Predicting Skin Sensitization Using 21st Century Toxicology

N Kleinstreuer[1], J Strickland[1], W Casey[2]

[1]ILS, RTP, NC, USA; [2]NICEATM/DNTP/NIEHS/NIH/HHS, RTP, NC, USA

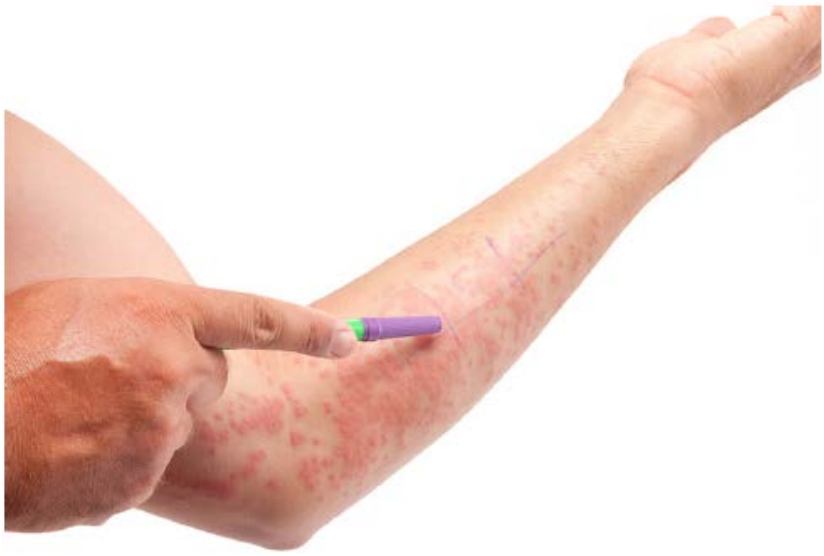
Abstract

Allergic contact dermatitis (ACD) is an adverse health effect that develops after repeated exposure to skin-sensitizing chemicals and products. To minimize the occurrence of ACD, regulatory authorities require testing using assays such as 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 initiated by covalent binding to proteins. In an effort to reduce or replace animal use, OECD is also pursuing the development of integrated testing strategies using novel *in vitro* and *in silico* approaches. The U.S. Tox21 and ToxCast projects include high-throughput screening (HTS) assays that map to key events in the skin sensitization AOP (e.g., oxidative stress, cytokine expression) from which data on hundreds of potential skin sensitizers have been generated. We built a cross-validated random forest model using ToxCast Phase II data and a balanced training set of 60 chemicals. The model predicted LLNA results with 80% accuracy. The assays with highest variable importance in the random forest model 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). Well-characterized AOPs like skin sensitization provide opportunities to use HTS data to develop efficient testing strategies that minimize the use of animals 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.
    - NICEATM supported the development and evaluation of the murine local lymph node assay (LLNA) and reduced LLNA test methods, which reduced the number of animals required for this testing (ICCVAM 1999, 2009).
    - NICEATM is currently developing integrated testing strategies incorporating *in vitro* assays and *in silico* models. These approaches could further reduce and eventually eliminate animal use.

Figure 1 ACD Rash on Human Skin



NICEATM Skin Sensitization Databases

- NICEATM is collecting and curating high-quality *in vivo* data and amassing them into 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 Overview of NICEATM *In Vivo* Skin Sensitization Databases Currently Under Development

Database	Number of chemicals	Number of studies per chemical [range]	% Positive (any study)	Tox21 chemical overlap	ToxCast chemical overlap
LLNA	668	[1-43]	65%	273	60
Human	127	[1-8]	59%	79	24

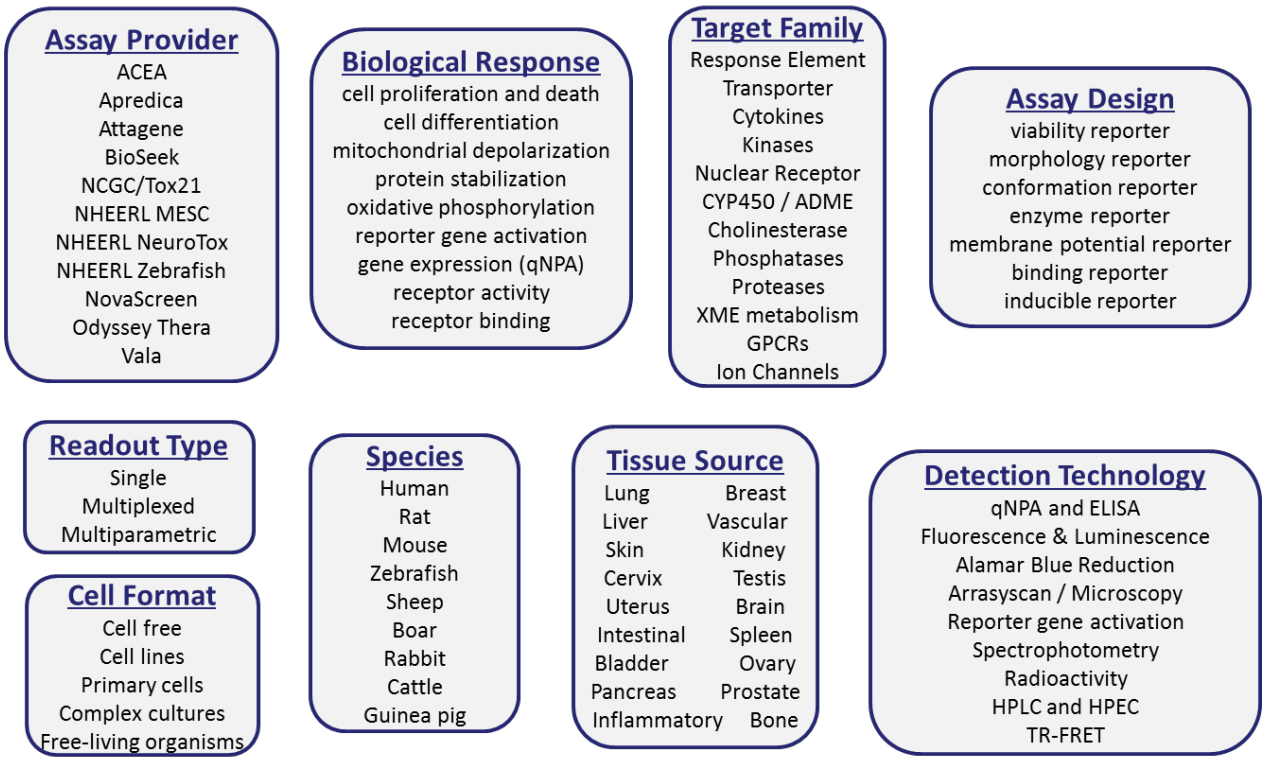
U.S. Programs for High-Throughput Chemical Testing

- The U.S. Tox21 and ToxCast projects include high-throughput screening assays that may be relevant to skin sensitization. Data on hundreds of potential skin sensitizers have been generated.
  - The 10,000 chemicals in the Tox21 library have been screened using all Tox21 assays, including assays measuring induction of IL-8 and TNFalpha. These data are still undergoing analysis (Tice et al. 2013).

  - The 1047 chemicals in the ToxCast Phase I and II libraries have been screened using over 700 ToxCast assays, with a top testing concentration of 100-200 µM depending on the system.

- The ToxCast assay portfolio, outlined in **Figure 2,** includes cell-free biochemical assays, small model organisms, and human primary cells such as skin cells and monocytes (Kavlock et al. 2012).

Figure 2 Overview of the ToxCast Assay Portfolio



(http://actor.epa.gov/actor/faces/ToxCastDB)

The Skin Sensitization Adverse Outcome Pathway

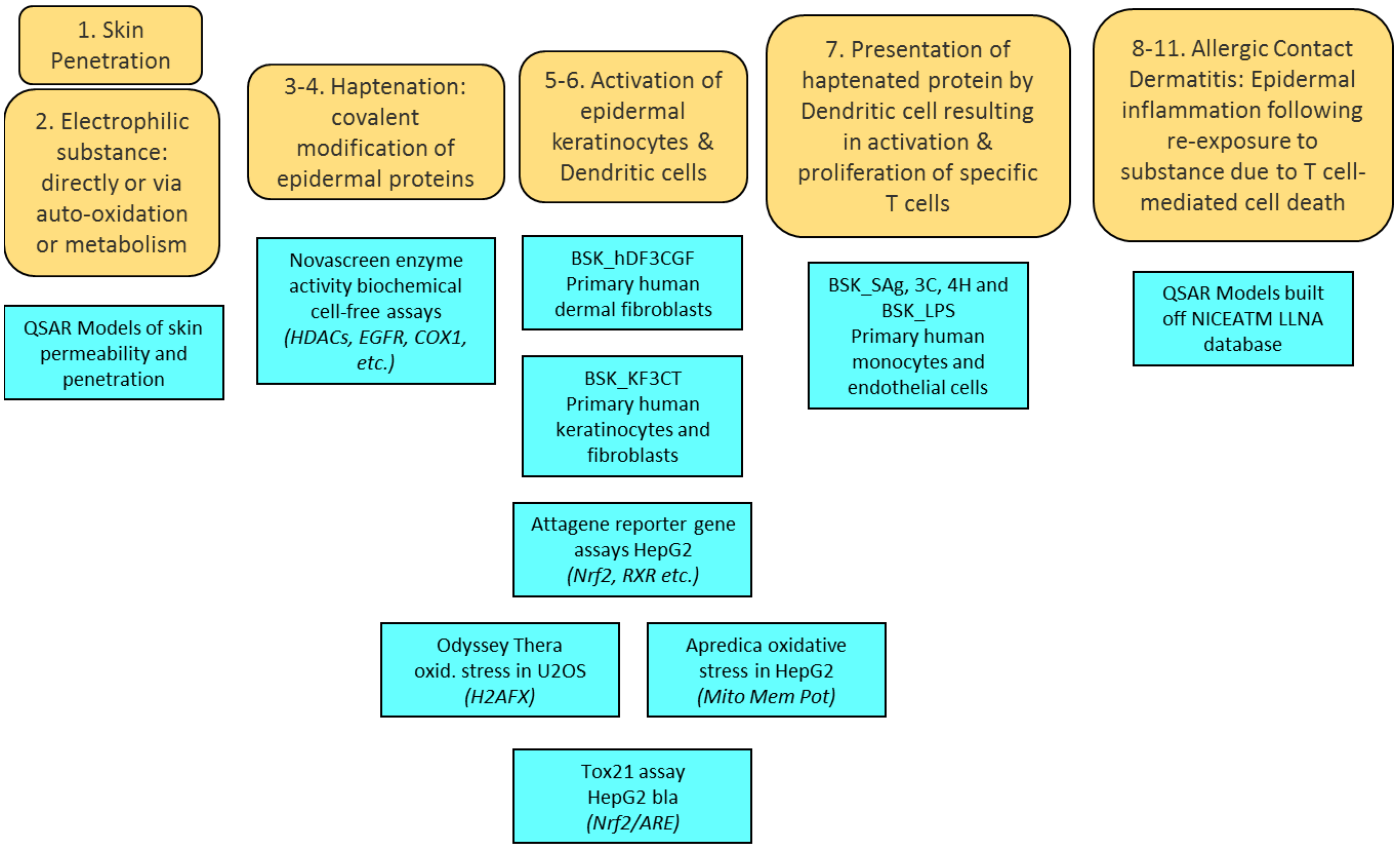
- The OECD report "The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins" (OECD 2012) identified the key events that occur after exposure to certain types of skin sensitizers that result in the development of ACD (**Figure 3**, yellow boxes).

  - The adverse outcome pathway (AOP) begins with penetration of the sensitizing substance into the viable skin layers (**Figure 3,** steps 1–2).

- The sensitizing substance modifies skin proteins (**Figure 3,** 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 3,** steps 5–6).
- The dendritic cells migrate to the local lymph nodes (**Figure 3,** step 7).
- Antigen is presented to specific naïve T-cells, causing T-cell proliferation and differentiation (**Figure 3,** 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 3,** 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 3 **Preliminary Mapping of *In Vitro* Assays and *In Silico* Models to the Skin Sensitization AOP**



- **Figure 3** shows how ToxCast and Tox21 *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. Collaborations are ongoing to incorporate more data and refine the models.
- The current project used data from 60 chemicals (36 LLNA positives and 24 LLNA negatives) 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 3**, as well as those not yet known to be biologically relevant.

Data Analysis

- We created a random forest (RF) model using the ToxCast assay results as descriptors and LLNA results as an endpoint for a set of 60 chemicals (36 skin sensitizers and 24 non-sensitizers).
  - 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:

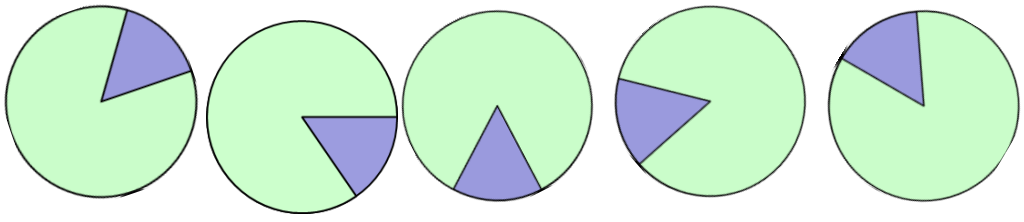
    (i) A bootstrap sample is performed on the entire set of N compounds to form a training set for the current tree. The compounds omitted from the training set are placed in the out-of-bag (OOB) set (size ~ N/3).

    (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 4.**

Figure 4 Five-Fold Cross-Validation of RF Model



- Each iteration produced a random forest model with features (ToxCast assays) ranked by variable importance in predicting the LLNA results.
- **Table 2** shows average statistics across model iterations**.**

Table 2 Five-Fold Cross-Validation of RF Model Predicting LLNA Results with ToxCast Data Using 60 Compounds

Model Run	Sensitivity	Specificity	PPV	NPV	BA
1	0.86	0.60	0.75	0.75	0.73
2	0.63	1.00	1.00	0.57	0.81
3	0.57	1.00	1.00	0.63	0.79
4	0.71	1.00	1.00	0.71	0.86
5	0.86	0.80	0.86	0.80	0.83
AVG	**0.73**	**0.88**	**0.92**	**0.69**	**0.80**

Abbreviations: BA = balanced accuracy; NPV = negative predictive value; PPV = positive predictive value.

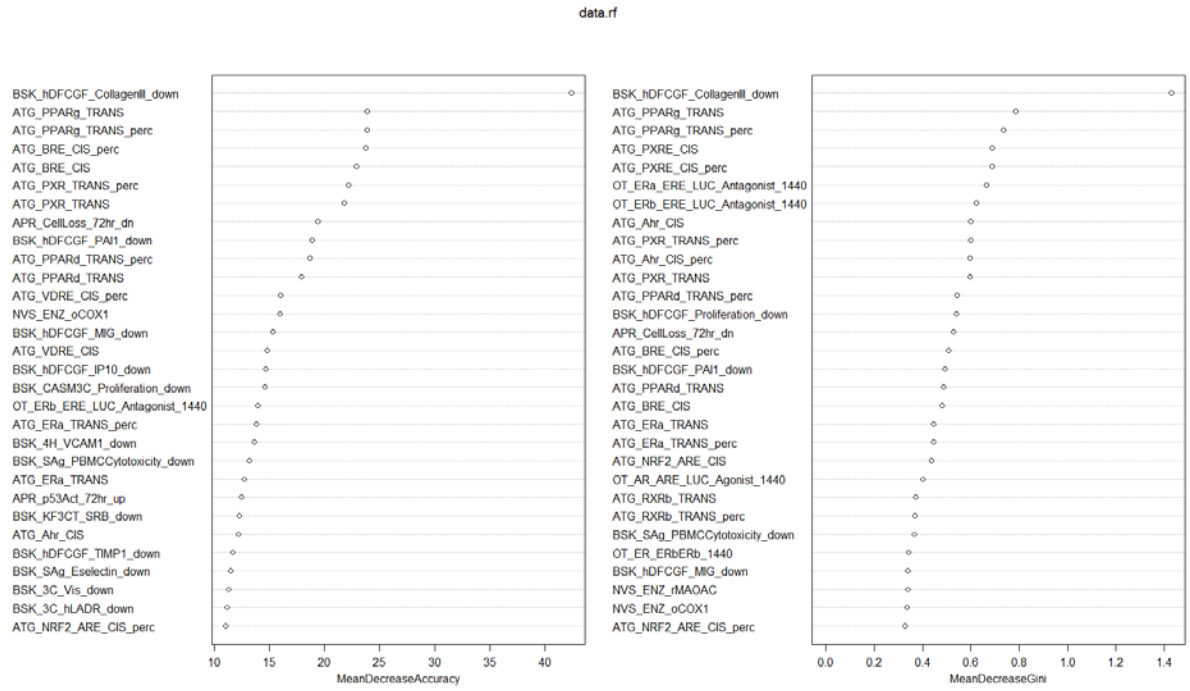
- To build the final model and to avoid overfitting due to dimensionality concerns, the top 100 features across all model iterations based on their variable importance scores were used. The final model was retrained on the entire set of 60 chemicals with ToxCast and LLNA data.
- The OOB error estimate was 20%, representing the accuracy of the model against all chemicals when they appear in the external test sets. **Table 3** shows the confusion matrix and the correct classification rates for sensitizers and nonsensitizers.

Table 3 Confusion Matrix for Final RF Model Showing Performance on External Test Sets

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

- The most important ToxCast assays for predicting LLNA results from the final RF model are shown in **Figure 5**. They are ranked by two different measures of their variable importance: the mean decrease in accuracy resulting when that feature is removed (left plot), and the mean decrease in Gini score, a measure of node purity and feature relevance (right plot).

Figure 5 RF Model: Variable Importance Plot



- The ToxCast assay with the highest variable importance in the model, by both statistical measures, was Collagen Type 3 (Col III) regulation in primary human dermal fibroblasts (BSK_hDFCGF_CollagenIII_down, the first assay listed in each plot).

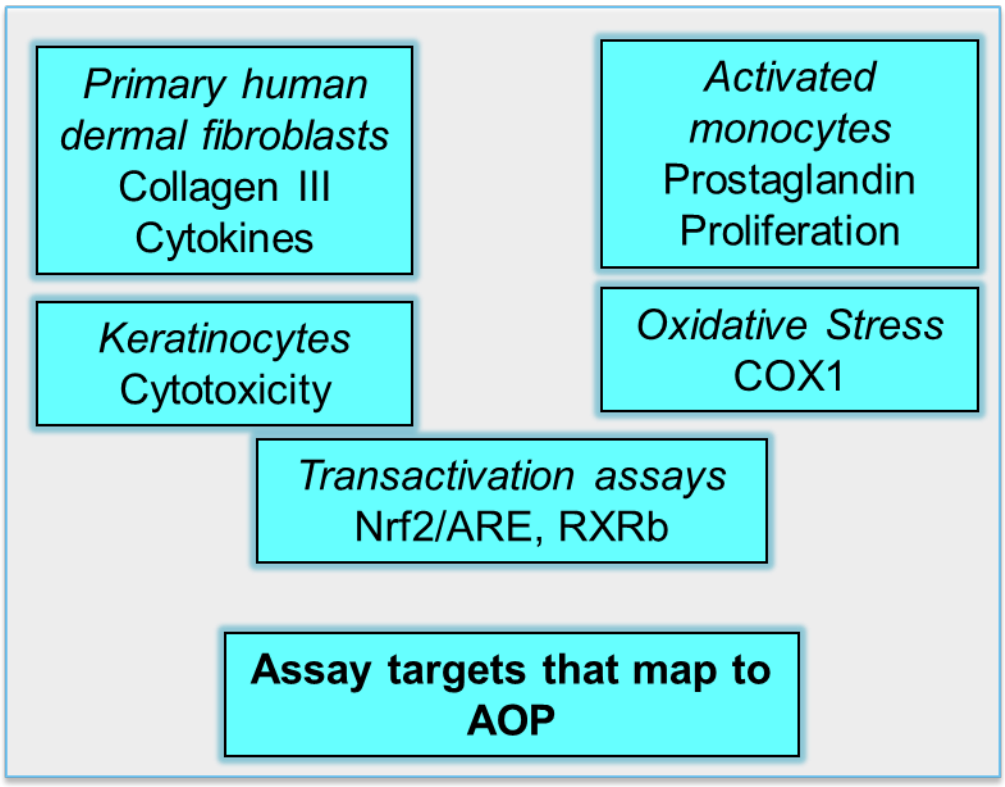
- Out of the 36 positive compounds, 23 produced at least 2-fold downregulation of Col III expression in this assay, while only 3 of the negative compounds showed this activity.
- When retrained on the entire data set and applied to make predictions against the training set, the RF model using data from the assays shown in **Figure 5** had an accuracy of 95%. The model failed to correctly classify three chemicals: diethyl sulfate, methyl methanesulfonate, and dimethyl sulfate.

Discussion

Most Important RF Model Features Map to AOP

- An RF modeling approach was taken to identify assays in the ToxCast program that may be predictive of skin sensitization results in the LLNA assay. Despite the limited overlap between the ToxCast chemical library and the NICEATM LLNA database (n=60chemicals), the unsupervised modeling process identified a number of ToxCast assays that had already been mapped to the skin sensitization AOP based on biological relevance (**Figure 6**).

Figure 6 Targets With Potential Biological Relevance to Skin Sensitization AOP



- **Primary human dermal fibroblasts:**

  - Collagen production and fibroblast proliferation are controlled by lymphocyte signaling and are known to change under specific antigenic challenges.
  - Fibroblast proliferation and extracellular matrix protein production are involved in oxidative stress signaling in skin sensitization processes (Wondrak et al. 2003).
  - Collagen in chronically inflamed tissue has altered biochemical characteristics and functions, which may impact the pathogenesis of chronic dermatitis (Hirota et al. 2003).
  - Human dermal fibroblasts are components of various 2-D and 3-D *in vitro* culture systems under development for assessing skin sensitization.

- **Keratinocytes:**
  - Various chemokines are elevated in atopic dermatitis via enhanced production by keratinocytes and self-perpetuating inflammatory mechanisms.
  - Keratinocyte actively contribute to the pathogenesis of skin sensitization by producing an array of cytokines including interleukins and TNFalpha, often in response to cellular injury or cytotoxicity.

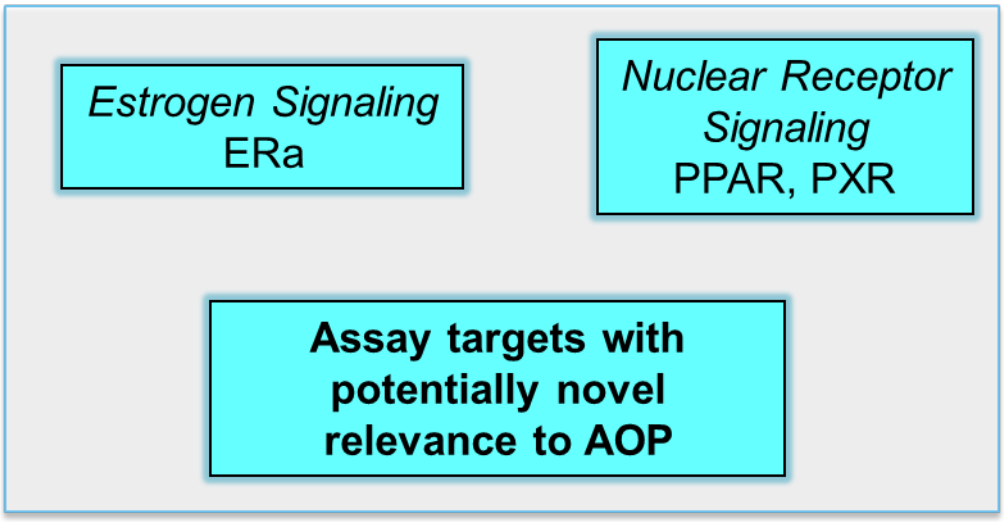
- **Activated monocytes:**
  - Monocyte signaling regulates activation and proliferation of T cells, a key event in the AOP.
  - Endothelial adhesion molecules such as VCAM1 can be upregulated by CD40L on activated T cells and are critical for memory T-cell infiltration.
  - Cutaneous prostaglandin signaling was shown to promote allergic skin inflammation in response to cutaneous exposure to antigen in previously sensitized mice via the T-cell receptor mediated prostaglandin responsive chemotaxis (He at al. 2010).

- **Oxidative stress/transactivation assays:** Oxidative stress leads to the activation of transcription factors and signaling pathways, including NF-kB and p38 MAPK, which leads to the release of cytokines and chemokines. Reactive oxygen species serve as essential second messengers mediating cellular responses and resulting in immune cell activation (Corsini et al. 2013).

The Model Identified Some Potential Novel Targets

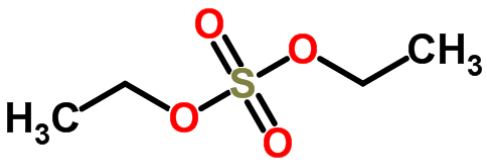
- A number of novel targets were identified that do not map to the current AOP (**Figure 7**).

Figure 7 Targets Not Known to be Biologically Relevant to AOP

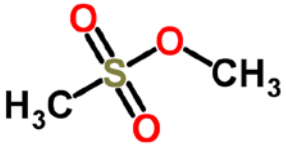


- **Estrogen signaling:** Estrogens have significant effects on skin physiology and pathophysiology, and human skin fibroblasts express estrogen receptor alpha and beta (Haczynski et al. 2002).
- **Nuclear receptor signaling:**
  - All three PPAR isotypes are expressed in rodent and human skin and play an important role in inflammatory responses and signaling of keratinocytes and other skin cells (Sertznig and Reichrath 2011).
  - PXR is known to regulate oxidative stress and to control cell proliferation. PXR is expressed in skin, and in highest levels in proliferating keratinocytes.

Chemicals That Were Not Correctly Classified by the Final Model

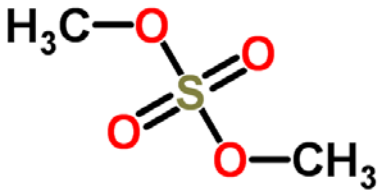


- **Diethyl sulfate** is used as an ethylating agent and as a chemical intermediate. According to LLNA data, it would be classified as a moderate sensitizer. (ICCVAM 2011)



- **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. 1999. The Murine Local Lymph Node Assay: A Test Method for Assessing the Allergic Contact Dermatitis Potential of Chemicals/Compounds. The Results of an Independent Peer Review Evaluation Coordinated by ICCVAM and NICEATM. NIH Publication No. 99-4494. Research Triangle Park, NC: National Institute of Environmental Health Sciences.

ICCVAM. 2009. ICCVAM Test Method Evaluation Report. The Reduced Murine Local Lymph Node Assay: An Alternative Test Method Using Fewer Animals to Assess the Allergic Contact Dermatitis Potential of Chemicals and Products. NIH Publication No. 09-6439. Research Triangle Park, NC: National Institute of Environmental Health Sciences.

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.

Kavlock R, Chandler K, Houck K, et al. 2012. Chem Res Toxicol 25: 1287–1302.

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.

Tice RR, Austin CP, Kavlock RJ, Bucher JR. 2013. Environ Health Perspect 121: 756–765.

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.