

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

---

Food and Drug Administration Papers

U.S. Department of Health and Human Services

---

2020

## Development of a Nicotinic Acetylcholine Receptor nAChR $\alpha 7$ Binding Activity Prediction Model

Sugunadevi Sakkiah

*U.S. Food and Drug Administration, Jefferson*

Carmine Leggett

*U.S. Food and Drug Administration, Calverton, carmine.leggett@gmail.com*

Bohu Pan

*U.S. Food and Drug Administration, Jefferson*

Wenjing Guo

*U.S. Food and Drug Administration, Jefferson*

Luis G. Valerio Jr.

*U.S. Food and Drug Administration, Calverton*

*See next page for additional authors*

Follow this and additional works at: <https://digitalcommons.unl.edu/usfda>



Part of the [Dietetics and Clinical Nutrition Commons](#), [Health and Medical Administration Commons](#), [Health Services Administration Commons](#), [Pharmaceutical Preparations Commons](#), and the [Pharmacy Administration, Policy and Regulation Commons](#)

---

Sakkiah, Sugunadevi; Leggett, Carmine; Pan, Bohu; Guo, Wenjing; Valerio, Luis G. Jr.; and Hong, Huixiao, "Development of a Nicotinic Acetylcholine Receptor nAChR  $\alpha 7$  Binding Activity Prediction Model" (2020). *Food and Drug Administration Papers*. 56. <https://digitalcommons.unl.edu/usfda/56>

This Article is brought to you for free and open access by the U.S. Department of Health and Human Services at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Food and Drug Administration Papers by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

---

**Authors**

Sugunadevi Sakkiah, Carmine Leggett, Bohu Pan, Wenjing Guo, Luis G. Valerio Jr., and Huixiao Hong

# Development of a Nicotinic Acetylcholine Receptor nAChR $\alpha 7$ Binding Activity Prediction Model

Sugunadevi Sakkiah, Carmine Leggett,\* Bohu Pan, Wenjing Guo, Luis G. Valerio, Jr., and Huixiao Hong\*



Cite This: *J. Chem. Inf. Model.* 2020, 60, 2396–2404



Read Online

ACCESS |



Metrics & More

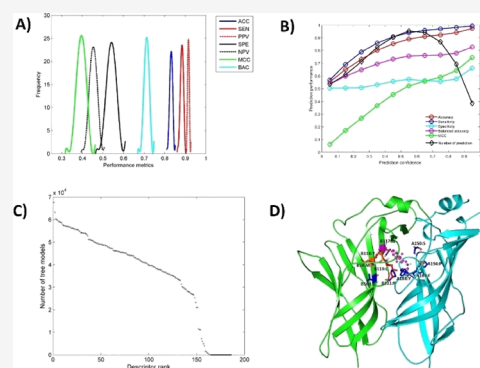


Article Recommendations



Supporting Information

**ABSTRACT:** Despite the well-known adverse health effects associated with tobacco use, addiction to nicotine found in tobacco products causes difficulty in quitting among users. Nicotinic acetylcholine receptors (nAChRs) are the physiological targets of nicotine and facilitate addiction to tobacco products. The nAChR- $\alpha 7$  subtype plays an important role in addiction; therefore, predicting the binding activity of tobacco constituents to nAChR- $\alpha 7$  is an important component for assessing addictive potential of tobacco constituents. We developed an  $\alpha 7$  binding activity prediction model based on a large training data set of 843 chemicals with human  $\alpha 7$  binding activity data extracted from PubChem and ChEMBL. The model was tested using 1215 chemicals with rat  $\alpha 7$  binding activity data from the same databases. Based on the competitive docking results, the docking scores were partitioned to the key residues that play important roles in the receptor–ligand binding. A decision forest was used to train the human  $\alpha 7$  binding activity prediction model based on the partition of docking scores. Five-fold cross validations were conducted to estimate the performance of the decision forest models. The developed model was used to predict the potential human  $\alpha 7$  binding activity for 5275 tobacco constituents. The human  $\alpha 7$  binding activity data for 84 of the 5275 tobacco constituents were experimentally measured to confirm and empirically validate the prediction results. The prediction accuracy, sensitivity, and specificity were 64.3, 40.0, and 81.6%, respectively. The developed prediction model of human  $\alpha 7$  may be a useful tool for high-throughput screening of potential addictive tobacco constituents.



## INTRODUCTION

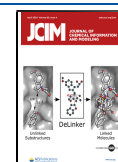
It has been established that the use of tobacco products is harmful and causes deleterious health effects such as respiratory disease, heart disease, multiple cancers, and many other adverse effects on organ systems as detailed in the U.S. Surgeon General Report.<sup>1,2</sup> It is noted that tobacco use is the most preventable cause of premature death in the United States. However, 14.0% of all adults (34.3 million people: 15.8% of men and 12.2% of women) were current cigarette smokers in 2017 ([https://www.cdc.gov/vitalsigns/youth-tobacco-use/?s\\_cid=osh-stu-home-slider-005](https://www.cdc.gov/vitalsigns/youth-tobacco-use/?s_cid=osh-stu-home-slider-005)). Also concerning is that according to the CDC, in 2018, there were 4.9 million youth that used tobacco products, and in 2017–2018, the use of any tobacco products grew by 38.3% among high school students, and this increase can be attributed largely due to the use of electronic cigarettes. These striking statistics underscore the importance of better understanding nicotine dependence; particularly how nicotinic acetylcholine receptors (nAChRs), the physiological targets of nicotine, facilitate addiction to tobacco products.

Nicotine dependence and reinforcement are mediated by a family of pentameric ligand-gated ion channels, the nicotinic acetylcholine receptors.<sup>3</sup> Generally, the pentameric complex of

nAChRs consists of an extracellular domain or ligand binding domain (LBD), a transmembrane region with four helices, and an intracellular domain.<sup>4</sup> The rodent neuronal and muscle-type nAChR are the two-major types in the nAChR family. There are various subunits reported for the neuronal nAChRs such as  $\alpha 2$ – $\alpha 10$  and  $\beta 2$ – $\beta 4$ . Among these subunits,  $\alpha 7$ – $\alpha 10$  are able to form a homo-pentamer complex and the remaining subunit  $\alpha 2$ – $\alpha 6$  forms heteromeric complexes by combining with  $\beta 2$ – $\beta 4$ .<sup>4–7</sup> The  $\alpha 7$  pentameric complex is highly present in the human brain and plays an important role in addiction. Hence, predicting the binding activity of tobacco constituents to nAChR  $\alpha 7$  is important for understanding the addictive potential of tobacco constituents.<sup>8</sup> More than 8000 chemicals are identified in tobacco smoke.<sup>9</sup> However, only a few tobacco constituents have been tested for nAChR  $\alpha 7$  binding activity. Experimental evaluation of the remaining tobacco constituents

Received: February 7, 2020

Published: March 11, 2020

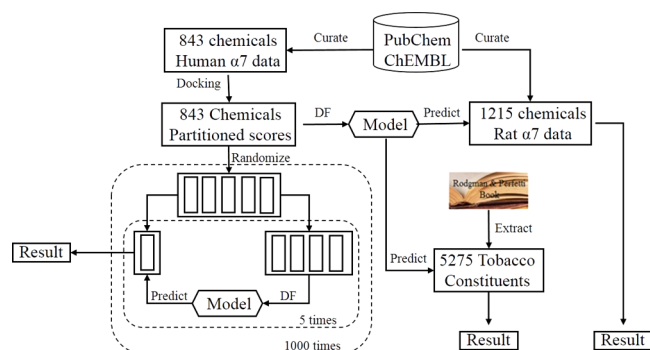


is costly and time consuming. Well-established *in silico* techniques such as molecular docking<sup>9–15</sup> and machine learning<sup>16–21</sup> have been used for developing prediction models of various biological activities. We previously applied molecular docking and molecular dynamic simulations to reveal the important residues of  $\alpha 7$  nAChR that forms interactions with tobacco constituents.<sup>22</sup> In this study, we combined molecular docking and machine learning to develop an  $\alpha 7$  binding activity prediction model based on a training data set of 930 chemicals with known human  $\alpha 7$  binding activity data. Based on the competitive docking results, the docking scores were then partitioned to the key residues that play important roles in the receptor–ligand binding. The developed model was used to predict the potential human  $\alpha 7$  binding activity for 5275 tobacco constituents.

This study presents a comprehensive assessment of tobacco constituents that may exhibit addiction potential by binding to the nAChR  $\alpha 7$  receptors. The results from a validation test of 125 chemicals with known nAChR  $\alpha 7$  receptor binding activity indicate that prediction accuracy, sensitivity, and specificity were 77.4, 57.7, and 86.2%, respectively. The development of a prediction model of human  $\alpha 7$  binding activity shows that tobacco products may have unusually high proportions of constituents relative to nicotine that may induce physiological effects mediated through nAChR receptor subtypes.

## MATERIAL AND METHODS

**Study Design.** The study design is depicted in Figure 1. In brief, the human and rat  $\alpha 7$  nAChR binding activity data were



**Figure 1.** Study design. Human and rat  $\alpha 7$  nAChR binding data were curated from PubChem and ChEMBL databases as training and external validation sets. Tobacco constituents were taken from the literature as the target set. Competitive docking was applied to all three data sets, and the resulting docking scores were partitioned to the amino acid residues near the binding pocket. The decision forest algorithm was used to build the models in the 1000 iterations of 5-fold cross validations. The model built with the entire training set was externally validated and then used to predict  $\alpha 7$  nAChR binding for the target set.

curated from the public databases PubChem (<https://pubchem.ncbi.nlm.nih.gov/>) and ChEMBL (<https://www.ebi.ac.uk/chembl/>). The human and rat  $\alpha 7$  nAChR binding activity data were used as the training and external validation data sets, respectively. The target set contains the tobacco constituents that were curated from a publication.<sup>9</sup> Competitive molecular docking<sup>10</sup> was used to generate the descriptors. All chemicals of the three data sets were subjected to competitive molecular docking to two-, three-dimensional (3D) structures of human  $\alpha 7$  nAChR constructed using

homology modeling and molecular dynamic (MD) simulations. The resultant docking scores and the partitioned scores to the residues in the active site of the 3D structures of  $\alpha 7$  nAChR and their energy terms were used as the descriptors in the model development. A decision forest algorithm<sup>21,23</sup> was used to generate the predictive model based on the training set. The 5-fold cross validation was performed 1000 times to determine the robustness and predictive power of the models. In addition, the 5-fold cross validations were used to examine the effect of prediction confidence and to identify the important descriptors that indicate the key residues and their interactions with ligands. The permutation was conducted to estimate the predictivity of the model. The developed model was validated using the external validation set of 1215 chemicals. The validated model was used to predict human  $\alpha 7$  nAChR binders or nonbinders for the tobacco constituents from which 84 were selected to experimentally validate the prediction results.

**Data Sets. Training Data Set.** The compounds with human  $\alpha 7$  nAChR binding activity data were collected from the popular chemical databases PubChem and ChEMBL by searching target CHRNA7. There were 803 and 974 compounds with human  $\alpha 7$  nAChR binding activity data in PubChem and ChEMBL, respectively. After removal of the compounds that have a conflict with  $\alpha 7$  nAChR binding activity data between the two databases and the compounds of which 3D structures could not be generated, 843 compounds remained and were used as the training compounds. Based on the  $\alpha 7$  nAChR binding activity data, the 843 compounds were classed as 715 binders and 128 nonbinders (Table S1). The LigPrep module (<https://www.schrodinger.com/ligprep>) from Maestro was used to generate the 3D structures of the 843 training compounds.

**External Validation Data Set.** The compounds, which had rat  $\alpha 7$  nAChR binding activity data, were collected from PubChem and ChEMBL. In total, 1501 and 1512 compounds were collected from PubChem and ChEMBL, respectively. In the same way, the compounds with conflicted experimental data and the compounds with difficulties to construct 3D structures were discarded. The remaining compounds were then classed into  $\alpha 7$  nAChR binders and nonbinders using their rat  $\alpha 7$  nAChR binding activity data. The resultant 1290 binders and 122 nonbinders were used as the external validation data set (Table S2). The 3D structures of the compounds in the external validation set were generated using LigPrep from Maestro.

Comparison of the training data set and the external validation data set found 197 compounds in both data sets. The 197 common compounds were removed from the external validation data set. The remaining 1215 compounds were used as the external validation data set (1096 binders and 119 nonbinders) to validate the model that was constructed using the training data set.

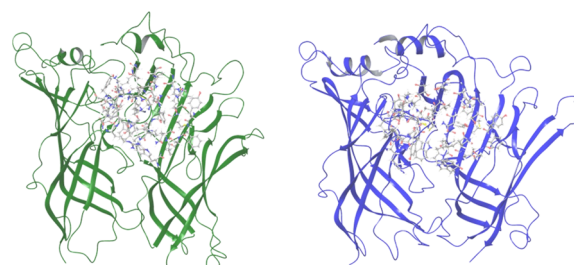
**Tobacco Constituents.** It is documented that there are more than 8000 tobacco constituents.<sup>9</sup> The developed model was used to predict the  $\alpha 7$  nAChR binding activity of these tobacco constituents. Three preprocessing steps were applied to construct structures of these tobacco constituents. First, 8391 compounds with chemical abstracts service registry numbers (CASNs) were extracted from public literature<sup>9</sup> using a Perl script. Second, the CASNs of 8391 compounds were searched in the PubChem database to get the structural details. Among the 8391 tobacco constituents, 6199 unique com-

pounds were found in the PubChem database using the CASNs. Finally, 5275 tobacco constituents remained after removing the small inorganic compounds, mixtures, and ions. The two-dimensional (2D) structures of 5275 tobacco constituents were downloaded from PubChem. Molconvert (<https://chemaxon.com/marvin-archive/3.3.3/marvin/doc/user/molconvert.html>) was used to convert the 2D structures into 3D structures with hydrogens. The LigPrep module from Maestro was used to validate and minimize the 3D structures.

**Competitive Molecular Docking.** We previously constructed two 3D structures of human  $\alpha 7$  nAChR using homology modeling and MD simulations.<sup>22</sup> They were used as the target proteins for the molecular docking approach in this study. Glide ([www.schrodinger.com/Glide](http://www.schrodinger.com/Glide)) from Maestro was used to dock the compounds in the active site of the proteins. First, a grid box was generated around the active site of each of the two 3D structures of human  $\alpha 7$  nAChR (hereafter, the two structures are termed structure\_1 and structure\_2). The 3D coordinates of the structure\_1 and structure\_2 are provided in a pdb format in Tables S3 and S4. The compounds of the training set and the external validation set, as well as the tobacco constituents, were then docked inside the grid boxes generated around the active sites of structure\_1 and structure\_2 using the extra precision (XP) method. The default docking parameters (0.80 for a van der Waals radius scaling factor and a partial charge cutoff of 0.50 in the Glide module) were applied. Five poses were selected for each pair of the ligand and protein for subsequently post-docking minimization. Of the five minimized poses, the best and its docking score were output. The competitive docking yielded three types of results: (1) the compounds were successfully docked to one structure, structure\_1 or structure\_2; (2) the compounds were successfully docked to both structures, the structure with lower docking scores was selected for subsequent analysis; (3) the compounds failed in docking to both structure\_1 and structure\_2. The first and second types of results assumed that the compounds are  $\alpha 7$  nAChR binders, and the third type of results indicated that the compounds are  $\alpha 7$  nAChR nonbinders.

**Partition of Docking Scores.** The overall docking score for each complex was partitioned using its Glide output file to the amino acid residues and energy terms. The energy terms include hydrogen bonding, coulomb interaction, van der Waals interaction, the sum of these three terms, and the minimum distance between the ligand and residue. The residues within 12 Å from the grid center were used in the partitioning of the overall docking scores. There are 14 residues from chain A (148S, 149W, 150S, 151Y, 152G, 186R, 187F, 188Y, 190C, 192K, 193E, 194P, and 195Y) and 23 residues from chain B (32Y, 33F, 34S, 55W, 56L, 57Q, 58M, 59S, 60W, 77T, 79R, 107W, 108V, 109L, 110V, 111N, 115H, 116C, 117Q, 118Y, 119L, 120P, and 121P) within this space in  $\alpha 7$  nAChR structures (Figure 2) that were used in the partition.

**Decision Forest Algorithm.** A decision forest, a pattern recognition algorithm we previously developed<sup>23</sup> and applied to analyze diverse types of data,<sup>17,26,26–28</sup> was used to develop the model to predict binders or nonbinders of human  $\alpha 7$  nAChR for chemicals. The decision forest combines individual trees to give final predictions based on the consensus of predictions from the individual trees. More specifically, the algorithm generated a decision tree based on the pool of partitioned docking scores. The partitioned docking scores, which were used in the previous tree, were then removed, and



**Chain A:** 148S, 149W, 150S, 151Y, 152G, 186R, 187F, 188Y, 190C, 191C, 192K, 193E, 194P, 195Y

**Chain B:** 32Y, 33F, 34S, 55W, 56L, 57Q, 58M, 59S, 60W, 77T, 79R, 107W, 108V, 109L, 110V, 111N, 115H, 116C, 117Q, 118Y, 119L, 120P, 121P

**Figure 2.** Amino acid residues used in the partition. The top panel gives the structures of human  $\alpha 7$  nAChR constructed using templates 3S06 (left) and 2XYT (right). The proteins were drawn in the ribbon model and the residues near the binding pockets in stick models. The amino acids were listed separately for chain A and chain B in the bottom panel.

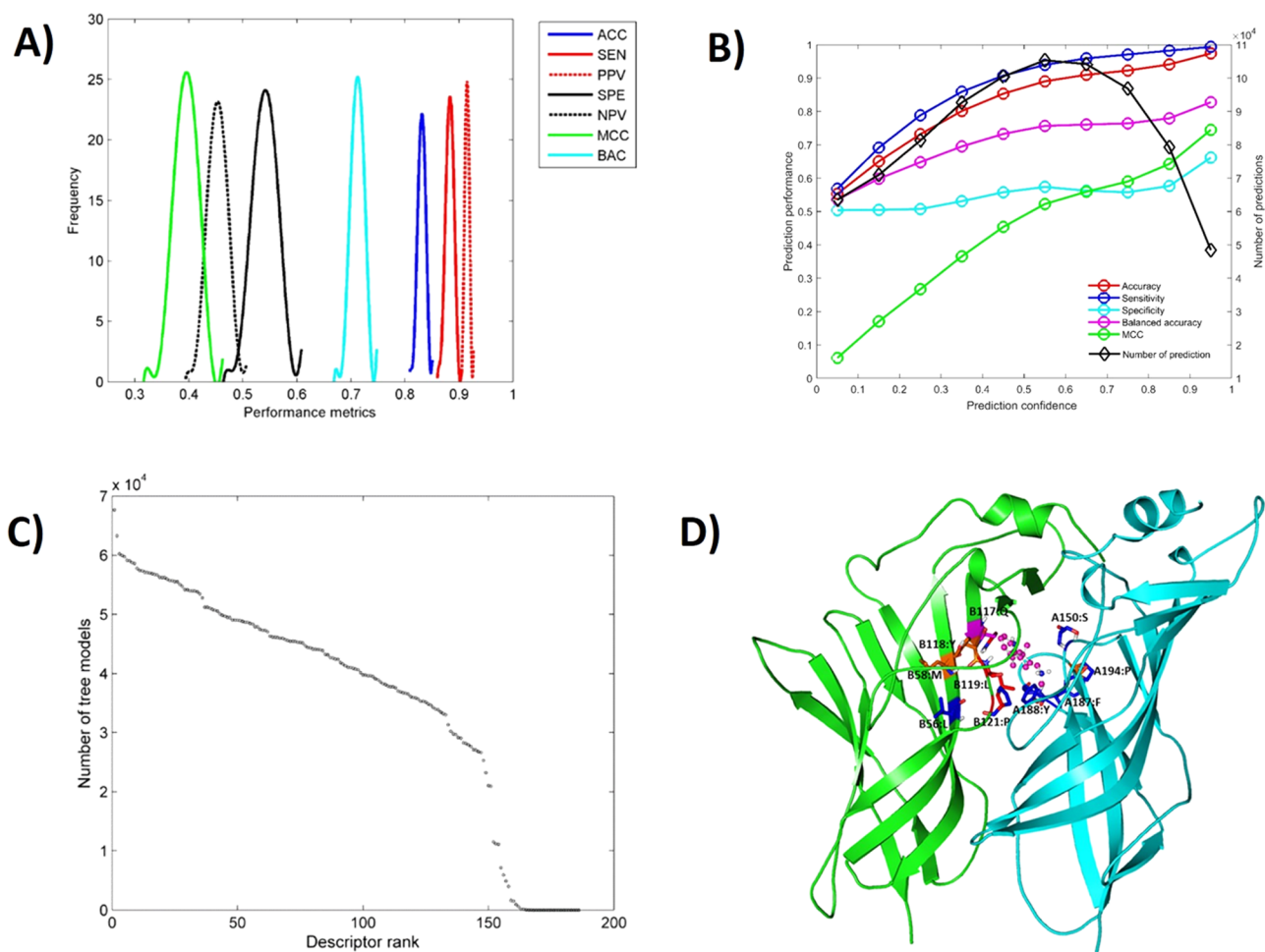
the remaining pool of partitioned docking scores was used to construct a new tree. The training result from combining all trees was evaluated. This process was repeated until no improvement in the training result could be achieved by adding a new decision tree. The following parameters were used to construct the individual trees: the maximum number of trees was set to 5, the minimum size of the node to be split was fixed at 10, the maximum number of levels to be pruned was 3, and Gini's diversity index was used for node splitting.

**Performance Metrics.** Seven prediction performance metrics (sensitivity, positive predictive rate, specificity, negative predictive rate, accuracy, Matthew's correlation coefficient (MCC), and balanced accuracy) were used to estimate performance of the predictive models. The detailed calculation equations for the seven metrics are provided in the Supporting Information. These seven metrics measure different aspects of a model. Sensitivity measures the capability of a model to predict active chemicals, while specificity measures its predictive capability on inactive chemicals. The positive predictive rate and negative predictive rate indicate goodness of positive and negative predictions for a model, respectively. The accuracy measures the overall goodness of a model. However, it depends at prevalence and is biased to the training majority. Balance accuracy and MCC reduce the dependency of the overall performance of a model at prevalence.

**Internal Validation.** The 5-fold cross-validation approach was used as the internal validation to validate the performance of the model for predicting  $\alpha 7$  nAChR binding activity. In a 5-fold cross validation, the training data set was randomly divided into five equal parts. Four parts were used to develop a predictive model and the remaining one part was used to validate the developed model. This process was repeated five times so that each of the five parts was used once and only once to challenge the model constructed using the other four parts. The performance of the predictions was measured by seven metrics calculated using all results of the five models. The 5-fold cross validation was iterated for 1000 times to a statistically robust performance estimation for the internal validation.

**External Validation.** The performance of the developed model using the decision forest was assessed using the external validation data set. A decision forest model was first constructed using the whole training data set. The model





**Figure 3.** Cross validation results. (A) Performance of the predictive model. The *x*-axis indicates performance metrics value and the *y*-axis gives the frequency of the models with the performance among the 1000 iterations of 5-fold cross validations. (B) Prediction confidence analysis results. The *x*-axis depicts the prediction confidence level. The number of predictions within a prediction confidence level from the 1000 iterations of 5-fold cross validations was plotted as a black diamond indicated at the right *y*-axis, and the corresponding performance metrics were plotted as circles (color codes in the color legend) with their values are indicated at the left *y*-axis. (C) The frequency of energy terms of amino acid residues of  $\alpha 7$  nAChR where the *x*-axis gives the rank and the *y*-axis shows the frequency. (D) The top energy terms of amino acid residues in  $\alpha 7$  nAChR.  $\alpha 7$  nAChR was drawn in a ribbon model, the ligand in ball models, and the informative residues in a stick model. ACC: accuracy, SEN: sensitivity, PPV: positive predictive value, SPE: specificity, NPV: negative predictive value, MCC: Mathews' correlation coefficient, and BAC: balance accuracy.

was then used to predict binders or nonbinders of  $\alpha 7$  nAChR for the chemicals in the external validation data set. The seven metrics were calculated based on the prediction result to estimate performance of the model.

**Prediction Confidence Analysis.** A multilabel classification predictive model generated using the decision forest not only predicts the class a sample belongs but also gives a probability that estimates the confidence of the sample belonging to the predicted class. Hence, the predictive model could be evaluated not only using overall prediction performance metrics but also by prediction confidences of the predictions. The prediction confidence from a decision forest model is likely a value between 0 (least confident) and 1 (higher confident) that is calculated using the equation below

$$\text{prediction confidence} = \frac{|P - 0.5|}{0.5}$$

where *P* is the prediction probability output from the decision forest model.

The confidence levels of the predictions in the 5-fold cross-validation were conducted to assess the performance of the  $\alpha 7$  nAChR binding activity prediction model. First, the prediction confidence values were calculated using the prediction probability values for the predictions in the 1000 iterations of 5-fold cross validation. The predictions were then assembled into 10 groups based on their prediction confidence values that were divided into 10 even bins. Lastly, the performance metrics values were calculated separately for the predictions in those 10 groups. The relationship between the performance of predictions and their prediction confidence levels was analyzed.

**Identification of Informative Amino Acids.** The 1000 iterations of 5-fold cross-validation constructed 5000 decision forest models that consists of tens of thousands of decision trees. The frequency of a partitioned docking score in those trees is an indication of the importance of the residue and associated energy term. The higher the frequency the more decision trees used the residue and the energy term. Analyses of the frequency values of the partitioned docking scores were

conducted to identify the important amino acids that play key roles in the interaction between ligands and  $\alpha 7$  nAChR. We first calculated the frequency values for all partitioned docking scores. Then, the partitioned docking scores were ranked by their frequency values, and a relative frequency was calculated. The residues and interaction energy terms that had a relative frequency greater than 1.5 in the decision forest (DF) models in the 5-fold cross validations were more informative than those with lower relative frequencies to the DF models. Subsequently, the top partitioned docking scores (amino acids and their energy terms) with a relative frequency greater than 1.5 were identified as informative descriptors that play key roles in the binding of human  $\alpha 7$  nAChR for chemicals.

**Prediction of  $\alpha 7$  nAChR Binding of Tobacco Constituents.** The predictive model that was constructed using the whole training data set was used to classify the 5275 tobacco constituents as binders and nonbinders of human  $\alpha 7$  nAChR. To assess the reliability of the prediction results of the 5275 tobacco constituents, 84 were randomly selected to be experimentally validated using the previously developed assays after considering compounds availability and similar distribution of predictions.<sup>25</sup> The 84 experimentally tested tobacco constituents and their measured activity values are given in Table S5.

Here, a brief description of the assay is given. The nAChR  $\alpha 7$  was expressed in Chinese Hamster Ovary (CHO) cell lines as described previously.<sup>24</sup> The CHO cells were maintained in Ham's F-12 media and transfected with the linearized cDNA constructs by nucleofection. The transfected cells were incubated with  $\alpha 7$ /RIC3 (antibiotics) for 2 weeks to select the stably expressed cells. The transfected cells were passed to the medium lacking antibiotics before 2 to 4 days to start the experiment. The test chemicals were dissolved in dimethyl sulfoxide (DMSO) and prepared in 384-well compound plates using an automated liquid handling system. Then, the cell suspension was added into the wells, and membrane currents were recorded by on-board patch clamp amplifiers. Inward current peak amplitudes and charge movement were measured. The  $\alpha 7$  nAChR activation was calculated as the ratio of the agonist-elicited current signal to the mean current signal elicited with 300  $\mu$ M nicotine.

**Generation of Chemical Space.** To better understand the performance of the model, the chemical spaces of the training chemicals, the chemicals in the external validation set, and the tobacco constituents were generated. First, the partitioned docking scores for each of the three sets of chemicals were transformed by principal component analysis using MatLab (<https://www.mathworks.com/products/matlab.html>). Then, the top three principal components were used to present the chemical spaces. At last, the three chemical spaces were visually compared using scatter plots.

## RESULTS

**Internal Validation.** One thousand 5-fold cross validations were conducted to estimate the robustness and accuracy of the models generated for predicting of human  $\alpha 7$  nAChR binding activity of chemicals. The performance metrics values of the models from the 5-fold cross validations were calculated and are given in Figure 3A. The average predicted accuracy, sensitivity, specificity, MCC, and balanced accuracy values for 1000 times iteration of 5-fold cross validation are 0.831, 0.883, 0.541, 0.395, and 0.712, respectively. The metric values indicated that the predictive model has an excellent predictive

power to differentiate the binder and nonbinder of human nAChR.

**Prediction Confidence.** The prediction confidence was analyzed for the predictions from the 1000 iterations of 5-fold cross validation. Figure 3B shows the accuracy, sensitivity, specificity, balanced accuracy, and MCC calculated for the predictions against their prediction confidence. It was observed that performance of predictions is dependent at their prediction confidence: the higher likelihood value in a prediction, the better is the performance of the prediction. For the predictions with confidence higher than 0.8, the accuracy is higher than 90%. Furthermore, most of the predictions were at the middle (0.3–0.7) or high (>0.7) confidence level. The prediction confidence analysis demonstrated the prediction confidence from the models provides an additional parameter that facilitates the application of the human  $\alpha 7$  nAChR binding activity prediction model.

**Identification of Key Amino Acids.** The amino acids and their energy terms that are important to human  $\alpha 7$  nAChR binding were selected based on the frequency of the descriptors used in the predictive models generated in the 1000 iterations of 5-fold cross validations. Figure 3C plots the ranked amino acid energy terms against the frequency used in the decision trees. The average frequency of all residue energy terms was calculated. The relative frequency (frequency divided by the average frequency) was calculated for each amino acid energy term. The energy terms with a relative frequency greater than 1.5 (used in >54,500 trees) were deemed more informative than those with lower relative frequencies to the DF models. Of the 186 amino acid energy terms used in the constructions of the human  $\alpha 7$  nAChR predictive models in the cross validations, 28 were used by more than 54,500 decision trees and are from 14 amino acids. Among the 14 amino acids, L119 and C193 were reported to form interactions with nicotine.<sup>29</sup> The key amino acids and their energy terms identified by the 5-fold cross validations in the complex structure of the human  $\alpha 7$  nAChR with CHEMBL1835619 are depicted in Figure 3D. The site-directed mutagenesis studies found that the mutations of glycine 189 of the chicken  $\alpha 7$ -subunit to aspartic acid, glutamic acid, asparagine, and glutamine impacted neonicotinoid sensitivity of the  $\alpha 7$  nAChR.<sup>30</sup> Interestingly, this glycine is not in the list of key amino acids identified in our modeling. This may indicate that this smallest residue may not directly interact with ligands but is in the binding region. Mutation of it to a bulkier amino acid changes the conformation of its adjacent residues (the key amino acid proline 121 identified in this study) that in turn affect interactions with ligands. However, the evident interaction mechanism of this glycine deserves further investigation.

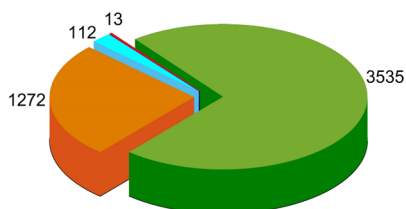
**External Validation.** Table 1 summarizes the prediction results of an external test set (1215 chemicals with rat  $\alpha 7$

Table 1. Validation Results of the Predictive Model

| validation results | actual    | predicted positives | predicted negative | total |
|--------------------|-----------|---------------------|--------------------|-------|
| external           | positives | 962                 | 134                | 1096  |
|                    | negatives | 99                  | 20                 | 119   |
|                    | total     | 1061                | 154                | 1215  |
| experimental       | positives | 14                  | 21                 | 35    |
|                    | negatives | 9                   | 40                 | 49    |
|                    | total     | 23                  | 61                 | 84    |

nAChR binding experimental data) using the human  $\alpha 7$  nAChR binding predictive model. The accuracy, sensitivity, specificity, balanced accuracy, MCC, positive predictive rate, and negative predictive rate for the external test set are 0.808, 0.878, 0.168, 0.523, 0.132, 0.907, and 0.130, respectively. The external validation results confirmed the predictive power of the human  $\alpha 7$  nAChR binding predictive model.

**Prediction on Tobacco Constituents.** The 5275 tobacco constituents were docked in the binding sites of homology-modelled human  $\alpha 7$  nAChR, structure\_1 and structure\_2, based on two template structures. Among the 5275 tobacco constituents, 5132 were successfully docked in the binding site of structure\_1 or structure\_2. The remaining compounds failed to dock in the binding site of structure\_1 or structure\_2. The docking scores were partitioned in different energy terms to the residues (12 Å from the center of the grid) in the binding sites of structure\_1 and structure\_2. The human  $\alpha 7$  nAChR binding predictive model was used to identify the binders and nonbinders of human  $\alpha 7$  nAChR from 5132 tobacco constituents. Among the 5132 tobacco constituents, 1597 were predicted as binders and 3535 were predicted as nonbinders of the human  $\alpha 7$  nAChR (Figure 4). The



**Figure 4.** Predictions on  $\alpha 7$  nAChR binding activity for the tobacco constituents. The types of the predictions were given by the slices of the pie. Predicted  $\alpha 7$  nAChR nonbinders are color coded in blue, high confidence binders in purple, moderate confidence binders in green, and low confidence binders in red. The number of tobacco constituents is given for each slice.

prediction confidence analysis was calculated for the 1597 predicted binders. Among the 1597 binders, 13 tobacco constituents showed a higher confidence value (greater than 0.6), 112 had prediction confidence values between 0.3 and 0.6, and the remaining 1272 had low prediction confidence values less than 0.3.

**Validation of Prediction Results.** To confirm the predictive power and robustness of the human  $\alpha 7$  nAChR binding predictive model, 84 tobacco constituents were randomly selected to experimentally test for human  $\alpha 7$  nAChR binding. Among the 84 compounds, 23 and 61 were predicted as binders and nonbinders, respectively (Table S5). The experimental validation of the 84 tobacco constituents showed that 35 and 49 compounds were binders and nonbinders of human  $\alpha 7$  nAChR, respectively. The comparison between the experimental and prediction results is summarized in Table 1. The prediction accuracy, sensitivity, specificity, balanced accuracy, MCC, positive prediction rate, and negative prediction rate for the 84 compounds were 64.3, 40.0, 81.6, 60.8, 23.9, 60.8, and 65.5%, respectively.

## DISCUSSION

Knowing the 3D information of a target protein has the potential to improve the predictive model for ligand binding to the target. Understanding the interactions between tobacco

constituents and residues in the binding site of human  $\alpha 7$  nAChR is important to identify its potential binders and nonbinders given the role in nicotine addiction that this target protein plays. Until now, a 3D structure of human  $\alpha 7$  nAChR was not available. Hence, homology modelling methodology was used to generate the 3D structures of human  $\alpha 7$  nAChR based on two templates. The template proteins were selected based on the binding pocket of the AChR.

Molecular docking is one of the well-known methods to estimate binding potential in the binding site of a protein for chemicals. Rigid and flexible docking methods are widely used by researchers. Rigid docking does not allow flexibility for the protein. Flexible docking gives flexibility to the protein; however, it is time consuming to run flexible docking. Hence, it is difficult to apply flexible docking in the screening of a large number of compounds, such as those that are known to be present as constituents in tobacco smoke. To overcome this difficulty, more than one experimentally determined 3D structures of a protein are used in molecular docking. Hence, the competitive docking method was developed previously<sup>10,22</sup> and used to dock the compounds in the binding site of human  $\alpha 7$  nAChR in this study to accommodate the conformations induced by chemical binding.

In this investigation, the predictive model was developed based on the binding data set of the human  $\alpha 7$  nAChR. The predictive model was validated using rigorous cross-validation and external validation test sets. However, it should be noted that the human  $\alpha 7$  nAChR binding predictive model has limitations when it is applied to assess addiction potential for tobacco constituents. One limitation is due to the inaccurate descriptors generated based on the homology models of the human  $\alpha 7$  nAChR optimized using molecular dynamics simulations. Hence, using experimentally determined crystal structures of the human  $\alpha 7$  nAChR could improve the human  $\alpha 7$  nAChR predictive model. Another limitation is caused by the imbalance of binders and nonbinders in the training set. The training set contains a higher number of binders (85%) than nonbinders (15%). Consequently, as demonstrated in the cross-validation and external validation tests, the model performed with a higher skill in the prediction of binders than nonbinders. Therefore, the model has a limitation in its prediction of specificity. Inclusion of more negative compounds in the training data set is expected to improve the specificity of the predictive model. However, the high level of sensitivity that was observed based on external validation with rat  $\alpha 7$  nAChR binding activity data that were collected from PubChem and ChEMBL for a robust set of 1215 chemicals (>80%) is promising. From a public health and knowledge development standpoint of identifying the harm of compound addiction liability, being able to predict a tobacco constituent, ingredient, or additive to bind to the nAChR is advantageous. The advantage of a high sensitivity performing predictive model is that the model would tend to error on the side of predicting a false positive rather than a false negative. For obvious reasons of protecting public health from addiction, erroring on the side of a false positive is more desirable than a false negative. Additionally, a highly sensitive model supports efforts with compound priority setting for further testing to confirm the ligand–protein interaction. Therefore, the model higher sensitivity compared to high specificity is a desirable performance characteristic from a regulatory perspective.

A model is trained by extracting the relationship among the chemicals in the training set. Therefore, the chemical space of



the training chemicals is important for suitable applications of the model. If the chemicals in prediction are not covered and far away from the training chemical space, the model is not suitable for the prediction and the prediction results may not be reliable. The chemical space of the training set, test set, and the target tobacco constituents was calculated using PCA based on Mold2 descriptors.<sup>31</sup> The chemical spaces shown in Figure S1 revealed that the chemical training space of the model covered most of the chemicals in the test set the tobacco constituents in the prediction. Therefore, the prediction results on the tobacco constituents are expected to be useful. Moreover, the chemical coverage for the test set and tobacco constituents are similar, and the performance of the prediction results on the tobacco constituents is close to the external validation performance.

Comparison of the training set with the external test set found that 197 compounds have been tested for both human and rat  $\alpha 7$  nAChR. Hence, the 197 compounds were used to estimate the consistency between the experimental data of human and rat  $\alpha 7$  nAChR. The consistency was used to assess the performance of the predicted model that was generated based on human  $\alpha 7$  nAChR assay data on the compounds tested with rat  $\alpha 7$  nAChR assay data. Among 197 compounds tested in human  $\alpha 7$  nAChR had shown 167 and 0 compounds as positive and negative, respectively, in rat  $\alpha 7$  nAChR. The metrics such as accuracy, sensitivity, specificity, balanced accuracy, MCC, and positive and negative prediction rate values are similar with those from external validation results (Table 1), indicating that the developed model reached the performance of experiments.

Prediction confidence analysis showed that the human  $\alpha 7$  nAChR predictive model can accurately predict the binding activity for most compounds but not so well for some compounds. Increase in the prediction confidence value improves the prediction accuracy. Hence, combining the activity prediction by the human  $\alpha 7$  nAChR predictive model with its prediction confidence analysis could better utilize the prediction results.

Although in product development efforts, there may be a focus on generating a model that performs with a high positive predictivity rate (i.e., models with high specificity),<sup>32,33</sup> it is also within reason to support strategies in tobacco harm reduction using a predictive model geared toward a high negative predictive value (i.e., models with high sensitivity), which is a metric that indicates how frequent negative predictions are correct. False negatives may be a great concern for regulatory purposes because there is an addiction liability associated with binding to the human  $\alpha 7$  nAChR, and a prediction that tends to make erroneous negative predictions (i.e., high false negative rate) would miss detecting this liability. Models with high specificity have high false negative rates, and although models with high sensitivity have high false positive rates, an error in prediction with a high sensitivity performing model, such as the one in this study, would side on ensuring that there is predictive power in detecting the potential harm of binding to the human  $\alpha 7$  nAChR. The 5-fold cross validations and external validation showed higher sensitivity than specificity, which is a favorable observation in predictive performance for regulatory science. Although the model needs improvement in specificity, the high sensitivity is highly promising and makes it more suitable for screening of positives than predicting negatives. As a model development strategy, this study not only undertook a thorough computational cross

validation through performing 1000 iterations in the cross-validation test but it also performed external validation using randomly selected test set of compounds with known activity to bind to human  $\alpha 7$  nAChR. The external validation assessment then went even further to learn of the model predictive performance. From a structure–activity and structure–property relationship with respect to nAChR alpha7, it appears that the nAChR alpha receptor has marginal-to-good specificity for this test set given that 30% (1597 of 5275) of the modeled tobacco chemicals bind to the receptor, and the chemical space of the model showed a good coverage of the chemicals. Yet, only 13 tobacco constituents showed a higher confidence value (greater than 0.6), and 112 had prediction confidence values between 0.3 and 0.6, with the remaining 1272 showing low prediction confidence values (< 0.3). It performed an additional layer of external validation by conducting experimental testing of a subset of the external validation test set to verify binding or nonbinding activity to the human  $\alpha 7$  nAChR. Thus, there was extensive validation and rigorous testing of the model that provides evidence on how we could expect the model to perform. Since high sensitivity was a characteristic feature of the predictive performance of the model during validation tests, there is utility for it in screening tobacco constituents for their potential to contribute to addiction liability of tobacco products.

Among the 186 amino acid energy terms used in the models, 28 were identified as informative descriptors to the human  $\alpha 7$  nAChR predictive models (Figure 3D). Therefore, these residues and their interaction energy terms are considered as the most important structural features to differentiate the binder and nonbinder of human  $\alpha 7$  nAChR for tobacco constituents based on their interactions with the human  $\alpha 7$  nAChR. The 28 structural features of the human  $\alpha 7$  nAChR are expected to help in understanding the binding interactions of tobacco constituents with human  $\alpha 7$  nAChR.

## CONCLUSIONS

In this study, the human  $\alpha 7$  nAChR predictive model was developed and extensively validated using large data sets collected from the public chemical databases as well as experimental testing post computational prediction. The competitive docking method used in this study revealed the important residues and their energy terms involved in the binding of the compounds in the binding site of human  $\alpha 7$  nAChR. The results obtained from the competitive docking were used as descriptors to build a predictive model using the decision forest algorithm to differentiate tobacco constituents as potential binders or nonbinders to the human  $\alpha 7$  nAChR. The 1000 iterations of 5-fold cross-validation and permutation tests confirmed that the predictive model has good predictive ability and robustness. The predictive power of the model was further demonstrated by the external validation testing. The validated predictive model was used to differentiate the binder and nonbinder for the tobacco constituents. The predictive performance of the human  $\alpha 7$  nAChR predictive model on the tobacco constituents was experimentally validated on the randomly selected 84 tobacco constituents. In addition, the prediction confidence from the model provides information on the utilization of the prediction results. In conclusion, the developed predictive model for ligand binding to the human  $\alpha 7$  nAChR demonstrated a performance with utility for

screening tobacco constituents for harm with regard to potential addiction liability.

## ■ ASSOCIATED CONTENT

### SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jcim.0c00139>.

Training data set (Table S1), external validation data set (Table S2), 3D structure<sub>1</sub> of  $\alpha 7$  nAChR (Table S3), 3D structure<sub>2</sub> of  $\alpha 7$  nAChR (Table S4), and the 84 experimentally tested compounds and prediction results (Table S5) (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Authors

**Carmine Leggett** – Division of Nonclinical Science, Office of Science, Center for Tobacco Products, U.S. Food and Drug Administration, Calverton, Maryland 20705, United States; Email: [carmine.leggett@gmail.com](mailto:carmine.leggett@gmail.com)

**Huixiao Hong** – Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, U.S. Food and Drug Administration, Jefferson, Arkansas 72079, United States; [orcid.org/0000-0001-8087-3968](https://orcid.org/0000-0001-8087-3968); Email: [huixiao.hong@fda.hhs.gov](mailto:huixiao.hong@fda.hhs.gov)

### Authors

**Sugunadevi Sakkiah** – Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, U.S. Food and Drug Administration, Jefferson, Arkansas 72079, United States

**Bohu Pan** – Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, U.S. Food and Drug Administration, Jefferson, Arkansas 72079, United States

**Wenjing Guo** – Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, U.S. Food and Drug Administration, Jefferson, Arkansas 72079, United States

**Luis G. Valerio, Jr.** – Division of Nonclinical Science, Office of Science, Center for Tobacco Products, U.S. Food and Drug Administration, Calverton, Maryland 20705, United States

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.jcim.0c00139>

### Author Contributions

H.H. and C.L. designed and led the project. S.S. collected the data and implemented the methods. S.S., B.P., W.G., C.L., L.V., and H.H. discussed the data analysis and the results. S.S., H.H., L.V., and C.L. wrote the manuscript.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

This study was funded by the Center for Tobacco Products (CTP) of the U.S. Food and Drug Administration (FDA). This publication represents the views of the author(s) and does not represent FDA/CTP position or policy. This research was supported in part by an appointment to the Research Participation Program at the National Center for Toxicological Research (Sugunadevi Sakkiah, Bohu Pan and Wenjing Guo) administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and the U.S. Food and Drug Administration.

## ■ REFERENCES

- (1) Rostron, B. Smoking-attributable Mortality in the United States. *Epidemiology* **2011**, *22*, 350–355.
- (2) US Department of Health and Human Services. *The Health Consequences of Smoking—50 Years of Progress: A Report of the Surgeon General*; U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health, 2014.
- (3) Fowler, C. D.; Kenny, P. J. Nicotine Aversion: Neurobiological Mechanisms and Relevance to Tobacco Dependence Vulnerability. *Neuropharmacology* **2014**, *76*, 533–544.
- (4) Albuquerque, E. X.; Pereira, E. F. R.; Alkondon, M.; Rogers, S. W. Mammalian Nicotinic Acetylcholine Receptors: From Structure to Function. *Physiol. Rev.* **2009**, *89*, 73–120.
- (5) Karlin, A. Emerging Structure of the Nicotinic Acetylcholine Receptors. *Nat. Rev. Neurosci.* **2002**, *3*, 102–114.
- (6) Anand, R.; Conroy, W. G.; Schoepfer, R.; Whiting, P.; Lindstrom, J. Neuronal Nicotinic Acetylcholine Receptors Expressed in *Xenopus* oocytes have a Pentameric Quaternary Structure. *J. Biol. Chem.* **1991**, *266*, 11192–11198.
- (7) Cooper, E.; Couturier, S.; Ballivet, M. Pentameric Structure and Subunit Stoichiometry of a Neuronal Nicotinic Acetylcholine Receptor. *Nature* **1991**, *350*, 235–238.
- (8) Fowler, C. D.; Arends, M. A.; Kenny, P. J. Subtypes of Nicotinic Acetylcholine Receptors in Nicotine Reward, Dependence, and Withdrawal: Evidence from Genetically Modified Mice. *Behav. Pharmacol.* **2008**, *19*, 461–484.
- (9) Rodgman, A.; Perfetti, T. *The Chemical Components of Tobacco and Tobacco Smoke*; Second Edition. CRC Press Taylor & Francis Group: Florida, 2013.
- (10) Ng, H. W.; Zhang, W.; Shu, M.; Luo, H.; Ge, W.; Perkins, R.; Tong, W.; Hong, H. Competitive Molecular Docking Approach for Predicting Estrogen Receptor Subtype  $\alpha$  Agonists and Antagonists. *BMC Bioinf.* **2014**, *15*, S4.
- (11) Sakkiah, S.; Arullaperumal, V.; Hwang, S.; Lee, K. W. Ligand-based Pharmacophore Modeling and Bayesian Approaches to Identify c-Src Inhibitors. *J. Enzyme Inhib. Med. Chem.* **2013**, *29*, 69–80.
- (12) Sakkiah, S.; Baek, A.; Lee, K. W. Pharmacophore Modeling and Molecular Dynamics Simulation to Identify the Critical Chemical Features Against Human Sirtuin 2 Inhibitors. *J. Mol. Struct.* **2012**, *1011*, 66–75.
- (13) Drake, R. R.; Neamati, N.; Hong, H.; Pilon, A. A.; Suntharankar, P.; Hume, S. D.; Milne, G. W. A.; Pommier, Y. Identification of a Nucleotide Binding Site in HIV-1 Integrase. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 4170–4175.
- (14) Salmaso, V.; Moro, S. Bridging Molecular Docking to Molecular Dynamics in Exploring Ligand-Protein Recognition Process: An Overview. *Front. Pharmacol.* **2018**, *9*, 923.
- (15) de Ruyck, J.; Brysbaert, G.; Blossey, R.; Lensink, M. F. Molecular Docking as a Popular Tool in Drug Design, an in silico Travel. *Adv. Appl. Bioinform. Chem.* **2016**, *Volume 9*, 1–11.
- (16) Sakkiah, S.; Guo, W.; Pan, B.; Kusko, R.; Tong, W.; Hong, H. Computational Prediction Models for Assessing Endocrine Disrupting Potential of Chemicals. *J. Environ. Sci. Health, Part C: Environ. Carcinog. Ecotoxicol. Rev.* **2018**, *36*, 192–218.
- (17) Sakkiah, S.; Selvaraj, C.; Gong, P.; Zhang, C.; Tong, W.; Hong, H. Development of Estrogen Receptor Beta Binding Prediction Model Using Large Sets of Chemicals. *Oncotarget* **2017**, *8*, 92989–93000.
- (18) Hong, H.; Tong, W.; Fang, H.; Shi, L.; Xie, Q.; Wu, J.; Perkins, R.; Walker, J. D.; Branham, W.; Sheehan, D. M. Prediction of Estrogen Receptor Binding for 58,000 Chemicals Using an Integrated System of a Tree-based Model with Structural Alerts. *Environ. Health Perspect.* **2002**, *110*, 29–36.
- (19) Liu, J.; Mansouri, K.; Judson, R. S.; Martin, M. T.; Hong, H.; Chen, M.; Xu, X.; Thomas, R. S.; Shah, I. Predicting Hepatotoxicity Using ToxCast in vitro Bioactivity and Chemical Structure. *Chem. Res. Toxicol.* **2015**, *28*, 738–751.

(20) Ng, H. W.; Doughty, S. W.; Luo, H.; Ye, H.; Ge, W.; Tong, W.; Hong, H. Development and Validation of Decision Forest Model for Estrogen Receptor Binding Prediction of Chemicals Using Large Data Sets. *Chem. Res. Toxicol.* **2015**, *28*, 2343–2351.

(21) Tong, W.; Hong, H.; Fang, H.; Xie, Q.; Perkins, R. Decision Forest: Combining the Predictions of Multiple Independent Decision Tree Models. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 525–531.

(22) Ng, H. W.; Leggett, C.; Sakkiah, S.; Pan, B.; Ye, H.; Wu, L.; Selvaraj, C.; Tong, W.; Hong, H. Competitive Docking Model for Prediction of the Human Nicotinic Acetylcholine Receptor  $\alpha 7$  Binding of Tobacco Constituents. *Oncotarget* **2018**, *9*, 16899–16916.

(23) Hong, H.; Tong, W.; Xie, Q.; Fang, H.; Perkins, R. An in silico Ensemble Method for Lead Discovery: Decision Forest. *SAR QSAR Environ. Res.* **2005**, *16*, 339–347.

(24) Kirsch, G. E.; Fedorov, N. B.; Kuryshv, Y. A.; Liu, Z.; Armstrong, L. C.; Orr, M. S. Electrophysiology-Based Assays to Detect Subtype-Selective Modulation of Human Nicotinic Acetylcholine Receptors. *Assay Drug Dev. Technol.* **2016**, *14*, 333–344.

(25) Wible, B. A.; Kuryshv, Y. A.; Smith, S. S.; Liu, Z.; Brown, A. M. An Ion Channel Library for Drug Discovery and Safety Screening on Automated Platforms. *Assay Drug Dev. Technol.* **2008**, *6*, 765–780.

(26) Hong, H.; Thakkar, S.; Chen, M.; Tong, W. Development of Decision Forest Models for Prediction of Drug-Induced Liver Injury in Humans Using A Large Set of FDA-approved Drugs. *Sci. Rep.* **2017**, *7*, 17311.

(27) Hong, H.; Rua, D.; Sakkiah, S.; Selvaraj, C.; Ge, W.; Tong, W. Consensus Modeling for Prediction of Estrogenic Activity of Ingredients Commonly Used in Sunscreen Products. *Int. J. Environ. Res. Public Health* **2016**, *13*, 958.

(28) Hong, H.; Shen, J.; Ng, H. W.; Sakkiah, S.; Ye, H.; Ge, W.; Gong, P.; Xiao, W.; Tong, W. A Rat  $\alpha$ -Fetoprotein Binding Activity Prediction Model to Facilitate Assessment of the Endocrine Disruption Potential of Environmental Chemicals. *Int. J. Environ. Res. Public Health* **2016**, *13*, 372–372.

(29) Morales-Perez, C. L.; Noviello, C. M.; Hibbs, R. E. X-ray Structure of the Human  $\alpha 4\beta 2$  Nicotinic Receptor. *Nature* **2016**, *538*, 411–415.

(30) Matsuda, K.; Buckingham, S. D.; Kleier, D.; Rauh, J. J.; Grauso, M.; Sattelle, D. B. Neonicotinoids: Insecticides Acting on Insect Nicotinic Acetylcholine Receptors. *Trends Pharmacol. Sci.* **2001**, *22*, 573–580.

(31) Hong, H.; Xie, Q.; Ge, W.; Qian, F.; Fang, H.; Shi, L.; Su, Z.; Perkins, R.; Tong, W. Mold<sup>2</sup>, Molecular Descriptors from 2D Structures for Chemoinformatics and Toxicoinformatics. *J. Chem. Inf. Model.* **2008**, *48*, 1337–1344.

(32) Valencia, A.; Prous, J.; Mora, O.; Sadrieh, N.; Valerio, L. G., Jr. A Novel QSAR Model of Salmonella Mutagenicity and its Application in the Safety Assessment of Drug Impurities. *Toxicol. Appl. Pharmacol.* **2013**, *273*, 427–434.

(33) Valerio, L. G., Jr. Tenth Anniversary of Expert Opinion on Drug Metabolism & Toxicology. *Expert Opin. Drug Metab. Toxicol.* **2014**, *10*, 767–768.