

DIFFERENTIAL GENE-EXPRESSION ANALYSIS DEFINES A MOLECULAR PATTERN RELATED TO OLIVE POLLEN ALLERGY

M. AGUERRI¹, D. CALZADA¹, D. MONTANER², M. MATA^{3,4}, F. FLORIDO⁵, J. QUIRALTE⁶, J. DOPAZO², C. LAHOZ^{1,4} and B. CÁRDABA^{1,4}

¹Immunology Department, IIS-Fundación Jiménez Díaz, Madrid, Spain; ²Functional Genomic Unit Bioinformatics Department. Centro de Investigación Príncipe Felipe (CIPF), Valencia, Spain; ³Genomic Department, Fundación de Investigación, Hospital General Universitario, Valencia, Spain; ⁴CIBER de Enfermedades Respiratorias, CIBERES, Spain; ⁵Allergy Department, Hospital Universitario San Cecilio, Granada, Spain; ⁶Allergy Department, Hospital Universitario Virgen del Rocío, Sevilla, Spain

Received October 2, 2012 – Accepted February 8, 2013

The first two authors have contributed equally to this work.

Analysis of gene-expression profiles by microarrays is useful for characterization of candidate genes, key regulatory networks, and to define phenotypes or molecular signatures which improve the diagnosis and/or classification of the allergic processes. We have used this approach in the study of olive pollen response in order to find differential molecular markers among responders and non-responders to this allergenic source. Five clinical groups, non-allergic, asymptomatic, allergic but not to olive pollen, untreated-olive-pollen allergic patients and olive-pollen allergic patients (under specific-immunotherapy), were assessed during and outside pollen seasons. Whole-genome gene expression analysis was performed in RNAs extracted from PBMCs. After assessment of data quality and principal components analysis (PCA), differential gene-expression, by multiple testing and, functional analyses by KEGG, for pathways and Gene-Ontology for biological processes were performed. Relevance was defined by fold change and corrected *P* values (<0.05). The most differential genes were validated by qRT-PCR in a larger set of individuals. Interestingly, gene-expression profiling obtained by PCA clearly showed five clusters of samples that correlated with the five clinical groups. Furthermore, differential gene expression and functional analyses revealed differential genes and pathways in the five clinical groups. The 93 most significant genes found were validated, and one set of 35 genes was able to discriminate profiles of olive pollen response. Our results, in addition to providing new information on allergic response, define a possible molecular signature for olive pollen allergy which could be useful for the diagnosis and treatment of this and other sensitizations.

Allergic diseases are highly complex adverse reactions of the immune system against otherwise harmless substances. The reason why exposure to common environmental antigens induces allergic

Key words: microarray, gene expression, Gene set enrichment, genomic and allergy, molecular signature, olive pollen sensitization

Mailing address: Blanca Cárdaa,
Immunology Department,
IIS- Fundación Jiménez Díaz,
Avda Reyes Católicos, 2,
28040 Madrid Spain
Tel.: +34 1 5494445 Fax: +34 1 5448246
e-mail: bcardaba@fjd.es

0393-974X (2013)

Copyright © by BIOLIFE, s.a.s.

This publication and/or article is for individual use only and may not be further reproduced without written permission from the copyright holder.

Unauthorized reproduction may result in financial and other penalties

DISCLOSURE: ALL AUTHORS REPORT NO CONFLICTS OF INTEREST RELEVANT TO THIS ARTICLE.

responses in some people and not others remains unexplained.

To date the search for specific protective or risk biomarkers in allergic diseases is ongoing. Even though significant findings regarding susceptibility chromosome regions and genes have been achieved (1, 2), these studies contribute only slightly to the understanding of these diseases. However, new tools such as genomics, proteomics or transcriptomics have induced a revolutionary change in the general concept and study of complex biological systems. Major applications of this methodology have been related with cancer or single-gene disorders but in recent years, this approach has yielded some interesting results in complex disorders, such as allergic diseases, as has been recently revised (3). Three major kinds of advances can be pointed out: firstly, the discovery and verification of differentially expressed genes, with the description of many potential biomarkers and/or therapeutic targets (4-6); secondly, the description of molecular signatures associated with clinical subphenotypes, as has been demonstrated in asthma, where the degree of TH2 inflammation defined at least two distinct molecular phenotypes (7); lastly, the possibility of discovering deregulated pathways through the analysis of modules that include tightly interacting genes, which are usually functionally related (8).

We used this experimental approach to analyze differential molecular mechanisms implicated in allergic response to olive pollen, one of the major causes of allergic diseases in Mediterranean countries and some parts of Australia and North America. Olive pollen induces mainly nasal and conjunctivae symptoms, although it may also induce asthma exacerbations in areas with high airborne levels of pollen. We previously described specific genetic and environmental factors associated with olive pollen allergy (9-12). In this study we analyzed whole-genome expression profile in peripheral blood cells of responders and non-responders subjects to this allergen in situations of both high and low exposure. The subjects were selected from a region in southern Spain with particularly high pollen counts during the pollen season and a high prevalence of asthma (10). Our specific purpose was to analyze the underlying differential molecular mechanisms in sensitization and natural or induced tolerance to olive pollen,

checking the influence of olive pollen antigenic load in order to identify biomarkers or deregulated pathways that could prove useful in the diagnosis and/or prophylaxis of allergic diseases.

MATERIALS AND METHODS

Subjects

The study population comprised 30 subjects selected from a previous immunological study (13). These subjects were recruited in 2 olive-pollen-exposure situations: inside (April-June 2006) and outside pollen season (October-December 2006). We established 5 groups, and 6 subjects from each group were selected for gene-expression analysis (Table I): Group 1, non-allergic subjects (skin tests negative and no history of respiratory allergies); Group 2, asymptomatic subjects (diagnosed with olive pollen allergy by skin testing, with no seasonal respiratory symptoms [rhinitis and/or asthma]); Group 3, patients who were allergic, but not to olive pollen; Group 4, untreated olive-pollen-allergic; and Group 5, treated olive-pollen-allergic patients (receiving olive-pollen specific immunotherapy).

Olive-pollen-allergic patients fulfilled the following criteria: seasonal rhinitis and/or asthma from April to June, a positive skin prick test result for *O. europaea* pollen extract (ALK Abelló, Madrid, Spain), and no previous immunotherapy except for Group 5 (13).

Informed consent was obtained from each subject. Ethical approval for the study was obtained from the Ethics and Research Committee of the participating hospitals. Biological samples were obtained from Biobanco-FJD, IIS-Fundación Jiménez Díaz Madrid.

Isolation of peripheral blood mononuclear cells (PBMCs)

PBMCs were isolated from heparin-containing peripheral blood samples taken inside and outside of pollen season, by gradient centrifugation on Lymphoprep (Comercial Rafer, Zaragoza, Spain) following the manufacturer's instructions.

Microarray preparation

Total RNA was isolated from PBMCs (10^6 cells) using Trizol (Invitrogen, Carlsbad, CA, USA) method. An equimolar amount of RNA from 2 independent donors was pooled prior to labelling and hybridization, strategy previously described for microarray studies (14). Three independent pools were created for each clinical group and pollen-exposure condition. A total of 30 RNAs were analyzed by microarrays, fifteen for each exposure condition, but two samples were rejected due to technical reasons (problems with the RNA quality).

Microarray targets were prepared from 200 ng of total RNA using the Affymetrix kit for reverse transcription and a double round of *in vitro* transcription. HuGe U133 plus 2.0 Gene Chips were hybridized with 10 µg of cDNA and processed according to the manufacturer's protocol. This process was performed in Scientific Park of Complutense, Madrid, Spain.

Image and data analysis

Preliminary microarray analysis included assessment of data quality by standard quality checks and principal component assays (PCA). First-line check, background corrections, and quantile normalization of the data were performed using the Robust Multiarray Averaging (RMA) method. Data analysis was performed using Partek Genomics Suite software.

Microarrays analysis

All primary data discussed in this work have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE37157 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE37157>). The analyses were carried out using the Babelomics web site (15).

Differential expression

Differential gene expression assessment of all comparisons was carried out using limma moderated *t*-statistics. The *P*-values were corrected for multiple testing using the correction of Benjamini Hochberg (16).

Functional profiling

Gene set analysis was carried out by the logistic regression models previously described (17). Functional blocks from two data-bases, GO Biological Process and KEGG Pathways were used. For each comparison, the number of functional statistical significant blocks (adjusted *P* value < 0.05), with differential genes enrichment, expressed by positive or negative LOR (Log Odds Ratio) was obtained.

Validation of differential gene-expression by qRT-PCR

Ninety three genes (Table II) were selected for validation by qRT-PCR in a bigger population. Selection criteria for these genes included a fold change >2 or <-2, an adjusted *P*-value <0.05, and representativeness of the higher number possible of differential pathways and biological processes.

A total of 80 subjects, 8 from each group and each pollen season, were selected from the same area and with the same criteria as the initial population (13). Demographic and clinical characteristics of these subjects were similar that the previously studied population (data

not shown). RNA of these 80 new samples was analyzed by duplicated. We used 384-well micro fluidic cards (Applied Biosystem, Foster City, CA, USA) with the 93 genes (Table II) and 3 reference genes (18S, GADPH, β_2 -Microglobuline). Briefly, two-hundred nanograms of total RNA (extracted from PBMCs) were sent to the Scientific Park of Cantoblanco, Madrid, Spain, to be analyzed by qRT-PCR using the TaqMan Gene Expression System (Applied Biosystems). Specific gene mRNA expression from each gene was normalized to three endogens genes (18S rRNA, GADPH and β_2 -Microglobuline) and calculated according to the cycle threshold (C_t) method. Statistical analyses were performed with StatMiner program.

Functional analyses of 35 genes selected were performed using the DAVID Bioinformatics Resources 6.7 program (<http://david.abcc.ncifcrf.gov/>) and the analysis of protein interactions was carried out with the STRING 9.0 (Known and Predicted Protein-Protein Interactions) free program (<http://string-db.org/>).

RESULTS

Patients and controls

Demographic and clinical parameters of 30 subjects studied, classified according clinical conditions, appear in Table I. Mean age was homogeneous, and there were slightly more women in the olive-pollen-allergic groups. The mean levels of *O. europaea* specific IgE antibodies and total IgE were measured during and outside of the pollen season. Total IgE was increased in all groups compared to the control group, being statistically significant ($P < 0.05$) compared with those from olive-pollen-allergic subjects. The highest specific IgE antibody levels were obtained in the two olive-pollen-allergic groups but only, the untreated olive-pollen-allergic group showed statistically significant differences compared to healthy controls, asymptomatic patients and those who were allergic but not to olive pollen.

Gene expression clustering according clinical conditions

Whole-genome gene expression analysis (38,500 well-characterized human genes) was performed on 28 RNA samples (2 RNA samples were excluded for low RNA quality).

After data normalization, the one-way ANOVA supervised analysis identified 1,746 transcripts

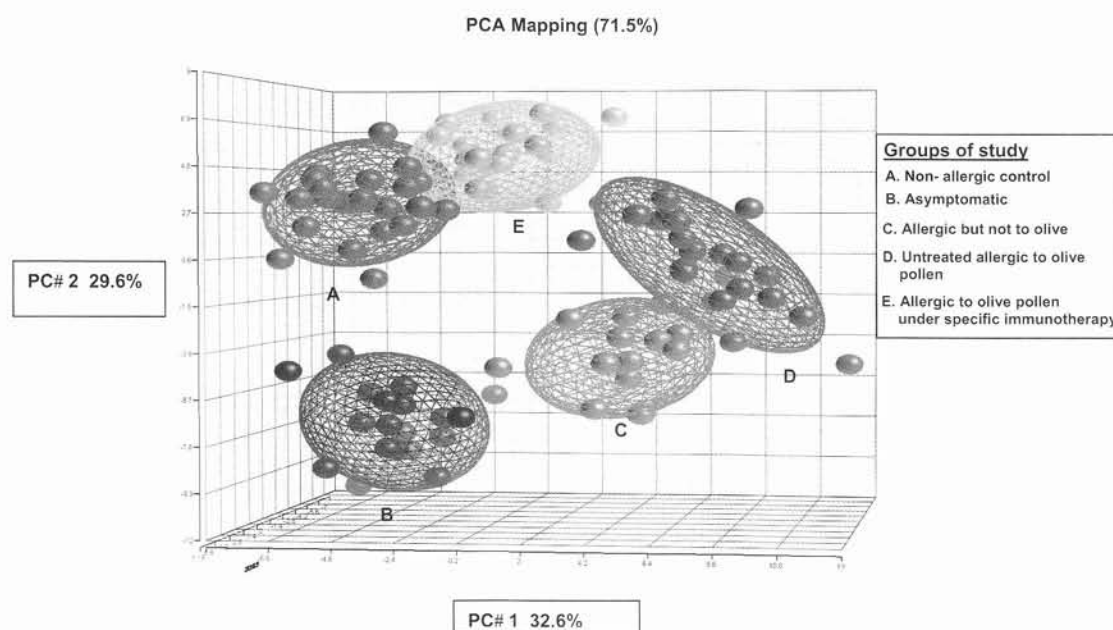


Fig. 1. Principal component analysis (PCA) to detect global genetic variation among samples from five clinical conditions. The gene-expression analysis by unsupervised PCA showed how it was possible to draw a clear division and separate the data points from the five clinical groups of study ($p < 0.001$).

differentially expressed ($P < 0.05$) from among the five selected groups. The false discovery rate threshold was 5%. Following this, a principal component analysis (PCA) was performed. A plot of the results of unsupervised clustering is shown in Fig. 1. As is clearly depicted, five different gene expression clusters were defined, correlating with the five clinical situations established in the study. The clear separation observed was consistent with P -values lesser than 0.05 and 0.001.

To determine the functional implications of differential gene expression, two kinds of gene set enrichment analysis were performed, KEGG for characterizing differential functional pathways and Gene Ontology for biological processes. A total number of 20 comparisons were carried out: 10 during and the other 10 outside of the pollen season. Thus, for each comparison, a list of up- and down-expressed genes (fold change > 2 or < -2 and adjusted P -values < 0.05), as well as a list of functional pathways and biological processes differentially regulated (with adjusted P -values < 0.05 and LOR) were obtained. In general, the greatest differences

were found in allergic groups, and most of the differences were up-regulated pathways and processes. According to pathways, the highest number of differences was found among untreated patients allergic to olive pollen and controls, with up to 31 pathways found as being up-regulated outside of pollen season and 24 during pollen season. Fig. 2 shows the specific pathways up- and down-regulated in untreated olive-pollen-allergic patients versus non-allergic controls during and outside of pollen season. Although there are many different kinds of differential pathways, many similar up-regulated routes exists during and outside of pollen season (marked in Fig. 2); some of these are related to immune response, such as antigen processing and presentation, graft-versus-host disease, autoimmune thyroid disease or B-cell receptor signalling pathway, and others which are not directly related, such as Parkinson's disease, oxidative phosphorylation or pathogenic *E. coli* infection. Interestingly, the only pathway that exhibited a statistically significant decrease in olive-pollen allergic subjects compared with controls, outside of pollen season, was the

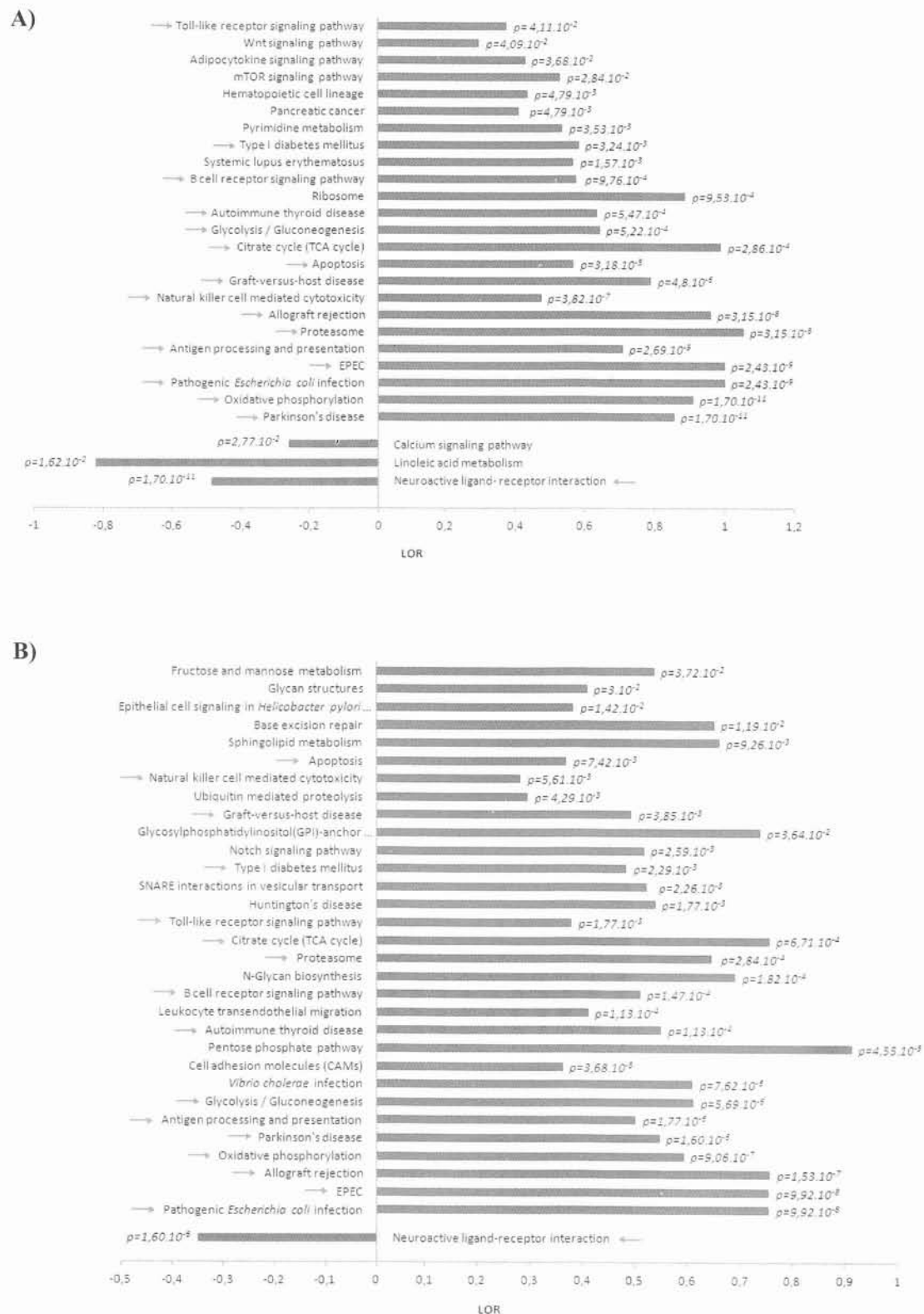
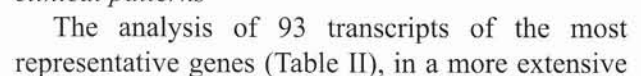


Fig. 2. Differential pathways ($P_{\text{adjusted}} < 0.05$) between untreated olive pollen allergic patients and non-allergic controls. **A:** Inside pollen season, **B:** Outside pollen season. LOR: Odds Ratio Logarithm. Arrows indicate differential pathways shared in both pollen seasons.



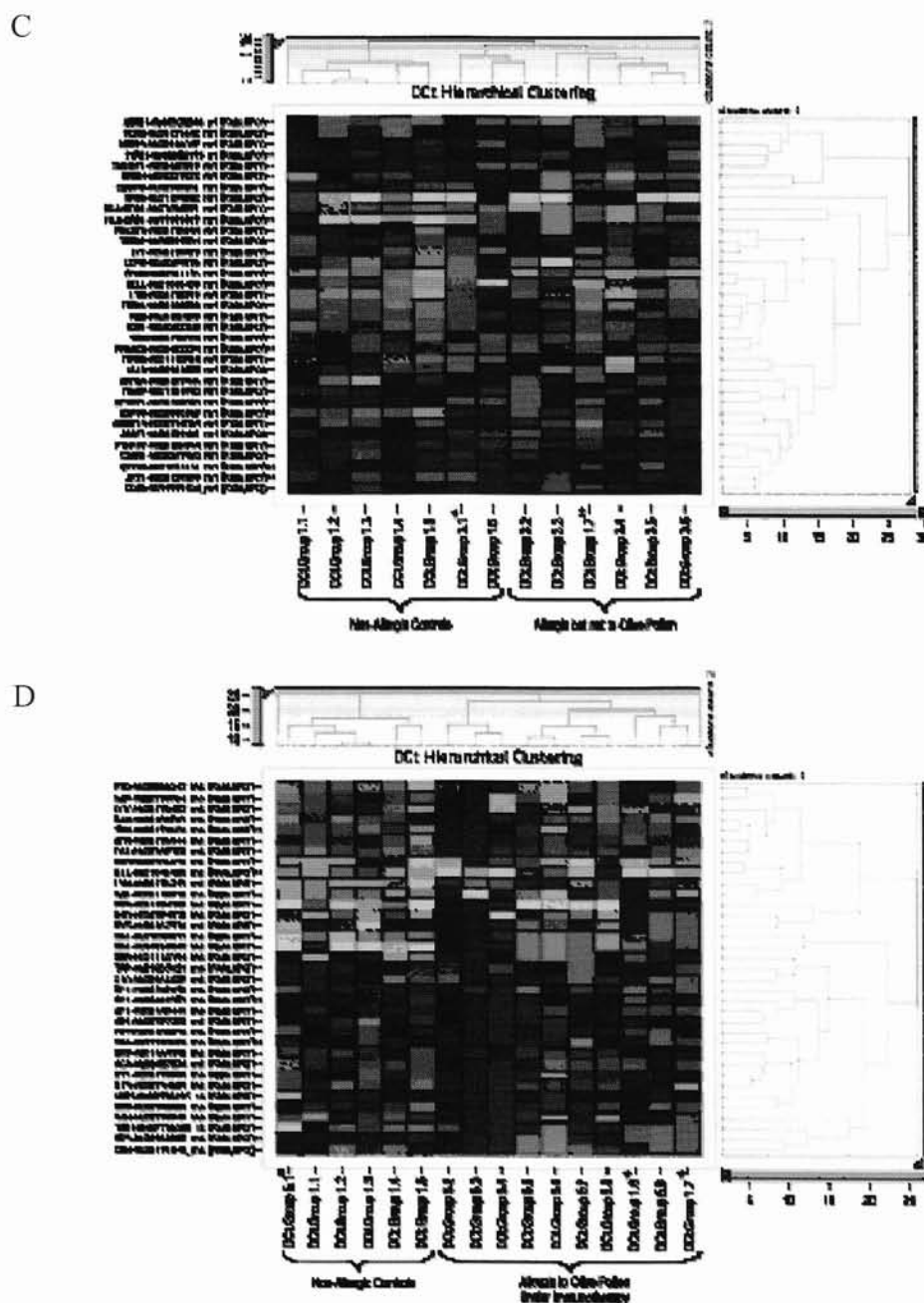


Fig. 3. Unsupervised hierarchical model separate clinical conditions according gene-expression of the 35 most differential genes. **A)** Classification of untreated olive- pollen-allergic patients and non-allergic controls. **B)** Classification of asymptomatic and non-allergic controls, **C)** Classification of allergic but not to olive pollen and non-allergic controls and **D)** Classification of olive-pollen-allergic patients with treatment and non-allergic controls.

population by qRT-PCR, confirmed 35 genes as being differentially expressed: ADRB1, AKT1, ALOX5, APAF1, BAX, CD40, CD48, CD86, C3AR1, CRTAP, FPR3, GZMH, HLA-DRB1, HLA-

DQB1, IL-10, ITGAL, ITGB7, ITGB8, LCK, LYN, MAPK13, NFTAC1, NLRP3, PRKACA, PRKACB, PTGER2, S1PR5, SELL, SMURF1, SOS1, SPN, TAGAP, TRIM37, VCAN, and ZAP70. When

Table I. Study subjects.

	N	Mean age (years)	Sex (%) female/male	INSIDE SEASON		OUTSIDE SEASON		Clinical diagnosis
				Total IgE	Specific IgE (kUA/l)	Total IgE	Specific IgE (kUA/l)	
Group 1 (Non-allergic control)	6	35±10	50/50	42.27 ± 41.68	0.30 ± 0.00	39.62 ± 38.98	0.30 ± 0.00	Non-allergic
Group 2 (Asymptomatic)	6	32±10	50/50	214.40 ± 234.77	1.48 ± 1.36	290.00 ± 310.00	2.83 ± 3.20	Subclinical
Group 3 (Allergic but not to olive pollen)	6	27±16	50/50	121.72 ± 95.33	0.31 ± 0.04	141.58 ± 106.20	0.58 ± 0.53	16.6% Anaphylaxis 16.6% Insect sting reaction 33.3% Asthma/Rhinitis 33.3% Rhinitis
Group 4 (Untreated allergic to olive pollen)	6	33±11	83/17	162.67 ± 48.97*	20.09 ± 10.91**	203.83 ± 84.37*	50.62 ± 36.38**	100% Asthma/Rhinitis
Group 5 (Allergic to olive pollen under specific immunotherapy)	6	28±7	83/17	527.34 ± 756.35	30.40 ± 41.05	534.56 ± 607.95	38.84 ± 39.56	75% Asthma/Rhinitis 25% Rhinitis

*Statistical significant differences ($p < 0.05$) compared with G1.

**Statistical significant differences ($p < 0.05$) compared with G1, G2 and G3.

controls (Fig. 3D), with a sensitivity of 90%, and a specificity of 77.7%.

These data are very challenging because they open the possibility of discriminating, with a unique genes set, between five different clinical situations.

These 35 differential genes were reanalyzed from a functional point of view using the DAVID program. This analysis of gene set enrichment according to functional pathways and biological processes, showed seven main functional annotation clustering's (Table III): cell activation, cell adhesion, inflammatory response, apoptosis, MAPK signalling pathway, neuroactive ligand-receptor interaction, ubiquitin mediated proteolysis, and a final group classified as miscellaneous, with two genes (CRTAP and TAGAP) that have never before been related to allergic disorders.

DISCUSSION

This report supports the gene-expression analysis relevance for defining molecular profiles

able to discriminate different clinical situations and describes new elements implicated in complex disorders such as allergic diseases.

Firstly, our results have demonstrated by PCA, an unsupervised method, the potential to classify five different clinical situations by analyzing gene expression distribution. Although PCA is one of the first steps in microarray analysis, it is not usual to find such as clear-cut result in five clinically different situations, thus illustrating how proper clinical classifications may be validated by gene-expression clustering. In addition, this point highlights the accuracy of the clinical classification of our population and increases the likelihood that differences among the groups will be found at a genomic level.

On the other hand, besides finding differential pathways and biological processes among the clinical conditions studied, related and not directly related with immune allergic response (Fig. 2), we have found how the analysis of 35 genes makes it possible to define a molecular profile able to discriminate

Table II. Summary of selected genes and validation assay conditions.

LocusLink Gene Symbol	LocusLink Gene Name	Assay ID	LocusLink Gene Symbol	LocusLink Gene Name	Assay ID	LocusLink Gene Symbol	LocusLink Gene Name	Assay ID
ADAM32, Hcg1810775	ADAM metalloproteinase domain 32	Hs00376607_m1	CRHR1, hCG28513	Corticotropin releasing hormone receptor 1	Hs00366363_m1	IFNG, hCG15987	Interferon, gamma,	Hs00989291_m1
ADRA1A, hCG15476	Adrenergic, alpha-1A-	Hs00169124_m1	CRTAP, hCG1996245	Cartilage associated protein, Gene	Hs01035151_m1	IL10, hCG22208	Interleukin 10	Hs00961622_m1
ADRB1, hCG39839	Adrenergic, beta-1-Receptor	Hs02330048_s1	CUL4B, hCG40765	Cullin 4B	Hs00186086_m1	IL13, hCG24111	Interleukin 13	Hs99999038_m1
AKT1, hCG96740	V-akt murine thymoma viral oncogene homolog 1	Hs00178289_m1	EPC1, hCG18888	Enhancer of polycomb homolog (Drosophila)	Hs00228677_m1	IL13RA1, hCG1741666	Interleukin 13 receptor, alpha 1	Hs006069817_m1
ALOX5, hCG18416	Arachidonate 5-lipoxygenase	Hs00167536_m1	FCER1A, hCG39426	Fc fragment of IgE, high affinity I, receptor for; alpha polypeptide	Hs00175232_m1	IL17A, hCG20863	Interleukin 17A	Hs00174383_m1
APAF1, hCG32373	Apoptotic peptidase activating factor 1	Hs00559441_m1	FOX P3, hCG19830	Forkhead box P3	Hs01085830_m1	IL4, hCG24113	Interleukin 4	Hs00174122_m1
BAX, hCG1811614	BCL2-associated X protein	Hs00180269_m1	FPR3, hCG91958	Formyl peptide receptor 3	Hs01574392_m1	IL4R, hCG1984906	Interleukin 4 receptor	Hs00166237_m1
BDKRB1, hCG24512	Bradykinin receptor B1	Hs00664201_s1	GPR83, hCG20931	G protein-coupled receptor 83	Hs00173906_m1	IL5, hCG24116	Interleukin 5 (colony-stimulating factor, eosinophil)	Hs99999031_m1
B2M	Beta-2-microglobulin	Hs99999907_m1	GRIK1, hCG401170	Glutamate receptor, Ionotropic, kainate 1	Hs00168165_m1	IQUB	IQ motif and ubiquitin domain containing	Hs00401552_m1
C3AR1, hCG22220	Complement component 3a receptor 1	Hs00377780_m1	GRM5, hCG2043000	Glutamate receptor, metabotropic 5	Hs00168275_m1	ITGAL, hCG18670	Integrin, alpha L (antigen CD11A (p180), lymphocyte function-associated antigen 1; alpha polypeptide)	Hs00158218_m1
CD247, hCG1769040	CD247 molecule, Gene	Hs00167901_m1	GZMH, hCG1639886	Granzyme H (cathepsin G-like 2, protein h-CCPX)	Hs00277212_m1	ITGB7, hCG43756	Integrin, beta 7	Hs01565750_m1
CD40, hCG40016	CD40 molecule, TNF receptor superfamily member 5	Hs00374176_m1	HLA-DPB1, hCG17514, hCG2043389	Major histocompatibility complex, class II, DP beta 1	Hs03045105_m1	ITGB8, hCG37311	Integrin, beta 8	Hs0110394_m1
CD48, hCG20849	CD48 molecule	Hs00381156_m1	HLA-DQB1	Major histocompatibility complex, class II, DQ beta 1	Hs03054971_m1	KIR3DL1, hCG2039725	Killer cell immunoglobulin-like receptor, three domains, long cytoplasmic tail, 1	Hs01592437_m1
CD86, hCG14759	CD86 molecule	Hs99999104_m1	HLA-DRB1	Major histocompatibility complex, class II, DR beta 1	Hs99999917_m1	LCK, hCG41609	Lymphocyte-specific protein tyrosine kinase	Hs00178427_m1
CDH2, hCG22518	Cadherin 2, type 1, N-cadherin (neuronal)	Hs00983062_m1	HRH4, hCG38323	Histamine receptor H4	Hs00222094_m1	LCP2, hCG37753	Lymphocyte cytosolic protein 2 (SH2 domain containing leukocyte protein of 76kDa)	Hs00175501_m1
CLEC2B, hCG38144	C-type lectin domain family 2, member B	Hs00192860_m1	IFNB1, hCG28967	Interferon, beta 1, fibroblast	Hs01077958_s1	LEPR, hCG20174	Leptin receptor	Hs00174497_m1
LocusLink Gene Symbol	LocusLink Gene Name	Assay ID	LocusLink Gene Symbol	LocusLink Gene Name	Assay ID	LocusLink Gene Symbol	LocusLink Gene Name	Assay ID
LYN, hCG1810958	V-yes-1 Yamauchi sarcoma viral related oncogene homolog	Hs00176719_m1	PRKCB, hCG1810821	Protein kinase C, beta	Hs00176998_m1	TAC1, hCG16982	Tachykinin, precursor 1	Hs00243225_m1
MAPK11, hCG2039436	Mitogen-activated protein kinase 11	Hs00177101_m1	PRL, hCG36753	Prolactin, Gene	Hs00168730_m1	TAGAP, hCG2029776	T-cell activation RhoGTPase activating protein	Hs00299284_m1
MAPK13, hCG15774	Mitogen-activated protein kinase 13	Hs00234085_m1	PTGER2, hCG2013357	Prostaglandin E receptor 2 (subtype EP2), 53kDa	Hs00168754_m1	TAPBP, hCG2029776	TAP binding protein (tapasin)	Hs00175269_m1
MCAM, hCG15250	Melanoma cell adhesion molecule	Hs00174838_m1	PTGER3, hCG1811960	Prostaglandin E receptor 3 (subtype EP3)	Hs00168755_m1	TBXA2R, hCG23625	Thromboxane A2 receptor	Hs00169054_m1
MSR1, hCG27149	Macrophage scavenger receptor 1	Hs00234007_m1	PTPRC, hCG1811178	Protein tyrosine phosphatase, receptor type, C	Hs00365634_g1	TGFB1, hCG22321	Transforming growth factor, beta 1.	Hs00998133_m1
NCR2, hCG2000284	Natural cytotoxicity triggering receptor 2	Hs00183113_m1	PTX3, hCG26914	Pentraxin-related gene, rapidly induced by IL-1 beta	Hs00173615_m1	THRB, hCG15525	Thyroid hormone receptor, beta (erythroblastic leukemia viral (v-erb-a) oncogene homolog 2, avian)	Hs00230861_m1
NFATC1, hCG15098	Nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 1.	Hs00542678_m1	RAET1E, hCG21012	Retinoic acid early transcript 1E	Hs00373622_m1	TLR4, hCG27399	Toll-like receptor 4	Hs00152939_m1
NFATC3, hCG28103	Nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 3	Hs00190046_m1	SIPR5, hCG1644153	Sphingosine-1-phosphate receptor 5	Hs00928195_s1	TOLLIP, hCG40380	Toll interacting protein	Hs01553188_m1
NFIB, hCG1742080	Nuclear factor I/B	Hs01029176_m1	SELL, hCG37088	Selectin L	Hs01046459_m1	TRD@, hCG1794860	T cell receptor delta locus	Hs00413262_g1
NFKB1, hCG39336	Nuclear factor of kappa light polypeptide gene enhancer in B-cells 1	Hs00765730_m1	SKP1, hCG1980915	S-phase kinase-associated protein 1	Hs00749532_s1	TRIM37, hCG1819317	Tripartite motif-containing 37	Hs00248701_m1
NLRP3, hCG1982559	NLR family, pyrin domain containing 3	Hs00918082_m1	SMURF1, hCG41327	SMAD specific E3 ubiquitin protein ligase 1	Hs00410929_m1	VCAN, hCG2015974	Versican,	Hs00171642_m1
PARD3, hCG21413	Par-3 partitioning defective 3 homolog (C. elegans)	Hs00971409_m1	SOS1, hCG1641052	Son of sevenless homolog 1 (Drosophila)	Hs00362308_m1	VNN1, hCG20013	Vanin 1	Hs00190582_m1
PDE1A, hCG2004696	Phosphodiesterase 1A, calmodulin-dependent	Hs00395232_m1	SPN, hCG21624	Sialophorin	Hs00174604_m1	WWP1, hCG2010044	WW domain containing E3 ubiquitin protein ligase 1	Hs00366927_m1
PECAM1, hCG2002097	Platelet/endothelial cell adhesion molecule	Hs00169777_m1	STAT1, hCG25794	Signal transducer and activator of transcription 1, 91kDa	Hs01013996_m1	ZAP70	Homo sapiens zeta-chain (TCR) associated protein kinase 70kDa (ZAP70), transcript variant 2, mRNA	Hs00896347_m1
PRKACA, hCG27521	Protein kinase, cAMP-dependent, catalytic, alpha	Hs00427274_m1	SULF1, hCG18956	Sulfatase 1, Gene hCG18956 Celera Annotation	Hs00290918_m1	18S	Eukaryotic 18S rRNA	Hs99999901_s1
PRKACB, hCG23660	Protein kinase, cAMP-dependent, catalytic, beta	Hs00176944_m1	TAAR2, hCG1786551	trace amine associated receptor 2	Hs00274365_s1			

from a genomic point of view, olive-pollen-allergic from non-allergic subjects. Furthermore, the differential gene-expression of these same genes is able to discern between the five clinical situations,

with a good specificity (Fig. 3), finding the highest sensitivity and specificity for classifying olive-pollen-allergic subjects versus non-allergic control. To our knowledge, this is the first report which

Table III. Functional annotation clustering of the most relevant genes.

Functional Annotation Clustering	Genes
Cell Activation	CD86, CD40, CD48, HLA-DRB1, HLA-DQB1, IL10, NFATC1, LCK, SPN, LYN, ZAP70, ITGAL, BAX.
Cell adhesion molecules	SELL, VCAN, ITGAL, ITGB7, ITGB8, CD40, CD86, HLA-DRB1, HLA-DQB1
Inflammatory response	CD40, ALOX5, IL10, LYN, C3AR1, ITGAL, AKT1
Apoptosis	BAX, APAF1, PRKACA, PRKACB , AKT1, NLPR3, GZMH
MAPK signalling pathway	MAPK13, PRKACA, PRKACB, SOS1 , AKT1
Neuroactive ligand-receptor interaction	ADRB1, C3AR1, FPR3, PTGER2, S1PR5
Ubiquitin mediated proteolysis	SMURF1 , TRIM37
Miscellaneous	CRTAP , TAGAP

Bold type indicates genes that have never before been associated with allergic disorders.

describes a molecular profile to define five clinical situations.

The correlation analysis of these data using specific informatics programs for proteins interactions (Fig. 4) shows a group of proteins around a central core, CD86 and CD40, which are related to immune response. But interestingly there are also some others, which have not been previously associated with this type of responses.

In order to reach a fuller understanding of the results, these 35 differential genes were reanalyzed from a functional point of view. This analysis showed 7 main functional annotations clustering (Table III) according to functional pathways and biological processes (some of the genes are related to more than one functional pathway). Two main clustering, cell activation and cellular adhesion molecules included 13 and 9 genes, respectively. All of these genes are codifying by proteins previously implicated in immune-allergic responses.

CD86, a membrane receptor expressed in antigen-presenting cells and CD40, a member of TNF-receptors super-family, which are implicated in adaptative immune response, antigen presentation, Th1/Th2 differentiation and inflammatory response.

CD48, a glycosyl-phosphatidylinositol-anchored protein, is involved in lymphocyte adhesion, activation, and costimulation, and has been described as a potential target for the suppression of asthma (18). HLA class II antigens (DQB1 and DRB1) are related to antigen presentation and activation, and extensively

associated with allergic diseases. IL-10, a cytokine with anti-inflammatory properties, is considered as one of the main T-cell regulatory cytokines. The transcription factor NFATC1 is involved in the inducible expression of numerous genes as such Th2 cytokines and related genes (19).

Directly related with these genes are LCK, LYN and ZAP70, three tyrosine kinases implicated in the earliest detectable signalling response upon antigen-induced immune-receptor activation in inflammatory cells.

Also, we have found as differentials some other genes related to essential steps of allergic responses: different adhesion molecules such as SELL, which plays an important role in rolling and leukocyte extravasations and is implicated in remodelling; VCAN, extracellular matrix protein implicated in human epithelial-to-mesenchymal transition and cell growth; SPN, which regulates Th2 differentiation and inflammation; ITGAL (integrin, alpha L); ITGB8 (integrin, beta 8) implicated in the mediation of TGF- β in the control of human airways epithelial proliferation (20); and ITGB7 (integrin, beta 7) which plays an important role in cellular adhesion, migration and B lymphocytes homing.

The next significant functional clustering (Table III) were related with inflammatory response, apoptosis, MAPK signalling pathway, and neuroactive ligand-receptor interaction. These clustering include ALOX5 or 5- Lipxygenase, an enzyme that catalyzes the production of leukotrienes

from arachidonic acid and has been associated with asthma and allergic diseases. Regarding C3AR1, receptor for complement C3 protein, has been linked to airway hyperresponsiveness and recently, has been described as a potential biomarker of acute respiratory allergen exposure (21). Several genes of protein kinases such as PRKACA (protein kinase, cAMP-dependent, catalytic, alpha) that may contribute to asthma pathogenesis by regulation of matrix metalloproteinase-9 associated with airway remodelling, and aid in leukotriene D4 signalling (22) and PRKACB that encodes the catalytic subunit beta of the protein kinase A and never before has been related with allergic diseases.

Other kinases associated were the AKT1, protein serine-threonine-kinase, involved in NF-Kappa Beta-pathway, with anti-apoptotic activity and associated with toll-like receptor signalling and inflammatory response; and MAPK13, a member of the MAP kinase family, p38 δ isoform. MAP kinases are involved in a wide variety of cellular processes such as proliferation, differentiation, transcription and regulation. Related with MAPK we have found SOS1. Few studies relate SOS1 to human disease, but recently it has been reported that this gene was induced by dioxin, an environmental toxin, pointing to the possibility that SOS1 gene expression alterations may affect toxicological endpoints and lead to clinical disease (23). Interestingly, no study has yet dealt with allergy diseases.

Other interesting genes were related with apoptosis: BAX. The protein codified by this gene has been implicated in apoptotic processes of cells from the respiratory tracts of asthmatic patients (24) and APAF1, described as related with pathophysiology of lung diseases. NLRP3, NOD-like receptor (NLR) signalling (inflammasome), related to apoptosis and innate response, has recently been reported as critical in driving lung inflammation (25) and may represent an interesting therapeutic target in allergy. Finally, in this cluster we have found GZMH a serine protease that is highly expressed in NK cells and suggested as implicated in triggering an alternative cell-death pathway in innate immunity. Our finding is the first association described with allergic diseases.

One of the most surprising differential functional cluster and, the main down-expressed pathway, was related to neuroactive-ligand receptor interaction.

Here we have found two predictable, but interesting genes, ADRB1 (adrenergic beta-1- receptor) and PTGER2 (E-prostanoid 2 receptor), both related with asthma treatment. Another differential gene found was FPR3, N-formyl peptide receptor, an innate immunity receptor that could be implicated in nasal tissue remodelling (26). On the other hand, other interesting gene is S1PR5, a G protein-coupled receptor, expressed mainly in dendritic and NK cells and is required for movement of NK to inflammation sites. Antagonism of S1PRs abrogates experimental asthma, and it also suppresses the development of allergic conjunctivitis in mice (27).

The next functional clustering was Ubiquitin mediated proteolysis and it included two interesting genes: SMURF1 (Smad-ubiquitin regulatory factor 1) that codify by an E3 ubiquitin ligase that regulates signalling of the TGF- β Superfamily and never before has been associated with allergic diseases, and TRIM37, a gene which encodes a member of the tripartite motif (TRIM) family, involved in diverse cellular functions, and specifically TRIM37 has been reported very recently as useful biomarker for asthma (28).

Finally, we have found two genes that have never before been associated with allergic diseases and, were classified as miscellaneous: CRTAP (cartilage-associated protein) related with collagen and TAGAP, a gene expressed at high levels in the immune system, including B and T lymphocytes. The protein product of this gene is predicted to function as a Rho-GTPase-activating protein, but although little is known about their function in the immune system, it has been associated with the risk of three immunological autoimmune disorders, celiac disease (29), Type 1 of diabetes (29) and very recently, with rheumatoid arthritis (30).

In summary, in spite of the complexity of the results and, the need of going deeper into their biological relevance, to analyze the protein expression, to increase the number of subjects and to confirm these results in some sorted cell population, this study shows how, a good clinical classification permits the clustering of five clinical allergic conditions, according to gene-expression. We have also been able to define, using 35 genes, a genetic profile that discriminates the five clinical situations with a specificity of 87.5 to 77.7% (Fig. 3). Finally

and most importantly, we have not only validated several gene/pathways that play an important role in allergy pathogenesis, but we have also identified novel genes/pathways which should be relevant in this and other allergic diseases. This is the first study of this kind, performed in an allergic population sensitized to *O. europaea*. We have found potential biomarkers and/or therapeutic targets (validated in two different populations) which should be very important in the diagnosis and treatment of allergy diseases in the future.

ACKNOWLEDGEMENTS

We are grateful to the F.I.S. FJD-UTE (CP05/0183), "Incorporación de Investigadores al SNS" Program, for financial support, to Drs J. Delgado and A. Miranda for providing patients for the study, to Oliver Shaw for editing the manuscript in English and, to J. Cantalejo and R. Ramos from Scientific Park of Madrid for their technical help.

The funding sources for this paper are supported in part by research grants CP05/00183, PI07/0024, PI10/0025, CIBERES (ISCIII, 0013) and RETIC (RD09/0076/00101) from the Fondo de Investigación Sanitaria (Ministerio de Sanidad y Consumo, Spain), Fundación Ramón Areces 2007 (5236/001) and SEAIC (Research grant 2007). M. Aguerri and D. Calzada are supported by Fundación Conchita Rábago, Madrid, Spain.

REFERENCES

- Holloway JW, Yang IA, Holgate ST. Genetics of allergic disease. *J Allergy Clin Immunol* 2010; 125:(S)81-94.
- Vercelli D. Gene-environment interactions in asthma and allergy: the end of the beginning? *Curr Opin Allergy Immunol* 2010; 10:145-8.
- Cárdaba B, Aguerri M, Calzada D, Lahoz C. New challenges for old diseases: The impact of -OMICS technologies in the understanding of allergic diseases. In: C. Pereira, editor. *Allergic Diseases - Highlights in the Clinic, Mechanisms and Treatment*. INTECH Open Access Publishers; 2012; pp3-30.
- Hansel NN, Cheadle C, Diette GB, Wright J, Thompson KM, Barnes KC, Georas SN. Analysis of CD4+ T-cell gene expression in allergic subjects using two different microarray platforms. *Allergy* 2008; 63:366-9.
- Woodruff PG, Boushey HA, Dolganov GM, et al. Genome-wide profiling identifies epithelial cell genes associated with asthma and with treatment response to corticosteroids. *Proc Natl Acad Sci USA* 2007; 104:15858-63.
- Hakonarson H, Bjornsdottir US, Halapi E, et al. Profiling of genes expressed in peripheral blood mononuclear cells predicts glucocorticoid sensitivity in asthma patients. *Proc Natl Acad Sci USA* 2005; 102:14789-94.
- Woodruff PG, Modrek B, Choy DF, et al. T-helper type 2-driven inflammation defines major subphenotypes of asthma. *Am J Respir Crit Care Med* 2009; 180:388-95.
- Mobini R, Andersson BA, Erjefält J, Hahn-Zoric M, Langston MA, Perkins AD, Cardell LO, Benson M. A module-based analytical strategy to identify novel disease-associated genes shows an inhibitory role for interleukin 7 Receptor in allergic inflammation. *BMC Syst Biol* 2009; 3:19.
- Cárdaba B, Cortegano I, Florido F, et al. Update in the understanding of genetic predisposition to olive pollen sensitization. *Allergy* 2002; 57(S)71:41-6.
- Quiralte J, Llanes E, Barral P, et al. Ole e 2 and Ole e 10: new clinical aspects and genetic restrictions in olive pollen allergy. *Allergy* 2005; 60:360-5.
- Cárdaba B, Llanes E, Chacártégui M, et al. Modulation of allergic response by gene-environment interaction: olive pollen allergy. *J Invest Allergy Clin Immunol* 2007; 17:31-5.
- Llanes E, Quiralte J, López E, et al. Analysis of polymorphisms in olive pollen allergy: IL13, IL4RA, IL5 and β_2 Adrenergic Receptor genes. *Int Arch Allergy Immunol* 2009; 148:228-38.
- Aguerrri M, Calzada D, Martín E, et al. FOXP3 and TGF- β : Differential regulatory molecules between sensitization and tolerance to olive pollen. *Eur J Inflammation* 2012; 10:193-202.
- Kendzioriski C, Irizarry RA, Chen KS, Haag JD, Gould MN. On the utility of pooling biological samples in microarray experiments. *Proc Natl Acad Sci USA* 2005; 102:4252-257.
- Medina I, Carbonell J, Pulido J et al. Babelomics: an integrative platform for the analysis of

- transcriptomics, proteomics and genomic data with advanced functional profiling. *Nucl Acids Res* 2010; 38:W210-13.
16. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A practical approach to multiple testing. *J Royal Statistic Society Series B (Method)* 1995; 57:289-300.
 17. Montaner D, Dopazo J. Multidimensional Gene Set Analysis of Genomic Data. *PLoS ONE* 2010; 5:e10348.
 18. Munitz A, Bachelet I, Finkelman FD, Rothenberg ME, Levi-Schaffer F. CD48 is critically involved in allergic eosinophilic airway inflammation. *Am J Respir Crit Care Med* 2007; 175:911-8.
 19. Sumpter TL, Payne KK, Wilkes DS. Regulation of the NFAT pathway discriminates CD4+CD25+ regulatory T cells from CD4+CD25- helper T cells. *J Leukoc Biol* 2008; 83:708-17.
 20. Fjellbirkeland L, Cambier S, Broaddus VC, Hill A, Brunetta P, Dolganov G. Integrin α v β 8-mediated activation of transforming growth factor- β inhibits human airway epithelial proliferation in intact bronchial tissue. *Am J Pathol* 2003; 163:533-42.
 21. Pucheu-Haston CM, Copeland LB, Vallanat B, Boykin E, Ward MD. Biomarkers of acute respiratory allergen exposure: screening for sensitization potential. *Toxicol Appl Pharmacol* 2010; 244:144-55.
 22. Hoshino M, Izumi T, Shimizu T. Leukotriene D4 activates mitogen-activated protein kinase through a protein kinase C α -Raf-1-dependent pathway in human monocytic leukemia THP-1 cells. *J Biol Chem* 1998; 273:4878-82.
 23. Pierre S, Bats AS, Coumoul X. Understanding SOS (Son of Sevenless). *Biochem Pharmacol* 2011; 82:1049-56.
 24. Abdulmir AS, Kadhim HS, Hafidh RR, Ali MA, Faik I, Abubakar F, Abbas KA. Severity of asthma: the role of CD25+, CD30+, NF-kappaB, and apoptotic markers. *J Investig Allergol Clin Immunol* 2009; 19:218-24.
 25. Besnard AG, Guillou N, Tschopp J, Erard F, Couillin I, Iwakura Y, Quesniaux V, Ryffel B, Togbe D. NLRP3 inflammasome is required in murine asthma in the absence of aluminum adjuvant. *Allergy* 2011; 66:1047-57.
 26. Prevete N, Salzano FA, Rossi FW, et al. Role(s) of formyl-peptide receptors expressed in nasal epithelial cells. *J Biol Regul Homeost Agents* 2011; 25:553-64.
 27. Sumi T, Fukushima A, Nishino K, Fukuda K, Kumagai N, Nishida T, Ueno H. Treatment with FTY720 during the induction or effector phase suppresses the development of experimental allergic conjunctivitis in mice. *Cell Biol Int* 2009; 33:534-41.
 28. Shin SW, Oh TJ, Park SM, et al. Asthma-predictive genetic markers in gene expression profiling of peripheral blood mononuclear cells. *Allergy Asthma Immunol Res* 2011; 3:265-72.
 29. Smyth DJ, Plagnol V, Walker NM, et al. Shared and distinct genetic variants in type 1 diabetes and celiac disease. *N Engl J Med* 2008; 359:2767-77.
 30. Chen R, Stahl EA, Kurreeman FA, et al. Fine mapping the TAGAP risk locus in rheumatoid arthritis. *Genes Immun* 2011; 12:314-8.