



Development of an *in vitro* test to identify respiratory sensitizers in bronchial epithelial cells using gene expression profiling

Sander Dik^{a,b}, Jeroen L.A. Pennings^a, Henk van Loveren^{a,b}, Janine Ezendam^{a,*}

^a Centre for Health Protection, National Institute for Public Health and the Environment, 3720 BA Bilthoven, The Netherlands

^b Department of Toxicogenomics, Maastricht University, 6200 MD Maastricht, The Netherlands

ARTICLE INFO

Article history:

Received 19 August 2015

Received in revised form 21 October 2015

Accepted 24 October 2015

Available online 28 October 2015

Keywords:

Respiratory sensitization

Occupational asthma

Epithelial cells

Gene expression

Pulmonary barrier function

ABSTRACT

Chemicals that induce asthma at the workplace are substances of concern. At present, there are no widely accepted methods to identify respiratory sensitizers, and classification of these substances is based on human occupational data. Several studies have contributed to understanding the mechanisms involved in respiratory sensitization, although uncertainties remain. One point of interest for respiratory sensitization is the reaction of the epithelial lung barrier to respiratory sensitizers. To elucidate potential molecular effects of exposure of the epithelial lung barrier, a gene expression profile was created based on a DNA microarray experiment using the bronchial epithelial cell line 16HBE14o[−]. The cells were exposed to 12 respiratory sensitizers and 10 non-sensitizers. For statistical analysis, we used a class prediction approach that combined three machine learning algorithms, leave-one-compound-out cross validation, and majority voting per tested compound. This approach allowed for a prediction accuracy of 95%. Identified predictive genes were mainly associated with the cytoskeleton and barrier function of the epithelial cell. Several of these genes were reported to be associated with asthma as well. Taken together, this indicates that pulmonary barrier function is an important target for respiratory sensitizers and associated genes can be used to predict the respiratory sensitization potential of chemicals.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Asthma induced by low molecular weight (LMW) chemicals (Sastre et al., 2003) at the workplace is a common type of allergy, as 15–20% of all asthma cases are due to this type of exposure. Occupational asthma is described for several work environments, such as hospitals (Vellore et al., 2006) and hair salons (Moscato et al., 2005), and the socioeconomic- and medical impact is significant (Mapp et al., 2005; Toren and Blanc, 2009). There are currently no widely accepted tests available to identify respiratory sensitizers. This is partly due to the absence of validated predictive *in vitro* and *in vivo* tests. This can be attributed to the lack of mechanistic insight in the pathways involved.

By reviewing all relevant data regarding the mechanism of respiratory sensitization, Kimber et al. (2014) aimed to apply the general principles of modeling an Adverse Outcome Pathway (AOP) for chemical respiratory allergy. The molecular initiating event of LMW respiratory allergy is covalent binding of the chemical to a protein after exposure, forming a hapten that can be recognized by dendritic cells (Kimber et al., 2014). There are two types of methods currently available to determine the chemical reactivity of LMW chemicals to covalently bind proteins. The first are (quantitative) structure activity relationship

(QSAR) models using chemical reactivity alerts associated with protein binding and respiratory sensitization. A study implementing QSAR models to predict respiratory sensitization was performed by Dik et al. (2014). The second method is the direct peptide reactivity assay (DPRA) for respiratory sensitizers (Lalko et al., 2012), which correlates peptide depletion by a chemical with the reactivity of the chemical.

The cellular response to a hapten in case of a respiratory sensitizer is similar to skin sensitization (OECD, 2012); dendritic cells recognize the hapten and start to produce inflammatory cytokines (Kimber et al., 2014). In case of respiratory sensitization, this usually leads to the initiation of Th2-type responses (individual references are evaluated by (Kimber et al., 2014)). The Th2-response then causes sensitization through an IgE mediated mechanism. The involvement of IgE in the sensitization process remains unclear however, as some patients have detectable IgE antibodies specific for the chemical, whereas other patients do not (individual references are evaluated by Kimber et al. (2014)). Although much research has focused on the process of respiratory sensitization, several biological aspects are still not (fully) understood.

Next to the cellular response of dendritic cells, cellular responses in keratinocytes are identified as a key event in the AOP for skin sensitization (OECD, 2012). Keratinocytes are considered essential for the generation of danger signal, necessary for the adaptive immune response towards skin sensitizers (Natsch, 2010). In case of respiratory sensitization, once a respiratory sensitizer is inhaled, it comes into contact with

* Corresponding author at: National Institute for Public Health and the Environment (RIVM), PO Box 1, NL-3720BA Bilthoven, The Netherlands.
E-mail address: janine.ezendam@rivm.nl (J. Ezendam).

the pulmonary epithelium. Holgate (2008) found that airway epithelium is fundamentally abnormal in asthma and suggested that the barrier function of the epithelium is compromised through defective tight junction formation. Studies that are more recent also show the importance of pulmonary epithelial cells and their barrier function in relation to asthma (Georas and Rezaee, 2014; Rezaee and Georas, 2014). Therefore, lung epithelial cells are considered to be an important target of *in vitro* research. Verstraelen et al. (2008a) studied the applicability of different epithelial cell types as models for the identification of respiratory sensitizers. Among these are bronchial cells (for example BEAS-2B cells) or alveolar cells (for example A549 cells). Both cell lines have already been used to identify biomarkers associated with respiratory sensitizers (Verstraelen et al., 2009a, 2009b). These cells have various benefits compared to other cell lines, as discussed by Verstraelen et al. (2008a). However, these cells do not or only slightly have the ability to form tight junctions. As a decreased barrier function is an important aspect of asthma, we selected a cell line capable of forming proper tight junctions. The 16HBE14o[−] cell line was considered appropriate in this regard as it has the ability to form tight junctions and a proper barrier as measured by the trans-epithelial electrical resistance (Cozens et al., 1994).

The aim of this study was to evaluate the predictive performance of a set of gene profiles to identify respiratory sensitizers. Additionally, these gene profiles were evaluated to identify molecular pathways induced by respiratory sensitizers to elucidate the role of bronchial epithelial cells in chemical respiratory sensitization.

2. Methods

2.1. Cell culture

The human bronchial epithelial cell line 16HBE14o[−] was obtained as a kind gift from Dr. Gruenert (National Institutes of Health [NIH], San Francisco, CA). Cells were grown to 80–90% confluence in DMEM/F12 cell culture media supplemented with 100 U/ml Penicillin/Streptomycin solution (Gibco) and 10% Fetal Bovine Serum (FBS; Greiner) (complete medium). Cells were cultured in plastic culture flasks (Greiner) and were maintained at 37 °C, in a humidified atmosphere of 5% CO₂ in air.

2.2. Chemical exposure

The chemicals used in this study were twelve respiratory sensitizers (belonging to five different chemical subclasses) and ten respiratory non-sensitizers that are not skin sensitizers. The chemicals tested in this study can be found in Table 1, which shows the identity of the chemical, whether the chemical is a respiratory sensitizer (RS), respiratory irritant (RI) or non-sensitizer (NS) (based on our previous work (Dik et al., 2014)), the vehicle in which the test chemicals were dissolved and the final exposure concentration. The respiratory sensitizers consisted of five chemical classes: Isocyanates, acid anhydrides, amines, salts and aldehydes. All chemicals were obtained from Sigma Aldrich (Zwijndrecht, The Netherlands). The final exposure concentrations were determined by calculating the concentration that results in a decrease of cell viability to 80% (CV80), based on a colorimetric assay using the tetrazolium dye MTT (Morgan, 1998) (data not shown). The highest concentration tested was 4 mM. If a chemical did not show toxicity at this concentration, experiments were conducted using 4 mM.

Before exposure, cells were diluted to a concentration of 6×10^5 cells/ml fresh complete medium, seeded into 6-wells plates (4 ml per well; Costar, VWR) and incubated at 37 °C in a humidified atmosphere of 5% CO₂ in air for 24 h to form a monolayer. The cells were subsequently washed with HBSS and exposed to chemicals in medium without FBS or solvent controls without FBS for 4 h. The DMSO or HBSS content in all samples was 1%. At the end of the exposure, 700 µl RNAlater cell reagent (Qiagen, Westburg, The Netherlands) was

Table 1

Overview of all chemicals tested in this study, identified by name and CAS number. The vehicle in which each chemical was dissolved (dimethylsulfoxide (DMSO) or Hanks Balanced Salt Solution (HBSS)) as well as the exposure concentration (CV80) is also provided.

Chemical	Abbreviation	CAS number	Vehicle	CV80 (mM) ^a
<i>Respiratory sensitizers</i>				
Hexamethylene diisocyanate	HDI	822–06–0	DMSO	1.10
Toluene diisocyanate	TDI	584–84–9	DMSO	0.15
Maleic anhydride	MA	108–31–6	DMSO	2.65
Phthalic anhydride	PA	85–44–9	DMSO	1.33
Trimellitic anhydride	TMA	552–30–7	DMSO	2.73
Ethylene diamine	ED	107–15–3	HBSS	4.00
Triethylenetetramine	TETA	112–24–3	HBSS	4.00
Ethanolamine	Etha	141–43–5	HBSS	4.00
Sodium hexachloroplatinate	HcPt	16,929–58–7	HBSS	0.42
Chloramine-T	Chlor-T	127–65–1	HBSS	1.83
Glutaraldehyde	Glut	111–30–8	HBSS	0.19
Formaldehyde	Form	50–00–0	HBSS	0.11
<i>Respiratory irritants</i>				
Acrolein	Acro	107–02–8	HBSS	0.04
Methyl salicylate	MS	119–36–8	DMSO	1.17
Epichlorohydrin	EPI	106–89–8	HBSS	0.08
Sodium dodecyl sulfate	SDS	151–21–3	HBSS	0.11
<i>Respiratory non-sensitizers</i>				
Glycerol	GLY	56–81–5	HBSS	4.00
Lactose	LAC	63–42–3	HBSS	4.00
Mannitol	MAN	69–65–8	HBSS	4.00
Vanillin	VAN	121–33–5	DMSO	4.00
Saccharin	SAC	81–07–2	DMSO	4.00
Lactic acid	LAA	50–21–5	HBSS	4.00

^a In all cases where the CV80 is 4.00, the chemical was not found to be toxic at this concentration.

added to each well. Cells were resuspended and stored at −20 °C until further analysis within the same week.

2.3. RNA isolation

RNA was isolated by using an RNeasy Plus Mini Kit (cat.no. 74134, Qiagen) according to the manufacturer's instructions. RNA concentrations and qualities were determined using a NanoDrop Spectrophotometer (Isogen-Life Sciences) and an Agilent 2100 Bioanalyzer (Agilent), respectively. All RNA samples were diluted in RNase-free water to a concentration of 200 ng/µl in 15 µl and stored at −80 °C. Control RNA samples from HBSS and DMSO exposed cells were included in the analysis as well.

2.4. DNA microarray analysis

RNA samples were further processed for hybridization to Affymetrix HT HG-U133 + PM plates at the Microarray department (MAD) of the University of Amsterdam, The Netherlands. Amplification, labeling and hybridization were performed according to Affymetrix protocols, using an automated Affymetrix Genechip console. Samples were randomized prior to processing and hybridization. Array plates were scanned with a Genechip HT array plate scanner and analyzed with the Affymetrix HT software suite. Quality control and normalization of Affymetrix CEL files were performed using the ArrayAnalysis website (www.arrayanalysis.org/) (Eijssen et al., 2013), using the Robust Multichip Average (RMA) method (Bolstad et al., 2003) and the MBNI custom CDF version 14 (Dai et al., 2005). Normalized data consisted of log₂ transformed signal values for 18,909 genes.

2.5. Statistical analysis

Further statistical analysis including classification was performed in R (www.r-project.org) (version 3.0.3) (R Development Core Team,

2010). Values were corrected for the vehicle-exposed samples to obtain log2-ratios for compound exposure data.

Next, the chemicals were classified as either respiratory sensitizers or non-sensitizers (respiratory irritants + non-sensitizers). Classification was done using a leave-one-compound-out cross-validation approach, employing methods described previously (van der Veen et al., 2013; Vandebriel et al., 2010). In this approach, each time the replicate samples for one compound are used as a test set, while the data for the other compounds serve as a training set. A classification algorithm is trained on the training set data, including gene selection and parameter settings, and this algorithm is then applied to predict the class of the test set samples. This is repeated until each compound has been classified. The classification of a compound was based on the prediction for the majority of the replicate samples.

To obtain a predictive gene set that is not only dependent on a single classification algorithm, three different algorithms were applied to the data, each of which working in a different way. The first is Support Vector Machine (SVM) (Rifkin et al., 2003), which creates a separating hyperplane in a multidimensional space. The second algorithm used is Random forests (RF) (Breiman, 2001), which uses a combination of prediction trees. Lastly, Prediction Analysis for Microarrays in R (PAM) (Tibshirani et al., 2002), uses shrunken centroids to classify samples. Finally, to reduce the algorithm dependent class prediction, the outcomes of the three classifier algorithms were combined to a single overall class prediction. The prediction accuracy for the test set samples was calculated as the average of all compound predictions.

2.6. Gene signature identification

For each of the 22 training sets a set of predictive genes was obtained using the leave-one-compound-out approach. Although the gene lists

differ for each compound that is left out, overlap in the genes exists. This overlap is used to identify consistent predictive genes for each algorithm. Therefore, for each algorithm, genes that were present in at least 11 of the gene lists were selected as a consensus gene set for that algorithm. Additionally, other consensus gene sets were defined using more stringent approaches, namely genes present in two and three of the lists, thus forming consensus gene sets across multiple algorithms.

Functional annotation and overrepresentation analysis was performed using DAVID (Huang da et al., 2009), using functional databases such as Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG). Functional terms were clustered using the DAVID clustering option.

For genes identified as part of at least one consensus gene set, expression changes compared to controls were visualized in a heatmap combined with hierarchical clustering (Euclidean distance, Ward linkage method).

3. Results

3.1. Classification and biomarker selection

The three statistical algorithms (SVM, RF and PAM) were used individually to classify the test chemicals as respiratory sensitizers or non-sensitizers. SVM correctly identified all respiratory sensitizers and misclassified only the irritant acrolein as respiratory sensitizer. Therefore, the accuracy of SVM is 95% (Table 2A). The corresponding consensus gene set for SVM contained 83 genes.

RF identified a consensus set of 75 genes to discriminate respiratory sensitizers from non-sensitizers. As SVM, RF did not correctly classify acrolein. In addition, the respiratory sensitizer chloramine-T was incorrectly predicted as non-sensitizer. This resulted in an accuracy of 91% (Table 2B).

PAM misclassified acrolein, chloramine-T and the respiratory sensitizer glutaraldehyde. The accuracy of PAM is 86% (Table 2C), with a corresponding consensus gene set of 6 genes.

The three individual statistical algorithms in total found 117 unique genes (Fig. 1) that were able to discriminate respiratory sensitizers from non-sensitizers.

Next, the algorithms were combined by majority voting class prediction, which resulted in an accuracy of 91%. Chloramine-T and acrolein were both misclassified (Table 3A). In addition, the robustness of the classifications was evaluated by reducing the number of genes. For this purpose, only genes that were selected by at least two algorithms were used to predict the test chemicals, again by majority voting of the three algorithms into a single predictor. In total, 42 genes are overlapping between at least two of the three signature sets (Fig. 2). The predictive performance of these genes on our test set is 100% accurate (Table 3B). Finally, predictions on all test set chemicals were also generated based on the five genes that are selected by all three algorithms. This resulted in a slightly lower accuracy (95%), as acrolein was not correctly identified as a non-sensitizer (Table 3C).

To confirm the statistical robustness of the consensus gene sets, the above steps were repeated for a scrambled dataset, in which each chemical was randomly assigned as a respiratory sensitizer or non-sensitizer. Chemicals in the same chemical subclass could not be assigned the same random label (e.g. if maleic anhydride was assigned the label of non-sensitizer, phthalic anhydride was assigned the label of respiratory sensitizer). Then the analysis was performed again, with the aim to find consensus genes for each statistical algorithm to distinguish randomly assigned respiratory sensitizers from randomly assigned non-sensitizers. The outcome for all three algorithms was significantly lower compared to the original performance (SVM 23%, RF 27%, PAM 5%). Due to inconsistency between predictions per algorithm, majority voting resulted in an accuracy performance of only 5%. Based on the significantly lower performance of all algorithms including the majority voting approach, and the consistent results obtained in the

Table 2

Predictions by statistical algorithm, based on the consensus genes found by A) SVM, B) RF and C) PAM, to discriminate respiratory sensitizers from non-sensitizers. RS = respiratory sensitizer, NS = non-sensitizer.

				SVM (83 genes)		
				RS	NS	
A				13	9	100% Sensitivity
	Human	RS	12	12	0	90% Specificity
		NS	10	1	9	92% PPV
Accuracy 95%						100% NPV
False prediction:						
False negative						False positive
-						// Acrolein
				RF (75 genes)		
				RS	NS	
B				12	10	92% Sensitivity
	Human	RS	12	11	1	90% Specificity
		NS	10	1	9	92% PPV
Accuracy 91%						90% NPV
False prediction:						
False negative						False positive
Chloramine-T						// Acrolein
				PAM (6 genes)		
				RS	NS	
C				11	11	83% Sensitivity
	Human	RS	12	10	2	90% Specificity
		NS	10	1	9	91% PPV
Accuracy 86%						82% NPV
False prediction:						
False negative						False positive
Chloramine-T						// Acrolein
Glutaraldehyde						

original analysis, it can be concluded that the observed consensus genes in the original analysis are robust.

3.2. Biological function of genes

DAVID was used to functionally annotate the 117 unique genes found by at least one of the algorithms. The most prominent associated processes were (Table 4): cytoskeleton (20 genes, 17%), DNA binding (19 genes, 16%), protein transport (6 genes, 5%).

Table 4 summarizes the genes that were detected in at least two of the three algorithms. These genes were associated with cytoskeleton, DNA binding, protein transport and with NF- κ B signaling. Only five genes were retrieved by all three algorithms (GCC2, KIAA0355, FCHSD2, ANKRD12, BAZ2B). A biological function is known for GCC2 and BAZ2B. GCC2 is involved in Golgi structure maintenance and AP-1 mediated vesicle tethering (Brown et al., 2011) and BAZ2B is involved in chromatin remodeling (Jones et al., 2000) and was categorized by DAVID to DNA binding. A general overview of all 117 genes with their fold-changes can be found in the supplemental data.

4. Discussion

Our study aimed to generate a set of genes to discriminate respiratory sensitizers from non-sensitizers. By using three classification algorithms combined with majority voting, an accuracy of 86% was achieved based on 117 genes. The accuracy could be increased to >95% by using only genes that were identified as discriminating by two (45 genes) or even three (5 genes) algorithms. After randomly assigning the test chemicals as respiratory (non-) sensitizers, the observed high accuracy was lost. Overall, it can be concluded that our predictions are robust and the benefit of combining multiple prediction algorithms was shown. However, these results should be validated using for example a PCR experiment.

The other aim of this study was to identify molecular pathways induced by respiratory sensitizers in bronchial epithelial cells. Several of the genes found to be regulated after exposure in this study could be associated with the cytoskeleton. Interestingly, of these genes, three (PPP1R12A, SOS1 and AKAP9) can be associated with the function of the epithelial barrier and asthma. PPP1R12A was found to be involved in the regulation of the actin cytoskeleton through the IL-33 signaling pathway (Pinto et al., 2015), which is implicated in airway

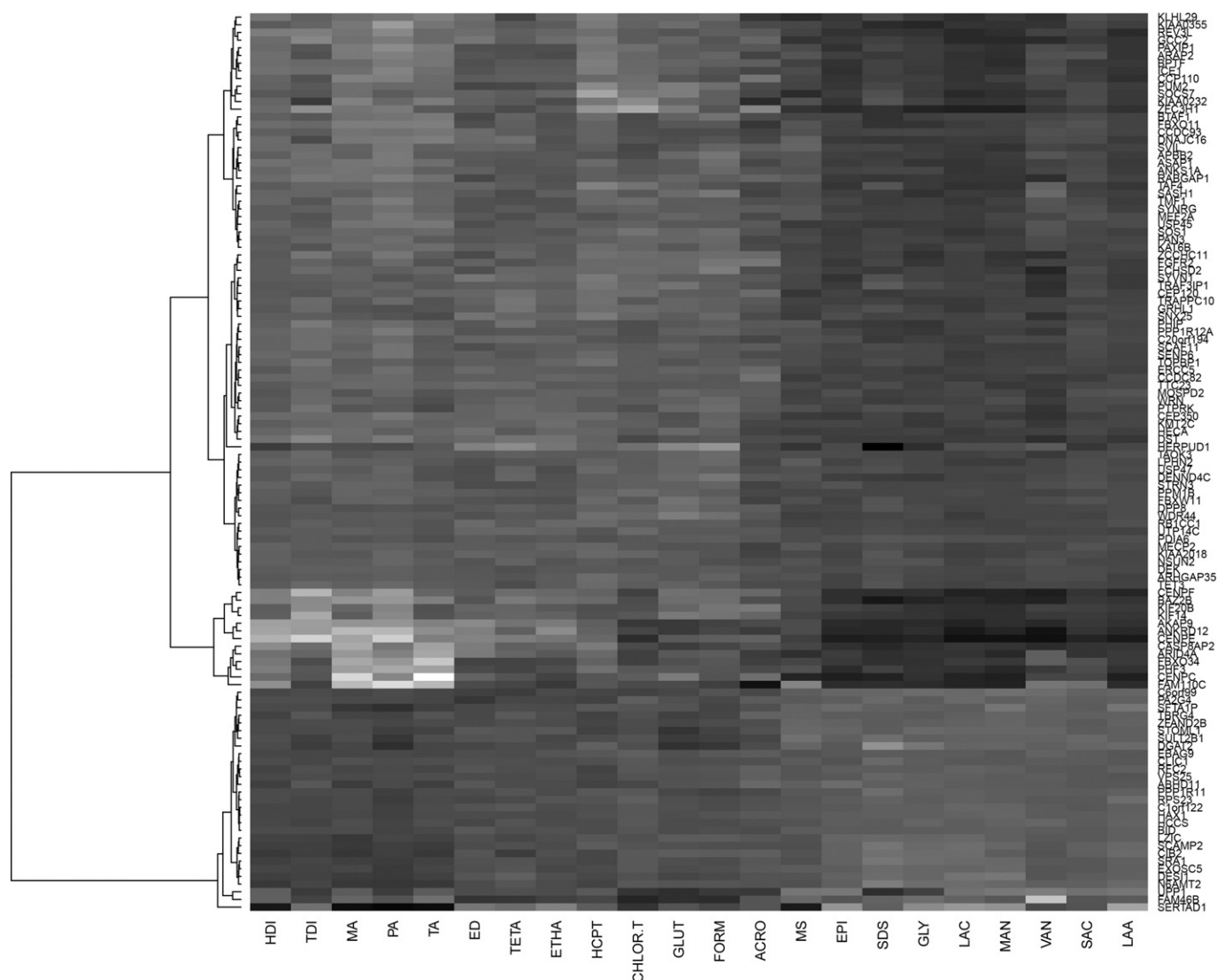


Fig. 1. Heat map of all 117 unique genes found by one or more statistical algorithms. Up-regulation of genes is colored in red, down-regulation of genes is colored in blue. Abbreviations of the chemicals can be found in Table 1.

Table 3

Prediction results using majority voting of the three algorithms (SVM, RF and PAM) based on A) all 117 unique genes selected by at least one of the algorithms; B) genes selected by at least two of the three algorithms; C) the five genes selected by all three algorithms. RS = respiratory sensitizer, NS = non-sensitizer.

A				Prediction		
				RS	NS	
Majority voting based on 117 genes				12	10	92% Sensitivity 90% Specificity 92% PPV 90% NPV
Human	RS	12	11	1		
	NS	10	1	9		
Accuracy 86%						
False prediction:						
False negative				False positive		
Chloramine-T				// Acrolein		

B				RF (75 genes)		
				RS	NS	
Majority voting based on 42 genes				12	10	100% Sensitivity 100% Specificity 100% PPV 100% NPV
Human	RS	12	12	0		
	NS	10	0	10		
Accuracy 100%						
False prediction:						
False negative				False positive		
-				// -		

C				Prediction		
				RS	NS	
Majority voting based on 5 genes				13	9	100% Sensitivity 90% Specificity 92% PPV 100% NPV
Human	RS	12	12	0		
	NS	10	1	9		
Accuracy 95%						
False prediction:						
False negative				False positive		
-				// Acrolein		

inflammation (Oboki et al., 2011). IL-33 stimulation resulted in phosphorylation of PPP1R12A, which is a regulatory subunit of myosin phosphatase (Oboki et al., 2011). Phosphorylation of PPP1R12A results in the inhibition of its myosin phosphatase activity (Feng et al., 1999). The consequence is increased myosin light chain phosphorylation. This results in smooth muscle contraction, which is one of the causes of bronchoconstriction in asthmatics. Interestingly, the actin cytoskeleton regulation pathway described by Pinto et al. (2015) contains the Sos peptide which is encoded by SOS1/2. SOS1 has been shown to be implicated in the PI3K/AKT pathway (Guittard et al., 2015), which is also triggered by the IL-33 signaling pathway (Pinto et al., 2015).

The gene AKAP9 was found to be involved in the functionality of the epithelial barrier after exposure to cigarette smoke extract (CSE) (Oldenburger et al., 2014). In that study, CSE-exposed 16HBE cells had reduced barrier function and showed reduced protein expression of E-cadherin and AKAP9. AKAP9 mRNA expression was also decreased in primary bronchial epithelial cells from current smokers compared with non/ex-smokers (Oldenburger et al., 2014).

Although the immune system plays a major role in occupational asthma, our data indicates that respiratory sensitizers do not affect immunological pathways in epithelial cells. Our data are in line with other studies that have shown that impairment of the epithelial barrier is a major factor in asthma (Georas and Rezaee, 2014; Rezaee and Georas, 2014; Xiao et al., 2011). Combined, these findings suggest that for predicting respiratory sensitizers, next to genes with a direct immunological function, a predictive gene expression profile can be based on genes representing barrier function. This should however be evaluated in an experimental set-up that represents the biological environment of the pulmonary barrier more accurately, such as in an air-liquid interface. In addition, sensitization induced by respiratory sensitizers is thought to be possible through the skin as well (individual references

evaluated in for example (Kimber et al., 2014)). Therefore, evaluation of keratinocytes exposed to respiratory sensitizers could also result in an increased understanding of the biological mechanisms involved in respiratory sensitization. It would also allow comparing the affected pathways between skin and pulmonary epithelial cells, gaining insights in the different and/or similar effects due to route of exposure.

As with all animal-free testing and non-testing methods, it can be argued that they only provide information on only one or a few aspects of the biological pathway(s) involved. The notion that multiple animal-free tests should be combined to come to a conclusion on a toxicological end-point received significant amounts of attention after publication of the report *Toxicity Testing in the 21st century* (2007). The use of integrated testing methods that address different aspects of a biological pathway was clearly shown, for example in the related field of skin sensitization (Bauch et al., 2012; van der Veen et al., 2014). Based on the AOP for skin sensitization, several testing methods that provide information on one of the key elements of the AOP were successfully combined. In contrast to the field of skin sensitization, which already has a well-defined AOP (OECD, 2012), no accepted AOP currently exists for respiratory sensitization. However, Kimber et al. (2014) evaluated several key elements deemed important for the process of respiratory sensitization. The role of the epithelial barrier was, however, not evaluated by that group. We, nevertheless, felt that the epithelial barrier is an important target of the toxicological effects of respiratory sensitizers. One of the aspects that was described by Kimber et al. (2014) is the involvement of dendritic cells. The only other study, at the time of writing this paper, that used genes to classify respiratory sensitizers was done by Forreryd et al. (2015), using dendritic cells. They generated a set of 389 discriminating genes in the MUTZ-3 cell line, which was subsequently used to classify chemicals as respiratory sensitizers or non-sensitizers with an accuracy of 84%. Both our study (using epithelial cells) and that of Forreryd et al. (2015) (using dendritic cells) achieved an accuracy of ~90% when predicting respiratory sensitization potential of chemicals. This may suggest that both approaches could be used within a testing strategy that incorporates multiple tests each addressing different parts of the (partly unknown) respiratory sensitization pathway. Several other studies also aimed to identify biomarkers that can discriminate respiratory sensitizers from respiratory non-sensitizers (Huang et al., 2013; Lauenstein et al., 2014; Remy et al., 2014; Verstraelen et al., 2009a, 2009b, 2009c; Verstraelen et al., 2008b). Their results can contribute to the development of an AOP for respiratory sensitization. However, our own study together with these studies did not apply their findings to predict either a subset of their test chemicals or a new set of chemicals, so their applicability within a testing strategy for respiratory sensitization will need to be further evaluated.

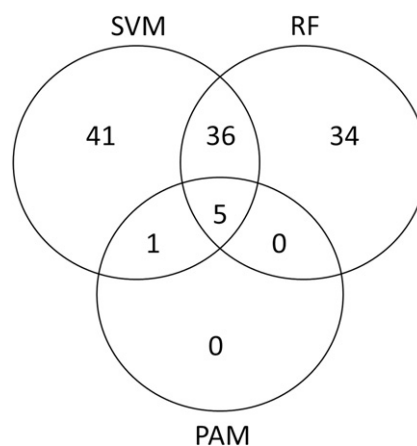


Fig. 2. Venn diagram showing the number of genes and overlap found by the different algorithms.

Table 4

Overview of all genes identified by one or more of the statistical algorithms that are associated with cytoskeleton, DNA binding, protein transport, NF- κ B and the remaining genes that are identified by all three algorithms. Genes are identified by Gene ID, symbol and full name.

Gene ID	Symbol	Found by which algorithm(s)	Full name
<i>Genes associated with cytoskeleton</i>			
1062	CENPE	SVM, PAM	Centromere protein E, 312 kDa
2909	ARHGAP35	SVM, RF	Rho GTPase activating protein 35
4659	PPP1R12A	SVM, RF	Protein phosphatase 1, regulatory subunit 12A
6654	SOS1	SVM, RF	Son of sevenless homolog 1 (<i>Drosophila</i>)
7486	WRN	SVM, RF	Werner syndrome, RecQ helicase-like
9928	KIF14	SVM, RF	Kinesin family member 14
153241	CEP120	SVM, RF	Centrosomal protein 120 kDa
667	DST	SVM	Dystonin
1063	CENPF	SVM	Centromere protein F, 350/400 ka (mitosin)
9738	CCP110	SVM	CP110 protein
9857	CEP350	SVM	Centrosomal protein 350 kDa
10142	AKAP9	SVM	A kinase (PRKA) anchorprotein (yotiao) 9
10456	HAX1	SVM	HCLS1 associated protein X-1
23637	RABGAP1	SVM	RAB GTPase activating protein 1
6840	SVIL	RF	Supervillin
9585	KIF20B	RF	Kinesin family member 20B
11073	TOPBP1	RF	Topoisomerase (DNA)II binding protein 1
26146	TRAF3IP1	RF	TNF receptor-associated factor 3 interacting protein 1
54521	WDR44	RF	WD repeat domain 44
642273	FAM110C	RF	Family with sequence similarity 110, member C
<i>Genes associated with DNA binding</i>			
29994	BAZ2B	SVM, RF, PAM	Bromodomain adjacent to zinc finger domain, 2B
4204	MECP2	SVM, RF	Methyl CpG binding protein 2
4205	MEF2A	SVM, RF	Myocyte enhancer factor 2A
5926	ARID4A	SVM, RF	A T rich interactive domain 4A (RBP1-like)
205717	KIAA2018	SVM, RF	KIAA2018
5980	REV3L	SVM	REV3-like, catalytic subunit of DNA polymerase zeta (yeast)
5982	RFC2	SVM	Replication factor C (activator 1) 2, 40 kDa
7110	TMF1	SVM	TATA element modulatory factor 1
7913	DEK	SVM	DEK oncogene
9044	BTAF1	SVM	BTAF1 RNA polymerase II, B-TFIID transcription factor associated, 170 kDa
10011	SRA1	SVM	Steroid receptor RNA activator 1
23522	KAT6B	SVM	K (lysine) acetyltransferase 6B
29841	GRHL1	SVM	Grainyhead-like 1 (<i>Drosophila</i>)
29966	STRN3	SVM	Striatin, calmodulin binding protein 3
58508	KMT2C	SVM	Myeloid/lymphoid or mixed-lineage leukemia 3
1060	CENPC	RF	Centromere protein C1
2073	ERCC5	RF	Excision repair cross-complementing rodent repair deficiency, complementation group 5
2186	BPTF	RF	Bromodomain PHD finger transcription factor
6874	TAF4	RF	TAF4 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 135 kDa
<i>Genes associated with protein transport</i>			
10066	SCAMP2	SVM, RF	Secretory carrier membrane protein 2
11276	SYNRG	SVM, RF	Synergism, gamma
83891	SNX25	SVM, RF	Sorting nexin 25
84313	VPS25	SVM, RF	Vacuolar protein sorting 25 homolog (<i>S. cerevisiae</i>)
2263	FGFR2	SVM	Fibroblast growth factor receptor 2
7109	TRAPP10	SVM	Trafficking protein particle complex 10
<i>Genes associated with NF-κB</i>			
5495	PPM1B	SVM, RF	Protein phosphatase, Mg ²⁺ /Mn ²⁺ dependent, 1B
23,328	SASH1	SVM, RF	SAM and SH3 domain containing 1
<i>Other genes found by all three algorithms</i>			
9648	GCC2	SVM, RF, PAM	GRIP and coiled-coil domain containing 2
9710	KIAA0355	SVM, RF, PAM	KIAA0355
9873	FCHSD2	SVM, RF, PAM	FCH and double SH3 domains 2
23,253	ANKRD12	SVM, RF, PAM	Ankyrin repeat domain 12

5. Conclusion

This paper described the search for biomarkers that can identify respiratory sensitizers *in vitro*. Our findings show that the respiratory sensitization potential of chemicals can be accurately predicted using an epithelial cell line. Based on our study and that of the only other study to date that aimed to predict respiratory sensitization potential of chemicals (Forreryd et al., 2015), it can be suggested that single cell line studies are suitable to predict respiratory sensitization, even if the genes are not necessarily related to an

immunological function. Based on our results and on the other studies on epithelial cells discussed in this paper, the pulmonary barrier should be considered an important target of research in the field of respiratory sensitization.

Transparency document

The Transparency document associated with this article can be found, in online version.

6. Acknowledgments

The authors would like to acknowledge Eric Gremmer, Leonie Lautz and Fariza Bouallala for their practical contributions to this work.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.tiv.2015.10.010>.

References

- Anon. 2007. Toxicity Testing in the 21st Century: a Vision and a Strategy. The National Academies Press, Washington, DC.
- Bauch, C., Kolle, S.N., Ramirez, T., Eltze, T., Fabian, E., Mehling, A., Teubner, W., van Ravenzwaay, B., Landsiedel, R., 2012. Putting the parts together: combining in vitro methods to test for skin sensitizing potentials. *Regul. Toxicol. Pharmacol.* 63, 489–504.
- Bolstad, B.M., Irizarry, R.A., Astrand, M., Speed, T.P., 2003. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 19, 185–193.
- Breiman, L., 2001. Random forests. *Mach. Learn.* 45, 5–32.
- Brown, F.C., Schindelhaim, C.H., Pfeffer, S.R., 2011. GCC185 plays independent roles in Golgi structure maintenance and AP-1-mediated vesicle tethering. *J. Cell Biol.* 194, 779–787.
- Cozens, A.L., Yezzi, M.J., Kunzelmann, K., Ohnui, T., Chin, L., Eng, K., Finkbeiner, W.E., Widdicombe, J.H., Gruenert, D.C., 1994. CFTR expression and chloride secretion in polarized immortal human bronchial epithelial cells. *Am. J. Respir. Cell Mol. Biol.* 10, 38–47.
- Dai, M., Wang, P., Boyd, A.D., Kostov, G., Athey, B., Jones, E.G., Bunney, W.E., Myers, R.M., Speed, T.P., Akil, H., Watson, S.J., Meng, F., 2005. Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Res.* 33, e175.
- Dik, S., Ezendam, J., Cunningham, A.R., Carrasquer, C.A., van Loveren, H., Rorije, E., 2014. Evaluation of in silico models for the identification of respiratory sensitizers. *Toxicol. Sci.* 142, 385–394.
- Eijssen, L.M., Jaillard, M., Adriaens, M.E., Gaj, S., de Groot, P.J., Muller, M., Evelo, C.T., 2013. User-friendly solutions for microarray quality control and pre-processing on ArrayAnalysis.org. *Nucleic Acids Res.* 41, W71–W76.
- Feng, J., Ito, M., Ichikawa, K., Isaka, N., Nishikawa, M., Hartshorne, D.J., Nakano, T., 1999. Inhibitory phosphorylation site for rho-associated kinase on smooth muscle myosin phosphatase. *J. Biol. Chem.* 274, 37385–37390.
- Forreryd, A., Johansson, H., Albrekt, A.S., Borrebaeck, C.A., Lindstedt, M., 2015. Prediction of chemical respiratory sensitizers using GARD, a novel in vitro assay based on a genomic biomarker signature. *PLoS One* 10, e0118808.
- Georas, S.N., Rezaee, F., 2014. Epithelial barrier function: at the front line of asthma immunology and allergic airway inflammation. *J. Allergy Clin. Immunol.* 134, 509–520.
- Guittard, G., Kortum, R.L., Balagopal, L., Cuburu, N., Nguyen, P., Sommers, C.L., Samelson, L.E., 2015. Absence of both Sosl-1 and Sosl-2 in peripheral CD4 T cells leads to PI3K pathway activation and defects in migration. *Eur. J. Immunol.*
- Holgate, S.T., 2008. The airway epithelium is central to the pathogenesis of asthma. *Allergol. Int.* 57, 1–10.
- Huang da, W., Sherman, B.T., Lempicki, R.A., 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57.
- Huang, S., Wiszniewski, L., Constant, S., Roggen, E., 2013. Potential of in vitro reconstituted 3D human airway epithelia (MucilAir) to assess respiratory sensitizers. *Toxicol. in Vitro* 27, 1151–1156.
- Jones, M.H., Hamana, N., Nezu, J., Shimane, M., 2000. A novel family of bromodomain genes. *Genomics* 63, 40–45.
- Kimber, I., Dearman, R.J., Basketter, D.A., Boverhof, D.R., 2014. Chemical respiratory allergy: reverse engineering an adverse outcome pathway. *Toxicology* 318, 32–39.
- Lalko, J.F., Kimber, I., Gerberick, G.F., Foertsch, L.M., Api, A.M., Dearman, R.J., 2012. The direct peptide reactivity assay: selectivity of chemical respiratory allergens. *Toxicol. Sci.* 129, 421–431.
- Lauenstein, L., Switala, S., Prenzler, F., Seehase, S., Pfennig, O., Forster, C., Fieguth, H., Braun, A., Sewald, K., 2014. Assessment of immunotoxicity induced by chemicals in human precision-cut lung slices (PCLS). *Toxicol. in Vitro* 28, 588–599.
- Mapp, C.E., Boschetto, P., Maestrelli, P., Fabbri, L.M., 2005. Occupational asthma. *Am. J. Respir. Crit. Care Med.* 172, 280–305.
- Morgan, D.M.L., 1998. Tetrazolium (MTT) assay for cellular viability and activity. In: Morgan, D.M.L. (Ed.), *Polyamine Protocols*. Humana Press, pp. 179–184.
- Moscato, G., Pignatti, P., Yacoub, M.R., Romano, C., Spezia, S., Perfetti, L., 2005. Occupational asthma and occupational rhinitis in hairdressers. *Chest* 128, 3590–3598.
- Natsch, A., 2010. The Nrf2-Keap1-ARE toxicity pathway as a cellular sensor for skin sensitizers—functional relevance and a hypothesis on innate reactions to skin sensitizers. *Toxicol. Sci.* 113, 284–292.
- Oboki, K., Nakae, S., Matsumoto, K., Saito, H., 2011. IL-33 and airway inflammation. *Allergy Asthma Immunol. Res.* 3, 81–88.
- OECD, 2012. The adverse outcome pathway for skin sensitisation initiated by covalent binding to proteins. OECD Environment, Health and Safety publications; Series on testing and assessment.
- Oldenburger, A., Poppinga, W.J., Kos, F., de Bruin, H.G., Rijks, W.F., Heijink, I.H., Timens, W., Meurs, H., Maarsingh, H., Schmidt, M., 2014. A-kinase anchoring proteins contribute to loss of E-cadherin and bronchial epithelial barrier by cigarette smoke. *Am. J. Physiol. Cell Physiol.* 306, C585–C597.
- Pinto, S.M., Nirujogi, R.S., Rojas, P.L., Patil, A.H., Manda, S.S., Subbannayya, Y., Roa, J.C., Chatterjee, A., Prasad, T.S., Pandey, A., 2015. Quantitative phosphoproteomic analysis of IL-33-mediated signaling. *Proteomics* 15, 532–544.
- R Development Core Team, 2010. A language and environment for statistical computing. R Foundation for Statistical Computing.
- Remy, S., Verstraelen, S., Van Den Heuvel, R., Nelissen, I., Lambrechts, N., Hooyberghs, J., Schoeters, G., 2014. Gene expressions changes in bronchial epithelial cells: markers for respiratory sensitizers and exploration of the NRF2 pathway. *Toxicol. in Vitro* 28, 209–217.
- Rezaee, F., Georas, S.N., 2014. Breaking barriers. New insights into airway epithelial barrier function in health and disease. *Am. J. Respir. Cell Mol. Biol.* 50, 857–869.
- Rifkin, R., Mukherjee, S., Tamayo, P., Ramaswamy, S., Yeang, C., Angelo, A., Reich, M., Poggio, T., Lander, E., Golub, T., Mesirov, J., 2003. An analytical method for multiclass molecular cancer classification. *SIAM Rev.* 45, 706–723.
- Sastre, J., Vandenplas, O., Park, H.S., 2003. Pathogenesis of occupational asthma. *Eur. Respir. J.* 22, 364–373.
- Tibshirani, R., Hastie, T., Narasimhan, B., Chu, G., 2002. Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc. Natl. Acad. Sci. U. S. A.* 99, 6567–6572.
- Toren, K., Blanc, P.D., 2009. Asthma caused by occupational exposures is common — a systematic analysis of estimates of the population-attributable fraction. *BMC Pulm. Med.* 9, 7.
- van der Veen, J.W., Pronk, T.E., van Loveren, H., Ezendam, J., 2013. Applicability of a keratinocyte gene signature to predict skin sensitizing potential. *Toxicol. in Vitro* 27, 314–322.
- van der Veen, J.W., Rorije, E., Emter, R., Natsch, A., van Loveren, H., Ezendam, J., 2014. Evaluating the performance of integrated approaches for hazard identification of skin sensitizing chemicals. *Regul. Toxicol. Pharmacol.* 69, 371–379.
- Vandebriel, R.J., Pennings, J.L., Baken, K.A., Pronk, T.E., Boorsma, A., Gottschalk, R., Van Loveren, H., 2010. Keratinocyte gene expression profiles discriminate sensitizing and irritating compounds. *Toxicol. Sci.* 117, 81–89.
- Vellore, A.D., Drought, V.J., Sherwood-Jones, D., Tunnicliffe, B., Moore, V.C., Robertson, A.S., Burge, P.S., 2006. Occupational asthma and allergy to sevoflurane and isoflurane in anaesthetic staff. *Allergy* 61, 1485–1486.
- Verstraelen, S., Bloemen, K., Nelissen, I., Witters, H., Schoeters, G., Van Den Heuvel, R., 2008a. Cell types involved in allergic asthma and their use in in vitro models to assess respiratory sensitization. *Toxicol. in Vitro* 22, 1419–1431.
- Verstraelen, S., Nelissen, I., Hooyberghs, J., Witters, H., Schoeters, G., Van Cauwenberge, P., Van Den Heuvel, R., 2009a. Gene profiles of a human alveolar epithelial cell line after in vitro exposure to respiratory (non-)sensitizing chemicals: identification of discriminating genetic markers and pathway analysis. *Toxicol. Lett.* 185, 16–22.
- Verstraelen, S., Nelissen, I., Hooyberghs, J., Witters, H., Schoeters, G., Van Cauwenberge, P., Van Den Heuvel, R., 2009b. Gene profiles of a human bronchial epithelial cell line after in vitro exposure to respiratory (non-)sensitizing chemicals: identification of discriminating genetic markers and pathway analysis. *Toxicology* 255, 151–159.
- Verstraelen, S., Nelissen, I., Hooyberghs, J., Witters, H., Schoeters, G., Van Cauwenberge, P., Van Den Heuvel, R., 2009c. Gene profiles of THP-1 macrophages after in vitro exposure to respiratory (non-)sensitizing chemicals: identification of discriminating genetic markers and pathway analysis. *Toxicol. in Vitro* 23, 1151–1162.
- Verstraelen, S., Wens, B., Hooyberghs, J., Nelissen, I., Witters, H., Schoeters, G., Cauwenberge, P.V., Heuvel, R.V., 2008b. Gene expression profiling of in vitro cultured macrophages after exposure to the respiratory sensitizer hexamethylene diisocyanate. *Toxicol. in Vitro* 22, 1107–1114.
- Xiao, C., Puddicombe, S.M., Field, S., Haywood, J., Broughton-Head, V., Puxeddu, I., Haitchi, H.M., Vernon-Wilson, E., Sammut, D., Bedke, N., Cremin, C., Sones, J., Djukanovic, R., Howarth, P.H., Collins, J.E., Holgate, S.T., Monk, P., Davies, D.E., 2011. Defective epithelial barrier function in asthma. *J. Allergy Clin. Immunol.* 128, 549–556, e541–512.