Identification of biomarkers in food allergy using metabolomics

Haerin Jang

Department of Medical Science
The Graduate School, Yonsei University

Identification of biomarkers in food allergy using metabolomics

Directed by Professor Kyung Won Kim

The Master's Thesis
submitted to the Department of Medical Science,
the Graduate School of Yonsei University
in partial fulfillment of the requirements for
the degree of Master of Medical Science

Haerin Jang

December 2020

This certifies that the Master's Thesis of
Haerin Jang is approved.

_____

Thesis Supervisor : Kyung Won Kim

_____

Thesis Committee Member#1 : Myung Hyun Sohn

_____

Thesis Committee Member#2 : Jae-Hyun Lee

The Graduate School
Yonsei University

December 2020

ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere gratitude to my advisor Professor Kyung Won Kim. Her guidance helped me through every stage of my studies and the writing of this thesis. I could not have imagined having a better advisor for my master's study. Her enthusiasm and motivation towards research have been a true inspiration. It has been a privilege to learn from her.

I would also like to thank the thesis committee, Professor Myung Hyun Sohn and Professor Jae-Hyun Lee for their encouragement and insightful comments. The constructive feedback pushed me to sharpen my thinking and brought my work to a higher level.

I owe my deepest gratitude to Professor Junjeong Choi for believing in me and convincing me that I have what it takes to continue research. Without her guidance and persistent encouragement, I would not have decided to seek my potential. All my success depends on her words of encouragement. I am truly grateful for everything she has done for me.

I am also extremely fortunate to have had the warm-hearted support from friends and colleagues. I wish to extend my special thanks to my friends: Gyubin Min, TaeWoo Kim, JeongA Lee, Woojong Kang and Yeaeun Lee for the wonderful time we shared, especially the tea-breaks and dinners that gave me the time I

needed to get back on my feet. I also thank my colleagues, Eunroc Seo and Minsoo Shim, for making me look forward to going to the lab each morning. It was a great pleasure to study next to them.

Last but not least, I would like to thank my caring family for their constant support and comfort provided. I want to thank them for their patience and for all of the opportunities I was given to further my studies.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Identification of biomarkers in food allergy using metabolomics

Haerin Jang

Department of Medical Science

The Graduate School, Yonsei University

(Directed by Professor Kyung Won Kim)

Food allergy (FA) is affecting an increasing proportion of children, with its emergence to an important health concern. The natural history of childhood FA occasionally involves a spontaneous resolution of symptoms after years of adversity. However, the mechanism behind the resolution of FA is not studied in-depth. This study aims at identifying biomarkers for the development and prognosis of FA through a metabolomics approach.

The metabolomic profile of 20 children with FA and 20 healthy control subjects were studied using liquid-chromatography coupled with

mass spectrometry. Serum samples for FA are collected from both the time of diagnosis and resolution to identify the change in metabolite levels. We conducted a comparative analysis to discover metabolites associated with FA and FA resolution. With the integration of genotype data, we conducted a metabolite QTL analysis to analyze associated pathways.

Patients with FA exhibited a unique metabolomic profile compared to control subjects. A marked increase of sphingolipid metabolites and a decrease in acylcarnitine metabolites were associated to FA. In resolving FA subjects, a significantly high level of omega-3 metabolites were observed. The level of omega-3 metabolites decreased with FA resolution as compared to an increase over the same period in persistent FA. Also, platelet-activating factor, and lysophosphatidylcholine levels were significantly associated to FA resolution.

Several lipid metabolites are closely related to FA and FA resolution in children. These results suggest potential predictive biomarkers and provides insight in disease mechanisms of childhood FA.

Key words: food allergy, metabolomics, omega-3, biomarker, childhood

Identification of biomarkers in food allergy using metabolomics

Haerin Jang

Department of Medical Science

The Graduate School, Yonsei University

(Directed by Professor Kyung Won Kim)

I. INTRODUCTION

1. Food allergy

Food allergy (FA) is a disease where allergic reactions repeatedly occur after consuming certain foods. Symptoms involve tingling or itching in the mouth, rashes or hives of the skin, angioedema, difficulty swallowing, shortness of breath, or gastrointestinal symptoms such as abdominal pain and diarrhea[1]. In severe cases, anaphylaxis may occur, which results in trouble breathing or speaking and can lead to life-threatening situations. Immediate medical treatment is required in most cases, and if not treated properly it can be fatal[2]. However, the current standard treatment of FA is allergen avoidance but this

does not change the natural course of FA[3].

FA is affecting an increasing proportion of people, with the prevalence increasing especially in developed countries. The prevalence of FA is especially high in Australia, the United States, and Europe with a prevalence of close to 10% in all ages[4]. In children, the prevalence of FA has also been high, with a prevalence of 8.0% in the United States[5]. Children experiencing severe reactions are 38.7% of those with FA[5], indicating that a high proportion of patients are at risk of anaphylaxis. In Korean children, the prevalence is relatively low at about 4% in school children but it has increased in the recent years[6].

Common causative allergens are cow's milk, hen's egg, peanuts, tree nuts, soy, wheat, and fruits[7]. Particular proteins found in each food causes the immune system to overreact, even when it is consumed in small amounts. The severity of FA is commonly associated with the allergen or number of different food allergen one reacts to[8]. Some food allergies cause cross-reaction to other proteins that are similar to allergy causing-proteins. This commonly occurs in pollen-food allergy syndrome where one is allergic to multiple foods due to the similarity in protein structure[9].

The mechanism of IgE-mediated FA can be explained in two steps. When a food allergen is introduced to the body for the first time, it binds to receptors on tissue basophils and mast cells and stimulates the production of allergen-specific IgE. This process is called "initial sensitization". Following

this process, the introduction of the food allergen induces an immediate hypersensitivity in which specific IgE antibodies bound to mast cells and basophils activate a chemical cascade and trigger a release of mediator molecules. Histamine is a distinctive mediator, which causes an immediate pathophysiological response in the blood vessel and several organs. Prostaglandins and leukotrienes are also involved in this response, causing clinical reactivity and allergic symptoms. Non-IgE medicated FA is less common and its reaction is more likely to be related to T cells instead of leukocytes[10].

Several risk factors are associated with FA. First, a family history or coexistence of allergic diseases, including asthma, eczema, or allergic rhinitis, increases the risk of FA[11]. Young children who are sensitized to food allergens are also more likely to develop allergic rhinitis and asthma[12]. Also, age is a contributing risk factor. FA is more common in children, especially toddlers and infants. FA mostly begins in the first two years of life and either continue through adulthood or resolve during childhood and adolescence[13].

The natural history of childhood FA occasionally involves a spontaneous resolution of symptoms after years of adversity. Many patients with FA will naturally outgrow their symptoms, with resolution occurring at different ages[13]. Many factors contribute to resolution including the causative allergen and severity of symptoms. Food allergens such as cow's milk and hen's egg are outgrown often during childhood or adolescence, whereas peanut and tree nut

allergies more often persist through adulthood, rarely resolving with time[14].

The mechanism behind the resolution of FA is not studied in-depth, possibly due to heterogeneous phenotypes and the point of resolution being unclear. Multiple factors are known to be involved and there has been controversy on the predicting factors, one being the decrement of allergen-specific IgE levels[13]. However, this does not happen in all cases[15]. Some patients become tolerant even with a persistently elevated allergen-specific IgE level. Thus, the loss of specific IgE is not a necessary factor for FA resolution. Serial testing is needed to determine if a FA is resolved. *In vitro* tests and skin prick tests are most common, and oral food challenges are the most determinant, but oral food challenges have great risk of anaphylaxis and full emergency equipment is required.

2. Omics analysis

Recent omics studies have shown great potential in revealing the missing link between the genome and disease phenotypes[16]. Omics researches aim at a systems approach, where a collective characterization of biological molecules shows the full picture. Common molecules quantified and analyzed are DNAs, RNAs, proteins, and metabolites, which translate to genomics, transcriptomics, proteomics, and metabolomics. It is becoming more and more common to integrate different types of omics data on a large scale as computational techniques allow further investigation.

Metabolomics is the study of metabolites, or small biochemical molecules within cells, biofluids, tissues, or organisms[17]. The metabolome is the whole of the metabolites in a biological system. Metabolites are produced and modulated by enzyme-mediated reactions and contribute to crucial cellular functions. Metabolomics provides a direct and sensitive measure of disease phenotype at the molecular level[18]. It is downstream of the genome, after the transcriptome and proteome. Hence it reflects both the genetic basis and environmental changes. Many low-weight biochemicals, including hydrocarbons, amino acids, and lipids are measured using metabolomics.

Metabolomics employs analytical methodologies such as liquid chromatography-mass spectrometry analysis to either identified or unknown compounds. It identifies the mass profile of each metabolite within samples. The peak area of each metabolite determined is used to define the concentration differences between different biological samples being measured. When unknown metabolites are identified, it is called “untargeted metabolomics” and is used to identify a large range of metabolites[19]. Methods used in metabolomics continue to evolve and improve, yet there is no standardized method for analysis.

Metabolomics is used in hypothesis-generating studies, where the metabolome for diseased or specific individuals are screened to find predictive biomarkers or mechanistic explanations. Improvements in analytical equipment and data analysis software have driven the development of metabolomic

research. It has been used in studies for many complex diseases, such as diabetes[20], cardiovascular diseases[21], and Alzheimer's disease[22], producing promising results. Most of these studies have been interpreted using its relevance to the gut microbiome, genome, and environmental factors.

There have been several attempts to perform metabolomic studies on allergic diseases. Metabolomic approaches have identified several metabolites that differ by disease severity and phenotype in various biofluids. One pilot study identified volatile organic compound markers of childhood asthma in exhaled breath, where 8 compounds could distinguish asthmatic children from healthy children with high accuracy[23]. In another study, metabolic profiles were used to classify "non-severe profilin-allergic patients" to mild and moderate phenotypes[24]. They discovered that arachidonic acid precursors were increased, suggesting a possible explanation regarding consumption due to the synthesis of inflammatory mediators.

However, in FA, there are only limited studies that perform metabolomic analysis on human samples. There have been reports on differences in baseline metabolite state between patients with peanut allergy and tolerant patients[25]. Also, children with FA were shown to have lower levels of sphingolipid, and ceramides[26]. In a mouse model study, an arachidonic acid metabolite, prostaglandin D2 in urine was correlated with the severity of oral allergy symptoms[27].

Compared to metabolomics, genomic research for FA has been done

more actively with multiple genome-wide association studies and candidate gene studies. FLG, HLA, IL10, IL13 are most often mentioned as genes associated with FA[28]. Many of these genes are also associated with other allergic diseases, such as atopic eczema and asthma[29]. Despite promising results, genetic studies of FA have their limitations. There are few studies with adequate sample size and analysis with multiple testing correction. However, results show that there is a common root behind allergic diseases that may contribute to the inflammation state.

Integrated omics studies are becoming more and more common in complex diseases[16]. There is a potential for use in patient subgrouping, associating omics-based molecular measurements with clinical outcomes, or understanding the biological system under diseases. Integrated approaches combine individual omics data, sequentially or simultaneously, to understand the interplay of molecules[30]. Compared to seeing only one type of data, this provides a holistic view that may result in higher accuracy of prediction or describing disease phenotype. For example, an integrated study of metabolomics and transcriptome yielded molecular perturbations underlying prostate cancer. This research suggests that integrated studies of omics data have the potential to improve diagnosis and treatment in clinical settings. [31]

3. Research purpose

This study aims at identifying biomarkers for the development and

prognosis of FA through an integrated metabolomics and genomics approach. We use liquid-chromatography coupled with mass spectrometry to identify metabolites in serum samples from children with FA. We then conduct a comparative analysis with healthy controls to discover metabolites that are associated with FA. Also, to discover metabolites associated with FA resolution, those with resolving FA symptoms are compared to those with persistent symptoms. Serum samples of FA patients are collected from both the time of diagnosis and resolution to identify the change in metabolite levels. Since metabolites are influenced by the genome and the environment, this will further our knowledge of the epidemiological mechanism of FA and its resolution.

II. MATERIALS AND METHODS

1. Study population and sample collection

A total of 40 subjects were included in this study: 20 children with FA and 20 healthy control subjects. All children with FA had confirmed symptoms of hen's egg, with several children also having FA reactions to multiple food allergens. FA reactions include tingling or itching in the mouth, hives, swelling of the face, mouth or other body areas, wheezing, shortness of breath, nausea, vomiting, abdominal pain and diarrhea. All FA subjects had confirmed allergic sensitization, defined by specific IgE levels of greater than 0.7 kU/L to egg white and at least one of the following food or airborne allergens: milk, peanut, soybean, wheat, *Dermatophagoides pteronyssinus*, *Dermatophagoides farinae*, Alternaria species, or *Blattella germanica*. Resolution of FA was defined as the resolution of symptoms from all food allergens that have previously caused FA reactions, often confirmed by oral food challenges. If symptoms persisted in at least one food allergen, it was defined as persistent FA.

Serum samples were collected from FA children at both the time of diagnosis and at follow-up, hence resulting in 20 pairs of samples for the FA subjects. Follow-up was after confirmed FA resolution or matching time progress in persistent patients. Serum samples were stored at -20℃ until metabolite profiling was performed. Whole blood was also collected from the

40 subjects for genotype analysis at the same time point. Control subjects did not have any history of allergic disease or evidence of allergic sensitization. All children were recruited from Severance Children's Hospital, Seoul, Korea, and the study was approved by the institutional review board of Severance Hospital (Seoul, Korea; IRB no. 4-2019-1271). All subjects were unrelated and written informed consent was provided by them or their parents.

2. LC-MS profiling

Metabolome analysis was performed in 60 samples of human serum using liquid chromatography-tandem mass spectrometry (LC-MS) in two modes for cationic and anionic metabolites. Each 100 μL sample was mixed with 300 μL of 0.1% formic acid in methanol containing internal standards and centrifuged (9,100 x g, 4℃, 10 mins). Then, 250 μL of supernatant and 550 μL of 0.1% formic acid in water were mixed and filtrated by using SPE columns (MonospinC18, 5010-2170, GL Sciences Inc., Tokyo, Japan). The filtrate was purified by 0.1% formic acid solution and 0.1% formic acid-25% methanol solution. Then, purified lipid metabolite was dissolved by 200 μL of 0.1% formic acid in methanol immediately before measurement. Peaks detected in LC-MS analysis were extracted using automatic integration software (MultiQuant, AB Sciex) to obtain peak information including *m/z*, retention time (RT), and peak area. The peak area was then converted to relative peak area.

From the mass spectrometric data acquisition, 344 metabolites were identified from multiple classes (including free fatty acid (FFA), acylcarnitine (AC), oxylipin, lysophosphatidylcholine (LPC), lysophosphatidylethanolamine (LPE), lysophosphatidyllinositol (LPI), lysophosphatidylserine (LPS), lysophosphatidylglycerol (LPG), lysophosphatidic acid (LPA), platelet-activating factor (PAF), acylethanolamine (AEA), sphinganine, sphingosine, ganglioside (GM, GD, GT), glucosylceramide, lactosylceramide, ceramide-1P, and steroids). These metabolites were further classified based on the number of total acyl chain carbon atoms and double-bond contents and annotated as follows: [Lipid class] [Number of acyl chain carbon atoms]:[Number of double bonds in fatty acid moieties]. For quality control, we removed metabolites that were not quantified in over 10% of the samples.

3. Metabolome analysis

Missing values for the metabolite data were imputed using the k-nearest neighbor (KNN) algorithm. It uses the values of the neighbors and obtains a weighted average of their values to fill in the unknowns. The median of the 10 nearest neighbors was used for computation. The probabilistic quotient normalization (PQN) method was used for normalization of the metabolite data. PQN is commonly used in metabolomics data for its robustness and accuracy in data with a large number of metabolites. It is based on the calculation of a probable dilution factor by looking at the distribution of the

quotients of the amplitudes of a test spectrum by those of a reference spectrum[32].

To observe data patterns and identify any outliers, principal component analysis (PCA), sparse partial least squares-discriminant analysis (sPLS-DA), and orthogonal partial least squares-discriminant analysis (orthoPLS-DA) were performed using Metaboanalyst[33]. Only the initial data was used in this process to find distribution patterns at the early stage. PLS-DA is a supervised form of discriminant analysis, where class labels are introduced to the algorithm. It is used to identify class differences from a multivariate dataset[34]. This overcomes the limitation of using PCA in metabolomic data, where within-group variation can be bigger than between-group variation. 5 components were used to control the sparseness of the sPLS-DA model and 5-fold cross-validation was used. sPLS-DA and orthoPLS-DA are extensions of PLS-DA which maximizes the explained variables. Each analysis was performed on 2 different group classifications; ‘FA-control’, ‘persistent FA-resolving FA’. Subjects with continued FA reactions to any allergen were classified as persistent FA, while resolving FA was defined as resolution from all food allergens.

Correlation between each metabolite and the total IgE level or age was sought by computing the Pearson’s correlation coefficient using log-transformed metabolite data. Metabolites significantly associated with either total IgE levels or age were excluded from further analysis. The initial data of each group were compared using Welch’s two-sample t-test with normal

distributions and the Mann-Whitney U test was used instead for non-normal distributions. A p-value under 0.05 was considered statistically significant. For HDoHE metabolites that were significantly associated with FA resolution, the change in metabolite levels between initial data and follow up data was assessed. Also, heatmaps were built and hierarchical clustering was performed using Metaboanalyst to identify patterns in significant metabolites. Euclidean distance measure and Ward clustering algorithm were used on the normalized data for the heatmaps. Data were analyzed using R version 3.3 (R Foundation for Statistical Computing, Vienna, Austria) and SAS version 9.4 (SAS Institute Inc, Cary, NC, U.S.A).

4. Genotyping

Genotyping was performed using Infinium Omni5Exome-4 Bead Chip (Illumina Inc., San Diego, CA, USA) on whole blood samples. For quality control, single-nucleotide polymorphisms (SNPs) were excluded if missing rates were higher than 5%, Hardy-Weinberg equilibrium p-values were lower than 10^{-5}, missing rates between cases and controls were significantly different (p-value < 0.001), or minor allele frequencies were lower than 1%. With the SNPs after quality control, SNPs that have reported association to FA in previous studies were chosen for further analysis. Suaini et al.[28] provided a list of 112 SNPs that were previously published in either genome-wide studies or targeted gene studies of FA. These were curated from cross-sectional studies,

case-control studies, prospective, retrospective longitudinal studies (cohorts and case-control studies), family linkage studies, sibling-pair studies, and randomized control trials on FA. SNPs in linkage disequilibrium (LD) with those 112 SNPs were identified using SniPA[35] proxy search using a threshold of $r^2 > 0.8$. We used PLINK[36] to extract the SNPs and the SNPs in LD from the whole-genome SNP data. From the 350 SNPs, 91 SNPs were included in the Infinium Omni5Exome-4 Bead Chip and passed quality control. Fisher's exact test was used to test association with each phenotype and also with PAF metabolite levels. A p-value under 0.05 was considered statistically significant.

5. Metabolite QTL

Metabolite levels are known to be heritable and might be controlled by multiple genes, which qualifies the use of metabolite quantitative trait loci (metabolite QTL) analysis. There are several gene loci known to have an association with allergic inflammation and FA. However, little is known about how the biological pathways represented by these genes affect the metabolic signature and disease phenotypes. Like the mechanism behind gene expression quantitative trait (eQTL) loci analysis, metabolite QTL discovers novel genetic risk variants that are associated with specific metabolites. Metabolite QTL analysis was conducted using the R-package Matrix eQTL. This package searches for an association between each SNP and residual metabolite levels by modeling the additive effects of genotypes in the least-squares model[37]. It

performs a separate test for each metabolite–SNP pair and corrects for multiple comparisons by calculating the false discovery rate (FDR). From the significant metabolite QTL pairs, the genes associated with HDoHE metabolites were used for pathway analysis conducted with the Reactome Pathway Database. This procedure provides a comprehensive understanding of the molecular mechanisms that cause complex diseases[38].

6. Quantification of metabolites

80 subjects were chosen for the quantification of selected metabolites. Among the 80 subjects for quantification, 5 subjects with FA were also included in the previous metabolomics analysis. This cohort consisted of 50 FA subjects and 30 healthy controls. Inclusion criteria were identical to the metabolomics analysis. A total of 80 serum samples were used to quantify the concentration of 2 lipid metabolites. Serum levels of prostaglandin H2 (PGH2) and PAF were measured using an enzyme-linked immunosorbent assay (ELISA) kit (Abbexa Biologics, Arlington, Texas, USA) according to the manufacturers' protocols. Groups were compared using Welch's two-sample t-test with parametric statistics and Mann-Whitney U test for nonparametric statistics. A p-value under 0.05 was considered statistically significant.

III. RESULTS

1. Characteristics of study subjects

The demographics of the study subjects and disease attributes are summarized in Table 1. The mean age of FA subjects were slightly higher than control subjects were not statistically significant. The median age of FA subjects was 2.74 while for control, it was 2.70. The standard deviation of the age of FA subjects and control subjects was 2.30 and 2.28 respectively, showing a variation in subject age. The proportion of male subjects were similar between the groups. Total serum IgE levels were available for all subjects with a higher number in FA, which is a common characteristic of FA patients. In addition to the higher mean level, the variance within the group was also larger with a standard deviation of 1059.48 kU/L. The mean follow-up period between the blood sample collections for FA subjects was 16.5 months, with most children between ages 4 and 5. Among the children with FA, 7 subjects had only egg allergy, while 4 subjects had 2 food allergens and 8 subjects had more than two food allergies. All subjects experience urticarial after food ingestion, and 55% of subjects experienced angioedema as a food-related symptom. 35% of subjects had a history of anaphylaxis with symptoms ranging from shortness of breath, vomiting, diarrhea and low blood pressure.

Table 1. Subject characteristics

	FA (n=20)	Control (n=20)	p-value
Age, years	2.74 ± 2.30	2.70 ± 2.28	0.95
Male, n (%)	11 (55)	10 (50)	0.75
Total serum IgE, kU/L	676.49 ± 1059.48	17.72 ± 14.73	0.01
Follow-up period (months)	16.5 ± 7.4		
FA resolution, n (%)*	10 (55.6)		
Egg allergy resolution, n (%)*	7 (38.9)		
Multiple food allergies, n (%)[+]	12 (63.2)		
AD comorbidity, n (%)	10 (50)		
Food-related symptoms			
Urticaria, n (%)	20 (100)		
Angioedema, n (%)	11 (55)		
Anaphylaxis, n (%)	7 (35)		

FA, food allergy; IgE, immunoglobulin E; AD, atopic dermatitis.

Data given as number (%) or mean (± standard deviation), as appropriate

p-value calculated with chi-squared test for categorical values and Welch's t-test for continuous values.

[+]Percentage calculated with a total of 18 FA subjects due to limited information on FA resolution.

*Percentage calculated with a total of 19 FA subjects due to limited information on multiple allergies.

Children with FA were grouped using 3 different standards; FA resolution, egg allergy resolution, and multiple allergies. FA and egg allergy resolution were assessed using hospital records up to June 2020. Those with no symptoms, local or systemic, after consumption of the culprit food, were considered as resolution. Over half the FA subjects experience resolution in at least one food, while those with the resolution of egg allergy were less. 18 FA subjects had definite information on FA resolution because the other 2 subjects had indefinite reactions to food allergens. Also, there was 1 subject with unclear information on the number of allergic foods with definite symptoms. We removed those subjects for the respective analysis.

2. Metabolic profiling

A total of 344 metabolites were identified from LC-MS profiling. After removing metabolites that had over 10% missing values, 292 metabolites were included in further analysis. The data distribution pattern after normalization is shown in Figure 1. The distribution of the data before normalization is skewed to the left, but after normalization, it is close to a normal bell curve.

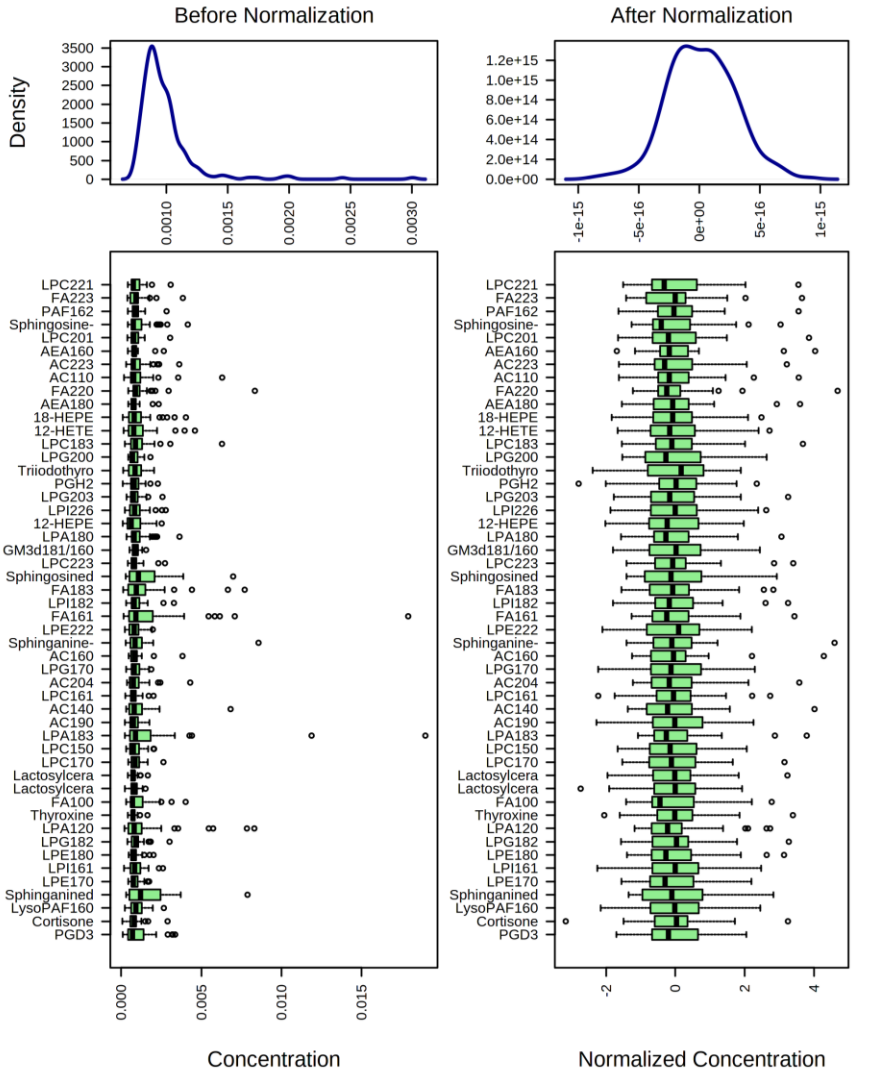


Figure 1. Data distribution results of metabolite data after normalization. The density plot after probabilistic quotient normalization shows an approximate bell curve. The box and whisker plots shows only the first 50 metabolites while the density plots are based on all data. Plots drawn using Metaboanalyst.

PCA results for the metabolite data are shown in Figure 2. The first dimension explains over 40% of the variation within the data set while the second dimension explains approximately 12% of the variation. Although the first and second dimensions explain over half the variation within the data, the two groups were not separated. However, because the original data does not follow a strict multivariate normal distribution and the whole variation is not present in the plots, it is hard to make definite conclusions about the results. Also, PCA often considers low variance components in the data as noise. However, considering the nature of metabolomics data, the slight difference in metabolites can have a great biological impact. Also, the PCA plots prove that there are no severe outliers within the data.

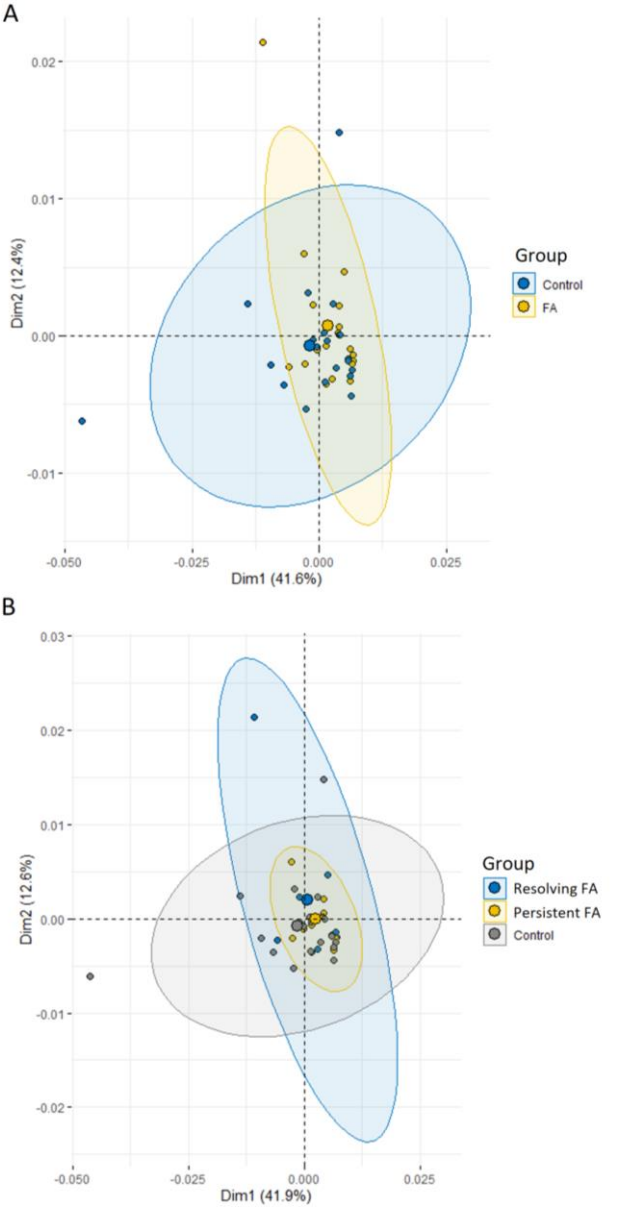


Figure 2. 2D Principal component analysis (PCA) plots with metabolite levels for FA and FA resolution. Unsupervised multivariate analysis of LC-MS data with (A) FA (yellow) and healthy controls (blue). (B) Unsupervised multivariate analysis with persistent FA (yellow), resolving FA(blue), and healthy controls (gray), FA; food allergy, Dim; dimension.

The results in Figure 3 for sPLS-DA and ortho PLS-DA show data patterns in supervised models. Compared to the FA samples, a higher variance was observed in the controls. The weight of each component in the sPLS-DA models was low with 16.5% and 13% for FA and 9.5% and 8.4% for FA resolution. The variance explained by the first two components were higher in FA. Resolving FA and persistent FA were relatively well separated, with boundaries barely overlapping. Ortho PLS-DA showed better separation results. An outlier was observed in the control group but it was not removed for further analysis.

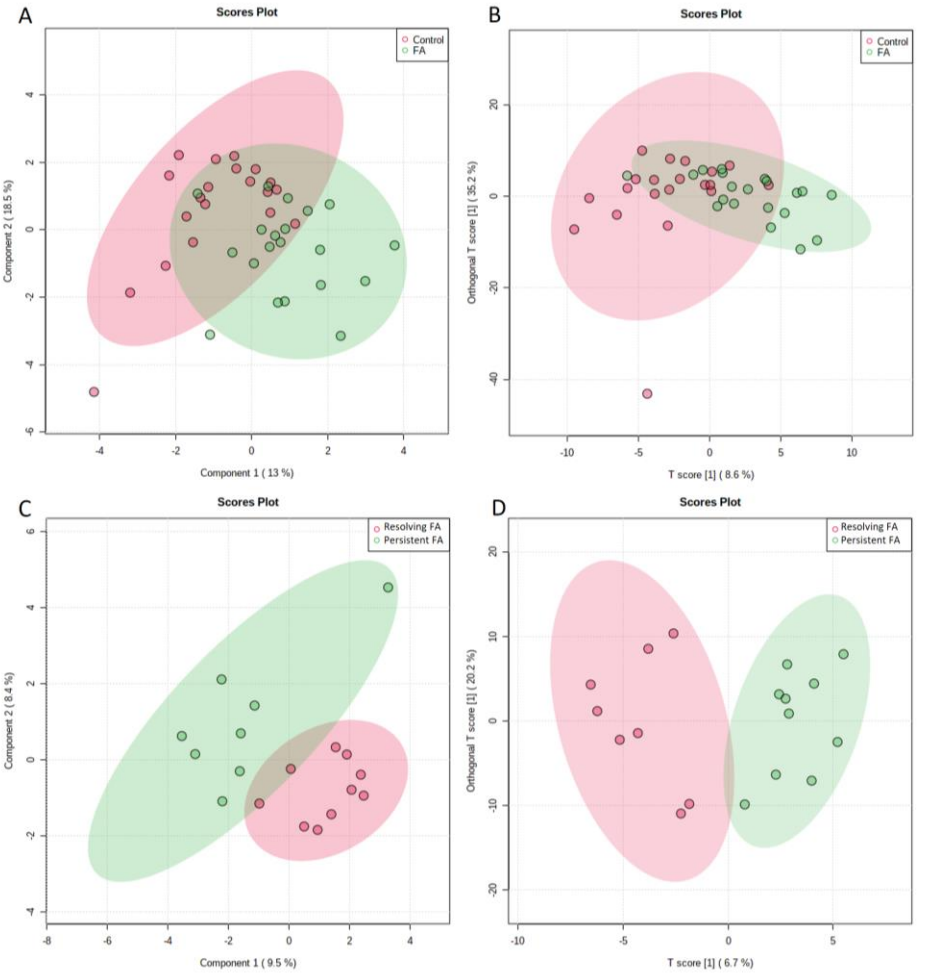


Figure 3. Partial least squares discriminant analysis (PLS-DA) plots with metabolite levels for FA and FA resolution. Supervised multivariate analysis of LC-MS data was applied. Scores scatter plot of model on FA (green) and healthy controls (red) with (A) sPLS-DA and (B) ortho PLS-DA and on persistent FA (green) and resolving FA (red) with (C) sPLS-DA and (D) ortho PLS-DA. Center scaling was used for the models. Plots drawn using Metaboanalyst. FA; food allergy.

3. Metabolites in correlation with age and total IgE levels

Table 2 shows the metabolites that are significantly correlated ($p < 0.01$) with age or total IgE levels. 15 and 9 metabolites had a significant correlation with age and total IgE levels. Most metabolites had a negative correlation with age and a positive correlation with total IgE level. These metabolites were removed from further analysis.

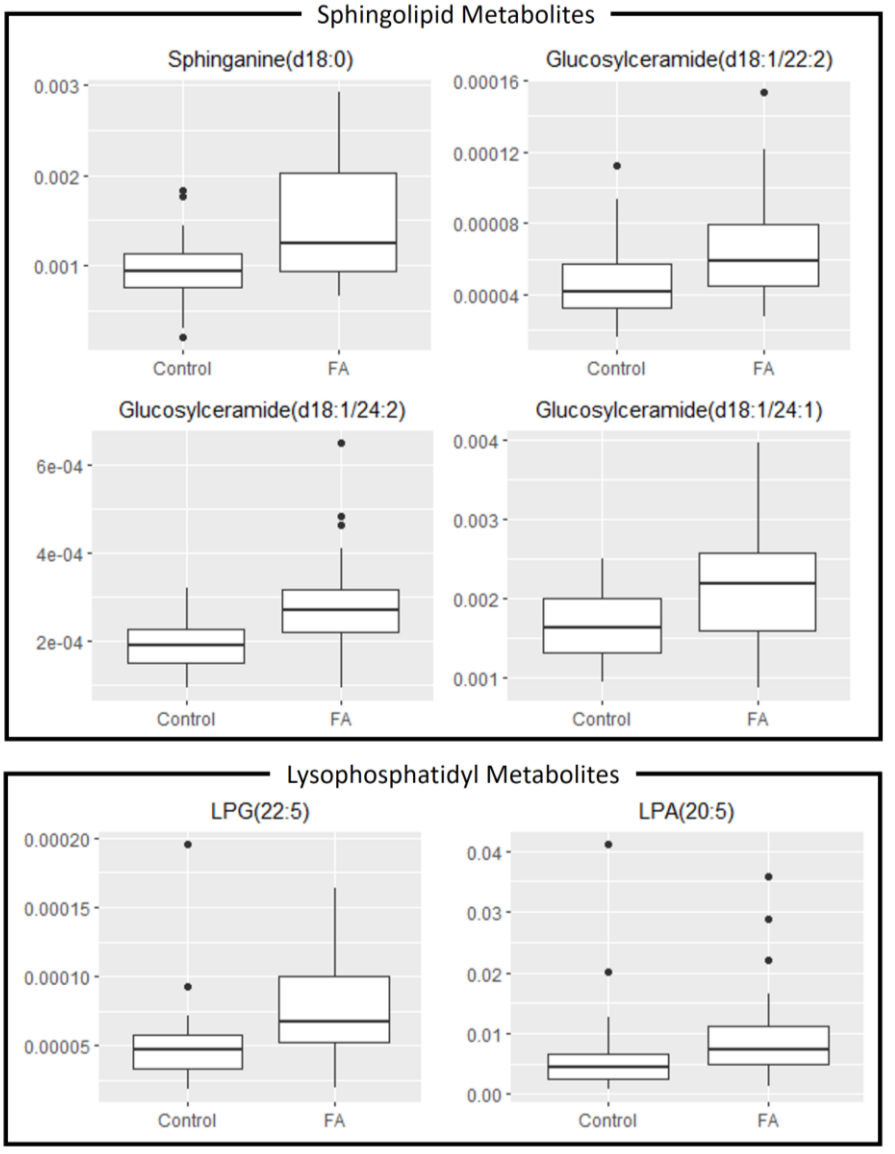
Table 2. Pearson's correlation between metabolite levels and age or total IgE levels

	Correlation Coefficient	Correlation p-value
Age		
FFA(10:0)	-0.40572	0.009393
FFA(12:0)	-0.41384	0.007942
Glucosylceramide(d18:1/14:0)	-0.48004	0.001721
GM3(d18:1/12:0)	-0.42841	0.005816
GM3(d18:1/14:0)	-0.55314	0.000214
Lactosylceramide(d18:1/16:0)	0.572128	0.000115
Lactosylceramide(d18:1/24:1)	0.411492	0.00834
Lactosylceramide(d18:1/24:2)	0.407275	0.009099
LPS(18:1)	-0.41726	0.007391
LPS(20:3)	-0.47532	0.001938
LPS(20:4)	-0.49013	0.001326
LPS(22:4)	-0.56264	0.000157
LPS(22:6)	-0.43521	0.005006
Sphingosine-1P(d16:1)	-0.45419	0.003241
TXB2	-0.53315	0.000397
Total IgE		
AC(19:0)	-0.41678	0.007466
AC(22:5)	0.439473	0.00455
LPC(16:1)	0.63386	0.0000113
LPC(22:5)	0.641033	0.0000083
LPE(16:1)	0.583873	0.0000765
LPE(20:5)	0.44593	0.003928
LPE(22:5)	0.750858	0.00000002
PAF(20:5)	0.661117	0.0000034
PGD3	0.430175	0.005596

FFA; free fatty acid, GM3; monosialodihexosylganglioside, LPS; lysophosphatidylserine, TXB2; thromboxane B2, AC; acylcarnitine, LPC; lysophosphatidylcholine, LPE; lysophosphatidylethanolamine, PAF; platelet activating factor, PGD3; prostaglandin D3, IgE; Immunoglobulin E.

4. Metabolites associated with FA and FA resolution

Pairwise comparison of children with FA and healthy control subjects revealed 15 metabolites significantly different (p-value < 0.05) between the groups. Figure 4 shows 12 of these metabolites with box plots to show the metabolite level differences. All sphingolipid metabolites were higher in FA. Five acylcarnitine metabolites with high carbon numbers and low double bonds were significantly associated with FA. For all five acylcarnitines, a common trend was observed, with a lower level in FA compared to control. Prostaglandin H2, a precursor for many other prostanoids, was higher in the control group.



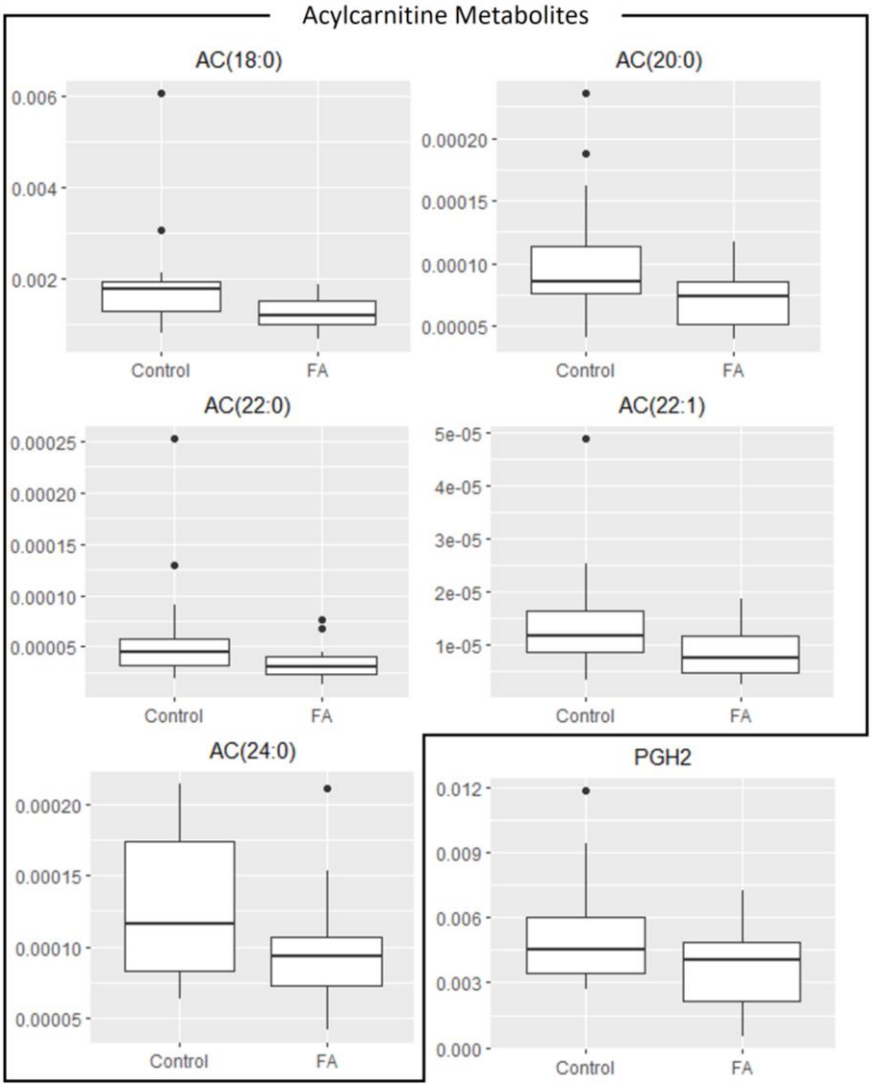
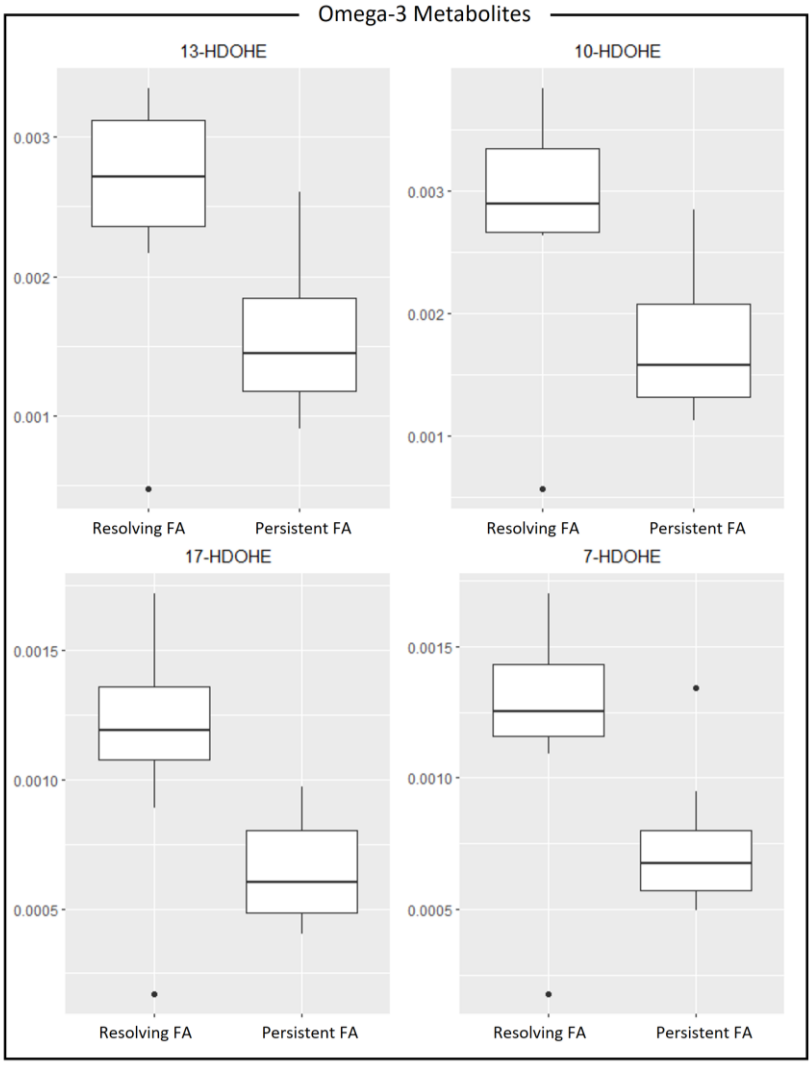


Figure 4. Metabolites significantly associated with FA. Metabolite levels were measured in children with FA (n=20) and healthy controls (n=20) by LC-MS. Sphingolipid metabolites, lysophosphatidyl metabolites and acylcarnitine metabolites are grouped together. The level of metabolites for the FA group and control group are significantly different ($p < 0.05$). LPG; lysophosphatidylglycerol, LPA; lysophosphatidic acid, AC; acylcarnitine, PGH2; prostaglandin H2.

Figure 5 shows metabolites significantly associated with FA resolution. Four different omega-3 metabolites, hydroxydocosahexaenoic acids (HDoHE), were highly associated with FA resolution, with a lower metabolite level observed in persistent FA for all cases. The level of omega-3 metabolites in the resolving FA group was an average 1.66 times higher compared to the persistent FA group. On the other hand, several lysophosphatidylcholines were higher in the resolving FA group with a consistent trend. Also, the level of PAF(16:1) was higher in persistent FA with a fold change of 1.2.



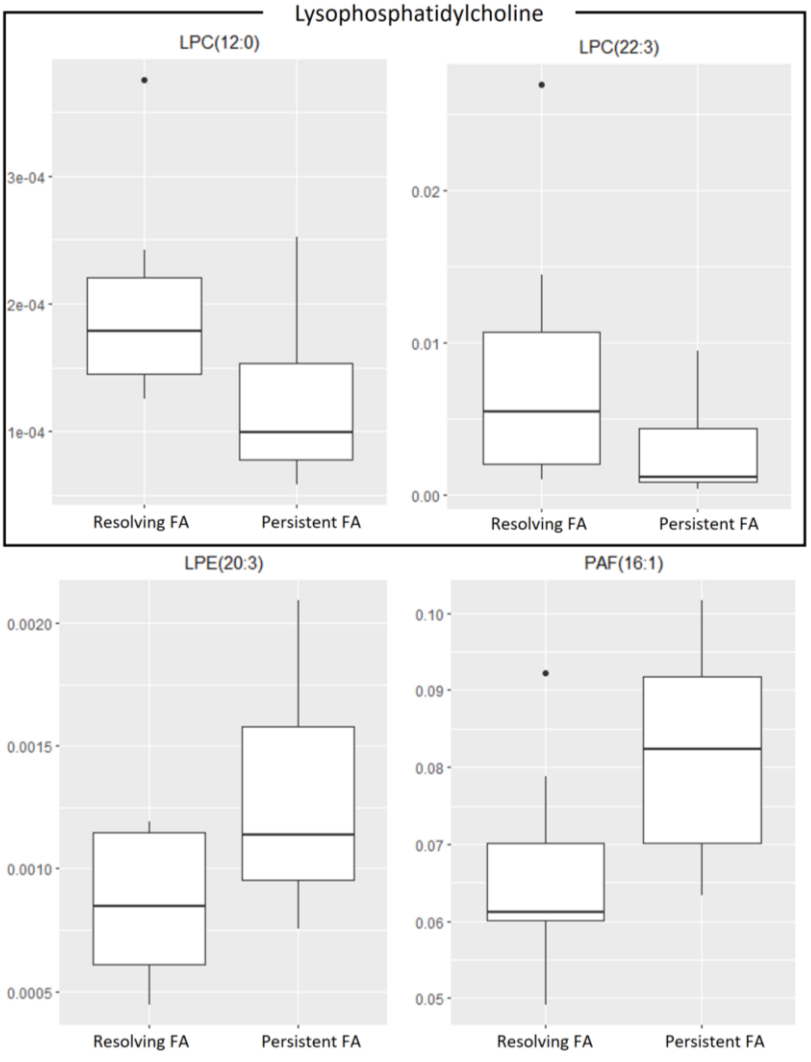


Figure 5. Metabolites significantly associated with FA resolution. Metabolite levels were measured in children with resolving FA (n=10) and persistent FA (n=8) by LC-MS. Omega-3 metabolites and lysophosphatidylcholine metabolites are grouped together. The level of metabolites for the resolving FA group and persistent FA group are significantly different ($p < 0.05$). HDoHE; hydroxydocosahexanoic acid, LPC; lysophosphatidylcholine, LPE; lysophosphatidylethanolamine, PAF; platelet-activating factor

To further examine the omega-3 metabolites, we compared the initial serum samples to the follow-up samples in order to observe changes through FA resolution. Figure 6 shows the four hydroxydocosahexanoic acids (HDoHE) metabolites significantly associated with FA resolution in the initial samples. In persistent FA, the metabolite level generally increases at follow up. However, in resolving FA, the levels show a decreasing trend. This trend is observed in all four HDoHE metabolites. The average fold change from initial to follow up in persistent FA was 1.38, 1.42, 1.77, 1.34 as compared to -1.47, -1.49, -1.44, -1.40 in resolving FA for 10-HDoHE, 13-HDoHE, 17-HDoHE, 7-HDoHE, respectively.

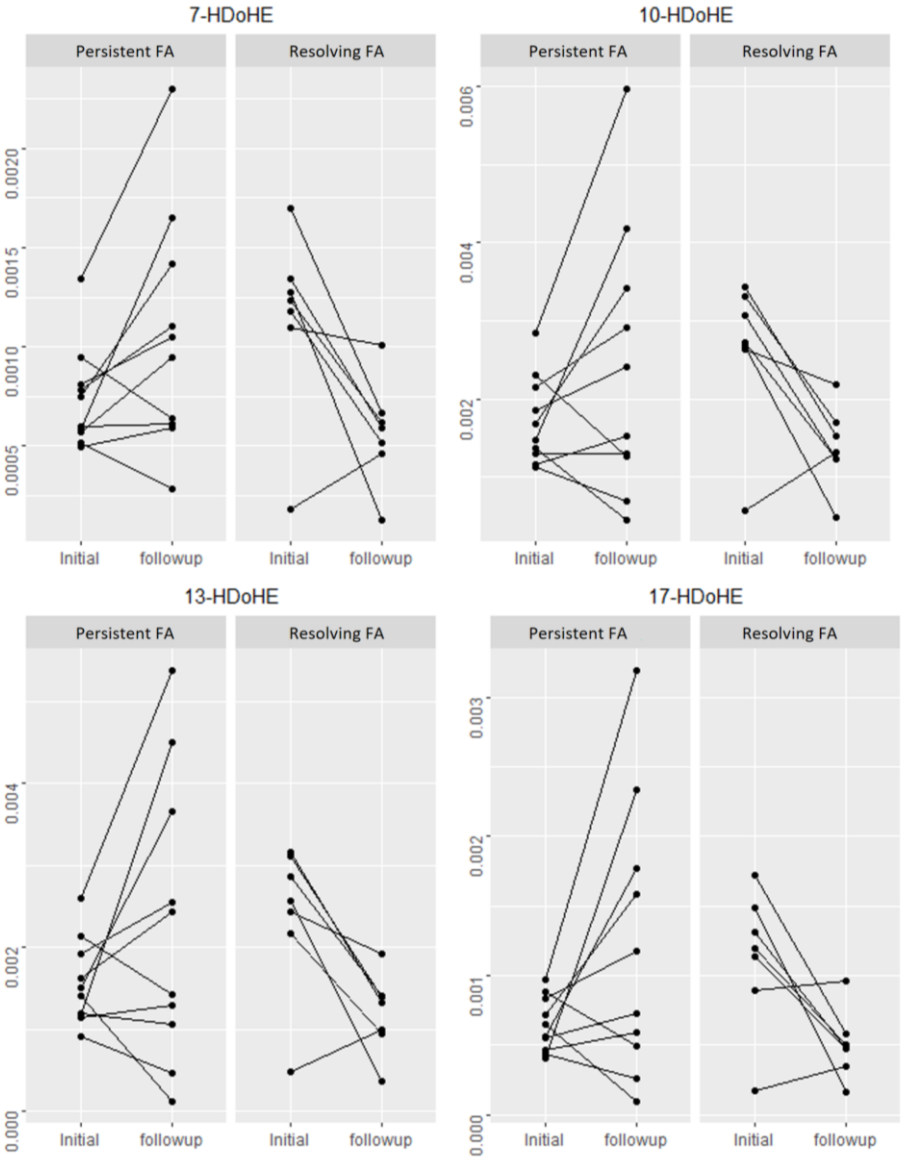


Figure 6. Omega-3 metabolite level changes between initial time point and follow-up. Metabolite levels were measured in 10 subjects with persistent FA and 7 subjects with resolving FA. In resolving FA, the follow up time point is after confirmed resolution of FA symptoms. HDoHE; hydroxydocosahexanoic acid.

Figure 7 shows heat maps drawn with the 15 and 11 metabolites significantly associated with FA and FA resolution, respectively. Hierarchical clustering applied to the metabolites shows that metabolites within the same metabolite class exhibit a similar level within samples. For example, in Figure 7A, glucosylceramide metabolites and acylcarnitines are clustered in the same group. Also, in Figure 7B, all the HDoHE metabolites, 17-HDoHE, 10-HDoHE, 13-HDoHE, 7-HDoHE, are clustered in the same group and show similar levels. Figure 7A also shows a pattern difference between resolving FA and persistent FA, with HDoHE metabolites high in resolving FA and several lysophosphatidic metabolites high in persistent FA. However, the distinction was not as clear with the FA phenotype.

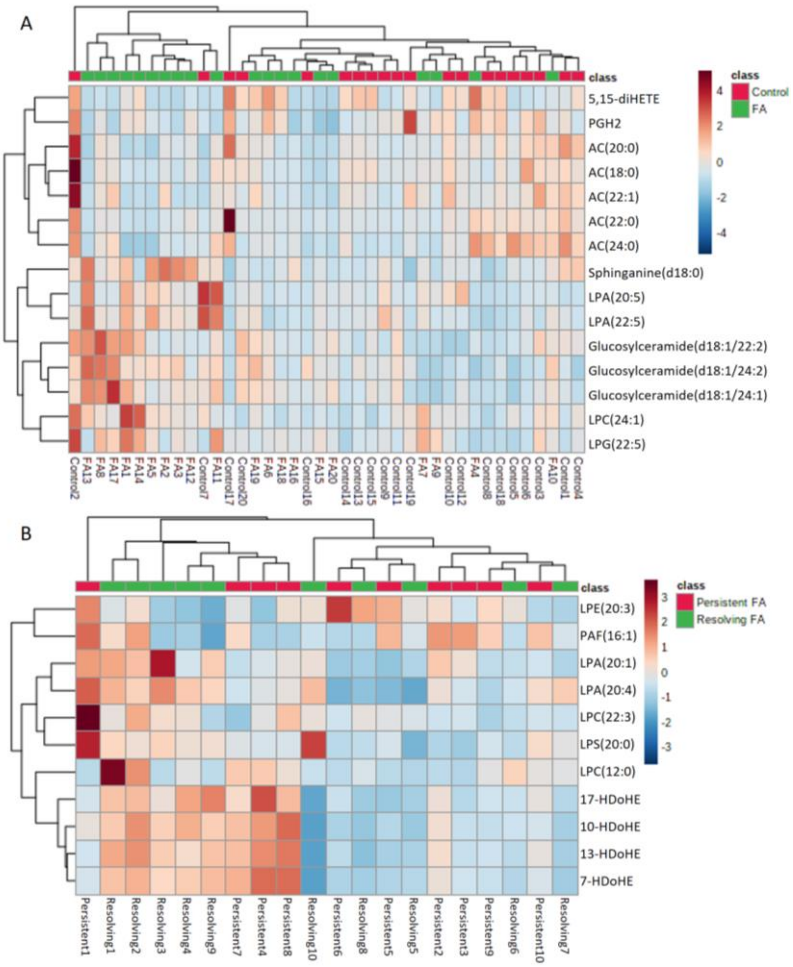


Figure 7. Heatmap with hierarchical clustering on metabolites significantly associated with FA and FA resolution. Heatmap with metabolites associated with (A) FA and (B) FA resolution. Both rows and columns are clustered using Euclidean distance measure and Ward clustering algorithm. Plots drawn using Metaboanalyst. diHETE; dihydroxyicosatetraenoic acid, PGH2; prostaglandin H2, AC; acylcarnitine, LPA; lysophosphatidic acid, LPC; lysophosphatidylcholine, LPG; lysophosphatidylglycerol, LPE; lysophosphatidylethanolamine, PAF; platelet activating factor, LPS; lysophosphatidylserine, HDoHE; hydroxydocosahexaenoic acid.

5. Association to FA related SNPs

From the 91 SNPs related to FA in previously published GWAS studies, 10 SNPs were associated ($p < 0.05$) with FA in our study subjects (Table 3). Nine of the SNPs were in linkage disequilibrium and located on the *Platelet Activating Factor Acetylhydrolase 1b Regulatory Subunit 1 (PAFAH1B1)* gene. It is located on chromosome 17 and encodes an acetylhydrolase which catalyzes platelet-activating factors. Two SNPs were associated with serum PAF levels. SNPs were located on the *Indoleamine 2,3-Dioxygenase 2* (IDO2) gene and *NLR Family Pyrin Domain Containing 3* (NLRP3) gene.

Table 3. SNPs significantly associated with FA phenotype or PAF levels

SNP	Position*	Alleles	MAF	p-value	Gene
FA					
rs4790356	17:2682935	A>G	0.163339	0.02337	PAFAH1B1
rs4790355	17:2680682	G>A	0.202077	0.02337	PAFAH1B1
rs4790353	17:2675354	G>A	0.216254	0.02337	PAFAH1B1
rs2317297	17:2673797	C>A	0.174121	0.02337	PAFAH1B1
rs3213697	17:2670564	C>G	0.107628	0.02337	PAFAH1B1
rs3213696	17:2670358	C>G	0.107628	0.02337	PAFAH1B1
rs12450722	17:2658859	G>A	0.117612	0.02337	PAFAH1B1
rs3785957	17:2657875	T>C	0.107827	0.02337	PAFAH1B1
rs8077351	17:2642179	A>G	0.072085	0.02337	PAFAH1B1
rs17389644	4:122576542	G>A	0.119609	0.04808	
PAF					
rs2955903	8:40015785	G>A	0.370607	0.03978	IDO2
rs4925654	1:247432755	A>G	0.177915	0.04021	NLRP3

*Position is based on GRCh38, MAF (minor allele frequency) based on 1000Genomes.

6. Metabolite QTL analysis

Metabolite QTL was done on 7-HDoHE, 10-HDoHE, 13-HDoHE, 17-HDoHE to identify 4, 29, 34, 25 metabolite-SNP pairs, respectively. Table 4 shows the Reactome pathway analysis results from the genes at the SNP locations. Several pathways significantly associated were affiliated to cell-cell communication and citric acid cycle metabolism. Also, 6 pathways were related to the cell cycle.

Table 4. Reactome pathway analysis results for HDoHE metabolite-SNP pairs in metabolite QTL results

Pathway	Parent pathway	p-value
Regulation of cytoskeletal remodeling and cell spreading by IPP complex components	Cell-Cell communication	0.004448
Activation of RAC1	Developmental Biology	0.007219
Cell-extracellular matrix interactions	Cell-Cell communication	0.009984
Citric acid cycle (TCA cycle)	Metabolism	0.012191
Pyruvate metabolism and citric acid (TCA) cycle	Metabolism	0.030245
Signaling by ERBB4	Signal Transduction	0.031873
GABA receptor activation	Neuronal System	0.032957
Loss of Nlp from mitotic centrosomes	Cell Cycle	0.037821
Loss of proteins required for interphase microtubule organization	Cell Cycle	0.037821
AURKA Activation by TPX2	Cell Cycle	0.039438
Recruitment of mitotic centrosome proteins and complexes	Cell Cycle	0.044275
Centrosome maturation	Cell Cycle	0.044275
Regulation of PLK1 Activity at G2/M Transition	Cell Cycle	0.047489

7. Quantification of metabolites

The ELISA results for PAF and PGH2 is presented in Figure 8. Two outliers were removed because their values were over 2 times the interquartile range above the third quartile for both metabolites. Of the FA subjects, 25 patients had persistent FA and 23 patients had resolving FA. Although the mean values of PAF levels were different between resolving FA and persistent FA, the groups did not have a significant difference. When only egg allergy resolution was considered, persistent egg allergy had a significantly higher level of PAF, with a mean value of 141.04 pg/mL compared to 121.23 pg/mL in resolving egg allergy. There were no differences in the level of PGH2 observed between FA and control. With resolving FA and persistent FA, the average level of PGH2 was lower in persistent FA even when only egg allergy was considered. However, the results were not significant and the variance within the data was big.

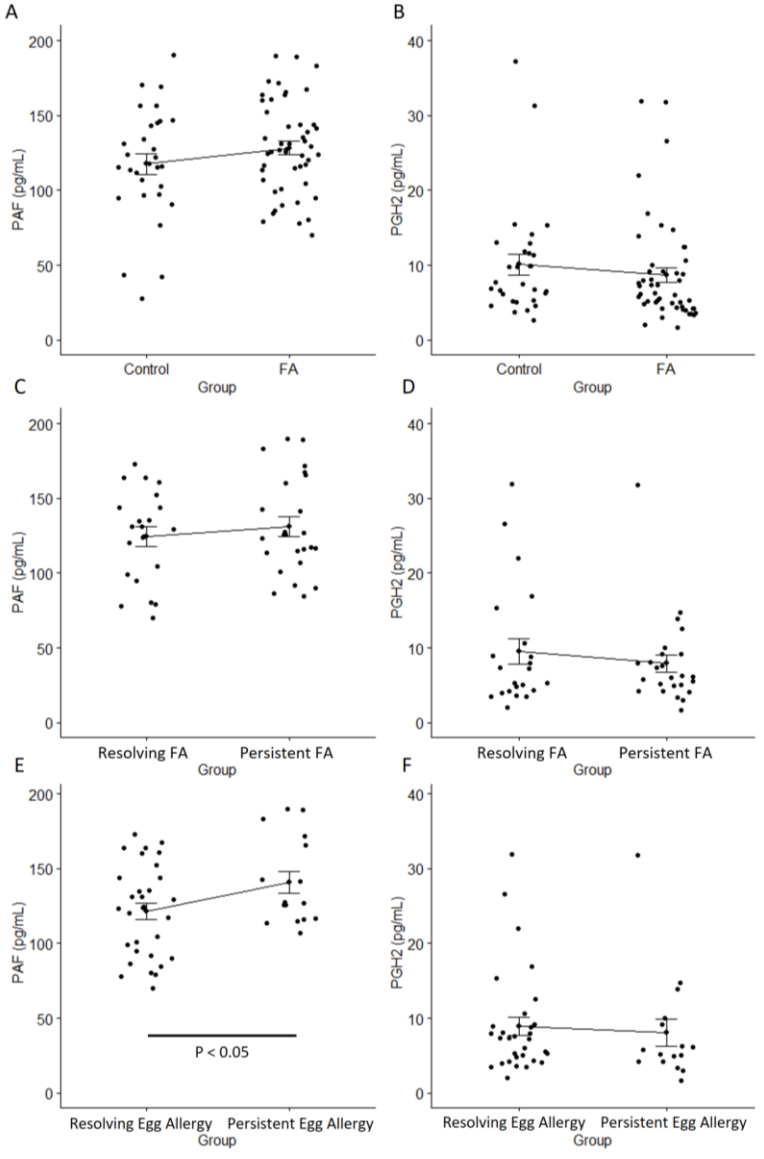


Figure 8. Level of PAF and PGH2 in FA and resolving FA patients. Metabolite levels were measured with ELISA using serum samples from 50 FA subjects and 30 healthy controls. Scatter plots are presented for (A) PAF levels in FA (B) PGH2 levels in FA (C) PAF levels in FA resolution (D) PGH2 levels in FA resolution (E) PAF levels in egg allergy resolution (F) PGH2 levels in egg allergy resolution. PAF; platelet activating factor, PGH2; prostaglandin H2.

IV. DISCUSSION

In this study, serum metabolic profiling was performed on children with FA. A unique metabolic signature was associated with FA and FA resolution. Through an integrated approach with both metabolomics and genomics, we identified several pathways that are associated with FA resolution. This study aimed to identify biomarkers for the development and prognosis of FA. Through the process, understanding the etiology and pathological mechanisms of FA can be made possible.

Several sphingolipid metabolites and acylcarnitine metabolites were significantly associated with FA. Sphingolipids constitute a class of lipids synthesized in the endoplasmic reticulum. They play significant roles in cell membranes and provide many bioactive metabolites that regulate cell signal transduction[39]. In our study, 4 types of sphingolipid metabolites, mainly glucosylceramides, were significantly increased in FA. In a previous study, a decrease in sphingomyelin and ceramide levels were observed in patients with FA[26]. As seen in Figure 9, both sphingomyelin and glucosylceramide are synthesized from ceramides. Hence, the increase in glucosylceramide levels might have been affected by the decrease in sphingomyelin and ceramides. Although the level of sphingomyelins and ceramides were not measured in our study, ceramide 1-P levels were measured and no difference was shown between FA and control subjects.

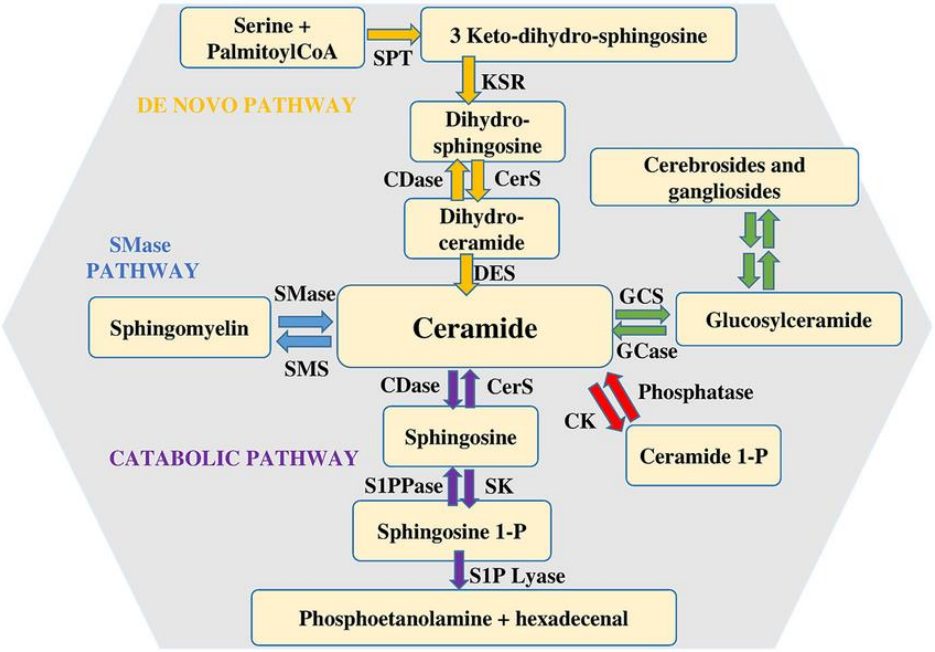


Figure 9. Metabolism of sphingolipids. Ceramide is the key metabolite that links the synthesis of sphingomyelin, glucosylceramide, sphingosine, and ceramide 1-P. Figure adapted from external source[40].

Also, changes in plasma sphingolipids have been reported in association with severe allergic respiratory responses[24]. In asthma, disruption in sphingolipid metabolism has been reported several times regarding the genotypic variation in the gene *ORMDL Sphingolipid Biosynthesis Regulator 3* (ORMDL3)[41,42]. Asthma is also an allergic disease that commonly coexists with FA in many children. In addition to the previous study, children with asthma had lower ceramide and sphingomyelin levels. Glucosylceramides might also be related to this process and have a role in allergic inflammation. The role of glucosylceramides is still ambiguous, but alterations in its level have been associated with cardiovascular disease, diabetes, and chronic inflammation[43].

In contrast to sphingolipid metabolites, acylcarnitine metabolites with long carbon chains and a low level of double bonds were significantly associated with FA. Long-chain acylcarnitines are by-products of incomplete beta-oxidation during fatty acid metabolism. It has been reported acylcarnitine activates pro-inflammatory signaling pathways, possibly through a mechanism involving pattern recognition receptors[44]. Plasma acylcarnitine concentrations were previously suggested as a biomarker for metabolism disorders and cardiovascular diseases[45], but there are no previous studies for allergic diseases.

Prostaglandin H2 was also significantly lower in FA. Prostaglandin H2 is synthesized from arachidonic acids and is a precursor for many other biologically significant molecules. It is a pro-inflammatory molecule and is related to platelet aggregation. However, the difference was not distinctive in

the quantification results from ELISA, possibly due to the large variation in metabolite levels.

Regarding FA resolution, the level of HDoHE metabolites were clearly higher in resolving FA compared to persistent FA. In addition to the difference in HDoHE levels between the FA and control group at the time of diagnosis, subjects with resolving FA showed a clear decrease with time-spanning through the resolution of FA. HDoHEs are essential lipid mediators involved in inflammation resolution. HDoHE is derived from docosahexaenoic acid (DHA), an omega-3 fatty acid. Exogenous DHA is converted by human platelets to 17-HDoHE and by human neutrophils mainly to 7- HDoHE[46]. 17-HDoHE produces resolvins and protectins, which act as active mediators of inflammation resolution.

Studies on omega-3 derivatives in allergic diseases have produced meaningful results. In mothers with low preexisting levels of DHA, reduced risk of FA, atopic dermatitis, and asthma was consistently observed[47]. Also in patients with severe asthma, decreased synthesis of protectin D1 in eosinophils was observed[48]. These omega-3 fatty acids act to oppose the actions of omega-6 fatty acids, particularly concerning eicosanoid synthesis, protecting against allergic manifestations[49]. Although HDoHE has a clear role as an intermediate product of DHA metabolism, whether it has a direct effect on anti-inflammatory mechanisms are still unknown. However, the high level of HDoHE may act as a marker for resolving FA, and a decreasing trend over time can solidify the

prediction.

Pathway analysis results of genes from HDoHE metabolite-SNP pairs highlights pathways related to cell-cell communication and metabolism. Especially, two pathways significantly associated with the genes are related to the citric acid cycle. The citric acid cycle has a central role in lipogenesis because acetyl-CoA from the cycle is synthesized to fatty acid and triglyceride. There have been previous reports on omega-3 fatty acids causing alterations in the citric acid cycle[50,51]. SNPs located within genes related to the citric acid cycle might cause dysregulation of omega-3 fatty acid metabolism.

Several lysophosphatydyl metabolites were associated with FA and FA resolution. Lysophophosphatidyl metabolites have a central role in cell membranes and many are presented as a phospholipid in the membrane acting in cell-mediated cell signaling and the activation of enzymes. From metabolites significantly associated with FA resolution, lysophosphatidylcholine has been reported to have a critical role in allergic airway disease manifestation, possibly via natural killer T cells[52]. It has been shown to aggravate contact hypersensitivity by promoting neutrophil infiltration and IL-17 expression[53]. Along with lysophosphatic acid, it has been considered several times for potential asthma biomarkers[54,55]. Since allergic asthma and FA have similar pathological mechanisms, lysophophatidic acid might act as a biomarker for FA and FA resolution.

The level of platelet-activating factor (16:1) also showed a significant

difference between subjects with resolving FA and persistent FA. The association between genotypic variation in the *Platelet Activating Factor Acetylhydrolase 1b Regulatory Subunit 1 (PAFAH1B1)* gene and FA phenotype might point towards a derangement of PAF activity. The level of PAF was also associated with rs2955903 and rs4925654, which encodes to IDO2 and NLRP3. PAFs are produced by various types of cells, including mast cells, neutrophils, eosinophils, and platelets, playing an important role in inflammatory diseases by amplifying acute inflammatory cascades[56,57]. Several published reports have linked the association between PAF and FA, especially food-induced severe anaphylaxis[58]. Circulating PAF levels correlated with the severity of anaphylaxis in humans[59] and possibly affect the biological system by increasing vascular permeability decreasing cardiac output, smooth muscle contraction of central airways, and circulatory collapse. A higher level of serum PAF in persistent FA might be related to this effect. However, ELISA results of PAF show that serum PAF levels are more associated with the resolution of egg allergy rather than a full resolution to all food allergens.

A key strength of our study is the recruitment of carefully phenotyped patients with children with confirmed egg allergy. Since allergy to different food allergens can induce different pathological mechanisms, unifying the allergen to egg white can control variability among subjects. Another strength is the availability of information on metabolite levels before and after FA resolution. This provides insight into the act of metabolites on resolution and

serves as an additional potential biomarker. Lastly, our study focused on lipid metabolites, which act as profound biological signals. Many metabolites associated with inflammatory pathways are lipids and we assessed these metabolites for a clear understanding of the crosswalk between inflammation signals.

There are also limitations to our study. First, the sample size was limited. We collected serum samples from 20 FA subjects and 20 control subjects, which may limit the interpretation of the results. Second, serum metabolite levels are affected by diet. Especially in children with FA, their restricted diet can cause an impact on metabolite levels and lipid metabolism due to microbiome variations. Further studies with a larger patient group and a controlled diet may provide more accurate insight into the metabolomic profile of FA.

V. CONCLUSION

In conclusion, our study discovered that children with FA, especially those with resolving FA exhibit a unique metabolomic signature. High levels of omega-3 metabolites might indicate resolving FA, especially when the level decrease with time. The combined analysis of genotypic and metabolomic data in FA patients shows distinct disease mechanisms that function in resolving FA, and this sheds light on the trajectory of FA pathology. Our observations suggest that metabolic profiling may be a powerful tool in providing insights into disease mechanisms and in identifying biomarkers for FA and FA resolution in children.

REFERENCES

1. Jones SM, Burks AW. Food Allergy. N Engl J Med 2017;377:1168-76.
2. Muraro A, Werfel T, Hoffmann-Sommergruber K, Roberts G, Beyer K, Bindslev-Jensen C, et al. EAACI food allergy and anaphylaxis guidelines: diagnosis and management of food allergy. Allergy 2014;69:1008-25.
3. Costa C, Coimbra A, Vitor A, Aguiar R, Ferreira AL, Todo-Bom A. Food allergy-From food avoidance to active treatment. Scand J Immunol 2020;91:e12824.
4. Renz H, Allen KJ, Sicherer SH, Sampson HA, Lack G, Beyer K, et al. Food allergy. Nat Rev Dis Primers 2018;4:17098.
5. Gupta RS, Springston EE, Warrier MR, Smith B, Kumar R, Pongracic J, et al. The prevalence, severity, and distribution of childhood food allergy in the United States. Pediatrics 2011;128:e9-17.
6. Kim M, Lee JY, Jeon HY, Yang HK, Lee KJ, Han Y, et al. Prevalence of Immediate-Type Food Allergy in Korean Schoolchildren in 2015: A Nationwide, Population-based Study. Allergy Asthma Immunol Res 2017;9:410-6.
7. Panel NI-SE, Boyce JA, Assa'ad A, Burks AW, Jones SM, Sampson HA, et al. Guidelines for the diagnosis and management of food allergy in the United States: report of the NIAID-sponsored expert panel. J Allergy Clin Immunol 2010;126:S1-58.
8. Sicherer SH, Sampson HA. Food allergy: A review and update on epidemiology, pathogenesis, diagnosis, prevention, and management. J Allergy Clin Immunol 2018;141:41-58.
9. Price A, Ramachandran S, Smith GP, Stevenson ML, Pomeranz MK, Cohen DE. Oral allergy syndrome (pollen-food allergy syndrome). Dermatitis 2015;26:78-88.
10. Cianferoni A. Non-IgE Mediated Food Allergy. Curr Pediatr Rev 2020;16:95-105.
11. Lack G. Update on risk factors for food allergy. J Allergy Clin Immunol 2012;129:1187-97.
12. Vermeulen EM, Koplin JJ, Dharmage SC, Gurrin LC, Peters RL, McWilliam V, et al. Food Allergy Is an Important Risk Factor for Childhood Asthma, Irrespective of Whether It Resolves. J Allergy Clin Immunol Pract 2018;6:1336-41 e3.
13. Savage J, Sicherer S, Wood R. The Natural History of Food Allergy. J Allergy Clin Immunol Pract 2016;4:196-203; quiz 4.
14. Thong BY, Hourihane JO. Monitoring of IgE-mediated food allergy

in childhood. Acta Paediatr 2004;93:759-64.
15. Mastrorilli C, Caffarelli C, Hoffmann-Sommergruber K. Food allergy and atopic dermatitis: Prediction, progression, and prevention. Pediatr Allergy Immunol 2017;28:831-40.
16. Sun YV, Hu YJ. Integrative Analysis of Multi-omics Data for Discovery and Functional Studies of Complex Human Diseases. Adv Genet 2016;93:147-90.
17. Johnson CH, Ivanisevic J, Siuzdak G. Metabolomics: beyond biomarkers and towards mechanisms. Nat Rev Mol Cell Biol 2016;17:451-9.
18. Li B, He X, Jia W, Li H. Novel Applications of Metabolomics in Personalized Medicine: A Mini-Review. Molecules 2017;22.
19. Schrimpe-Rutledge AC, Codreanu SG, Sherrod SD, McLean JA. Untargeted Metabolomics Strategies-Challenges and Emerging Directions. J Am Soc Mass Spectrom 2016;27:1897-905.
20. Guasch-Ferre M, Hruby A, Toledo E, Clish CB, Martinez-Gonzalez MA, Salas-Salvado J, et al. Metabolomics in Prediabetes and Diabetes: A Systematic Review and Meta-analysis. Diabetes Care 2016;39:833-46.
21. Cheng S, Shah SH, Corwin EJ, Fiehn O, Fitzgerald RL, Gerszten RE, et al. Potential Impact and Study Considerations of Metabolomics in Cardiovascular Health and Disease: A Scientific Statement From the American Heart Association. Circ Cardiovasc Genet 2017;10.
22. Veiga S, Wahrheit J, Rodriguez-Martin A, Sonntag D. Quantitative Metabolomics in Alzheimer's Disease: Technical Considerations for Improved Reproducibility. Methods Mol Biol 2018;1779:463-70.
23. Gahleitner F, Guallar-Hoyas C, Beardsmore CS, Pandya HC, Thomas CP. Metabolomics pilot study to identify volatile organic compound markers of childhood asthma in exhaled breath. Bioanalysis 2013;5:2239-47.
24. Obeso D, Mera-Berriatua L, Rodriguez-Coira J, Rosace D, Fernandez P, Martin-Antoniano IA, et al. Multi-omics analysis points to altered platelet functions in severe food-associated respiratory allergy. Allergy 2018;73:2137-49.
25. Peeters KA, Lamers RJ, Penninks AH, Knol EF, Bruijnzeel-Koomen CA, van Nesselrooij JH, et al. A search for biomarkers as diagnostic tools for food allergy: a pilot study in peanut-allergic patients. Int Arch Allergy Immunol 2011;155:23-30.
26. Crestani E, Harb H, Charbonnier LM, Leirer J, Motsinger-Reif A, Rachid R, et al. Untargeted metabolomic profiling identifies disease-specific signatures in food allergy and asthma. J Allergy

Clin Immunol 2020;145:897-906.
27. Maeda S, Nakamura T, Harada H, Tachibana Y, Aritake K, Shimosawa T, et al. Prostaglandin D2 metabolite in urine is an index of food allergy. Sci Rep 2017;7:17687.
28. Suaini NHA, Wang Y, Soriano VX, Martino DJ, Allen KJ, Ellis JA, et al. Genetic determinants of paediatric food allergy: A systematic review. Allergy 2019;74:1631-48.
29. Kim KW, Ober C. Lessons Learned From GWAS of Asthma. Allergy Asthma Immunol Res 2019;11:170-87.
30. Bersanelli M, Mosca E, Remondini D, Giampieri E, Sala C, Castellani G, et al. Methods for the integration of multi-omics data: mathematical aspects. BMC Bioinformatics 2016;17 Suppl 2:15.
31. Ren S, Shao Y, Zhao X, Hong CS, Wang F, Lu X, et al. Integration of Metabolomics and Transcriptomics Reveals Major Metabolic Pathways and Potential Biomarker Involved in Prostate Cancer. Mol Cell Proteomics 2016;15:154-63.
32. Dieterle F, Ross A, Schlotterbeck G, Senn H. Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics. Anal Chem 2006;78:4281-90.
33. Chong J, Wishart DS, Xia J. Using MetaboAnalyst 4.0 for Comprehensive and Integrative Metabolomics Data Analysis. Curr Protoc Bioinformatics 2019;68:e86.
34. Worley B, Powers R. Multivariate Analysis in Metabolomics. Curr Metabolomics 2013;1:92-107.
35. Arnold M, Raffler J, Pfeufer A, Suhre K, Kastenmuller G. SNiPA: an interactive, genetic variant-centered annotation browser. Bioinformatics 2015;31:1334-6.
36. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559-75.
37. Shabalin AA. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. Bioinformatics 2012;28:1353-8.
38. Wang K, Li M, Bucan M. Pathway-based approaches for analysis of genomewide association studies. Am J Hum Genet 2007;81:1278-83.
39. Gault CR, Obeid LM, Hannun YA. An overview of sphingolipid metabolism: from synthesis to breakdown. Adv Exp Med Biol 2010;688:1-23.
40. Romero-Guevara R, Cencetti F, Donati C, Bruni P. Sphingosine 1-phosphate signaling pathway in inner ear biology. New

therapeutic strategies for hearing loss? Front Aging Neurosci 2015;7:60.
41. Ono JG, Kim BI, Zhao Y, Christos PJ, Tesfaigzi Y, Worgall TS, et al. Decreased sphingolipid synthesis in children with 17q21 asthma-risk genotypes. J Clin Invest 2020;130:921-6.
42. Worgall TS. Sphingolipids, ORMDL3 and asthma: what is the evidence? Curr Opin Clin Nutr Metab Care 2017;20:99-103.
43. Messner MC, Cabot MC. Glucosylceramide in humans. Adv Exp Med Biol 2010;688:156-64.
44. Rutkowsky JM, Knotts TA, Ono-Moore KD, McCoin CS, Huang S, Schneider D, et al. Acylcarnitines activate proinflammatory signaling pathways. Am J Physiol Endocrinol Metab 2014;306:E1378-87.
45. Makrecka-Kuka M, Sevostjanovs E, Vilks K, Volska K, Antone U, Kuka J, et al. Plasma acylcarnitine concentrations reflect the acylcarnitine profile in cardiac tissues. Sci Rep 2017;7:17528.
46. Fischer S, von Schacky C, Siess W, Strasser T, Weber PC. Uptake, release and metabolism of docosahexaenoic acid (DHA, c22:6 omega 3) in human platelets and neutrophils. Biochem Biophys Res Commun 1984;120:907-18.
47. Venter C, Meyer RW, Nwaru BI, Roduit C, Untersmayr E, Adel-Patient K, et al. EAACI position paper: Influence of dietary fatty acids on asthma, food allergy, and atopic dermatitis. Allergy 2019;74:1429-44.
48. Miyata J, Fukunaga K, Iwamoto R, Isobe Y, Niimi K, Takamiya R, et al. Dysregulated synthesis of protectin D1 in eosinophils from patients with severe asthma. J Allergy Clin Immunol 2013;131:353-60 e1-2.
49. Miles EA, Calder PC. Can Early Omega-3 Fatty Acid Exposure Reduce Risk of Childhood Allergic Disease? Nutrients 2017;9.
50. Ahmed AA, Balogun KA, Bykova NV, Cheema SK. Novel regulatory roles of omega-3 fatty acids in metabolic pathways: a proteomics approach. Nutr Metab (Lond) 2014;11:6.
51. D'Alessandro A, D'Amici GM, Timperio AM, Merendino N, Zolla L. Docosohaexanoic acid-supplemented PACA44 cell lines and over-activation of Krebs cycle: an integrated proteomic, metabolomic and interactomic overview. J Proteomics 2011;74:2138-58.
52. Bansal P, Gaur SN, Arora N. Lysophosphatidylcholine plays critical role in allergic airway disease manifestation. Sci Rep 2016;6:27430.
53. Song MH, Gupta A, Kim HO, Oh K. Lysophosphatidylcholine

aggravates contact hypersensitivity by promoting neutrophil infiltration and IL17 expression. BMB Rep 2020.

54. Ackerman SJ, Park GY, Christman JW, Nyenhuis S, Berdyshev E, Natarajan V. Polyunsaturated lysophosphatidic acid as a potential asthma biomarker. Biomark Med 2016;10:123-35.
55. Yoder M, Zhuge Y, Yuan Y, Holian O, Kuo S, van Breemen R, et al. Bioactive lysophosphatidylcholine 16:0 and 18:0 are elevated in lungs of asthmatic subjects. Allergy Asthma Immunol Res 2014;6:61-5.
56. Palgan K, Bartuzi Z. Platelet activating factor in allergies. Int J Immunopathol Pharmacol 2015;28:584-9.
57. Zimmerman GA, McIntyre TM, Prescott SM, Stafforini DM. The platelet-activating factor signaling system and its regulators in syndromes of inflammation and thrombosis. Crit Care Med 2002;30:S294-301.
58. Vadas P. The platelet-activating factor pathway in food allergy and anaphylaxis. Ann Allergy Asthma Immunol 2016;117:455-7.
59. Vadas P, Gold M, Perelman B, Liss GM, Lack G, Blyth T, et al. Platelet-activating factor, PAF acetylhydrolase, and severe anaphylaxis. N Engl J Med 2008;358:28-35.

ABSTRACT(IN KOREAN)

대사체 연구를 통한 식품알레르기 바이오마커 발굴

< 지도교수 김경원 >

연세대학교 대학원 의과학과

장 혜 린

식품알레르기는 점점 더 많은 어린이들에게 영향을 미치게 되면서 중요한 건강 문제로 대두되면서, 소아 식품알레르기는 때때로 증상들이 자연적으로 호전되기도 한다. 그러나 식품알레르기 호전의 원리는 아직 심도 있게 연구되지 않았다. 본 연구는 대사체학을 통해 식품알레르기와 그의 호전을 예측할 수 있는 바이오마커를 개발하는 것을 목적으로 하였다.

식품알레르기 환자 20명과 대조군 20명의 대사체 정보를 액체 크로마토그래피 질량분석법을 사용하여 연구하였다.

식품알레르기 환자의 혈청 검체를 사용하였으며, 식품알레르기 진단 시점과 호전 시점에 검체를 채취하여 대사물 수치의 변화를 파악하였다. 또한, 유전자형 분석을 실시하여 대사과정과 연관된 경로를 분석하였다.

식품알레르기 환자들은 대조군과는 확연히 다른 대사체를 보였다. Sphingolipid의 증가와 acylcarnitine의 감소가 식품알레르기와 연관되어 있었다. 또한, 식품알레르기가 호전되는 환자들에게서는 오메가-3 대사물이 많이 관찰되었다. 오메가-3 대사물의 수준은 식품알레르기가 호전되면서 같이 감소했는데, 같은 기간 동안 식품알레르기가 호전되지 않은 환자에서는 증가하는 양상을 보였다. 또한, 혈소판 활성화 인자와 lysophosphatidylcholine은 식품알레르기 호전과 밀접한 관련이 있었다.

여러 지질 대사물이 소아 식품알레르기의 발병과 호전에 밀접한 관련이 있으며 이것은 식품알레르기를 예측할 수 있는 바이오마커로서 활용될 수 있음을 시사한다.

핵심되는 말: 식품알레르기, 대사체학, 오메가-3, 바이오마커