

# Identification of putative skin sensitizers using QSAR models



UNC  
ESHELMAN  
SCHOOL OF PHARMACY



UFG  
UNIVERSIDADE FEDERAL DE GOIÁS

A. Tropsha<sup>a</sup>, V. Alves<sup>b</sup>, E. Muratov<sup>a</sup>, D. Fourches<sup>a</sup>,  
J. Strickland<sup>c</sup>, N. Kleinstreuer<sup>c</sup>, C. Andrade<sup>b</sup>

<sup>a</sup>Laboratory for Molecular Modeling, UNC Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC, 27599, USA;

<sup>b</sup>Laboratory of Molecular Modeling and Design, Faculty of Pharmacy, Federal University of Goiás, Goiânia, GO, 74605, Brazil;

<sup>c</sup>ILS, Inc., P.O. Box 13501, Research Triangle Park, NC 27709, USA.

## RESEARCH OBJECTIVES

Skin sensitization is an immune reaction caused by the repetitive exposure to a chemical agent. Until now, there is no standard method to replace animal testing for the identification of skin sensitizing chemicals.

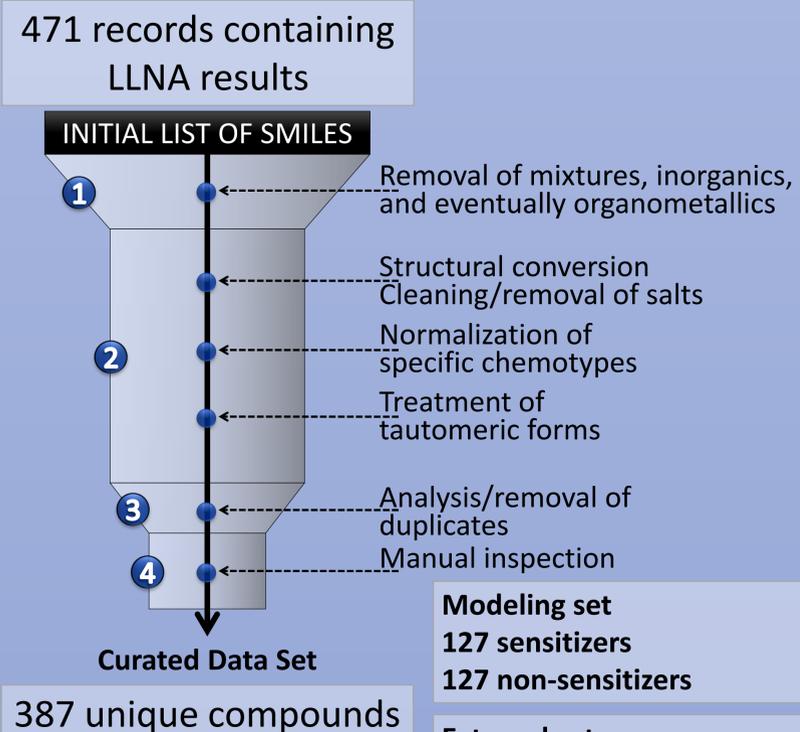


Sensitized skin

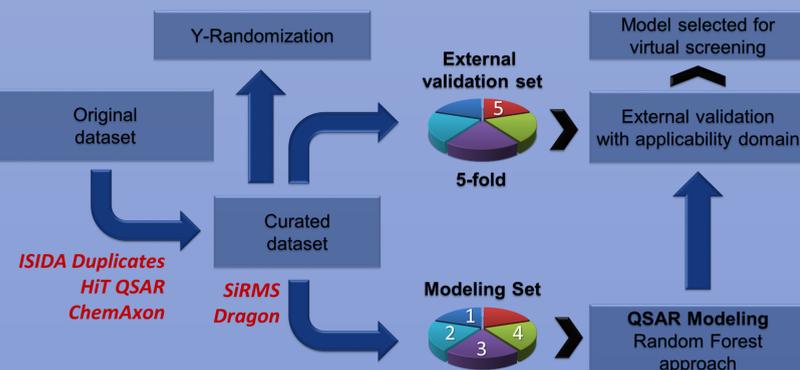
We have developed computational models capable of accurately assessing the skin sensitization potential of environmental chemicals. To this end, we have (i) compiled, curated, and integrated the largest publicly-available database of skin-sensitizing chemicals; (ii) used this data to generate and validate QSAR models for skin sensitization; and (iii) employed these models to identify putative sensitizers among chemicals in the Scorecard and Tox21 databases.

## MATERIALS AND METHODS

### Skin sensitization data set



### QSAR modeling workflow



## RESULTS AND DISCUSSION

### Developed QSAR models for skin sensitization.

Models	Sensitivity	PPV	Specificity	NPV	CCR	Coverage
SiRMS	0.83	0.85	0.83	0.81	0.83	0.58
Dragon	0.84	0.88	0.87	0.80	0.85	0.52
Consensus	0.79	0.85	0.85	0.79	0.82	0.70
Consensus Rigor	0.91	0.90	0.85	0.88	0.88	0.39

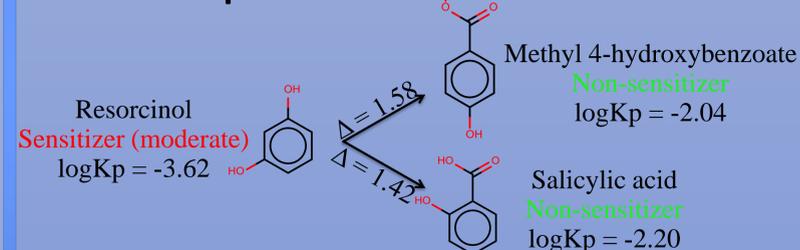
5-fold External CV results for QSAR models. CCR = Correct Classification Rate. Y-randomization = 0.5

### Comparison with OECD QSAR Toolbox

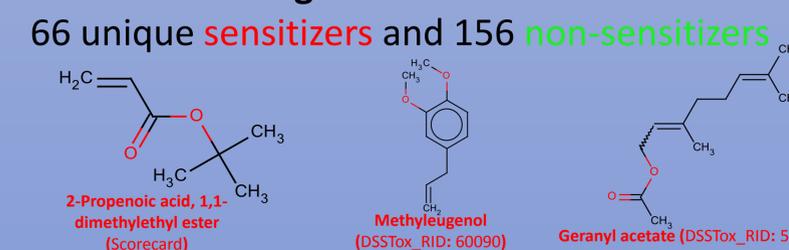
Models	Sensitivity	PPV	Specificity	NPV	CCR	Coverage
Consensus	0,48	0,95	0,98	0,71	0,73	0,60
QSAR Toolbox	0,82	0,51	0,20	0,52	0,51	0,92

163 compounds absent from QSAR Toolbox database. Consensus = predictions from external folds.

### Model interpretation



### Virtual screening of Scorecard and Tox21



## CONCLUSIONS

The overall classification accuracies of QSAR models discriminating sensitizers from non-sensitizers were 68-88% when evaluated on several external validation sets. When compared to the OECD QSAR Toolbox skin sensitization module, our models afforded significantly higher Positive and Negative Predictive Rates. When applied to chemicals within the applicability domains, the models could reliably identify positive and negative sensitizers with 94% and 71% certainty, respectively. Statistically significant descriptors from high-accuracy models yielded SAR rules that could guide structural optimization of chemicals of interest. Using these models, we have identified 66 putative skin sensitizers in the ScoreCard and Tox21 databases as primary hits for further experimental testing. Models developed in this study are publicly available for early identification of skin sensitizers through Chembench (<http://chembench.mml.unc.edu>).

**Acknowledgements :** NIH (GM66940, GM096967), EPA (RD 83499901), FAPEG, NIEHS (N01-ES35504).