

ORIGINAL ARTICLE

Basic and Translational Allergy Immunology

Understanding uncontrolled severe allergic asthma by integration of omic and clinical data

María Isabel Delgado-Dolset^{1,2}  | David Obeso^{1,2}  | Juan Rodríguez-Coira^{1,2,3}  | Carlos Tarín¹ | Ge Tan³  | José A. Cumplido⁴ | Ana Cabrera⁴ | Santiago Angulo⁵  | Coral Barbas²  | Milena Sokolowska³  | Domingo Barber¹  | Teresa Carrillo⁴ | Alma Villaseñor¹  | María M. Escribese¹ 

¹Institute of Applied Molecular Medicine (IMMA), Department of Basic Medical Sciences, Facultad de Medicina, Universidad San Pablo CEU, CEU Universities, Urbanización Montepríncipe, Madrid, Spain

²Centre for Metabolomics and Bioanalysis (CEMBIO), Department of Chemistry and Biochemistry, Facultad de Farmacia, Universidad San Pablo CEU, CEU Universities, Urbanización Montepríncipe, Madrid, Spain

³Swiss Institute of Allergy and Asthma Research (SIAF), University of Zurich, Zurich, Switzerland

⁴Hospital Universitario de Gran Canaria Doctor Negrin, Las Palmas de Gran Canaria, Spain

⁵Department of Applied Mathematics and Statistics, Universidad San Pablo-CEU, CEU Universities, Madrid, Spain

Correspondence

María M. Escribese, Department of Basic Medical Sciences, Faculty of Medicine, San Pablo CEU University, Campus Montepríncipe, Crtra, Boadilla del Monte km 5.3, CP 28668 Boadilla del Monte, Madrid, Spain.
Email: mariamarta.escribese@ceu.es

Funding information

Ministerio de Ciencia, Innovación y Universidades, Grant/Award Number: PCI2018-092930; ISCIII, Grant/Award Number: PI18/01467 and PI19/00044; Ministerio de Ciencia, Innovación y Universidades, Grant/Award Number: RTI2018-095166-B-I00; Swiss National Science Foundation (SNFS), Grant/Award Number: 310030_189334/1

Abstract

Background: Asthma is a complex, multifactorial disease often linked with sensitization to house dust mites (HDM). There is a subset of patients that does not respond to available treatments, who present a higher number of exacerbations and a worse quality of life. To understand the mechanisms of poor asthma control and disease severity, we aim to elucidate the metabolic and immunologic routes underlying this specific phenotype and the associated clinical features.

Methods: Eighty-seven patients with a clinical history of asthma were recruited and stratified in 4 groups according to their response to treatment: corticosteroid-controlled (ICS), immunotherapy-controlled (IT), biologicals-controlled (BIO) or uncontrolled (UC). Serum samples were analysed by metabolomics and proteomics; and classifiers were built using machine-learning algorithms.

Results: Metabolomic analysis showed that ICS and UC groups cluster separately from one another and display the highest number of significantly different metabolites among all comparisons. Metabolite identification and pathway enrichment analysis highlighted increased levels of lysophospholipids related to inflammatory pathways in the UC patients. Likewise, 8 proteins were either upregulated (CCL13, ARG1, IL15 and TNFRSF12A) or downregulated (sCD4, CCL19 and IFN γ) in UC patients compared to ICS, suggesting a significant activation of T cells in these patients. Finally, the machine-learning model built including metabolomic and clinical data was able to classify the patients with an 87.5% accuracy.

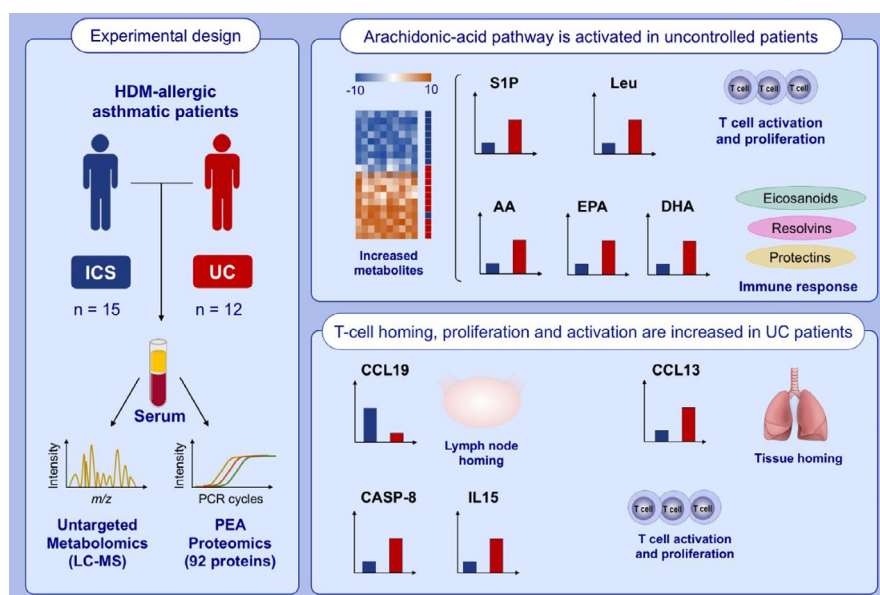
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2021 The Authors. *Allergy* published by European Academy of Allergy and Clinical Immunology and John Wiley & Sons Ltd.

Conclusions: UC patients display a unique fingerprint characterized by inflammatory-related metabolites and proteins, suggesting a pro-inflammatory environment. Moreover, the integration of clinical and experimental data led to a deeper understanding of the mechanisms underlying UC phenotype.

KEYWORDS

allergy, asthma, machine learning, metabolomics, proteomics



GRAPHICAL ABSTRACT

Severe uncontrolled HDM-allergic asthma (UC) displays an increased T-cell activation and proliferation (IL-15, CASP-8, S1P, Leu) and an increased T-cell tissue recruitment (CCL13) compared to corticosteroid-controlled HDM-allergic asthma (ICS). UC shows an exacerbated inflammatory response with increased levels of inflammatory mediators (AA, EPA, DHA, ...). Integration of clinical and metabolomic data is the best strategy to stratify patients by severity.

1 | INTRODUCTION

Asthma is a heterogeneous disease characterized by chronic airway inflammation and a clinical history of respiratory symptoms such as wheeze, shortness of breath, chest tightness and cough that vary over time and in intensity, together with variable expiratory airflow limitation.¹

Asthma can have multiple phenotypes and endotypes, depending on the clinical features, histopathology, genetics or response to treatment, along with other characteristics.² We can distinguish between allergic, intrinsic, neutrophilic, aspirin-intolerant and extensive remodelling asthma,³ among others. Patients with allergic asthma are usually sensitized to house dust mites (HDM). In fact, IgE titres to HDM correlate with exacerbations in asthma.⁴

HDM are the most relevant inducers of allergic diseases worldwide.⁵ HDM allergens include cysteine-proteases (Der p 1, Blo t 1) and chitin-associated proteins (Der p 23), involved in epithelial disruption and remodelling; proteins able to activate TLR4 (Der p 2, Eur

m 2), and lipid-binding proteins (Der p 5, Der p 7 and Der p 21), linked with the establishment of the Th2 response.⁵ The Canary Islands in Spain is one of the regions in Europe with higher mite exposure levels and sensitization prevalence to HDM due to its almost tropical climate of high humidity and warm temperatures.^{6,7} These features are common with other high-exposure HDM areas like Singapore or New Zealand.⁵ Additionally, the exposome in the Canary Islands has been deeply studied⁸; and it is very stable along the island, which makes it a perfect model to minimize variability in the study.

Due to the complexity of asthma, effective treatment is also a challenge. The GINA guideline sets a step-by-step approach, in which treatment with increasing doses and add-on medications are given to patients. For allergic asthma, therapeutic strategies include corticosteroids, bronchodilators, immunotherapy and biological drugs (i.e., anti-IgE or anti-IL5).¹

Classification of asthmatic allergic patients is a difficult task. Several endotypes and phenotypes have been described. According to their response to treatment, patients have severe asthma if they

either needed treatment from steps 4–5 of the GINA,¹ or 5–6 of the GEMA⁹ guidelines (which includes inhaled corticosteroids coupled to long term acting beta-agonists (inhaled CS/LABA) in high doses + antileukotrienes (LTRA) + theophylline (T)) or had to take systemic corticosteroids to achieve control in the previous year.

However, despite having received the treatment previously mentioned, around 3%–10% of patients display severe uncontrolled asthma. These patients often present comorbidities such as obesity, nasal polyposis, allergic rhinitis or food allergy; resulting in a worse quality of life and higher mortality rates.^{1,10}

Particularly, in the Canary Islands several studies report frequent respiratory problems in children and high rates of asthma prevalence (69.6%). Moreover, 16% of the asthmatic patients have severe uncontrolled asthma,¹¹ which is similar to other European countries, such as the United Kingdom where it is close to 18%.¹²

Regardless of this significant prevalence, there is no available effective treatment for these patients. Their treatment options rely on the development of new personalized medicine strategies.

In this study, all the patients were included by the Allergy Service at Hospital Universitario de Gran Canaria Dr Negrin, a group with deep knowledge on asthma management that works in an area with high HDM exposure and with a well-known exposome, resulting in a very homogenous cohort of asthmatic patients stratified according to their response to treatment. This is a unique clinical model to study a complex disease such as allergic asthma.

Here, we used untargeted metabolomics and a bioinformatic analysis, in addition to a proteomics approach, to generate a specific fingerprint associated with uncontrolled allergic asthma that conforms a set of biomarkers with a potential use to develop the personalized intervention strategies that these patients need.

2 | METHODS

2.1 | Patients

Patient recruitment was conducted within one full year. Eighty-seven patients with a clinical history of asthma, classified by the GINA guideline¹ and with a positive skin-prick test to HDM were enrolled in the Allergy Department of Hospital Universitario de Gran Canaria Dr Negrin (Las Palmas de Gran Canaria, Spain).

Recruited patients were characterized by a very complete clinical history, since they were monitored for at least 5 years prior to the study. Detailed information about patient sensitization profiling and clinical characteristics can be found in Supplementary. All patients signed informed consent and the study was approved by the Ethics Committee of the hospital on the 4th/February/2016 (code: 160009). Serum samples were obtained as described in Supplementary.

Clinical characteristics were compared between the 4 groups. For continuous variables, Shapiro-Wilk test was applied to assess normality of the data. Age and BMI passed the normality test, so we performed ANOVA test with Tukey's post hoc. Onset age and total

IgE, however, did not pass the Shapiro-Wilk test, and were therefore analysed using Kruskal-Wallis with Dunn's post hoc. For categorical variables, chi-square analysis was performed, and for groups with data lower than 5, Fisher's exact was used.

2.2 | Metabolomic analysis

Serum samples were prepared and measured in batches, in a randomized order using a liquid chromatography coupled to mass spectrometry (LC-MS) with a quadrupole-time of flight (Q-TOF) analyser (Agilent series 6520). The experiment, as previously described,¹³ was measured in electrospray ionization in positive and negative modes (ESI+ and ESI-, respectively). A quality control (QC) sample was prepared by mixing equal volumes of a representative set of samples and was analysed throughout the analyses to ensure instrumental reproducibility. Metabolite annotation was performed using CEU Mass Mediator 3.0 online data base¹⁴ and confirmed through tandem mass experiments (MS/MS) with a fragmentation energy of 20 eV. Full description of sample preparation, instrumental characteristics, data quality, data treatment, statistical analysis and metabolite annotation are stated in Supplementary.

2.3 | Proteomic analysis

Serum proteins were analysed by Proximity Extension Assay (PEA) (Olink®, Uppsala). PEA is based on the probes linked to paired antibodies recognizing the same protein. Upon the antibodies binding to the target, the primers hybridize and create a PCR product, which is then amplified as the read-out, giving the Normalized Protein Expression (NPX) as result. For this assay, 10 µl of serum of corticosteroid-controlled (ICS) ($n = 15$) and uncontrolled (UC) patients ($n = 11$) were analysed by Olink. Therefore, proteomic analysis was not performed in all samples. We used Target 96 Immuno-Oncology panel (95311, Olink®, Uppsala) consisting in 92 proteins involved in processes such as promotion and inhibition of tumour immunity, chemotaxis, vascular & tissue remodelling, apoptosis & cell killing and metabolism & autophagy. Detailed information about quality control, data treatment and statistical analysis can be found in Supplementary.

2.4 | Machine-learning model construction

To gather novel insights about the molecular mechanisms involved in the UC allergic asthma phenotype, we employed multi-omic data analysis. The integration and analysis of omics and clinical data with machine-learning (ML) methods offer novel techniques enabling the discovery of new biomarkers.¹⁵ After data pre-processing and down sampling, clinical (50 variables after the exclusion of medication) and metabolomic (86 identified significant metabolites) data were mixed

to create a ML predictor that could classify ICS and UC patients in separate groups. For the ML model, *caret*, *mlbench*, *sjmisc* and *class* R packages were used for k-Nearest Neighbour (k-NN) algorithm. Proteomic data were not included in this analysis because the samples used for metabolomics and proteomics were not always the same. Full description of the data treatment steps, training and testing can be found in Supplementary.

3 | RESULTS

3.1 | Patient classification

Patients were classified in 4 groups according to the medication that controlled their symptoms: (1) Mild asthma patients controlled with inhaled or topic corticosteroids, and never with systemic corticosteroids (ICS, $n = 15$); (2) mild-moderate asthma controlled with immunotherapy (IT, $n = 44$); (3) moderate-severe asthma patients controlled with omalizumab (anti-IgE) (BIO, $n = 16$); and (4) severe asthma patients that do not respond to any treatment, or

uncontrolled patients (UC, $n = 12$). More information about these groups can be found in Supplementary.

Moreover, their clinical history was thoroughly analysed. There were no differences related to sex, ethnicity, BMI, smoking status, total IgE levels and type of HDM sensitization among the groups ($p > 0.05$) (Table 1). However, age was found to be significantly higher in the UC group than in the ICS.

Looking at the medication, patients from the ICS group were treated with antihistamine (AH) and topical corticosteroids (Topic CS), and most of them (93%) administered inhaled CS/LABA. Three of these patients were prescribed LTRA and/or short-effect bronchodilators (SABA). IT patients had a medication pattern like that of ICS patients, although only about a third of them were taking inhaled CS/LABA, while most of them (75%) were prescribed SABA. BIO patients, on the other hand, had more in common with the UC, although the percentage of patients taking topic CS doubled that of UC patients taking this drug (68.8%). The exact treatment and sensitizations of each patient are shown in Table S1. Regarding the UC, all patients were taking both inhaled CS/LABA and SABA, and most of them (75%) were taking LTRA as well. 50% of these

TABLE 1 Clinical information of the patients

	ICS	IT	BIO	UC
N	15	44	16	12
Age (years)	37.2 ± 2.3	37.3 ± 1.5 [†]	43.3 ± 2.4	48.5 ± 3.6*
Onset Age (years)	15.8 ± 2.9	16.1 ± 1.7	12.1 ± 1.8	11.9 ± 2.7
Gender (%F/%M)	66.7/33.3	77.3/22.7	87.5/12.5	58.3/41.7
BMI	26.7 ± 1.4	26.0 ± 0.6	26.8 ± 1.1	28.0 ± 1.2
Smoker (%)	13.3	0	6.3	0
Ex-smoker (%)	6.7	2.3	6.3	25
Total IgE (U)	483.7 ± 164.7	396.2 ± 64.6	503.9 ± 161.5	536.3 ± 181.7
AC (%)	0	0 [†]	43.8*	50*
AH (%)	100	90.9 [†]	37.5*	41.7*
LTRA (%)	13.3	11.4 [†]	75*	75*
Inhaled CS (%)	0	2.3	0	0
Inhaled CS/LABA (%)	93.3	34.1* [†]	100	100
Oral CS (%)	0	0	0	16.7
Topic CS (%)	100	97.7 [†]	68.8*	33.3*
SABA (%)	20	75*	93.8*	100*
T (%)	0	0	0	8.3
NSAID-HS (%)	13.3	6.8	6.3	16.7
<i>D. pteronyssinus</i> (%)	93.3	97.7	93.8	91.7
<i>D. farinae</i> (%)	93.3	93.2	87.5	91.7
<i>L. destructor</i> (%)	66.7	50	43.8	58.3
<i>B. tropicalis</i> (%)	86.7	77.3	81.3	66.7
<i>A. siro</i> (%)	26.7	20.5	0	8.3
<i>T. putrescentiae</i> (%)	53.3	45.5	56.3	58.3

Abbreviations: AC, anticholinergic; AH, antihistaminic; BMI, Body mass index; CS, corticosteroids; Inhaled CS/LABA, inhaled corticosteroids combined with Long-acting beta-adrenoceptor agonist; LTRA, antileukotriene; NSAID-HS, NSAID-hypersensitivity; SABA, short-acting beta-adrenoceptor agonist; T, theophylline; U, ISAC units.

* $p < .05$ against ICS. [†] $p < .05$ against UC.

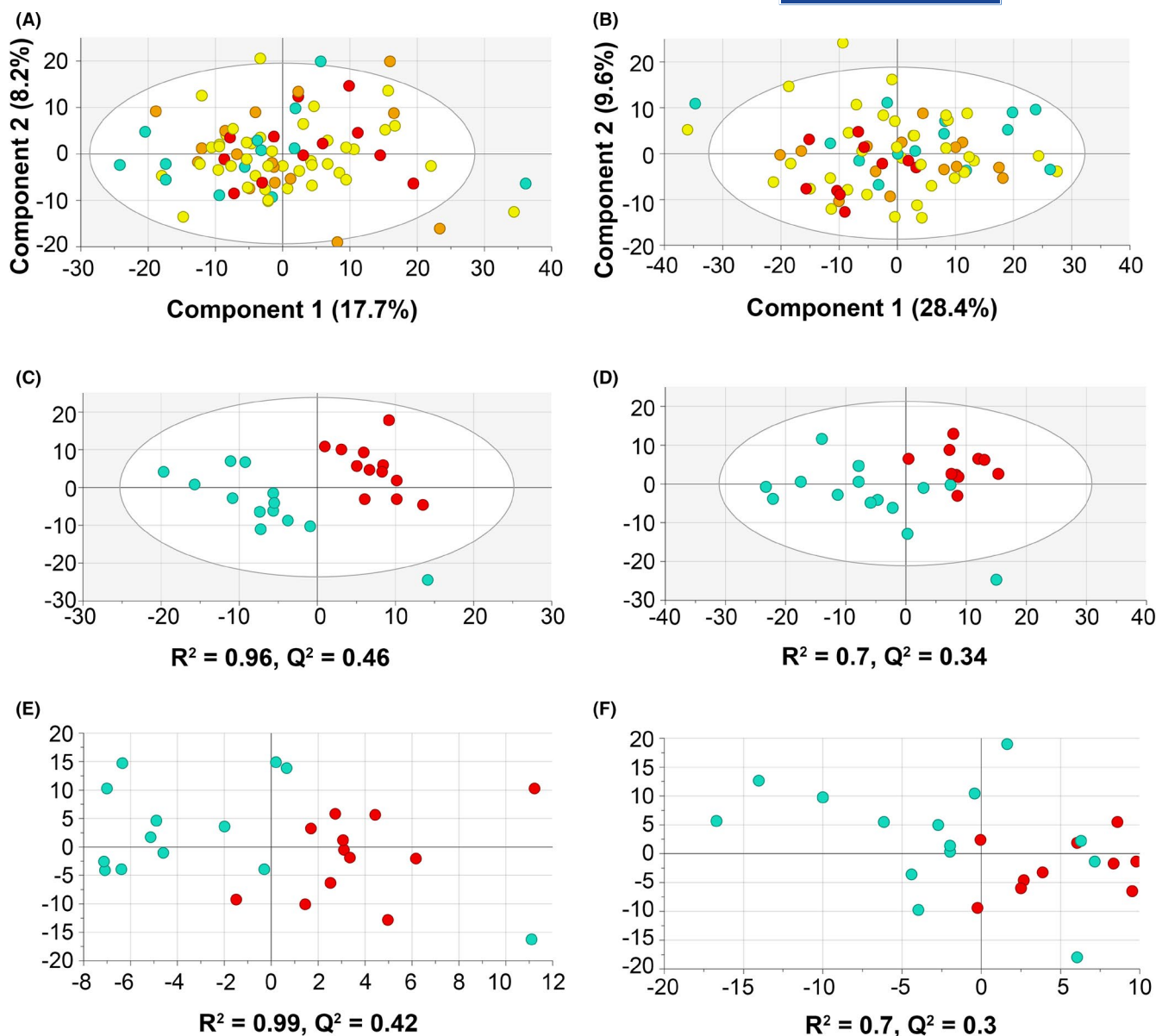


FIGURE 1 Serum metabolomic profile of all patients. A. Unsupervised PCA model of all groups (ICS, IT, Bio, UC) built using the 833 signals for ESI+ that compiled with quality criteria. B. Unsupervised PCA model of all groups built using the 565 signals for ESI- that compiled with quality criteria. C. PLS-DA model for patients controlled with non-systemic corticosteroids (ICS) vs patients with uncontrolled asthma (UC) comparison in LC-MS (ESI+). D. PLS-DA model for ICS vs UC comparison in LC-MS (ESI-). E. OPLS-DA model for ICS vs UC comparison in LC-MS (ESI+). F. OPLS-DA model for ICS vs UC comparison in LC-MS (ESI-). All data were UV scaled. (Key: 'ICS': blue, 'IT': yellow, 'BIO': orange, 'UC': red). R^2 is the capability of the model to classify the samples; Q^2 is the capability of the model to predict the class of a new sample

patients were prescribed with anticholinergics (AC); 41.7%, with AH; and 16.7% were having oral corticosteroid (Oral CS), while 8.3% were having theophylline (T). 33.3% were taking topic CS.

3.2 | Severe uncontrolled asthmatic patients display a unique metabolic fingerprint

The metabolomic signature of each patient was obtained by LC-MS. A total of 833 signals for ESI+ and 565 for ESI- complied with the quality criteria for metabolomics. Clustering of the QC injections in

the non-supervised (PCA) model indicated the high quality of the data, while dispersion of the samples showed the biological variability of the patients (Figure S1).

Patterns of the groups were analysed by PCA models for ESI+ and ESI- modes, respectively (Figure 1A,B). UC patients clustered together, especially in the ESI- mode. Additionally, we performed PCA models to see the differences between the UC group and the others, but only differences between ICS and UC patients were observed (Figure S2).

More restrictive analysis was performed to assess the metabolic differences between ICS and UC patients, the ones that differed most

in terms of asthma control. To evaluate real differences between ICS and UC groups, PLS-DA and OPLS-DA models were performed (Figure 1C–F). The robustness of OPLS-DA models was assessed by a seven-fold cross-validation.¹⁶ The resulting models showed good quality parameters: R^2 (0.70–0.96) for the variance explained, and Q^2 (0.3–0.46) for the prediction capability by the model. The cross-validated OPLS-DA scores plot showed that the percentage of accuracy was of 84% and 75% for ESI+ and ESI– modes (Figure 1D–E), respectively. From these models two ICS patients were observed to be outliers (bottom right in the panels from Figure 1C,F). The outlier patient from the ESI+ mode has the lowest FEV1 and FVC of the group ($\approx 70\%$), while the outlier patient from the ESI– mode has hypersensitivity to NSAIDs drugs. However, we did not find any remarkable reason to exclude these patients from the study.

Furthermore, as age was found significant between ICS and UC groups, PCA models were performed to illustrate their significance (Figure S3). We found that only for the ESI– mode there is a slightly trend of clustering inside ICS group between younger or older than 35 years old.

Together, these results reveal that the UC group is different from the ICS group in terms of their metabolic fingerprint. However, potential differences with IT and BIO might be expected.

3.3 | Severe uncontrolled asthma metabolic fingerprint is characterized by increased levels of lipid mediators

Univariate analysis was carried out to obtain significant metabolic signals in the ICS vs UC comparison, finding 280 with $p < 0.05$ (204 of them had a FDR < 0.2). Using MS/MS fragmentation experiments, we could annotate (identify) 89 of these metabolites. Most of them had a FDR < 0.2 (83 out of 89); which is valid in exploratory studies as it allows to find more potential biomarkers.^{17–19}

Next, aiming to see if the metabolic signals were able to stratify the patients according to their group, we built a heatmap with hierarchical clustering analysis (HCA) (Figure 2) using the 280 statistically significant signals ($p < 0.05$). We found that ICS patients (in light blue) are stratified at the bottom of the heatmap, while UC patients (in red) are located in the top. Accordingly, ICS and UC patients are able to be clustered separately using these 280 signals, as expected. IT and BIO patients did not cluster in any specific part of the heatmap, although most BIO patients seemed to be closer to UC patients. Therefore, we demonstrate significant metabolic features that differentiate ICS and UC and point out specific features for IT and BIO patients, although no significant clustering was demonstrated.

Since age was significant between ICS and UC groups (Table 1 and Figure S2), ANCOVA was used to remove those metabolites that were affected by age. Therefore, we ended up with 86 annotated metabolites that were not affected by age. ANCOVA test, physico-chemical and statistical details of annotated metabolites are shown in Tables S2–S4, respectively. From these, we obtained 33 metabolites that were uniquely identified.

Among the significant annotated metabolites there were amino acids, bile acids, fatty acids, phospholipids, sphingolipids and vitamins; with phospholipids being the greatest group. Changes in the abundance of compounds between ICS and UC are presented in Figure 3 and Figure S4.

Most of these compounds, such as lysophosphocholines (LPC), lysophosphoethanolamines (LPE), lysophosphatidylinositols (LPI), leucine, arginine, arachidonic acid, sphingosine-1-phosphate (S1P) and retinol were increased in the UC group; with the exception of bilirubin, deoxycholic acid (Figure S4) and phosphocholine 16:0/20:5 (Figure 3), which were decreased in the UC group.

We then performed a pathway analysis of the unique annotated metabolites ($n = 33$) to look for the routes in which they are involved (Table S5). We found an altered phospholipase A2 (PLA2) pathway (100% of affected metabolites in this route, whose changes are shown in Figure 4), along with arachidonic acid metabolism (79.81%), lipoprotein remodelling (52%), phospholipid biosynthesis (95.83%), sphingolipid metabolism (100%) and biosynthesis (82.86%), and other lipid-related routes (Table S5). There are also changes in signalling and transport pathways. Interestingly, there is an increase in metabolites related to a deficiency in leukotriene C4 production (91.5%). These routes are significantly enriched in the UC compared to the ICS patients.

Figure 5 represents metabolites from oleic acid and arachidonic acid pathways. We observed that all measured metabolites that participate in these pathways are increased in UC patients. Thus, we conclude that UC patients have an altered lipid metabolism, with increased levels of lipid mediators in serum compared to the ICS group.

3.4 | UC patients have a distinctive expression of inflammatory-related proteins

In addition to the metabolic fingerprint, we studied the protein serum profile in ICS and UC patients. We measured the expression of 92 proteins (Table S6) and found that 8 of them were significantly different between ICS and UC ($p < 0.05$) (Figure 6A). From these, one of them (CASP-8) also had FDR < 0.2 (Table S6). A STRING analysis was performed in order to find relations between the differentially expressed proteins. We were able to link 7 out of the 8 proteins (Figure 6B). We found that CCL13, ARG1, IL15, TNFRSF12A and CASP-8 were increased in the UC patients, whereas sCD4, CCL19 and IFN γ were decreased (Figure 6C).

Finally, we looked for biological processes in which these proteins were involved using gene ontology (GO) pathway enrichment analysis (Table S7). We found that these proteins are linked to immune response including cytokine response, leukocyte differentiation and adhesion, T cell activation, and nitrogen metabolism, among other pathways.

Therefore, proteomic results support that UC patients have an altered inflammatory response, and point out immune cell activation as a possible key mechanism.

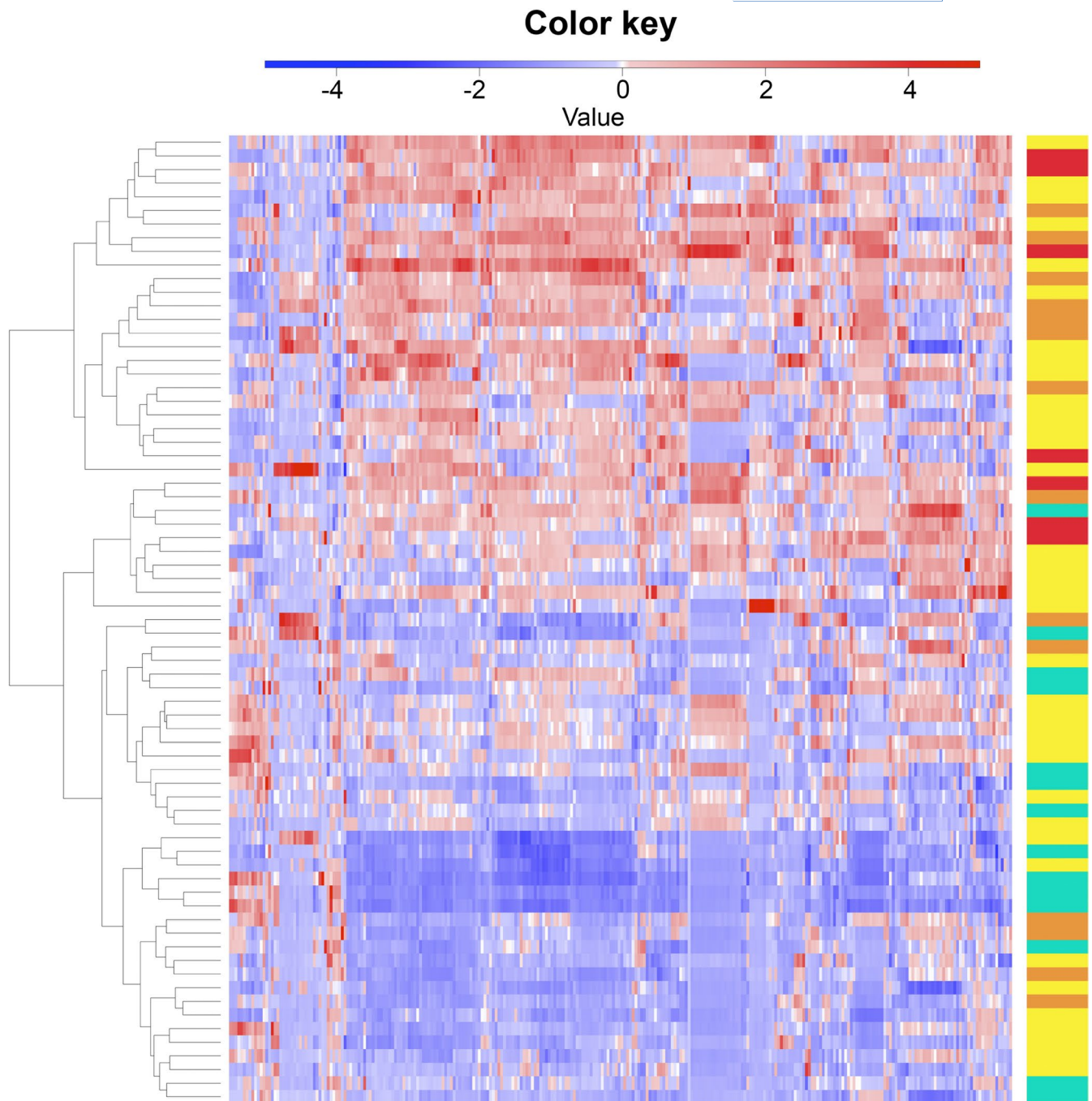


FIGURE 2 Hierarchical clustering heatmap of the patients from the four groups (in rows) and the 280 significant metabolites between ICS and UC (in columns). Red cells represent an increase of the specific metabolite in that sample, whereas blue cells represent decreased metabolites. Samples and metabolites are clustered according to their similarity, and the group they belong to is showed with colours at the right. (Key: 'ICS': light blue; 'IT': yellow; 'BIO': orange; 'UC': red)

3.5 | Machine-learning models for UC classification

Aiming to obtain a more accurate profile of the UC asthmatic phenotype, we integrated both the metabolomic and clinical data in a ML model. Since the proteomic analysis was performed in a limited number of samples, it could not be included in the model. The ML model aim was to find which variables could be used to classify UC and ICS patients. The k-NN model is a ML algorithm that classify a sample

based on all the variables given, which has been tested for breast cancer prediction model.²⁰ The k-NN learning algorithm requires us to specify how many neighbours (k) are going to be considered. To obtain this parameter we used the elbow method. The lack of elbow when using only clinical or metabolomic data alone suggests that these datasets are not good enough to predict classification of UC and ICS patients on their own. When using both datasets combined, the elbow test suggested that the k parameter should be either 2 or

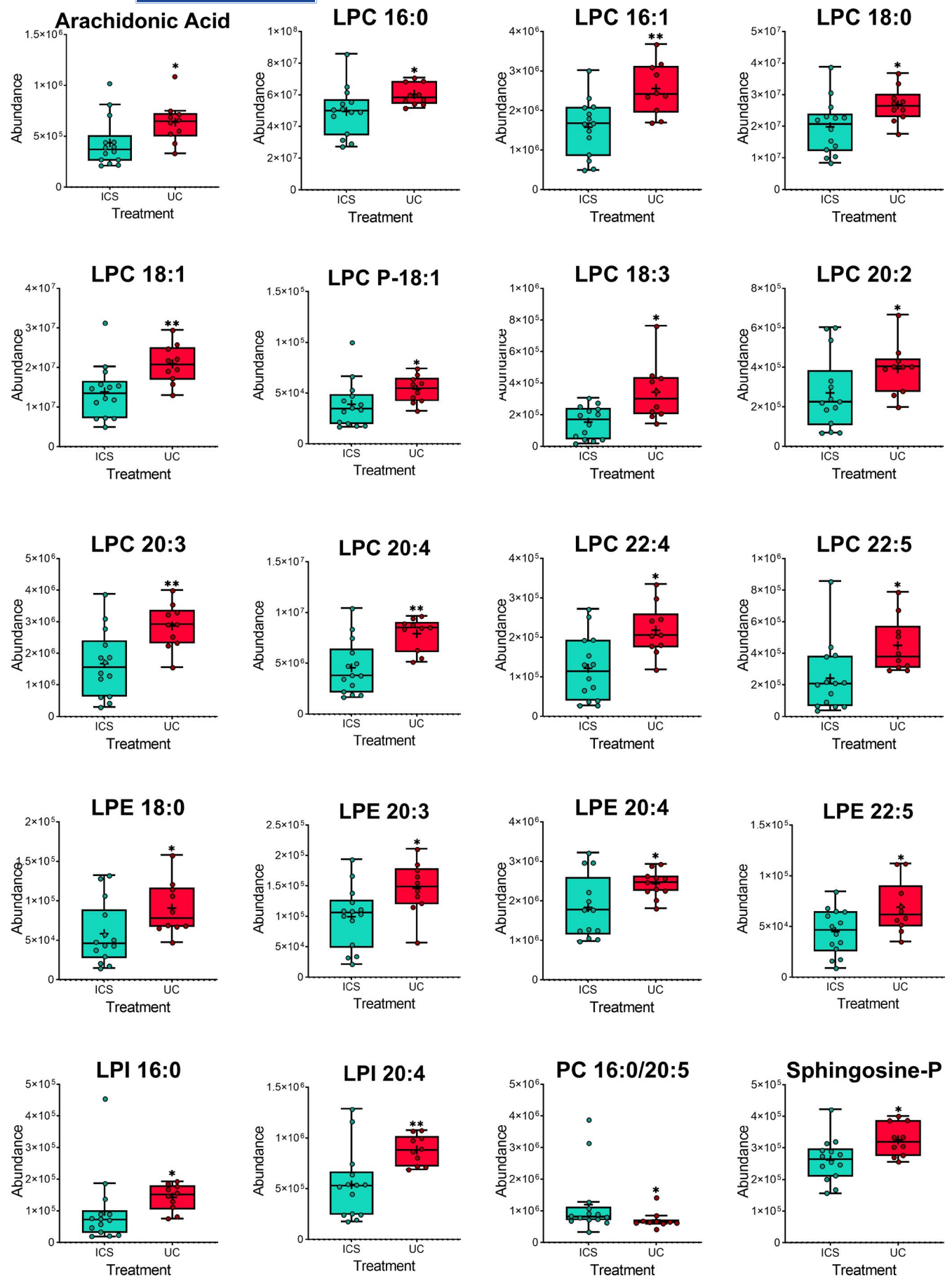
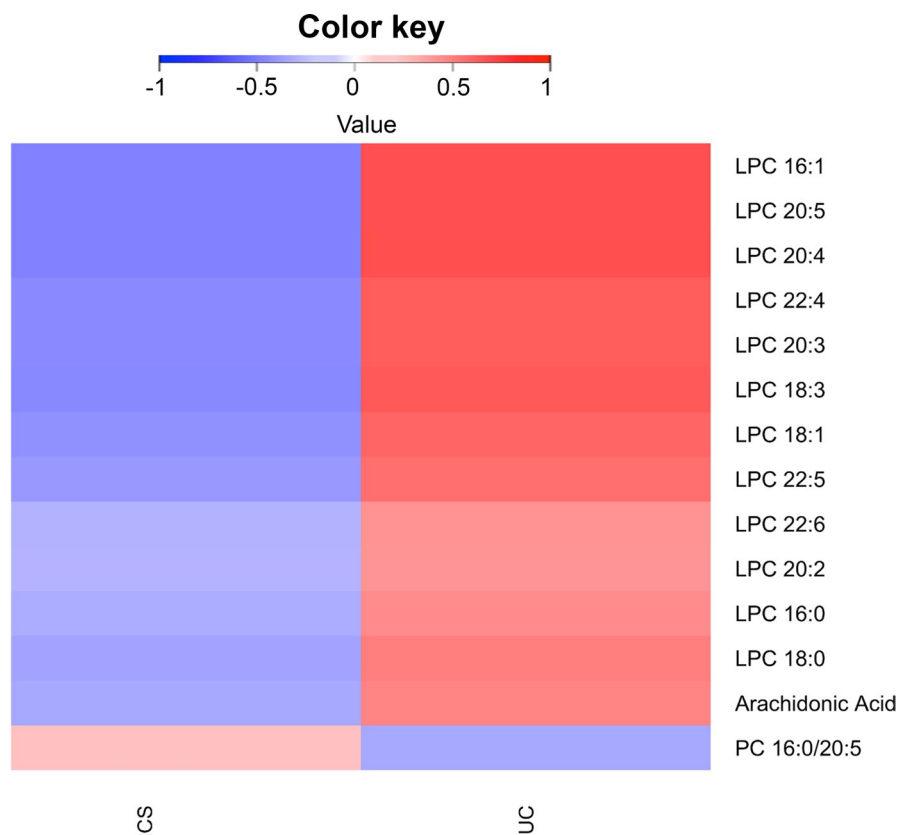


FIGURE 3 Legend on next page

FIGURE 3 Trajectories of the most relevant identified compounds. Box and whiskers plots with mean are represented for corticosteroid-controlled (ICS, blue) and uncontrolled (UC, red) groups. For metabolites that had more than one m/z or that were detected in both ESI modes, the most abundant m/z overall was represented. Mean is represented with '+' inside the boxes and individual data points are shown as blue (ICS) or red (UC) dots. Mann-Witney U test was used to calculate significant differences. * $p < .05$; ** $p < .01$. Other trajectories can be found in Figure S4

FIGURE 4 Changes in the abundance of metabolites in the PLA2 pathway between ICS and UC patients. Red cells represent an increase in the amount of a metabolite, whereas blue cells represent a decrease. LPC: lysophosphocholine, PC: phosphocholine



3, so we built one model with each value to select the one with a better performance. However, both models ended up being identical.

From the original 136 variables (50 clinical variables and 86 metabolites), we ended with a k-NN model that used 39 (16 clinical and 23 metabolomic variables). This model was able to correctly classify all patients but one. The accuracy, specificity, sensibility, positive predicted value (PPV) and negative predicted value (NPV) were all ≥ 0.75 (Table S8). We also performed a variable selection analysis, but the resultant model had worst performance (data not shown).

4 | DISCUSSION

Management of uncontrolled severe asthma entails a significant challenge for healthcare professionals. These patients are difficult to identify in the first place and need personalized treatment because they do not respond to conventional pharmacological interventions. We believe that a better knowledge regarding the metabolic, proteomic and biological pathways associated with this specific phenotype will provide new insights and will be helpful towards the management of these patients.

Here, we studied a cohort of HDM-allergic asthmatic patients from a high-exposure HDM region. The Canary Islands are located near the west coast of North Africa, close to the Sahara Desert, and thus, are impacted by Saharan Dust (calima) during approximately 30% of the year. The most common HDM sensitizers in the Canary Islands are *Dermatophagoides pteronyssinus* and *Blomia tropicalis*, which are linked to asthma prevalence.^{7,21} Moreover, prevalence of persistent rhinitis and asthma is higher in this area than in the mainland of Spain, as previously reported,^{6,11} probably due to the higher sensitization profile to perennial allergens and favourable ecological conditions supporting mite growth. Similar conditions have been described in other insular areas like the United Kingdom, Australia or New Zealand.^{7,12} Thus, even if this study was carried out in a specific region, the results can be extrapolated to many Atlantic climate areas.

Moreover, this cohort of patients is characterized by a well-known exposome and a recruitment process by a single clinical group with more than 30 years of experience in the field, which assures the uniformity of inclusion criteria and treatment strategy. This fact minimizes variability factors, making this a unique asthma model that is representative of the severe allergic asthma phenotype.

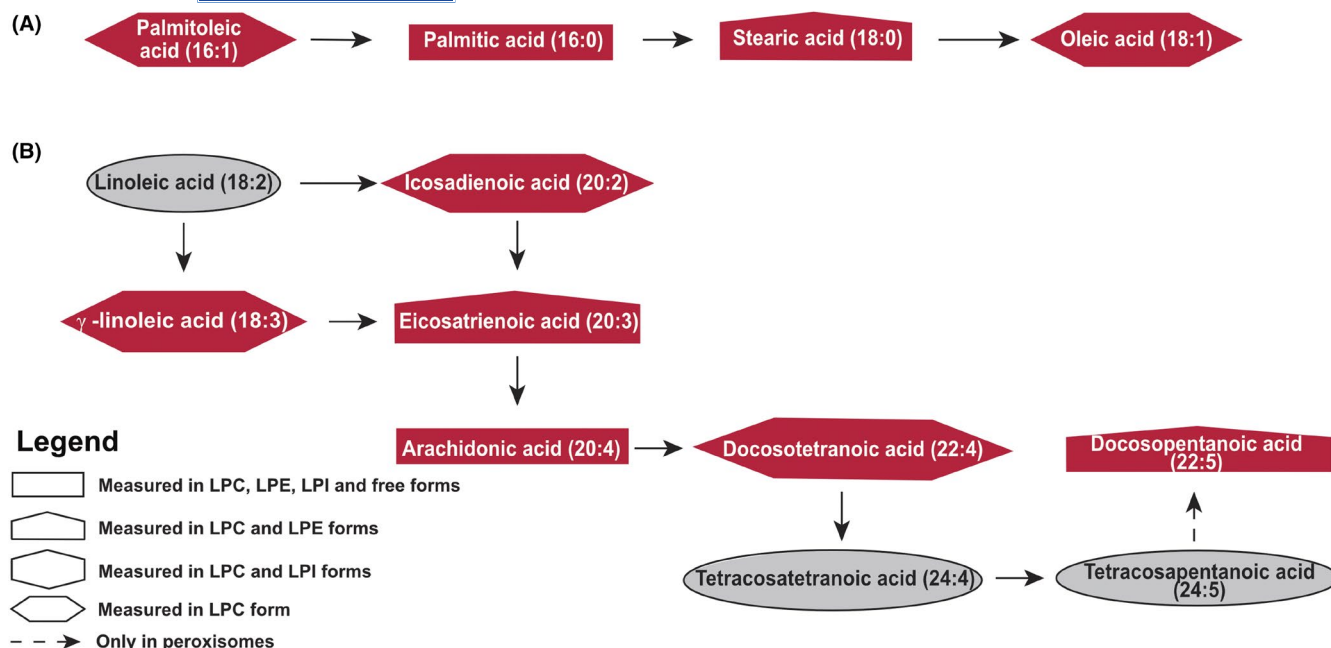


FIGURE 5 Pathways of lipid metabolism in uncontrolled patients. A. Oleic acid pathway; B. Arachidonic acid pathway. In red, metabolites increased in uncontrolled patients (UC) with respect to corticosteroid-controlled patients (ICS); in grey, metabolites that were not significantly changed between UC and ICS patients

According to their response to treatment, patients were classified into 4 groups. Patients that controlled asthma with non-systemic CS were considered mild HDM-allergic asthmatic patients (ICS); and patients that do not respond to none of the available treatments are considered uncontrolled HDM-allergic asthmatic patients (UC). Additionally, patients controlled with IT and biological drugs were also included, showing a disperse distribution probably due to several factors, like time of treatment, administration pathway or response to treatment, that might need further analysis. Moreover, BIO patients were only treated with omalizumab, since it was the only biological drug approved for the treatment of asthma in the Canary Islands at the time. Therefore, it will be relevant to perform further studies considering other biologics such as mepolizumab or benralizumab.

Here, we combine an appropriate clinical classification, focusing on disease control, with a wide range of techniques to identify metabolites and proteins.

Metabolomics is a promising tool in the detection of dynamic changes and alterations in the metabolism associated to a pathology.²² In fact, the metabolome has been described to be closely linked to the phenotype of a disease and can be an extremely useful tool when evaluating the effect of treatments. From a practical point of view, it uses very sensitive and specific techniques, such as LC-MS, which allows the simultaneous detection of a great variety of metabolites in each biological sample.^{23,24}

In this study, we performed an untargeted metabolomic analysis in all samples from each experimental group, what results, after a complex and laborious statistical analysis, into the identification of several significant signals that conform the metabolomic signature of UC group. Therefore, this study should be understood as

exploratory. In addition to the metabolites, we have also identified the biological pathways associated to those metabolic changes. However, further analyses in prospective studies are needed to validate these metabolites as potential biomarkers in a new and larger cohort of patients, including other subtypes of asthma.

Most of the identified compounds (94%) have a FDR < 0.2. For exploratory studies, setting higher cut-offs allows them to find more potential biomarkers resulting in a more extensive knowledge¹⁷⁻¹⁹; although validation in further studies is critical to ensure that the resulting biomarkers can properly classify the patients.²⁵

Our metabolomic results show that LPC, LPE and LPI, along with arachidonic acid, are increased in the UC compared to the ICS. Lipid metabolism has been previously associated with asthma, as lipid mediators are related to inflammatory signalling pathways. Yoder *et al*²⁶ previously reported increased levels of LPC 16:0 and LPC 18:0 in the bronchoalveolar lavage fluid (BALF) of patients with moderate asthma compared to non-asthmatic or mild-asthmatic patients. This was also accompanied by an increase in PLA2 activity, a pathway that we report increased in UC patients as well. Moreover, Pang *et al*²⁷ compared eosinophilic and non-eosinophilic asthma, finding that LPC 18:1 was increased in the eosinophilic phenotype. As a matter of fact, LPC have been reported to have a role in eosinophil regulation.²⁸⁻³⁰ Therefore, the observed increase in LPC seems to be directly associated with an exacerbated inflammatory response taking place in UC patients.

LPC have also been associated to arachidonic acid release in human derived monocytes.³¹ Arachidonic acid (C 20:4; AA), eicosapentaenoic acid (C 20:5; EPA) and docosahexaenoic acid (C 22:6; DHA) were detected either as free fatty acids or in LPC, LPE and/or LPI forms. An enrichment of these metabolites and their implied

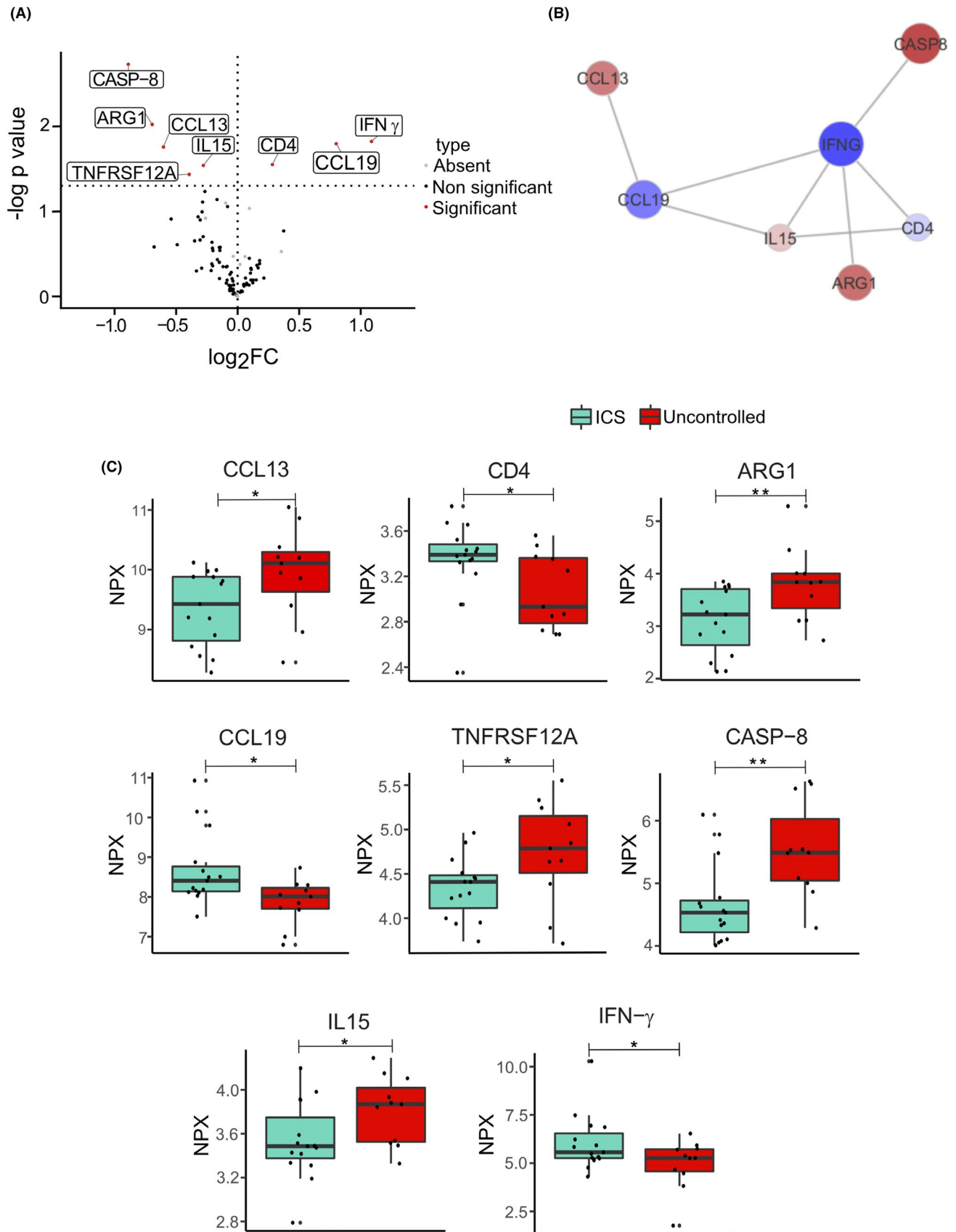


FIGURE 6 Specific serum proteomic profile of ICS and UC patients. A. Volcano plot of the differences in proteomic profile between ICS and uncontrolled patients. B. STRING analysis of significant proteins. An increase in the size of the circles and intensity of colour signals an increase in logFC. Red signals an increase of the protein expression, while blue means that the protein is decreased. C. Boxplot of the relative abundances of significant proteins between ICS (light blue) and UC (red). Significance * $p < .05$, ** $p < .01$

pathways has been previously reported in asthma.^{27,32-34} These fatty acids act as precursors of eicosanoids (prostaglandins, leukotrienes, lipoxins), resolvins and protectins through the activity of enzymes such as cyclooxygenase (COX), lipoxygenase (LOX) or cytochrome P450.^{35,36} This suggests that UC patients have a significant activation of these pathways, which are involved in the regulation of the immune response.

S1P was also increased in UC patients. This metabolite has been previously associated with asthma and other lung inflammatory diseases,³⁷ along with asthma severity.^{32,38} Interestingly, we have reported an increase of S1P levels in plasma of patients with severe allergy.³⁹ Moreover, S1P also plays a key role in regulation of T cells lifespan and recruitment.⁴⁰ Thus, this increase might point to the activation of T cells in UC patients, which may play a role in uncontrolled asthma pathogenesis.

Other metabolites significantly altered in UC patients are arginine, which is related to the nitric oxide (NO) pathway^{38,41-44}, and leucine, associated with the activation of mTORC1 pathways responsible for T cells activation, proliferation and differentiation.⁴⁵

Furthermore, proteomic analysis reveal that UC patients seems to have an increased pro-inflammatory T cell response, which is also supported but the changes in S1P and leucine abovementioned. We found an increase in CCL13 and IL-15 in the UC patients, which have been shown to attract monocytes and lymphocytes and promote T cell proliferation, suggesting an activation of the immune response.⁴⁶⁻⁴⁹ This activation would be mainly produced in the mucosa, as there was a reduction of CCL19, responsible of lymphoid organ homing, and soluble CD4.⁵⁰⁻⁵² Moreover, this is a type 2 response, as shown by the decrease of IFN γ . This inflammation was accompanied by complementary epithelial and endothelial permeability and growth, as shown in the increase of proangiogenic factor TNFRSF12A and ARG 1, which regulates NOS activity.^{53,54} Finally, this inflammation might be maintained in time by the increased levels of CASP-8, as this caspase is important to maintain T cell proliferation and has been shown to inactivate RIPK1, which limits TNF and inflammatory responses.^{55,56}

Together, metabolomic and proteomic data demonstrate unique biological mechanisms associated with HDM-allergic asthmatic uncontrolled asthma patients (UC) when compare with the mild phenotype (ICS), which are mainly associated with a pro-inflammatory environment and T cell activation and proliferation.

Next, aiming to integrate clinical and demographical with metabolic data and to identify potential classifiers for UC patients, we performed a ML approach. K-NN algorithm is one of the simplest ML algorithms, but it has a very efficient performance.⁵⁷ Using this model, we could properly classify UC patients with more than 80% of accuracy. In fact, our results demonstrate that inclusion of clinical data in the ML model significantly improved patient's classification, demonstrating the need to develop more bioinformatic models that allow the integration of clinical and omics data to generate a complete picture of the patient phenotype.

Currently, high throughput techniques are already being used in other disciplines, like cancer or even in allergy (e.g. protein arrays).

The next step in precision medicine will be to have a more detailed and accurate vision of the patient through the integration of data from one or more high throughput techniques (multi-omics) with clinical and demographical data. This type of analysis will provide a better comprehension of the pathology, and will help develop better systems to classify patients, design personalized treatments and reach a more accurate diagnosis.

Together, these data show a set of specific biological features associated with the UC phenotype. Metabolomics, proteomics and clinical data all point towards a systemic and sustained inflammatory response, underlined by an exacerbated lipid, S1P and NO metabolism, and an increase in T cells activation and proliferation. Moreover, our results support the need of exhaustive clinical criteria to better characterize and classify these patients. These observations shed light into the molecular and cellular mechanisms underlying this phenotype and provide a starting point to design novel personalized treatments for these patients.

ACKNOWLEDGEMENTS

We would like to thank all institutions involved: Institute of Applied Molecular Medicine (IMMA, Universidad CEU San Pablo, CEU Universities, Madrid), Centre of Metabolomics and Bioanalysis (CEMBIO, Universidad CEU San Pablo, CEU Universities, Madrid), and Swiss Institute of Allergy and Asthma Research (SIAF); as well as the Hospital Universitario de Gran Canaria Dr Negrín (Las Palmas de Gran Canaria, Spain). This work was supported by ISCIII (PI18/01467 and PI19/00044), cofounded by FEDER 'Investing in your future' for the thematic network and co-operative research centres ARADyAL RD16/0006/0015; as well as by the grant from Ministerio de Ciencia, Innovación y Universidades co-financed with FEDER RTI2018-095166-B-I00. This work was also supported by the Ministry of Science, Innovation and Universities in Spain (PCI2018-092930) co-funded by the European program ERA HDHL - Nutrition & the Epigenome, project Dietary Intervention in Food Allergy: Microbiome, Epigenetic and Metabolomic interactions DIFAMEM. M.I.D.D. and J.R.-C. are supported by FPI-CEU predoctoral fellowships, D.O. is funded by JPI-DIFAMEM PCI2018-092930 and A.V. is funded by a postdoctoral research fellowship from ARADyAL. MS is funded by the Swiss National Science Foundation (SNFS) grant 310030_189334/1.

CONFLICT OF INTEREST

Delgado-Dolset has nothing to disclose. Dr. Obeso has nothing to disclose. Rodríguez-Coira has nothing to disclose. Dr. Tarín has nothing to disclose. Dr. Tan has nothing to disclose. Dr. Cumplido Bonny has nothing to disclose. Dr. Cabrera Santana has nothing to disclose. Dr. Angulo Díaz-Parreño has nothing to disclose. Dr. Barbas reports grants from Ministerio de Ciencia, Innovación y Universidades co-financed with FEDER RTI2018-095166-B-I00, during the conduct of the study. Dr. Sokolowska reports grants from Swiss National Science Foundation (SNSF), grants from GSK, outside the submitted work. Dr. Barber reports grants from ALK, Allero Therapeutics, personal fees from ALK, AIMMUNE, outside the submitted work and

the grant from the Ministry of Science, Innovation and Universities in Spain (PCI2018-092930) co-funded by the European program ERA HDHL - Nutrition & the Epigenome, project Dietary Intervention in Food Allergy: Microbiome, Epigenetic and Metabolomic interactions DIFAMEM. Dr. Carrillo has nothing to disclose. Dr. Villaseñor has nothing to disclose. Dr. Escribese has nothing to disclose.

AUTHOR CONTRIBUTIONS

MME was the PI and together with DB, AV, CB and MS designed and supervised the research. TC, JAC and AC recruited the patients and obtained the samples. MIDD, JR-C, AV and DO performed the metabolomics analysis and data treatment. JR-C and GT performed the proteomic analysis and data treatment. MIDD performed analysis of the results together with CT and SA. All authors contributed to the writing of the manuscript and have given approval to the final version of the manuscript.

ORCID

María Isabel Delgado-Dolset  <https://orcid.org/0000-0001-7924-5454>

David Obeso  <https://orcid.org/0000-0001-7875-7327>

Juan Rodríguez-Coira  <https://orcid.org/0000-0003-0517-7078>

Ge Tan  <https://orcid.org/0000-0003-0026-8739>

Santiago Angulo  <https://orcid.org/0000-0002-7868-7442>

Coral Barbas  <https://orcid.org/0000-0003-4722-491X>

Milena Sokolowska  <https://orcid.org/0000-0001-9710-6685>

Domingo Barber  <https://orcid.org/0000-0002-5488-5700>

Alma Villaseñor  <https://orcid.org/0000-0002-6652-2739>

María M. Escribese  <https://orcid.org/0000-0001-5057-5150>

REFERENCES

- Global Initiative for Asthma. *Global strategy for asthma management and prevention*. 2020. www.ginasthma.org
- Campo P, Rodríguez F, Sánchez-García S, et al. Phenotypes and endotypes of uncontrolled severe asthma: new treatments. *J Investig Allergol Clin Immunol*. 2013;23:76-88. quiz 1 p. follow 88.
- Agache I, Akdis C,utel M, Virchow JC. Untangling asthma phenotypes and endotypes. *Allergy*. 2012;67:835-846.
- Kennedy JL, Heymann PW, Platts-Mills TAE. The role of allergy in severe asthma. *Clin Exp Allergy*. 2012;42:659-669.
- Sánchez-Borges M, Fernandez-Caldas E, Thomas WR, et al. International consensus (ICON) on: clinical consequences of mite hypersensitivity, a global problem. *World Allergy Organ J*. 2017;10:14.
- Julià-Serdà G, Cabrera Navarro P, Acosta Fernández O, et al. High prevalence of asthma symptoms in the Canary Islands: climatic influence? *J Asthma*. 2005;42:507-511.
- Julià-Serdà G, Cabrera-Navarro P, Acosta-Fernández O, et al. High prevalence of asthma and atopy in the Canary Islands, Spain. *Int J Tuberc Lung Dis*. 2011;15:536-541.
- Barber D, Arias J, Boquete M, et al. Analysis of mite allergic patients in a diverse territory by improved diagnostic tools. *Clin Exp Allergy*. 2012;42:1129-1138.
- Plaza MV. GEMA4.0. Guía española para el manejo del asma. *Arch Bronconeumol*. 2015;51:2-54.
- Peters SP, Ferguson G, Deniz Y, Reisner C. Uncontrolled asthma: a review of the prevalence, disease burden and options for treatment. *Respir Med*. 2006;100:1139-1151.
- Suárez-Lorenzo I, Cruz-Niesvaara D, Rodríguez-Gallego C, Rodríguez de Castro F, Carrillo-Díaz T. Epidemiological study of the allergic population in the north of Gran Canaria. *J Investig Allergol Clin Immunol*. 2018;28:212-215.
- European Respiratory Society. *European Lung White Book*. 2003;16-25.
- Rodríguez-Coira J, Delgado-Dolset M, Obeso D, et al. Troubleshooting in large-scale LC-ToF-MS metabolomics analysis: solving complex issues in big cohorts. *Metabolites*. 2019;9:247.
- Gil-de-la-Fuente A, Godzien J, Saugar S, et al. CEU mass mediator 3.0: a metabolite annotation tool. *J Proteome Res*. 2019;18:797-802.
- Reel PS, Reel S, Pearson E, Trucco E, Jefferson E. Using machine learning approaches for multi-omics data analysis: a review. *Biotechnol Adv*. 2021;49:107739.
- Triba MN, Le Moyec L, Amathieu R, et al. PLS/OPLS models in metabolomics: the impact of permutation of dataset rows on the K-fold cross-validation quality parameters. *Mol Biosyst*. 2015;11:13-19.
- Chen M, Zhou K, Chen X, et al. Metabolomic analysis reveals metabolic changes caused by bisphenol A in rats. *Toxicol Sci*. 2014;138:256-267.
- Fiehn O, Garvey WT, Newman JW, Lok KH, Hoppel CL, Adams SH. Plasma metabolomic profiles reflective of glucose homeostasis in non-diabetic and type 2 diabetic obese African-American women. *PLoS One*. 2010;5:e15234.
- Putluri N, Shojai A, Vasu VT, et al. Metabolomic profiling reveals potential markers and bioprocesses altered in bladder cancer progression. *Cancer Res*. 2011;71:7376-7386.
- Hajiloo M, Damavandi B, HooshSadat M, et al. Breast cancer prediction using genome wide single nucleotide polymorphism data. *BMC Bioinformatics*. 2013;14:S3.
- Julià-Serdà G, Cabrera-Navarro P, Acosta-Fernández O, Martín-Pérez P, García-Bello MA, Antó-Boqué J. Prevalence of sensitization to *Blomia tropicalis* among young adults in a temperate climate. *J Asthma*. 2012;49:349-354.
- Villaseñor A, Rosace D, Obeso D, et al. Allergic asthma: an overview of metabolomic strategies leading to the identification of biomarkers in the field. *Clin Exp Allergy*. 2017;47:442-456.
- Crestani E, Harb H, Charbonnier L-M, et al. Untargeted metabolomic profiling identifies disease-specific signatures in food allergy and asthma. *J Allergy Clin Immunol*. 2020;145:897-906.
- Delgado-Dolset MI, Obeso D, Sánchez-Solares J, et al. Understanding systemic and local inflammation induced by nasal polyposis: role of the allergic phenotype. *Front Mol Biosci*. 2021;8:409.
- Vinaixa M, Samino S, Saez I, Duran J, Guinovart JJ, Yanes O. A guideline to univariate statistical analysis for LC/MS-based untargeted metabolomics-derived data. *Metabolites*. 2012;2:775-795.
- Yoder M, Zhuge Y, Yuan Y, et al. Bioactive lysophosphatidylcholine 16:0 and 18:0 are elevated in lungs of asthmatic subjects. *Allergy Asthma Immunol Res*. 2014;6:61.
- Pang Z, Wang G, Wang C, Zhang W, Liu J, Wang F. Serum metabolomics analysis of asthma in different inflammatory phenotypes: a cross-sectional study in Northeast China. *Biomed Res Int*. 2018;2018:1-14.
- Knuplex E, Curcic S, Theiler A, et al. Lysophosphatidylcholines inhibit human eosinophil activation and suppress eosinophil migration in vivo. *Biochim Biophys Acta - Mol Cell Biol Lipids*. 2020;1865:158686.
- Zhu X, Learoyd J, Butt S, et al. Regulation of eosinophil adhesion by lysophosphatidylcholine via a non-store-operated Ca²⁺ channel. *Am J Respir Cell Mol Biol*. 2007;36:585-593.
- Kariyawasam HH, Robinson DS. The role of eosinophils in airway tissue remodelling in asthma. *Curr Opin Immunol*. 2007;19:681-686.
- Oestvang J, Anthonsen MW, Johansen B. LysoPC and PAF trigger arachidonic acid release by divergent signaling mechanisms in monocytes. *J Lipids*. 2011;2011:1-11.

32. McGeachie MJ, Dahlin A, Qiu W, et al. The metabolomics of asthma control: a promising link between genetics and disease. *Immunity, Inflamm Dis*. 2015;3:224-238.
33. Nie X, Wei J, Hao Y, et al. Consistent biomarkers and related pathogenesis underlying asthma revealed by systems biology approach. *Int J Mol Sci*. 2019;20:4037.
34. Sokolowska M, Chen L-Y, Liu Y, et al. Dysregulation of lipidomic profile and antiviral immunity in response to hyaluronan in patients with severe asthma. *J Allergy Clin Immunol*. 2017;139:1379-1383.
35. Miyata J, Arita M. Role of omega-3 fatty acids and their metabolites in asthma and allergic diseases. *Allergol Int*. 2015;64:27-34.
36. Sokolowska M, Rovati GE, Diamant Z, et al. Current perspective on eicosanoids in asthma and allergic diseases - EAACI Task Force consensus report, part I. *Allergy*. 2021;76(1):114-130. 10.1111/all.14295
37. Petrache I, Berdyshev EV. Ceramide signaling and metabolism in pathophysiological states of the lung. *Annu Rev Physiol*. 2016;78:463-480.
38. Reinke SN, Gallart-Ayala H, Gómez C, et al. Metabolomics analysis identifies different metabolotypes of asthma severity. *Eur Respir J*. 2017;49:1601740.
39. Obeso D, Mera-Berriatua L, Rodríguez-Coira J, et al. Multi-omics analysis points to altered platelet functions in severe food-associated respiratory allergy. *Allergy*. 2018;73:2137-2149.
40. Mendelson K, Evans T, Hla T. Sphingosine 1-phosphate signalling. *Development*. 2014;141:5-9.
41. Kertys M, Grendar M, Kosutova P, Mokra D, Mokry J. Plasma based targeted metabolomic analysis reveals alterations of phosphatidylcholines and oxidative stress markers in guinea pig model of allergic asthma. *Biochim Biophys Acta - Mol Basis Dis*. 2020;1866:165572.
42. Comhair SAA, McDunn J, Bennett C, Fattig J, Erzurum SC, Kalhan SC. Metabolomic endotype of asthma. *J Immunol*. 2015;195:643-650.
43. Jung J, Kim S-H, Lee H-S, et al. Serum metabolomics reveals pathways and biomarkers associated with asthma pathogenesis. *Clin Exp Allergy*. 2013;43:425-433.
44. Farraia M, Cavaleiro Rufo J, Paciência I, et al. Metabolic interactions in asthma. *Eur Ann Allergy Clin Immunol*. 2019;51:196.
45. Ananieva EA, Powell JD, Hutson SM. Leucine metabolism in T cell activation: mTOR signaling and beyond. *Adv Nutr Int Rev J*. 2016;7:798S-805S.
46. Garcia-Zepeda EA, Combadiere C, Rothenberg ME, et al. Human monocyte chemoattractant protein (MCP)-4 is a novel CC chemokine with activities on monocytes, eosinophils, and basophils induced in allergic and nonallergic inflammation that signals through the CC chemokine receptors (CCR)-2 and -3. *J Immunol*. 1996;157:5613-5626.
47. Li CW, Zhang KK, Li TY, et al. Expression profiles of regulatory and helper T-cell-associated genes in nasal polyposis. *Allergy*. 2012;67:732-740.
48. Morris SR, Chen B, Mudd JC, et al. 'Inflammascent' CX3CR1+CD57+ CD8 T cells are generated and expanded by IL-15. *JCI Insight*. 2020;5. 10.1172/jci.insight.132963
49. Grabstein K, Eisenman J, Shanebeck K, et al. Cloning of a T cell growth factor that interacts with the beta chain of the interleukin-2 receptor. *Science*. 1994;264:965-968.
50. He H, Suryawanshi H, Morozov P, et al. Single-cell transcriptome analysis of human skin identifies novel fibroblast subpopulation and enrichment of immune subsets in atopic dermatitis. *J Allergy Clin Immunol*. 2020;145:1615-1628.
51. Kaiser A, Donnadiou E, Abastado J-P, Trautmann A, Nardin A. CC chemokine ligand 19 secreted by mature dendritic cells increases naive T cell scanning behavior and their response to rare cognate antigen. *J Immunol*. 2005;175:2349-2356.
52. Takamura K, Fukuyama S, Nagatake T, et al. Regulatory role of lymphoid chemokine CCL19 and CCL21 in the control of allergic rhinitis. *J Immunol*. 2007;179:5897-5906.
53. Mendez-Barbero N, Yuste-Montalvo A, Nuñez-Borque E, et al. The TNF-like weak inducer of the apoptosis/fibroblast growth factor-inducible molecule 14 axis mediates histamine and platelet-activating factor-induced subcutaneous vascular leakage and anaphylactic shock. *J Allergy Clin Immunol*. 2020;145:583-596.e6.
54. Suwanpradit J, Shih M, Pontius L, et al. Arginase1 deficiency in monocytes/macrophages upregulates inducible nitric oxide synthase to promote cutaneous contact hypersensitivity. *J Immunol*. 2017;199:1827-1834.
55. Oberst A, Dillon CP, Weinlich R, et al. Catalytic activity of the caspase-8-FLIP L complex inhibits RIPK3-dependent necrosis. *Nature*. 2011;471:363-368.
56. Bell BD, Leverrier S, Weist BM, et al. FADD and caspase-8 control the outcome of autophagic signaling in proliferating T cells. *Proc Natl Acad Sci USA*. 2008;105:16677-16682.
57. Venables WN, Ripley BD. *Modern Applied Statistics with S-Plus*. Springer; 2013.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Delgado-Dolset MI, Obeso D, Rodríguez-Coira J, et al. Understanding uncontrolled severe allergic asthma by integration of omic and clinical data. *Allergy*. 2021;00:1-14. doi:[10.1111/all.15192](https://doi.org/10.1111/all.15192)

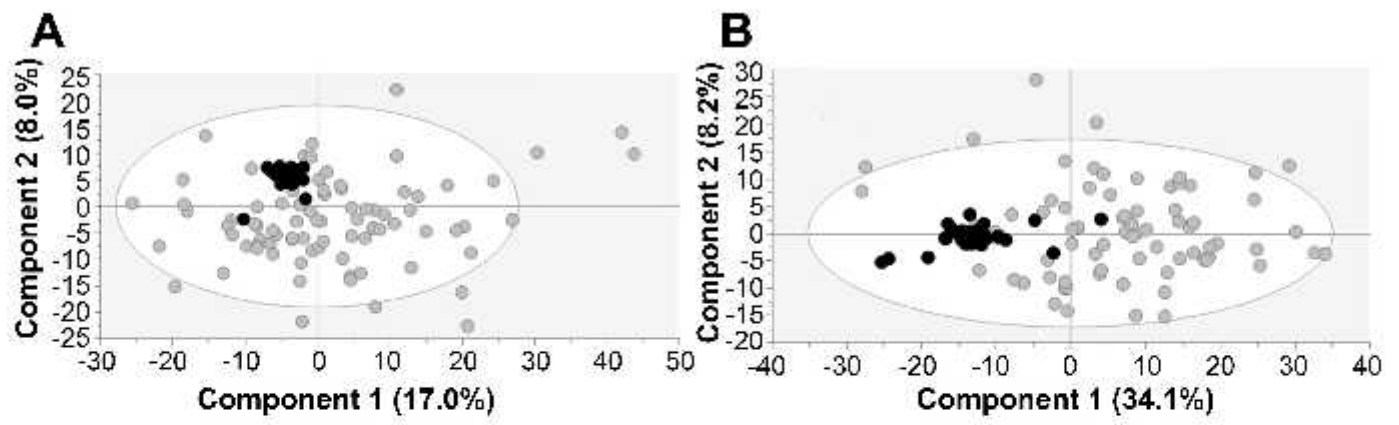


Figure S1_Delgado Dolset et al

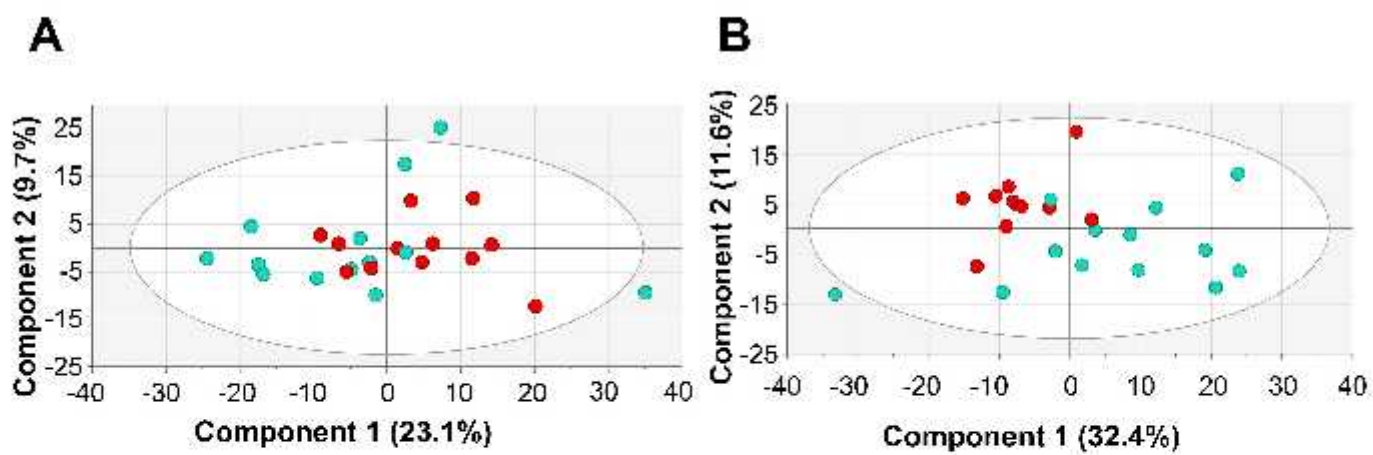


Figure S2_Delgado Dolset et al

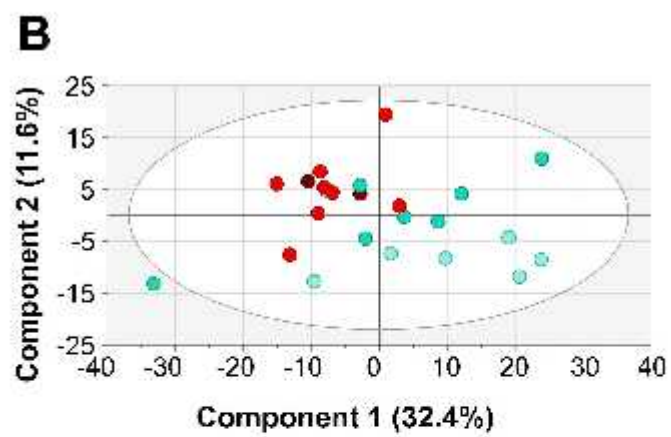
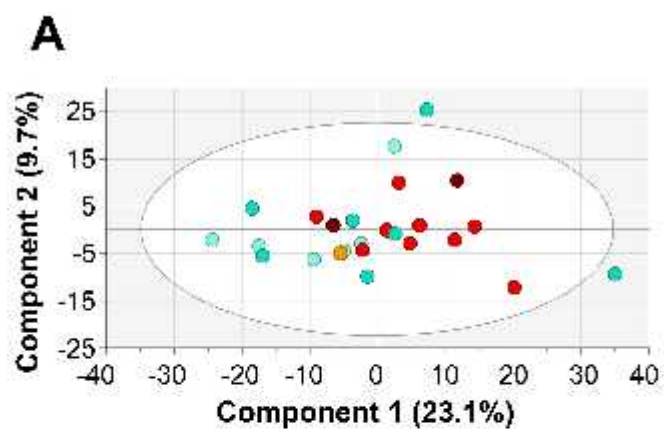


Figure S3_Delgado Dolset et al

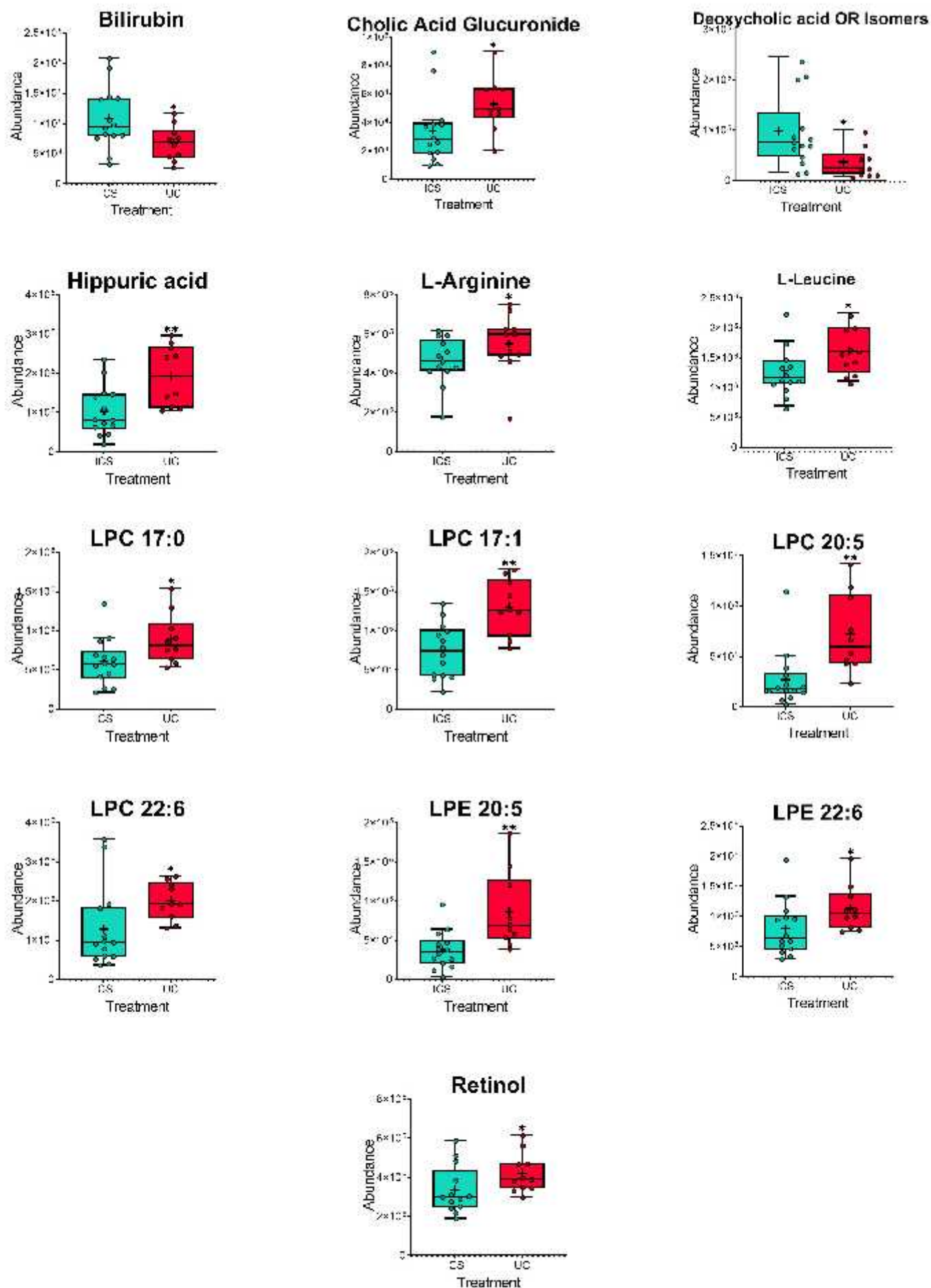


Figure S4_Delgado Dolset et al

Online Data Supplement

Understanding uncontrolled severe allergic asthma by integration of omic and clinical data.

María I Delgado-Dolset^{1,2}, David Obeso^{1,2}, Juan Rodríguez-Coira^{1,2,3}, Carlos Tarín¹, Ge Tan³, José A Cumplido⁴, Ana Cabrera⁴, Santiago Angulo⁵, Coral Barbas², Milena Sokolowska³, Domingo Barber¹, Teresa Carrillo⁴, Alma Villaseñor¹, María M Escribese^{*1}

1. Institute of Applied Molecular Medicine (IMMA), Department of Basic Medical Sciences, Facultad de Medicina, Universidad San Pablo CEU, CEU Universities, Urbanización Montepríncipe, 28660 Boadilla del Monte, Madrid, Spain.
2. Centre for Metabolomics and Bioanalysis (CEMBIO), Department of Chemistry and Biochemistry, Facultad de Farmacia, Universidad San Pablo CEU, CEU Universities, Urbanización Montepríncipe, 28660 Boadilla del Monte, Madrid, Spain.
3. Swiss Institute of Allergy and Asthma Research (SIAF), University of Zurich, Switzerland
4. Hospital Universitario de Gran Canaria Doctor Negrin, Las Palmas de Gran Canaria, Spain
5. Department of Applied Mathematics and Statistics, Universidad San Pablo-CEU, CEU Universities, Madrid, Spain

Corresponding Author:

María M Escribese, PhD

Department of Basic Medical Sciences

Faculty of Medicine. San Pablo CEU University.

Content:

Part 1. Patient profiling and clinical characteristics

Part 2. Serum samples collection

Part 3. Detailed description of sample preparation, instrumental description and data pre-treatment for metabolomics analysis.

Part 4. Proteomics analysis: Olink approach

Part 5. Machine learning model construction

Supplementary Tables:

Table S1. Complete information of patients recruited in the study.[Errore. Il segnalibro non è definito.](#)¹⁰

Table S 2. Results from ANCOVA analysis between ICS and UC groups considering age as a covariable [Errore. Il segnalibro non è definito.](#)¹⁵

Table S3. Physicochemical characteristics of significant annotated signals in metabolomics through LC-MS/MS between UC and ICS groups. [Errore. Il segnalibro non è definito.](#)¹⁷

Table S4. Pairwise comparisons showing the significance and % of change among experimental groups. Significant differences are marked bold, while the percentage of changes for significant metabolites is marked in bold red for increasing metabolites or bold blue for decreasing metabolites. [Errore. Il segnalibro non è definito.](#)²⁰

Table S5. Enriched metabolic pathways in UC patients compared to ICS group ($FDR < 0.2$). [Errore. Il segnalibro non è definito.](#)²³

Table S6. Differential expression of serum proteins between ICS and UC groups.[Errore. Il segnalibro non è definito.](#)²⁴

Table S7. Top 25 most significantly enriched biological Gene Ontology (GO) processes linked to significantly expressed proteins in serum of UC patients.[Errore. Il segnalibro non è definito.](#)²⁶

Table S8. Machine-Learning parameters for k-NN models.[Errore. Il segnalibro non è definito.](#)²⁷

Supplementary Figures:

Figure S 1. Unsupervised PCA models showing the quality of the data for LC-MS positive (A) and negative (B) polarities. (Key: QC injections (●; black). samples (●; grey). The clustering of the QCs injections proved that the data from metabolomics is of high-quality, so the differences found later between the samples are due to the condition of the patients nor the technique.

[Errore. Il segnalibro non è definito.29](#)

Figure S 2. PCA models of ICS and UC groups in both ESI modes (Key: “ICS”: blue; “UC”: red).

[Errore. Il segnalibro non è definito.30](#)

Figure S 3. PCA models of ICS and UC samples coloured in both ESI modes by age range (keynote: red: UC samples, from which orange-red: 18-35 years old, light-red: 35-60 years old and dark-red: ≥ 60 years old; blue: ICS samples, from which light-blue: 18-35 years old; dark-blue: 35-60 years old).

[Errore. Il segnalibro non è definito.31](#)

Figure S 4. Trajectories of the remaining identified compounds. Box and whiskers plots with mean are represented for corticosteroid-controlled (ICS, blue) and uncontrolled (UC, red) groups. For metabolites that had more than one m/z or that were detected in both ESI modes, the most abundant m/z overall was represented. Mean is represented with “+” inside the boxes and individual patients are shown as blue (ICS) or red (UC) circles. Mann-Whitney U test was used to calculate significant differences. * $p < 0.05$; ** $p < 0.01$.

[Errore. Il segnalibro non è definito.32](#)

Supplementary References

Part 1. Patient profiling and clinical characteristics

Patients were tested using standardized extracts of the following HDM: *Dermatophagoides pteronyssinus*, *Dermatophagoides farinae*, *Blomia tropicalis*, *Acarus siro*, *Lepidoglyphus destructor* and *Tyrophagus putrescentiae*. Skin prick test was considered as positive when the larger diameter of the wheal was at least 3 mm bigger than that of the negative control.

Regarding patient stratification, patients were considered severe if they either needed treatment from steps 4-5 of the GINA, or 5-6 of the GEMA guidelines (ICS/LABA in high doses + LTRA + theophylline) or had to take systemic corticosteroids to achieve control in the previous year. Patients whose asthma remained uncontrolled despite receiving high doses of corticosteroids (both inhaled and systemic) or biologicals were also included.

Asthma control was assessed using both Asthma Control test (ACT) and objective factors such as symptom control, cease of exacerbations in the prior year, reduction in bronchodilator usage, etc. Inclusion criteria for controlled groups was that patients had not suffered any exacerbations for a year prior to their enrolment in this study. The uncontrolled patients had roughly 5 exacerbations per year.

Patients from the IT group were considered to have reached control of their asthma after receiving IT treatment. Asthma control was assessed after one year of treatment, and patients were treated with IT for at least 2 years prior to sample collection. Patients received a subcutaneous immunotherapy with a composition of those HDM to which the patients were sensitised, according to both the skin prick test and the specific IgE levels found in serum of each patient. For the BIO group, patients were in treatment with omalizumab for at least 1 year prior to the sample collection. Asthma control was evaluated after 12 to 16 weeks of treatment. To assess asthma control, symptom control, reduction in the number of exacerbations, reduction in bronchodilators usage and Asthma Control Test, among other parameters, were evaluated.

Part 2. Serum samples collection

Whole blood was collected and incubated with a clotting agent using Vacutainer SST II tubes. Samples were placed at room temperature for 30 min and then centrifuged at 2000 x g for 10 min. Serum was recovered and stored at -80° C until the metabolomic and proteomic analysis was performed.

Part 3. Detailed description of sample preparation, instrumental description and data pre-treatment for metabolomics analysis.

1. Sample and QC preparation. Composition of Internal standard mixture

Samples were measured following an approach by batches to avoid analytical signal decay. At the day of analysis, 100 µl of serum from each sample were mixed with 300 µl of cold (-20°C) methanol:ethanol (1:1) and 20 µl of an internal standard (IS) mixture. The IS mix was selected to cover a wide range of different retention time (RTs) and physic-chemical properties. This was composed of: 1) Carnitine-D3 ($C_7H_{12}D_3NO_3$, mass = 144.1454 Da, RT: 0.70 min, 0.02 mM); 2) Isoleucine- ^{13}C , ^{15}N ($^{13}C_6H_{13}^{15}NO_2$, molecular mass = 138.1118 Da, RT: 0.77 min, 0.18 mM); 3) Sphingosine-D₇ ($C_{18}H_{30}D_7NO_2$, molecular mass = 306.3264 Da, RT: 14.38 min, 0.02 mM); 4) LPC 18:1-D₇ ($C_{26}H_{45}D_7NO_7P$, molecular mass = 528.3921 Da, 2 peaks at RTs: 19.30 and 20.00 min, 0.01 mM) and 5) Stearic acid-D₅ ($C_{18}H_{31}D_5O_2$, molecular mass = 289.3029 Da, RT: 34.54 min, 0.17 mM).

Samples were then vortex-mixed, kept on ice for 5 min and centrifuged at 16110 x g for 20 min at 4°C. Supernatant was transferred into LC vial for analysis. For blank preparation, 100 µl of water and 300 µl of cold methanol:ethanol with 20 µl of the IS mixture were used and followed the same steps as the sample. Quality control sample (QC) was prepared by pooling equal volumes of a set of samples. QC was analysed throughout the run to measure system stability.

128

129 **2. Instrumental analysis**

130 Serum samples were measured using an Agilent HPLC system (1200 series) coupled with
131 quadrupole-time of flight analyser system (Q-ToF MS 6520) (Agilent Technologies,
132 Waldbronn, Germany).

133 **a. LC-Q-ToF MS Analysis**

134 The HPLC system was equipped with a degasser, two binary pumps, and a thermostated
135 autosampler. Briefly, 10 μ L of sample were injected into a Discovery HS C18 column (2.1 $\times$
136 150 mm, 3.0 μ m; Supelco, Sigma Aldrich, Germany), with a guard column Discovery® HS
137 C18 (2.1 $\times$ 20 mm, 3 μ m; Supelco), both maintained at 40 °C. The flow rate was set at 0.6
138 mL/min. The elution gradient involved a mobile phase consisting of: (A) 0.1% v/v formic acid
139 (FA) in water and (B) 0.1% v/v FA in acetonitrile. The initial conditions were set at 25% phase
140 B, which increased linearly to 95% phase B in 35 min. Then it returned to the initial
141 conditions in 1 min, which were held for 9 min for column reconditioning. Samples were
142 analysed in both ESI+ and ESI- modes in separate injections. The capillary voltage was set
143 at 3500 for ESI+ and 4000V for ESI-. The drying gas flow rate was 10.5 L/min at 330°C and
144 gas nebulizer at 52 psi; fragmentor voltage was 175 V; skimmer and octopole radio
145 frequency voltages were set to 65 and 750 V, respectively. MS spectra were collected in the
146 centroid mode at a scan rate of 1.2 Hz. The MS detection window was performed in full scan
147 from 100 to 1200 m/z for both modes. Automatic MS recalibration during batch analysis was
148 carried out introducing a reference standard into the source via a reference sprayer valve.
149 Reference masses for ESI+ were purine (m/z = 121.0508) and HP-0921 (m/z = 922.0097),
150 whereas for ESI- TFA NH₄ (m/z = 119.0363) and HP-0921 (m/z = 966.0007).

151 **b. Compound annotation**

152 Compound annotation (ID) for LC-MS was carried out for statistically significant features (p <
153 0.05) in two steps. First, a putative identification was performed in all the significant features

by searching the accurate masses against the online available databases such as KEGG, METLIN, LipidMaps and HMDB using the CEU Mass Mediator tool (1). Then, identity was confirmed through LC-MS/MS experiments by using a QTOF (model 6520, Agilent). Experiments were repeated in identical chromatographic conditions as the primary analyses. Ions were targeted using the narrow m/z window (1.3 Da) and 20 eV of energy for fragmentation on the quadrupole. Comparison of the structure proposed against the obtained fragments led to the confirmation of the identity.

For those metabolites where a commercial standard was available, identification was further confirmed by comparing RT and fragmentation pattern distribution.

3. Data treatment and data quality

Full processing of acquired signals has been published (2). Briefly, data were refined from the background noises using MassHunter Profinder (B.08.00; SP3, Agilent Technologies) software. Molecular Feature Extraction (MFE) and Find by Ion (Fbi) algorithms were applied to reduce size and complexity of the data. Afterwards, 1382 and 887 chemical signals for ESI- and ESI+ modes were found, respectively. Raw data was filtered by keeping features that were not present in the blanks, were detected in >50% of QCs and >75% in patient samples. The rest of the signals were excluded from the analyses. Missing values were replaced using k-nearest neighbours (kNN) algorithm through an in-house script developed in Matlab® (3). To correct the drift of the MS signal within batches, data were normalized using a QC-SVRC algorithm (4). For the normalization between batches, a sequential inter-batch correction called QC-norm (5) was applied. Afterwards, features with a relative standard deviation (RSD) <20% in QCs were used for the analysis. Quality of MS analysis and normalization strategy was assured by measuring a set of samples in all batches and showing their clustering in a non-supervised principal component analysis (PCA) model (2). The work illustrates a strategy to generate high-quality data from untargeted large-scale metabolomics studies.

4. Statistical analysis

Multivariate analysis was performed using SIMCA v.16.0 (Sartorius Stedim Data Analytics). A PCA model was used to evaluate data quality and find patterns in samples. Likewise, Partial Least Square Discriminant Analysis (PLS-DA) and Orthogonal Projections to Latent Structures Discriminant Analysis (OPLS-DA) supervised models were used to classify the samples and to evaluate differences between groups, respectively. OPLS-DA models were cross-validated by splitting the data into 7 cross-validation subsets. Six out of these were used to build a first model, which was then used to predict the subset that was left out. This process continues until all subsets are used to build a model and have been predicted (6). The cross-validation subsets are selected randomly, and cross-validated samples are plotted into the resulting OPLS-DA model. UV scaling was used in all models. These were evaluated using R^2 and Q^2 parameters which are the classification and prediction capacity, respectively.

Significant signals were selected after pair-wise comparisons using two-tailed Mann-Whitney U test with an FDR correction using an in-house script for Matlab R2015a software (Mathworks, Natick, Massachusetts, USA); and used to build a Hierarchical Clustering Heatmap with RStudio 1.3.1075. Additionally, as age was a significant factor between UC and ICS groups, we performed an ANCOVA test to exclude from the significant ($p < 0.05$) and annotated compounds, those that were affected by age. We discarded those that were significant either for age or for both factors (age and group); leaving the identified metabolites related only to the group using SPSS (IBM SPSS statistics version 24). Pathways of the significant identified compounds were obtained using IMPaLA (v 12.0) online tool (<http://impala.molgen.mpg.de/>).

Part 4. Proteomics analysis: Olink approach

1. Olink quality control

Olink measured samples were mixed with 3 different internal standards during the analysis to assess the integrity of the antibody binding, primer amplification and PCR DNA read out. If the quality control samples of any sample deviated more than $\pm 0.3\text{NPX}$ of the median of the plate that sample was discarded. In this project, 2 samples were discarded due to not meeting the quality criteria leading to $n=14$ for ICS and $n=10$ for UC patients.

2. Data treatment and statistical analysis

The Olink PEA data are reported in NPX values (normalized protein expression levels) which are on a log2 scale. Olink data in NPX format were imported, processed and compared between groups with OlinkR package (<https://github.com/ge11232002/OlinkR>). A protein was considered expressed if it was measured above the detection limit in more than half of the samples in either of the two compared groups. Fold change, p -value and FDR were estimated by fitting the expression of each protein to a linear model (7). Statistically significant proteins were analysed with STRING to detect the possible interaction between themselves and significantly associated Gene Ontology (GO) biological processes (FDR < 0.2)

Part 5. Machine learning model construction

1. Data pre-processing

First, clinical data were codified and prepared in order to be compatible with the metabolic and proteomic ones (converting the categorical variables into numeric variables). For the pre-processing of the data, we decided to remove those variables in which there were no values. Then, we computed the correlation matrix of all the variables and removed those that

provided the same information for the model (those which had a correlation coefficient greater than or equal to 0.75). Finally, we eliminated variables that either had one unique value (*i.e.* are zero variance predictors) or that had both very few unique values relative to the number of samples and a large ratio of the frequency of the most common value to the frequency of the second most common value. In this case, we set a cut-off value of 95/5. After this process, proteomic variables were discarded due to them not being present in all samples.

2. Downsampling and Learning Parameter Selection

We tried to build a machine learning model for the clinical data alone, the metabolomic data alone and to integrate both datasets.

To avoid bias towards the biggest (ICS) group, we performed a downsampling of the ICS samples using the function `downSample` of `caret` package and ended with just 10 samples for each group. Then, to avoid overfitting and select the optimal k -value, the elbow method was applied. The elbow method runs k -means clustering on the dataset for a range of 1 to 10 values for k , then for each value of k computes an average score for all clusters. The number of clusters is determined based on the Total Within Sum of Squares (TWSS) (8). The best candidates of the number of clusters are chosen when the graph of TWSS shows a sharp turn, called elbow (9). We found no elbow when using only metabolomic or clinical data. When using both datasets combined, we found an elbow of $k = 2$ or 3.

3. Evaluation

To estimate the quality of the classifier, we centered and scaled the variables and randomly created two groups (training and testing groups, in a proportion of 60/40) using `caret` package, to evaluate the k -NN algorithm. However, we also test the performance of a variable selection using a wrapper filter for feature selection of predictors by ANOVA, random forest and `cv-repeat`, in order to minimize the number of variables to create a simpler classifier.

257 The code used to build the k -NN algorithm can be found in
258 <https://github.com/ctarin77/Allergy>.

Supplementary Tables

Table S1. Complete information of patients recruited in the study.

Patient	Group	Age (years)	Onset Age (years)	Gender	BMI	Smoker	Treatment	Other Sensitizations	Total IgE (U)
1	ICS	30	5	F	30.0	no	AH + Inhaled CS/LABA + Topic CS	dog, shellfish	160
2	ICS	31	14	F	20.4	no	AH + Topic CS + SABA	--	346
3	ICS	28	8	F	19.4	no	AH + Inhaled CS/LABA + Topic CS	--	110
4	ICS	26	7	M	20.8	no	AH + Inhaled CS/LABA + Topic CS	cat, dog, aspergillus	1297
5	ICS	22	20	F	25.9	no	AH + Inhaled CS/LABA + Topic CS	--	76
6	ICS	42	15	M	30.3	no	AH + Inhaled CS/LABA + Topic CS	seafood	358
7	ICS	41	9	F	34.3	yes	AH + Inhaled CS/LABA + Topic CS	--	89
8	ICS	41	21	F	34.9	no	AH + Inhaled CS/LABA + Topic CS	--	611
9	ICS	40	11	F	26.0	yes	AH + Inhaled CS/LABA + Topic CS	grass pollen	365
10	ICS	41	22	F	34.6	no	AH + Inhaled CS/LABA + Topic CS	cat, dog	237
11	ICS	32	6	M	24.6	no	AH + Inhaled CS/LABA + Topic CS + LTRA	cat, dog, olive pollen	2500
12	ICS	43	40	M	25.7	no	AH + Inhaled CS/LABA + Topic CS	cat, cockroach	116
13	ICS	41	13	M	27.7	no	AH + Inhaled CS/LABA + Topic CS + LTRA	cat, dog	238
14	ICS	57	40	F	19.7	no	AH + Inhaled CS/LABA + Topic CS + SABA	barnacle, cat, dog, horse	547
15	ICS	43	6	F	26.1	ex-smoker	AH + Inhaled CS/LABA + Topic CS + SABA	LTPs (vegetables, pollen)	206

16	IT	39	25	F	21.6	no	AH + Inhaled CS/LABA + Topic CS	cat, olive pollen	208
17	IT	39	20	F	23.7	no	AH + Inhaled CS/LABA + Topic CS	cat, dog, horse	231
18	IT	31	12	F	25.0	no	AH + Inhaled CS/LABA + Topic CS	dog	182
19	IT	35	29	F	25.8	no	AH + Topic CS + SABA	dog	94
20	IT	28	3	F	22.0	no	AH + Topic CS + SABA	--	191
21	IT	43	28	F	28.9	no	Topic CS + SABA	--	331
22	IT	31	16	F	37.9	no	AH + Topic CS + SABA	paracetamol	285
23	IT	29	14	M	28.1	no	Topic CS + SABA	mugwort	741
24	IT	47	32	F	29.3	no	AH + Topic CS + SABA	cat, dog	279
25	IT	44	15	F	24.5	no	AH + Topic CS + SABA	cat, dog	981
26	IT	56	47	M	32.6	no	AH + Topic CS + SABA	--	96
27	IT	38	15	F	26.8	no	AH + Topic CS + SABA	cat, mugwort	170
28	IT	34	28	F	35.1	no	AH + Topic CS + SABA	--	41
29	IT	43	14	F	30.3	no	AH + Inhaled CS/LABA + Topic CS	cat, dog, kiwi	303
30	IT	44	34	M	26.6	no	AH + Topic CS + SABA	--	256
31	IT	42	30	F	23.7	no	AH + Topic CS + LTRA + SABA	--	49
32	IT	42	34	F	20.8	no	AH + Topic CS + SABA	--	581
33	IT	24	3	F	22.0	no	AH + Topic CS + SABA	avocado, cat, dog	1217
34	IT	22	10	M	18.7	no	AH + Topic CS + SABA	cat, dog	778
35	IT	45	11	F	29.7	no	Inhaled CS + Topic CS + SABA	--	135

36	IT	22	7	M	24.2	no	AH + Inhaled CS/LABA + Topic CS	--	1360
37	IT	42	5	M	28.5	no	AH + Inhaled CS/LABA + Topic CS	--	687
38	IT	21	4	F	22.8	no	AH + Topic CS + SABA	--	182
39	IT	39	12	F	22.0	no	AH + Topic CS + SABA	--	149
40	IT	59	35	F	24.4	no	Inhaled CS/LABA + Topic CS + LTRA	--	165
41	IT	36	2	F	24.7	no	AH + Topic CS + SABA	--	2000
42	IT	40	27	F	22.1	no	AH + Inhaled CS/LABA + Topic CS	--	168
43	IT	23	5	F	21.9	no	AH + Topic CS + SABA	grass pollen	424
44	IT	44	25	M	29.6	no	AH + Topic CS + SABA	cat, dog	244
45	IT	60	26	F	26.2	no	AH + Inhaled CS/LABA + Topic CS + LTRA + SABA	--	219
46	IT	34	14	F	32.0	no	AH + Topic CS + SABA	molluscs	192
47	IT	40	17	F	23.6	no	AH + Topic CS + SABA	cat, dog, horse	184
48	IT	42	12	F	29.7	no	AH + Topic CS + SABA	dog, latex	292
49	IT	49	13	F	22.1	no	AH + Inhaled CS/LABA + Topic CS + LTRA + SABA	--	52
50	IT	22	4	M	20.5	no	AH + Topic CS + LTRA + SABA	cat, dog, shellfish	1502
51	IT	41	5	M	24.8	no	AH + Topic CS + SABA	cat, dog	220
52	IT	31	8	F	20.1	no	AH + Inhaled CS/LABA + Topic CS + SABA	--	178
53	IT	22	4	F	29.4	no	AH + Inhaled CS/LABA + Topic CS + SABA	cat, dog	61
54	IT	41	18	F	24.3	no	AH + Topic CS + SABA	cat, dog, bunny	660
55	IT	24	15	F	21.2	no	AH + Inhaled CS/LABA + Topic CS	cat, dog	96

56	IT	33	7	M	27.8	no	AH + Inhaled CS/LABA + Topic CS	cat, dog	77
57	IT	38	5	F	27.8	no	AH + SABA	--	532
58	IT	37	6	F	31.2	no	AH + Inhaled CS/LABA + Topic CS	cat, dog	154
59	IT	44	11	F	29.3	ex-smoker	AH + Topic CS + SABA	cat, dog	487
60	Bio	43	5	M	28.7	ex-smoker	Inhaled CS/LABA + LTRA + SABA	dog	39
61	Bio	40	6	F	20.2	no	Inhaled CS/LABA + LTRA + SABA	--	358
62	Bio	61	11	F	26.9	no	Inhaled CS/LABA + SABA	--	93
63	Bio	49	14	F	34.0	no	Inhaled CS/LABA + LTRA + SABA	cat, dog	170
64	Bio	38	--	F	25.6	no	AC + Inhaled CS/LABA + LTRA + SABA	--	--
65	Bio	55	20	F	31.5	no	AH + Inhaled CS/LABA + Topic CS + LTRA + SABA	cat, dog	306
66	Bio	47	24	F	30.5	no	Inhaled CS/LABA + Topic CS + LTRA + SABA	--	2000
67	Bio	40	23	F	33.9	no	AC + Inhaled CS/LABA + Topic CS + LTRA + SABA	cat, shellfish	281
68	Bio	32	4	F	19.6	no	Inhaled CS/LABA + Topic CS	cat, dog, grass pollen	375
69	Bio	37	6	F	26.7	yes	AC + AH + Inhaled CS/LABA + Topic CS + LTRA + SABA	dog	606
70	Bio	22	14	F	23.9	no	AC + AH + Inhaled CS/LABA + Topic CS + LTRA + SABA	dog, horse	321
71	Bio	39	11	F	23.3	no	AC + Inhaled CS/LABA + Topic CS + LTRA + SABA	--	43
72	Bio	44	12	F	21.3	no	AC + Inhaled CS/LABA + Topic CS + LTRA + SABA	dog, seafood	2000
73	Bio	54	21	F	24.9	no	AH + Inhaled CS/LABA + Topic CS + SABA	--	210
74	Bio	39	5	F	28.8	no	AH + Inhaled CS/LABA + Topic CS + SABA	dog	386

75	Bio	53	6	M	28.7	no	AC + AH + Inhaled CS/LABA + Topic CS + LTRA + SABA	cat, dog	370
76	Uncontrolled	54	5	M	31.6	no	AC + Inhaled CS/LABA + LTRA + SABA	α -amylase, flour	12
77	Uncontrolled	53	15	M	22.7	no	AC + Inhaled CS/LABA + Oral CS + LTRA + SABA	crustaceans	679
78	Uncontrolled	70	20	M	27.1	ex-smoker	AC + Inhaled CS/LABA + LTRA + SABA	--	176
79	Uncontrolled	64	15	F	27.6	no	Inhaled CS/LABA + S + SABA + T	--	160
80	Uncontrolled	44	12	M	23.1	no	AC + Inhaled CS/LABA + LTRA + SABA	--	202
81	Uncontrolled	36	4	F	26.8	no	AH + Inhaled CS/LABA + Topic CS + LTRA + SABA	--	2000
82	Uncontrolled	41	9	F	36.4	no	AH + Inhaled CS/LABA + LTRA + SABA	--	245
83	Uncontrolled	43	4	F	28.6	no	AC + Inhaled CS/LABA + LTRA + SABA	peach	647
84	Uncontrolled	24	2	F	23.9	no	AH + Inhaled CS/LABA + Topic CS + SABA	dog, wheat	1585
85	Uncontrolled	48	5	F	29.5	ex-smoker	AH + Inhaled CS/LABA + Topic CS + SABA	dog	455
86	Uncontrolled	56	18	F	30.1	no	AH + Inhaled CS/LABA + Topic CS + LTRA + SABA	--	45
87	Uncontrolled	49	34	M	39.2	ex-smoker	AC + Inhaled CS/LABA + Oral CS + SABA	--	230

Legend: *BMI*: Body mass index; *U*: ISAC units; *AC*: anticholinergic; *AH*: antihistaminic; *CS*: corticosteroids; *LABA*: Long-acting beta-adrenoceptor agonist; *LTRA*: antileukotriene; *SABA*: short-acting beta-adrenoceptor agonist; *T*: theophylline; *ICS*: Patients controlled with inhaled or topic corticosteroids without the need for systemic corticosteroids; *IT*: Patients controlled with immunotherapy, *BIO*: Patients controlled with Omalizumab.

Table S 2. Results from ANCOVA analysis between ICS and UC groups considering age as a covariable

No	Metabolite	Corrected Model (<i>p</i> value)	Intercept (<i>p</i> value)	Age (<i>p</i> value)	Group (<i>p</i> value)	No	Metabolite	Corrected Model (<i>p</i> value)	Intercept (<i>p</i> value)	Age (<i>p</i> value)	Group (<i>p</i> value)
1	Arachidonic Acid	0.104	0.120	0.456	0.222	26	LPC 18:0	0.072	0.001	0.558	0.034
2	Arachidonic Acid	0.085	0.184	0.302	0.295	27	LPC 18:0	0.080	0.001	0.379	0.030
3	Bilirubin	0.069	0.006	0.326	0.234	28	LPC 18:0	0.072	0.001	0.341	0.027
4	Cholic Acid Glucuronide OR Isomers	0.138	0.125	0.674	0.176	29	LPC 18:0	0.111	0.002	0.521	0.049
5	Deoxycholic acid OR isomers	0.050	0.053	0.351	0.171	30	LPC 18:0	0.096	0.001	0.582	0.046
6	Deoxycholic acid OR isomers	0.049	0.053	0.373	0.157	31	LPC 18:0	0.069	0.001	0.545	0.032
7	Hippuric acid	0.020	0.330	0.883	0.180	32	LPC 18:1	0.015	0.071	0.253	0.101
8	LPC 16:0	0.156	0.001	0.076	0.169	33	LPC 18:1	0.007	0.074	0.181	0.082
9	LPC 16:0	0.161	0.007	0.856	0.149	34	LPC 18:1	0.022	0.030	0.304	0.109
10	LPC 16:0	0.178	0.002	0.960	0.140	35	LPC 18:1	0.013	0.112	0.204	0.114
11	LPC 16:0	0.137	0.001	0.979	0.103	36	LPC 18:1	0.012	0.110	0.201	0.110
12	LPC 16:0	0.234	0.006	0.827	0.213	37	LPC 18:1	0.022	0.091	0.257	0.129
13	LPC 16:0	0.166	0.001	0.795	0.168	38	LPC 18:1	0.022	0.087	0.259	0.128
14	LPC 16:0	0.158	0.005	0.805	0.158	39	LPC 18:1	0.015	0.107	0.214	0.122
15	LPC 16:0	0.241	0.002	0.895	0.199	40	LPC 18:1	0.019	0.074	0.238	0.129
16	LPC 16:1	0.016	0.014	0.594	0.038	41	LPC 18:1	0.023	0.053	0.213	0.165
17	LPC 16:1	0.009	0.045	0.385	0.040	42	LPC 18:1	0.057	0.058	0.309	0.215
18	LPC 16:1	0.008	0.047	0.346	0.043	43	LPC 18:3	0.107	0.022	0.687	0.057
19	LPC 16:1	0.008	0.047	0.376	0.040	44	LPC 18:3	0.016	0.092	0.990	0.018
20	LPC 16:1	0.009	0.044	0.391	0.041	45	LPC 20:2	0.191	0.013	0.483	0.080
21	LPC 16:1	0.008	0.045	0.386	0.037	46	LPC 20:3	0.029	0.022	0.923	0.026
22	LPC 16:1	0.008	0.050	0.332	0.045	47	LPC 20:3	0.047	0.042	0.940	0.048
23	LPC 16:1	0.009	0.041	0.404	0.039	48	LPC 20:3	0.023	0.031	0.910	0.028
24	LPC 17:0	0.082	0.004	0.448	0.034	49	LPC 20:3	0.001	0.623	0.131	0.014
25	LPC 17:1	0.005	0.004	0.739	0.004	50	LPC 20:3	0.045	0.034	0.990	0.041
No	Metabolite	Corrected Model (<i>p</i>	Intercept (<i>p</i> value)	Age (<i>p</i>	Group (<i>p</i> value)	No	Metabolite	Corrected Model (<i>p</i>	Intercept (<i>p</i> value)	Age (<i>p</i> value)	Group (<i>p</i> value)

		value)		value)				value)			
51	LPC 20:4	0.017	0.044	0.685	0.033	71	LPC P-18:1	0.219	0.049	0.794	0.212
52	LPC 20:4	0.015	0.128	0.389	0.062	72	LPE 18:0	0.102	0.251	0.393	0.256
53	LPC 20:4	0.005	0.099	0.302	0.037	73	LPE 20:3	0.081	0.006	0.653	0.043
54	LPC 20:4	0.007	0.123	0.264	0.054	74	LPE 20:4	0.009	0.022	0.366	0.045
55	LPC 20:4	0.005	0.044	0.336	0.030	75	LPE 20:5	0.010	0.456	0.320	0.055
56	LPC 20:4	0.056	0.004	0.751	0.075	76	LPE 22:5	0.053	0.171	0.320	0.195
57	LPC 20:4	0.012	0.131	0.355	0.059	77	LPE 22:6	0.047	0.467	0.097	0.554
58	LPC 20:4	0.014	0.110	0.311	0.075	78	LPI 16:0	0.489	0.354	0.697	0.494
59	LPC 20:4	0.016	0.125	0.305	0.087	79	LPI 20:4	0.016	0.145	0.271	0.096
60	LPC 20:5	0.005	0.854	0.172	0.071	80	PC 16:0/20:5	0.150	0.016	0.174	0.719
61	LPC 20:5	0.003	0.711	0.163	0.045	81	Sphingosine Phosphate	0.078	0.001	0.890	0.078
62	LPC 20:5	0.005	0.679	0.203	0.053	82	L-Arginine	0.274	0.001	0.516	0.113
63	LPC 22:4	0.023	0.070	0.740	0.037	83	L-Leucine	0.051	0.006	0.125	0.315
64	LPC 22:5	0.032	0.168	0.471	0.087	84	LPE 20:4	0.047	0.010	0.219	0.169
65	LPC 22:5	0.259	0.081	0.905	0.165	85	LPE 20:4	0.023	0.008	0.200	0.100
66	LPC 22:5	0.061	0.301	0.503	0.132	86	Retinol	0.163	0.004	0.422	0.251
67	LPC 22:5	0.063	0.100	0.839	0.071	87	Palmitoleic Acid	0.001	0.346	0.007	0.369
68	LPC 22:6	0.098	0.376	0.294	0.335	88	Palmitoleyl lynoleate (18:2/16:1)	0.010	0.001	0.046	0.384
69	LPC 22:6	0.208	0.546	0.303	0.558	89	Phe-Ile OR Phe- Leu	0.015	0.001	0.046	0.005
70	LPC 22:6	0.220	0.505	0.325	0.545						

Legend: LPC: lysophosphocholine; LPE: lysophosphoethanolamine; LPI: lysophosphatidylinositol; PC: Phosphatidilcholine.

Table S3. Physicochemical characteristics of significant annotated signals in metabolomics through LC-MS/MS between UC and ICS groups.

No.	Technique	Biological category	Compound	<i>m/z</i>	Mass (Da)	RT (min)	Formula	Error (ppm)	Adduct	CV in QCs (%)
1	LC-MS-	Fatty Acids	Arachidonic Acid ‡	371.2198	304.2402	27.79	C20H32O2	-0.1	M+HCOONa-H	17.2
2	LC-MS-	Fatty Acids	Arachidonic Acid ‡	303.2328	304.2406	27.79	C20H32O2	1.3	M-H	9.1
3	LC-MS-	Others	Bilirubin ‡	583.255	584.2628	15.77	C33H36N4O6	-1.1	M-H	10.3
4	LC-MS-	Bile Acids	Cholic Acid Glucuronide	565.3008	584.3192	6.86	C30H48O11	-0.8	M-H-H2O	7.3
5	LC-MS-	Others	Hippuric acid ‡	178.0533	179.0611	1.04	C9H9NO3	16.1	M-H	5.9
6	LC-MS-	Phospholipids	LPC 16:0 ‡	584.3557	495.3318	18.37	C24H50NO7P	-1.3	M+C3H6O3-H	8.3
7	LC-MS-	Phospholipids	LPC 16:0 ‡	480.309	495.3325	19.16	C24H50NO7P	0	M-CH3-H	5.5
8	LC-MS-	Phospholipids	LPC 16:0 ‡	530.3025	495.3336	18.38	C24H50NO7P	2.3	M+Cl	6.3
9	LC-MS-	Phospholipids	LPC 16:0 ‡	540.3306	495.3329	18.38	C24H50NO7P	0.9	M+FA-H	7.4
10	LC-MS-	Phospholipids	LPC 16:0 ‡	608.317	495.3320	18.38	C24H50NO7P	-1.1	M+TFA-H	14.9
11	LC-MS-	Phospholipids	LPC 16:0 ‡	480.3089	495.3324	18.37	C24H50NO7P	-0.2	M-CH3-H	8.0
12	LC-MS-	Phospholipids	LPC 16:0 ‡	540.3306	495.3329	19.16	C24H50NO7P	0.9	M+FA-H	5.6
13	LC-MS-	Phospholipids	LPC 16:0 ‡	530.3032	495.3343	19.16	C24H50NO7P	3.8	M+Cl	5.5
14	LC-MS-	Phospholipids	LPC 16:1	538.3145	493.3168	16.27	C24H48NO7P	0	M+FA-H	8.5
15	LC-MS-	Phospholipids	LPC 16:1	538.3145	493.3168	16.27	C24H48NO7P	0	M+FA-H	8.9
16	LC-MS-	Phospholipids	LPC 16:1	538.3145	493.3168	16.27	C24H48NO7P	0	M+FA-H	8.6
17	LC-MS-	Phospholipids	LPC 16:1	538.3145	493.3168	16.27	C24H48NO7P	0	M+FA-H	8.6
18	LC-MS-	Phospholipids	LPC 16:1	538.3143	493.3166	15.58	C24H48NO7P	-0.4	M+FA-H	8.5
19	LC-MS-	Phospholipids	LPC 16:1	538.3143	493.3166	15.58	C24H48NO7P	-0.4	M+FA-H	8.7
20	LC-MS-	Phospholipids	LPC 16:1	538.3145	493.3168	16.27	C24H48NO7P	0	M+FA-H	8.5
21	LC-MS-	Phospholipids	LPC 16:1	528.2877	493.3188	16.27	C24H48NO7P	4.1	M-Cl	9.3
22	LC-MS-	Phospholipids	LPC 17:0 ‡	554.3458	509.3481	21.24	C25H52NO7P	0	M+FA-H	8.1
23	LC-MS-	Phospholipids	LPC 17:1	552.3292	507.3315	18.16	C25H50NO7P	-1.9	M+FA-H	15.5
24	LC-MS-	Phospholipids	LPC 18:0 ‡	508.3402	523.3637	23.40	C26H54NO7P	-0.2	M-CH3-H	7.3
25	LC-MS-	Phospholipids	LPC 18:0 ‡	568.3615	523.3638	22.58	C26H54NO7P	0.1	M+FA-H	8.6
26	LC-MS-	Phospholipids	LPC 18:0 ‡	560.3306	523.3461	23.41	C26H54NO7P	-33.8	M+Cl	5.1
27	LC-MS-	Phospholipids	LPC 18:0 ‡	568.3619	523.3642	23.40	C26H54NO7P	0.9	M+FA-H	7.8
28	LC-MS-	Phospholipids	LPC 18:0 ‡	558.3314	523.3625	22.58	C26H54NO7P	-2.4	M+Cl	8.2

29	LC-MS-	Phospholipids	LPC 18:0‡	612.387	523.3631	23.40	C26H54NO7P	-1.3	M+C3H6O3-H	7.2
30	LC-MS-	Phospholipids	LPC 18:1 ‡	506.3246	521.3481	19.43	C26H52NO7P	-0.1	M-CH3-H	11.1
31	LC-MS-	Phospholipids	LPC 18:1 ‡	566..3458	521..3481	19..43	C26H52NO7P	0	M+FA-H	11..5
32	LC-MS-	Phospholipids	LPC 18:1 ‡	566..3458	521..3481	19..43	C26H52NO7P	0	M+FA-H	11..5
33	LC-MS-	Phospholipids	LPC 18:1 ‡	566..3463	521.3486	20.13	C26H52NO7P	1	M+FA-H	10.8
34	LC-MS-	Phospholipids	LPC 18:1 ‡	610.3711	521.3472	20.13	C26H52NO7P	-1.7	M+C3H6O3-H	10.4
35	LC-MS-	Phospholipids	LPC 18:1 ‡	506.3244	521.3479	20.13	C26H52NO7P	-0.5	M-CH3-H	12.7
36	LC-MS-	Phospholipids	LPC 18:1 ‡	566.3463	521.3486	20.13	C26H52NO7P	1	M+FA-H	10.6
37	LC-MS-	Phospholipids	LPC 18:1 ‡	566.3463	521.3486	20.13	C26H52NO7P	1	M+FA-H	10.5
38	LC-MS-	Phospholipids	LPC 18:1 ‡	556.3169	521.3480	20.13	C26H52NO7P	-0.2	M+Cl	11.1
39	LC-MS-	Phospholipids	LPC 18:1 ‡	634.3326	521.3476	20.13	C26H52NO7P	-1.1	M+TFA-H	18.0
40	LC-MS-	Phospholipids	LPC 18:1 ‡	634.3336	521.3486	19.43	C26H52NO7P	0.8	M+TFA-H	14.8
41	LC-MS-	Phospholipids	LPC 18:3	562.3139	517.3162	15.97	C26H48NO7P	-1.1	M+FA-H	12.6
42	LC-MS-	Phospholipids	LPC 18:3	562.3137	517.3160	15.63	C26H48NO7P	-1.5	M+FA-H	17.7
43	LC-MS-	Phospholipids	LPC 20:2	592.361	547.3633	21.19	C28H54NO7P	-0.8	M+FA-H	18.4
44	LC-MS-	Phospholipids	LPC 20:3	590.345	545.3473	20.03	C28H52NO7P	-1.5	M+FA-H	15.7
45	LC-MS-	Phospholipids	LPC 20:3	590.3453	545.3476	18.47	C28H52NO7P	-0.9	M+FA-H	19.9
46	LC-MS-	Phospholipids	LPC 20:3	590.3453	545.3476	18.47	C28H52NO7P	-0.9	M+FA-H	18.9
47	LC-MS-	Phospholipids	LPC 20:3	590.3459	545.3482	19.06	C28H52NO7P	0.2	M+FA-H	18.3
48	LC-MS-	Phospholipids	LPC 20:3	590.3459	545.3482	19.06	C28H52NO7P	0.2	M+FA-H	19.1
49	LC-MS-	Phospholipids	LPC 20:4	588.3305	543.3328	17.70	C28H50NO7P	0.7	M+FA-H	14.4
50	LC-MS-	Phospholipids	LPC 20:4	588.3301	543.3324	17.13	C28H50NO7P	-0.1	M+FA-H	16.6
51	LC-MS-	Phospholipids	LPC 20:4	578.301	543.3321	17.13	C28H50NO7P	-0.6	M+Cl	15.5
52	LC-MS-	Phospholipids	LPC 20:4	588.3301	543.3324	17.13	C28H50NO7P	-0.1	M+FA-H	15.4
53	LC-MS-	Phospholipids	LPC 20:4	656.3182	543.3332	17.13	C28H50NO7P	1.2	M+TFA-H	27.1
54	LC-MS-	Phospholipids	LPC 20:4	528.3086	543.3321	17.71	C28H50NO7P	-0.8	M-CH3-H	13.6
55	LC-MS-	Phospholipids	LPC 20:4	656.3182	543.3332	17.13	C28H50NO7P	1.2	M+TFA-H	27.4
56	LC-MS-	Phospholipids	LPC 20:4	578.301	543.3321	17.71	C28H50NO7P	-0.6	M+Cl	10.7
57	LC-MS-	Phospholipids	LPC 20:4	580.3009	543.3164	17.70	C28H50NO7P	-29.6	M+Cl	11.2
58	LC-MS-	Phospholipids	LPC 20:5	586.3143	541.3166	15.78	C28H48NO7P	-0.4	M+FA-H	14.5
59	LC-MS-	Phospholipids	LPC 20:5	586.3145	541.3168	15.23	C28H48NO7P	0	M+FA-H	10.8
60	LC-MS-	Phospholipids	LPC 20:5	586.3145	541.3168	15.23	C28H48NO7P	0	M+FA-H	11.7

61	LC-MS-	Phospholipids	LPC 22:4	616.3609	571.3632	20.62	C30H54NO7P	-0.9	M+FA-H	21.3
62	LC-MS-	Phospholipids	LPC 22:5	614.3462	569.3485	18.59	C30H52NO7P	0.7	M+FA-H	21.4
63	LC-MS-	Phospholipids	LPC 22:5	614.3483	569.3506	18.59	C30H52NO7P	4.4	M+FA-H	20.9
64	LC-MS-	Phospholipids	LPC 22:5	614.3435	569.3458	18.04	C30H52NO7P	-4	M+FA-H	18.7
65	LC-MS-	Phospholipids	LPC 22:5	614.347	569.3493	19.61	C30H52NO7P	2.1	M+FA-H	20.1
66	LC-MS-	Phospholipids	LPC 22:6	612.33	567.3323	17.63	C30H50NO7P	-0.3	M+FA-H	14.1
67	LC-MS-	Phospholipids	LPC 22:6	612.3297	567.3320	17.13	C30H50NO7P	-0.8	M+FA-H	16.6
68	LC-MS-	Phospholipids	LPC 22:6	612.3297	567.3320	17.13	C30H50NO7P	-0.8	M+FA-H	16.0
69	LC-MS-	Phospholipids	LPC P-18:1	550.3506	505.3529	21.20	C26H52NO6P	-0.5	M+FA-H	12.8
70	LC-MS-	Phospholipids	LPE 18:0 ‡	462.2984	481.3168	21.04	C23H48NO7P	-0.1	M-H-H2O	21.4
71	LC-MS-	Phospholipids	LPE 20:3	502.293	503.3008	18.93	C25H46NO7P	-0.7	M-H	10.2
72	LC-MS-	Phospholipids	LPE 20:4	500.2778	501.2856	17.54	C25H44NO7P	0.2	M-H	6.4
73	LC-MS-	Phospholipids	LPE 20:5	498.262	499.2698	15.65	C25H42NO7P	-0.1	M-H	6.1
74	LC-MS-	Phospholipids	LPE 22:5	526.2929	527.3007	18.46	C27H46NO7P	-0.9	M-H	12.0
75	LC-MS-	Phospholipids	LPE 22:6	524.2775	525.2853	17.49	C27H44NO7P	-0.4	M-H	5.9
76	LC-MS-	Phospholipids	LPI 16:0 ‡	571.2892	572.2970	22.67	C25H49O12P	1.5	M-H	21.3
77	LC-MS-	Phospholipids	LPI 20:4	619.288	620.2958	20.50	C29H49O12P	-0.6	M-H	17.9
78	LC-MS-	Bile Acids	Deoxycholic acid OR isomers ‡	437.2901	392.2924	14.73	C24H40O4	-0.5	M+FA-H	5.4
79	LC-MS-	Bile Acids	Deoxycholic acid OR isomers ‡	391.2848	392.2926	14.73	C24H40O4	-0.1	M-H	6.7
80	LC-MS-	Phospholipids	PC 16:0/20:5	824.5453	779.5476	34.14	C44H78NO8P	1.5	M+FA-H	19.7
81	LC-MS-	Sphingolipids	Sphingosine-1-Phosphate ‡	378.2412	379.2490	14.85	C18H38NO5P	0.7	M-H	5.9
82	LC-MS+	Aminoacids	L-Arginine ‡	175.1186	174.1108	0.75	C6H14N4O2	-5.2	M+H	16.9
83	LC-MS+	Aminoacids	L-Leucine ‡	132.1006	131.0928	0.94	C6H13NO2	-14.1	M+H	9.8
84	LC-MS+	Phospholipids	LPE 20:4	502.2947	501.2869	17.58	C25H44NO7P	2.7	M+H	6.4
85	LC-MS+	Phospholipids	LPE 20:4	502.2947	501.2869	17.58	C25H44NO7P	2.7	M+H	6.8
86	LC-MS+	Vitamins	Retinol	269.2266	286.2293	27.26	C20H30O	-1.1	M+H-H2O	6.4

Legend: LPC: lysophosphocholine; LPE: lysophosphoethanolamine; LPI: lysophosphatidylinositol; PC: Phosphatidilcholine; FA: Formic Acid; TFA: Trifluoroacetic acid; CV: Coefficient of variation; QC: Quality control. The coefficient of variation (CV) in the injections of the quality control sample (QC), assures that the metabolites were measured throughout the experiment in a constant and reproducible manner. In this sense, if the variations in QCs are lower than 30%, it means that the variations found in a particular metabolite are due to the true increase or decrease between the comparison groups. Thus, the % of CV in QCs should be always less than the % change between groups for each metabolite. In this way, we assessed that the results of this experiment are reliable.

Table S4. Pairwise comparisons showing the significance and % of change among experimental groups. Significant differences are marked bold, while the percentage of changes for significant metabolites is marked in bold red for increasing metabolites or bold blue for decreasing metabolites.

No	Compound	Mass (Da)	RT (min)	ICS vs UC			ICS vs IT			ICS vs BIO			IT vs BIO			IT vs UC			BIO vs UC		
				% change (UC)	p value	p value (BH)	% change (IT)	p value	p value (BH)	% change (Bio)	p value	p value (BH)	% change (Bio)	p value	p value (BH)	% change (UC)	p value	p value (BH)	% change (UC)	p value	p value (BH)
1	Arachidonic Acid ‡	304.2402	27.79	45.4	0.0281	0.1352	17.4	0.1389	0.7152	1.8	0.4484	0.8998	-13.3	0.4303	0.9686	23.9	0.1131	0.4013	42.9	0.0177	0.4079
2	Arachidonic Acid ‡	304.2406	27.79	46.7	0.0326	0.1415	18.6	0.1277	0.7078	0.4	0.5351	0.9301	-15.4	0.3047	0.9686	23.7	0.1015	0.3785	46.1	0.0242	0.4079
3	Bilirubin ‡	584.2628	15.77	-37.6	0.0376	0.1491	19.1	0.9397	0.9908	-7.7	0.4763	0.9180	-22.5	0.3363	0.9686	-47.6	0.0342	0.3084	-32.4	0.2787	0.5898
4	Cholic Acid Glucuronide	584.3192	6.86	54.7	0.0177	0.1097	29.6	0.0820	0.7078	61.3	0.0041	0.8682	24.5	0.0118	0.9686	19.3	0.0604	0.3288	-4.1	0.6187	0.8110
5	Hippuric acid ‡	179.0611	1.04	87.5	0.0077	0.0864	129.3	0.0087	0.7078	192.2	0.0695	0.8682	27.5	0.5669	0.9686	-18.2	0.7192	0.8709	-35.8	0.7922	0.9040
6	LPC 16:0 ‡	495.3318	18.37	28.4	0.0242	0.1256	4.2	0.8543	0.9488	7.3	0.8005	0.9767	3.0	0.8883	0.9810	23.2	0.0039	0.1871	19.6	0.0207	0.4079
7	LPC 16:0 ‡	495.3325	19.16	30.6	0.0242	0.1256	9.6	0.3363	0.7736	5.0	0.5972	0.9394	-4.3	0.7377	0.9686	19.1	0.0503	0.3288	24.4	0.0242	0.4079
8	LPC 16:0 ‡	495.3336	18.38	26.1	0.0376	0.1491	8.4	0.4303	0.7976	1.7	1.0000	1.0000	-6.1	0.2848	0.9686	16.3	0.0569	0.3288	23.9	0.0065	0.4079
9	LPC 16:0 ‡	495.3329	18.38	25.8	0.0177	0.1097	8.1	0.2390	0.7734	2.8	0.7304	0.9596	-4.9	0.3363	0.9686	16.4	0.0262	0.3046	22.4	0.0177	0.4079
10	LPC 16:0 ‡	495.3320	18.38	22.6	0.0376	0.1491	7.3	0.4178	0.7895	6.3	0.6961	0.9510	-1.0	0.6578	0.9686	14.3	0.0390	0.3199	15.4	0.0434	0.4079
11	LPC 16:0 ‡	495.3324	18.37	21.8	0.0151	0.1034	6.1	0.2848	0.7734	5.1	0.6625	0.9497	-0.9	0.4692	0.9686	14.8	0.0366	0.3084	15.9	0.0376	0.4079
12	LPC 16:0 ‡	495.3329	19.16	21.6	0.0207	0.1160	7.4	0.2947	0.7734	5.3	0.5972	0.9394	-2.0	0.5669	0.9686	13.2	0.0320	0.3084	15.5	0.0434	0.4079
13	LPC 16:0 ‡	495.3343	19.16	18.6	0.0281	0.1352	9.2	0.1911	0.7466	6.2	0.7304	0.9596	-2.7	0.3815	0.9686	8.6	0.1787	0.4952	11.7	0.0570	0.4079
14	LPC 16:1	493.3168	16.27	63.1	0.0038	0.0864	9.7	0.5099	0.8493	21.7	0.2802	0.8682	11.0	0.6423	0.9686	48.7	0.0021	0.1592	34.0	0.0570	0.4079
15	LPC 16:1	493.3168	15.58	61.0	0.0092	0.0875	6.7	0.6894	0.9002	21.4	0.3462	0.8682	13.7	0.5817	0.9686	50.9	0.0013	0.1592	32.7	0.0651	0.4079
16	LPC 16:1	493.3168	16.27	62.0	0.0038	0.0864	9.1	0.5380	0.8556	21.1	0.3012	0.8682	11.0	0.6423	0.9686	48.4	0.0023	0.1592	33.8	0.0570	0.4079
17	LPC 16:1	493.3168	16.27	62.5	0.0054	0.0864	8.3	0.5669	0.8658	22.6	0.2413	0.8682	13.2	0.5524	0.9686	50.0	0.0018	0.1592	32.6	0.0841	0.4079
18	LPC 16:1	493.3166	16.27	61.6	0.0065	0.0864	8.5	0.5817	0.8658	20.7	0.3462	0.8682	11.3	0.6116	0.9686	48.9	0.0027	0.1592	33.8	0.0841	0.4079
19	LPC 16:1	493.3166	15.58	60.8	0.0065	0.0864	6.8	0.6423	0.8816	21.0	0.3232	0.8682	13.3	0.5966	0.9686	50.6	0.0012	0.1592	32.9	0.0570	0.4079
20	LPC 16:1	493.3168	16.27	60.2	0.0045	0.0864	6.0	0.7215	0.9010	21.0	0.2603	0.8682	14.2	0.5239	0.9686	51.1	0.0016	0.1592	32.3	0.0651	0.4079
21	LPC 16:1	493.3188	16.27	52.0	0.0065	0.0864	7.6	0.7377	0.9039	15.6	0.4213	0.8821	7.4	0.6894	0.9686	41.2	0.0035	0.1871	31.5	0.0570	0.4079
22	LPC 17:0 ‡	509.3481	21.24	47.9	0.0281	0.1352	19.8	0.1224	0.7078	12.1	0.4213	0.8821	-6.4	0.4303	0.9686	23.5	0.1467	0.4423	32.0	0.0570	0.4079
23	LPC 17:1	507.3315	18.16	73.9	0.0045	0.0864	28.7	0.1634	0.7250	34.6	0.1753	0.8682	4.6	1.0000	1.0000	35.1	0.0148	0.2980	29.2	0.0841	0.4079
24	LPC 18:0 ‡	523.3637	23.40	38.4	0.0281	0.1352	18.2	0.0899	0.7078	17.3	0.3232	0.8682	-0.8	0.7377	0.9686	17.1	0.1324	0.4234	18.0	0.0952	0.4196
25	LPC 18:0 ‡	523.3638	22.58	38.1	0.0281	0.1352	17.6	0.1332	0.7152	16.3	0.2802	0.8682	-1.2	0.7705	0.9686	17.4	0.1324	0.4234	18.8	0.1354	0.4582
26	LPC 18:0 ‡	523.3461	23.41	31.6	0.0376	0.1491	16.7	0.1224	0.7078	13.4	0.2603	0.8682	-2.9	0.5817	0.9686	12.8	0.2577	0.5725	16.1	0.1073	0.4250
27	LPC 18:0 ‡	523.3642	23.40	35.6	0.0326	0.1415	16.6	0.1224	0.7078	16.1	0.3703	0.8682	-0.5	0.8543	0.9810	16.3	0.1542	0.4488	16.9	0.1207	0.4311
28	LPC 18:0 ‡	523.3625	22.58	33.4	0.0326	0.1415	23.4	0.0616	0.7078	24.8	0.0849	0.8682	1.1	0.9569	0.9981	8.1	0.4971	0.7456	6.9	0.5387	0.7703
29	LPC 18:0‡	523.3631	23.40	33.9	0.0434	0.1629	12.0	0.3256	0.7734	18.6	0.2413	0.8682	5.9	0.7871	0.9747	19.6	0.0473	0.3274	12.9	0.1207	0.4311
30	LPC 18:1 ‡	521.3481	19.43	54.6	0.0077	0.0864	22.9	0.0940	0.7078	9.3	0.3703	0.8682	-11.1	0.3150	0.9686	25.9	0.1015	0.3785	41.5	0.0498	0.4079
31	LPC 18:1 ‡	521.3481	19.43	55.4	0.0065	0.0864	29.2	0.0858	0.7078	12.8	0.4213	0.8821	-12.7	0.3150	0.9686	20.3	0.0535	0.3288	37.7	0.0376	0.4079
32	LPC 18:1 ‡	521.3481	19.43	55.1	0.0077	0.0864	28.4	0.1123	0.7078	13.2	0.3703	0.8682	-11.9	0.3473	0.9686	20.7	0.0444	0.3268	37.0	0.0651	0.4079
33	LPC 18:1 ‡	521.3486	20.13	53.5	0.0038	0.0864	29.3	0.0531	0.7078	14.3	0.3703	0.8682	-11.6	0.3585	0.9686	18.8	0.0960	0.3726	34.3	0.0498	0.4079
34	LPC 18:1 ‡	521.3472	20.13	48.9	0.0177	0.1097	22.7	0.1172	0.7078	15.5	0.2603	0.8682	-5.9	0.6735	0.9686	21.3	0.0416	0.3268	28.9	0.0570	0.4079
35	LPC 18:1 ‡	521.3479	20.13	52.0	0.0065	0.0864	28.6	0.1634	0.7250	11.9	0.4213	0.8821	-13.0	0.4178	0.9686	18.2	0.0366	0.3084	35.9	0.0741	0.4079
36	LPC 18:1 ‡	521.3486	20.13	51.3	0.0054	0.0864	27.2	0.0712	0.7078	13.3	0.3703	0.8682	-10.9	0.3815	0.9686	18.9	0.0811	0.3549	33.5	0.0651	0.4079
37	LPC 18:1 ‡	521.3486	20.13	51.0	0.0054	0.0864	27.1	0.0712	0.7078	13.2	0.3953	0.8682	-10.9	0.3363	0.9686	18.9	0.0960	0.3726	33.4	0.0570	0.4079
38	LPC 18:1 ‡	521.3480	20.13	43.7	0.0109	0.0916	27.1	0.0858	0.7078	15.2	0.3953	0.8682	-9.3	0.3150	0.9686	13.1	0.1394	0.4339	24.7	0.0651	0.4079
39	LPC 18:1 ‡	521.3476	20.13	42.2	0.0128	0.0969	20.4	0.1277	0.7078	11.5	0.3232	0.8682	-7.4	0.4178	0.9686	18.1	0.0858	0.3595	27.5	0.0570	0.4079
40	LPC 18:1 ‡	521.3486	19.43	39.2	0.0376	0.1491	17.6	0.3363	0.7736	4.8	0.7652	0.9703	-10.9	0.4303	0.9686	18.3	0.0681	0.3504	32.8	0.0376	0.4079

41	LPC 18:3	517.3162	15.97	57.7	0.0207	0.1160	6.4	0.5966	0.8658	14.9	0.3953	0.8682	8.0	0.4055	0.9686	48.3	0.0063	0.2306	37.3	0.0376	0.4079
42	LPC 18:3	517.3160	15.63	124.5	0.0109	0.0916	29.9	0.4055	0.7895	41.2	0.3012	0.8682	8.7	0.6735	0.9686	72.8	0.0074	0.2406	59.0	0.0498	0.4079
43	LPC 20:2	547.3633	21.19	45.1	0.0498	0.1768	27.7	0.1277	0.7078	5.2	0.6961	0.9510	-17.6	0.2752	0.9686	13.6	0.3445	0.6467	37.9	0.0952	0.4196
44	LPC 20:3	545.3473	20.03	173.9	0.0009	0.0864	73.2	0.0434	0.7078	73.1	0.0229	0.8682	0.0	0.4961	0.9686	58.2	0.0262	0.3046	58.2	0.0741	0.4079
45	LPC 20:3	545.3476	18.47	67.1	0.0065	0.0864	32.3	0.1701	0.7250	24.2	0.3012	0.8682	-6.1	0.8205	0.9810	26.4	0.0960	0.3726	34.6	0.0952	0.4196
46	LPC 20:3	545.3476	18.47	67.6	0.0092	0.0875	31.3	0.1701	0.7250	25.2	0.3012	0.8682	-4.6	0.9054	0.9924	27.7	0.0722	0.3504	33.8	0.0841	0.4079
47	LPC 20:3	545.3482	19.06	69.3	0.0128	0.0969	36.9	0.1277	0.7078	29.5	0.2802	0.8682	-5.4	0.7871	0.9747	23.7	0.1131	0.4013	30.7	0.1514	0.4923
48	LPC 20:3	545.3482	19.06	71.2	0.0077	0.0864	34.8	0.1447	0.7199	24.6	0.3462	0.8682	-7.6	0.6894	0.9686	27.0	0.0811	0.3549	37.4	0.0841	0.4079
49	LPC 20:4	543.3328	17.70	74.2	0.0031	0.0864	37.1	0.0456	0.7078	22.7	0.2064	0.8682	-10.5	0.4431	0.9686	27.1	0.0444	0.3268	42.0	0.0498	0.4079
50	LPC 20:4	543.3324	17.13	73.4	0.0109	0.0916	31.7	0.0820	0.7078	19.4	0.3232	0.8682	-9.4	0.5239	0.9686	31.6	0.0342	0.3084	45.2	0.0376	0.4079
51	LPC 20:4	543.3321	17.13	69.9	0.0109	0.0916	26.5	0.0983	0.7078	13.2	0.3703	0.8682	-10.6	0.3699	0.9686	34.3	0.0244	0.3046	50.1	0.0092	0.4079
52	LPC 20:4	543.3324	17.13	72.6	0.0109	0.0916	30.8	0.0858	0.7078	20.0	0.2802	0.8682	-8.3	0.6578	0.9686	32.0	0.0300	0.3084	43.8	0.0651	0.4079
53	LPC 20:4	543.3332	17.13	64.2	0.0092	0.0875	27.6	0.0899	0.7078	15.2	0.3232	0.8682	-9.7	0.4825	0.9686	28.7	0.0228	0.3046	42.6	0.0128	0.4079
54	LPC 20:4	543.3321	17.71	70.6	0.0045	0.0864	35.4	0.0371	0.7078	21.9	0.1904	0.8682	-9.9	0.4178	0.9686	26.0	0.0569	0.3288	39.9	0.0281	0.4079
55	LPC 20:4	543.3332	17.13	63.6	0.0092	0.0875	27.2	0.0820	0.7078	15.0	0.3232	0.8682	-9.6	0.5239	0.9686	28.5	0.0366	0.3084	42.2	0.0128	0.4079
56	LPC 20:4	543.3321	17.71	65.1	0.0045	0.0864	29.7	0.0333	0.7078	20.9	0.1478	0.8682	-6.8	0.5966	0.9686	27.3	0.0366	0.3084	36.6	0.0207	0.4079
57	LPC 20:4	543.3164	17.70	36.4	0.0109	0.0916	16.2	0.2222	0.7734	8.0	0.5657	0.9301	-7.1	0.4961	0.9686	17.4	0.0766	0.3549	26.3	0.0326	0.4079
58	LPC 20:5	541.3166	15.78	168.4	0.0012	0.0864	91.4	0.0940	0.7078	62.2	0.0935	0.8682	-15.3	0.9569	0.9981	40.2	0.0171	0.2980	65.5	0.0151	0.4079
59	LPC 20:5	541.3168	15.23	155.2	0.0009	0.0864	87.8	0.0531	0.7078	75.2	0.1129	0.8682	-6.7	0.9569	0.9981	35.9	0.0184	0.2980	45.7	0.0281	0.4079
60	LPC 20:5	541.3168	15.23	154.0	0.0021	0.0864	76.6	0.1332	0.7152	65.6	0.1611	0.8682	-6.3	0.8543	0.9810	43.8	0.0171	0.2980	53.4	0.0281	0.4079
61	LPC 22:4	571.3632	20.62	79.4	0.0128	0.0969	27.6	0.1769	0.7250	23.1	0.3232	0.8682	-3.5	0.9225	0.9943	40.6	0.0244	0.3046	45.7	0.0651	0.4079
62	LPC 22:5	569.3485	18.59	85.3	0.0109	0.0916	29.6	0.2062	0.7597	22.6	0.1753	0.8682	-5.4	0.9569	0.9981	43.0	0.0228	0.3046	51.1	0.1073	0.4250
63	LPC 22:5	569.3506	18.59	76.0	0.0151	0.1034	28.7	0.2062	0.7597	17.3	0.3012	0.8682	-8.9	0.8543	0.9810	36.8	0.0159	0.2980	50.1	0.0741	0.4079
64	LPC 22:5	569.3458	18.04	74.4	0.0077	0.0864	28.0	0.1701	0.7250	20.8	0.1354	0.8682	-5.6	0.9397	0.9981	36.3	0.0148	0.2980	44.3	0.0741	0.4079
65	LPC 22:5	569.3493	19.61	53.0	0.0281	0.1352	41.0	0.0559	0.7078	35.1	0.2413	0.8682	-4.2	0.5239	0.9686	8.5	0.2467	0.5586	13.2	0.0651	0.4079
66	LPC 22:6	567.3323	17.63	47.1	0.0151	0.1034	22.4	0.0456	0.7078	3.0	0.3232	0.8682	-15.8	0.4055	0.9686	20.1	0.0722	0.3504	42.7	0.0741	0.4079
67	LPC 22:6	567.3320	17.13	55.5	0.0151	0.1034	30.0	0.0616	0.7078	11.8	0.3953	0.8682	-14.0	0.3473	0.9686	19.6	0.0908	0.3722	39.1	0.0841	0.4079
68	LPC 22:6	567.3320	17.13	48.1	0.0128	0.0969	23.5	0.0782	0.7078	3.6	0.4213	0.8821	-16.1	0.3815	0.9686	20.0	0.1015	0.3785	43.0	0.0841	0.4079
69	LPC P-18:1	505.3529	21.20	38.3	0.0281	0.1352	24.3	0.1570	0.7250	11.4	0.5657	0.9301	-10.4	0.5966	0.9686	11.3	0.2691	0.5887	24.2	0.2787	0.5898
70	LPE 18:0 ‡	481.3168	21.04	55.4	0.0326	0.1415	50.5	0.0480	0.7078	13.4	0.5657	0.9301	-24.7	0.3363	0.9686	3.2	0.4804	0.7299	37.1	0.2081	0.5479
71	LPE 20:3	503.3008	18.93	47.1	0.0177	0.1097	10.5	0.7871	0.9236	25.7	0.5053	0.9301	13.8	0.7705	0.9686	33.1	0.0148	0.2980	17.0	0.1207	0.4311
72	LPE 20:4	501.2856	17.54	52.6	0.0065	0.0864	19.7	0.1839	0.7433	15.1	0.3953	0.8682	-3.8	0.5966	0.9686	27.5	0.0087	0.2668	32.6	0.0177	0.4079
73	LPE 20:5	499.2698	15.65	129.2	0.0038	0.0864	56.7	0.3256	0.7734	55.2	0.1904	0.8682	-1.0	0.6269	0.9686	46.3	0.0102	0.2819	47.7	0.0570	0.4079
74	LPE 22:5	527.3007	18.46	53.3	0.0498	0.1768	-7.5	0.5817	0.8658	15.4	0.3953	0.8682	24.8	0.0456	0.9686	65.8	0.0018	0.1592	32.9	0.1073	0.4250
75	LPE 22:6	525.2853	17.49	42.4	0.0242	0.1256	12.3	0.2566	0.7734	8.8	0.4484	0.8998	-3.1	0.6894	0.9686	26.8	0.0858	0.3595	30.8	0.1073	0.4250
76	LPI 16:0 ‡	572.2970	22.67	44.6	0.0151	0.1034	24.4	0.0983	0.7078	11.9	0.2413	0.8682	-10.1	0.5966	0.9686	16.2	0.2156	0.5163	29.3	0.3055	0.6195
77	LPI 20:4	620.2958	20.50	63.6	0.0065	0.0864	24.1	0.2477	0.7734	16.9	0.4484	0.8998	-5.8	0.7053	0.9686	31.7	0.0262	0.3046	39.9	0.0326	0.4079
78	Deoxycholic acid OR isomers ‡	392.2924	14.73	-62.4	0.0128	0.0969	-1.3	0.2947	0.7734	-10.1	0.1753	0.8682	-8.9	0.9054	0.9924	-61.9	0.1965	0.5143	-58.1	0.3055	0.6195
79	Deoxycholic acid OR isomers ‡	392.2926	14.73	-63.3	0.0128	0.0969	-10.3	0.1075	0.7078	-9.6	0.1753	0.8682	0.8	0.8038	0.9810	-59.1	0.3052	0.6233	-59.4	0.3055	0.6195
80	PC 16:0/20:5	779.5476	34.14	-40.2	0.0326	0.1415	-34.0	0.2305	0.7734	-26.6	0.5053	0.9301	11.2	0.6269	0.9686	-9.5	0.6222	0.8120	-18.6	0.5387	0.7703
81	Sphingosine-1-Phosphate ‡	379.2490	14.85	24.2	0.0242	0.1256	7.6	0.3699	0.7857	0.9	0.8722	0.9847	-6.2	0.4178	0.9686	15.4	0.0444	0.3268	23.1	0.0242	0.4079
82	L-Arginine ‡	174.1108	0.75	19.0	0.0426	0.3035	19.9	0.0478	0.4563	28.3	0.1195	0.8001	7.1	0.9253	0.9986	-0.7	0.6062	0.9221	-7.3	0.7861	0.9711
83	L-Leucine ‡	131.0928	0.94	24.7	0.0426	0.3035	17.4	0.0572	0.4664	14.4	0.2278	0.8001	-2.6	0.6392	0.9986	6.2	0.3903	0.8716	9.1	0.2083	0.8588
84	LPE 20:4	501.2869	17.58	37.3	0.0370	0.3035	19.1	0.0913	0.4808	18.5	0.1958	0.8001	-0.5	0.8982	0.9986	15.3	0.0204	0.5008	15.8	0.0717	0.6872
85	LPE 20:4	501.2869	17.58	32.9	0.0426	0.3035	18.0	0.1115	0.4808	18.3	0.2114	0.8001	0.2	0.7916	0.9986	12.6	0.0588	0.5134	12.3	0.1323	0.7587

86	Retinol	286.2293	27.26	25.0	0.0426	0.3035	10.3	0.1870	0.5647	-0.6	0.6452	0.8754	-9.9	0.2569	0.9986	13.3	0.2051	0.7565	25.8	0.0318	0.6148
----	---------	----------	-------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------

Legend: LPC: lysophosphocholine; LPE: lysophosphoethanolamine; LPI: lysophosphatidylinositol; PC: Phosphatidilcholine; p value (BH): p-value after applying the Benjamini-Hochberg procedure to control the False Discovery Rate (FDR); ‡: ID confirmed using a commercial standard; ICS: Patients controlled with inhaled corticosteroids; IT: Patients controlled with immunotherapy, BIO: Patients controlled with Omalizumab, UC: Patients with uncontrolled asthma. Benjamin-Hochberg multiple test correction (pBH) is also known as False Discovery Rate. The pBH value correction was applied to the significant metabolites.

Table S5. Enriched metabolic pathways in UC patients compared to ICS group (*FDR* < 0.2).

Pathway name	<i>p</i> -value (Fisher test)	<i>p</i> -Value (BH)	Metabolites present in the pathway (%)	Pathway Source
phospho-PLA2 pathway	1.43E-06	0.00609	100.00	Reactome
Ca-dependent events	6.67E-06	0.0142	91.67	Reactome
PLC beta mediated events	2.72E-05	0.0232	89.47	Reactome
G-protein mediated events	2.72E-05	0.0232	89.47	Reactome
Opioid Signalling	5.27E-05	0.0322	87.50	Reactome
GPCR downstream signalling	9.70E-05	0.0443	69.86	Reactome
Signaling by GPCR	1.04E-04	0.0443	69.48	Reactome
Leukotriene biosynthesis	1.43E-04	0.0553	96.67	HumanCyc
Acyl chain remodeling of CL	2.01E-04	0.0661	50.00	Reactome
Acyl chain remodelling of PC	2.01E-04	0.0661	46.15	Reactome
Fc gamma R-mediated phagocytosis - Homo sapiens (human)	2.81E-04	0.0800	87.50	KEGG
COPI-independent Golgi-to-ER retrograde traffic	2.81E-04	0.0800	87.50	Reactome
Signal Transduction	3.62E-04	0.0966	70.85	Reactome
Golgi-to-ER retrograde transport	4.80E-04	0.1060	90.00	Reactome
G alpha (i) signalling events	5.48E-04	0.1060	70.44	Reactome
HDL remodeling	5.99E-04	0.1060	58.82	Reactome
Plasma lipoprotein remodeling	5.99E-04	0.1060	50.00	Reactome
Choline metabolism in cancer - Homo sapiens (human)	5.99E-04	0.1060	90.91	KEGG
Intra-Golgi and retrograde Golgi-to-ER traffic	7.30E-04	0.1060	91.67	Reactome
Plasma lipoprotein assembly_ remodeling_ and clearance	0.001	0.1060	52.00	Reactome
Leukotriene C4 Synthesis Deficiency	0.0016	0.1060	91.55	SMPDB
Retrograde endocannabinoid signaling - Homo sapiens (human)	0.0018	0.1100	89.47	KEGG
Alpha Linolenic Acid and Linoleic Acid Metabolism	0.0018	0.1100	94.44	SMPDB
Celecoxib Action Pathway	0.0019	0.1150	90.79	SMPDB
Linoleic acid metabolism - Homo sapiens (human)	0.0022	0.1340	67.86	KEGG
Membrane Trafficking	0.0025	0.1450	66.67	Reactome
Linoleate metabolism	0.0025	0.1450	83.33	EHMN
Transport of small molecules	0.0031	0.1670	78.32	Reactome
Arachidonic acid metabolism	0.0032	0.1670	79.81	Reactome
Phospholipid Biosynthesis	0.0033	0.1670	95.83	SMPDB
Glycerophospholipid metabolism - Homo sapiens (human)	0.0033	0.1670	44.23	KEGG

Legend: *p* value (BH): *p*-value after applying the Benjamini-Hochberg procedure to control the False Discovery Rate (FDR)

Table S6. Differential expression of serum proteins between ICS and UC groups.

Protein	p-value	p-value (BH)	logFC in UC
CASP-8	0.0019	0.1722	0.8887
ARG1	0.0095	0.3219	0.6945
IFN γ	0.0150	0.3219	-1.0880
CCL19	0.0160	0.3219	-0.8013
CCL13	0.0175	0.3219	0.6027
CD4	0.0282	0.379	-0.2826
IL15	0.0288	0.379	0.2786
TNFRSF12A	0.0367	0.4222	0.3939
CD40	0.0583	0.5963	0.2662
TIE2	0.0725	0.6059	0.1675
HGF	0.0778	0.6059	0.2847
FGF2	0.0796	0.6059	0.0972
PDGF subunit B	0.0878	0.6059	0.0820
CXCL12	0.0922	0.6059	-0.0957
ICOSLG	0.1013	0.6104	0.2889
TNFSF14	0.1159	0.6104	0.3216
IL13	0.1202	0.6104	0.2609
GZMB	0.1230	0.6104	0.5404
IL8	0.1261	0.6104	0.3042
LAMP3	0.1700	0.7819	-0.3742
CXCL1	0.1981	0.8525	0.2798
VEGFA	0.2169	0.8525	0.3049
PTN	0.2224	0.8525	0.3497
FASLG	0.2305	0.8525	0.2086
EGF	0.2464	0.8525	0.4909
MIC-A/B	0.2614	0.8525	0.6787
CX3CL1	0.2643	0.8525	0.1422
CD244	0.2657	0.8525	0.1485
IL7	0.2707	0.8525	0.2045
ADA	0.2879	0.8525	0.1445
CCL23	0.2890	0.8525	0.2007
IL5	0.2965	0.8525	-0.3551
IL2	0.3370	0.8764	0.0362
CD28	0.3393	0.8764	-0.0596
VEGFR-2	0.3564	0.8764	-0.0951
MCP-1	0.3692	0.8764	0.1647
CXCL5	0.3789	0.8764	-0.1786
GZMA	0.3969	0.8764	0.1487
PGF	0.4114	0.8764	-0.1782
IL6	0.4123	0.8764	0.2100
IL-1 alpha	0.4189	0.8764	-0.0197
CRTAM	0.4298	0.8764	-0.1564
ADGRG1	0.4366	0.8764	0.2031
MMP7	0.4433	0.8764	0.0854
ANGPT1	0.4466	0.8764	0.1188
Gal-9	0.4495	0.8764	0.0919
CAIX	0.4602	0.8764	-0.1663
CCL20	0.4647	0.8764	0.2925
CD27	0.4668	0.8764	-0.1079

PD-L1	0.4808	0.8847	-0.1106
GZMH	0.4952	0.8859	0.2153
CD40-L	0.5007	0.8859	0.3346
IL12RB1	0.5204	0.9034	0.0520
TNFRSF9	0.5323	0.9069	-0.1042
DCN	0.5511	0.9218	0.0619
CD83	0.6067	0.9372	-0.0637
KIR3DL1	0.6069	0.9372	-0.2133
MCP-2	0.6147	0.9372	0.1473
Gal-1	0.6168	0.9372	0.0449
CXCL9	0.6393	0.9372	-0.1510
TNF	0.6401	0.9372	-0.1260
CD70	0.6405	0.9372	-0.1081
CD8A	0.6426	0.9372	0.0857
TNFRSF21	0.6597	0.9372	0.0384
CXCL13	0.6886	0.9372	-0.1233
IL18	0.6958	0.9372	0.0917
ANGPT2	0.7009	0.9372	-0.0732
PD-L2	0.7227	0.9372	-0.0488
CD5	0.7245	0.9372	-0.0736
CSF-1	0.7309	0.9372	0.0322
IL12	0.7336	0.9372	-0.0927
PDCD1	0.7375	0.9372	0.0706
LAP TGF-beta-1	0.7437	0.9372	0.0539
TNFRSF4	0.7549	0.9385	-0.0516
LAG3	0.7671	0.941	0.0602
MMP12	0.8168	0.9677	0.0520
TWEAK	0.8183	0.9677	0.0360
CXCL10	0.8219	0.9677	0.0644
CXCL11	0.8486	0.9677	0.0543
CCL17	0.8524	0.9677	0.0640
IL10	0.8562	0.9677	0.0373
HO-1	0.8628	0.9677	-0.0292
KLRD1	0.8731	0.9677	-0.0379
TRAIL	0.8924	0.9774	-0.0202
MCP-3	0.9123	0.9874	0.0256
MUC-16	0.9304	0.9903	0.0262
CCL3	0.9365	0.9903	0.0369
NCR1	0.9653	0.9964	0.0072
NOS3	0.9692	0.9964	0.0119
CCL4	0.9781	0.9964	0.0072
IL4	0.9856	0.9964	0.0037
IL33	1	1	0

Normalized protein expression values (NPX) between the ICS and UC were analyzed using a *t*-test comparison. **Legend:** *FC*: Fold change; *p value (BH)*: *p*-value after applying the Benjamini-Hochberg procedure to control the False Discovery Rate (FDR)

Table S7. Top 25 most significantly enriched biological **Gene Ontology (GO)** processes linked to significantly expressed proteins in serum of UC patients.

#term ID	Pathway	Observed gene count	Background gene count	p-value (BH)
GO:0034097	Response to cytokine	11	1035	1.46E-11
GO:1902107	Positive regulation of leukocyte differentiation	7	139	2.70E-10
GO:0071345	Cellular response to cytokine stimulus	10	953	4.21E-10
GO:0048584	Positive regulation of response to stimulus	11	2054	4.43E-09
GO:0051707	Response to other organism	9	835	5.43E-09
GO:0071310	Cellular response to organic substance	11	2219	5.56E-09
GO:1903037	Regulation of leukocyte cell-cell adhesion	7	278	5.56E-09
GO:0097191	Extrinsic apoptotic signaling pathway	5	93	7.98E-08
GO:0033993	Response to lipid	8	825	8.91E-08
GO:0050776	Regulation of immune response	8	873	1.29E-07
GO:0002684	Positive regulation of immune system process	8	882	1.35E-07
GO:0007166	Cell surface receptor signaling pathway	10	2198	1.63E-07
GO:0045597	Positive regulation of cell differentiation	8	908	1.63E-07
GO:0050863	Regulation of T cell activation	6	302	2.78E-07
GO:0034612	Response to tumor necrosis factor	5	217	2.01E-06
GO:0032268	Regulation of cellular protein metabolic process	9	2486	6.02E-06
GO:0043123	Positive regulation of I-kappab kinase/NF-kappab signaling	4	122	7.81E-06
GO:0046425	Regulation of receptor signaling pathway via JAK-STAT	4	151	1.57E-05
GO:0051173	Positive regulation of nitrogen compound metabolic process	9	2946	2.09E-05
GO:0070374	Positive regulation of ERK1 and ERK2 cascade	4	196	3.78E-05
GO:0006959	Humoral immune response	4	252	8.97E-05
GO:0032879	Regulation of localization	8	2524	8.97E-05
GO:0032700	Negative regulation of interleukin-17 production	2	12	2.30E-04
GO:0035723	Interleukin-15-mediated signaling pathway	2	13	2.60E-04
GO:0045834	Positive regulation of lipid metabolic process	3	135	4.10E-04

Legend: *p value (BH)*: p-value after applying the Benjamini-Hochberg procedure to control the False Discovery Rate (FDR). Note: There were 345 pathways with FDR < 0.2

Table S8. Machine-Learning parameters for k-NN models.

Algorithm	Accuracy	Specificity	Sensitivity	PPV	NPV
k-NN (k=2 OR k=3)	0.875	0.75	1.00	0.8	1.00

Legend. *PPV*: Positive predictive value; *NPV*: Negative predictive value

Supplementary Figures

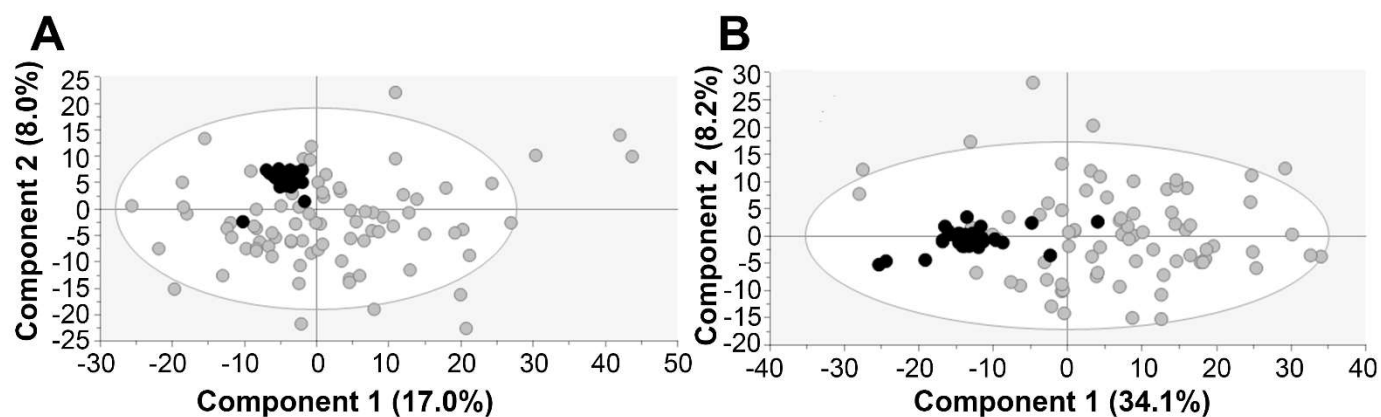


Figure S 11. Unsupervised PCA models showing the quality of the data for LC-MS positive (A) and negative (B) polarities. (Key: QC injections (●; black). samples (●; grey)). The clustering of the QCs injections proved that the data from metabolomics is of high-quality, so the differences found later between the samples are due to the condition of the patients nor the technique.

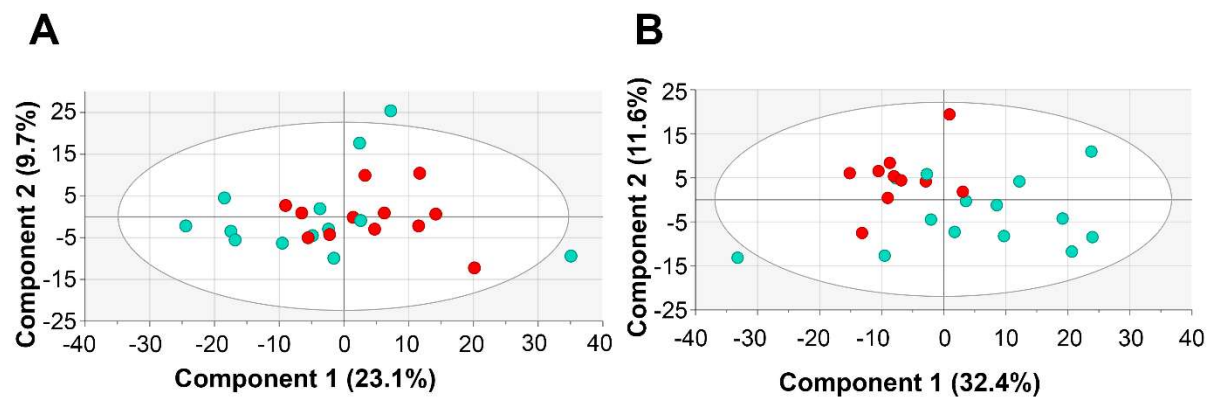


Figure S 2. PCA models of ICS and UC groups in both ESI modes (Key: “ICS”: blue; “UC”: red).

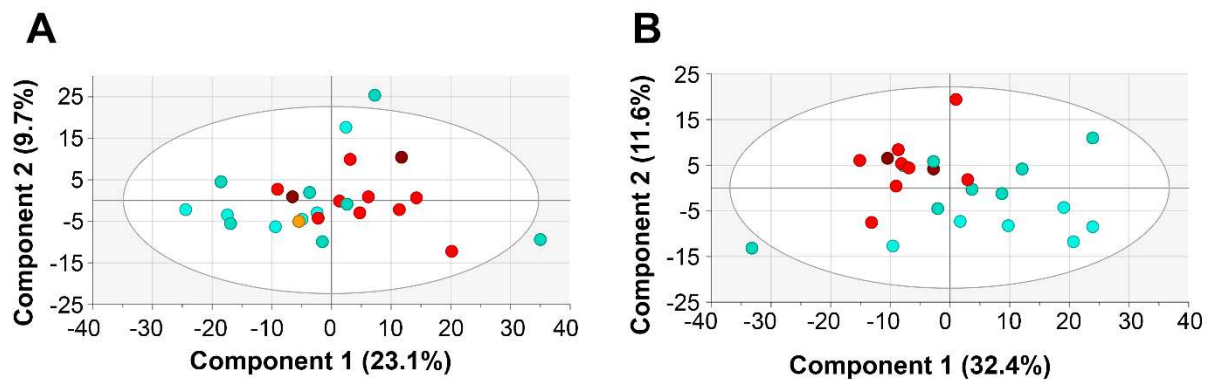


Figure S 3. PCA models of ICS and UC samples coloured in both ESI modes by age range (keynote: red: UC samples, from which orange-red: 18-35 years old, light-red: 35-60 years old and dark-red: ≥ 60 years old; blue: ICS samples, from which light-blue: 18-35 years old; dark-blue: 35-60 years old).

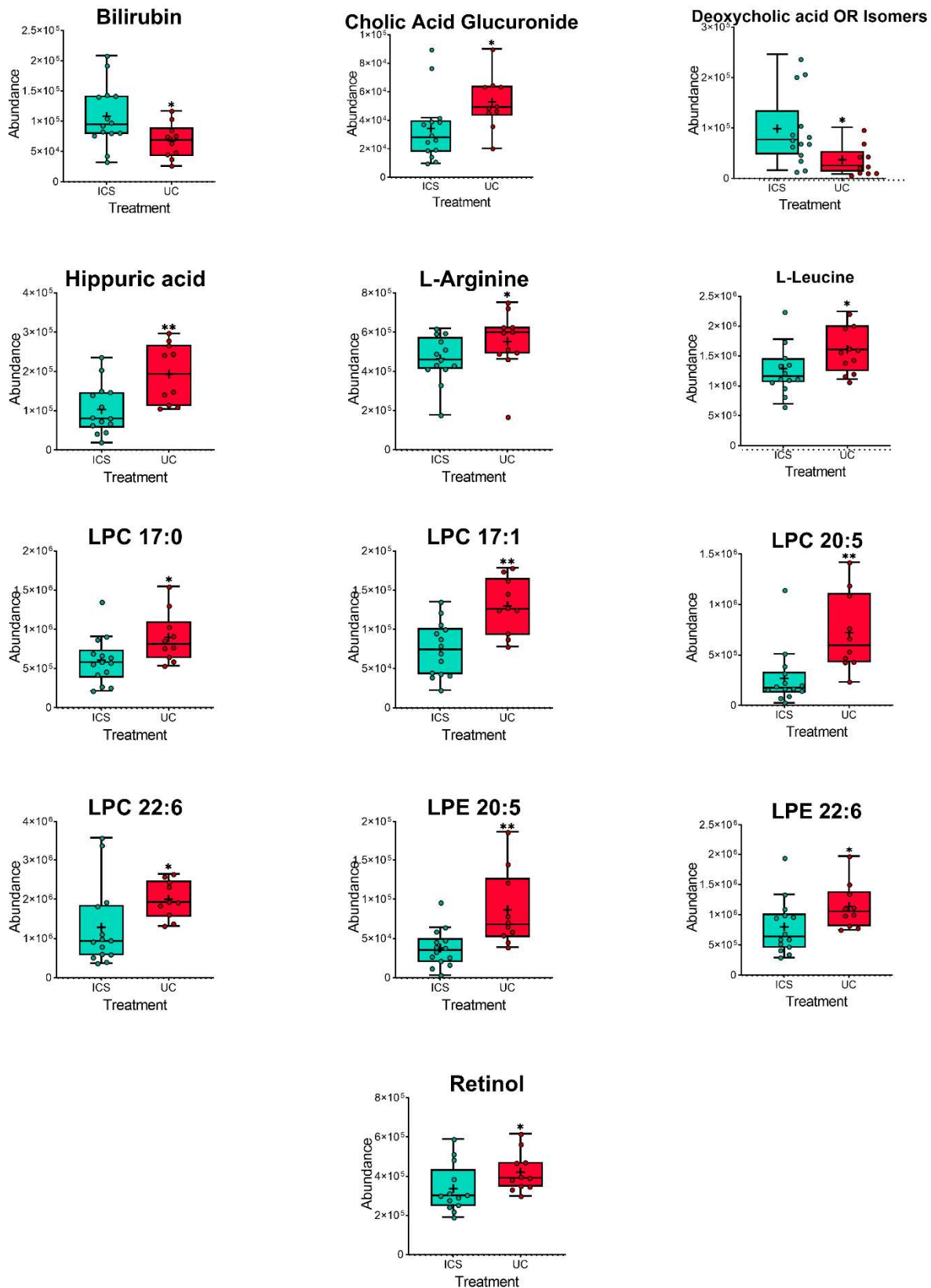


Figure S 4. Trajectories of the remaining identified compounds. Box and whiskers plots with mean are represented for corticosteroid-controlled (ICS, blue) and uncontrolled (UC, red) groups. For metabolites that had more than one m/z or that were detected in both ESI modes, the most abundant m/z overall was represented. Mean is represented with “+” inside the boxes and individual patients are shown as blue (ICS) or red (UC) circles. Mann-Whitney U test was used to calculate significant differences. * $p < 0.05$; ** $p < 0.01$.

Supplementary References

1. Gil de la Fuente A, Godzien J, Fernández López M, Rupérez FJ, Barbas C, Otero A. Knowledge-based metabolite annotation tool: CEU Mass Mediator. *J Pharm Biomed Anal* 2018;**154**:138–149.
2. Rodríguez-Coira J, Delgado-Dolset MI, Obeso D, Dolores-Hernández M, Quintás G, Angulo S et al. Troubleshooting in Large-Scale LC-ToF-MS Metabolomics Analysis: Solving Complex Issues in Big Cohorts. *Metabolites* 2019;**9**:247.
3. Gromski P, Xu Y, Kotze H, Correa E, Ellis D, Armitage E et al. Influence of Missing Values Substitutes on Multivariate Analysis of Metabolomics Data. *Metabolites* 2014;**4**:433–452.
4. Kuligowski J, Sánchez-Illana Á, Sanjuán-Herráez D, Vento M, Quintás G. Intra-batch effect correction in liquid chromatography-mass spectrometry using quality control samples and support vector regression (QC-SVRC). *Analyst* 2015;**140**:7810–7817.
5. Sánchez-Illana Á, Piñeiro-Ramos JD, Sanjuan-Herráez JD, Vento M, Quintás G, Kuligowski J. Evaluation of batch effect elimination using quality control replicates in LC-MS metabolite profiling. *Anal Chim Acta* 2018;**1019**:38–48.
6. Triba MN, Le Moyec L, Amathieu R, Goossens C, Bouchemal N, Nahon P et al. PLS/OPLS models in metabolomics: the impact of permutation of dataset rows on the K-fold cross-validation quality parameters. *Mol Biosyst* 2015;**11**:13–19.
7. Mitamura Y, Schulz D, Oro S, Li N, Kolm I, Lang C et al. Cutaneous and systemic hyperinflammation drives maculopapular drug exanthema in severely ill COVID-19 patients. *Allergy Eur J Allergy Clin Immunol* 2021;:1–14.
8. Calinski T, Harabasz J. A dendrite method for cluster analysis. *Commun Stat - Theory Methods* 1974;**3**:1–27.
9. Kodinariya T, Makwana PR. Review on determining number of Cluster in K-Means Clustering. *Int J Adv Res Comput Sci Manag Stud* 2013;**1**:90–95.