941	18381565	
	Hydrochlorothiazide		933	14610216	
	Lamivudine		140	18174163	
	Loxoprofen		170	18789319	
	Norfloxacin		558	18381565	
	Ofloxacin		745	18381565	
	Salicylate		519	14722319	

	Tenofovir		302	18174163	
	Tiaromide		450	18789319	
	Zalcitabine		125.9	18174163	
	Acyclovir		1460	12660303	Low Affinity
	Cephaloridine		1140	12005172	
	Metoclopramide		1376	19737926	
	Nicotine		1713	19737926	
	Quinidine		1675	19737926	
	Stavudine		1664	18174163	
	Acetylsalicylate		5000	18789319	Very Low Affinity
	Cefoselis		2914	16098483	
	Histamine		10230	19737926	
	Procainamide		4873	19737926	
OCT1	Abacavir ABC		7.2 x10^-5	19141712	High Affinity
	Acebutolol		95.8	9655880	
	Acyclovir	151		11861798	
	Amantadine		38.1	19833842	
	Amiloride		39.9	18788725	
	Amitriptyline		11.9	18788725	
	Amsacrine		3.5	18788725	
	Apomorphine		14.7	18788725	
	Atropine		8.4	18788725	
	Berberine	14.8		18157518	
	Bucindolol		18.9	18788725	
	Butylscopolamine		16	16263091	
	Chlorpromazine		18.9	18788725	
	Chlorprotixen		54.5	18788725	
	Cimetidine		66	15389554	
	Citalopram		13.3	18788725	
	Clemastine		3.4	18788725	
	Clomipramine		13.3	18788725	
	Clonidine		16.1	18788725	
	Cocaine		85	20170649	
	Cyclo(His-Pro)	655		17460754	
	D-Amphetamine		202	20170649	
	Denopamine		32.9	18788725	
	Desipramine		39.9	18788725	
	Diltiazem		8.4	18788725	
	Diphenylhydramine		3.4	16263091	
	Disopyramide		57.3	18788725	
	Disprocynium 24		0.4	20170649	
	Dopamine		487	12606755	
	Emtricitabine FTC		2x10^-5	19141712	
	Etilefrine		447	16263091	
	Famotidine		27	16141367	
	Flecainide		42	18686197	
	Flupentixol		63	18788725	
	Fluphenazine		76.9	18788725	
	Furaminidine two positive charges	6.1	7.4	18971316	
	Ganciclovir	516		11861798	
	Haloperidol		99.3	18788725	
	Imipramine		11.9	18788725	
	Indinavir		37	18490433	
	Ketamine		115	20170649	
	Lamivudine		17	18490433	
	Loperamide		16.8	18788725	
	Memantine		18.9	18788725	
	Mepenzolate		45.5	18788725	
	Metoclopramide		66.4	18788725	
	Metoprolol		268	18686197	
	Midazolam		3.6	9655880	
	Morphine		19.6	18788725	
	Nandrolone		24.3	18788725	
	Nelfinavir		7	18490433	
	Nicotine		185.2	15817714	
	Ondansetron		14	18788725	

	Orphenadrine		9	18788725	
	Oxprenolol		20.3	18788725	
	Pentamidine two positive charges	36	16	18971316	
	Phencyclidine		4.4	20170649	
	Phenoxybenzamine		10.5	18788725	
	Pindolol		39	18686197	
	Prazosin		1.5	19833842	
	Procainamide		48.6	19833842	
	Prochlorperazine		35	18788725	
	Promethazine		24.5	18788725	
	Propafenone		7.7	18788725	
	Propranolol		44.1	18788725	
	Pyrimethamine		3.4	20065018	
	Quinidine		79.7	18788725	
	Quinine		36.4	18788725	
	Ranitidine		28	16263091	
	Repaglinide		6.4	18788725	
	Ritonavir		14	18490433	
	Rosiglitazone		15	18314419	
	Salsolinol		440	17460754	
	Saquinavir		37	18490433	
	Tenofovir TDF		8.5 x10^-4	19141712	
	Terazosine		16.8	18788725	
	Tetraethylammonium (TEA)	229		16330770	
	Thiamine		434	12606755	
	Tramadol		37.1	18788725	
	Trimethoprim		39.9	18788725	
	Trimipramine		19.6	18788725	
	Vecuronium		123.2	9655880	
	Verapamil		1.14	19833842	
	YM155	22		19833842	
	Zalcitabine		24	18490433	
	Zidovudine		1.6x10^-4	19141712	
	Diclofenac		<1956	12388633	
	Ibuprofen		<1956	12388633	
	Indomethacin		<1956	12388633	
	Ketoprofen		<1956	12388633	Low Affinity
	Mefenamic acid		<1956	12388633	
	N1-Methyl-nicotinamide		1035	12606755	
	Piroxicam		<1956	12388633	
	Sulindac		<1956	12388633	
	Aminoguanidine		<9784.7	19426682	
	Histamine		3007	12606755	Very Low Affinity
	Metformin	2160		19591196	
OCT2	Abacavir		0.000041	19141712	
	Amantadine		27.9	20170649	
	Amiloride	95		16394027	
	Amitriptyline		6.1	19002438	
	Atropine		29	16263091	
	Beclomethasone		4.3	15817714	
	Berberine	4.4		18157518	
	Bisoprolol		2.4	19740083	
	Budesonide		7.1	15817714	
	Carvedilol		63.4	19002438	
	Chlorpromazine		6.1	19002438	
	Cimetidine	73		16006492	High Affinity
	Clonidine		7	19002438	
	Cyclo(His-Pro)	74		17460754	
	D-Amphetamine		11	20170649	
	Desipramine		16	9260930	
	Desloratidine		26	19002438	
	Diphenylhydramine		14.99	16263091	
	Disprocynium 24		0.28	20170649	
	Dopamine	390		9687576	
	Doxepin		5.7	19002438	
	Emtricitabine		0.0024	19141712	

	Epinephrine	420		20170649	
	Famotidine	56		16006492	
	Fenfluramine		4.4	19002438	
	Flecainide		<83.8	19002438	
	Flurazepam		26.3	19002438	
	Histamine	940		20170649	
	Imipramine		2.6	19002438	
	Ipratropium		6.6	19002438	
	Ketamine		22.9	20170649	
	Lamivudine		32.6	18490433	
	Memantine	34		9687576	
	Metformin	990		16272756	
	Metoprolol		49.5	19740083	
	Mexiletine		24	19002438	
	Nafamostat		20	14705184	
	Nelfinavir		12.8	18490433	
	Nicotine		41	15817714	
	Pentamidine two positive charges		3.8	18490433	
	Phencyclidine		25	20170649	
	Phenoxybenzamine		4.9	12110607	
	Prazosin		58.4	19833842	
	Procainamide		67	19833842	
	Propafenone		10.9	19002438	
	Pyrimethamine		6.1	20065018	
	Quinidine		38.2	19002438	
	Quinine		23	16263091	
	Ranitidine	65		16006492	
	Ritonavir		24.7	18490433	
	Sibutramine		12.7	19002438	
	Tamoxifen		38.2	19002438	
	Tenofovir		0.00057	19141712	
	Tetraethylammonium	46		15630081	
	Trimethoprim		21	15212162	
	Varenicline	370		17971819	
	Verapamil		37.3	19002438	
	Zidovudine		0.00027	19141712	
	Butylscopolamine		762	16263091	
	Cocaine		112.7	20170649	
	Disopyramide		142	19002438	
	Furamidine two positive charges		161	18971316	
	Indinavir		271	18490433	
	Levofloxacin		127	17072098	
	Lidocaine		291	18686197	Mid Affinity
	Mefloquine		201.4	19002438	
	N1-Methyl-nicotinamide		248.4	15630081	
	Propranolol		100.4	19002438	
	Saquinavir		202.4	18490433	
	Sulindac		<901.7	12388633	
	Zalcitabine		129.3	18490433	
	Ibuprofen		1803-4509	12388633	
	Mefenamic acid		<1803	12388633	
	Chloroquine		1081.9	19002438	
	Diclofenac		<1802	12388633	Low Affinity
	Indomethacin		<1803	12388633	
	Ketoprofen		<1803	12388633	
	Norepinephrine	1500		20170649	
	Piroxicam		<1803	12388633	
	Aminoguanidine	4100		19426682	Very Low Affinity
	Etilefrine		3998	16263091	

SUPPLEMENTAL TABLE 2
[a]Machine-learning Models: OAT1 vs OCT1
The validation performance of various machine-learning tools
used for the comparison of OAT1/OCT1 high-affinity drugs.

Model	Correctly Classified	ROC Area
Decision Tree	86.57%	0.905
Support Vector Machine	93.28%	0.934
Neural Network – Multilayer Perception	88.06%	0.945
Naïve Bayes	88.81%	0.933
Decision Rule	86.57%	0.918

[a]Models based on different machine-learning tools generated well-validated results. However, since decision trees provide the most logical way to analyze and discuss how physicochemical properties affect substrate preferences between transporters, we focused on decision trees.

Supplemental Table 3
Performance validation of various decision tree analyses
For separate analyses of drugs with Ki and with Km

	Drugs With Ki				Drugs With Km	
	High Affinity		Medium Affinity			
Transporters Compared	**Correctly classified**	**ROC area**	**Correctly classified**	**ROC area**	**Correctly classified**	**ROC area**
OAT1/OCT1	88.33	0.893	88.46	0.958	41.18	0.371
OAT1/OCT2	86.87	0.937	86.84	0.933	95.24	0.921
OAT3/OCT1	85.25	0.896	80	0.825	61.54	0.586
OAT3/OCT2	91.75	0.973	66.67	0.671	95.24	0.921
OAT1/OAT3	70	0.684	66.67	0.364	71.43	0.781
OCT1/OCT2	62	0.52	ND	ND	58.33	0.529