EUROPEAN COMMISSION
JOINT RESEARCH CENTRE

Institute for Health and Consumer Protection
European Centre for the Validation of Alternative Methods (ECVAM)

ECVAM
Test Submission Template (TST)

KeratinoSens assay for the testing of skin sensitizers

TABLE OF CONTENTS

Submission of confidential information

We follow a strategy of full transparency in our program to develop alternatives to animal testing, and thus most information submitted in this document is not confidential and may be made available as deemed appropriate by ECVAM. The assay, the standard operating procedure and the results of the interlaboratory study have been published .

The only documents which we require to treat confidential are:
Attachment 12c_PC-extended list_new.pdf
Attachment 12d_PC-surfactants_new.pdf
Attachment 12e_PC-epoxides_new.pdf

These are unpublished documents and partly made in collaboration with partners. We have a approval from these partners to present the data confidentially.
.

List of abbreviations

ARE:	antioxidant response element
Keap1:	Kelch-like ECH-associated protein 1
Nrf2:	nuclear factor (erythroid-derived 2)-like 2
AKR1C2:	Aldoketoreductase
DMSO:	Dimethylsulfoxide
FCS:	Foetal calf serum
PBS:	Phosphate buffered saline
MTT:	Thiazolyl Blue Tetrazolium bromide
EC1.5:	Extrapolated concentration for a 1.5 fold of luciferase induction
I_{max}:	Maximal induction of luciferase activity relative to solvent control over the complete dose-response range measured
IC50:	Concentration for reduction of cellular viability by 50% as determined with the MTT assay

Submitter suggestion for entering the ECVAM validation process

	Please tick where appropriate
Prevalidation	**(if data not considered sufficient for peer review)**
Prospective validation	
Retrospective validation	
Validation based on PS	
Peer-review	**Peer-review of the data as an external pre-validation study (preferred)**

1. INFORMATION ON VALIDATION MODULES

1.1 MODULE 1: TEST DEFINITION

1.1.1 Human health, environmental or other biological effects addressed by the test method

Skin sensitization / allergic contact dermatitis caused by chemicals

1.1.2 Scientific basis – biological and/or mechanistic relevance

The only feature all skin sensitizers have in common is their intrinsic electrophilicity, or their potential to be metabolically transformed to electrophilic chemicals.[1]
The signaling pathway with the repressor protein Keap1 and the transcription factor Nrf2, which binds to the antioxidant / electrophile response element (ARE / EpRE), is known to respond to electrophilic chemicals. The sensor protein Keap1 (Kelch-like ECH-associated protein 1) contains highly reactive Cys residues. In un-induced conditions, Keap1 is bound to Nrf2 (nuclear factor-erythroid 2-related factor 2), which targets Nrf2 for proteolytic degradation. Covalent modification of crucial Cys residues by small electrophilic molecules leads to dissociation of Keap1 from the transcriptional regulator Nrf2, which then activates genes (mainly genes coding for phase II detoxifying enzymes) having an antioxidant response element (ARE) in their promoter sequence.[2]

This pathway was found to be a valuable cellular endpoint to detect skin sensitizers *in vitro* in our laboratory[3-5] and this result was confirmed by several independent laboratories[6-8]. Measurement of the induction of this signaling pathway in a reporter cell line provides a high-throughput cell-based *in vitro* test to screen for the skin sensitization potential of novel chemicals.

The initial results indicated an empirically observed (correlative) relationship. Yet recently, the *in vivo* relevance of this signalling pathway for contact allergy and in particular for the T_H1 response, has been established by Kim et al. [9] by using Nrf2-knockout mice. The response to skin sensitizers, in particular the T_H1 response, was reduced in these mice, thus indicating that the approach is also mechanistically and biologically relevant.

Many phase II detoxification genes contain an ARE-element in their promoter. One particular gene is AKR1C2 coding for an aldo-ketoreductase. This particular gene was identified as one of the target genes up-regulated by contact sensitizers in dendritic cells by gene-chip and RT-PCR analysis [10,11].

The rapidly accumulating evidence from different laboratories for the up-regulation of Nrf2-regulated genes by skin sensitizers has recently been reviewed.[12] After this review had been published, another detailed gene-chip study in HaCaT cells has again confirmed that the ARE-regulated genes form a key group of genes specifically responding to sensitizer challenge [6].

1.1.3 Intended purpose of the test method

This method is a screening method to test for the potential of chemicals to be skin sensitizing: The main focus is the testing of chemicals evaluated under REACH and under the Cosmetic Directive.
If used as stand-alone method, the use of the method may be restricted to classification and labelling, relevant to "REGULATION (EC) No 1272/2008 on classification, labelling and packaging of substances and mixtures".
If used as part of an integrated testing strategy and combined with read-across, the quantitative results generated should facilitate at least a basic risk assessment and safety prediction, e.g. definition of sensitisation potency and setting of safe use levels.

1.1.4 Evidence demonstrating the need for the test method

No validated *in vitro* assay is currently available for skin sensitization, although three methods are in the prevalidation process. Methods to replace the currently used local lymph node assay (LLNA) for skin sensitization are needed in the cosmetic industry by 2013.

Within the REACH legislation a large number of chemicals need to be tested for skin sensitization. For animal welfare reasons on the one hand, but also for a more economic implementation of the REACH legislation on the other hand, a high-throughput assay for skin sensitization testing is of high importance. Whereas for the Cosmetics a full replacement is needed, for REACH also a partial replacement (reduction) by a high throughput test for a first-tier screening is of interest.
The proposed assay is a high-throughput assay, as both endpoints (luciferase induction and cellular viability) can be rapidly assessed in a 96-well format for many chemicals, and thus the assay offers advantages in effectiveness and throughput. If done on a routine basis the costs per chemical are less than 20% of the currently done LLNA.

1.1.5 Technical specifications

1.1.5.1 Protocol(s) of the test method

The protocol is appended as Attachment 1. It has been written *de novo* and is not directly related to an existing INVITOX protocol. The protocol has already been used in the interlaboratory study and proven to be sufficiently detailed in order that other laboratories could execute the test without a direct training.

a) Brief description of the test system

The cell line KeratinoSens contains a 56-base-pair insertion containing the ARE sequence from the AKR1C2 gene. This genetic element was inserted in front of a SV40 promotor, and the combination of the ARE-element with the SV40 promotor was inserted upstream of the luciferase gene in the vector pGL4 from Promega. The resulting vector was transfected into HaCaT keratinocytes, and clones with a stable insertion were selected in the presence of Geneticin / G418. The selected clone 8 (termed KeratinoSens) was further propagated as a reporter cell line. These adherent cells can directly be grown by the testing lab, and cells can be used without any further modification or differentiation steps.

Note: Aliquots of an early passage of the cell line have been deposited at DSMZ as safe deposits and disaster recovery stocks. Multiple stock cultures of early passages have been prepared and are stored at Givaudan. Aliquots of these early pasasages have been used to make daughther cultures which have been frozen and subcultured again. No change in the performance was noted in these cultures. Hence Givaudan is confident, that the performance of the cells can be assured over time and due to stocks at DSMZ and in various partner labs we should have no risk of losing the cells / or the original performance.

b) Parameters and endpoints measured

Induction of luciferase is the read-out / endpoint evaluated to determine sensitization potential. Luciferase induction directly indicates activation of the AKR1C2-ARE element. The cell line expresses a stable background of luciferase activity due to the basal activity of the SV40 promotor. Fold-induction of luciferase activity relative to this background is determined for a full dose-response curve between 0.98 µM and 2000 µM. Based on these data the EC1.5 value is determined, which indicates the concentration needed for gene induction 50% over the background. Cytotoxicity is measured in parallel and expressed as IC50 value. If the luciferase induction is only observed at cytotoxic concentrations, this is indicative of a false-positive gene-induction generated by a skin irritant.

c) Quality criteria applied to the test system

Care is taken that cells are never grown to too high densities. If cell density is controlled no further quality criteria need to be applied prior to the assay. Quality is assured by including in each test plate:
(i) six wells with the negative solvent controls and (ii) the positive control in a dose-response analysis.

d) Positive control, negative control, benchmarks

Positive Control(s)

Cinnamic aldehyde, a well known moderate skin sensitizer, is used as a positive control in each test plate. Cinnamic aldehyde needs to be positive in each accepted test, and quantitative criteria are applied to discard results if the induction by cinnamic aldehyde is outside the acceptable range.

Negative Control(s)

The negative control is DMSO. It is used as a solvent, and each test plate contains 6 wells of DMSO solvent control to have a statistically sound amount of data to calculate fold-induction values. Variation of the negative controls must be under 20%

Benchmarks *(if applicable)*

No further Benchmark is used. The positive control cinnamic aldehyde itself is an item giving a mid-range induction. Based on the extensive ring-study data, a further useful benchmark may now be defined for future studies, but it is not considered necessary.

e) Acceptance criteria applied to the results

The standard deviation of the DMSO background is calculated over the three triplicate plates in each test (in total 18 wells). The test is accepted if the variation of this background is below 20%. Cinnamic aldehyde needs to be positive in each accepted test.
It is recommended, that for accepted tests, the fold-induction by the positive control cinnamic aldehyde at a concentration of 64 µM needs to be between 2 and 8 and the EC1.5 for cinnamic aldehyde needs to be in the range between 7 and 30 µM.

1.1.5.2 Data analysis and Prediction Model (PM)

Data analysis:
The data processing is conducted automatically with an Excel template forming part of the SOP. This Excel file is attached as Attachment 1b.
The raw data are directly pasted into this file, and then the following parameters are calculated automatically:
- Gene induction relative to DMSO control in each well
- Wells with statistically significant induction over a given threshold (default value set to 1.5 = 50% enhanced gene activity)
- I_{max} (the maximal induction over the full dose-response)
- The EC1.5 value (extrapolated concentration for statistically significant induction above threshold, calculated similarly to the LLNA)
- IC50 value and IC70 value for cytotoxicity

As detailed in the SOP (Attachment 1a), these automatic calculations are in most cases reliable and sufficient. A careful inspection of the dose response curves, which are also automatically plotted in the Excel file, is nevertheless recommended to check whether the graphical dose-response appears consistent with the calculated EC1.5 values.

In order to better understand how the Excel-template of the SOP evaluates the raw data, we added Attachment 1c, which gives the actual raw data file from Master Plate 1 of Experiment 1 of the additional intralaboratory study recommended by ECVAM (see below).

Prediction model: Based on the above calculations, chemicals can be rated positive or negative. A chemical is positive if it (i) gives a statistically significant induction of luciferase activity > 1.5-fold in at least 2 of the three experiments, (ii) the EC1.5 value is below 1000µM and (iii) the viability at the EC1.5 determining concentration (i.e. the concentration which yields the significant induction > 1.5-fold) is > 70%.
Accordingly, chemicals are rated negative if no significant induction is recorded, or (i) only in one repetition, or (ii) the induction occurs above 1000 µM or (iii) at cytotoxic concentrations only.

1.1.5.3 Expertise required for performing the test method protocol

The method is based on very classical cell biological test set-up in 96-well plates. Thus any technician or scientist with expertise in running cell biological experiments, and in particular in performing dose-response tests on chemicals in a 96-well format (such as classical cytotoxicity assays) should be able to run this method.

1.1.6 Test items used for developing and optimising the test method (protocol and PM)

The optimal clone was selected and the SOP parameters were defined with the set of 6 chemicals given in Table 1 in Attachment 2. The prediction model was established based on a screening with a test set of 67 chemicals tested according to the SOP and referred to as the 'Silver List' shown in Table 2 of Attachment 2. This prediction model is a slight modification of the one used in the AREc32 assay before (Natsch and Emter, 2008). The tested chemicals (i) span a range of molecular weights from 30 to 388, (ii) cover cLogP ranges from -4.8 to 5.2 (iii) cover the full range of skin sensitizer potency classes and (iv) all the key published reference lists of test chemicals for the assessment of skin sensitization test methods. The details on how the 'Silver list' was set up and the rationale for selecting the chemicals is given in the publication[3] and in Table 2 of Attachment 2 . The results in the KeratinoSens assay according to the SOP along with physical data are also summarized in Table 2 of Attachment 2.

1.1.7 Cost and time estimates per test item

Materials:
The costs to test 1 compound in three repetitions each with three replicates at 12 concentrations are: Media and FCS 2.4 €, the luciferase substrate: 72 €, Lysis reagent 2.5 €, test plates 6.5 €, other disposables (pipettes, tissue culture dishes, MTT, PBS, DMSO, test tubes) estimated at maximal 10 €. Thus the cost per compound excluding labour and fixed equipment is: 93.5 €.
Time required:
Without the use of robotics, a trained experimenter can run at least 21 compounds in one run in triplicate and can perform at least two runs in one week. Thus three 40 h weeks are needed for the final results in three repetitions on 42 compounds (full dose response at 12 concentrations) Therefore the labour per compound (notably, once the test has become laboratory routine!) is 3 h. By using robotics, the throughput might be significantly enhanced. As the method is based on a very standardized cell biology setup, and uses an adherent cell line grown in 96-well plates, it should be possible to run the test on current laboratory robots.

The full cost per compound for labour, overhead costs and disposables is therefore around 500€, which is 20% of the current cost of an LLNA.

1.1.8 Occurrence of non-qualified tests

In our laboratory, 188 triplicate runs were performed since the introduction of cinnamic aldehyde as positive control. The performance criteria for a variability of < 20% in the control wells was fulfilled in 179 of the 188 runs (Median variability 10.4%). The performance criteria that cinnamic aldehyde was significantly positive (EC1.5 < 64 µM) was fulfilled in all of the 188 runs. The quantitative performance criteria for the induction by cinnamic aldehyde (a) Induction at 64 µM between 2 and 8 was fulfilled in 171 of the 188 runs (91%), and (b) EC1.5 between 7 µM and 30 µM was fulfilled in 168 of the 188 runs (89%). In 177 of the 188 runs

one of the two quantitative performance criteria was fulfilled. These data are listed in Attachment 17a for the 188 runs.

In the ring study, once the test was set up and running in the different laboratories, the performance criteria for a variability of < 20% in the control wells was fulfilled in 42 of the 48 runs in the four external labs, and it was between 20 and 26% in the remaining 6 runs. The median variability for all runs was 11.7 %. The performance criteria that cinnamic aldehyde was significantly positive (EC1.5 < 64 µM) was fulfilled in all of the 48 runs. The quantitative performance criteria for the induction by cinnamic aldehyde (a) Induction at 64 µM between 2 and 8 was fulfilled in 39 of the 48 runs, and (b) EC1.5 between 7 µM and 30 µM was fulfilled in 38 of the 48 runs**.** These data are given in Attachment 17b. These data are summarized in the below Table.
In summary, overall over 80% of the experiments do meet the performance criteria.

Acceptance criteria	**Target**	**Result of 188 runs in lead lab** [1)]	**Result of 4 external labs in Ring study**
Cinnamic aldehyde positive	statistically significant induction	100 % (188/188)	100 % (48/48)
EC 1.5 cinnamic aldehyde	7 – 30 µM	89 % (168/188), Median 19.5	79.2 % (38/48), Median 4 labs 12.8
Induction at 64 µM cinnamic aldehyde	2 – 8 fold	91% (171/188). Median 2.82	81.2 % (39/48) Median 4 labs 4.4
One of the two quantitative criteria fulfilled		94% (177/188)	91.6 % (44/48)
Variability in solvent wells	below 20%	95 % (179/188), Median 10.4 %	87.5 % (42/48) Median 11.7 %

In the ring study, results not fulfilling the criterium for a variability of < 20% in the control wells were discarded in the transferability phase. Not sufficiently standardized luminescence readings were the reason for most non-qualified tests (plate formats not matching the luminometer, too low sensitivity of the selected luminometer, inappropriate combination of luminescence substrate and plate reader; see below).

Note 1: Once the performance of the luminescence readings was assured and laboratories proceeded to phase II of the ring study (blind coded chemicals), all the experiments were accepted, even if slightly outside of the performance criteria. With this approach, analysis of the data now indicates how sensitive the assay is relative to the set performance criteria. This may help to eventually adjust the performance criteria for full validation studies, whereas if the criteria had been applied too strictly in the ring-study, we would not have learnt how narrowly they need to be defined in order to obtain reproducible results.
This is easily understood based on two examples presented in Attachment 17. In Figures 17a and 17b the two repetitions on the same chemicals in the same lab with the biggest difference in control value variability are shown. Overall, both repetitions gave similar results, but the

dose-response curves in the run with control variability of 8.4% are smoother as compared to the data in the run with 26.8% variability. This indicates that one should in the future strictly adhere to the 20% variability criterium and this may even further improve the data quality in future studies.

Note 2: Generally the results summarized above indicate that cinnamic aldehyde is a very robust positive control to verify that the test is working and it is reliably positive in all the 236 runs summarized in Attachment 17a and 17b. Yet the quantitative variability for cinnamic aldehyde is clearly higher as compared to some other chemicals studied in the ring study.

1.1.9 Known limitations and drawbacks of the test methods

There are no technical drawbacks, as the nature of the testing is technically straightforward. As the optimal detection device is readily available for an economic price, the cost of entry into this technology is low. Key drawbacks are the limitations regarding prediction of potency and other limitations discussed under 'applicability domain'.

1.1.10 Intellectual Property Rights (IPRs)

The luciferase gene in the KeratinoSens cell line is patent protected by Promega Corp. It can be used by any laboratory for research use with the proviso that the substrate used for the assay is purchased from Promega. If the assay is used to offer commercial service, a licence fee needs to be paid to Promega, but this is not prohibitive as it is the current business model of Promega.

A patent from the year 2001 (EP1130086 A1) on the general use of reporter cell lines derived from HaCaT cells for the use of toxicity screening has been abandoned, and thus it does not pose any limitations. We are not aware of any other patents which may limit the use of the assay for the intended purpose.

Givaudan has decided to follow a no-patent strategy for new developments in alternative assays, and thus, has so far not filed any patent applications on either KeratinoSens nor on our initial findings on the general principle of using Nrf2-regulated genes as screening targets in order to not hamper validation and regulatory acceptance. Givaudan has applied for a Trade mark on the name 'KeratinoSens', but this is not limiting its use.
Givaudan will share the recombinant cell line KeratinoSens with third laboratories under a material transfer agreement. No licence fee will be reimbursed for research and validation studies. A licence fee may be asked for commercial testing, but this fee will be below 5-10% of the current cost of an LLNA.

1.1.11 History of test method development

Initial testing started in the reporter cell line AREc32[4, 13], a cell line derived from the breast cancer cell line MCF-7. To bring the approach closer to conditions in human skin, the new KeratinoSens cell line was developed. This cell line has the added advantage of higher reproducibility of test results. A range of different genetic constructs were made and tested to finally select the construct used to generate this cell line. Part of this optimization work is summarized in the publication[3].

Note: The genetic construct finally chosen contains an SV40 promotor upstream of the luciferase gene. This leads to a relatively high but very stable background signal, against which gene induction can be compared. As luciferase measurements do not have a standardized scale (i.e. there is no official unit measured by different luminometers), the use of scale-independent fold-induction values was found crucial for the development of a transferable method which is not dependent on a specific measurement device. Initial genetic constructs were based on a minimal promotor contained in the PGL4 plasmids. With these constructs almost no background signal was found, which then led to very nice induction curves, but fold induction could not reliably be calculated. Therefore this approach was abandoned in order to develop a transferable method with a scale-independent read-out.

1.1.12 Quality system(s) of the developing laboratory

The Givaudan *in vitro* toxicology laboratory is a research laboratory not certified according to GLP. Nevertheless an SOP was developed early in the assay development, and this SOP was not further modified once all the parameters were determined. The laboratory has strictly followed this SOP in order to make all results and all screenings fully comparable.

1.2 MODULE 2: WITHIN-LABORATORY REPRODUCIBILITY (WLR)

1.2.1 Rationale for the selection of the test items used for assessing WLR

The WLR was assessed based on (i) the data generated on the screening of the 'Silver List' summarized in Attachment 2 and (ii) the repetitions run on a fraction of these chemicals as part of the detailed ring study. Therefore the chemicals used to assess WLR are the same as discussed in Module 3 and 4. Attachment 3 lists these chemicals along with their commercial source.

The chemicals selected for the ring study cover (i) all the chemicals in the publication by Casati *et al.* [14] (which is largely overlapping with the list used by the Sens-it-iv consortium) and (ii) all the chemicals in the performance standards for alternative endpoints in the LLNA published by ICCVAM. Diethylphthalate was added as additional negative control from the Sens-it-iv list to make up the total number to 28 chemicals.

Note: Since the chemicals for the ring study were not selected by an independent 'chemicals selection comitee', but by the lead lab, it was decided to stick to the chemicals of these two published lists, in order not to introduce a bias (i.e. particularly favoring chemicals which can be tested easily with the test method).

1.2.2 Assessment of within-laboratory reproducibility of experimental data

Database for WLR analysis: The chemicals were first tested in the initial screening with 2 - 4 independent repetitions, each in triplicate. The individual results of these repetitions are given in the rows 'GIV_hist' in Attachment 4a. The 28 chemicals listed in Attachment 3 were then repeated with three repetitions in the ring study, and the results are given in the rows 'GIV_RS'. These two studies were run on different batches of chemicals, but they were performed by the

same operator. There was a time lapse of around 6 months between these two studies. Overall, for each chemical thus 5 to 7 independent experiments were performed.
As suggested by ECVAM, an additional study on WLR was performed. This covered 8 blind – coded chemicals supplied by ECVAM and 6 chemicals selected by the test submitter. This set of chemicals was run in three full experiments with each 3 runs (Attachment 4c and Appendix to 4c).

WLR for yes/no prediction:
(1) Comparision of historical data and ring study, Attachment 4a and 4b.
With the current prediction model of rating chemicals positive if (i) EC1.5 values is below 1000 µM and (ii) reached at non-cytotoxic concentrations, 23 of the 28 chemicals gave congruent yes/no prediction in all the 5 – 7 runs, and 3 chemicals gave one single diverging result in the 7 runs. Imidazolidinyl urea, which induces luciferase activity just below cytotoxic concentrations, was positive in 5 of 7 runs. Based on the prediction model the rating of the chemicals as positive or negative in these two full experiment was congruent for 27 of the 28 chemicals (column 'positive / negative in full experiment' in Attachment 4a).

(2) Additional WLR study (attachment 4). Congruent results were obtained in all the three full experiments for 12 of the 14 chemicals. As discussed in Attachment 4c, the two chemicals with incongruent yes/no results overall gave reproducible results but were very close to the threshold (in one case fold-induction threshold and in one case cytotoxicity threshold) which resulted in different ratings.

WLR for EC1.5 values. For all the positive chemicals an evaluation of the variation of the EC1.5 values was made. As detailed in Attachment 4b, these calculations were made with LOG_2 - transformed values.
Excluding the outlier MCI (see below), the geometric standard deviation has, on the average for the 15 evaluated positive chemicals, a value of 1.38. This is below the square root of 2, indicating that the within lab 95.4% confidence interval for determination of the EC1.5 value is slightly below a variation of one well up and down of the geometric mean in the microtiter plate dilutions.
For details on the rationale and results of these calculations see Attachment 4b. A similar analysis of the WLR was also performed on the data from the second WLR study, and are presented and discussed in Attachment 4c. The geometric standard deviation, both within and between the experiments was, on the average, 1.23 in this second study, and reproducibility of EC1.5 values was thus even higher.
In this context it is interesting to compare to the intralaboratory variability of the LLNA EC3 values. Detailed results have been published for Isoeugenol[15], and based on these results the geometric standard deviation can be calculated and it is 1.6.

WLR for the IC50 values. In Attachment 4c also an analysis of the reproducibility of the IC50 values is presented. The geometric standard deviation within the experiments was, on the average, 1.35 and between the experiments 1.23 and thus reproducibility of IC50 values is similar to the reproducibility of the EC1.5 values.

Note: Above assessment of WLR was all made in the laboratory developing the method. Each individual repetition done in an experiment itself has a high statistical power (full dose-

response in triplicate). Hence even if the prediction model for a full assessment considers the result of all the three repetitions together, comparisons between these three independent repetitions themselves may be used to evaluate within-laboratory reproducibility of the quantitative data in the other four laboratories. An assessment of this within laboratory reproducibility / variance of the three repetitions in all the five laboratories is included in the Attachment 10d and 10e on the BLR, and it is discussed under Module 1.4.2. The individual values of the repetitions in each of the labs is also shown in Attachment 8b, 10b and 10c.

1.2.3 Identification and discussion of outlying values

Discordant results for yes/no prediction Hexyl-cinnamic aldehyde proved to be a borderline chemical, with 3 of 5 experiments being positive. For this chemical, cytotoxicity starts just above the luciferase inducing concentration, as already observed in the initial screening on the silver list and indicated in the publication[3].
Outlier for EC1.5 values There was one clear outlier with a high variation: (5-chloro)-Methylisothiazolinone (MCI). This chemicals also had a high variation between labs in the ring study (see below and in attachment 10g.) This chemical is the only one tested at maximal concentration of 500 µM instead of 2000 µM (due to the fact that it is only sold as a diluted solution). A possible explanation is that a calculation error was made (correcting the different concentration twice, once at the plate setup, and once at the calculation level). Nevertheless even with this variation, MCI consistently has a very low EC1.5 value which indicates its strong sensitization potential.

1.2.4 Quality system(s) of the testing laboratory

See 1.1.12.

1.3 MODULE 3: TRANSFERABILITY (TF)

1.3.1 Rationale for the selection of the test items used for assessing the TF

The test items in Attachment 5 used for the transferability study are selected from the items listed in Attachment 3:
- DNCB was selected as the most classical test item used in many *in vitro* studies.
- The other chemicals were selected from those being in the ICCVAM performance standards but not in the list of Casati et al.[14]. The reason for this selection was to be able to include the chemicals of the latter list in the blind-coded phase of the study.
- The study set included three clear negatives (chlorobenzene, methyl salicylate, sulfanilamide), three clear positives (DNCB, citral and ethyleneglycol dimethacrylate) and hexyl cinnamic aldehyde as borderline test item.

1.3.2 Training required for transferring the test method

No face-to-face training was organized to transfer the method, and the SOP was found sufficiently detailed to perform the test in all the four external laboratories.
There was no training protocol defined.

1.3.3 Obstacles pertaining to transferability that are specific to the test method

There are no significant technical obstacles specific to the method. The key issues identified in the transfer phase are specific to luciferase measurements in general. These may easily be overcome and they were due to three reasons:

(a) The use of not sufficiently sensitive luminometers

(b) The use of a Glow-substrate instead of a Flash-substrate for the luciferase measurements. Flash-substrates are added by the luminometer to individual wells immediately before reading the well, while Glow-substrates are added to the entire plate, before the plate is read. Glow-substrates emit low levels of light for prolonged periods. This can contribute to lower sensitivity and thus higher variability. Long reading times for plates with the Glow-substrate caused a gradient over the assay plate in one laboratory. Nevertheless, one laboratory could successfully use the Glow-approach.

(c) The use of assay plates which do not fit the height of the luminometer. Especially using the Glow-approach this can lead to light contamination into the adjacent rows, and thus false-positive results.

(i) a good sensitivity, (ii) a low variability (and thus a stable background) and (iii) no gradient over the plate are absolutely critical to ensure accurate determinations of EC1.5 values.

If using (a) one of the luminometers recommended based on this experience and indicated in the INVITOX SOP (Attachment 1a) and (b) a substrate and plate format matching these luminometers, no further technical obstacles have to be expected.
Based on the experience of the ring study, a simple experiment was designed, which should be performed in the future as an initial part of the method transfer phase to new laboratories, to ensure that the luminescence readings yield the desired performance. This is not a training protocol already used, but defined for future studies. It is attached as Attachment 6.

Since some of the initial hurdles in the transferability appeared to be due to different luminometers, two labs rented an external luminometer and compared it to the luminometer used. These data are presented in Attachment 17d. The results indicated that the different Tecan devices give very similar results as the Promega device, and the accurate luminometer measurement is not specific to a particular brand, and also not specific to specific luminescence substrate. Thus there is clearly flexibility in these parameters as long as the conditions (i) – (iii) defined above are met.

1.3.4 Organisation of the transfer phase

Each laboratory received the cell line on dry ice along with the SOP in the format of Attachment 7a. This previous format of the SOP is very similar to the INVITOX format given in Attachment 1a, but it is included here for completeness.
The laboratories also received the ring study set-up as detailed in Attachment 7b giving some more details. Specific questions were then directly answered by phone and e-mail by the lead lab, but no meetings were organized.

Note: Although the ring-study was not under the auspices of ECVAM, the Attachment 7a, 7b and the content of Attachment 3 had been shared with S. Casati from ECVAM prior to the ring study to ensure full transparency on what was planned and finally has been done.
All the labs performed three repetitions consisting of three replicates on the substances of the transferability study and sent these data to the lead lab. Data quality was assessed by the lead lab whether they fulfil the following criteria: (i) Variability in the 18 DMSO-control wells for each triplicate experiment is below 20 % in all three repetitions (ii) dose-response curves are reproducible within the laboratory (i.e. increasing gene activation with increasing concentration up to the cytotoxic levels; EC1.5 values which are no more then one well up and down in the dilution series from the average), and (iii) the positive control cinnamic aldehyde (contained in each test plate) gives a statistically significant induction above 1.5 fold below 64 µM in all three repetitions. Once this criteria were met by the three consecutive repetitions done by a particular lab, these data were taken as the final data (shown in attachments 8a and 8b) of the transferability study, and the individual labs were allowed to move into the between laboratory reproducibility phase (described below).

1.3.5 Assessment of the transferability to another laboratory

Once the obstacles under 1.3.3 were overcome by all labs, the results provided were of good quality.
The dose-response results on the 7 chemicals tested in this phase are summarized in Attachment 8a, and the numerical evaluations are given as a Table in Attachment 8b.
The positive chemicals DNCB, citral and ethylenglycol-dimethacrylate were positive in all 4 external laboratories, and in all three repetitions.
The three negative chemicals chlorobenzene, methyl salicylate and sulfanilamide were overall negative in all labs, with few cases with a borderline induction in one of the three repetitions.
The dose-response curves clearly confirm the positive rating for the three positive chemicals and the negative rating of the three negatives (see Attachment 8a).
The borderline chemical hexyl cinnamic aldehyde was consistently positive in two labs, negative in one lab and gave a mixed result in one lab, and thus the between laboratory variability nicely reproduced a variability already observed in the lead lab.
The EC1.5 values for DNCB and citral did vary little in the four external labs and were similar to the historical and new data of the lead lab. A somewhat higher variation was observed for ethylenglycoldimethacrylate. Also for hexyl cinnamic aldehyde, the EC1.5 values in the positive repetitions are very close to each other (see Attachment 8b).

1.3.6 Quality system(s) of the other laboratory

The laboratories performed the study under non-GLP conditions, but a high standard was ensured by the fact that all the laboratories worked with the same batch of the test chemicals.

1.4 MODULE 4: BETWEEN-LABORATORY REPRODUCIBILITY (BLR)

1.4.1 Rationale for the selection of test items used for assessing BLR

As detailed above, an initial assessment of the between-laboratory reproducibility could already be made based on the results on the first seven test items tested in the transfer phase.
In the second phase, 21 more test items were tested to further evaluate the BLR. These test items were blind coded, and therefore these same results will be used below in Module 1.5 for the initial evaluation of the predictive capacity.
Attachment 9 lists these chemicals (a subset of the chemicals in Attachment 3). The rationale for selecting them has been given in section 1.2.1. and 1.3.1.

1.4.2 Assessment of reproducibility

Each lab performed three repetitions consisting of three replicates on the substances in Attachment 9 and these data were directly used as the final data, and no data were rejected at this stage. This approach was chosen in order to avoid repeated testing and to gain experience on how robust the assay is and how narrowly the acceptance criteria should be defined in the future. More narrow criteria for the EC1.5 value of the positive control cinnamic aldehyde were initially also defined, but experiments were accepted even if these criteria were not met, also to gain experience on how narrow the criteria should be applied in the future.

Attachment 10a lists all the dose-response curves for the 21 chemicals tested in the BLR phase on blind coded chemicals. Attachment 10b lists the luciferase induction data (I_{max} and EC1.5) in a tabular form. Attachment 10c gives a summary of the reproducibility of the cytotoxicity measurements. Attachment 10d gives a comparison of the WL variation for the EC1.5 values within the different laboratories and the BL variation based on the geometric standard deviation as detailed in Attachment 4. Attachment 10e presents the same analysis for the IC50 values (from both study phases).

<u>BLR for yes/no prediction:</u>
Among the 15 sensitizers tested in the BLR phase, 11 where rated positive in all three repetitions in all five labs, and for 2 there was one single negative repetition in one lab. For all these 13 chemicals the dose response curves are very clear. Phenyl-benzoate is a false-negative in the historical data and this was confirmed with three negative repetitions in four labs, with one single positive repetition in one lab.
The only chemical which gave contradicting results in the BLR phase is Eugenol. In three labs, this test item was negative in 2 out of 3 repetitions and it was negative in the historical data. Nevertheless, it was rated positive in all repetitions in two labs with very reproducible dose-response curves.
Among the 6 non-sensitizers, 4 were rated negative in all three repetitions in all the five labs. Diethylphthalate was negative in four labs, but positive in one lab. SDS was negative in four labs with 1 to 3 repetitions giving significant induction at cytotoxic concentrations only. There was one exception: in one lab induction was observed at the same concentration as in the other labs, but in this lab the cells were still fully viable in the MTT plates at the inducing concentration.

<u>BLR for EC1.5 values:</u>
As exemplified by the dose response graphs (Attachment 10a), similar overall dose-response results were obtained in the different labs. The biggest variation is observed in the dynamic

range, i.e. the maximal gene induction I_{max} can vary significantly between the laboratories. However, the EC1.5, which had been found the key parameter to quantify the luciferase response [5], appears more reproducible (see Attachment 10b) and this can be better quantified by looking at the analysis of the variance (Attachment 10d).
(Note: This assessment also includes the three positive chemicals of the transfer phase for completeness.)

- For each chemical and each laboratory the geometric standard deviation from the three repetitions was calculated.
- The average of these geometric standard deviations for WLR was then calculated for each chemical (column WLR).
- This value was in the range between 1.1. and 1.8 for the different chemicals, with an average for all positive chemicals of 1.45
- This is close to the square root of 2, and thus indicates that the 95.4% confidence interval for WLR lies, on the average, within one well up and down of the geometric mean in the dilution series.
- This value is close to 1.38 found above for the WLR in the Lead Lab and thus, on the average, the WLR is similar in all laboratories.
- Next for each chemical the geometric standard deviation of the EC1.5 over the geometric means of the individual laboratories was calculated (column BLR).
- This parameter, which is a true measure of the between laboratory avariability, varied in the range between 1.2 and 2.6 with an average of 1.64 for all chemicals (1.56 without the outlier MCI), indicating that the between laboratory variability is slightly higher as compared to the within laboratory variability. Still, for the majority of the chemicals it is below 1.41. Therefore, also for the BLR the 95.4% confidence interval of the EC1.5 value is, for most chemicals, within one well up and down from the geometric mean.

In attachment 10g a Box-plot analysis of the normalized logarithmic averages of all the EC1.5 values for all laboratories is shown. This analysis indicate that (5-chloro)-Methyl-isothiazolinone (MCI) clearly is an outlier with the highest variability in the BL-comparison as observed also in the WL comparison. Still, even with this variability and as discussed above, all results clearly indicate that this chemical induces the luciferase gene at very low concentration.

<u>BLR for cytotoxicity:</u>
The average IC50 values for all chemicals and all repetitions in each lab are listed in Attachment 10c, and in general congruent results between the labs were obtained. To quantify this, the within and between lab geometric standard deviations of the IC50 values were calculated (as done for the EC1.5 values above). These values could be calculated for those 18 test substances with a reduction of viability > 50% at any of the test concentrations and they are listed in Attachment 10e. The average of these geometric standard deviations within individual labs ranged between 1.09 and 1.81 for all of the test substances, with an average for all cytotoxic substances of 1.35, indicating that also for the WLR of the cytotoxicity values the 95.4% confidence interval is confined by a factor of less than 2. The BLR geometric standard deviations were between 1.08 and 1.94 for all test substances, with an average of 1.42. Thus, the variability of the IC50 values between the laboratories is even slightly lower as compared to the variability of the EC1.5 values.

Besides these descriptive statistics of Attachement 10d and 10e, the BLR for the IC50 values, I_{max} values and the EC1.5 values was further evaluated in a statistical analysis to evaluate which lab contributed most to error variance in the overall results. The statistical report from our statistician is attached as Attachment 10f. From this analysis it becomes apparent, that Lab 1 had a significantly lower performance in the three quantitative parameters as compared to the other labs. This is also reflected by the highest geometric within lab standard deviation for both EC1.5 and IC50 in Lab 1 as shown in Attachments 10d and 10e. However, these larger quantitative variations did not affect the yes/no prediction, and Lab 1 had for all chemicals the same yes/no prediction as the lead lab in the ring study. Also the overall dose-response curves for this lab are not very different from the other labs.

Note: It is noteworthy that Lab 1 performed poorest in this analysis. This lab had no problems with the luciferase readings in the transfer phase, and as a consequence did not have to repeat the experiments of the transfer phase several times. Hence, the research scientists in this lab had aquired little experience in running the test during the transfer phase and could quickly move to the phase II. This might be reflected in a poorer performance for the quantitative parameters, whereas the other labs had a longer training in running the method until the issues with the luminescence readings were settled. Based on this, in further validation studies, a minimal number of training experiments in naïve labs may be requested before the 'real' data are generated, yet it should be kept in mind that the observed differences are minor.

1.4.3 Identification and discussion of outlying values

Eugenol: Eugenol is a reported weak sensitizer. It was correctly identified positive in two labs with very similar dose-response curves in the two labs. It is consistently negative in the lead lab and in the other two testing labs. The reason for this is unknown.
SDS: For this chemical a positive gene induction is observed at a very narrow range of partly cytotoxic concentrations. This has been described with a detailed dose-response analysis in the publication[3]. This result was nicely reproduced in the ring study: Several labs found occasional positive results if a partly cytotoxic concentration was hit at any test concentration and the derived EC1.5 values from these experiments are very reproducible. There was one discordant result: Lab 4 reported the same EC1.5 value, but there was no cytotoxicity observed at the EC1.5 determining concentration in the parallel MTT plates, and, according to the prediction model, SDS must thus be rated as positive in this lab.
Diethylphthalate: One Lab found a positive induction. However, this induction was paralleled by a strong increase in the MTT values, which may indicate that also cell numbers increased or that the metabolic activity of the cells was increased by DEP. The reason for this is unknown.

1.4.4 Quality system(s) of the testing laboratories

The laboratories performed the study under non-GLP conditions, but a high standard was ensured by the fact that all the laboratories worked with the same batch of the test chemicals.

1.5 MODULE 5: PREDICTIVE CAPACITY (PC)

1.5.1 Rationale for the selection of test items used for assessing PC

The predictive capacity was determined at three levels:

a) based on the screening in the lead lab on the silver list of 67 chemicals used for evaluation of the prediction model. This list of chemicals is given in Table 2 of Attachment 2 and the rationale of their selection is given in section 1.1.6. Attachment 11 lists the same chemicals, but also including the information (mainly from the ICCVAM database) on the sensitization evidence based on the LLNA, GPMT, Buehler test and human evidence. The predictivity for this dataset has already been published (Emter et al., 2010).

b) based on the screening of the 28 chemicals in Attachment 3 in the ring study to evaluate the predictive capacity in the different laboratories. The rationale for selecting these chemicals is given in section 1.2.1.

c) Based on an extended list of 114 chemicals. The rationale for selecting these chemicals as well as their available *in vivo* data are all given in Attachment 12c.

d) Based on two specific studies addressing the chemical classes of epoxides and surfactants (Attachment 12d and 12e).

1.5.2 Assessment of the predictive capacity of the test method

Predictive capacity on the silver list:
Table 2 in Attachment 2 lists the positive / negative rating in the KeratinoSens test of each chemical in the silver list. The chemicals are classified as sensitizers and non-sensitizers in this Table based on the weight of evidence for the sensitization potential (based on human and animal experience summarized in Attachment 11).
Attachment 12a lists the Cooper statistics based on the data provided in Attachment 2.
Note: These Cooper statistics are slightly different from the published ones, as they are calculated with the refined prediction model rating chemicals only positive if the EC1.5 value is below 1000 µM.
The overall accuracy for this set of chemicals is 85.1%.
Attachment 12a also discusses the false-positives and false negatives. As discussed, the predictivity will be further improved by an ITS approach integrating a peptide reactivity assay.

Predictive capacity in the interlaboratory study:
Attachment 12b lists the positive (red and orange) and negative (dark and faint green) results for all the 28 chemicals in the five laboratories and the historical (silver list) data. Based on these results the Coopers statistics were calculated for all laboratories.
The accuracy was between 85.4 and 96.7% in the different laboratories. The main reason for the difference between the laboratories is the different rating of the borderline chemical α-hexyl-cinnamic aldehyde and the different results for eugenol discussed in section 1.4.3.

Predictive capacity on the extended list of chemicals
Attachment 12c gives detailed results on an extended list of 114 chemicals. This list contains some more pro-haptens and more acyl-transfer agents, but also other chemicals of diverse mechanistic applicability domains. In attachment 12c the detailed results are shown and

discussed. This list also contains the chemicals of the silver list, thus there is redundancy with the data presented in Attachment 12a.

Predictive capacity for surfactants
Attachment 12d reports an unpublished study, in which the KeratinoSens assay was used to rate a number of exemplary surfactants which give discordant results in the GPMT and the LLNA, and which are, based on a weight-of –evidence assessment, considered irritant non-sensitizers. These chemicals are correctly negative in the KeratinoSens at non-cytotoxic concentrations.

Predictive capacity for epoxides
Attachment 12e reports an unpublished study, in which the KeratinoSens assay was used to perform a structure-activity-study on a series of sensitizing epoxides. It reports a case-study how the KeratinoSens data may also be used for a read-across potency prediction within a close structural class.

1.5.3 Identification and discussion of false predictions

False predictions for the chemicals in the Silver list
False positives:
Propylparaben, known weak human sensitizer from epidemiological studies, but negative in GPMT and LLNA, thus a disputable false-positive
Ethylvanillin, contains structural alerts for skin sensitizers (catechol precursor; aromatic aldehyde), but it is negative in animal and human studies, true false positive
Dimethylisophthalate, slight induction only at 1000µM, borderline but positive with current prediction model
Tween 80, generally considered non-sensitizer, but positive in some guinea pig tests. True false positive.
False negatives:
Benzoyl peroxide;
True false-negative, positive/adduct forming in LC-MS assay
Diethylenetriamine, clear dose response in all repetitions but EC1.5 > 1000 µM, negative in current prediction model
Thioglycerol, disulfide-bridge forming chemical, positive/adduct forming in LC-MS assay, might be rapidly metabolized in the cells
Resorcinol, putative prohapten requiring P450 activation
Eugenol, putative prohapten requiring P450 activation
Phenyl benzoate, acyl transfer agent, rated positive/adduct forming in LC-MS based peptide reactivity assay, probably due to reaction with amine groups

False predictions for the extended list
Attachment 12c presents an in depth discussion on false-positives and false-negatives in the extended list, which may help to further clarify the applicability domain. The individual results are thus not further detailed here in the TST.

1.6 MODULE 6: APPLICABILITY DOMAIN (AD)

1.6.1 Applicability of the test method identified through testing

Technically, all chemicals soluble in either water or DMSO can be tested in the current protocol. Compounds with a cLogP of up to 5 were successfully tested.

Chemicals of widely differing structural classes were tested, including: Heterocyclic compounds, quinones, aryl halides, aromatic amines, aldehydes, dialdehydes, halogenated alkyl compounds, esters, α,β-unsaturated esters, -ketones, -aldehydes and -lactones, phenols, catechols, aliphatic amines, disulfides, thiols, epoxides, metals, glycols, acids, surfactants, alcohols, sugars, and amides.

1.6.2 Limitations of the test method identified through testing

Extremely hydrophobic molecules with a cLogP > 7 cannot be tested due to solubility issues in DMSO and water, testing in these cases is sometimes possible at lower maximal concentration. Only very few chemicals fall into this category.

There is no experience so far for molecules of a cLogP between 5 and 7, but these are rather rare chemicals.

Chemicals with an extremely high reactivity with DMSO or water cannot be reliably tested, as these chemicals react during preparation and handling of the solutions and the finally tested preparation is no longer the test item. This has been observed for some acid chlorides, which are used as intermediates in chemical synthesis. However, the very high reactivity of such materials is known (they are exactly used for this purpose in chemical manufacturing) and their high sensitization risk can normally be assessed with direct reactivity measurements and no cell-based sensitive test method is needed for their evaluation.

1.6.3 Suggested Applicability Domain (AD) of the test method

The majority of skin sensitizers appear to be detected based on the current dataset, but few potential skin sensitizers with an exclusive chemical reactivity towards Lysine-residues turn out to be false-negatives. This has been observed for amine-reactive chemicals such as phthalic anhydride, 3,4-dihydrocoumarine and phenyl-benzoate. These chemicals react with lysine by acyl-transfer reactions, and this is discussed in more detail in Attachment 12c. Also a minority of the skin sensitizing aldehydes, such as Lilial, which are thought to react with amine groups of proteins through Schiff-base formation, are negative.
Thus, the assay cannot be used for anhydrides as a general group of chemicals. Otherwise there is no clear *a priori* criterium to predict whether a chemical is selectively amine-reactive, and thus no clear structural limitation of the applicability domain in relation to this criterium can be given. Yet, purely amine-reactive chemicals generally are a small minority of the known reactive chemicals, as amine-reactivity in most cases is accompanied with an even stronger reactivity with the more reactive thiol group. Nevertheless, it is recommended that in an integrated testing strategy, the test should be combined with a reactivity assay, which recognizes specific amine-reactive chemicals (see Attachment 12c).

The test can recognize a variety of pro-haptens such as α,β-unsaturated alcohols like cinnamic alcohol and geraniol or amines such as ethylene-diamine. Some phenolic pro-haptens which are thought to require an activation step by P450 enzymes are not detected with the test. Typical examples are eugenol, resorcinol and creosol. This is discussed in more detail in Attachment 12c. Attempts to overcome this limitation by inducing endogenous P450 enzyme in the KeratinoSens assay have failed so far. CYP1A1 activity was successfully induced, but this did not affect the results for these putative prohaptens.
The test can recognize a variety of pre-haptens such as 1,4-phenylenediamine, hydroquinone and isoeugenol. Pre-haptens with a slower rate of spontaneous oxidation such as limonene are reliably recognized by including an oxidation step (exposure to O_2 for 3 – 7 days) prior to the experiment (See Attachment 4c for an illustrative example).
We expect that the applicability domain is wider than experimentally tested so far. Thiol-reactive skin sensitizers, which are stable and bioavailable under the test conditions, will most likely be recognized, even if they have a reaction mechanism and structural features which are different from the chemicals tested. The rationale for this being, that the Nrf2-pathway is a sensor mechanisms not responding to specific spatial features of chemicals (lock-and-key recognition) but it rather responds to reactivity of chemicals, and therefore the diversity of chemicals activating this pathway is probably much wider then currently known.

1.7 MODULE 7: PERFORMANCE STANDARDS (PS)

Not applicable at current stage

2. ESSENTIAL INFORMATION FOR A SPECIFIC VALIDATION PROCESS

Not applicable

3. ADDITIONAL INFORMATION

3.1 Additional information

None provided

3.2 List of references

1. Roberts, D. W.and Aptula, A. O., (2008) Determinants of skin sensitisation potential. Journal of Applied Toxicology 28, 377-387.
Review highlighting the importance of reactivity for skin sensitization

2. Dinkova-Kostova, A. T.,Holtzclaw, W. D.and Kensler, T. W., (2005) The role of Keap1 in cellular protective responses. Chemical Research in Toxicology 18, 1779-1791.
Review on the Nrf2-Keap1-ARE pathway

3. Emter, R.,Ellis, G.and Natsch, A., (2010) Performance of a novel keratinocyte-based reporter cell line to screen skin sensitizers in vitro. Toxicol Appl Pharmacol 245, 281-290.
Description of the development of the KeratinoSens assay

4. Natsch, A.and Emter, R., (2008) Skin sensitizers induce antioxidant response element dependent genes: Application to the in vitro testing of the sensitization potential of chemicals. Toxicol. Sci. 102, 110-119.
The first report on the value of using an ARE-reporter cell line to screen skin sensitizers

5. Natsch, A.,Emter, R.and Ellis, G., (2009) Filling the concept with data: Integrating data from different in vitro and in silico assays on skin sensitizers to explore the battery approach for animal-free skin sensitization testing. Toxicol. Sci. 107, 106-121.
An extended database on results in the AREc32 cell line and the concept of combining the assay results with results from the DPRA in an ITS

6. Vandebriel, R. J.,Pennings, J. L.,Baken, K. A.,Pronk, T. E.,Boorsma, A.,Gottschalk, R.and Van Loveren, H., Keratinocyte gene expression profiles discriminate sensitizing and irritating compounds. Toxicol Sci.
A very detailed gene chip study in the HaCaT cellular background proving that Nrf2-regulated genes are the most stable markers discriminating sensitizers from irritants in keratinocytes

7. Ade, N.,Leon, F.,Pallardy, M.,Peiffer, J. L.,Kerdine-Romer, S.,Tissier, M. H.,Bonnet, P. A.,Fabre, I.and Ourlin, J. C., (2009) HMOX1 and NQO1 genes are upregulated in response to contact sensitizers in dendritic cells and THP-1 cell line: Role of the Keap1/Nrf2 pathway. Toxicol. Sci. 107, 451-460.
Detailed mechanistic proof, that sensitizers (i) directly regulate Nrf2 accumulation in the nucleus and (ii) that the pathway is also activated in dendritic cells

8. Python, F.,Goebel, C.and Aeby, P., (2009) Comparative DNA microarray analysis of human monocyte derived dendritic cells and MUTZ-3 cells exposed to the moderate skin sensitizer cinnamaldehyde. Toxicol Appl Pharmacol.

A detailed gene chip study in primary dendritic cells and the THP-1 cell line. Only very few genes are common markers between the two models, and only four genes were found as reliable markers, two of these are Nrf2-regulated.

9. Kim, H. J.,Barajas, B.,Wang, M.and Nel, A. E., (2008) Nrf2 activation by sulforaphane restores the age-related decrease of T(H)1 immunity: role of dendritic cells. J Allergy Clin Immunol 121, 1255-1261 e7.

This paper shows that Nrf2 is essential in vivo to fully express the skin sensitization reaction, and esp. for the induction of the key mediator IFN-γ by sensitizers.

10. Gildea, L. A.,Ryan, C. A.,Foertsch, L. M.,Kennedy, J. M.,Dearman, R. J.,Kimber, I.and Gerberick, G. F., (2006) Identification of gene expression changes induced by chemical allergens in dendritic cells: Opportunities for skin sensitization testing. Journal of Investigative Dermatology 126, 1813-1822.

This paper reports the first gene chip study on sensitizer induced gene changes. Among the identified markers, AKR1C2 and a few others are Nrf2 regulated, although this link was not yet made at the time of publication.

11. Ryan, C. A.,Gildea, L. A.,Hulette, B. C.,Dearman, R. J.,Kimber, I.and Gerberick, G. F., (2004) Gene expression changes in peripheral blood-derived dendritic cells following exposure to a contact allergen. Toxicology Letters 150, 301-316.

This paper reports a follow up on ref 10 with RT-PCR. AKR1C2 was found as one of the best markers.

12. 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, 284-92.

Review article, specifically highlighting the importance of the Nrf2 pathway and the potential difference of SH and NH_2 reactive chemicals when it comes to activation of innate responses by sensitizers.

13. Wang, X. J.,Hayes, J. D.and Wolf, C. R., (2006) Generation of a stable antioxidant response element-driven reporter gene cell line and its use to show redox-dependent activation of nrf2 by cancer chemotherapeutic agents. Cancer Res 66, 10983-94.

Report on the generation of the AREc32 cell line

14. Casati, S.,Aeby, P.,Kimber, I.,Maxwell, G.,Ovigne, J. M.,Roggen, E.,Rovida, C.,Tosti, L.and Basketter, D., (2009) Selection of chemicals for the development and evaluation of in vitro methods for skin sensitisation testing. Altern Lab Anim 37, 305-12.

List of standard chemicals for a first evaluation of in vitro assays

15. Basketter, D. A.and Cadby, P., (2004) Reproducible prediction of contact allergenic potency using the local lymph node assay. Contact Dermatitis 50, 15-7.

4. ATTACHMENTS

Attachment Number	Description	Tick if attached	File name
Attachment 1	Protocol(s) of the test method Example results evaluated with the SOP template	X	Attachment1a_INVITTOXProtocolKeratinoSens_revised_clean.DOC Attachment1b_SOP_calculation_revised.xls Attachment1c_SOP_calculation_IntraLabStudy_Plate1_Experiment1_new.xls
Attachment 2	List of test items used to develop and optimise the test method, including their CAS number and basic physical/chemical properties	X	Attachment2_Test Chemicals Results KeratinoSens.DOC
Attachment 3	List of test items used to assess WLR, including their CAS number and relevant properties	X	Attachment3_WLR_chemicals.pdf
Attachment 4	Data used for WLR assessment (Tables and/or Figures)	X	Attachment4a_WLR_revised.pdf Attachment4b_WLR_stat-analysis_revised.pdf Attachment4c_second intralab study_new.pdf Attachment4c_second intralab study_appendix-dose-response_new.pdf
Attachment 5	List of test items used to assess transferability, including their CAS number and relevant properties	X	Attachment5_transferability_chemicals.pdf
Attachment 6	Training protocol	X	Attachment6_Transfer_phase experiment.doc
Attachment 7	Transfer protocol	X	Attachment7a_RingStudy_SOP_KeratinoSens_ Natsch_Version1.1.doc Attachment7b_Ring study Setup.doc
Attachment 8	Transfer report	X	Attachment8a_Transferability_dose-response_revised.pdf Attachment8b_Transferability_Table_revised.pdf
Attachment 9	List of test items used to assess BLR, including their CAS number and relevant properties	X	Attachment9_BLR_chemicals.pdf
Attachment 10	Data used for BLR assessment (Tables and/or Figures)	X	Attachment10a_BLR_dose-response_revised.pdf Attachment10b_BLR_Table_revised.pdf

			Attachment10c_BLR_Cytotoxicity_revised.pdf Attachment10d_BLR_comparision of variance_revised.pdf Attachment10e_BLR_comparision of variance_IC50_new.pdf Attachment10f_BLR_Statistical_report_new.pdf Attachment10g_BLR_comparision of variance_outlier analysis_new.pdf
Attachment 11	List of test items used to assess PC, including their CAS number and relevant properties	X	Attachment11_Chemicals_for_PC.pdf
Attachment 12	Data used for PC assessment (Tables and/or Figures)	X	Attachment12a_PC-SilverList.pdf Attachment12b_BL-PC.pdf Attachment 12c_PC-extended list_new.pdf Attachment 12d_PC-surfactants_new.pdf Attachment 12e_PC-epoxides_new.pdf
Attachment 13	Project plan(s)		
Attachment 14	Information/data collected for a retrospective validation study		
Attachment 15	Performance standards for validation of a similar or updated test method.		
Attachment 16	References	X	Attachment16_ring_study_paper_reprint_new.pdf Attachment16_Ref1-13; 13 pdf attachments
Attachment 17	Additional attachments: Analysis of the quality criteria data (17a-c) and comparison of luminometers (17d)	X	Attachment17a_Positive control_WLR_revised.pdf Attachment17b_Positive control_BLR_revised.pdf Attachment17c_Positive control_BLR_outliers.pdf Attachment17d_luminometer_comparison.pdf