

## Predicting Skin Sensitization Using ToxCast Assays

N Kleinstreuer<sup>1</sup>, J Strickland<sup>1</sup>, D Allen<sup>1</sup>, W Casey<sup>2</sup>

<sup>1</sup>ILS/NICEATM, RTP, NC, USA; <sup>2</sup>NIH/NIEHS/DNTP/NICEATM, RTP, NC, USA

### Abstract

Allergic contact dermatitis (ACD) is an adverse health effect from repeated exposure to skin-sensitizing chemicals and products. To minimize ACD, regulatory authorities require tests like the murine local lymph node assay (LLNA) to identify potential skin sensitizers. The Organisation for Economic Co-operation and Development (OECD) established an adverse outcome pathway (AOP) for skin sensitization. Many organizations, including the OECD and NICEATM, are pursuing integrated testing strategies using novel *in vitro* and *in silico* approaches to reduce or replace animal use. The U.S. EPA's ToxCast project includes high-throughput screening (HTS) assays in human primary skin cells and other systems that map to key events in the AOP (e.g., oxidative stress, cytokines). We built a cross-validated random forest model using ToxCast data and a balanced training set of 60 chemicals with *in vivo* LLNA data. The model predicted LLNA results with 80% accuracy, representing the performance against all chemicals when they appear in external test sets. The assays with the highest variable importance included known AOP targets (e.g., Nrf2, T-cell proliferation) as well as targets outside of the current AOP (e.g., Coll III, PPAR, PXR, ER). Compounds mispredicted by the model were found to be structurally similar, and we will discuss potential enrichment of this approach by incorporating molecular descriptors. Well-characterized AOPs like skin sensitization provide opportunities to use ToxCast HTS data to identify critical biological targets and develop efficient testing strategies that minimize animal use in regulatory testing.

## Introduction

- Allergic contact dermatitis (ACD) is a skin reaction characterized by localized redness, swelling, blistering, or itching after direct contact with a skin allergen (**Figure 1**). Workers and consumers can develop ACD when exposed to skin-sensitizing chemicals and products, which include substances such as nickel and formaldehyde.
- ACD is a common condition and is difficult to treat, so prevention of ACD is an important public health challenge.
- National and international regulatory authorities require testing of pesticides, personal care products, and other chemical products to assess their potential to cause ACD. The results of these tests are used to determine appropriate labeling for safe use and handling.
- The National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) evaluates alternative test methods for assessing the potential of chemicals to cause ACD.

**Figure 1** ACD Rash on Human Skin



## NICEATM Skin Sensitization Databases

- NICEATM is collecting and curating high-quality *in vivo* data in endpoint-specific databases. These data will provide bases for development of mechanistic models of skin sensitization and other toxicities. These models will in turn facilitate validation of relevant *in vitro* and *in silico* approaches that may replace animal tests.

- The NICEATM skin sensitization databases include *in vivo* data from:
  - **Mice:** The LLNA is the current preferred animal test, and data from the LLNA represent the broadest and highest-quality dataset available to support development of assays that can eventually replace animal tests. The LLNA database is available on the NTP website at <http://ntp.niehs.nih.gov/go/40498>.
  - **Humans:** Data from the human repeat insult patch test and maximization patch test are currently under curation. These data will better support relevance of *in vitro* and *in silico* approaches to human risk assessment.

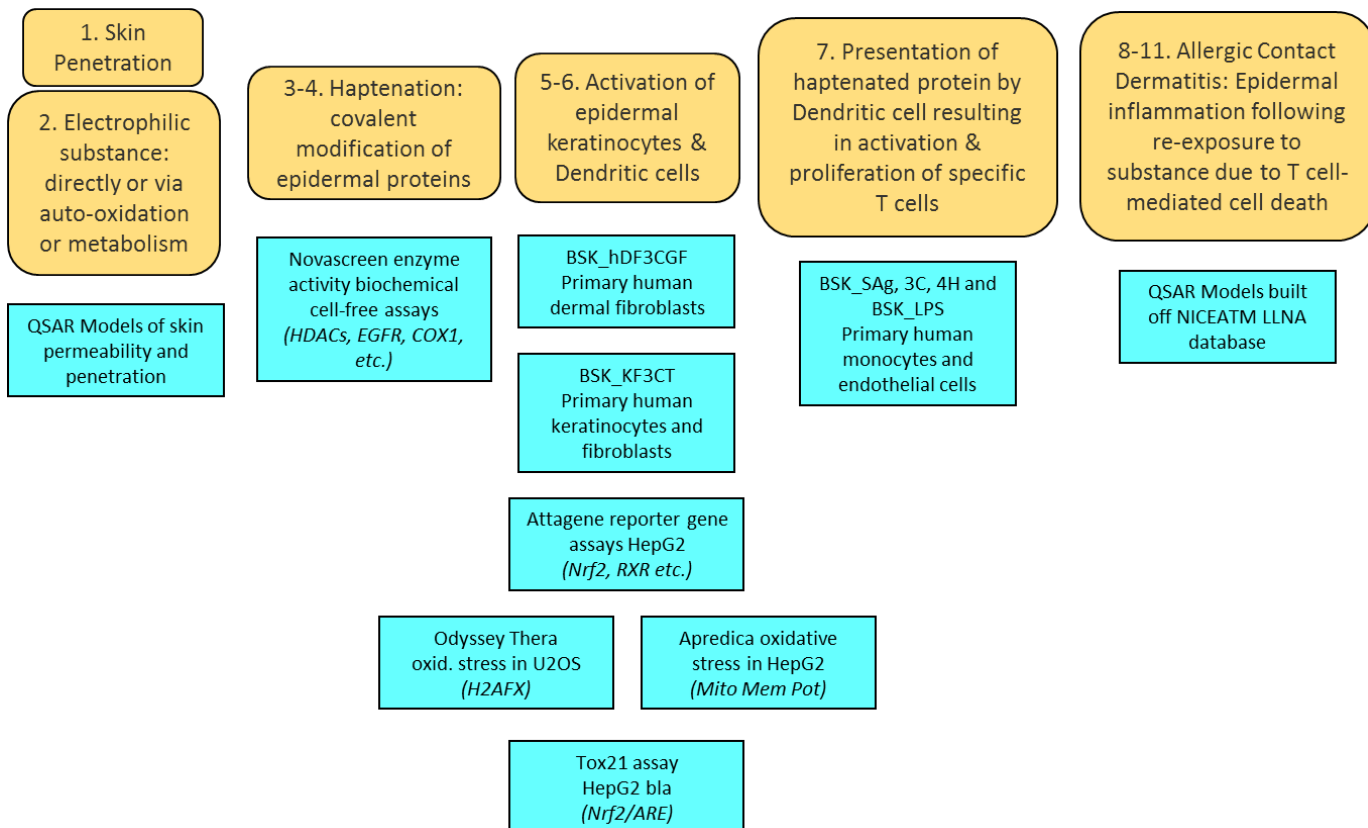
**Table 1 NICEATM *In Vivo* Skin Sensitization Databases Currently Under Development**

<b>Database</b>	<b>Number of chemicals</b>	<b>Number of studies per chemical [range]</b>	<b>% Positive (any study)</b>	<b>ToxCast chemical overlap</b>
<b>LLNA</b>	668	[1-43]	65%	60
<b>Human</b>	127	[1-8]	59%	24

## Alignment of ToxCast Assays with the Skin Sensitization Adverse Outcome Pathway

- The OECD report “The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins” (OECD 2012) identified the key events that occur after exposure to certain types of skin sensitizers that result in the development of ACD (, yellow boxes).
  - The AOP begins with penetration of the sensitizing substance into the viable skin layers (**Figure 2**, steps 1–2).
  - The sensitizing substance modifies skin proteins (**Figure 2**, steps 3–4). Keratinocytes, fibroblasts and dendritic cells produce reactive oxygen species and inflammatory mediators in response. This is followed by uptake, processing, and presentation of modified protein (antigen) by dendritic cells (**Figure 2**, steps 5–6).
  - The dendritic cells migrate to the local lymph nodes (**Figure 2**, step 7).
  - Antigen is presented to specific naïve T-cells, causing T-cell proliferation and differentiation (**Figure 2**, step 7).
  - If re-exposure occurs and the population of antigen-specific memory T cells has reached a critical number, the T cells mediate an elicitation response at the site of re-exposure, observed clinically in humans as ACD (**Figure 2**, steps 8–11).
- The ToxCast assay portfolio includes assays that use these primary human skin cell types. These and additional ToxCast assays measure many of these oxidative stress and inflammatory signaling processes.

**Figure 2 Preliminary Mapping of *In Vitro* Assays and *In Silico* Models to the Skin Sensitization AOP**

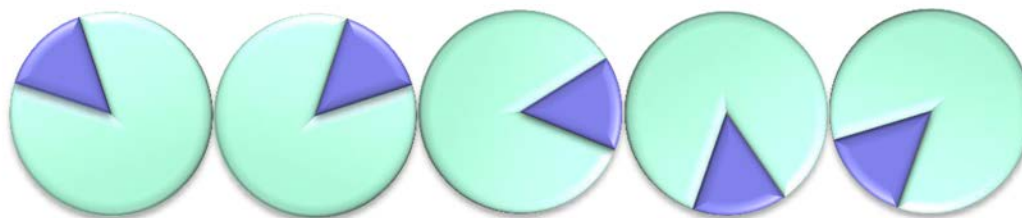


- **Figure 2** shows how ToxCast *in vitro* assays map preliminarily to the skin sensitization AOP based on the known biological relevance of the assay targets. Quantitative structure activity relationship (QSAR) models have also been mapped to the AOP (steps 1, 2, and 8–11). For example, a QSAR model for predicting LLNA results was built using an earlier version of the NICEATM database as a training set (Tropsha et al. 2014). Collaborations are ongoing to incorporate more data and refine the models.
- The current project used data from 60 chemicals screened in ToxCast assays to build a model that would predict the LLNA results for those chemicals in the NICEATM LLNA database (**Table 1**). The ToxCast assay suite included targets that mapped to the skin sensitization AOP, shown in **Figure 2**, as well as those not yet known to be biologically relevant.

## Developing a Model to Use ToxCast Data to Predict LLNA Results

- ToxCast assay data from 60 chemicals (36 LLNA positives and 24 LLNA negatives) were used to build a model to predict LLNA results for chemicals in the NICEATM LLNA database (**Table 1**).
  - A random forest (RF) model was created using the ToxCast assay results as descriptors and LLNA results as an endpoint.
    - RF is an ensemble machine learning technique based on randomized decision trees (Breiman 2001). The outputs of all trees are aggregated to obtain one final prediction.
    - Each tree is grown as follows:
      - (i) A bootstrap sample is performed on the entire set of N compounds to form a training set for the current tree. The compounds omitted from the training set are placed in the out-of-bag (OOB) set (size  $\sim N/3$ ).
      - (ii) The best split among the randomly selected descriptors from the entire pool at each node is chosen.
      - (iii) Each tree is grown to the largest possible extent without pruning. The predicted classification values are defined by majority voting, and each tree predicts values for only those compounds in the OOB set.
- The final model is chosen by the lowest error for prediction of the OOB set.
- Five-fold cross-validation was performed by dividing the data independently into training sets (80% of data) and test sets (20% of data) five times, as illustrated in **Figure 3**.

**Figure 3** Five-Fold Cross-Validation of RF Model



## Results

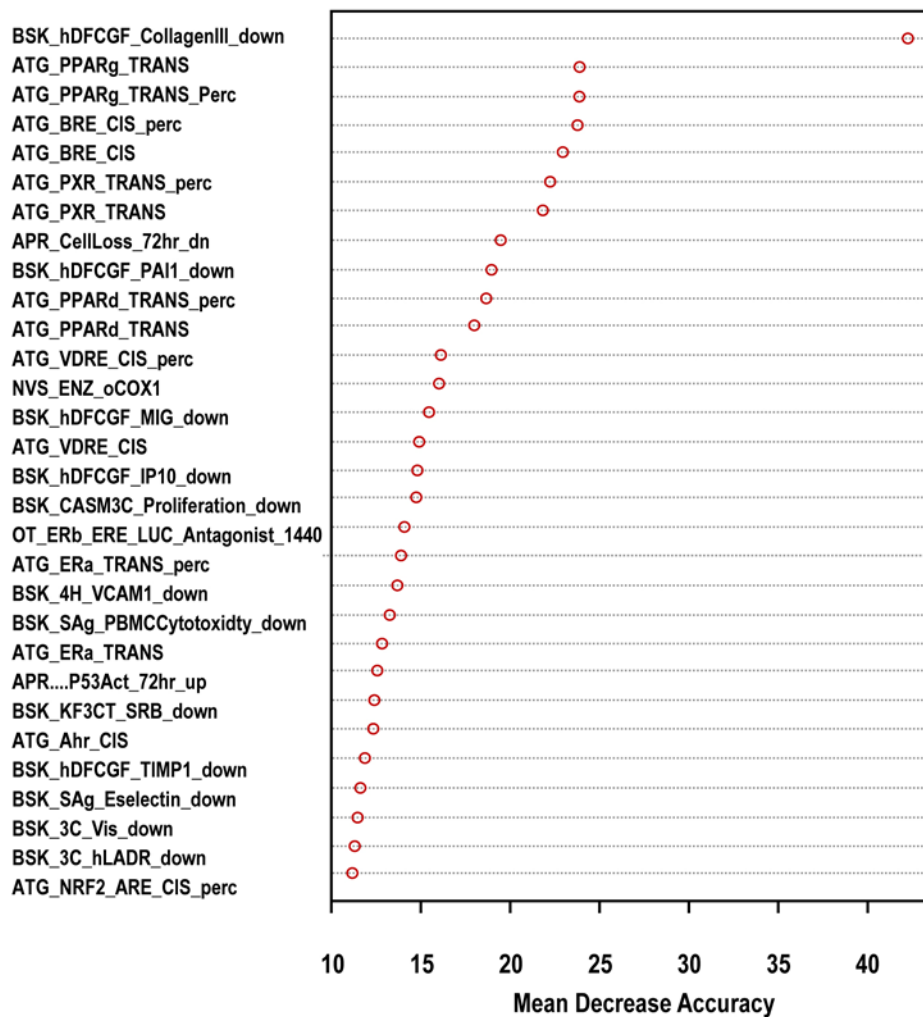
- Each iteration produced a random forest model with features (ToxCast assays) ranked by variable importance in predicting the LLNA results.
- The cross-validated model produced an average specificity of 73% and an average sensitivity of 88%, with a negative predictive value of 69% and a positive predictive value of 92%.
- To build the final model and to avoid overfitting due to dimensionality concerns, the top 100 features across all model iterations based on their variable importance scores were used. The final model was retrained on the entire set of 60 chemicals with ToxCast and LLNA data.
- The final OOB error estimate was 20% (12/60), representing the accuracy of the model (80% [48/60]) against all chemicals when they appear in the external test sets. **Table 2** shows the confusion matrix and the correct classification rates for sensitizers and nonsensitizers.

**Table 2 Confusion Matrix for Final RF Model: External Test Set Performance**

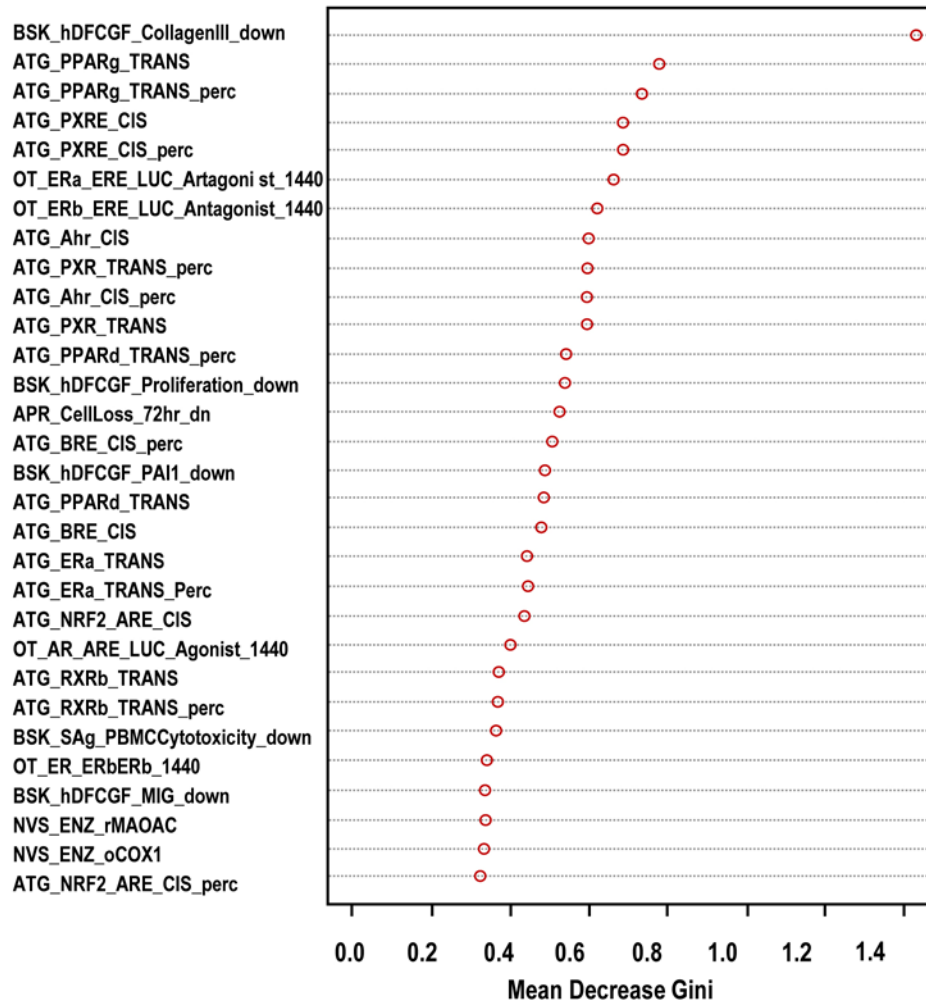
	Nonsensitizer	Sensitizer	Classification Error
Nonsensitizer	20	4	0.17
Sensitizer	8	28	0.22

**Figure 4 RF Model: Variable Importance Plot**

- The graphs below show the 30 most important ToxCast assays for predicting LLNA results from the final RF model. Assays are ranked by the mean decrease in accuracy resulting when that feature is removed (upper plot), and the mean decrease in Gini score, a measure of node purity and feature relevance (lower plot).

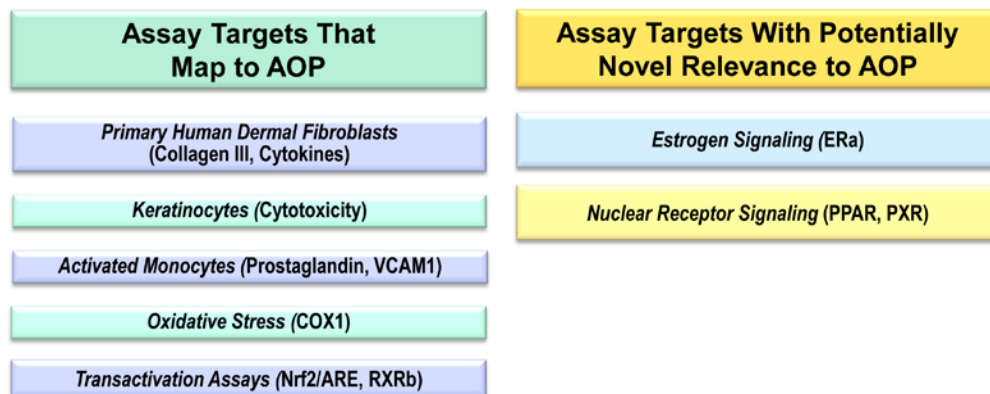






## Discussion

**Figure 5 Summary of Most Important ToxCast Assays for Predicting LLNA Results**



Abbreviations: AOP = adverse outcome pathway; LLNA = murine local lymph node assay; oxid. = oxidative; QSAR = quantitative structure–activity relationship.

### ***Most Important RF Model Features Map to AOP***

- Despite the limited overlap (n=60 chemicals) between the ToxCast chemical library and the NICEATM LLNA database, the unsupervised modeling process identified a number of ToxCast assays (**Figure 4**) that map to the skin sensitization AOP based on biological relevance (**Figure 5**, left panel, and **Figure 2**).
  - **Assays using primary human dermal fibroblasts:**
    - Fibroblast proliferation is controlled by lymphocyte signaling and is affected by specific antigenic challenges.
    - Fibroblast proliferation is involved in oxidative stress signaling in skin sensitization (Wondrak et al. 2003).
    - Collagen in chronically inflamed tissue has altered biochemical characteristics and functions (Hirota et al. 2003).
  - **Assays using keratinocytes:**
    - Keratinocytes contribute to skin sensitization by producing cytokines, for example interleukins and TNF-alpha, in response to cellular injury or cytotoxicity.
    - Various chemokines are elevated in atopic dermatitis via enhanced production by keratinocytes and self-perpetuating inflammatory mechanisms.
  - **Assays using activated monocytes:**
    - Monocyte signaling regulates activation and proliferation of T cells.

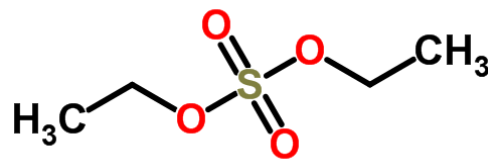
- Endothelial adhesion molecules such as VCAM1 can be upregulated by CD40L on activated T cells and are critical for memory T-cell infiltration.
- Prostaglandin signaling promotes allergic skin inflammation in response to cutaneous exposure to antigen via T-cell receptor-mediated prostaglandin-responsive chemotaxis (He et al. 2010).
- **Oxidative stress/transactivation assays:**
  - Oxidative stress activates transcription factors and signaling pathways, including NF- $\kappa$ B and p38 MAPK, leading to the release of cytokines and chemokines.
  - Reactive oxygen species serve as messengers mediating cellular responses and resulting in immune cell activation (Corsini et al. 2013).

***The Model Identified Some Potential Novel Targets***

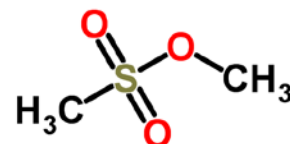
- A number of novel targets were identified that do not map to the current AOP (**Figure 5**, right panel).
  - **Estrogen signaling assays:**
    - Estrogens affect on skin physiology and pathophysiology; human skin fibroblasts express estrogen receptors alpha and beta (Haczynski et al. 2002).
  - **Nuclear receptor signaling assays:**
    - Human and rodent skin cells express all three PPAR isotypes, which play an important role in inflammatory responses and signaling (Sertznig and Reichrath 2011).
    - PXR, expressed in skin and especially in proliferating keratinocytes, is known to regulate oxidative stress and to control cell proliferation.

***Chemicals That Were Not Correctly Classified  
by the Final Model***

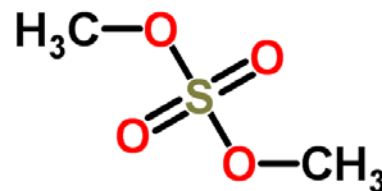
- **Diethyl sulfate** is used as an ethylating agent and chemical intermediate. According to LLNA data, it would be classified as a moderate sensitizer. (ICCVAM 2011)



- **Methyl methanesulfonate** is an alkylating agent that would be classified as a moderate sensitizer according to LLNA data (ICCVAM 2011).



- **Dimethyl sulfate** is a methylating agent used in the manufacture of dyes and perfumes. According to LLNA data, it would be classified as a strong sensitizer. (ICCVAM 2011)



These compounds were largely inactive across the ToxCast assays and were therefore incorrectly predicted as nonsensitizers by the model. The chemicals are structurally similar and hybrid models that incorporate molecular descriptors may assist in identifying them as sensitizers.

## Conclusions

- The AOP for skin sensitization has been very well characterized and thus provides opportunities to identify *in silico* models and *in vitro* assays that cover the relevant biology associated with this common human hazard.
- Well-curated reference databases combined with large HTS datasets allow for model building to develop efficient testing strategies that minimize the use of animals in regulatory testing.
- We developed a model that identified assays relevant to known AOP targets as well as targets outside of the current AOP. The model predicts LLNA results with 80% accuracy.
- Future goals include:
  - Collection and curation of additional *in vivo* and *in vitro* data
  - Developing multiclass and continuous models to predict sensitization potential and potency
  - Using human data to develop models to support the design of a testing strategy that will accurately predict clinical responses
  - Collaborating with QSAR researchers to develop hybrid models that incorporate structural descriptors and *in vitro* data for improved predictivity

## **References**

- Breiman L. 2001. *Machine Learning* 45: 5–32.
- Corsini E, Galbiati V, Nikitovic D, Tsatsakis AM. *Food Chem Toxicol* 61: 74–81.
- Haczynski J, Tarkowski R, Jarzabek K, et al. 2002. *Int J Mol Med* 10: 149–153.
- He R, Oyoshi MK, Wang JY, Hodge MR, Jin H, Geha RS. 2010. *J Allergy Clin Immunol* 126: 784–790.
- Hirota A, Ebihara T, Kusubata M, et al. 2003. *J Invest Dermatol* 121: 1317–1325.
- ICCVAM. 2011. ICCVAM Test Method Evaluation Report: Usefulness and Limitations of the Murine Local Lymph Node Assay for Potency Categorization of Chemicals Causing Allergic Contact Dermatitis in Humans. NIH Publication No. 11-7709. Research Triangle Park, NC: National Institute of Environmental Health Sciences.
- OECD. 2012. OECD Series on Testing and Assessment No. 168. The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins. Part 1: Scientific Assessment. Paris:OECD Publishing. Available: <http://www.oecd.org/env/ehs/testing/adverse-outcome-pathways-molecular-screening-and-toxicogenomics.htm> [accessed 2 Dec 2013]
- Sertznig P, Reichrath J. 2011. *Dermatoendocrinol* 3: 130–135.
- Tropsha A, Alves V, Muratov E, Fourches D, Strickland J, Kleinstreuer N, Andrade C. 2014. *Toxicologist* 138: 571. Abstract 2182c.
- Wondrak GT, Roberts MJ, Cervantes-Laurean D, Jacobson MK, Jacobson EL. 2003. *J Invest Dermatol* 121: 578–586.

## **Acknowledgements**

The Intramural Research Program of the National Institute of Environmental Health Sciences (NIEHS) supported this poster. Technical support was provided by ILS under NIEHS contracts N01-ES 35504 and HHSN27320140003C.

The views expressed above do not necessarily represent the official positions of any Federal agency. Since the poster was written as part of the official duties of the authors, it can be freely copied.