UNIVERSITÄTS-
KINDERSPITAL
ZÜRICH
Medizinische Klinik
Prof. Dr. med. Michael Grotzer

Respiratory Medicine
Prof. Dr. med. Alexander Möller
Head Respiratory Medicine, Cystic Fibrosis Center and Pediatric Sleep Center

Kinderspital Zürich - Eleonorenstiftung
Steinwiesstrasse 75
CH-8032 Zürich
www.kispi.uzh.ch

Sekretary Ms Yesim Karaman
Telefon +41 (0)44 266 7203
Telefax +41 (0)44 266 7670
yesim.karaman@kispi.uzh.ch

Kinderspital Zürich, Pneumologie, Postfach, 8032 Zürich

To the editor of The Lancet Respiratory Medicine Journal

Zurich, Wednesday, 18 May 2022 / MOA

Our manuscript: Asthma in one breath - Metabolic signatures for allergic asthma in children by online breath analysis: an observational study

Hereby we submit our original research manuscript titled: “Asthma in one breath - Metabolic signatures for allergic asthma in children by online breath analysis: an observational study”

We confirm that this work is original and has not been submitted or published anywhere else.

The analysis of breath as a non-invasive method to assess inflammation and metabolic processes within not only the airways and lungs but also the entire body has gained a lot of interest in recent years. The use of high-resolution mass spectrometry (HRMS) allows the simultaneous detection of several thousands of molecules in exhaled breath. We are the first to apply secondary electrospray ionization high-resolution mass spectrometry (SESI-HRMS) in children with allergic asthma and healthy controls. We detected 375*m/z* features that were significantly different ($p < 0.05$) between children with asthma and the controls, 134 of which could be identified due to the high resolution of the mass spectrometer. Many of these could be grouped to metabolites of common pathways or chemical families. We found several pathways that are well-represented by the significant metabolites, for example lysine degradation elevated in the asthmatic group and two arginine pathways elevated in the healthy control group.
We believe that our results are newsworthy and very interesting for the readers of the Lancet Respiratory Medicine Journal, since such an extensive metabolic signature for allergic asthma has so far not been described in exhaled breath. Until now, only a few identified molecules were reported in various studies. As we put a large effort into compound identification, our data allows for in-depth pathway analysis revealing the underlying connections between the identified metabolites.
Many of the volatile organic compounds (VOCs) that we were able to identify in breath are linked to well-described metabolic pathways involved in the pathophysiology of asthma. We think that our findings represent a relevant step forward in understanding the value of breath analysis and are optimistic that a subset of the identified VOCs has a high potential for future clinical diagnostic applications.

Thank you for your consideration of this manuscript.

Prof. Dr. med. Alexander Moeller

Asthma in one breath - Metabolic signatures for allergic asthma in children by online breath analysis: an observational study

R. Weber[1†], B. Streckenbach[2,3†], L. Welti[1], D. Inci[1], M. Kohler[3], N. Perkins[4], R. Zenobi[2], S. Micic[1*], A. Moeller[1*]
[1]University Children's Hospital Zurich, Department of Respiratory Medicine and Children's Research Center, Zurich, Switzerland
[2]ETH Zurich, Department of Chemistry and Applied Biosciences, Zurich, Switzerland
[3]Department of Pulmonology, University Hospital Zurich, Zurich, Switzerland
[4]University Children's Hospital Zurich, Division of Clinical Chemistry and Biochemistry, Zurich, Switzerland

*Corresponding authors:

Alexander Moeller
Steinwiesstrasse 75
8032 Zurich
Switzerland
+41(0)442667079
alexander.moeller@kispi.uzh.ch

Srdjan Micic
Steinwiesstrasse 75
8032 Zurich
Switzerland
+41(0)442667079
srdjan.micic@kispi.uzh.ch

†These authors contributed equally to this work.

Abstract

Background: There is a need to improve the diagnosis and management of pediatric asthma. Breath analysis aims to address this by non-invasively assessing altered metabolism and disease-associated processes. Our goal was to identify exhaled metabolic signatures that distinguish children with allergic asthma from healthy controls using secondary electrospray ionization high-resolution mass spectrometry (SESI-HRMS).

Methods: Breath analysis was performed via a SESI source linked to a high-resolution mass spectrometer. Significant *m/z* features in breath were extracted using the empirical Bayes

moderated t-statistics test. Corresponding molecules were identified by MS^2 database matching and pathway analysis.

Findings: 48 allergic asthmatics and 56 healthy controls were included in the study. Among 375 significant *m/z* features, 134 were putatively identified. Many of these could be grouped to metabolites of common pathways or chemical families. We found several pathways that are well-represented by the significant metabolites, for example lysine degradation elevated in the asthmatic group and two arginine pathways in the healthy group. Assessing the ability of breath profiles to classify samples as asthmatic or healthy in a leave-one-out cross-validation revealed an area under the curve of 0·85.

Interpretation: For the first time, a large number of breath-derived volatile organic compounds (VOCs) that discriminate children with allergic asthma from healthy controls were identified. Many are linked to well-described metabolic pathways and chemical families involved in pathophysiological processes of asthma. Furthermore, a subset of these VOCs showed high potential for clinical diagnostic applications.

Funding: Evi Diethelm-Winteler Foundation, Childhood Research Center, Heidi Ras Stiftung, Zurich Foundation, Lotte und Adolf Hotz Stiftung

Research in context

Evidence before this study

Several previous studies applied breath analysis in children with allergic asthma, but none of them used secondary electrospray ionization high-resolution mass spectrometry (SESI-HRMS) as the methodology. These prior studies have reported a small set of significant, but mostly unrelated volatile organic compounds (VOCs) and only a handful of them were detected in more than one study. A literature search was performed on PubMed on the 21st of February 2022 containing the keyword “asthma” in the title and the following keywords in the title or abstract: “breath”, “volatile organic compounds” and “children”. The search resulted in only 18 hits, seven of which were excluded due to the following criteria: electronic nose (2), not in children (1), assessing environmental VOCs (1), unrelated to this study (3). Additionally, we considered publications that compared children with asthma and a healthy group, and were featured in two recent reviews about biomarkers for pediatric asthma. We used the studies obtained from this search to compare the significant compounds from our study to previous results with different methodologies. We also compared our detected significant *m/z* features to the ones previously identified in publications by SESI-HRMS.

Added value of this study

This is the first reported study that applies online breath analysis by mass spectrometry to compare molecular breath signatures of children with allergic asthma and healthy controls. Compared to previous breath analysis studies, we report a larger amount of identified compounds that match to multiple metabolic pathways and chemical families, for both study groups. Interestingly, several of these upregulated pathways align well with known pathophysiologic processes of asthma. Additionally, we show that the metabolic profiles associated with allergic asthma can possibly be used for diagnostic purposes.

Implications of all the available evidence

Breath analysis is ideal for pediatric applications and has the potential to revolutionize the diagnosis of asthma in young children, as it provides a non-invasive way of assessing inflammatory and metabolic changes caused by the disease. This is based on the various metabolites that can be detected in breath by SESI-HRMS simultaneously and with high sensitivity, as has been demonstrated in several studies. Our study adds value to the molecular pathophysiology of asthma by exploring easily-accessible metabolic information from exhaled breath by SESI-HRMS. Additionally, we were able to detect almost entire metabolic pathways and chemical families, which is a clear novelty compared to previous studies that reported predictive biomarkers, often without a relevant pathophysiological connection to the underlying disease. The main challenges for the future will be to validate the relevance of the identified compounds and the transition into clinical diagnostic routine.

Introduction

Asthma is the most frequent chronic condition in children in the developed world. The disease is very heterogeneous in its presentation and clinical course. Due to the lack of a well recognized and easy to apply diagnostic gold-standard, misdiagnosis is very common. This has negative impacts on asthma related morbidity, quality-of-life, medication side-effects, prognosis and health costs. Therefore, the investigation of pediatric asthma and its associated molecular processes including airway inflammation is of high importance for the development of novel, much-needed diagnostic and monitoring applications.

Breath is known to contain several hundreds of metabolites that reflect metabolism as well as disease-specific mechanisms such as airway inflammation.[1] Therefore, there is great interest in discovering endogenous exhaled volatile organic compounds (VOCs) that are linked to diseases and their pathophysiological processes.[2] One of the few clinical tests taking advantage of this is the quantification of exhaled fractional nitric oxide (FeNO), which can be measured in all age groups. FeNO is a biomarker for eosinophilic airway inflammation that is related to allergic asthma.[1] This exemplifies the potential of applying breath analysis to further study allergic asthma and improve the diagnostic power of exhaled biomarkers.

Several breath analysis studies attempted to distinguish children with asthma from healthy controls by different techniques. Dallinga and colleagues compared exhaled breath of children with asthma and a healthy group by gas chromatography mass spectrometry (GC-MS) and identified a small set of discriminatory VOCs that is potentially related to lipid peroxidation, including various hydrocarbons, xylene, benzoic acid, and butanoic acid.[3] A pilot study from van Mastrigt *et al.* identified VOCs discriminating children with asthma, cystic fibrosis (CF) and healthy controls by using a broadband quantum cascade laser spectroscopy technique.[4] The distinguishing compound classes included different carboxylic acids, esters, and ethers. Altogether, there is only little overlap between the detected metabolites of different studies and sometimes even conflicting results are reported. Therefore standardization as well as external validation are challenges that need further research as summarized in recent reviews.[1,2,5]

Secondary electrospray ionization high-resolution mass spectrometry (SESI-HRMS) is a technology applied for online breath analysis that links real-time measurements without sample preparation to high mass resolution.[6] The latter strongly improves the confidence in compound identification of the detected mass-to-charge features (*m/z* features). Previous studies confirmed the potential of this technology to identify relevant VOCs, including biological metabolites, and reveal altered molecular pathways for different respiratory diseases.[6,7,8] Furthermore, its applicability in children was confirmed in our previous study on cystic fibrosis.[8]

The aim of this study was to identify metabolic signatures in exhaled breath consisting of discriminating VOCs specific to allergic asthma in children by SESI-HRMS. Moreover, we were interested in gaining a deeper understanding of the biological significance of our data. Breath analysis has the potential to improve diagnosis and management of asthma by providing a rapid and non-invasive approach of assessment, which is highly attractive for pediatric applications. Additionally, it might also reveal novel findings about the underlying molecular pathophysiology of asthma.

Methods

Study design, participants and clinical data

This observational study included children with allergic asthma and healthy controls, aged 5 to 18 years. Asthmatic patients from the outpatient clinic of the University Children's Hospital Zürich, Switzerland, were recruited for this study. Only patients with a definite diagnosis of allergic asthma were considered, *i.e.* bronchodilator reversibility (BDR) testing > 12% and/or 200 mL, and/or a positive bronchial challenge test according to the recently published practice guidelines for asthma diagnosis in school children.[9] Further, eligible patients were stable enough to temporarily stop the inhalation of asthma medication at least one week before the measurements and the patients and/or parents gave their written informed consent in advance. Clinical data was collected on the same day as breath analysis and is summarized together with anthropometric data in Table 1. The measurement and recruitment period was in parallel and lasted for 15 months. Efforts were put into recruiting participants of both cohorts at a randomized schedule across daytime and throughout the study period. The sample size was based on our previous study with a similar design.[8] The study was approved by the local ethics committee (KEK-ZH ID2018-00441) and was conducted in accordance with the Declaration of Helsinki.

Breath analysis

Online breath analysis was performed using a secondary electrospray ionization source (SuperSESI, FIT FossilionTech, Madrid, Spain) connected to a high-resolution time-of-flight (TOF) mass spectrometer (TripleTOF 5600+, AB Sciex, Concord, ON, Canada). In brief, children were exhaling directly into the instrument in a sitting position. The breathing maneuver consisted of at least three long exhalations at a constant pressure of 5 mbar with short breaks in between. Measurements were recorded in positive and negative ionization mode (*m/z*-range 50 - 500 Da). Methodological details and instrumental settings were previously described by our group[8], minor adaptations are specified in the supplementary material.

Data preprocessing

All data were recalibrated in PeakView 2.2 (AB Sciex, Concord, Canada) and processed in R version 4.1.1 (R Foundation for Statistical Computing, Vienna, Austria). The conversion and preprocessing of the raw data was done in the same way as described in our previous work.[8] In brief, the raw mass spectra were resampled by interpolation (Δ*m/z*: 0·0005, *m/z*-range: 50 - 500 Da), peak picking was performed on the average mass spectra associated with exhalation and signal intensities of the *m/z* features were determined by trapezoidal integration. The intensities of the *m/z* features were normalized to the total ion current (TIC), $\log_2$-transformed and arranged into a data matrix of breath profiles for further analysis. More details on data preprocessing are given in the supplemental information.

Statistical analysis

To account for unmodeled factors, batch adjustment was performed by applying surrogate variable analysis (SVA)[10] on the data matrix of breath profiles. Identification of differentially expressed *m/z* features when comparing cases and controls was assessed by the empirical Bayes moderated t-statistics test[11]. Correction for multiple hypothesis testing was conducted using Benjamin-Hochberg procedure[12] with significance threshold set to the adjusted p-value of 0·05 to determine the statistically significant features. Additionally, supervised machine learning was used together with Boruta feature selection[13] in a leave-one-out cross-validation (LOOCV) to assess the ability of breath profiles to classify the samples as asthmatic or healthy controls. Details on statistical analysis are given in the supplemental information.

Compound identification

For the 100 most significantly discriminative *m/z* features per study group, compound identification was based on MS^2 spectra that were recorded directly from exhaled breath by SESI-HRMS. The MS^2 spectra were analyzed by a workflow adapted from Kaeslin *et al.*[14] to detect isotopes, adducts, and losses, and with the SIRIUS software (v4.9.9[15]) to assign putative molecular formulae and chemical structures. The putatively identified compounds were screened for their biological context and subgrouped into metabolic pathways or chemical families. Additionally, pathway enrichment analysis using the mummichog algorithm (MetaboAnalyst, v5.0[16]) was performed for further identification on all significant features including those without recorded MS^2 spectra, without compound suggestions or with excluded MS^2 spectra (exclusion criteria: Table S1). Lastly, the detected *m/z* features were compared with previously identified compounds from literature. The certainty of identification was indicated by an identification (ID) confidence level ranging from ID 1 to 5, as described by Schymanski and colleagues.[17] More details on the identification approach are included in the supplementary material, including a schematic overview (Figure S1).

Results

Participants and clinical data

Exhaled breath samples of 48 allergic asthma patients and 56 healthy control participants, in total 104 children, were included in this study. The age and body mass index (BMI) values of the two cohorts were comparable, whereas the asthmatic group contained more male patients than the healthy one. Detailed clinical characteristics of the two individual study cohorts are shown in Table 1. All asthmatic patients stopped inhalation of long-acting asthma medication for at least one week prior to the study visit. The use of short-acting beta-agonists was allowed until the day before measurements were done. All children with asthma had a known allergic sensitization to at least one aero-allergen and the asthma severity ranged from mild to moderate. The FeNO values were significantly elevated in the allergic asthma cohort. Additionally, the forced expiratory volume in 1 second (FEV1) of the asthmatics was significantly lower, while the forced vital capacity (FVC) of the groups was comparable. 12

children from the healthy control group showed an allergic sensitization according to the skin prick test.

Table 1. Baseline characteristics.

	Allergic asthma (n = 48)	Healthy controls (n = 56)	p-value
Age [y]	12·1 ± 3·1	10·8 ± 4·0	0·07
male sex [n]	33 (68·8%)	24 (42·9%)	0·014
BMI [kg/m^2]	19·3 ± 4·2	18·3 ± 3·3	0·2
FEV1 [z-score]	-0·6 ± 1·1	-0·1 ± 1·0 *	0·01
FVC [z-score]	0·1 ± 1·0	0·1 ± 0·9 *	0·86
FeNO [ppm]	40·6 ± 35·5	11·4 ± 13·0	< 0·001
Allergic sensitization [n]	48 (100%)	12 (21·4%)	< 0.001

Data are presented as mean ± SD or n (%). BMI = body mass index, pre-bronchodilator FEV1 = forced expiratory volume in 1 second, pre-bronchodilator FVC = forced vital capacity, FeNO = fractional exhaled nitric oxide. p-values were determined by the two sample t-test and the Mann-Whitney U test for FeNO values (no normal distribution). *: 16 spirometries were excluded because of poor quality.

Discriminative breath patterns and their metabolic associations

The preprocessing of the acquired mass spectra of the study subjects revealed 2315 *m/z* features associated with exhaled breath. 375 *m/z* features were found to be significantly different between the two groups (Benjamin-Hochberg adjusted $p < 0·05$), of which 179 were upregulated and 196 downregulated in the allergic asthma group (Figure 1A-1B). Among those, 134 were successfully assigned to compounds. Inspection of the first two principal components (PCs) of the 134 putatively identified features revealed a moderate separation between the groups along the first PC (24·22% variance in the data, Figure 1C).

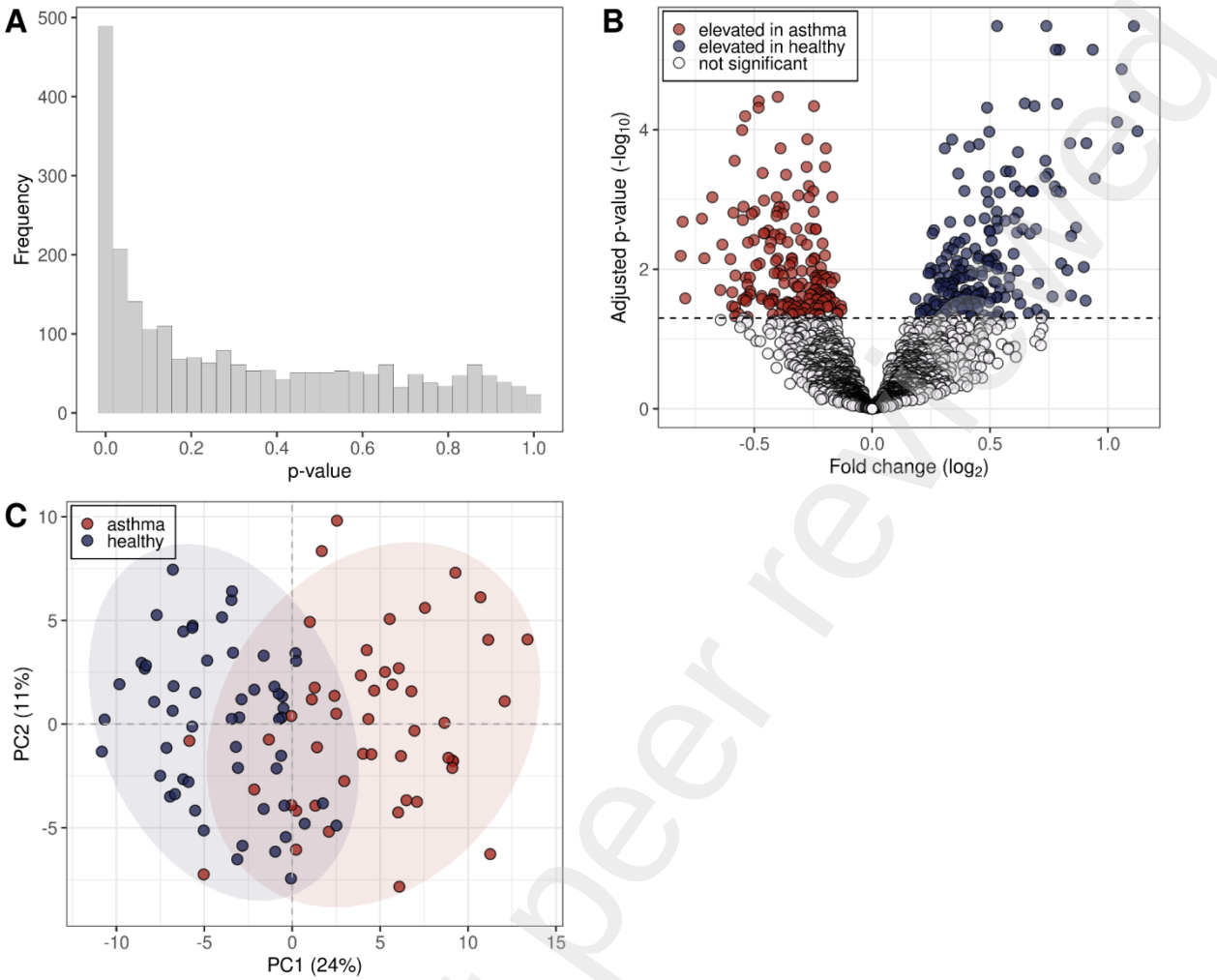


Figure 1: Statistical analysis of *m/z* features in breath profiles. A) Distribution of p-values (empirical Bayes moderated t-statistics test) of all 2315 *m/z* features. **B**) Volcano plot representing all detected 2315 *m/z* features. Dashed line: Benjamin-Hochberg adjusted p-value of 0·05. **C**) First two principal components (PCs) scores plot of the 134 putatively identified *m/z* features. 95% data ellipses were added per group for visual depiction.

Compound identification revealed several specific metabolic pathways and chemical families with many representatives for both study cohorts (Table 2-3). For the allergic asthma group, the chemical families of fatty acid metabolites and monosaccharides as well as the 2-oxocarboxylic acid metabolism and two amino acid pathways, *i.e.* lysine degradation and tyrosine metabolism, were elevated (Table 2). The relations of metabolites involved in these elevated pathways are visualized in Figure 3. For the healthy control group, arginine pathways were found to be well represented, including both arginine biosynthesis and the arginine and proline metabolism. Further, several compounds of the linoleic acid metabolism and of the chemical groups of aldehydes, amides, and fatty acids were identified (Table 3, Figure 4).

Table 2. Pathway related metabolites elevated in the allergic asthma cohort.

m/z	Charge	Adj. p-value	Molecular formula	Ionization	Δm (ppm)	Compound	Ann.	ID level
Lysine degradation								
131·035	neg	0·0002	C5H8O4	[M-H]-	0·1	Glutarate (pentanedioic acid)*	MS^2, Lit.	ID1
131·035	neg	0·0002	C5H8O4	[M-H]-	0·1	Glutarate (pentanedioic acid)*	MS^2, Lit.	ID1
117·019	neg	0·0004	C4H6O4	[M-H]-	-2·8	Succinate (butanedioic acid)*	MS^2, Lit.	ID1
129·019	neg	0·0008	C5H8O5	[M-H2O-H]-	-2·6	2-Hydroxyglutarate (2-hydroxypentanedioc acid)	MS^2	ID3
131·033	pos	0·0031	C5H6O4	[M+H]+	-6·8	Glutaconate (2-pentenedioic acid)	MS^2	ID3
161·0435	pos	0·0032	C6H8O5	[M+H]+	-5·9	2-Oxoadipate (2-oxohexanedioic acid)*	MS^2	ID3
162·0755	pos	0·0070	C6H11NO4	[M+H]+	-3·6	2-Aminoadipate (2-aminohexanedioic acid)	MS^2	ID3
97·029	neg	0·0137	C5H8O3	[M-H2O-H]-	-5·2	Glutarate semialdehyde (ω-oxopentanoaic acid)	MS^2	ID3
Tyrosine metabolism								
192·0285	neg	0·0001	C9H7NO4	[M-H]-	-9·0	5,6-Dihydroxyindole-2-carboxylate	MS^2	ID3
177·075	pos	0·0017	C7H10O4	[M+H2O+H]+	-4·2	Succinylacetone	MS^2	ID3
183·0295	neg	0·0038	C8H10O6	[M-H2O-H]-	-2·2	Succinylacetoacetate	MS^2	ID3
215·052	pos	0·0076	C9H8O5	[M+H2O+H]+	-14·0	3,4-Dihydroxyphenylpyruvate	MS^2	ID3
181·0505	neg	0·0123	C9H10O4	[M-H]-	-0·7	4-Hydroxyphenyllactate	MS^2	ID3
163·039	neg	0·0134	C9H8O3	[M-H]-	-6·6	4-Coumarate	MS^2	ID3
149·0245	neg	0·0163	C8H8O4	[M-H2O-H]-	0·6	3,4-Dihydroxymandelaldehyde	MS^2	ID3
179·036	neg	0·0214	C9H8O4	[M-H]-	5·7	4-Hydroxyphenylpyruvate, 4-Hydroxy-enol-phenylpyruvate	MS^1	ID4
			C9H10O5	[M-H2O-H]-	5·7	3-Methoxy-4-hydroxymandelate	MS^1	ID4
199·025	neg	0·0225	C8H8O6	[M-H]-	0·9	4-Maleylacetoacetate, 4-Fumarylacetoacetate	MS^1	ID4
197·046	neg	0·0268	C9H10O5	[M-H]-	2·3	3-Methoxy-4-hydroxymandelate*	MS^1	ID4
167·0345	neg	0·0271	C8H8O4	[M-H]-	-2·9	Homogentisate	MS^1	ID4
			C8H8O4	[M-H]-	-2·9	3,4-Dihydroxymandelaldehyde	MS^1	ID4
			C8H8O4	[M-H]-	-2·9	3,4-Dihydroxyphenylacetate	MS^1	ID4
2-Oxocarboxylic acid metabolism								
169·05	neg	0·0006	C8H12O5	[M-H2O-H]-	-3·7	2-Oxosuberate (2-oxooctanedionic acid)	MS^2	ID3
199·058	pos	0·0010	C9H12O6	[M-H2O+H]+	-10·5	cis-(Homo)3-aconitate	MS^2	ID3
159·0645	pos	0·0015	C7H12O5	[M-H2O+H]+	-4·3	3-Isopropylmalate	MS^2	ID3
161·0435	pos	0·0032	C6H8O5	[M+H]+	-5·9	2-Oxoadipate (2-oxohexanedioic acid)*	MS^2	ID3

m/z	Charge	Adj. p-value	Molecular formula	Ionization	Δm (ppm)	Compound	Ann.	ID level
162·0755	pos	0·0070	C6H11NO4	[M+H]+	-3·6	2-Aminoadipate (2-aminohexanedioic acid)	MS^2	ID3
148·06	pos	0·0077	C5H9NO4	[M+H]+	-2·9	Glutamate	MS^2	ID3
146·0545	neg	0·0404	C5(13C)H10O4	[M(C13)-H]-	7·3	2-Aceto-2-hydroxybutanoate	MS^1	ID4
Fatty acid metabolites								
117·019	neg	0·0004	C4H6O4	[M-H]-	-2·8	Butanedioic acid (succinate)*	MS^2, Lit.	ID1
131·035	neg	0·0002	C5H8O4	[M-H]-	0·1	Pentanedioic acid (glutarate)*	MS^2, Lit.	ID1
147·0645	pos	0·0070	C6H10O4	[M+H]+	-4·7	Hexanedioic acid (adipic acid)	MS^2	ID3
131·033	pos	0·0031	C5H6O4	[M+H]+	-6·8	Pentenedioic acid (glutaconate)	MS^2	ID3
143·0345	neg	0·0363	C6H8O4	[M-H]-	-3·7	Hexenedioic acid	Lit.	ID4
157·0505	neg	0·0230	C7H10O4	[M-H]-	-0·8	Heptenedioic acid*	Lit.	ID4
97·029	neg	0·0137	C5H8O3	[M-H2O-H]-	-5·2	ω-Oxopentanoaic acid (glutarate semialdehyde)	MS^2	ID3
125·06	neg	0·0195	C7H12O3	[M-H2O-H]-	-6·4	ω-Oxoheptanoic acid	MS^2	ID3
113·024	neg	0·0424	C5H8O4	[M-H]-	-3·7	ω-Oxopentenoic acid	Lit.	ID4
155·071	neg	0·0268	C8H12O3	[M-H]-	-2·1	ω-Oxooctenoic acid	Lit.	ID4
167·071	neg	0·0264	C9H12O3	[M-H]-	-1·9	ω-Oxononadienoic acid	Lit.	ID4
181·086	neg	4·83E-05	C10H14O3	[M-H]-	-5·6	ω-Oxodecadienoic acid	Lit.	ID4
87·0445	neg	0·0471	C4H8O2	[M-H]-	-7·5	Butanoic acid*	Lit.	ID2
101·0605	neg	0·0214	C5H10O2	[M-H]-	-0·2	Pentanoic acid*	Lit.	ID2
197·081	neg	0·0004	C10H14O4	[M-H]-	-4·6	2,7-Dimethyl-2,4-octadienedioic acid	MS^2	ID3
129·019	neg	0·0008	C5H8O5	[M-H2O-H]-	-2·6	2-Hydroxypentanedioc acid (2-hydroxyglutarate)	MS^2	ID3
178·0355	neg	0·0129	C5H9NO6	[M-H]-	-1·2	2-Amino-3,4-dihydroxypentanedioic acid	MS^2	ID3
161·0435	pos	0·0032	C6H8O5	[M+H]+	-5·9	2-Oxohexanedioic acid (2-oxoadipate)*	MS^2	ID3
162·0755	pos	0·0070	C6H11NO4	[M+H]+	-3·6	2-Aminohexanedioic acid (2-aminoadipate)	MS^2	ID3
133·05	neg	0·0104	C5H10O4	[M-H]-	-4·8	2,3-Dihydroxypentanoic acid	MS^2	ID3
Monosaccharides and metabolites								
163·024	neg	0·0002	C5H8O6	[M-H]-	-5·0	2-Dehydro-xylonate	MS^2	ID3
151·0585	pos	0·0013	C5H10O5	[M+H]+	-10·6	Arabinose	MS^2	ID3
163·0595	pos	0·0158	C6H12O6	[M-H2O+H]+	-3·7	Galactose*	MS^2	ID3
193·035	neg	0·0196	C6H10O7	[M-H]-	-2·0	Glucuronate	MS^2	ID3

m/z	Charge	Adj. p-value	Molecular formula	Ionization	Δm (ppm)	Compound	Ann.	ID level
209·03	neg	0·0244	C6H10O8	[M-H]-	-1·4	Glucarate	MS[1]	ID4
91·04	neg	0·0261	C3H8O3	[M-H]-	-0·7	Glycerol	MS[1]	ID4
119·0345	neg	0·0319	C4H8O4	[M-H]-	-4·1	Erythrulose	MS[1]	ID4

Putatively identified compounds elevated in the allergic asthma cohort, grouped by metabolic pathways or chemical families and ordered by their adjusted p-value. Exception: fatty acid metabolites are sorted based on their chemical relation. Mass error (Δm) in ppm, annotation (Ann.) e.g. based on literature (Lit.), references for literature-based identification are included in Table S2, *: compounds that were detected several times in different ionization forms (listed in Table S2), ID: identification confidence level ranging from ID1 (high) to ID5 (low).[17]

Table 3. Pathway related metabolites elevated in the healthy control cohort.

m/z	Charge	Adj. p-value	Molecular formula	Ionization	Δm (ppm)	Compound	Ann.	ID level
Arginine and proline metabolism								
104·07	pos	0·0002	C4H9NO2	[M+H]+	-5·8	4-Aminobutanoate*	MS[2], Lit.	ID3
60·0805	pos	0·0083	C4H9NO	[M-CO+H]+	-4·7	4-Aminobutanal	MS[1]	ID4
			C4H9NO2	[M-CO2+H]+	-4·7	4-Aminobutanoate	MS[1]	ID4
116·07	pos	0·0122	C5H9NO2	[M+H]+	-5·2	Proline	MS[1], Lit.	ID4
			C5H12N2O2	[M-NH3+H]+	-5·2	Ornithine	MS[1]	ID4
193·13	pos	0·0251	C6H14N4O2	[M+H2O+H]+	2·5	Arginine	MS[1]	ID4
118·086	pos	0·0252	C6H11NO3	[M-CO+H]+	-2·2	4-Acetamidobutanoate	MS[1]	ID4
114·0545	pos	0·0264	C5H7NO2	[M+H]+	-4·0	1-Pyrroline-2-carboxylate*	MS[1]	ID4
			C5H9NO3	[M-H2O+H]+	-4·0	Hydroxyproline, Glutamate 5-semialdehyde	MS[1]	ID4
112·075	pos	0·0313	C6H11NO2	[M-H2O+H]+	-6·2	N4-Acetylaminobutanal	MS[1]	ID4
102·0545	pos	0·0381	C5H9NO4	[M-HCOOH+H]+	-4·5	4-Hydroxyglutamate semialdehyde	MS[1]	ID4
			C5H7NO3	[M-CO+H]+	-4·5	1-Pyrroline-3-hydroxy-5-carboxylate*	MS[1]	ID4
Arginine biosynthesis								
61·039	pos	0·0076	CH4N2O	[M+H]+	-10·5	Urea	MS[2]	ID3
96·9925	neg	0·0113	C4H4O4	[M-H2O-H]-	-6·4	Fumarate	MS[2]	ID3
			C4H4O4	[M-H2O-H]-	-6·4	Maleate	MS[2]	ID3
116·07	pos	0·0122	C5H9NO2	[M+H]+	-5·2	Proline	MS[1], Lit.	ID4
			C5H12N2O2	[M-NH3+H]+	-5·2	Ornithine	MS[1]	ID4
193·13	pos	0·0251	C6H14N4O2	[M+H2O+H]+	2·5	Arginine	MS[1]	ID4

m/z	Charge	Adj. p-value	Molecular formula	Ionization	Δm (ppm)	Compound	Ann.	ID level
102·0545	pos	0·0381	C5H9NO4	[M-HCOOH+H]+	-4·5	4-Hydroxyglutamate semialdehyde	MS[1]	ID4
			C5H7NO3	[M-CO+H]+	-4·5	1-Pyrroline-3-hydroxy-5-carboxylate	MS[1]	ID4
Linoleic acid metabolism								
281·2475	pos	7·12E-06	C18H32O2	[M+H]+	-0·04	Linoleate*	MS[2]	ID3
295·225	pos	2·09E-04	C18H32O4	[M-H2O+H]+	-6·0	13(S)-HPODE*	MS[1]	ID4
297·242	pos	4·73E-04	C18H32O3	[M+H]+	-1·4	13(S)-HODE*, 12(13)-EpOME*, 9(10)-EpOME*	MS[1]	ID4
Aldehydes								
115·075	pos	0·0491	C6H10O2	[M+H]+	-3·1	4-Hydroxy-2-hexenal*	Lit.	ID2
146·117	pos	0·0210	C7H12O2	[M+NH4]+	-3·8	4-Hydroxy-2-heptenal	Lit.	ID4
143·106	pos	0·0179	C8H14O2	[M+H]+	-4·6	4-Hydroxy-2-octenal*	Lit.	ID4
258·243	pos	0·0225	C15H28O2	[M+NH4]+	1·0	4-Hydroxy-2-pentadecenal	Lit.	ID4
158·1175	pos	0·0082	C8H12O2	[M+NH4]+	-0·4	4-Hydroxy-2,6-octadienal	Lit.	ID4
172·133	pos	0·0004	C9H14O2	[M+NH4]+	-1·2	4-Hydroxy-2,6-nonadienal	Lit.	ID2
228·196	pos	0·0238	C13H22O2	[M+NH4]+	0·9	4-Hydroxy-2,6-tridecadienal	Lit.	ID2
283·191	neg	0·0292	C15H26O2	[M+HCOO]-	-1·7	4-Hydroxy-2,6-pentadecadienal	Lit.	ID4
253·2155	pos	0·0008	C16H28O2	[M+H]+	-2·8	4-Hydroxy-2,6-hexadecadienal	Lit.	ID4
Fatty amides								
200·201	pos	0·0008	C12H25NO	[M+H]+	0·5	Dodecanamide	MS[2]	ID3
256·263	pos	0·0008	C16H33NO	[M+H]+	-1·9	Hexadecanamide	MS[2]	ID3
302·305	pos	0·0093	C18H37NO	[M+H2O+H]+	-1·2	Octadecanamide	MS[2]	ID3
288·253	pos	0·0003	C16H33NO3	[M+H]+	-1·1	N,N-bis(2-hydroxyethyl)dodecanamide	MS[2]	ID3
316·2845	pos	0·0001	C18H35NO2	[M+H2O+H]+	-0·4	Palmitoleoylethanolamde	MS[2]	ID3
318·3	pos	0·0001	C18H37NO2	[M+H2O+H]+	-0·8	Palmitoylethanolamide	MS[2]	ID3
Fatty acids								
271·2265	pos	0·0006	C16H32O4	[M-H2O+H]+	-1·0	10,16-Dihydroxyhexadecanoic acid	MS[2]	ID3
220·1905	pos	0·0041	C11H23NO2	[M+H2O+H]+	-1·0	11-Aminoundecanoic acid	MS[2]	ID3
151·096	pos	0·0159	C6H12O3	[M+H2O+H]+	-3·2	6-Hydroxyhexanoic acid	MS[1]	ID4

Putatively identified compounds elevated in the healthy control cohort, grouped by metabolic pathways or chemical families and ordered by their adjusted p-value. Exception: aldehydes and fatty amides are sorted based on their chemical relation. Mass error (Δm) in ppm, annotation (Ann.) *e.g.* based on literature (Lit.), references for literature-based identification are included in Table S2, *: compounds that were detected several times in different ionization forms (listed in Table S2), ID: identification confidence level ranging from ID1 (high) to ID5 (low).[17] 12(13)-EpOME: 12,13-Epoxyoctadec-9(Z)-enoic acid; 9(10)-EpOME: 9,10-Epoxyoctadec-12(Z)-enoic acid; 13(S)-HPODE: 13(S)-Hydroperoxy-9Z,11E-octadecadienoic acid; 13(S)-HODE: 13(S)-Hydroxy-9Z,11E-octadeca-

dienoic acid

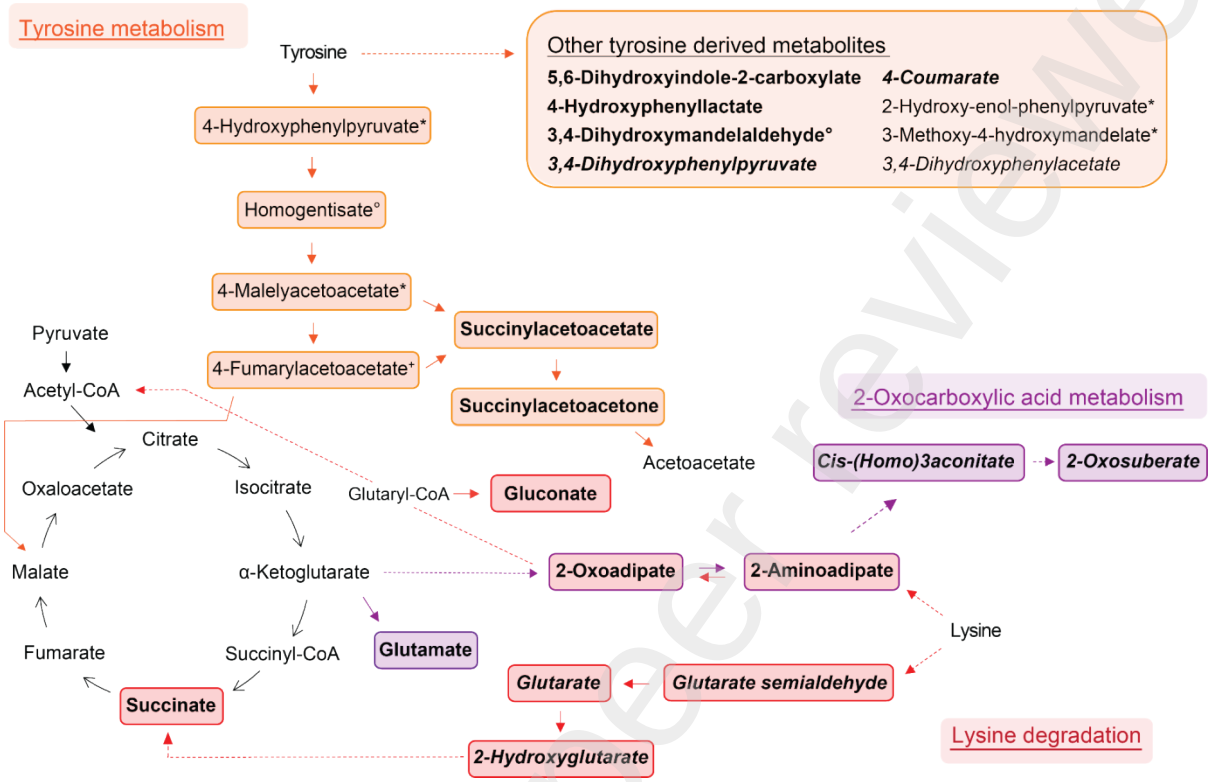


Figure 2. Schemes of metabolic pathways well-represented by compounds that were elevated in the allergic asthma group and putatively identified. Tyrosine: besides the main degradation pathway in humans other derived metabolites involved in catecholamine and melanin synthesis are summarized in the box. Two unrelated compounds of the 2-oxocarboxylic acid metabolism are not shown here (see Table 2). Arrows: direct metabolic relations, dashed lines: indirect metabolic relations (metabolites in between were not identified), coloured: putatively identified compound, bold: identified by MS^2 spectra matching, regular: identified based on exact mass and pathway mapping, or on literature, italic = metabolites from gut microbiota, *, °, $^+$: several possibilities for one *m/z* feature based on exact mass and pathway mapping.

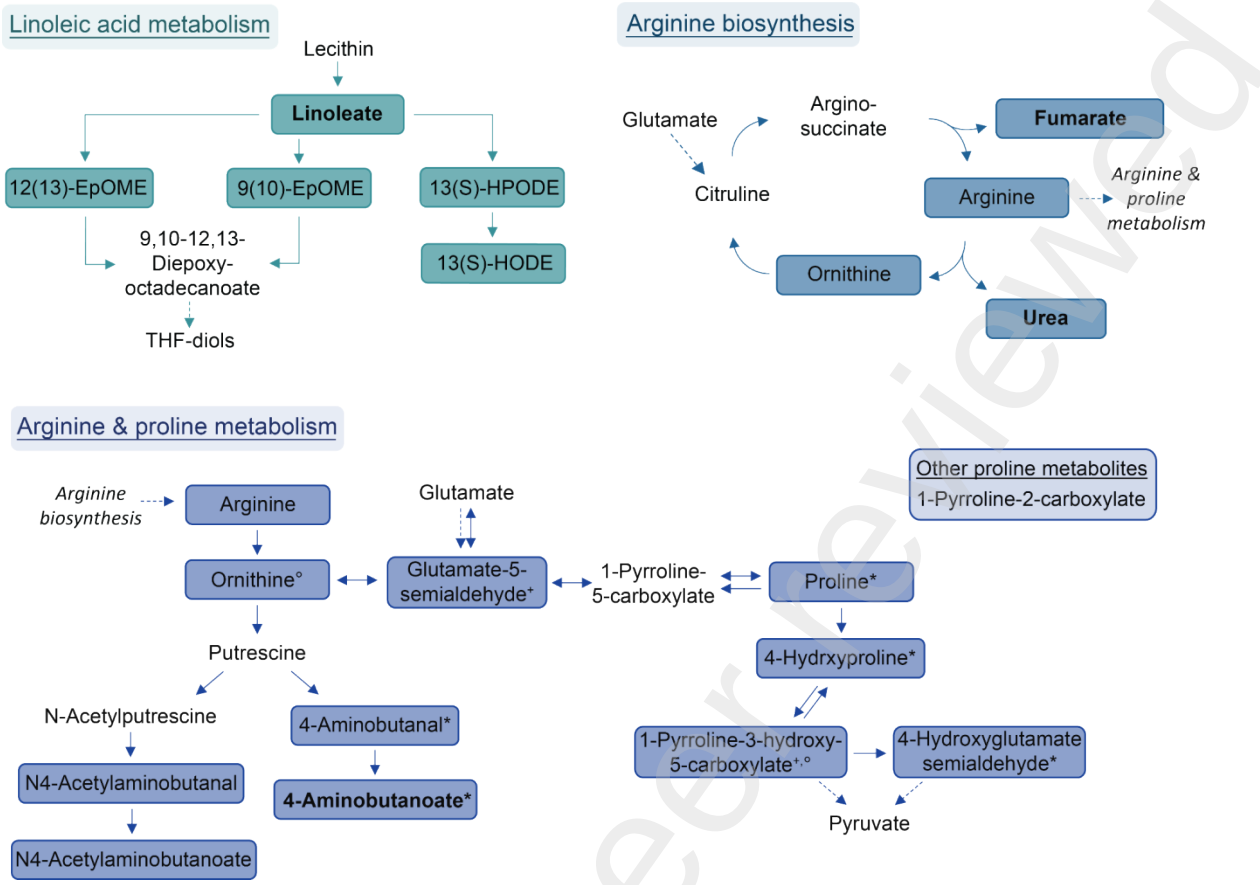


Figure 3. Schemes of metabolic pathways well-represented by compounds that were elevated in the healthy control group and putatively identified. Proline metabolism: One component is not directly connected to the displayed pathway and is shown in the box. Arrows: direct metabolic relations, dashed lines: indirect metabolic relations (metabolites in between were not identified), coloured: putatively identified compound, bold: identified by MS^2 spectra matching, regular: identified based on exact mass and pathway mapping, or on literature, *, °, $^+$, $^{+,\circ}$: several possibilities for one *m/z* feature based on exact mass and pathway mapping.

The assessment of the classification accuracy by performing LOOCV to discriminate between the asthmatic and the healthy samples resulted in an area under the curve (AUC) of 0·84, 95% CI: 0·73 - 0·94, (Figure 5A and Table S3). When examining feature selection by the Boruta scheme[13] within LOOCV, 51 (± 5·5) *m/z* features were selected on average in each cross-validation iteration, many of which were putatively identified with the compound identification workflow above (Figure 5B). Compounds which were most frequently selected in LOOCV are presented in Figure 6C (for box plots see Figure S2) and all the other selected metabolites can be found in Table S4.

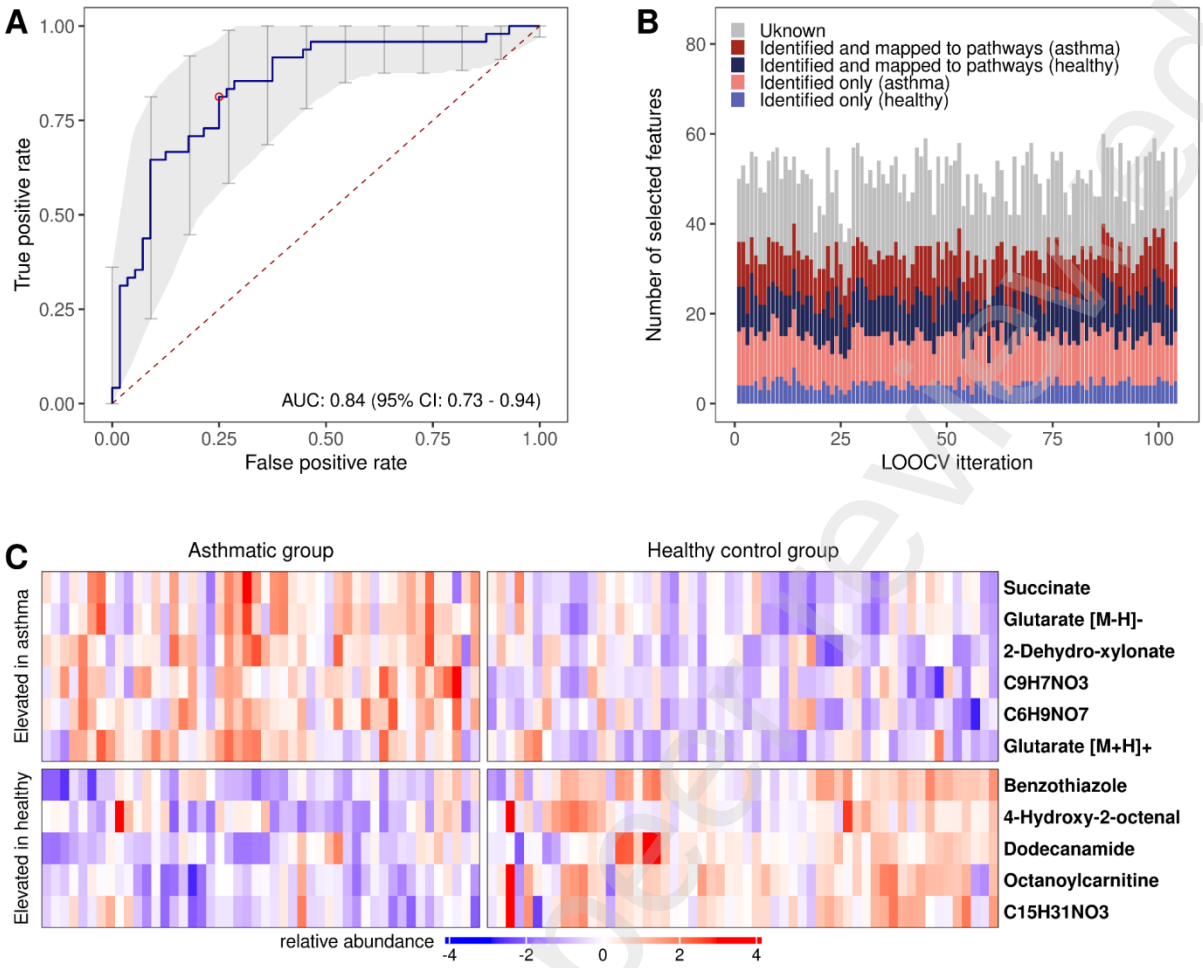


Figure 4. Disease prediction based on breath profiles A) Receiver operating characteristic curve (ROC) with an AUC of 0·85 resulting from LOOCV. Vertical gray bars: pointwise confidence intervals computed using bootstrapping (10.000 repetitions). Dashed line: line of no discrimination. **B**) Stacked bar plots of the selected features in each cross-validation iteration. Red/blue color scheme: upregulated features in the asthmatic/healthy group. **C**) Heat map of the most frequently chosen features (standardized intensities) as predictors in LOOCV (104 out of 104 times). Columns: study participants, rows: *m/z* features with putatively identified compounds (right), a chemical formula is provided if compound identification was not possible.

Discussion

We present the first online breath analysis study performed by SESI-HRMS on a pediatric population with allergic asthma. The study revealed group-specific breath patterns with a large number of discriminative *m/z* features, many of which were putatively identified and could be grouped to metabolic pathways or chemical families. Moreover, some of the relevant compounds and pathways were previously published in metabolomic studies in pediatric asthma[1,2,5] or reported in SESI-HRMS studies (see Table S2).

In general, an altered energy metabolism is expected in children with asthma due to the hypoxic environment, bronchoconstriction, and other associated changes as well as increased efforts for breathing.[5] The lung and gut microbiomes are also potential contributors to the pathophysiology of asthma.[18] Several previous breath analysis studies identified compounds and molecular pathways associated with pediatric asthma. However, some of the potential biomarkers were of exogenous origin and only a handful of them were consistently detected in more than one study.[1,2,5] The pathways and chemical families identified in our study are biologically relevant and reflect both an altered state of energy metabolism as well as changes in products from the microbiome.

The metabolism of lysine was the most significantly elevated pathway in asthma and all associated compounds were identified based on direct MS^2 spectra. Two different degradation pathways of lysine were found, one is taking place in humans and the other in the gut microbiota (Figure 3). The associated metabolites succinate and glutarate were unambiguously identified (ID1, see Table S2), and have been reported as associated with pediatric asthma in previous metabolomic studies in blood[19], urine[20], and breath[21]. Carraro *et al.* also reported a decreased level of oxoadipate in early asthma, which is not in line with our findings but could be explained by the different study design focusing on wheezing in preschool children.[21] However, a rather recent study linked an enzymatic complex involved in the lysine degradation pathway to the formation of reactive oxygen species from 2-oxoadipate[22], which could potentially be a link to asthma pathophysiology.

Tyrosine metabolism was also significantly upregulated in the allergic asthmatic group. As illustrated in Figure 3, some of the metabolites belong to the main human degradation pathway whereas other tyrosine-derived metabolites are of human or microbiotic origin. An increased level of tyrosine in asthmatic children was reported in previous metabolomics studies.[20,23,24] Additionally, the bacterial tyrosine metabolite 4-hydroxyphenylacetate was reported to be negatively correlated with the FEV1 in urine.[24] It is hypothesized that high levels of tyrosine metabolism might be related to inflammation and oxidative stress in asthma.[5] Also, tyrosine-derived catecholamines are important during conditions of stress and play a role in the regulation of the immune system.[25] In contrast to these findings, Carraro *et al.* reported a lower level of some tyrosine metabolites in children with early asthma compared to transient wheezers.[21]

The largest elevated group consisted of 20 fatty acid metabolites, including saturated and unsaturated dicarboxylic acids, *ω*-oxo-acids, hydroxy-acids, and alkanoic acids. All lysine metabolites are additionally fitting into this chemical family. A large part of these identified fatty acids were previously reported being decreased in chronic obstructive pulmonary disease exacerbations by SESI-HRMS and described as metabolites of the *ω*-oxidation, a minor pathway of the fatty acid oxidation.[6,26] Interestingly, important molecules in asthma pathophysiology including arachidonic acid, leukotrienes, and prostaglandins, although not detected in this study, are also common substrates of the cytochrome P450 *ω*-hydroxylases.[27] Therefore, our findings support the hypothesis that *ω*-oxidation might be upregulated in allergic asthma. Butanoic and pentanoic acid are both identified (ID2, Table S2) and were both reported to distinguish asthmatic from healthy children in previous studies.[3,4,28] Furthermore, the findings from previous metabolomic studies of an altered fatty acid metabolism in asthma is supported by our data.[1,2,5] Lastly, monosaccharides and derived metabolites were increased in allergic asthma. This difference in carbohydrate metabolism of asthmatic children is expected due to an altered energy demand and metabolism.[5]

In the healthy control group, the most prominent group of elevated metabolites was associated with arginine pathways, namely the arginine biosynthesis and the arginine and proline

metabolism. The enzyme arginase, in both pathways converting arginine into ornithine and urea, is an important contributor to the pathophysiology of asthma. Many studies have shown an increased arginase expression and activity in allergic asthmatics. This does not only result in decreased arginine bioavailability which in turn promotes inflammatory processes, but also in increased ornithine levels which has been associated with airway remodeling due to the increased proline production.[29]

Further, the linoleic acid metabolism was well-represented by upregulated compounds in the healthy group. While conjugated linoleic acid was consistently reported as having anti-inflammatory properties, the effect of linoleic acid especially on asthma is in dispute due to controversial observations in clinical trials.[30] Interestingly, a more recent study found genetically predicted linoleic acid to be associated with a lower risk for asthma, which is in line with our results.[31]

Within the group of amides, palmitoylethanolamide (PEA) was found to be elevated in the healthy controls. This is in line with several studies describing an anti-inflammatory effect of PEA in various organs. More recently, also an inhibitory effect for the development of allergic airway symptoms was reported for PEA in mice.[32]

The largest group of elevated compounds in the healthy group was assigned to aldehydes, which were annotated based on exact mass matches with previously identified compounds in breath (Table S2). All of these aldehydes were consistently elevated in the healthy group. A similar finding was recently reported on the oxidative stress biomarker 4-hydroxy-nonenal, which has been found to be decreased in asthmatic children by exhaled breath condensate (EBC).[33] However, a previous SESI-HRMS study reported for two of them, 4-hydroxy-heptenal and 4-hydroxy-octenal, to be increased in breath upon obstructive sleep apnea recurrence and hypothesized it being related to increased lipid peroxidation caused by oxidative stress.[7]

A recently published study comparing children with acute asthma exacerbations and healthy controls reported very similar results which confirms our findings. Despite investigating urine by high-performance liquid chromatography mass spectrometry (HPLC-MS), they also reported an elevated level of tyrosine metabolism including the metabolite gentisate and increased glucuronate as well as a downregulated linoleic acid metabolite and palmitic acid in children with acute asthma.[34]

All together, many of the enriched pathways that we reported either elevated or decreased in allergic asthma could be linked to previous findings of metabolomic studies using various methods for blood, urine or breath analysis. This strengthens the putative compound identification performed in this work and supports the possible biological and diagnostic value of these metabolites.

Assessing the predictability of the disease with supervised machine learning in LOOCV revealed an AUC of 0·85 (CI: 0·74 - 0·94), indicating that the metabolic profiles could be applied for potential diagnostic purposes. Some compounds that were allocated to subgroups of metabolic pathways or chemical families were frequently selected during cross-validation (Figure 6C and Table S4) suggesting that a smaller group of compounds might not only be pathophysiologically relevant, but also has potential for diagnostic models. The two dicarboxylic acids and lysine metabolites, succinate and glutarate, are promising candidates and were unambiguously identified. The sugar acid 2-dehydro-xylonate represents an additional potential diagnostic marker for allergic asthma. Predictors which were downregulated in asthma include the aldehyde 4-hydroxy-2-octenal, the fatty amide dodecanamide as well as benzothiazole and octanoylcarnitine.

Due to a rather large number of significant *m/z* features, a main focus was set on putative compound identification. We aimed at establishing an objective workflow that is based on matching direct MS^2 spectra with database fragment spectra, adapted from previous work[14], and was refined for a more extensive screening of suggested compounds, and expanded by additional pathway enrichment analysis to strengthen the feature annotation. A limitation in our identification approach is the lacking chromatographic separation in SESI-HRMS, which hinders the distinction of isomeric compounds. Further, with a minimum isolation window of 0·7 Da, co-fragmentation of several compounds with similar masses can occur, which complicates the annotation of fragment spectra. To address this, we excluded several MS^2 spectra with insufficient quality from further analysis, as specified in Table S1.
While asthma is a heterogeneous disease with different phenotypes, this study focused only on allergic asthma, the most frequent phenotype in children. Therefore our findings can not be extrapolated to all forms of pediatric asthma.
A clear strength of this study design is that all enrolled patients were taken off therapy at least one week prior to the study visit. While direct breath analysis by SESI bypasses any contamination during sample preparation, this adds up to also diminish confounders and signal interferences from medications in exhaled breath. However, the study was conducted over a period of 15 months on a highly sensitive instrument. Therefore, despite a strict adherence to standard operating procedures, we had to assume that unknown technical confounders might have impacted the *m/z* feature intensities. The surrogate variable analysis algorithm was successfully applied to adjust for any unmodeled factors, however, it also eliminated any heterogeneity in the data.[35] Therefore, no subgroup analysis or impact of other clinical parameters could be performed. Nevertheless, the pathway analysis indicated promising results which need to be validated on an independent cohort.

This study confirms the applicability of SESI-HRMS to a pediatric population and shows its potential to distinguish children with allergic asthma from healthy controls based on their breath signatures. Moreover, well-represented metabolic pathways that are potentially linked to the pathophysiology of allergic asthma in children could be identified. A smaller subset of the differentiating compounds could possibly be used for predictive modeling. These findings might set the path for much-needed, non-invasive clinical applications to improve early diagnosis of asthma.

Acknowledgements
We are very grateful to all the participants for participating in this study. We further like to thank the lung function team at the Children's University Hospital Zurich for their support and Jérôme Käslin for the helpful discussions on compound identification. This work is part of the Zurich exhalomics project, a flagship project of "Hochschulmedizin Zürich".

Author contributions
Study design and concept: A.M.
Data acquisition: B.S., R.W., L.W. D.I.
Data processing and analysis: S.M., B.S., R.W.
Data evaluation and interpretation: A.M., S.M., B.S., R.W.
Drafting of the manuscript: A.M., S.M., B.S., R.W.

Review and editing of the manuscript: all authors

Conflict of interest
The authors declare no conflict of interest.

Funding
Evi Diethelm-Winteler Foundation (A.M., S.M.), Childhood Research Center of the University Children's Hospital Zurich (R.W.), Heidi Ras Stiftung and Zurich Foundation (B.S.), Lotte and Adolf Hotz Stiftung (M.K.)

Data sharing statement
The study protocol is available online: https://www.kispi.uzh.ch/sites/default/files/2022-05/Study_Protocol_Asthma_Project.pdf. No further data available.

References

1 Ferraro V, Carraro S, Bozzetto S, Zanconato S, Baraldi E. Exhaled biomarkers in childhood asthma: old and new approaches. *Asthma Res. Pract*. 2018; **4**: 9.

2 Neerincx AH, Vijverberg SJH, Bos LDJ, Brinkman P, van der Schee MP, de Vries R, et al. Breathomics from exhaled volatile organic compounds in pediatric asthma. *Pediatr Pulmonol*. 2017; **52**: 1616–27.

3 Dallinga JW, Robroeks CMHHT, van Berkel JJBN, Moonen EJC, Godschalk RWL, Jöbsis Q, et al. Volatile organic compounds in exhaled breath as a diagnostic tool for asthma in children. *Clin Exp Allergy*. 2010; **40**: 68–76.

4 van Mastrigt E, Reyes-Reyes A, Brand K, Bhattacharya N, Urbach HP, Stubbs AP, et al. Exhaled breath profiling using broadband quantum cascade laser-based spectroscopy in healthy children and children with asthma and cystic fibrosis. *J Breath Res*. 2016; **10**: 026003.

5 Papamichael MM, Katsardis C, Sarandi E, Georgaki S, Frima E-S, Varvarigou A, et al. Application of Metabolomics in Pediatric Asthma: Prediction, Diagnosis and Personalized Treatment. *Metabolites*. 2021; **11**: 251.

6 Gaugg MT, Nussbaumer-Ochsner Y, Bregy L, Engler A, Stebler N, Gaisl T, et al. Real-Time Breath Analysis Reveals Specific Metabolic Signatures of COPD Exacerbations. *Chest*. 2019; **156**: 269–76.

7 Schwarz EI, Martinez-Lozano Sinues P, Bregy L, Gaisl T, Garcia Gomez D, Gaugg MT, et al. Effects of CPAP therapy withdrawal on exhaled breath pattern in obstructive sleep apnoea. *Thorax*. 2016; **71**: 110.

8 Weber R, Haas N, Baghdasaryan A, Bruderer T, Inci D, Micic S, et al. Volatile organic compound breath signatures of children with cystic fibrosis by real-time SESI-HRMS. *ERJ Open Res*. 2020; **6**: 00171–2019.

9 Gaillard EA, Kuehni CE, Turner S, Goutaki M, Holden KA, de Jong CCM, et al. European Respiratory Society clinical practice guidelines for the diagnosis of asthma in children aged 5–16 years. *Eur Respir J*. 2021; **58**: 2004173.

10 Leek JT, Storey JD. Capturing Heterogeneity in Gene Expression Studies by Surrogate Variable Analysis. *PLoS Genetics*. 2007; **3**: 1724–35.

11 Smyth GK. Linear Models and Empirical Bayes Methods for Assessing Differential Expression in Microarray Experiments. *Stat Appl Genet Mol Biol*. 2004; **3**.

12 Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc Series B Stat Methodol*. 1995; **57**: 289–300.

13 Kursa MB, Jankowski A, Rudnicki WR. Boruta – A System for Feature Selection. *Fundam Inform.* 2010; **101**: 271–85.

14 Kaeslin J, Micic S, Weber R, Müller S, Perkins N, Berger C, et al. Differentiation of Cystic Fibrosis-Related Pathogens by Volatile Organic Compound Analysis with Secondary Electrospray Ionization Mass Spectrometry. *Metabolites*. 2021; **11**: 773.

15 Dührkop K, Fleischauer M, Ludwig M, Aksenov AA, Melnik A v, Meusel M, et al. SIRIUS 4: a rapid tool for turning tandem mass spectra into metabolite structure information. *Nat Methods*. 2019; **16**: 299–302.

16 Pang Z, Chong J, Zhou G, de Lima Morais DA, Chang L, Barrette M, et al. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic Acids Res*. 2021; **49**: W388–96.

17 Schymanski EL, Jeon J, Gulde R, Fenner K, Ruff M, Singer HP, et al. Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. *Environ Sci Technol*. 2014; **48**: 2097–8.

18 Barcik W, Boutin RCT, Sokolowska M, Finlay BB. The Role of Lung and Gut Microbiota in the Pathology of Asthma. *Immunity*. 2020; **52**: 241–55.

19 Chang C, Guo Z, He B, Yao W. Metabolic alterations in the sera of Chinese patients with mild persistent asthma: a GC-MS-based metabolomics analysis. *Acta Pharmacol Sin*. 2015; **36**: 1356–66.

20 Saude EJ, Skappak CD, Regush S, Cook K, Ben-Zvi A, Becker A, et al. Metabolomic profiling of asthma: Diagnostic utility of urine nuclear magnetic resonance spectroscopy. *J Allergy Clin Immunol*. 2011; **127**: 757-764.e6.

21 Carraro S, Bozzetto S, Giordano G, el Mazloum D, Stocchero M, Pirillo P, et al. Wheezing preschool children with early-onset asthma reveal a specific metabolomic profile. *Pediatr Allergy Immunol*. 2018; **29**: 375–82.

22 Jordan F, Nemeria N, Gerfen G. Human 2-Oxoglutarate Dehydrogenase and 2-Oxoadipate Dehydrogenase Both Generate Superoxide/H2O2 in a Side Reaction and Each Could Contribute to Oxidative Stress in Mitochondria. *Neurochem Res*. 2019; **44**: 2325–35.

23 Tao JL, Chen YZ, Dai QG, Tian M, Wang S-C, Shan J-J, et al. Urine metabolic profiles in paediatric asthma. *Respirology*. 2019; **24**: 572–81.

24 Papamichael MM, Katsardis C, Erbas B, Itsiopoulos C, Tsoukalas D. Urinary organic acids as biomarkers in the assessment of pulmonary function in children with asthma. *Nutr Res*. 2019; **61**: 31–40.

25 Barnes MA, Carson MJ, Nair MG. Non-traditional cytokines: How catecholamines and adipokines influence macrophages in immunity, metabolism and the central nervous system. *Cytokine*. 2015; **72**: 210–9.

26 Gaugg MT, Bruderer T, Nowak N, Eiffert L, Martinez-Lozano Sinues P, Kohler M, et al. Mass-Spectrometric Detection of Omega-Oxidation Products of Aliphatic Fatty Acids in Exhaled Breath. *Anal Chem*. 2017; **89**: 10329–34.

27 Ni KD, Liu JY. The Functions of Cytochrome P450 ω-hydroxylases and the Associated Eicosanoids in Inflammation-Related Diseases. *Front Pharmacol*. 2021; **12**: 716801.

28 van Vliet D, Smolinska A, Jöbsis Q, Rosias PPR, Muris JWM, Dallinga JW, et al. Association between exhaled inflammatory markers and asthma control in children. *J Breath Res*. 2016; **10**: 016014.

29 Maarsingh H, Dekkers BGJ, Zuidhof AB, Bos IST, Menzen MH, Klein T, et al. Increased arginase activity contributes to airway remodelling in chronic allergic asthma. *Eur Respir J*. 2011; **38**: 318.

30 Wendell SG, Baffi C, Holguin F. Fatty acids, inflammation, and asthma. *J Allergy Clin Immunol*. 2014; **133**: 1255–64.

31 Zhao J v, Schooling CM. The role of linoleic acid in asthma and inflammatory markers: a Mendelian randomization study. *Am J Clin Nutr*; **110**: 685–90.

32 Roviezzo F, Rossi A, Caiazzo E, Orlando P, Riemma MA, Iacono VM, et al. Palmitoylethanolamide Supplementation during Sensitization Prevents Airway Allergic Symptoms in the Mouse. *Front Pharmacol*. 2017; **8**: 857.

33 Riscassi S, Corradi M, Andreoli R, Maccari C, Mercolini F, Pescollderungg L, et al. Nitric oxide products and aldehydes in exhaled breath condensate in children with asthma. *Clin Exp Allergy*. 2022; **52**: 561–4

34 Li J, Li X, Liu X, Wang X, Li J, Lin K, et al. Untargeted metabolomic study of acute exacerbation of pediatric asthma via HPLC-Q-Orbitrap-MS. *J Pharm Biomed Anal*. 2022; **215**: 114737.

35 Jaffe AE, Hyde T, Kleinman J, Weinbergern DR, Chenoweth JG, McKay RD, et al. Practical impacts of genomic data “cleaning” on biological discovery using surrogate variable analysis. *BMC Bioinform*. 2015; **16**: 372.

Asthma in one breath - Metabolic signatures for allergic asthma in children by online breath analysis: an observational study

Tables in main text:

Table 1. Baseline characteristics.

	Allergic asthma (n = 48)	Healthy controls (n = 56)	p-value
Age [y]	12·1 ± 3·1	10·8 ± 4·0	0·07
male sex [n]	33 (68·8%)	24 (42·9%)	0·014
BMI [kg/m²]	19·3 ± 4·3	18·3 ± 3·3	0·2
FEV1 [z-score]	-0·6 ± 1·1	-0·1 ± 1·0*	0·01
FVC [z-score]	0·1 ± 1·0	0·1 ± 0·9*	0·86
FeNO [ppm]	40·6 ± 35·5	11·4 ± 13·0	< 0·001
Allergic sensitization [n]	48 (100%)	12 (21·4%)	< 0·001

Data are presented as mean ± SD or n (%). BMI = body mass index, pre-bronchodilator FEV1 = forced expiratory volume in 1 second, pre-bronchodilator FVC = forced vital capacity, FeNO = fractional exhaled nitric oxide. p-values were determined by the two sample t-test and the Mann-Whitney U test for FeNO values (no normal distribution). * 16 spirometries were excluded because of poor quality.

Table 2. Pathway related metabolites elevated in the allergic asthma cohort.

m/z	Charge	Adj. p-value	Molecular formula	Ionization	Δm (ppm)	Compound	Ann.	ID level
Lysine degradation								
131·035	neg	0·0002	C5H8O4	[M-H]-	0·1	Glutarate (pentanedioic acid)*	MS^2, Lit.	ID1
131·035	neg	0·0002	C5H8O4	[M-H]-	0·1	Glutarate (pentanedioic acid)*	MS^2, Lit.	ID1
117·019	neg	0·0004	C4H6O4	[M-H]-	-2·8	Succinate (butanedioic acid)*	MS^2, Lit.	ID1
129·019	neg	0·0008	C5H8O5	[M-H2O-H]-	-2·6	2-Hydroxyglutarate (2-hydroxypentanedioc acid)	MS^2	ID3
131·033	pos	0·0031	C5H6O4	[M+H]+	-6·8	Glutaconate (2-pentenedioic acid)	MS^2	ID3
161·0435	pos	0·0032	C6H8O5	[M+H]+	-5·9	2-Oxoadipate (2-oxohexanedioic acid)*	MS^2	ID3
162·0755	pos	0·0070	C6H11NO4	[M+H]+	-3·6	2-Aminoadipate (2-aminohexanedioic acid)	MS^2	ID3
97·029	neg	0·0137	C5H8O3	[M-H2O-H]-	-5·2	Glutarate semialdehyde (ω-oxopentanoaic acid)	MS^2	ID3
Tyrosine metabolism								
192·0285	neg	0·0001	C9H7NO4	[M-H]-	-9·0	5,6-Dihydroxyindole-2-carboxylate	MS^2	ID3
177·075	pos	0·0017	C7H10O4	[M+H2O+H]+	-4·2	Succinylacetone	MS^2	ID3
183·0295	neg	0·0038	C8H10O6	[M-H2O-H]-	-2·2	Succinylacetoacetate	MS^2	ID3
215·052	pos	0·0076	C9H8O5	[M+H2O+H]+	-14·0	3,4-Dihydroxyphenylpyruvate	MS^2	ID3
181·0505	neg	0·0123	C9H10O4	[M-H]-	-0·7	4-Hydroxyphenyllactate	MS^2	ID3
163·039	neg	0·0134	C9H8O3	[M-H]-	-6·6	4-Coumarate	MS^2	ID3
149·0245	neg	0·0163	C8H8O4	[M-H2O-H]-	0·6	3,4-Dihydroxymandelaldehyde	MS^2	ID3
179·036	neg	0·0214	C9H8O4	[M-H]-	5·7	4-Hydroxyphenylpyruvate, 4-Hydroxy-enol-phenylpyruvate	MS^1	ID4
			C9H10O5	[M-H2O-H]-	5·7	3-Methoxy-4-hydroxymandelate	MS^1	ID4
199·025	neg	0·0225	C8H8O6	[M-H]-	0·9	4-Maleylacetoacetate, 4-Fumarylacetoacetate	MS^1	ID4
197·046	neg	0·0268	C9H10O5	[M-H]-	2·3	3-Methoxy-4-hydroxymandelate*	MS^1	ID4
167·0345	neg	0·0271	C8H8O4	[M-H]-	-2·9	Homogentisate	MS^1	ID4
			C8H8O4	[M-H]-	-2·9	3,4-Dihydroxymandelaldehyde	MS^1	ID4
			C8H8O4	[M-H]-	-2·9	3,4-Dihydroxyphenylacetate	MS^1	ID4
2-Oxocarboxylic acid metabolism								
169·05	neg	0·0006	C8H12O5	[M-H2O-H]-	-3·7	2-Oxosuberate (2-oxooctanedionic acid)	MS^2	ID3
199·058	pos	0·0010	C9H12O6	[M-H2O+H]+	-10·5	cis-(Homo)3-aconitate	MS^2	ID3
159·0645	pos	0·0015	C7H12O5	[M-H2O+H]+	-4·3	3-Isopropylmalate	MS^2	ID3

m/z	Charge	Adj. p-value	Molecular formula	Ionization	Δm (ppm)	Compound	Ann.	ID level
161·0435	pos	0·0032	C6H8O5	[M+H]+	-5·9	2-Oxoadipate (2-oxohexanedioic acid)*	MS^2	ID3
162·0755	pos	0·0070	C6H11NO4	[M+H]+	-3·6	2-Aminoadipate (2-aminohexanedioic acid)	MS^2	ID3
148·06	pos	0·0077	C5H9NO4	[M+H]+	-2·9	Glutamate	MS^2	ID3
146·0545	neg	0·0404	C5(13C)H10O4	[M(C13)-H]-	7·3	2-Aceto-2-hydroxybutanoate	MS^1	ID4
Fatty acid metabolites								
117·019	neg	0·0004	C4H6O4	[M-H]-	-2·8	Butanedioic acid (succinate)*	MS^2, Lit.	ID1
131·035	neg	0·0002	C5H8O4	[M-H]-	0·1	Pentanedioic acid (glutarate)*	MS^2, Lit.	ID1
147·0645	pos	0·0070	C6H10O4	[M+H]+	-4·7	Hexanedioic acid (adipic acid)	MS^2	ID3
131·033	pos	0·0031	C5H6O4	[M+H]+	-6·8	Pentenedioic acid (glutaconate)	MS^2	ID3
143·0345	neg	0·0363	C6H8O4	[M-H]-	-3·7	Hexenedioic acid	Lit.	ID4
157·0505	neg	0·0230	C7H10O4	[M-H]-	-0·8	Heptenedioic acid*	Lit.	ID4
97·029	neg	0·0137	C5H8O3	[M-H2O-H]-	-5·2	ω-Oxopentanoaic acid (glutarate semialdehyde)	MS^2	ID3
125·06	neg	0·0195	C7H12O3	[M-H2O-H]-	-6·4	ω-Oxoheptanoic acid	MS^2	ID3
113·024	neg	0·0424	C5H8O4	[M-H]-	-3·7	ω-Oxopentenoic acid	Lit.	ID4
155·071	neg	0·0268	C8H12O3	[M-H]-	-2·1	ω-Oxooctenoic acid	Lit.	ID4
167·071	neg	0·0264	C9H12O3	[M-H]-	-1·9	ω-Oxononadienoic acid	Lit.	ID4
181·086	neg	4·83E-05	C10H14O3	[M-H]-	-5·6	ω-Oxodecadienoic acid	Lit.	ID4
87·0445	neg	0·0471	C4H8O2	[M-H]-	-7·5	Butanoic acid*	Lit.	ID2
101·0605	neg	0·0214	C5H10O2	[M-H]-	-0·2	Pentanoic acid*	Lit.	ID2
197·081	neg	0·0004	C10H14O4	[M-H]-	-4·6	2,7-Dimethyl-2,4-octadienedioic acid	MS^2	ID3
129·019	neg	0·0008	C5H8O5	[M-H2O-H]-	-2·6	2-Hydroxypentanedioc acid (2-hydroxyglutarate)	MS^2	ID3
178·0355	neg	0·0129	C5H9NO6	[M-H]-	-1·2	2-Amino-3,4-dihydroxypentanedioic acid	MS^2	ID3
161·0435	pos	0·0032	C6H8O5	[M+H]+	-5·9	2-Oxohexanedioic acid (2-oxoadipate)*	MS^2	ID3
162·0755	pos	0·0070	C6H11NO4	[M+H]+	-3·6	2-Aminohexanedioic acid (2-aminoadipate)	MS^2	ID3
133·05	neg	0·0104	C5H10O4	[M-H]-	-4·8	2,3-Dihydroxypentanoic acid	MS^2	ID3
Monosaccharides and metabolites								
163·024	neg	0·0002	C5H8O6	[M-H]-	-5·0	2-Dehydro-xylonate	MS^2	ID3
151·0585	pos	0·0013	C5H10O5	[M+H]+	-10·6	Arabinose	MS^2	ID3
163·0595	pos	0·0158	C6H12O6	[M-H2O+H]+	-3·7	Galactose*	MS^2	ID3

m/z	Charge	Adj. p-value	Molecular formula	Ionization	Δm (ppm)	Compound	Ann.	ID level
193·035	neg	0·0196	C6H10O7	[M-H]-	-2·0	Glucuronate	MS^2	ID3
209·03	neg	0·0244	C6H10O8	[M-H]-	-1·4	Glucarate	MS^1	ID4
91·04	neg	0·0261	C3H8O3	[M-H]-	-0·7	Glycerol	MS^1	ID4
119·0345	neg	0·0319	C4H8O4	[M-H]-	-4·1	Erythrulose	MS^1	ID4

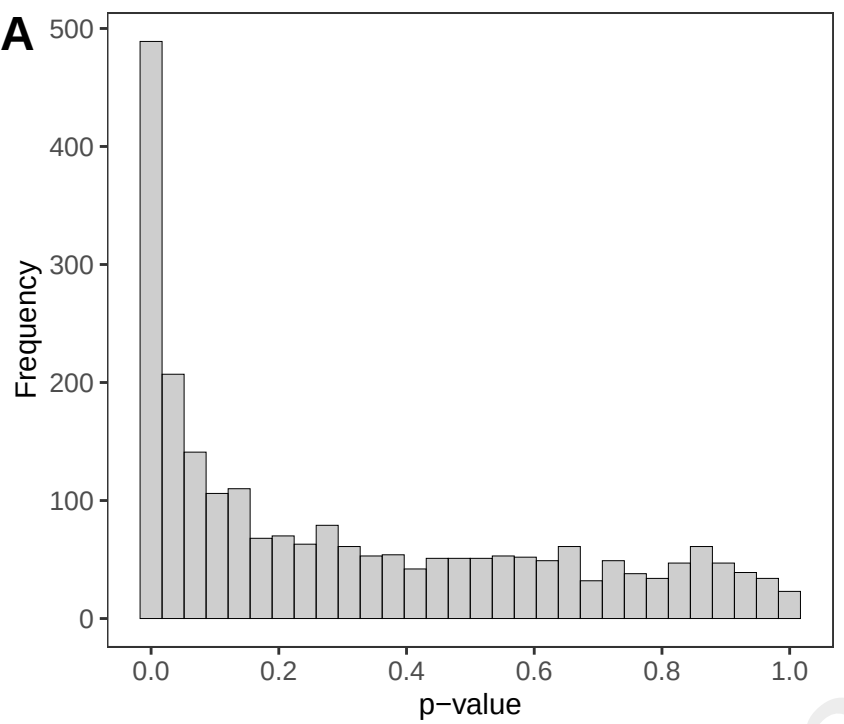
Putatively identified compounds elevated in the allergic asthma cohort, grouped by metabolic pathways or chemical families and ordered by their adjusted p-value. Exception: fatty acid metabolites are sorted based on their chemical relation. Mass error (Δm) in ppm, annotation (Ann.) e.g. based on literature (Lit.), references for literature-based identification are included in Table S2, *compounds that were detected several times in different ionization forms (listed in Table S2).

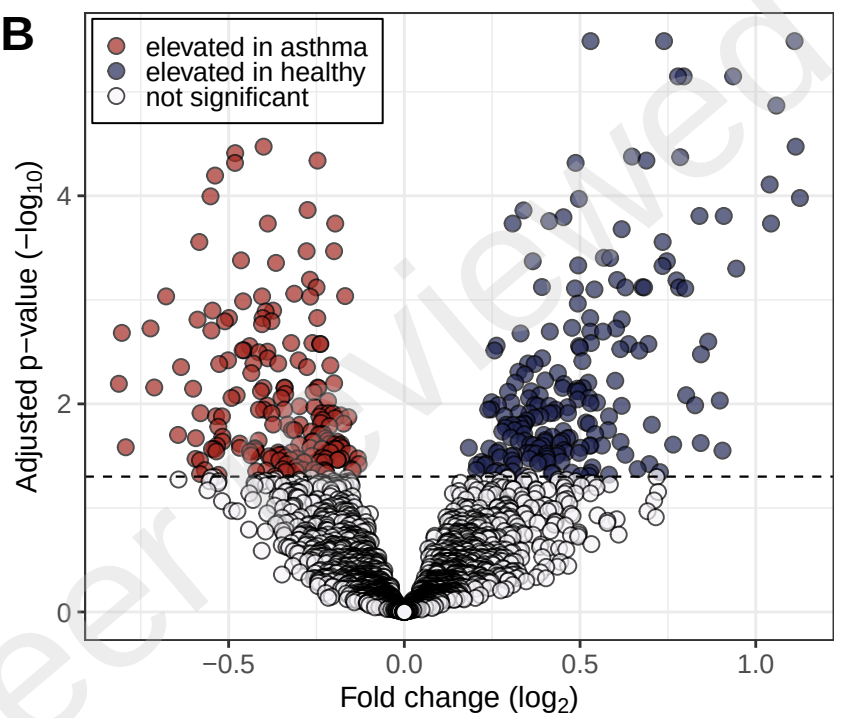
Table 3. Pathway related metabolites elevated in the healthy control cohort.

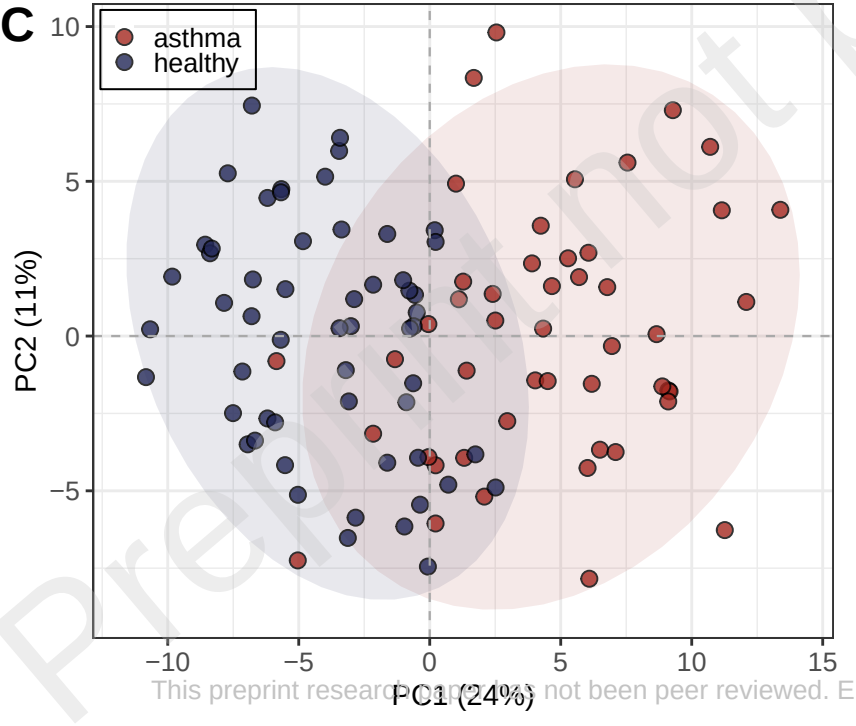
m/z	Charge	Adj. p-value	Molecular formula	Ionization	Δm (ppm)	Compound	Ann.	ID level
Arginine and proline metabolism								
104·07	pos	0·0002	C4H9NO2	[M+H]+	-5·8	4-Aminobutanoate*	MS^2, Lit.	ID3
60·0805	pos	0·0083	C4H9NO	[M-CO+H]+	-4·7	4-Aminobutanal	MS^1	ID4
			C4H9NO2	[M-CO2+H]+	-4·7	4-Aminobutanoate	MS^1	ID4
116·07	pos	0·0122	C5H9NO2	[M+H]+	-5·2	Proline	MS^1, Lit.	ID4
			C5H12N2O2	[M-NH3+H]+	-5·2	Ornithine	MS^1	ID4
193·13	pos	0·0251	C6H14N4O2	[M+H2O+H]+	2·5	Arginine	MS^1	ID4
118·086	pos	0·0252	C6H11NO3	[M-CO+H]+	-2·2	4-Acetamidobutanoate	MS^1	ID4
114·0545	pos	0·0264	C5H7NO2	[M+H]+	-4·0	1-Pyrroline-2-carboxylate*	MS^1	ID4
			C5H9NO3	[M-H2O+H]+	-4·0	Hydroxyproline, Glutamate 5-semialdehyde	MS^1	ID4
112·075	pos	0·0313	C6H11NO2	[M-H2O+H]+	-6·2	N4-Acetylaminobutanal	MS^1	ID4
102·0545	pos	0·0381	C5H9NO4	[M-HCOOH+H]+	-4·5	4-Hydroxyglutamate semialdehyde	MS^1	ID4
			C5H7NO3	[M-CO+H]+	-4·5	1-Pyrroline-3-hydroxy-5-carboxylate*	MS^1	ID4
Arginine biosynthesis								
61·039	pos	0·0076	CH4N2O	[M+H]+	-10·5	Urea	MS^2	ID3
96·9925	neg	0·0113	C4H4O4	[M-H2O-H]-	-6·4	Fumarate	MS^2	ID3
			C4H4O4	[M-H2O-H]-	-6·4	Maleate	MS^2	ID3
116·07	pos	0·0122	C5H9NO2	[M+H]+	-5·2	Proline	MS^1, Lit.	ID4
			C5H12N2O2	[M-NH3+H]+	-5·2	Ornithine	MS^1	ID4
193·13	pos	0·0251	C6H14N4O2	[M+H2O+H]+	2·5	Arginine	MS^1	ID4
102·0545	pos	0·0381	C5H9NO4	[M-HCOOH+H]+	-4·5	4-Hydroxyglutamate semialdehyde	MS^1	ID4
			C5H7NO3	[M-CO+H]+	-4·5	1-Pyrroline-3-hydroxy-5-carboxylate	MS^1	ID4
Linoleic acid metabolism								
281·2475	pos	7·12E-06	C18H32O2	[M+H]+	-0·04	Linoleate*	MS^2	ID3
295·225	pos	2·09E-04	C18H32O4	[M-H2O+H]+	-6·0	13(S)-HPODE*	MS^1	ID4
297·242	pos	4·73E-04	C18H32O3	[M+H]+	-1·4	13(S)-HODE*, 12(13)-EpOME*, 9(10)-EpOME*	MS^1	ID4
Aldehydes								
115·075	pos	0·0491	C6H10O2	[M+H]+	-3·1	4-Hydroxy-2-hexenal*	Lit.	ID2
146·117	pos	0·0210	C7H12O2	[M+NH4]+	-3·8	4-Hydroxy-2-heptenal	Lit.	ID4
143·106	pos	0·0179	C8H14O2	[M+H]+	-4·6	4-Hydroxy-2-octenal*	Lit.	ID4

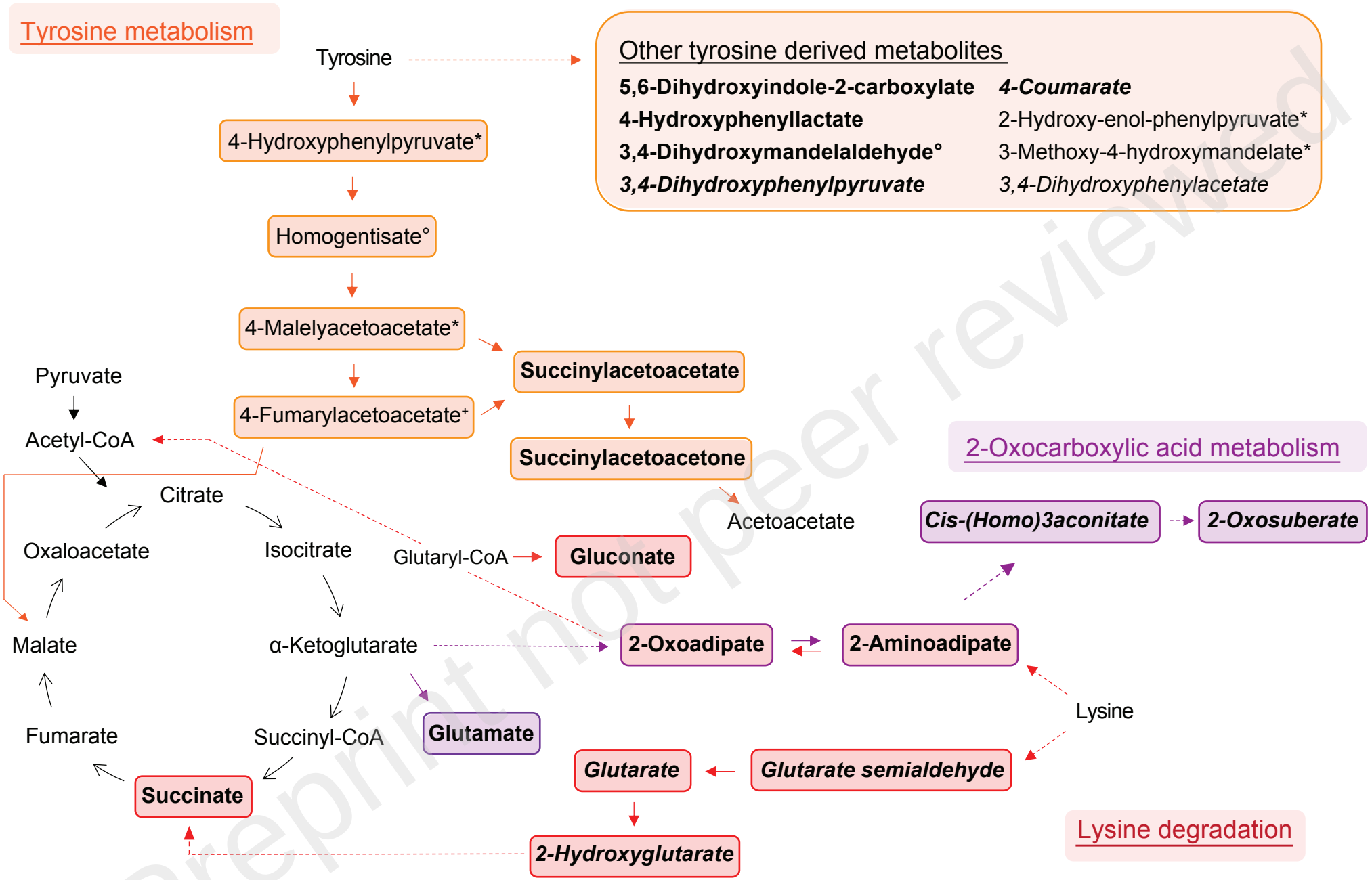
m/z	Charge	Adj. p-value	Molecular formula	Ionization	Δm (ppm)	Compound	Ann.	ID level
258·243	pos	0·0225	C15H28O2	[M+NH4]+	1·0	4-Hydroxy-2-pentadecenal	Lit.	ID4
158·1175	pos	0·0082	C8H12O2	[M+NH4]+	-0·4	4-Hydroxy-2,6-octadienal	Lit.	ID4
172·133	pos	0·0004	C9H14O2	[M+NH4]+	-1·2	4-Hydroxy-2,6-nonadienal	Lit.	ID2
228·196	pos	0·0238	C13H22O2	[M+NH4]+	0·9	4-Hydroxy-2,6-tridecadienal	Lit.	ID2
283·191	neg	0·0292	C15H26O2	[M+HCOO]-	-1·7	4-Hydroxy-2,6-pentadecadienal	Lit.	ID4
253·2155	pos	0·0008	C16H28O2	[M+H]+	-2·8	4-Hydroxy-2,6-hexadecadienal	Lit.	ID4
Fatty amides								
200·201	pos	0·0008	C12H25NO	[M+H]+	0·5	Dodecanamide	MS^2	ID3
256·263	pos	0·0008	C16H33NO	[M+H]+	-1·9	Hexadecanamide	MS^2	ID3
302·305	pos	0·0093	C18H37NO	[M+H2O+H]+	-1·2	Octadecanamide	MS^2	ID3
288·253	pos	0·0003	C16H33NO3	[M+H]+	-1·1	N,N-bis(2-hydroxyethyl)dodecanamide	MS^2	ID3
316·2845	pos	0·0001	C18H35NO2	[M+H2O+H]+	-0·4	Palmitoleoylethanolamde	MS^2	ID3
318·3	pos	0·0001	C18H37NO2	[M+H2O+H]+	-0·8	Palmitoylethanolamide	MS^2	ID3
Fatty acids								
271·2265	pos	0·0006	C16H32O4	[M-H2O+H]+	-1·0	10,16-Dihydroxyhexadecanoic acid	MS^2	ID3
220·1905	pos	0·0041	C11H23NO2	[M+H2O+H]+	-1·0	11-Aminoundecanoic acid	MS^2	ID3
151·096	pos	0·0159	C6H12O3	[M+H2O+H]+	-3·2	6-Hydroxyhexanoic acid	MS^1	ID4

Putatively identified compounds elevated in the healthy control cohort, grouped by metabolic pathways or chemical families and ordered by their adjusted p-value. Exception: aldehydes and fatty amides are sorted based on their chemical relation. Mass error (Δm) in ppm, annotation (Ann.) e.g. based on literature (Lit.), references for literature-based identification are included in Table S2, *compounds that were detected several times in different ionization forms (listed in Table S2).

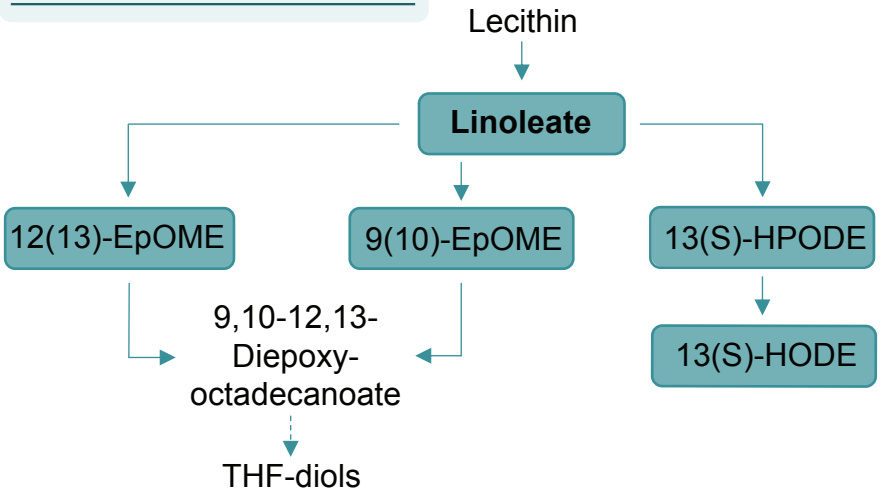




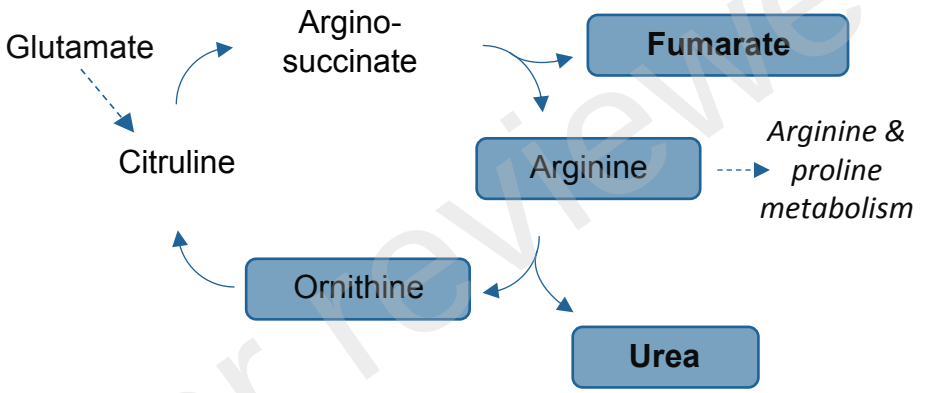




Linoleic acid metabolism



Arginine biosynthesis



Arginine & proline metabolism