

PROF. MARIE BAECK (Orcid ID : 0000-0003-0499-7939)

DR. MARC VOCANSON (Orcid ID : 0000-0002-0181-8862)

Article type : Original Article: Basic and Translational Allergy Immunology

Unique molecular signatures typify skin inflammation induced by chemical allergens and irritants

Manuscript Acceptance Date: 09-Jun-2021

Marine-Alexia LEFEVRE MD^{1-2†}, Audrey NOSBAUM MD/PhD^{1-3†}, Aurore ROZIERES PhD¹, Vanina LENIEF MSc¹, Amandine MOSNIER MSc¹, Angèle CORTIAL PhD¹, Margaux PRIEUX MSc¹, Simon De BERNARD PhD⁴, Julien NOURIKYAN PhD⁴, Pierre Emmanuel JOUVE PhD⁴, Laurent BUFFAT PhD⁴, Florence HACARD MD¹, Marie-Christine FERRIER-LEBOUEDEC MD⁵, Pauline PRALONG MD⁶, Charles DZVIGA MD², Anne HERMAN MD⁷, Marie BAECK MD/PhD⁷, Jean-François NICOLAS MD/PhD^{1-3†}, Marc VOCANSON PhD^{1†}

† These authors contributed equally to this work

Affiliations:

¹ CIRI, Centre International de Recherche en Infectiologie, (Team Immunology of Skin Allergy and Vaccination); INSERM, U1111; Univ Lyon; Université de Lyon 1; Ecole Normale Supérieure de Lyon; CNRS, UMR 5308, Lyon, France.

² Department of Dermatology and allergology, Centre Hospitalier Universitaire de Saint-Etienne, Saint-Priest-en-Jarez, France Department of Allergy and Clinical Immunology, Centre Hospitalier Lyon-Sud, Pierre-Benite, France.

³ Department of Allergy and Clinical Immunology, Centre Hospitalier Lyon-Sud, Pierre-Benite, France.

⁴ Altrabio, F-69007 Lyon, France.

⁵ Department of Dermatology, Centre Hospitalier Universitaire de Clermont-Ferrand, Clermont-Ferrand, France.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/ALL.14989](https://doi.org/10.1111/ALL.14989)

This article is protected by copyright. All rights reserved

⁶ Department of Dermatology, Allergology and Photobiology, Centre Hospitalier Universitaire Grenoble Alpes, La Tronche, France

⁷ Department of Dermatology, Cliniques universitaires Saint-Luc, Université Catholique de Louvain, Brussels, Belgium

Correspondence:

Marc Vocanson, CIRI-INSERM U1111, 21 avenue Tony Garnier 69007 Lyon, France

Tel: +33 437 28 23 48; e-mail address: marc.vocanson@inserm.fr

Short title: ACD and ICD molecular signatures.

Abbreviations

ACD: Allergic contact dermatitis

CD: Contact dermatitis

CTL: cytotoxic CD8⁺ T cell

DEG: Differentially expressed gene

Fc: Fold change

FDR: False discovery rate

GZMB: Granzyme B

GNLY: Granulysin

ICD: Irritant contact dermatitis

IFNG: Interferon gamma

MI: Methylisothiazolinone

PCA: Principal Component Analysis

PT: Patch test

SLS: Sodium lauryl sulphate

Teff: Effector T cells

Tmem: Memory T cells

ABSTRACT

Background

Skin exposure to chemicals may induce an inflammatory disease known as contact dermatitis (CD). Distinguishing the allergic and irritant forms of CD often proves challenging in clinic.

Methods

To characterize the molecular signatures of chemical-induced skin inflammation, we conducted a comprehensive transcriptomic analysis on the skin lesions of 47 patients with positive patch tests to reference contact allergens and non-allergenic irritants.

Results

A clear segregation was observed between allergen- and irritant-induced gene profiles. Distinct modules pertaining to the epidermal compartment, metabolism and proliferation were induced by both contact allergens and irritants; whereas only contact allergens prompted strong activation of adaptive immunity, notably of cytotoxic T cell responses. Our results also confirmed that: (i) unique pathways characterize allergen- and irritant-induced dermatitis; (ii) the intensity of the clinical reaction correlates with the magnitude of immune activation. Finally, using machine-learning approach, we identified and validated several minimal combinations of biomarkers to distinguish contact allergy from irritation.

Conclusion

These results highlight the value of molecular profiling of chemical-induced skin inflammation for improving the diagnosis of allergic versus irritant contact dermatitis.

KEYWORDS: Allergic contact dermatitis, contact allergy, patch testing, transcriptomic profiling, chemical allergens, chemical irritants.

INTRODUCTION

Contact Dermatitis (CD), or contact eczema, is a skin disease affecting approximately 15% to 20% of the general population in industrialized countries (1,2). CD is the most frequent occupational disease in many European countries (3). The direct (treatment and compensation) and indirect (sick leave and lack of productivity) costs associated with the disease have resulted in the economic burden of CD being estimated at over 5 billion Euros per year in Europe (4). CD typically presents as acute inflammation with itching, redness, edema, vesicles and oozing, as well as thickening and crusting in its chronic form. CD is induced by repeated skin contacts with external agents. A wide range of substances have been implicated in the disease, notably chemicals with irritant and/or allergenic properties, which are responsible for the irritant (ICD, (5)) and allergic (ACD, (6)) types of CD respectively. Available data indicate that ICD accounts for approximately 70% to 80% of all cases of occupational CD in industrialized countries (7).

The mechanisms by which chemicals cause ICD are poorly understood, and vary from disorganization of cell membrane lipid bilayers to damage of barrier proteins (8). ICD may be caused by acute exposure to corrosive substances such as acids and bases, which trigger intense cell necrosis and major skin barrier disruption, or by chronic exposure to mild irritants such as detergents (9). In both cases, irritation is most likely triggered by injured cells releasing reactive oxygen species (ROS) and ATP in response to stress, and generating damage-associated molecular patterns (DAMPs) involving high-mobility group box 1 (HMGB1), heat-shock proteins, and interleukin (IL)-1 α . These alterations are then detected by the innate receptors of surrounding healthy cells (10), resulting in the release of a myriad of inflammatory mediators (cytokines, chemokines, arachidonic acid derivatives, proteases, etc.) within minutes to hours after contact (11,12). The reaction is further amplified by infiltration from the blood of inflammatory monocytes, neutrophils, etc. (13).

ACD is the prototype of a type IV delayed-type hypersensitivity reaction, which develops in individuals who have previously been sensitized to chemical allergens (also referred to as haptens). The most common clinically relevant allergens associated with occupational or nonoccupational exposure include protein-reactive compounds, such as metals, fragrances, preservatives or rubbers (14). ACD is mediated by the recruitment and reactivation of allergen-specific effector/memory T cells (Teff/mem) – key cells involved in the immune response to allergens and skin inflammation (15). Akin to irritants, certain contact allergens have toxic/proinflammatory properties and activate danger recognition platforms (Toll-like and NOD-

like receptors) through the release of stress and DAMP signals to induce T-cell sensitization and ACD development (16).

Although ICD and ACD involve different pathogenic molecular mechanisms (17,18), their symptoms and presentation are often similar, making it difficult to distinguish between the two CD forms without performing allergology tests, such as patch tests (19). Although the robustness of patch tests has long been established, this method can sometimes give inconclusive results, leading to problems in disease management. This is particularly common for patients with highly sensitive skin: these patients mount positive responses both to reference allergens and to noninflammatory doses of irritant controls, leading to questions about the specificity of allergen-induced reactions. A recent 20-year retrospective analysis revealed that 30% to 40% of weak positive patch test reactions were not reproduced on retesting (20,21). This suggests a lack of persistence of skin allergy in some patients, or that some of the skin test results were false positives, occurring due to the irritant properties of chemicals applied for several days under occlusion (22).

There is now great interest in developing new diagnostic approaches that could help clinicians make a reliable diagnosis of CD. A recent study concluded that an immunohistochemical phenotyping approach, based on a minimal set of markers (CD3, CD4, CD8, etc.), was not reliable enough to discriminate between the cellular events shaping the ACD and the ICD reactions (23). Alternatively, several investigators have explored the molecular signatures induced by clinically relevant chemical allergens in human skin (24,25) (26–28). They identified pathways in CD that were distinct from those occurring in dermatoses such as psoriasis and atopic dermatitis (24,29), and Fortino *et al.* recently proposed that gene profiling could be of major value for discriminating between ICD and ACD (28).

In this study, we then conducted a comprehensive analysis to compare the reactions induced by a large set of allergens (n=6) and irritants (n=3), as well as different degrees of clinical reactions, in order to identify ACD and ICD biomarkers robust enough to allow for the variability of clinical contexts.

Our results identified the core signatures of allergen- and irritant-induced skin inflammation, opening the way for the development of new approaches to improve ACD diagnosis.

METHODS

A brief description of the Methods is provided below. Additional details about study design, gene arrays, qRT-PCR experiments and bioinformatic and biostatic analyses are provided in the Supplemental Methods section.

Patients suspected for skin allergy were enrolled in the study after being referred to the hospital for routine diagnosis of contact dermatitis. They were patch tested either with the contact allergens: nickel, methylisothiazolinone [MI] and linalool hydroperoxide, or the drug allergens: amoxicillin, piperacillin/tazobactam, and acetazolamide, and/or the irritants: SLS, nonanoic acid and cantharidin (**Tab.1**). At 72 hours, the patch test sites were evaluated for reactivity, and reactions were clinically graded as negative (-), doubtful (+/?) or positive (1+, 2+ or 3+) (19). Biopsies (3x3 mm) were collected from positive and negative patch-tested skin, as well as from control areas including nonlesional skin or vehicle (petrolatum or isopropanol) patch-tested skin. Total RNA was then extracted from the biopsies and used for Affymetrix PrimeView Human Gene Expression array and qRT-PCR analysis.

The study was approved by the local ethics committee (N°19-145 CHU Lyon), and written informed consent was obtained from each participant.

RESULTS

Forty-seven patients suspected for ACD (34 females and 13 males; aged 29–88 years, median 61 years) were enrolled in this study and received patch tests (PT) to different chemicals. Overall, thirty-six patients developed positive reactions to contact allergens (nickel, n=10; MI, n=6; linalool hydroperoxide, n=10) and/or chemical irritants (SLS, n=7; nonanoic acid, n=4; cantharidin, n=7). These patients were biopsied and their samples subjected to extensive genomic and molecular profiling. In parallel, we biopsied 11 patients with suspected drug allergies, who developed positive patch tests to drug allergens (amoxicillin, n=6; acetazolamide, n=1; piperacillin/tazobactam, n=1) and/or irritants (SLS; n=4). These patients were included in the study because patch test reactions to drugs are thought to exhibit clinical, histological and pathophysiologic mechanisms similar to those of reactions to contact allergens.

Patient clinical features and patch test scores are detailed in **Tab.1**. All patch test reactions were positive (graded as 1+, 2+, 3+), except for two patients with doubtful (+/?) reactions to linalool hydroperoxide and MI, and five patients with negative reactions to SLS.

Molecular profiling discriminates between contact allergy and irritation

We first executed a hypothesis-free principal component analysis (PCA) of data from the full set of genes measured on the arrays. This PCA demonstrated that all allergen and irritant samples segregated from their respective controls, with the exception of the negative SLS patch test reaction samples (**Fig.1**). This PCA also showed that the expression profiles of the majority of allergen samples (including those of the drug allergens) diverged from those of the irritant samples. All allergen and irritant samples were on the same side of PC1, suggesting that gene expression was partially shared. However, PC1/PC2 analysis revealed clear differences in expression (**Fig.1**). Interestingly, the expression profiles of nickel, linalool hydroperoxide and amoxicillin samples differed from each other. No differences were observed when nickel samples were compared with MI samples. Similar differential profiles were observed with the irritant molecules, notably between cantharidin and SLS (**Fig.1**).

Although differences in the magnitude of the respective patch test reactions between sample specimens (**Tab.1**) may have had an impact on the observed PCA variations, these results confirm that the molecular signatures of positive patch test reactions to allergens and irritants are distinct, and that molecule-dependent profiles exist within each group (24–28).

Main signaling pathways differentiating allergen- and irritant-induced inflammatory responses

We then searched to define the genes and signaling pathways that differed between the allergen and irritant samples. We first evaluated the number of differentially expressed genes (DEGs) between samples from each patient in each group and their respective controls. The five SLS samples that did not exhibit positive clinical responses were excluded. The expression of 2371 and 699 genes was found to be modulated consistently among the 34 allergen and 17 irritant samples, respectively (**Tab.S1 & Tab.S2**). Assuming a common response for all the tested molecules in a given group, we subsequently estimated the average expression level of each DEG, after correction for the donor effect, to establish whether this modulation was specific to the allergen group or the irritant group, or present in both groups. We detected 1956 DEGs predominantly modulated in the allergen group, 255 DEGs predominately modulated in the irritant group, and 457 DEGs that were modulated in both groups (**Fig.2**).

Gene ontology (GO) annotations revealed several commonly regulated pathways pertaining to the epidermal compartment (“peptide cross-linking” or “extracellular matrix assembly”), involved in metabolism (“cellular response to ketone” or “response to fatty acid”) or implicated in cell proliferation (“sister chromatid segregation”, “negative regulation of cell division” or “regulation of centrosome cycle”) (**Fig.3A and Tab.S3**). Pathways predominantly modulated in the irritant group were mainly related to cell proliferation. Among the most significant hits were “DNA replication”, “mitotic nuclear cell division” or “deoxyribonucleotide metabolic process” (**Fig.3B and Tab.S3**). Finally, GO terms most commonly associated with the allergen group were related to immune pathways (“inflammatory pathways”, “leukocyte activation”, “response to virus”, “regulation of apoptotic signaling pathway”, “regulation of type-I interferon” etc.), but GO terms associated with chemical detoxification and the epidermal compartment were also identified, with significant downregulation of genes associated with the “cornification” pathway (**Fig.3C and Tab.S3**).

Thus, distinct signaling pathways are activated in allergen and irritant patch test responses.

The core transcriptome of allergen-induced reactions mainly consists of T-cell-related transcripts

By comparing DEGs from the allergen and irritant samples, we were then able to define a core transcriptome specific for each reaction. We first selected genes with the highest positive fold

change (Fc) and false detection rate (FDR) responses, i.e. genes with the highest and most reproducible differential expression among all the samples from the respective groups. The allergen and irritant transcriptomes are illustrated in a heatmap (**Fig.4**).

The irritant transcriptome consisted of several cell cycle-related genes, such as E2F8, MNS1, PSCR1, MLF1IP, FAM64A, SAPCD2 or RRM1. It also included genes involved in epithelial cell differentiation, including GPSM2 or CLCA2. Among other irritant genes, we also detected genes associated with retinoic acid metabolism, such as CRABP2 and RDH12, and with cellular energy metabolism, like IGFL3 (**Fig.4 & Tab.S2**). Importantly, these markers were not upregulated in the five SLS-negative reactions.

The allergen transcriptome consisted of (i) several T-cell-related transcripts, including CD2, CD3D, CD3E, TRAC or TCRBC1 (**Fig.4 & Tab.S1**); (ii) DEGs down-stream of TCR signaling, such as LCK, THEMIS, FYB, TRAT1 or LCP2 and also (iii) numerous DEGs involved in the activation and differentiation of type-1 cytotoxic CD8⁺ T cells (CTLs), comprising CD8A, IFNG, CCL5, IL2RB, GZMA, GZMB, GZMK, GNLY, PRF1, TBX21, EOMES, RUNX3, STAT1, STAT4, IKZF1, BAFT3, CD69, CD300A, ICOS, CTLA4, KLRC1, KLRD1, NKG7, CST7, PLEK, SLAMF1 or SRGN. Among other DEGs, we found (iv) several markers known to be up-regulated by IFN- γ cytokine on surrounding cells, like HLA-DMA, HLA-DPA1, HLA-DQB1, HLA-DRA, CD40, CD74, CD86, IL-15, CSF2RB, IL10RA, CCL2, CXCL9, CXCL10, CXCL11, IDO1, ICAM1, VCAM1, IRF5, IRF8, PDE4B, TNFAIP3 or IFI30 (**Fig.4; Tab.S1**).

A previous gene-expression study associated the clinical intensity of patch test reactions to contact allergens with the immune response (25). To confirm these results and extend our analysis to irritant triggers, we estimated the average level of variation in core gene expression for each sample independently. We observed that core marker expression correlated with the intensity of the clinical symptoms, for both allergens (**Fig.5A**) and irritants (**Fig.5.B**). Of note, correlation was not only detected at the global level (including the 52 and the 15 genes of the allergen and irritant core signatures, respectively), but also at individual gene level (**Fig.S1**).

Taken together, these data confirm that unique biomarkers characterize the inflammatory reactions induced by chemical allergens and irritants, and that the expression of these markers strongly correlates with the magnitude of the clinical responses.

A small set of biomarkers allows for robust discrimination between allergen- and irritant-induced skin inflammation

Finally, we sought to identify a minimal set of biomarkers susceptible to discriminate between the two types of skin inflammation. To this end, we applied a biomarker discovery approach (using Boruta and Random Forest classification algorithms and k-fold cross validation) to select and assess the performance of different genes or gene combinations (here referred to as classifiers) (30). We identified more than 4000 classifiers capable of correctly classifying the different allergen or irritant samples with an accuracy ranging from 90% to 100%. The first 30 are shown in **Tab.2**. Each classifier combined between 1 to 5 different genes, identified among 15 of the major DEGs of the core transcriptomes: GPR183, IGFL3, PLEK, IL15, HLA-DRA, GPR65, SAMS1, HLA-DPA1, RAB8B, MCOLN2, CD69, HLA-DMA, CD74, HLA-DPB1.

To confirm the predictive potential of the respective models, we next evaluated the best performance classifiers on a previously published (27) and newly generated (Ljungberg Silic L. *et al*, manuscript in preparation) gene array datasets. Majority of the classifiers achieved very good prediction results, correctly classifying the allergen (nickel (27) and amerchol AL-101 (Ljungberg Silic L. *et al*, manuscript in preparation) or the irritant (SLS (27)) origin in approximately 90% and 66% of patch-test samples respectively (**Tab.3**). The predictive values of irritant samples from Wisgrill *et al.* were slightly lower, which could be explained by the fact that these investigators used different gene arrays (the sensitivity of Fold change measures vary depending on gene arrays) (27). We then plotted ROC (Receiver Operator Characteristic) curves, which compare the sensitivity and specificity metrics of the classifiers according to different, and not a unique, decision thresholds. The AUC values derived from ROC curves showed that nickel and SLS samples from Wisgrill *et al.* are in fact highly discriminated by most of the classifiers (**Tab.3**), as illustrated for the GPR183 / IGFL3 couple in **Fig.S2**.

Importantly, several of the biomarker genes detected in this study were validated by quantitative RT-PCR, confirming the differential expression recorded in a large majority of our gene array samples (**Fig.6 & Tab.S4**).

Collectively, these results therefore indicate that a small set of biomarkers is sufficient to distinguish between allergen and irritant patch test responses.

DISCUSSION

The main objective of our study was to compare the molecular responses induced by a range of chemicals with distinct proinflammatory properties, with the specific aim of identifying skin biomarkers for contact allergy. Our results clearly showed that the transcriptome of CD/skin inflammation induced by clinically relevant allergens differed from that induced by reference irritants, with core gene modules that were induced by both allergens and irritants. In addition, we confirmed that unique pathways are preferentially activated/inhibited depending on the nature of the tested molecules, with the magnitude of gene activation correlating with the intensity of the patch test-induced clinical inflammation. Finally, using machine-learning approach, we identified and validated several minimal biomarker combinations, which demonstrated excellent performance to distinguish between contact allergy from irritation.

Major hallmarks of the allergen signature

Skin allergy pathophysiology has been studied extensively over the past few decades, both in humans and in models. In contrast, far less is known about the skin inflammation induced by chemical irritants, especially in humans. Several studies had reported that transcripts from chemokines or innate sensors involved in antiviral responses and cell death, distinguished hapten- from irritant-induced responses (26–28,31–33). Our results confirmed the potential of CXCL9, CXCL10 and ZBP1 Zbp1, and, more generally, of the antiviral and type I interferon-related pathways (**Fig.3**) as factors for discriminating between allergic- and irritant-induced skin inflammation. As expected, our results revealed that allergen, but not irritant, core signatures were enriched with numerous T-cell-related transcripts, which we attributed to the differentiation of type-1 cytotoxic CD8⁺ T cells. Indeed, we detected a significant increase in CD8A transcripts, and to a lesser extent CD8B transcripts, in the allergen reactions, but detected relatively few CD4 transcripts (**Fig.2, Fig.4 & Fig.6**). Moreover, we found no differential expression for several NK or gamma delta T-cell-specific markers (such as NCR1, NCAM1, STYK1, TCRD or TCRG; **data not shown**), suggesting that the cytotoxic markers identified were mostly related to CD8⁺ T lymphocytes. These findings confirm those of previous studies performed in humans and mice, which showed that numerous cytotoxic CD8⁺ T cells infiltrate the skin of allergic individuals to instigate the formation of eczematous lesions (34–37). In another recent study, we showed that these cells persist for months in previously affected skin sites, as resident memory T cells, precipitating eczema relapses (38). Thus, the infiltration of cytotoxic CD8⁺ T cells represents a

major hallmark of the allergen signature, which, in addition, correlates strongly with the magnitude of the clinical response (**Fig.5A**).

Major hallmarks of the irritant signature

In contrast to allergic reactions, the most significant hits recorded for irritant-induced inflammation were associated with cell proliferation. This was expected as irritants trigger potent keratinocyte damage, followed by rapid and intensive tissue renewal (17,39,40). The occurrence of intense tissue remodeling was supported by our observation of strong upregulation of genes associated with cell metabolism (RDH12, CRABP2, IGFL3 etc.), peptide crosslinking (LCE5A, LCE3D, SPPR2A, etc.) and extracellular matrix assembly (MFAP4, MYH11, FBLN5, etc.), with the later genes being up-regulated in both irritant and allergen samples (**Tab.S1 & S2**). In contrast, we detected relatively few GO terms related to “inflammatory pathways”, which is surprising because SLS and, more commonly, cantharidin induced strong, and even bullous, reactions in some patients (**Tab.1**). This finding may be explained by the long time interval before biopsy collection (72 hours post application). In the hours following exposure, irritants trigger keratinocyte release of many cytokines (IL-1 β , IL-6 or TNF- α), chemokines (IL-8/CXCL8, CCL2, CCL3, CCL4, CCL17, and CCL20) and arachidonic acid derivatives (such as prostaglandin E2) (11,12,33,41). At 72 hours, the molecular inflammation would be tightly regulated. Nevertheless, we detected upregulation of several DEGs, such as IL-36G, IL-8/CXCL8, CCL2, CCL17, CCL18, MIF, and diverse HMGB or S100 proteins. All these mediators have been reported to exert crucial roles by stimulating leukocyte recruitment and cytokine production by keratinocytes (42–45), suggesting that they participate actively in late irritant inflammation. Of note, few cantharidin samples (patients 1, 4, 5) also showed marks of allergen signature. It will be interesting to determine whether it is the result of bystander recruitment and activation of cytotoxic T cells by the inflammatory environment and cytokine signaling, as reported upon microbial infection (46,47).

The predictive value of molecular signatures to improve the diagnosis of ACD

Beyond gene profiling of allergic and irritant reactions, another objective of this work was to identify a set of predictive biomarkers susceptible to serve as future platforms for improving clinical diagnosis. Using random forest classifier methodology, we identified and validated several classifier models, including 1 to 5 different genes of the major DEGs of the core transcriptomes, as minimal biomarkers that distinguish allergic and irritant reactions in human skin. The performance of these classifiers confirms a recent study conducted by Fortino al. on similar

material (including 89 biopsies collected from positive and negative patch-test reactions to 4 different allergens and 2 irritants), and which identified 22 potential biomarkers (CD47, BATF, FASLG, RGS16, SYNPO, SELE, PTPN7, WARS, PRC1, EXO1, RRM2, PBK, RAD54L, KIFC1, SPC25, PKMYT, HISTH1A, TPX2, DLGAP5, TPX2, CH25H, and IL37). Interestingly, if none of the biomarkers identified by Fortino et al. were present among our top 50 of the most discriminant genes, possibly due to key differences in experimental setups, most of them were significantly upregulated in the two sample groups (**Tab.S1-S3 and data not shown**). This illustrates that multiple biomarker combinations can be used to discriminate between the two types of skin inflammation.

Nevertheless, whatever the final combination chosen, it will be crucial to demonstrate in a future study that it is sufficiently efficient in detecting false positive patch-test reactions, which are triggered by the pro-inflammatory properties of allergens in non allergic individuals, and which remains the main problem encountered in the clinic. Indeed, all the analyses performed in this study were based on the assumption that none of collected data for chemical allergens were the results of false positive reactions. Thus, complementary trials are needed to determine whether biomarkers identified in this study are robust and sensitive enough to detect patients with clinically relevant ACD.

Finally, it will also be important to determine if the same gene signatures characterize both 72-hour patch test reactions and clinical lesions of active eczema. The detection of relevant allergy biomarkers in contact dermatitis lesions could help prevent patients being sent for unnecessary allergology workup. Currently, it is estimated that only one in five patients consulting for ACD is truly allergic (48), with the majority of these patients having ICD.

In conclusion, our study characterized the main features of allergic and irritant reactions to chemicals. We have identified several biomarkers that clearly distinguish the two types of inflammation. Our findings provide the basis for the development of new approaches to refine and improve the diagnosis of ACD.

REFERENCES

1. Svensson A, Ofenloch RF, Bruze M, Naldi L, Cazzaniga S, Elsner P et al. Prevalence of skin disease in a population-based sample of adults from five European countries. *Br J Dermatol* 2018;**178**:1111–1118.
2. Diepgen TL, Ofenloch RF, Bruze M, Bertuccio P, Cazzaniga S, Coenraads P-J et al. Prevalence of contact allergy in the general population in different European regions. *Br J Dermatol* 2016;**174**:319–329.
3. Ring J. Occupational skin disease - a major health problem in Europe. *J Eur Acad Dermatol Venereol* 2017;**31**:919–920.
4. Sartorelli P, Kezic S, Larese Filon F, John SM. Prevention of occupational dermatitis. *Int J Immunopathol Pharmacol* 2011;**24**:89S-93S.
5. Chew A-L, Maibach HI. Occupational issues of irritant contact dermatitis. *Int Arch Occup Environ Health* 2003;**76**:339–346.
6. Peiser M, Tralau T, Heidler J, Api AM, Arts JHE, Basketter DA et al. Allergic contact dermatitis: epidemiology, molecular mechanisms, in vitro methods and regulatory aspects. Current knowledge assembled at an international workshop at BfR, Germany. *Cell Mol Life Sci* 2012;**69**:763–781.
7. Carøe TK, Ebbenhøj N, Agner T. A survey of exposures related to recognized occupational contact dermatitis in Denmark in 2010. *Contact Derm* 2014;**70**:56–62.
8. Angelova-Fischer I. Irritants and Skin Barrier Function. *Curr Probl Dermatol* 2016;**49**:80–89.
9. Eskes C, Détape V, Koëter H, Kreysa J, Liebsch M, Zuang V et al. Regulatory assessment of in vitro skin corrosion and irritation data within the European framework: Workshop recommendations. *Regulatory toxicology and pharmacology : RTP* Published Online First: 2012. doi:10.1016/j.yrtph.2011.10.015
10. Patel S. Danger-Associated Molecular Patterns (DAMPs): the Derivatives and Triggers of Inflammation. *Curr Allergy Asthma Rep* 2018;**18**:63.
11. Clemmensen A, Andersen KE, Clemmensen O, Tan Q, Petersen TK, Kruse TA et al. Genome-wide expression analysis of human in vivo irritated epidermis: differential profiles induced by sodium lauryl sulfate and nonanoic acid. *J Invest Dermatol* 2010;**130**:2201–2210.
12. Fulzele SV, Babu RJ, Ahaghotu E, Singh M. Estimation of proinflammatory biomarkers of

skin irritation by dermal microdialysis following exposure with irritant chemicals. *Toxicology* 2007;**237**:77–88.

13. Calhoun KN, Luckett-Chastain LR, Frempah B, Gallucci RM. Associations Between Immune Phenotype and Inflammation in Murine Models of Irritant Contact Dermatitis. *Toxicol Sci* 2019;**168**:179–189.

14. Uter W, Werfel T, White IR, Johansen JD. Contact Allergy: A Review of Current Problems from a Clinical Perspective. *Int J Environ Res Public Health* 2018;**15**. doi:10.3390/ijerph15061108

15. Vocanson M, Hennino A, Rozières A, Poyet G, Nicolas J-F. Effector and regulatory mechanisms in allergic contact dermatitis. *Allergy* 2009;**64**:1699–1714.

16. Esser PR, Martin SF. Pathomechanisms of Contact Sensitization. *Curr Allergy Asthma Rep* 2017;**17**:83.

17. Nosbaum A, Vocanson M, Rozières A, Hennino A, Nicolas J-F. Allergic and irritant contact dermatitis. *Eur J Dermatol* 2009;**19**:325–332.

18. Leonard A, Guttman-Yassky E. The Unique Molecular Signatures of Contact Dermatitis and Implications for Treatment. *Clin Rev Allergy Immunol* Published Online First: 12 May 2018. doi:10.1007/s12016-018-8685-0

19. Johansen JD, Aalto-Korte K, Agner T, Andersen KE, Bircher A, Bruze M et al. European Society of Contact Dermatitis guideline for diagnostic patch testing - recommendations on best practice. *Contact Derm* 2015;**73**:195–221.

20. Dittmar D, Schuttelaar MLA. Contact sensitization to hydroperoxides of limonene and linalool: Results of consecutive patch testing and clinical relevance. *Contact Derm* 2019;**80**:101–109.

21. Bennike NH, Palangi L, Christensson JB, Nilsson U, Zachariae C, Johansen JD et al. Allergic contact dermatitis caused by hydroperoxides of limonene and dose-response relationship- A repeated open application test (ROAT) study. *Contact Derm* 2019;**80**:208–216.

22. Dittmar D, Ofenloch RF, Schuttelaar MLA. Persistence of contact allergy: a retrospective analysis. *Contact Derm* 2018;**78**:143–150.

23. Frings VG, Böer-Auer A, Breuer K. Histomorphology and Immunophenotype of Eczematous Skin Lesions Revisited-Skin Biopsies Are Not Reliable in Differentiating Allergic Contact Dermatitis, Irritant Contact Dermatitis, and Atopic Dermatitis. *Am J Dermatopathol* 2018;**40**:7–16.

24. Pedersen MB, Skov L, Menné T, Johansen JD, Olsen J. Gene expression time course in the human skin during elicitation of allergic contact dermatitis. *J Invest Dermatol* 2007;**127**:2585–2595.
25. Dhingra N, Shemer A, Correa da Rosa J, Rozenblit M, Fuentes-Duculan J, Gittler JK et al. Molecular profiling of contact dermatitis skin identifies allergen-dependent differences in immune response. *J Allergy Clin Immunol* 2014;**134**:362–372.
26. Inkeles MS, Scumpia PO, Swindell WR, Lopez D, Teles RMB, Graeber TG et al. Comparison of molecular signatures from multiple skin diseases identifies mechanisms of immunopathogenesis. *J Invest Dermatol* 2015;**135**:151–159.
27. Wisgrill L, Werner P, Jalonen E, Berger A, Lauerma A, Alenius H et al. Integrative transcriptome analysis deciphers mechanisms of nickel contact dermatitis. *Allergy* Published Online First: 24 July 2020. doi:10.1111/all.14519
28. Fortino V, Wisgrill L, Werner P, Suomela S, Linder N, Jalonen E et al. Machine-learning-driven biomarker discovery for the discrimination between allergic and irritant contact dermatitis. *Proc Natl Acad Sci U S A* 2020;**117**:33474–33485.
29. Quaranta M, Knapp B, Garzorz N, Mattii M, Pullabhatla V, Pennino D et al. Intraindividual genome expression analysis reveals a specific molecular signature of psoriasis and eczema. *Sci Transl Med* 2014;**6**:244ra90.
30. Breiman L. Random Forests. *Machine Learning* 2001;**45**:5–32.
31. Meller S, Lauerma AI, Kopp FM, Winterberg F, Anthoni M, Müller A et al. Chemokine responses distinguish chemical-induced allergic from irritant skin inflammation: memory T cells make the difference. *J Allergy Clin Immunol* 2007;**119**:1470–1480.
32. Ku H-O, Jeong S-H, Kang H-G, Pyo H-M, Cho J-H, Son S-W et al. Gene expression profiles and pathways in skin inflammation induced by three different sensitizers and an irritant. *Toxicol Lett* 2009;**190**:231–237.
33. Kamsteeg M, Jansen P a. M, van Vlijmen-Willems IMJJ, van Erp PEJ, Rodijk-Olthuis D, van der Valk PG et al. Molecular diagnostics of psoriasis, atopic dermatitis, allergic contact dermatitis and irritant contact dermatitis. *Br J Dermatol* 2010;**162**:568–578.
34. Kehren J, Desvignes C, Krasteva M, Ducluzeau M-T, Assossou O, Horand F et al. Cytotoxicity Is Mandatory for CD8+ T Cell-mediated Contact Hypersensitivity. *J Exp Med* 1999;**189**:779–786.
35. Vocanson M, Hennino A, Cluzel-Tailhardat M, Saint-Mezard P, Benetiere J, Chavagnac C

et al. CD8⁺ T cells are effector cells of contact dermatitis to common skin allergens in mice. *J Invest Dermatol* 2006;**126**:815–820.

36. Yawalkar N, Hunger RE, Buri C, Schmid S, Egli F, Brand CU et al. A comparative study of the expression of cytotoxic proteins in allergic contact dermatitis and psoriasis: spongiotic skin lesions in allergic contact dermatitis are highly infiltrated by T cells expressing perforin and granzyme B. *Am J Pathol* 2001;**158**:803–808.

37. Eyerich K, Böckelmann R, Pommer AJ, Foerster S, Hofmeister H, Huss-Marp J et al. Comparative in situ topoproteome analysis reveals differences in patch test-induced eczema: cytotoxicity-dominated nickel versus pleiotrope pollen reaction. *Exp Dermatol* 2010;**19**:511–517.

38. Gamradt P, Laoubi L, Nosbaum A, Mutez V, Lenief V, Grande S et al. Inhibitory checkpoint receptors control CD8⁺ resident memory T cells to prevent skin allergy. *J Allergy Clin Immunol* 2019;**143**:2147-2157.e9.

39. Luckett-Chastain LR, Gipson JR, Gillaspay AF, Gallucci RM. Transcriptional profiling of irritant contact dermatitis (ICD) in a mouse model identifies specific patterns of gene expression and immune-regulation. *Toxicology* 2018;**410**:1–9.

40. Novak-Bilić G, Vučić M, Japundžić I, Meštrović-Štefekov J, Stanić-Duktaj S, Lugović-Mihić L. IRRITANT AND ALLERGIC CONTACT DERMATITIS - SKIN LESION CHARACTERISTICS. *Acta Clin Croat* 2018;**57**:713–720.

41. Koppes SA, Ljubojevic Hadzavdic S, Jakasa I, Franceschi N, Jurakić Tončić R, Marinović B et al. Stratum corneum profiles of inflammatory mediators in patch test reactions to common contact allergens and sodium lauryl sulfate. *Br J Dermatol* 2017;**176**:1533–1540.

42. Corsini E, Galli CL. Cytokines and irritant contact dermatitis. *Toxicol Lett* 1998;**102–103**:277–282.

43. Russo RC, Garcia CC, Teixeira MM, Amaral FA. The CXCL8/IL-8 chemokine family and its receptors in inflammatory diseases. *Expert Rev Clin Immunol* 2014;**10**:593–619.

44. Bertheloot D, Latz E. HMGB1, IL-1 α , IL-33 and S100 proteins: dual-function alarmins. *Cell Mol Immunol* 2017;**14**:43–64.

45. Swindell WR, Beamer MA, Sarkar MK, Loftus S, Fullmer J, Xing X et al. RNA-Seq Analysis of IL-1B and IL-36 Responses in Epidermal Keratinocytes Identifies a Shared MyD88-Dependent Gene Signature. *Front Immunol* 2018;**9**. doi:10.3389/fimmu.2018.00080

46. Soudja SM, Ruiz AL, Marie JC, Lauvau G. Inflammatory monocytes activate memory CD8(+) T and innate NK lymphocytes independent of cognate antigen during microbial pathogen

invasion. *Immunity* 2012;**37**:549–562.

47. Raué H-P, Beadling C, Haun J, Slifka MK. Cytokine-mediated programmed proliferation of virus-specific CD8+ memory T cells. *Immunity* 2013;**38**:131–139.

48. Sasseville D. Occupational Contact Dermatitis. *Allergy Asthma Clin Immunol* 2008;**4**:59–

65.

CONFLICT OF INTEREST: All authors are public employees from the INSERM, universities, and French or Belgian hospitals.

AN reports other from Pfizer, during the conduct of the study; grants, personal fees and non-financial support from Sanofi Regereron, personal fees and non-financial support from Novartis, grants and personal fees from Pfizer, personal fees and non-financial support from Abbvie, grants and non-financial support from Pierre Fabre, personal fees and non-financial support from Medac, personal fees and non-financial support from Lilly, personal fees and non-financial support from Leo Pharma, grants, personal fees and non-financial support from Incyte, personal fees and non-financial support from Janssen, personal fees from Celgene, outside the submitted work.

SDB, JN, PEJ and LB reports grants from Association SIV, during the conduct of the study.

MAL, AR, VL, AM, AC, MP, FH, MCLB, PP, CD, AH, MB, JFN and MV have nothing to disclose.

ACKNOWLEDGEMENTS

We thank C.F.Gaudefroy, N.Gilmour, M.Elbaz, F.Boislève, G.Ellis and B.Smith for their support in the realization of this work in the frame of the MODAL project, and Marielle Romet and Emma Pilling (Santé Active Edition-Synergy Pharm) who provided medical writing assistance.

This study was supported by different institutional (Institut de la Santé et de la Recherche Medicale (INSERM)) and academic (Société Française d'Allergologie) grants.

AUTHOR CONTRIBUTIONS

MAL, AN, VL, AM and AC carried out the experiments and analysed the data.

MP, SDB, JN, PEJ and LB provide support for bioinformatic/biostatistics analysis.

MAL, AN, AR, FH, MCFL, PP, CD, AH and MB managed the clinical study, performed sample collections, obtained consents and/or participated in the interpretation of the data.

JFN and MV designed the experiments and supervised this study.

JFN and MV wrote the manuscript.

All authors approved the final submitted and published version.

TABLES

Table 1: Patient features and patch test results.

Patient		Patch test			
Patient number	Gender/age	Indication	Class	Compound	Result
1	M/78	Eczema	Irritant	Cantharidin	2+
			Vehicle	Petrolatum	-
2	F/57	Eczema	Irritant	Cantharidin	3+
			Vehicle	Petrolatum	-
3	F/66	Eczema	Irritant	Cantharidin	3+
			Vehicle	Petrolatum	-
4	F/43	Eczema	Irritant	Cantharidin	1+
			Vehicle	Petrolatum	-
5	M/67	Eczema	Irritant	Cantharidin	2+
			Vehicle	Petrolatum	-
6	F/52	Eczema	Allergen	Linalool hydroperoxide	+/?
			Vehicle	Petrolatum	-
7	F/50	Eczema	Allergen	Nickel	2+
			Vehicle	Petrolatum	-
8	M/66	Eczema	Allergen	Methylisothiazolinone	2+
			Nonlesional skin	-	-
9	F/50	Eczema	Irritant	Cantharidin	3+
			Allergen	Nickel	1+
			Vehicle	Petrolatum	-
10	F/61	Eczema	Irritant	Cantharidin	3+
			Vehicle	Petrolatum	-
11	F/40	Eczema	Allergen	Methylisothiazolinone	3+
			Allergen	Nickel	3+
			Vehicle	Petrolatum	-
12	F/68	Eczema	Allergen	Linalool hydroperoxide	2+
			Vehicle	Petrolatum	-
13	F/41	Eczema	Allergen	Nickel	3+
			Vehicle	Petrolatum	-
14	F/41	Eczema	Allergen	Nickel	2+
			Vehicle	Petrolatum	-
15	F/60	Eczema	Allergen	Methylisothiazolinone	3+
			Vehicle	Petrolatum	-
16	F/62	Eczema	Allergen	Nickel	1+
			Vehicle	Petrolatum	-
17	F/53	Eczema	Allergen	Nickel	2+
			Vehicle	Petrolatum	-
18	F/34	Eczema	Allergen	Methylisothiazolinone	2+

			Vehicle	Petrolatum	-
19	M/42	Eczema	Allergen	Linalool hydroperoxide	2+
			Vehicle	Petrolatum	-
20	F/66	Eczema	Allergen	Linalool hydroperoxide	1+
			Nonlesional skin	-	-
21	F/65	Eczema	Allergen	Linalool hydroperoxide	3+
			Vehicle	Petrolatum	-
22	F/67	Eczema	Allergen	Methylisothiazolinone	2+
			Vehicle	Petrolatum	-
23	F/57	Eczema	Allergen	Linalool hydroperoxide	2+
			Vehicle	Petrolatum	-
24	F/76	Eczema	Allergen	Methylisothiazolinone	+/?
			Vehicle	Petrolatum	-
25	F/35	Eczema	Irritant	SLS	3+
			Vehicle	Petrolatum	-
26	F/50	Eczema	Allergen	Nickel	1+
			Vehicle	Petrolatum	-
27	F/48	Eczema	Irritant	Nonanoic acid	1+
			Vehicle	Isopropanol	-
28	F/29	Eczema	Allergen	Linalool hydroperoxide	3+
			Irritant	Nonanoic acid	1+
			Nonlesional skin	-	-
29	M/44	Eczema	Irritant	Nonanoic acid	1+
			Vehicle	Isopropanol	-
30	M/62	Eczema	Irritant	Nonanoic acid	1+
			Vehicle	Isopropanol	-
31	F/29	Eczema	Allergen	Linalool hydroperoxide	1+
			Irritant	SLS	-
			Vehicle	Petrolatum	-
32	F/57	Eczema	Allergen	Nickel	1+
			Irritant	SLS	-
			Vehicle	Petrolatum	-
33	F/44	Eczema	Allergen	Nickel	1+
			Irritant	SLS	-
			Vehicle	Petrolatum	-
34	F/71	Eczema	Allergen	Linalool hydroperoxide	1+
			Irritant	SLS	-
			Vehicle	Petrolatum	-
35	M/54	Eczema	Allergen	Linalool hydroperoxide	2+
			Irritant	SLS	-
			Vehicle	Petrolatum	-
36	M/86	Eczema	Irritant	SLS	1+
			Vehicle	Petrolatum	-
37	M/88	cADRs	Allergen	Amoxicillin	1+
			Irritant	SLS	1+
			Nonlesional skin	-	-

38	M/75	cADRs	Irritant Nonlesional skin	SLS -	1+ -
39	F/78	cADRs	Allergen Vehicle	Acetazolamide Petrolatum	3+ -
40	M/43	cADRs	Allergen Vehicle	Amoxicillin Petrolatum	2+ -
41	F/66	cADRs	Allergen Vehicle	Amoxicillin Petrolatum	1+ -
42	F/68	cADRs	Irritant Vehicle	SLS Petrolatum	1+ -
43	F/73	cADRs	Allergen Nonlesional skin	Amoxicillin -	2+ -
44	M/64	cADRs	Allergen Vehicle	Amoxicillin Petrolatum	1+ -
45	F/68	cADRs	Allergen Nonlesional skin	Piperacillin/Tazobactam -	2+ -
46	M/73	cADRs	Allergen Nonlesional skin	Amoxicillin -	1+ -
47	F/68	cADRs	Irritant Vehicle	SLS Petrolatum	1+ -

Patch test reactions were scored at 72 hours in accordance with International Contact Dermatitis Research Group (ICDRG) criteria and European Society of Contact Dermatitis (ESCD) guidance.

-: negative reaction

+/? : doubtful reaction (erythema only, no infiltration)

1+: mild positive reaction (redness, induration, and possibly papules)

2+: strong positive reaction (erythema, induration, papules, and vesicles)

3+: extreme positive reaction (intense erythema, induration, and coalescing vesicles)

F=Female

M=Male

cADRs: cutaneous Adverse Drug Reactions

Table 2: Performance of the 30 best classifier models identified by Random Forest algorithm.

Gene sets	Accuracy	Precision- Allergen	Precision- Irritant	Recall- Allergen	Recall- Irritant	F1-Score- Allergen	F1-Score- Irritant	AUC
IGFL3, HLA-DRA	1.00	1.00	1.00	1.00	0.99	1.00	1.00	1.00
IGFL3, HLA-DRA, HLA-DMA	0.99	0.98	1.00	1.00	0.96	0.99	0.98	1.00

IGFL3, PLEK	0.98	0.98	1.00	1.00	0.95	0.99	0.97	1.00
IGFL3, HLA-DRA, HLA-DPA1	0.98	0.98	1.00	1.00	0.95	0.99	0.97	1.00
IGFL3, HLA-DPA1	0.98	0.99	0.98	0.99	0.98	0.99	0.97	1.00
IGFL3, HLA-DRA, CD74	0.98	0.98	1.00	1.00	0.95	0.99	0.97	1.00
IGFL3, PLEK, HLA-DRA, MCOLN2	0.98	0.98	1.00	1.00	0.95	0.99	0.97	1.00
IGFL3, PLEK, HLA-DRA, HLA-DMA	0.98	0.98	1.00	1.00	0.95	0.99	0.97	1.00
IGFL3, PLEK, HLA-DPA1	0.98	0.98	1.00	1.00	0.94	0.99	0.97	1.00
IGFL3, HLA-DRA, MCOLN2	0.98	1.00	0.96	0.97	1.00	0.99	0.97	1.00
IGFL3, HLA-DRA, RAB8B	0.98	0.99	0.96	0.98	0.99	0.99	0.97	1.00
HLA-DRA, CD69	0.98	0.97	1.00	1.00	0.94	0.99	0.96	0.99
IGFL3, HLA-DRA, CD69	0.98	0.97	1.00	1.00	0.94	0.99	0.97	1.00
GPR183, IGFL3, HLA-DRA	0.98	0.97	1.00	1.00	0.94	0.99	0.97	1.00
GPR183, IGFL3, CD69	0.98	0.97	1.00	1.00	0.94	0.99	0.96	1.00
IL15, CD69	0.98	0.97	1.00	1.00	0.94	0.99	0.96	0.98
GPR183, IGFL3	0.98	0.97	1.00	1.00	0.94	0.99	0.97	1.00
IGFL3, PLEK, HLA-DRA, RAB8B	0.98	0.98	1.00	1.00	0.95	0.99	0.97	1.00
IGFL3, IL15, CD69	0.98	0.97	1.00	1.00	0.94	0.99	0.96	1.00
IGFL3, PLEK, CD69	0.98	0.97	1.00	1.00	0.94	0.99	0.96	1.00
IGFL3, IL15, HLA-DRA	0.98	0.97	1.00	1.00	0.94	0.99	0.96	1.00
GPR183, IL15, HLA-DRA	0.98	0.97	1.00	1.00	0.94	0.99	0.96	1.00
IGFL3, PLEK, IL15	0.98	0.97	1.00	1.00	0.94	0.99	0.96	1.00
GPR65	0.98	0.97	1.00	1.00	0.94	0.99	0.96	0.99
GPR183, IGFL3, PLEK	0.98	0.97	1.00	1.00	0.94	0.99	0.96	1.00
IL15, HLA-DRA	0.98	0.97	1.00	1.00	0.94	0.99	0.96	1.00
PLEK, IL15	0.98	0.97	1.00	1.00	0.94	0.98	0.96	0.99
PLEK, CD69	0.97	0.97	0.98	0.99	0.94	0.98	0.96	0.99
IGFL3, IL15	0.97	0.97	0.98	0.99	0.94	0.98	0.95	1.00
PLEK, HLA-DRA	0.96	0.97	0.96	0.97	0.94	0.97	0.94	1.00
GPR183, CD69	0.96	0.95	1.00	1.00	0.88	0.97	0.93	0.98

Performance metrics:

- Accuracy is the rate of correct predictions among the total number of predictions the classifier performed with both allergen and irritant samples.
- Precision is the rate of correct predictions among the total number of positive predictions (considering both true positive and false positive predictions) for a given class sample (allergen or irritant).

- Recall is the rate of correct predictions among the total number of predictions (considering both true positive and false negative predictions) the classifier made with samples from a specific class (allergen or irritant).
- F1-Score for a given class is the harmonic mean of the Precision and Recall values.
- AUC is the area under the ROC curve calculated for a standard binary classifier. The ROC curve compares the sensitivity (here Recall Allergen) and the specificity (here Recall Irritant) of each binary classifier, when the decision threshold varies. The decision threshold is the value at which the binary classifier considers that a classification belongs to a specific class (allergen or irritant). An AUC value near to the 1.00 means an excellent discrimination of the two classes.

Table 3: Validation of the best classifiers using external gene array datasets.

Gene sets	Wisgrill et al. (27)			Ljungberg Silic et al. (in preparation)
	Recall Allergen (Nickel, n=11)	Recall Irritant (SLS, n=6)	AUC	Recall Allergen (Amerchol L-101, n=8)
GPR65	0.91	0.67	0.83	1
GPR183 IGFL3	0.91	0.67	0.94	0.88
PLEK IL15	0.91	0.67	0.85	1
PLEK CD69	0.91	0.67	0.86	1
IL15 HL-DRA	0.91	0.67	0.79	1
GPR183 IGFL3 PLEK	0.91	0.67	0.92	0.88
GPR183 IGFL3 CD69	0.91	0.67	0.92	1
IGFL3 PLEK CD69	0.91	0.67	0.88	1
IGFL3 IL15 CD69	0.91	0.67	0.86	1
IGFL3 PLEK HLA-DPA1	0.91	0.67	0.88	1

The classifiers shown in Table 2 were validated on two external gene array datasets, generated from biopsies of positive PT samples (i) to nickel and SLS (access code E-MTAB-8945) and (ii) to amerchol (access code GSE169573).

Different metrics, Recall and AUC of ROC curves, are shown for the 10 best performance classifiers.

FIGURE LEGENDS

Figure 1: Intra-individual comparison of gene expression in allergen and irritant samples.

Principal component analysis performed on the full set of genes measured on the arrays. To compensate for donor-induced variability, the gene expression level of the control sample from each donor was subtracted from that of the corresponding treated sample from the same donor. Results depict the sample location (symbols) in the first two principal components (PCs), with the controls a coordinate 0. Together, the first two components represented approximately 65% of the variance. Patient numbers are reported above each sample.

Figure 2: Common and distinct molecular responses induced by allergens and irritants

Each point of the graph illustrates the variation in the average expression levels of each gene in the allergen and irritant groups, after correction for the donor effect (mean fold changes).

Green dots indicate genes (n=457) significantly modulated in both groups.

Blue dots indicate genes (n=1956) predominantly modulated in the allergen group.

Yellow dots indicate genes (n=255) predominantly modulated in the irritant group.

Gray dots indicate non-significant differences (adjusted P value < 0.05).

The black square delimits fold changes >1.41 or <1/1.41.

Figure 3: Signaling pathway analysis of allergen and irritant patch test responses.

Shown are the most significant hits for the “biological process” GO terms for genes commonly (A) or predominantly regulated in the irritant (B) and allergen (C) groups.

The size of the bar indicates the level of the estimated frequency of activation of the respective pathway (from 0 to 1).

Figure 4: Unique molecular signatures for allergens and irritants.

A set of 67 DEGs was used to illustrate the core transcriptomes of the 52 positive allergen and irritant-induced reactions. Results are depicted as a heatmap, in which samples are grouped by allergens and irritants and stratified by clinical reaction scores (+/? to 3+) (top of the figure).

Blue, white and red colors represent low, moderate and high signal intensity, respectively.

Chemical compounds and patient numbers are reported at the bottom of the figure.

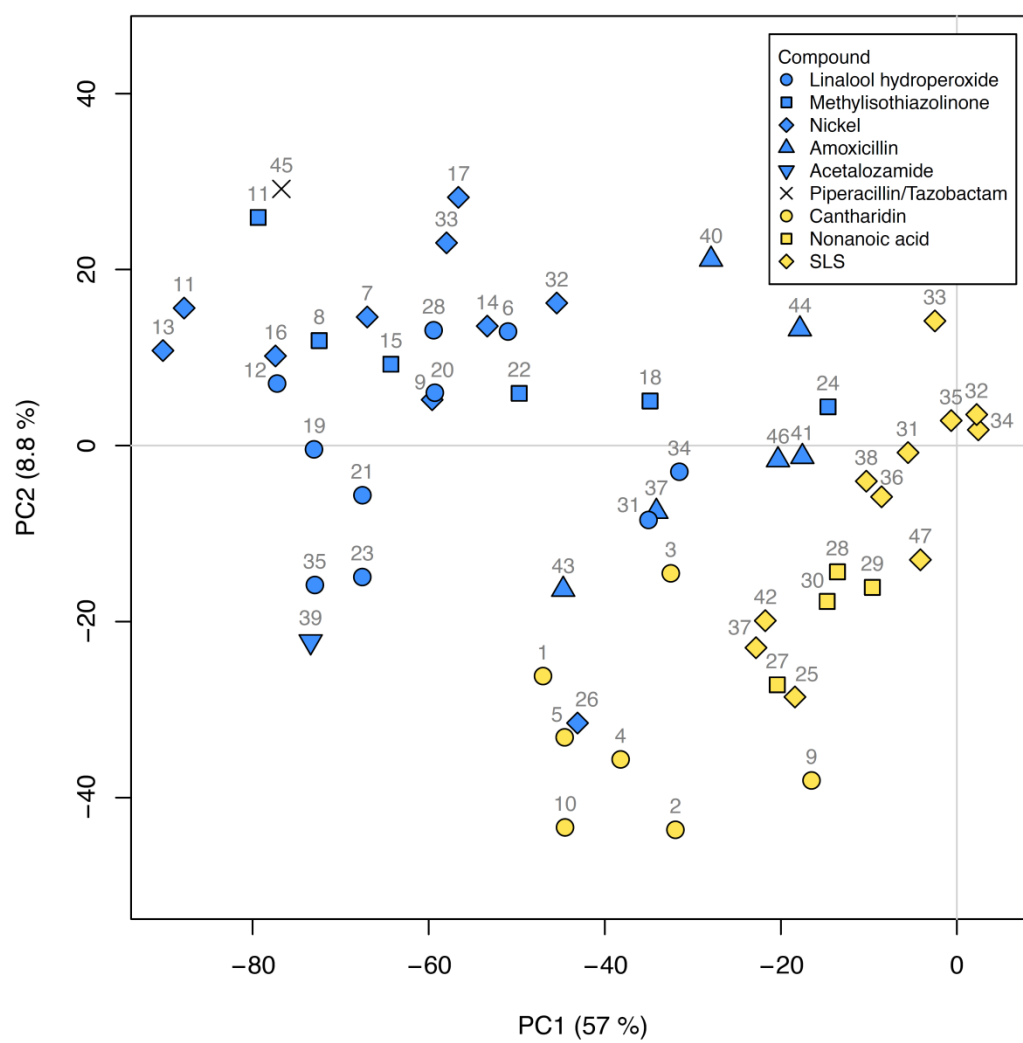
Figure 5: Gene expression correlates with the magnitude of the clinical response.

Box-and-whisker plots depict the distribution of the average levels of variation in core gene expression for the allergen (A) and irritant (B) groups depending on the clinical score, i.e. the intensity of the PT reactions. Each point represents the average expression level ($\log_2 F_c$) of core genes (including 52 genes for the allergen signature and 15 genes for the irritant signature) for one sample, after correction for the donor effect. Patient numbers are reported next to each sample.

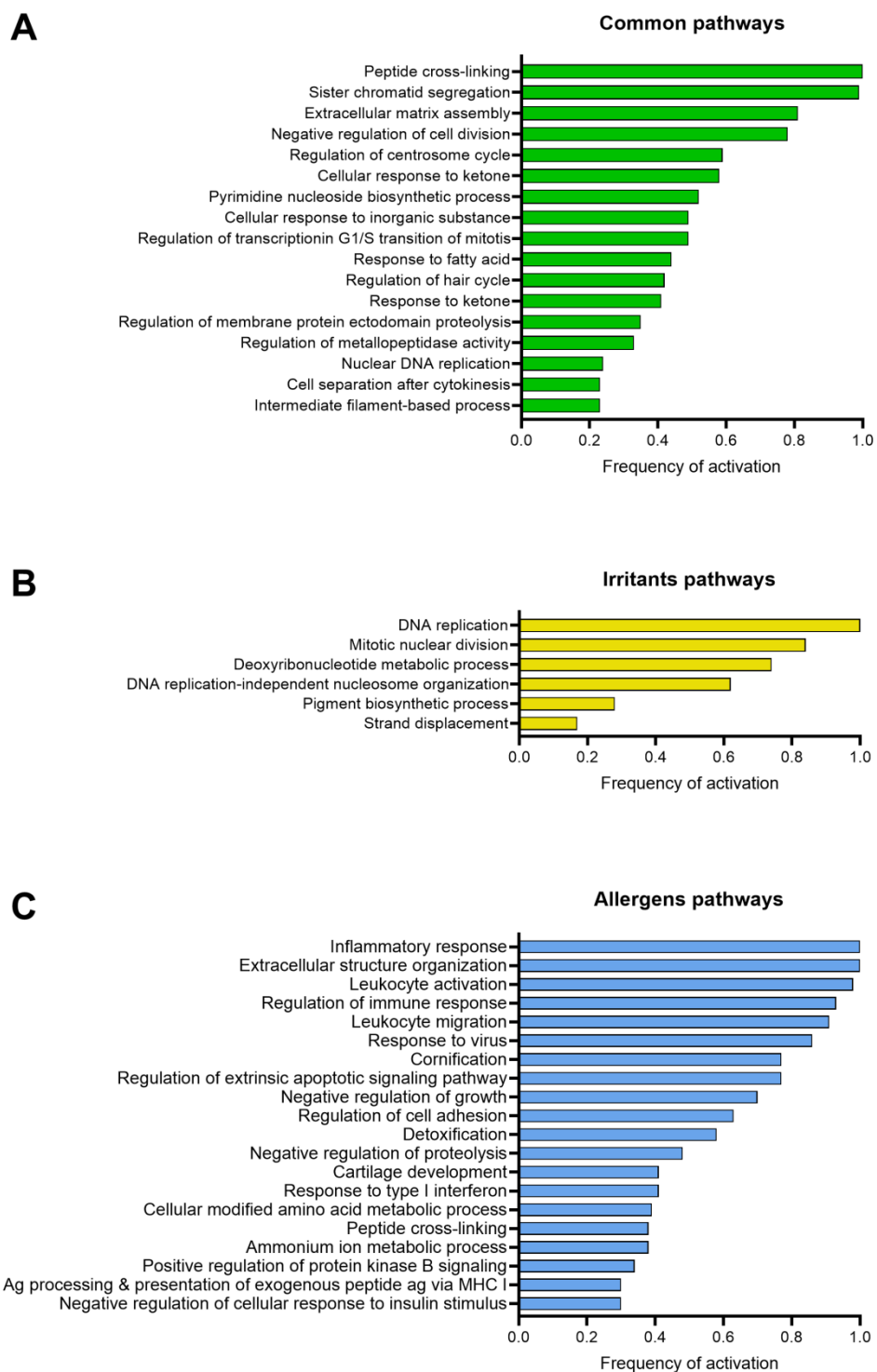
Figure 6: Validation of a small set of segregating biomarkers by qRT-PCR.

Gene expression of HLA-DRA, GPR183, IGFL3, PLEK, IL15, CD69, CD8A, CD4, GZMB, GNLY, IFNG and CXCL10 mRNA transcripts in allergen and irritant samples was analyzed by qRT-PCR. Results are expressed as fold changes versus control skin. Patient numbers are reported above each sample.

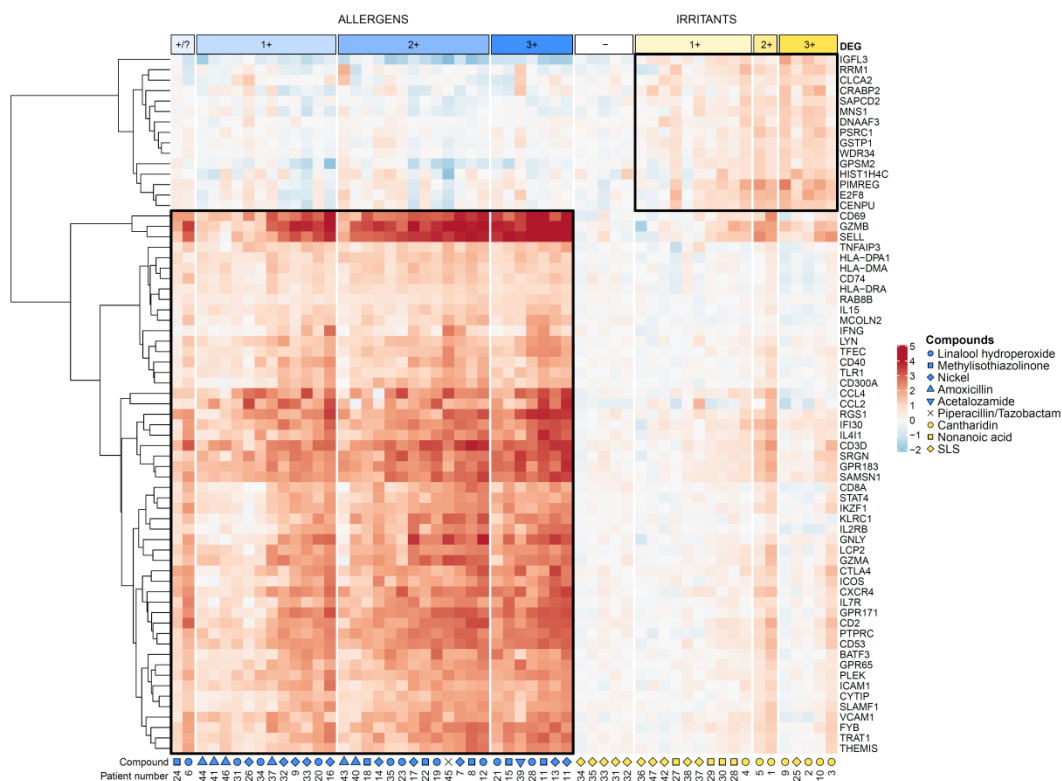
* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$; Mann-Whitney U test.



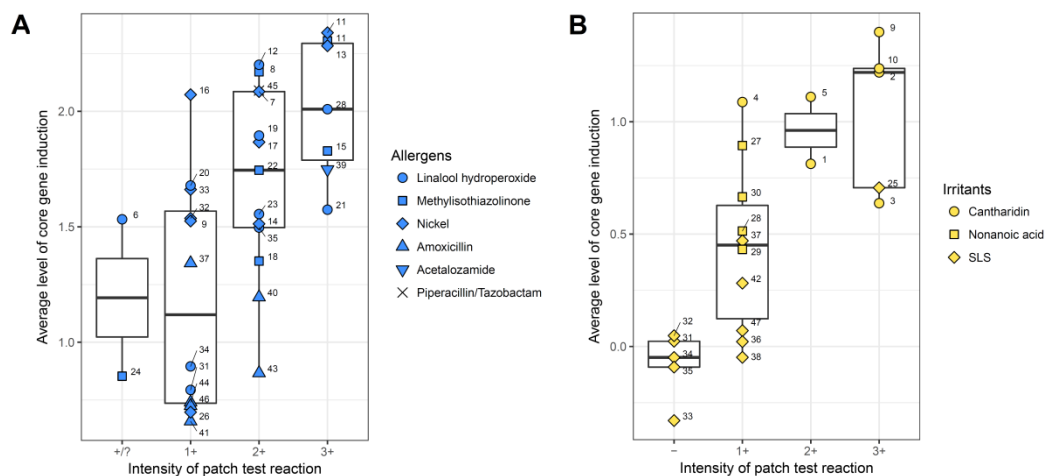
all_14989_f1.tif



all_14989_f3.tif



all_14989_f4.tif



all_14989_f5.tif

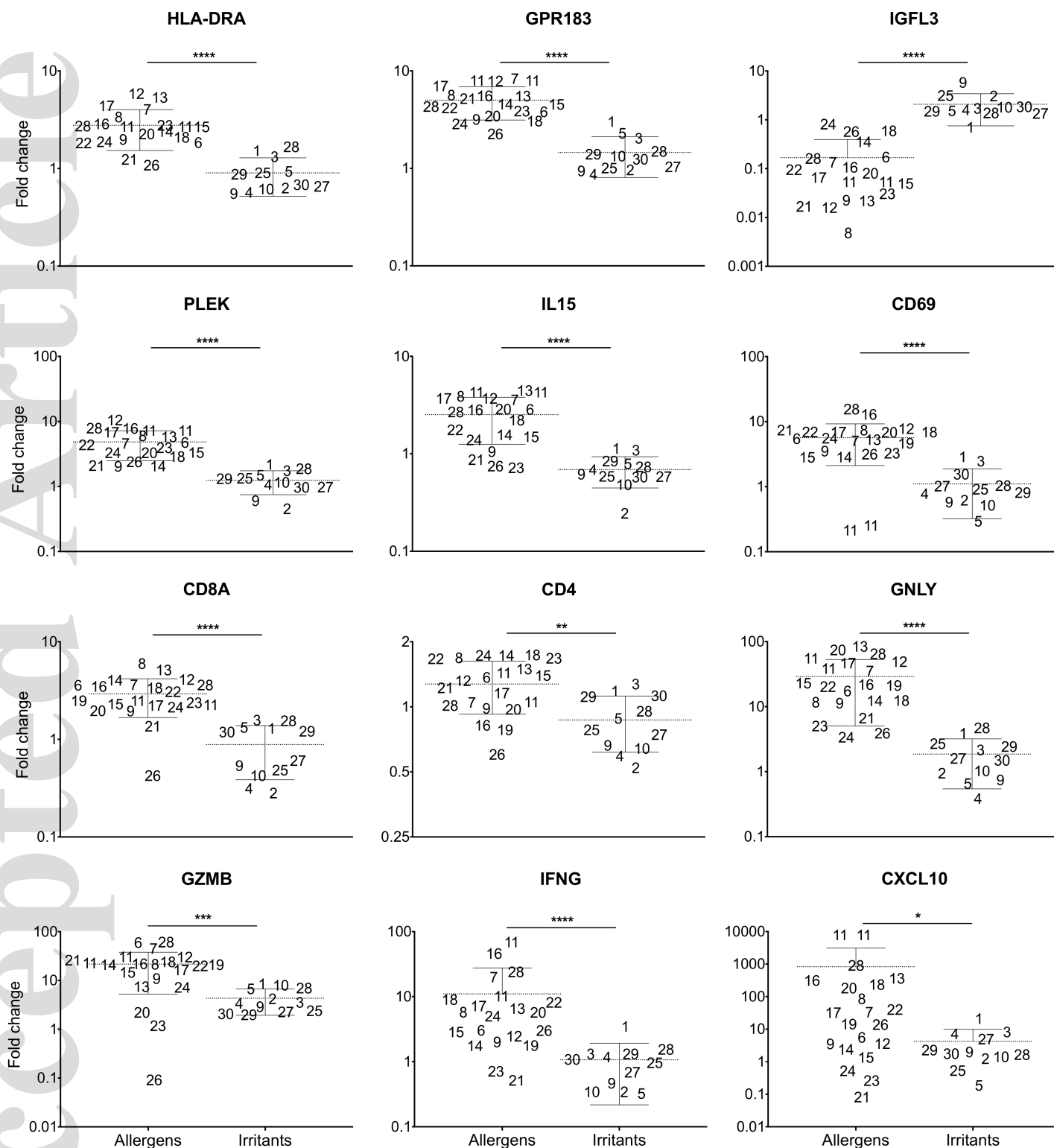


Figure 6

all_14989_f6.tif

SUPPLEMENTARY MATERIAL

Supplemental Materials and Methods

Chemicals

Chemical allergens used for patch testing were: nickel sulfate hexahydrate (5% in petrolatum), methylisothiazolinone (0.2% in distilled water) and/or linalool hydroperoxide (1% in petrolatum), all purchased from Chemotechnique Diagnostics (Grasse, France); as well as amoxicillin (2% in NaCl; Panpharma, Luitré, France), piperacillin/tazobactam (4%/0.5% in distilled water; Panpharma, Luitré, France) and/or acetazolamide (10% in distilled water; Cooper, Melun, France). Irritant reactions were induced using patches with sodium lauryl sulfate (0.5% in water; Sigma, St Quentin Fallavier, France), nonanoic acid (90% in isopropanol; Sigma, St Quentin Fallavier, France) and/or cantharidin (0.7% in acetone, collodion, castor oil and camphor; Dormer laboratories, Totonto, Canada). Vehicle (petrolatum and isopropanol) patches were prepared as controls.

Patch test procedure

Patch testing was performed using IQ Ultra chambers (Chemotechnique Diagnostics, Grasse, France). Skin reactions were scored at 72 hours in accordance with International Contact Dermatitis Research Group (ICDRG) criteria and European Society of Contact Dermatitis (ESCD) guidance (1): - negative reaction; +/- doubtful reaction (erythema only, no infiltration); 1+ weak positive reaction (redness, induration, and possibly papules); 2+ positive reaction (erythema, induration, papules, and vesicles); 3+ strong positive reaction (intense erythema, induration, and coalescing vesicles).

Gene arrays

Affymetrix PrimeView Human Gene Expression arrays (Affymetrix, Inc., Santa Clara, CA, USA) were used for the gene array study. They allowed the analysis of more than 36000 transcripts, including approximately 22 000 genes and unigenes.

Sample preparation for the array hybridization was carried out as described in the Affymetrix GeneChip 3' IVT Express Kit user manual. In brief, total RNA was extracted from skin biopsies using the Qiagen RNeasy Mini Kit (Courtaboeuf, France) and contaminating DNA was removed using the DNA-FreeTM DNA Removal kit (Invitrogen, Cergy-Pontoise, France). The RNA was then reverse-transcribed into double-stranded copy DNA (cDNA), and then underwent *in vitro* transcription to generate biotin-labeled amplified RNA (aRNA). Purified aRNA was hybridized to the Affymetrix GeneChip Human PrimeView arrays for 16 h at 45° C in a GeneChip hybridization oven. Hybridized arrays were washed and stained in an Affymetrix Fluidics Station FS450, and the fluorescent signals were measured with an Affymetrix GeneChip Scanner 3000 7G. Fluidics and scan functions were controlled by the Affymetrix GeneChip Command Console v4.1.3 software.

Bioinformatics and statistical analysis of gene array data

Quality control of the gene arrays was carried out using standard QC metrics and R package quality control. Data were preprocessed (background correction and normalization) by applying (i) function *backgroundCorrect* from the limma package on the perfect match (PM) signals, (ii) the variance stabilizing transformation algorithm (function *justvs* from the vsn package (2)), and (iii) a quantile normalization (function *normalize* from the affy package with parameter *method* = "quantiles").

To filter probe sets, they were marked as informative or noninformative based on the result of the I/NI method (3). Noninformative probe sets were excluded from unsupervised analyses.

All sets were included when the statistical models were applied, but only the informative sets were considered when correcting for multiple testing.

Single values for the probe signals for each probe set of each sample were obtained using the summarization step of the Robust Multichip Average (RMA (4)) approach (function `rmaPLM` from the `affyPLM` package, with parameters `background=FALSE` and `normalize=FALSE`). `rmaPLM` also returned standard error estimates for each estimated expression level: these were used to incorporate weight inputs into the linear regression models.

Expression values were modeled using different mixed linear regression models, with the group of chemicals or the type of molecules as fixed factor and a random effect for each patient. Mixed linear regression models were applied using function `lme` from the `nlme` package (5). Comparisons of interest were then analyzed through statistical contrasts, using the `emmeans` package (6). P values from paired t-tests were adjusted for multiple hypotheses using the Benjamini and Hochberg false discovery rate (FDR) controlling procedure (function `p.adjust` from package `stats` with parameter `method="BH"`).

Multiple analyses were performed on the gene array data to determine whether global gene expression levels differed between the different groups (allergens, irritants, controls) of samples. Comparisons were also performed to identify genes or pathways that were specifically and differentially expressed between all the samples from the allergen, irritant and/or control groups, or between the samples from each respective molecule (when the number of samples for a specific molecule was < 2 , they were excluded from the analysis).

In **Fig.1**, the PCA of differential expression was performed with probe set fold changes (F_c) neither centered nor scaled, using function `prcomp` from package `stats`.

In **Fig.2** and **Tab.S1&2**, to determine the effect of each group of molecules (allergens or irritants), the statistical model used was $signal \sim -1 + Condition$ where *Condition* represents the group of molecules. To account for the paired design, an additional random effect was

included ($\sim 1 \mid \text{Donor}$). The model estimated the average gene expression level in each condition of the experimental design after correction for the donor effect. The statistical contrasts applied allowed analysis of the following comparisons of interest: (Allergen - Control), (Irritant - Control), and (Irritant - Allergen). Genes were then considered as follows:

- positively correlated ($\sim$ on the $Y = X$ line) if they were significant (adjusted $P < 0.05$) and displayed $F_c > 1.41$ or $< 1/1.41$ for both effects.
- anti-correlated ($\sim$ on the $Y = -X$ line) if they were significant (adjusted $P < 0.05$) for both effects and displayed a $F_c > 1.41$ for one effect and $< 1/1.41$ for the other effect (or vice-versa).
- specific to the X-axis ($\log_2(Y) \simeq 0$) if they had a $F_c > 1.41$ or $< 1/1.41$ for the effect represented on the X axis and were significant (adjusted $P < 0.05$) for this effect but did not fulfill one of these criteria for the other effect.
- specific to the Y-axis ($\log_2(X) \simeq 0$) if they had a $F_c > 1.41$ or $< 1/1.41$ for the effect represented on the Y axis and were significant (adjusted $P < 0.05$) for this effect but did not fulfill one of these criteria for the other effect.

In **Fig.3** and **Tab.S3**, groups of genes identified by previous analysis were included in an input list for the MGSA algorithm (using the `mgsa` function from the `mgsa` package (7)). The model assumes that some gene sets are active and that all gene members of active sets are themselves active. Fitting the model amounts to inferring the probability of every set to be active given the input list (i.e., scores reported for each gene set are not P values and the closer to 1, the better). The approach was applied to Gene Ontology terms (from the Biological Process tree). Only sets with at least 10 genes and a maximum of 1000 genes were considered.

In **Fig.4**, a heatmap for gene expression was used to illustrate how the expression of a set of differentially expressed genes involved in the allergen response differed from that involved in

the irritant response. The most notable criteria for gene selection were a $F_c > 2$ and a $FDR < 10^{-6}$. Gene clustering was performed using the complete agglomerative method, passing the Euclidean distance between uncentered and scaled probe set signals to function `hclust` from package `stats`. To compensate for donor-induced variability, the gene expression level of the control sample of each donor was subtracted from that of the corresponding treated sample(s) of the same donor.

In **Fig.5**, to determine the correlation between gene expression levels and the clinical response, the average expression level ($\log_2 F_c$) of selected genes pertaining to the allergen or irritant core signatures were calculated for each sample after correction for the donor effect.

In **Fig.S1**, linear regression model was used to study the relationship between the expression of individual gene associated to the allergen and irritant core signatures and the intensity of the patch-test reactions.

qRT-PCR

Total RNA samples were also used for quantitative RT-PCR analysis. qRT-PCR investigations were performed using a StepOnePlus PCR system (ThermoFisher, Villebon-sur-Yvette, France), and primers from Qiagen (Courtaboeuf, France) for CXCL10, CXCL11, CCL2, CCL20, CD69 and GNLY: QuantiTect Primer Assays; or produced by Sigma (Saint-Quentin-Fallavier, France) for CD8a, CD4, IFNG, GZMB, Actin, GAPDH – designed using the publicly Universal Probe Library Assay Design Center (https://lifescience.roche.com/en_fr/brands/universal-probe-library.html#assaydesign-center). All the primers used are listed in **Tab.S4**.

Analysis of qRT-PCR data

Expression values for each mRNA transcript were normalized to values obtained for housekeeping genes (actin and GAPDH), and results from patch-tested samples were presented as the Fc in the test samples versus that in control skin.

Statistical analysis of qRT-PCR data

In **Fig.6**, the mRNA expression data were analyzed using the Kolmogorov–Smirnov test to determine if they followed a Gaussian distribution. If yes, statistical significance was determined using the two-tailed Student's t test. When the values did not assume a Gaussian distribution, the Mann–Whitney test was used. $P < 0.05$ was considered as statistically significant: * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$.

Biomarker discovery

A supervised machine-learning-based approach was used to identify biomarkers, whose gene expression profile can correctly predict/classify whether a given sample is an « allergen » or an « irritant » sample.

First, the Boruta algorithm (8) and the Random Forest classification algorithm (9) (using the permutation based importance measure of Random Forest) were successively run to select the genes, or the combination of genes, susceptible to provide the best classifications. At final, 15 biomarkers were selected among 694 relevant probe set genes, which represented 4943 different biomarker combinations including 1 to 5 genes.

We next employed a standard Random Forest-based classifier approach to assess the classification performances of each biomarker combination (referred to as “classifier”). To this end, classifiers were tested by repeated (50 times) 5-fold cross-validations in order to simulate different training sets and next evaluated these classifiers on unseen data. The classification performances were evaluated by creating a confusion matrix from the different

classification tests (a confusion matrix summarizes the counts of true positive, true negative, false positive and false negative classifications in predicted versus actual classes (here “allergens” and irritant”)). Several classification metrics were derived from the confusion matrix, such as the overall accuracy, as well as four class metrics: precision, recall, F1-score, and the area under curve (AUC) value of the ROC (Receiver Operator Characteristic) curve. The AUC of the ROC curve was calculated to account for variations in the decision threshold, that is the value at which the classifier considered that a prediction belongs to a specific class. **Tab.2** reports the list of the 30 best classifiers (selected on overall accuracy) calculated from the 50 runs of 5-fold cross-validation.

The performance of the trained classifiers were also tested on external datasets, derived from one previously published omics study, including 11 allergen and 6 irritant independent samples (data available from EBI ArrayExpress under accession no. E-MTAB-8945), as well as on unpublished study, including 8 allergen independent samples (data available from Gene Expression Omnibus under accession no. GSE169573). Results are reported in **Tab.3** and **Fig.S2**.

In **Fig.S2**, the ROC curve was created by plotting the sensitivity (here, Recall Allergen) against the specificity (here Recall Irritant) of HLA-DRA / IGFL3 classifier, at various threshold settings.

Data Availability

Expression microarray data have been deposited in Gene Expression Omnibus (<https://www.ncbi.nlm.nih.gov/geo>) under accession no. GSE168735.

Supplemental references

1. Johansen JD, Aalto-Korte K, Agner T, Andersen KE, Bircher A, Bruze M et al. European Society of Contact Dermatitis guideline for diagnostic patch testing - recommendations on best practice. *Contact Derm* 2015;**73**:195–221.
2. Huber W, von Heydebreck A, Sültmann H, Poustka A, Vingron M. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* 2002;**18 Suppl 1**:S96-104.
3. Talloen W, Clevert D-A, Hochreiter S, Amaratunga D, Bijmens L, Kass S et al. I/NI-calls for the exclusion of non-informative genes: a highly effective filtering tool for microarray data. *Bioinformatics* 2007;**23**:2897–2902.
4. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 2003;**4**:249–264.
5. Pinheiro JC, Bates DJ, DebRoy S, Sakar D. *The Nlme Package: Linear and Nonlinear Mixed Effects Models, R Version 3*. 2012
6. Lenth RV. Least-Squares Means: The R Package **lsmeans**. *J Stat Soft* 2016;**69**. doi:10.18637/jss.v069.i01
7. Bauer S, Gagneur J, Robinson PN. GOing Bayesian: model-based gene set analysis of genome-scale data. *Nucleic Acids Res* 2010;**38**:3523–3532.
8. Kursu MB, Rudnicki WR. Feature Selection with the Boruta Package. *Journal of Statistical Software* 2010;**36**:1–13.
9. Breiman L. Random Forests. *Machine Learning* 2001;**45**:5–32.

Supplemental tables

Table S1: Differentially expressed genes in the allergen group.

	Fc	Adj- p	Symbol	Description
1	x 16.1	5.04e-14	SERPINB4	serpin family B member 4
2	x 13.6	8.38e-20	MMP12	matrix metalloproteinase 12
3	x 9.56	1.18e-12	MMP1	matrix metalloproteinase 1
4	x 8.97	2.28e-19	GZMB	granzyme B
5	x 8.09	3.51e-10	CXCL10	C-X-C motif chemokine ligand 10
6	x 8.08	8.28e-22	SELE	selectin E
7	/ 8.04	1.92e-14	LCE5A	late cornified envelope 5A
8	x 7.9	2.6e-19	SELL	selectin L
9	x 7.81	3.12e-13	KRT6C	keratin 6C
10	x 7.8	1.09e-12	S100A7A	S100 calcium-binding protein A7A
11	x 7.6	2.4e-11	CXCL9	C-X-C motif chemokine ligand 9
12	x 6.9	5.43e-20	CCL19	C-C motif chemokine ligand 19
13	x 6.74	5.89e-12	S100A9	S100 calcium binding protein A9
14	x 6.69	5.3e-21	CD69	CD69 molecule
15	x 6.38	6.81e-20	LTB	lymphotoxin beta
16	x 5.94	2.6e-15	CCL18	C-C motif chemokine ligand 18
17	x 5.64	4.92e-22	CD3D	CD3d molecule
18	x 5.62	1.61e-11	KRT6A	keratin 6A
19	x 5.6	5.91e-15	LTF	lactotransferrin
20	x 5.57	1.24e-17	PLAC8	placenta specific 8
21	/ 5.14	1.47e-13	IL37	Interleukin-37
22	x 4.94	1.68e-08	PI3	peptidase inhibitor 3
23	x 4.93	6.11e-20	CD3E	CD3e molecule
24	x 4.9	5.91e-10	CXCL11	C-X-C motif chemokine ligand 11
25	x 4.88	1.65e-22	SRGN	serglycin
26	x 4.84	4.96e-17	PLA2G2A	phospholipase A2 group IIA
27	x 4.79	3.11e-24	SAMSN1	SAM domain, SH3 domain and nuclear localization
28	x 4.78	8.03e-25	GPR	1880 G protein-coupled receptor 183
29	x 4.68	5.47e-21	CD48	CD48 molecule
30	x 4.68	4.31e-09	SAA2	serum amyloid A2
31	x 4.67	7.64e-21	RGS1	regulator of G-protein signaling 1
32	x 4.66	5.11e-17	FCGR3A	Fc fragment of IgG receptor IIIa
33	/ 4.66	2.4e-14	KRT77	keratin 77
34	x 4.64	5.32e-18	UBD	ubiquitin D
35	x 4.57	5.22e-16	EPSTI1	epithelial stromal interaction 1
36	/ 4.57	5.83e-16	C5orf46	chromosome 5 open reading frame 46
37	x 4.49	2.54e-15	FAM26F	family with sequence similarity 26 member F
38	/ 4.4	9 1.52e-9	LOR	loricrin
39	x 4.44	1.05e-11	LCE3D	late cornified envelope 3D
40	/ 4.41	8.5e-16	CHP2	calcineurin-like EF-hand protein 2
41	x 4.26	1.46e-18	CCR7	C-C motif chemokine receptor 7
42	/ 4.23	2e-10	FLG2	filaggrin family member 2
43	x 4.23	1.86e-10	S100A8	S100 calcium binding protein A8
44	x 4.23	1.8e-19	GPR171	G protein-coupled receptor 171
45	x 4.22	4.92e-22	IFI30	IFI30, lysosomal thiol reductase
46	x 4.22	5.6e-18	TYMP	thymidine phosphorylase
47	x 4.19	5.43e-11	S100A7A	S100 calcium binding protein A7A
48	x 4.19	1.59e-22	CD53	CD53 molecule
49	x 4.19	3.86e-23	COTL1	coactosin-like F-actin binding protein 1
50	x 4.17	9.28e-21	ADAM19	ADAM metalloproteinase domain 19
51	x 4.16	1.19e-13	GBP5	guanylate binding protein 5
52	x 4.13	2.92e-20	CXCR4	C-X-C motif chemokine receptor 4
53	x 4.12	4.33e-13	AKR1B10P1	aldo-keto reductase family 1 member B10 pseudogene 1
54	x 4.11	2.1e-13	GBP1	guanylate binding protein 1

55	x 4.09	4.92e-22	PTPRC	protein tyrosine phosphatase, receptor type C
56	x 4.08	9.81e-22	CD52	CD52 molecule
57	x 4.07	7.3e-20	CD2	CD2 molecule
58	/ 4	3.69e-11	FLG	filaggrin
59	/ 3.98	2.54e-10	LCE2B	late cornified envelope 2B
60	x 3.97	6.89e-18	CCL4	C-C motif chemokine ligand 4
61	x 3.87	5.71e-07	SPRR2A	small proline-rich protein 2A
62	x 3.86	2e-13	ISG20	interferon stimulated exonuclease gene 20
63	x 3.84	1.36e-21	C1orf162	chromosome 1 open reading frame 162
64	x 3.81	2.59e-20	JAK3	Janus kinase 3
65	x 3.8	1.48e-13	TNC	tenascin C
66	x 3.78	4.85e-18	CLEC4A	C-type lectin domain family 4 member A
67	x 3.76	1.42e-18	SERPINA1	serpin family A member 1
68	x 3.76	1.38e-18	GZMA	granzyme A
69	/ 3.74	4.48e-10	C1orf68	chromosome 1 open reading frame 68
70	x 3.74	3.33e-10	SERPINB3	serpin family B member 3
71	x 3.72	9.59e-14	HAS3	hyaluronan synthase 3
72	x 3.72	1.73e-16	FPR1	formyl peptide receptor 1
73	x 3.67	4.38e-19	TRAF3IP3	TRAF3 interacting protein 3
74	x 3.66	2.92e-16	MMP9	matrix metalloproteinase 9
75	x 3.62	4.79e-16	BCL2A1	BCL2-related protein A1
76	x 3.6	5.92e-19	LGALS2	galectin 2
77	x 3.58	7.95e-21	RAC2	ras-related C3 botulinum toxin substrate 2
78	x 3.57	2.99e-18	LYZ	lysozyme
79	/ 3.57	1.91e-12	ELMOD1	ELMO domain containing 1
80	x 3.54	4.97e-15	GNLY	granulysin
81	x 3.53	2.7e-15	CCL2	C-C motif chemokine ligand 2
82	x 3.53	5.49e-19	P2RY8	purinergic receptor P2Y8
83	x 3.51	1.78e-20	CLEC10A	C-type lectin domain containing 10A
84	x 3.51	3.08e-15	LAMP3	lysosomal associated membrane protein 3
85	x 3.5	1.27e-10	CCL8	C-C motif chemokine ligand 8
86	/ 3.5	2.02e-10	LCE1C	late cornified envelope 1C
87	x 3.5	4.8e-20	TRAC	T-cell receptor alpha constant
88	/ 3.49	3.05e-13	KRT15	keratin 15
89	x 3.49	3.91e-19	IL4I1	Interleukin-4 induced 1
90	/ 3.48	2.79e-12	TSPAN8	tetraspanin 8
91	/ 3.48	2.17e-11	SERPINA12	serpin family A member 12
92	x 3.48	8.27e-19	CTLA4	cytotoxic T-lymphocyte-associated protein 4
93	x 3.48	1.52e-16	CH25H	cholesterol 25-hydroxylase
94	/ 3.48	7.72e-10	LCE2C	late cornified envelope 2C
95	x 3.47	7.52e-20	FYB	FYN binding protein
96	/ 3.45	8.02e-11	DCT	dopachrome tautomerase
97	x 3.45	7.06e-15	CD1B	CD1b molecule
98	/ 3.44	1.88e-13	CDHR1	cadherin related family member 1
99	x 3.43	1.23e-20	LCP2	lymphocyte cytosolic protein 2
100	x 3.41	3.6e-11	CMPK2	cytidine/uridine monophosphate kinase 2
101	x 3.39	5.05e-13	ALOX15	arachidonate 15-lipoxygenase
102	x 3.39	8.74e-18	GZMK	granzyme K
103	x 3.37	3.02e-12	CHI3L2	chitinase 3-like 2
104	x 3.36	2.29e-18	LCK	LCK proto-oncogene, Src family tyrosine kinase
105	x 3.36	2.18e-13	CCL17	C-C motif chemokine ligand 17
106	x 3.35	7.14e-15	TMPRSS4	transmembrane protease, serine 4
107	x 3.33	2.59e-20	TRAT1	T-cell receptor-associated transmembrane adaptator
108	x 3.33	6.71e-11	FPR2	formyl peptide receptor 2
109	x 3.33	3.03e-09	IDO1	indoleamine 2,3-dioxygenase 1
110	x 3.32	1.21e-17	CORO1A	coronin 1A
111	x 3.32	7.05e-22	CSF2RB	colony stimulating factor 2 receptor beta
112	/ 3.3	2.08e-09	LCE2D	late cornified envelope 2D
113	x 3.29	6.86e-19	TESPA1	thymocyte expressed, positive selection associated 1
114	x 3.28	6e-17	BIRC3	baculoviral IAP repeat-containing 3

115	x 3.27	6.05e-18	LILRB2	leukocyte immunoglobulin-like receptor B2
116	x 3.26	2.26e-09	ISG15	ubiquitin-like modifier
117	x 3.25	2.07e-20	KYNU	kynureninase
118	x 3.25	3.11e-24	PLEK	pleckstrin
119	x 3.23	1.87e-11	COL6A6	collagen type VI alpha 6 chain
120	x 3.23	8.06e-18	FCMR	Fc fragment of IgM receptor
121	x 3.23	4.22e-18	IL7R 3575	Interleukin-7 receptor
122	/ 3.19	8.77e-11	MSMB	microseminoprotein beta
123	x 3.19	7.28e-17	TCF7	transcription factor 7
124	x 3.17	1.3e-13	WARS	tryptophanyl-tRNA synthetase
125	x 3.17	1.61e-11	CXCL1	C-X-C motif chemokine ligand 1
126	x 3.17	3.94e-18	CD83	CD83 molecule
127	x 3.16	3.5e-16	SELL	selectin L
128	x 3.15	1.41e-09	MX1	MX dynamin-like GTPase 1
129	x 3.15	8.5e-19	TRBC1	T-cell receptor beta constant 1
130	/ 3.13	1.05e-10	KPRP	keratinocyte proline-rich protein
131	x 3.12	3.87e-18	PLSCR1	phospholipid scramblase 1
132	x 3.12	6.32e-15	SOD2	superoxide dismutase 2
133	/ 3.12	5.52e-13	AADACL2	arylacetamide deacetylase-like 2
134	x 3.12	1.88e-21	LCP1	lymphocyte cytosolic protein 1
135	x 3.12	2.36e-18	CTSC	cathepsin C
136	x 3.11	2.27e-21	ITGB2	integrin subunit beta 2
137	x 3.1	3.55e-16	PVRIG	poliovirus receptor-related immunoglobulin
138	x 3.1	2.29e-20	KLRB1	killer cell lectin-like receptor B1
139	/ 3.1	2.42e-12	WISP2	WNT1-inducible signaling pathway protein 2
140	x 3.1	8.17e-20	HCLS1	hematopoietic cell-specific Lyn substrate 1

Average expression levels of each gene from the allergen samples were compared to controls and expressed as the mean fold change (Fc). The Fc criteria were set at > 1.41 or $< 1/1.41$ and the false discovery rate (FDR) at < 0.05 . The statistical value of differences in expression levels was also calculated and expressed as an adjusted P value (adj-p).

The 140 highest differentially expressed genes are shown.

Gene symbols and gene descriptions are also indicated.

Table S2: Differentially expressed genes in the irritant group.

	Fc	Adj- p	Symbol	Description
1	/ 5.77	2.6e-08	LCE5A	late cornified envelope 5A
2	x 5.49	1.7e-05	ERPINB4	serpin family B member 4
3	x 5.03	8.67e-11	RRM2	ribonucleotide reductase regulatory subunit M2
4	x 4.53	1.31e-10	PCLAF	PCNA clamp-associated factor
5	x 4.51	6.67e-08	SERPINB3	serpin family B member 3
6	x 4.45	2.72e-06	S100A9	S100 calcium binding protein A9
7	/ 4.28	2.51e-08	IL37	interleukin 37
8	x 4.16	1.31e-10	CCNB1	cyclin B1
9	x 4.06	5.19e-10	KIF20A	kinesin family member 20A
10	x 3.86	4.85e-11	CDKN3	cyclin-dependent kinase inhibitor 3
11	x 3.75	2.31e-11	CKS2 1164	CDC28 protein kinase regulatory subunit 2
12	x 3.73	1e-09	UBE2C	ubiquitin-conjugating enzyme E2 C
13	x 3.68	4.97e-10	DLGAP5 9787	DLG-associated protein 5
14	x 3.62	1.02e-10	TYMS 7298	thymidylate synthetase
15	x 3.6	5.45e-09	TOP2A 7153	topoisomerase (DNA) II alpha
16	x 3.56	2.31e-11	ZWINT	ZW10-interacting kinetochore protein
17	x 3.56	1.57e-10	NCAPG	non-SMC condensin I complex subunit G
18	x 3.55	7.79e-11	CEP55	centrosomal protein 55
19	x 3.48	3.73e-11	TTK	TTK protein kinase
20	x 3.47	0.000318	PI3	peptidase inhibitor 3
21	x 3.45	3.33e-10	PTTG1	pituitary tumor-transforming 1
22	x 3.4	4.07e-06	SPRR1A	small proline-rich protein 1A
23	x 3.39	1.44e-08	S100A2	S100 calcium-binding protein A2
24	x 3.37	0.000374	SPRR2A	small proline-rich protein 2A
25	/ 3.36	6.82e-08	FCGBP	Fc fragment of IgG-binding protein
26	x 3.35	4.8e-10	CDK1	cyclin dependent kinase 1
27	x 3.34	2.05e-09	ANLN	anillin actin-binding protein
28	x 3.33	4.66e-07	SPRR1B	small proline-rich protein 1B
29	x 3.33	3.35e-06	LCE3D	late cornified envelope 3D
30	x 3.31	9.05e-10	ASPM	abnormal spindle microtubule assembly
31	x 3.25	6.58e-10	PBK	PDZ=binding kinase
32	/ 3.2	1.99e-07	CD207	CD207 molecule
33	x 3.19	5.97e-05	KRT6A	keratin 6A
34	x 3.19	1.89e-10	DEPDC1B	DEP domain-containing 1B
35	x 3.19	5.12e-10	CCNB2	cyclin B2
36	/ 3.17	9.37e-06	C1orf68	chromosome 1 open reading frame 68
37	x 3.16	4.93e-10	PRC1 9055	protein regulator of cytokinesis 1
38	x 3.15	0.000119	KRT6C	keratin 6C
39	x 3.15	3.52e-09	TOP2A	topoisomerase (DNA) II alpha
40	x 3.03	7.82e-10	NDC80	NDC80, kinetochore complex component
41	/ 3.02	8e-07	MSMB	microseminoprotein beta
42	x 3.02	5.19e-10	SPC25	SPC25, NDC80 kinetochore complex component
43	x 3	5.28e-09	NUSAP1	nucleolar and spindle-associated protein 1
44	x 3	5.4e-07	FGFBP1	fibroblast growth factor-binding protein 1
45	x 2.93	1e-09	TPX2	TPX2, microtubule nucleation factor
46	x 2.92	1.83e-09	BUB1B	BUB1 mitotic checkpoint serine/threonine kinase
47	x 2.9	1.51e-06	AKR1B10P1	aldo-keto reductase family 1 member B10 pseudogene 1
48	x 2.9	4.8e-10	CCNA2	cyclin A2
49	x 2.88	1.59e-10	TK1	thymidine kinase 1
50	x 2.87	5.66e-05	S100A8	S100 calcium-binding protein A8
51	/ 2.85	5.78e-05	LCE2B	late cornified envelope 2B
52	/ 2.84	2.31e-11	UGT3A2	UDP glycosyltransferase family 3 member A2
53	x 2.81	0.00227	SPRR2B	small proline-rich protein 2B
54	x 2.81	5.81e-10	KIF23	kinesin family member 23
55	x 2.78	2.31e-11	H2AFX	H2A histone family member X

56	x 2.78	6.24e-10	BUB1	BUB1 mitotic checkpoint serine/threonine kinase
57	x 2.76	1.26e-10	CDCA5	cell division cycle associated 5
58	x 2.75	1.64e-09	ESCO2	establishment of sister chromatid cohesion N-acetyltransferase 2
59	x 2.74	1.46e-05	SPP1	secreted phosphoprotein 1
60	x 2.72	2.31e-11	GIN52	GIN5 complex subunit 2
61	x 2.72	1.61e-10	AURKA	aurora kinase A
62	x 2.71	1.65e-10	PCLAF	PCNA clamp associated factor
63	x 2.7	9.09e-10	KNL1	kinetochore scaffold 1
64	x 2.69	5.73e-09	APOBEC3B	apolipoprotein B mRNA editing enzyme catalytic subunit 3B
65	x 2.69	1.58e-11	CDC6	cell division cycle 6
66	x 2.65	7.64e-10	MAD2L1	mitotic arrest deficient 2 like 1
67	x 2.63	7.79e-11	TRIP13	thyroid hormone receptor interactor 13
68	x 2.59	6.22e-10	HMMR	hyaluronan mediated motility receptor
69	x 2.57	4.41e-10	UBE2CP1	ubiquitin conjugating enzyme E2 C pseudogene 1
70	x 2.57	1.93e-06	IGFL1	IGF-like family member 1
71	/ 2.54	1.76e-05	WIF1	WNT inhibitory factor 1
72	x 2.54	2.44e-10	MND1	meiotic nuclear divisions 1
73	/ 2.54	1.99e-07	HSD11B1	hydroxysteroid 11-beta dehydrogenase 1
74	/ 2.51	3.16e-09	LRP4	LDL receptor-related protein 4
75	x 2.5	7.79e-11	UBE2T	ubiquitin-conjugating enzyme E2 T
76	x 2.5	3.16e-09	CDC20	cell division cycle 20
77	/ 2.49	1.12e-06	CDHR1	cadherin-related family member 1
78	x 2.49	7.37e-05	LCE3E	late cornified envelope 3E
79	x 2.48	1.89e-10	CCNE2	cyclin E2
80	x 2.48	9.87e-09	CENPA	centromere protein A
81	x 2.47	1.17e-07	HPSE	heparanase
82	x 2.47	4.8e-10	MELK	maternal embryonic leucine zipper kinase
83	x 2.46	1.93e-09	NUF2	NUF2, NDC80 kinetochore complex component
84	x 2.45	3.76e-10	SKA1	spindle and kinetochore associated complex subunit 1
85	/ 2.39	0.000411	LCE2C	late cornified envelope 2C
86	x 2.39	0.000102	S100A7	S100 calcium-binding protein A7
87	x 2.38	1.93e-09	KIF11	kinesin family member 11
88	x 2.38	1.56e-10	CDKN3	Cyclin-dependent kinase inhibitor 3
89	x 2.38	1.89e-10	ZWINT	ZW10-interacting kinetochore protein
90	x 2.37	1.05e-09	KIF2C	kinesin family member 2C
91	/ 2.36	6.03e-05	SERPINA12	serpin family A member 12
92	x 2.35	4.46e-11	CENPW	centromere protein W
93	x 2.34	1.02e-10	FAM111B	family with sequence similarity 111 member B
94	x 2.34	0.000588	S100A7A	S100 calcium-binding protein A7A
95	x 2.32	6.55e-10	OIP5	Opa-interacting protein 5
96	x 2.31	1.85e-10	CENPK	centromere protein K
97	x 2.31	2.03e-09	PIF1	PIF1 5'-to-3' DNA helicase
98	/ 2.29	0.000613	LCE2D	late cornified envelope 2D
99	x 2.29	3.35e-06	SELE	selectin E
100	x 2.28	6.09e-10	HJURP	Holliday junction recognition protein
101	/ 2.27	5.61e-09	BTC	betacellulin
102	x 2.26	9.22e-10	BIRC5	baculoviral IAP repeat-containing 5
103	/ 2.26	0.00235	FABP7	fatty acid-binding protein 7
104	x 2.25	2.18e-10	FANCI	Fanconi anemia complementation group I
105	x 2.25	1.81e-12	PCNA	proliferating cell nuclear antigen
106	x 2.25	2.63e-10	AURKB	aurora kinase B
107	x 2.24	0.000115	GJB2	gap junction protein beta 2
108	x 2.23	1.02e-10	DTL	denticleless E3 ubiquitin protein ligase homolog
109	x 2.22	2.03e-09	NEK2	NIMA-related kinase 2
110	x 2.21	0.00165	SPRR2D	small proline-rich protein 2D
111	x 2.21	2.71e-09	CENPF	centromere protein F
112	x 2.18	1.57e-10	CDCA3	cell division cycle associated 3
113	/ 2.17	0.000175	KPRP	keratinocyte proline-rich protein

114	/ 2.17	0.000106	LCE2A	late cornified envelope 2A
115	x 2.17	2.07e-11	HELLS	helicase, lymphoid-specific
116	x 2.15	1.27e-10	GGCT	gamma-glutamylcyclotransferase
117	x 2.15	2.03e-09	SMC4	structural maintenance of chromosomes 4
118	x 2.15	2.41e-10	PCLAF	PCNA clamp-associated factor
119	x 2.13	0.0124	S100A7A	S100 calcium-binding protein A7A
120	x 2.12	1.59e-07	FCN1	ficolin 1
121	/ 2.11	0.000974	LCE1C	late cornified envelope 1C
122	x 2.11	4.8e-10	CENPE	centromere protein E
123	/ 2.11	6.75e-06	CD1A	CD1a molecule
124	/ 2.11	0.00704	KRT2	keratin 2
125	x 2.1	1.6e-06	CH25H	cholesterol 25-hydroxylase
126	/ 2.1	0.000123	LCE6A	late cornified envelope 6A
127	x 2.09	3.32e-09	DEPDC1	DEP domain-containing 1
128	x 2.09	5.46e-11	STIL	SCL/TAL1 interrupting locus
129	/ 2.08	4.77e-09	F3	coagulation factor III, tissue factor
130	x 2.08	1.01e-09	KIF15	kinesin family member 15
131	x 2.08	6.7e-07	KRT78	keratin 78
132	/ 2.07	5.54e-08	CHL1	cell adhesion molecule L1-like
133	x 2.07	0.00998	CXCL8	C-X-C motif chemokine ligand 8
134	x 2.07	1.96e-09	MKI67	marker of proliferation Ki-67
135	x 2.07	5.7e-09	TIMP1	TIMP metalloproteinase inhibitor 1
136	x 2.06	5.66e-09	SLC7A5	solute carrier family 7 member 5
137	x 2.06	9.34e-07	TNFRSF12A	TNF receptor superfamily member 12A
138	x 2.05	6.19e-09	PIMREG	PICALM-interacting mitotic regulator
139	x 2.05	1.13e-10	MCM6	minichromosome maintenance complex component 6
140	x 2.05	9.6e-10	LMNB1	lamin B1

Average expression levels of each gene from the irritant samples were compared to controls and expressed as the mean fold change (Fc). The Fc criteria were set at > 1.41 or $< 1/1.41$ and the false discovery rate (FDR) at < 0.05 . The statistical value of differences in expression between the samples and controls was also calculated and expressed as an adjusted P value (adj-p).

The 140 highest differentially expressed genes are shown.

Gene symbols and gene descriptions are also indicated.

Table S3: Main signaling pathways in allergen and irritant patch test responses.

Common pathways

GO:0018149 peptide cross-linking (1) — 47 genes in gene set, 19 in common: LCE5A, PI3, LCE3D, SPRR2A, C1orf68, LCE2B, LCE2C, SPRR2B, LCE1C, LCE2D, SPRR1B, LCE3E, LCE2A, SPRR2D, KRT2, LCE1A, LCE1F, LCE1E, COL3A1

GO:0000819 sister chromatid segregation (0.99) — 204 genes in gene set, 34 in common: CCNB1, UBE2C, ZWINT, DLGAP5, TOP2A, NCAPG, TTK, NDC80, SPC25, PRC1, NUSAP1, BUB1, CDCA5, BUB1B, KIF23, CDC6, MAD2L1, KNL1, SKA1, CDC20, ...

GO:0085029 extracellular matrix assembly (0.81) — 20 genes in gene set, 5 in common: HAS3, MFAP4, MYH11, FBLN5, GAS6

GO:0051782 negative regulation of cell division (0.78) — 12 genes in gene set, 6 in common: BLM, SUSP2, ORC6, C10orf99, CHMP4C, TXNIP

GO:0046605 regulation of centrosome cycle (0.59) — 41 genes in gene set, 7 in common: AURKA, KIF11, STIL, NPM1, CHEK1, CHMP4C, RANBP1

GO:1901655 cellular response to ketone (0.58) — 28 genes in gene set, 7 in common: UGT3A2, SPP1, GNAI1, AQP1, POSTN, GAS6, ADCY2

GO:0046134 pyrimidine nucleoside biosynthetic process (0.52) — 20 genes in gene set, 5 in common: TYMP, TYMS, TK1, UPP1, CTPS1

GO:0071241 cellular response to inorganic substance (0.49) — 103 genes in gene set, 12 in common: CHP2, MMP9, TIGAR, ECT2, BLM, TIMELESS, MT2A, RYR3, AQP1, MT1G, MEF2A, CACNA1H

GO:0000083 regulation of transcription involved in G1/S transition of mitotic cell cycle (0.49) — 23 genes in gene set, 7 in common: RRM2, TYMS, CDK1, CDC6, PCNA, CDC45, FBXO5

GO:0070542 response to fatty acid (0.44) — 26 genes in gene set, 5 in common: CCL19, CCR7, DGAT2, TGFBR3, PDK4

GO:0042634 regulation of hair cycle (0.42) — 10 genes in gene set, 3 in common: CDH3, HPSE, PER1

GO:1901654 response to ketone (0.41) — 50 genes in gene set, 10 in common: CCL19, CCR7, UGT3A2, SPP1, GNAI1, AQP1, POSTN, TGFBR3, GAS6, ADCY2

GO:0051043 regulation of membrane protein ectodomain proteolysis (0.35) — 21 genes in gene set, 4 in common: TIMP1, TIMP3, APOE, TIMP2

GO:1905048 regulation of metalloproteinase activity (0.33) — 10 genes in gene set, 3 in common: TIMP1, TIMP3, TIMP2

GO:0033260 nuclear DNA replication (0.24) — 24 genes in gene set, 5 in common: GINS2, PCNA, BRCA2, CDC45, CDC7

GO:0000920 cell separation after cytokinesis (0.23) — 16 genes in gene set, 3 in common: KIF20A, CEP55, CHMP4C

GO:0045103 intermediate filament-based process; GO:0045104 intermediate filament cytoskeleton organization (0.23) — 29 genes in gene set, 4 in common: KRT6C, KRT6A, EPPK1, DES

Allergen pathways

GO:0043062 extracellular structure organization (1) — 238 genes in gene set, 67 in common: MMP12, MMP1, TNC, ITGB2, VCAM1, LCP1, ICAM1, CD47, HTRA1, SPOCK2, MMP3, COL4A5, ICAM3, CTSL, CTSS, LAMA2, ITGAM, COL14A1, CTSV, MMP2, ...

GO:0006954 inflammatory response (1) — 376 genes in gene set, 126 in common: CXCL10, CXCL9, CXCL11, CXCR4, GBP5, CCL4, JAK3, FPR1, CCL2, CCL8, ALOX15, FPR2, BIRC3, PLSCR1, CXCL1, ITGB2, CCL13, CCL5, CCL26, HCK, ...

GO:0045321 leukocyte activation (0.98) — 897 genes in gene set, 274 in common: LTF, GPR183, FLG2, CD53, CD2, PTPRC, JAK3, SERPINA1, FPR1, TRAC, CTLA4, CCL2, LCK, CORO1A, LILRB2, TESPA1, FPR2, TRBC1, SELL, GMFG, ...

GO:0050776 regulation of immune response (0.93) — 732 genes in gene set, 224 in common: LTF, GBP5, PTPRC, GBP1, JAK3, CLEC4A, FPR1, TRAC, CTLA4, FYB, LCK, LCP2, CD1B, TRAT1, ALOX15, LILRB2, TESPA1, FPR2, BIRC3, TRBC1, ...

GO:0050900 leukocyte migration (0.91) — 319 genes in gene set, 113 in common: MMP1, CXCL10, CXCL9, CXCL11, GPR183, CD2, CXCR4, CCL4, FPR1, RAC2, CCL2, CCL8, LCK, FPR2, SELL, CXCL1, ITGB2, VCAM1, CCL13, CCL5, ...

GO:0009615 response to virus (0.86) — 218 genes in gene set, 76 in common: CXCR4, PTPRC, ISG20, CCL4, ISG15, CCL8, LCK, MX1, BIRC3, PLSCR1, APOBEC3A, CCL5, IRF1, BATF3, DOCK2, HCK, APOBEC3G, CD247, MX2, RSAD2, ...

GO:2001236 regulation of extrinsic apoptotic signaling pathway (0.77) — 108 genes in gene set, 28 in common: ICAM1, CYLD, IL1B, TNFAIP3, MCL1, TNFSF10, THBS1, FAS, BID, GPX1, TRAF1, TCF7L2, PML, TNF, CFLAR, TNFRSF10A, SCG2, IGF1, FASLG, ATF3, ...

GO:0070268 cornification (0.77) — 107 genes in gene set, 28 in common: LOR, FLG, KRT15, DSC1, SPRR2F, KLK5, KRT80, KRT31, KRT19, KRT17, KRT23, DSG1, PKP1, CDSN, KRT6B, TGM5, SPRR2G, PPL, KRT1, CAPNS1, ...

GO:0045926 negative regulation of growth (0.7) — 147 genes in gene set, 28 in common: MT1M, SFRP2, BST2, RERG, FHL1, FRZB, SLIT3, MT1F, PML, BBS2, FGFR3, SMAD4, ALOX15B, GJA1, MT1E, ACVR1B, PPT1, TGFB1, RBP4, MT1A, ...

GO:0030155 regulation of cell adhesion (0.63) — 450 genes in gene set, 150 in common: CD2, PTPRC, GBP1, JAK3, RAC2, TRAC, CTLA4, CCL2, LCK, ALOX15, LILRB2, TESPA1, TRBC1, VCAM1, CCL5, ICOS, IRF1, ARHGDIB, CD47, HLA-DQA1, ...

GO:0098754 detoxification (0.58) — 22 genes in gene set, 10 in common: SOD2, BMP7, SESN1, GCH1, GSTM3, TNF, ATP7A, MARC2, MARC1, PRDX2

GO:0045861 negative regulation of proteolysis (0.48) — 150 genes in gene set, 36 in common: LTF, LAMP3, BIRC3, SERPINB9, SPOCK2, PLAUR, CTSZ, SERPINE2, IFI16, THBS1, RECK, TFAP2B, GPX1, LEF1, PML, AQP1, CR1, BEX3, SPOCK1, GAS1, ...

GO:0034340 response to type I interferon (0.41) — 77 genes in gene set, 34 in common: ISG20, ISG15, MX1, IRF1, IRF8, STAT1, IFI35, MX2, RSAD2, IRF7, XAF1, BST2, IFIT3, IFITM1, SAMHD1, OAS2, PSMB8, OAS1, IFITM3, IFITM2, ...

GO:0051216 cartilage development (0.41) — 77 genes in gene set, 20 in common: OSR2, SOX6, FGFR3, LRP6, RUNX3, RUNX2, SCIN, BARX2, GPLD1, TGFB1, MGP, ATP7A, PKD1, OSR1, RUNX1, NFIB, CHADL, TRPS1, SULF1, CD44

GO:0006575 cellular modified amino acid metabolic process (0.39) — 129 genes in gene set, 24 in common: PLSCR1, ETHE1, MTHFD1L, GSTA4, GPX1, GSTM3, GSTM5, GATM, GHR, PCYOX1, CRYM, PLA2G12A, GSTA3, SLC5A6, CKMT2, STAT5A, FASN, PLA1A, VNN2, SHMT1, ...

GO:0097164 ammonium ion metabolic process (0.38) — 154 genes in gene set, 32 in common: RARRES3, ACER1, BCHE, SLC44A5, GCH1, AKR1C2, LDLR, ODC1, APOC1, AKR1C3, PLA2G12A, MBOAT2, LPCAT1, SPTLC3, PLA2G3, GPLD1, HPRT1, BBOX1, ENPP2, ATP7A, ...

GO:0018149 peptide cross-linking (0.38) — 47 genes in gene set, 15 in common: LOR, FLG, SPRR2F, SPRR4, LCE1B, THBS1, LCE3A, LCE1D, COL3A1, F13A1, KRT1, ANXA1, LCE4A, KRT10, EVPL

GO:0051897 positive regulation of protein kinase B signaling (0.34) — 60 genes in gene set, 19 in common: HCLS1, CCL3, PIK3R5, CCL21, THBS1, IL18, TCF7L2, EGFR, AKR1C2, AKR1C3, PINK1, F10, TGFB1, INSR, C1QBP, SPRY2, PIK3CG, ADAM8, MYDGF

GO:1900077 negative regulation of cellular response to insulin stimulus (0.3) — 25 genes in gene set, 13 in common: CISH, SOCS1, PRKCQ, IL1B, PTPRE, KANK1, PTPN2, IRS1, SOCS3, PID1, SOCS2, GRB14, MZB1 D18346_SIV_JFN — Follow-up of project D14205_JFN_MV CONFIDENTIAL 9.1 Effect by Class of Compound 85 of 536

GO:0042590 antigen processing and presentation of exogenous peptide antigen via MHC class I (0.29) — 72 genes in gene set, 20 in common: PSMB10, FCER1G, CYBA, PSMB9, TAP2, TAP1, NCF2, CYBB, PSMB8, TAPBP, ITGB5, PSMA5, HLA-F, PSMA3, PSMD12, PSME1, NCF4, PSMB3, CALR, PSMA4

Irritant pathways

GO:0006260 DNA replication (1) — 235 genes in gene set, 33 in common: PIF1, PCLAF, RNASEH2A, MCM10, CENPS, POLQ, RFC4, MCM2, E2F8, GINS4, BRCA1, CDC25C, EXO1, MCM4, CDT1, RAD51, CDC25A, GINS1, RFC3, ORC1, ...

GO:0007067 mitotic nuclear division (0.84) — 266 genes in gene set, 35 in common: AURKB, NEK2, CENPF, CENPE, MKI67, KIF15, PRC1, KIFC1, CDCA8, KIF14, SMC4, RCC1, KIF4A, STMN1, SKA3, SPAG5, KNTC1, CDC25C, KIF18A, NCAPD2, ...

GO:0009262 deoxyribonucleotide metabolic process (0.74) — 25 genes in gene set, 4 in common: DTYMK, NUDT1, RRM1, NUDT15

GO:0034724 DNA replication-independent nucleosome organization (0.62) — 50 genes in gene set, 13 in common: CENPA, CENPW, OIP5, CENPS, ASF1B, CENPU, CENPH, RUVBL1, CENPL, NASP, CENPM, HIST1H4C, MIS18A

GO:0046148 pigment biosynthetic process (0.28) — 33 genes in gene set, 4 in common: TYRP1, TYR, MTHFD1, PAICS

GO:0000732 strand displacement (0.17) — 26 genes in gene set, 7 in common: RAD51AP1, BRCA1, EXO1, RAD51, RMI2, BRIP1, RAD51C

Gene Ontology terms (from the Biological Process tree) most associated with each group, according to the MGSA algorithm. The number in parentheses after each gene set name is the estimated frequency of activation (from 0 to 1). It is followed by the most differentially expressed genes associated with this term.

Table S4: Primer sequences, primer ID and detected transcripts for qRT-PCR analysis.

Gene name	Primer sequences	Amplicon length (bp)
<i>CD8A</i>	Forward: CAGGAACCGAAGACGTGTTTGC Reverse: AGTGGCTGTTGCACAGGGTTAG	103
<i>CD4</i>	Forward: CTAGTTTCCAGAGGCTTAATCACACC Reverse: AGGGTCAGAGAGAAATAATGAGAGAAGG	90
<i>GZMB</i>	Forward: CTTCTCCAACGACATCATGC Reverse: ACAGCTCTGGTCCGCTTG	60
<i>IFNG</i>	Forward: TGACCAGAGCATCCAAAAGAGTG Reverse: TTGCTTTGCGTTGGACATTCAAG	143
<i>Actin</i>	Forward: CCAACCGCGAGAAGATGA Reverse: CCAGAGGCGTACAGGGATAG	97
<i>GAPDH</i>	Forward: CACCCACTCCTCCACCTTTG AC Reverse: GTCCACCACCCTGTTGCTGTAG	112

Gene name	Primer ID	Detected transcript(s)	Amplicon length (bp)
<i>CCL2</i>	Hs_CCL2_1_SG	NM_002982	60
<i>CCL20</i>	Hs_CCL20_1_SG	NM_004591	102
<i>CXCL10</i>	Hs_CXCL10_1_SG	NM_001565	129
<i>CD69</i>	Hs_CD69_1_SG	NM_001781, NR_026671, NR_026672	90
<i>GNLY</i>	Hs_GRN_1_SG	NM_002087, XM_005257253	140

Supplemental figures

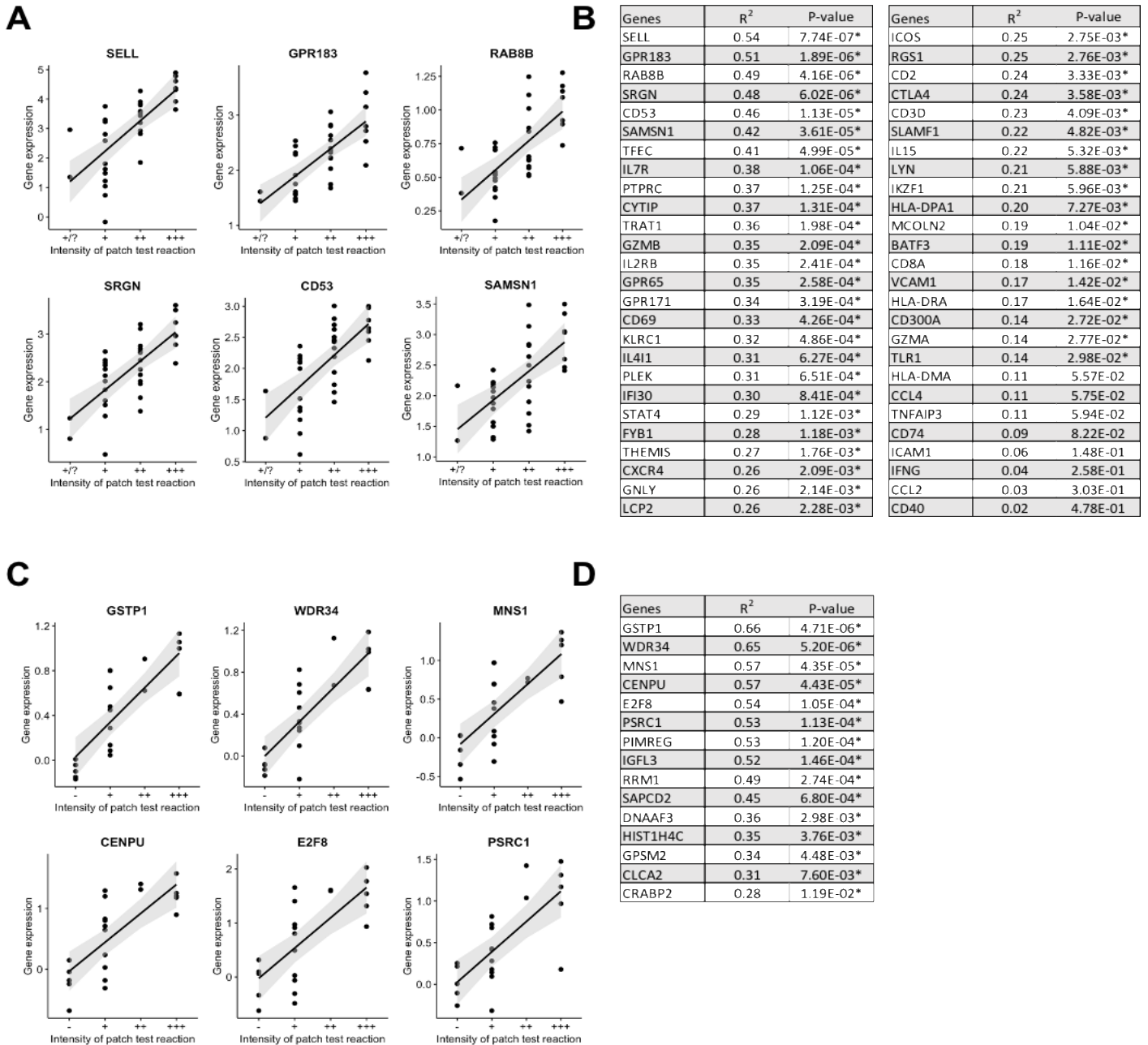


Figure S1: Variations in gene expression for individual genes of core allergen and irritant signatures correlate with the magnitude of the clinical response.

Correlations between the variations in gene expression of the 52 and 15 genes of the allergen (A & B) and irritant (C & D) core signatures and the intensity of the patch-test reactions were investigated by simple linear regression analysis. Results were illustrated as scatter plots for

the twelve most correlative genes (**A & C**), or depicted as coefficients of determination (R^2) and respective p-value (**B & D**). *P<0.05.

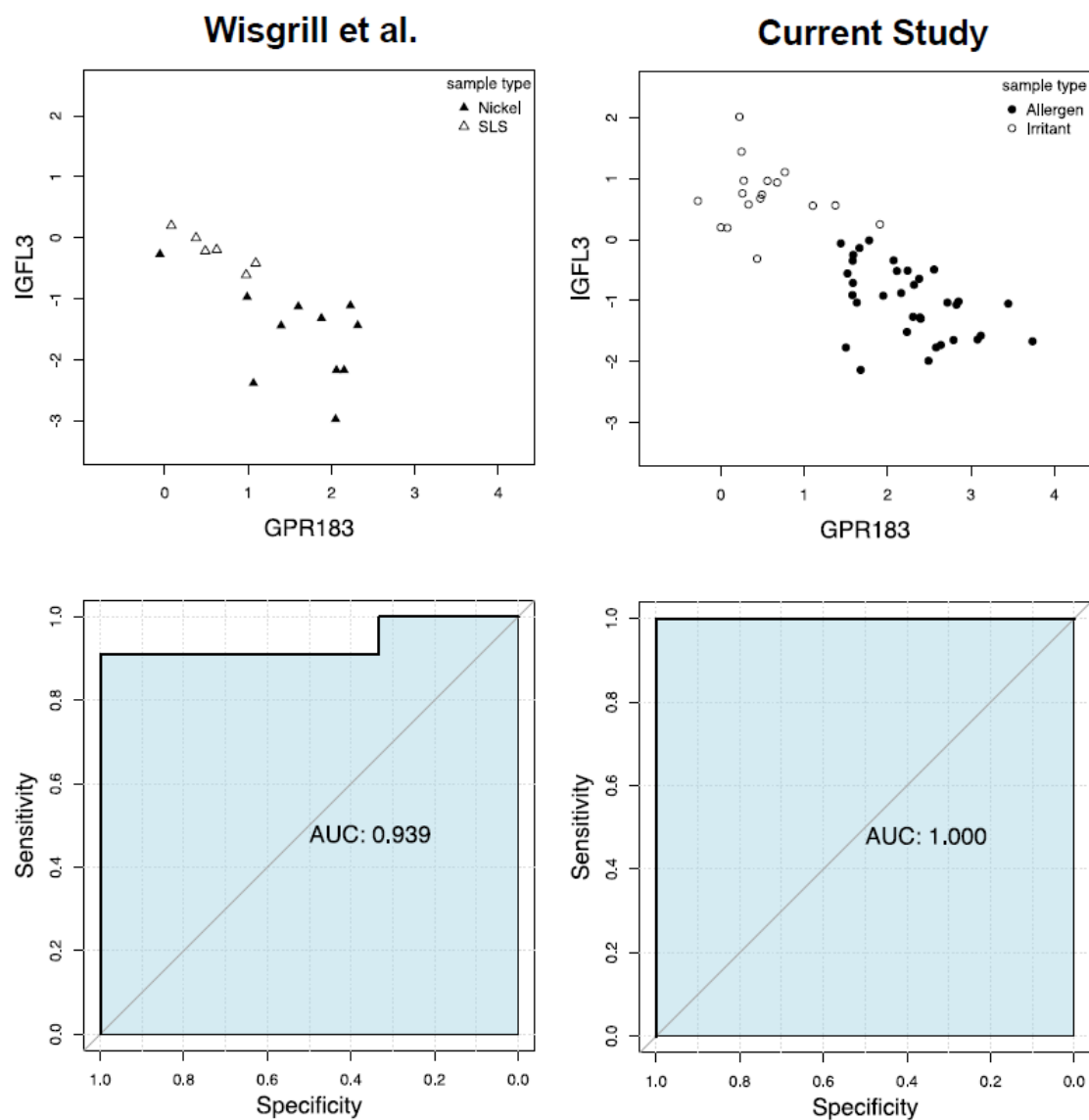


Figure S2: Validation of GPR183 / IGFL3 classifier using external gene array datasets.

GPR183 / IGFL3 classifier was validated on external gene array datasets, generated from biopsies of positive PT samples to nickel and SLS (access no. E-MTAB-8945).

The ability of GPR183 / IGFL3 classifier to discriminate between nickel and SLS samples is depicted, either as biplots representing differential gene expression (Fc, top) or as ROC curves (bottom). AUC is the area under the ROC curve.

Results obtained with the allergen and irritant samples of the current study are indicated for comparison.