Gut microbial metabolomics to understand allergies in early life

Savova, M.V.

Chapter 3

Fecal metabolome alterations in infants at risk of developing allergies during the first year of life

Based on:

Fecal metabolome alterations in infants at risk of developing allergies during the first year of life

Mariyana V. Savova[1], Pingping Zhu[1], Alida Kindt[1], the TEMPO study team, Harm Wopereis[2], Clara Belzer[3], Amy C. Harms[1], Thomas Hankemeier[1]

Metabolomics (Under revision)

Abstract

Background: Disturbances in the gut microbiome (GM) in infancy may contribute to the risk of allergy. This period is characterized by rapid microbial colonization, influenced by factors like delivery mode and infant feeding practices. The present study investigated changes in key GM taxa and fecal metabolites in relation to allergy development, delivery mode, age, and infant feeding practices during the first year of life.

Methods: Seventy-two infants at risk of allergies, exclusively breastfed for at least 16 weeks, were followed in their first year. During this period allergy manifestations were recorded and fecal samples were collected at three time points: before 16weeks, at 6 months, and at 12 months of age. The samples were subjected to metabolic profiling covering host and microbial metabolites and fluorescent in situ hybridization to quantify Bifidobacterium spp. and the *Eubacterium rectale/Clostridium coccoides* group.

Results: Strong age-associated metabolic shifts were observed, particularly in aromatic amino acid metabolites, bile acids, B vitamins, and short and long-chain fatty acids. Feeding practices, specifically the introduction of complementary feeding and the cessation of breastfeeding were significantly associated with changes to the fecal metabolome. Delivery mode had a pronounced impact on the metabolome, with differences between vaginal and Cesarean deliveries persisting until 6 months of age. Infants who developed an allergy (n=20) during this period had lower *Bifidobacterium* spp. and higher polyunsaturated fatty acid levels before the age of 16 weeks.

Conclusion: This study offers valuable insights into the longitudinal development of the fecal metabolome and factors influencing it during infancy, a critical period for immune system development.

Keywords

early life, birth mode, stool metabolomics, solid food introduction, infant, children, allergy

Abbreviations

HMOs: human milk oligosaccharides; **FISH**: Fluorescence *in situ* hybridization; **ER/CC**: *Eubacterium rectale/Clostridium coccoides*; **GC-FID**: gas chromatography coupled - flame ionization detector; **LC-MS**: liquid chromatography – mass spectrometry; **LMM**: linear mixed model; **AAs**: amino acids; **ppBMI** - pre-pregnancy BMI**, PUFA**: polyunsaturated fatty acids;

1. Introduction

Our guts are home to trillions of bacteria that live in a symbiotic relationship with us as hosts.[1] The first year of life is crucial for the development and maturation of the gut microbiome (GM).[2] This period also coincides with the development of the immune system[3] and is a key window in which GM colonization shapes the host's immune system.[4] An accumulating body of research links the disturbances of the GM composition in early life to a multitude of immune-mediated diseases,[5] including allergies.[6,7] Allergic disease often follows a temporal progression from atopic dermatitis and food allergy in infancy to allergic asthma and rhinitis in childhood, also known as "atopic march".[8] The study of early life GM composition and function in relation to allergy development is therefore a topic of considerable interest.

Many factors are known to influence the GM composition in infancy, including use of antibiotics, mode of delivery (vaginal versus C-section), milk feeding practices (breastfeeding versus formula feeding), and the transition to solid foods (complementary feeding).[9] The use of antibiotics and C-section have been associated with dysbiosis in early life and risk of developing of atopic dermatitis and other diseases later in life.[10,11] Even though the effect of breastfeeding on allergy is still a topic of debate, breastfeeding is the recommended infant nutrition for allergy prevention.[12] Breastmilk is considered the optimal nutrition for infants due to its balanced composition of macronutrients and bioactive compounds satisfying the infant's nutritional and physiological requirements.[13] It is also an important source of bifidobacteria and lactobacilli as well as human milk oligosaccharides (HMOs).[14] Bifidobacteria, e.g. *B. breve*, *B. bifidum*, *B. longum*, capable of utilizing HMOs and derivatives for energy, thrive in the guts of healthy breastfed infants and are crucial for immune system development.[15]

While the impact of the above-mentioned factors on the GM composition is relatively well-studied, their influence on GM activity remains understudied. Similarly, research examining the link between allergies and the GM have mainly focused on compositional analysis.[7] Since the GM influences the host's physiology via the production of metabolites, researchers are increasingly examining the metabolome to get insights into host-microbiota interactions.[16]

In this study, healthy breastfed infants at increased risk of developing allergies were followed during their first year. Data on delivery mode, allergy development, feeding practices were collected, and key gut microbial taxa along with the fecal metabolome were analyzed at three time points. This allowed us to assess microbiome and metabolomic changes associated with allergy, delivery mode, age, milk feeding practices (breastfeeding and formula feeding), and complementary feeding.

2. Experimental Section

2.1 Study design, sample collection and storage

The samples for this work arise from a randomized, double-blind, controlled, parallel-group, multi-country study called TEMPO (clinicaltrials.gov identifier: NCT03067714). Detailed information on ethics committees, institutional review boards, and regulatory authorities that approved the study was previously published.[17] TEMPO enrolled healthy term infants (age: <16 weeks) at increased risk of developing allergy based on family history. Subjects who began formula feeding before 16 weeks entered one of the two intervention arms, while those exclusively breastfed for at least 16 weeks comprised the breastfed reference group. Exclusive breastfeeding was defined as receiving only breastmilk, with no other liquids or solids except water or formula in the first 72 hours of life, disregarding vitamins, minerals, or medicines. All participants were followed for a year, during which events of allergic manifestations were diagnosed by qualified physicians and classified as skin, food, or respiratory allergies. Allergy manifestations were considered IgE-mediated if either the skin prick test to any tested allergen or specific IgE blood test was positive at 12m. In this study, we selected a subset of 72 subjects solely from the breastfed reference group based on the availability of fecal samples collected before 16 weeks (baseline), at 6 months (6m), and at 12 months (12m) of age. Sample collection and storage procedure is available in Supplementary materials.

2.2 Microbiome data acquisition

Fluorescence *in situ* hybridization (FISH) quantification of *Bifidobacterium* genus and *Eubacterium rectale/Clostridium coccoides* group (ER/CC) was performed on a subset of subjects as described previously.[18]

2.3 Metabolomic data acquisition

Liquid chromatography–mass spectrometry (LC-MS) metabolomic data acquisition and preprocessing were performed as previously described.[19] Briefly, wet fecal samples went through lyophilization and liquid-liquid extraction prior to the analysis by reverse phase LC-MS (RPLC-MS) using two separate assays one covering polar to semi-polar metabolites and a second covering bile acids (BAs) and long-chain fatty acids (LCFAs). In case of coelution, the targets were reported using the name or abbreviation of one of the targets followed by a "#" (**Table S1**). Data quality inspection, including between-batch correction and removal of metabolites with high technical variance (quality control RSD > 30%) was conducted using mzQuality.[20] The analysis of short-chain fatty acids (SCFAs) and lactic acid was conducted as already described.[21]

2.4 Data analysis

Data handling and statistical analyses were performed in R (version 4.3.3). After dry weight normalization, metabolites with a median signal below five times the mean signal of the procedure blanks were excluded. To detect group bias in missing data, the Fisher's exact test was applied to metabolites with any missing measurements. The 57 metabolites with missingness >20% were subjected to unpaired Mann-Whitney U test to assess the difference between visits and between the study groups (allergic vs non-allergic, vaginal vs C-section delivery, complementary-fed vs non-complementary-fed, formula-fed vs non-formula-fed, breastfed vs non-breastfed) at the relevant visits. Missing values of the 162 metabolites with missingness <20% were imputed after log2 transformation. Then, linear mixed models (LMMs) were used to examine the metabolomic difference between the study groups over time. Clinical characteristics were checked for associations to allergy, delivery mode, and feeding practices using Mann-Whitney U-test for numeric variables and the Fisher's exact test for binary variables. Differences in the microbiome data across visits and between the study groups at each visit were assessed using the Mann-Whitney U test. Spearman's correlation analysis was conducted to assess the relationship between LCFAs and microbiome taxa. Multiple testing correction was performed using the Benjamini-Hochberg method where Q<0.1 was considered as statistically significant. Further data analysis details are available in Supplementary Materials.

3. Results

3.1 Patient characteristics

Table 1 summarizes the characteristics of the 72 infants at risk of developing allergy who were followed throughout their first year. The associations between the clinical characteristics and allergy manifestation, delivery mode, and feeding practices, were examined (**Table S2-3**). Potential confounders excluded from this analysis include: i) clinical characteristics describing symptoms of allergy and its treatment, as well as gestational age and maternal pre-pregnancy body mass index (ppBMI) associated with C-section; ii) patient characteristics such as country and mineral supplementation which were excluded due to low sample size.

3.2 Age has a significant impact on the fecal metabolome

To explore the impact of age, diet, delivery mode, and allergy on the fecal metabolome, a range of host and gut microbial metabolites, including (aromatic) amino acids (AAs) and derivatives, vitamins, nucleobases, nucleosides, BAs, LCFAs, and SCFAs were examined (**Table S1**).

Table 1. Clinical characteristics. Numeric variables are presented as median [range]; categorical variables are presented as numbers of participants.

Variable	Whole cohort	Allergic	Non-Allergic	C-section	Vaginal
Sex (female/male)	35 / 37	9 / 11	26 / 26	12 / 8	23 / 29
Allergy manifestation (Allergic/Not allergic)	20 / 52†	-	-	7 / 13	13 / 39
Type of allergy (IgE/non-IgE)	10/10	10/10	-	3 / 4	7 / 6
Type of allergy† (skin/food/respiratory)	18 / 2 / 2	18 / 2 / 2	-	7 / 1 / 0	11 / 2 / 1
Age onset allergy (days)	-	126.5 [69, 299]	-	-	-
Mode of delivery (Vaginal/C-section)	52 / 20 ‡	13 / 7	39 / 13	-	-
Country					
Belgium	1	1	0	0	1
Czech Republic	36	15	21	8	28
United Kingdom	1	0	1	0	1
Hungary	14	2	12	5	9
Slovakia	20	2	18	7	13
Gestational age (weeks)	39.3 [37.6-41.9]	39.2 [37.6-41.7]	39.3 [37.6-41.9]	38.8 [37.6-40.9]	39.6 [37.6-41.9]
Birth head circumference (cm)	34.5 [32-39]	34.0 [33-38]	35.0 [32-39]	35.0 [33-39]	34.0 [32-38]
Birth weight (kg)	3.4 [2.6-4.2]	3.4 [2.9-4.2]	3.4 [2.6-4]	3.4 [2.9-4]	3.4 [2.6-4.2]
Birth length (cm)	50 ± 2.4	50 ± 1.8	50 ± 2.6	50 ± 1.8	50 ± 2.6
Birth length (cm)	50 [47-58]	50 [47-54]	50 [47-58]	50 [47-53]	50 [47-58]
Mother's ppBMI	23.3 [18.4-40]	22.2 [18.4-35]	23.9 [18.6-40]	24.6 [19.9-40]	22.8 [18.4-35]
Age (days)					
baseline	41.5 [1-111]	23 [1-111]	52.5 [2-111]	53 [2-108]	37 [1-111]
6 months	180 [166-227]	179 [166-192]	180 [167-227]	180 [168-227]	180 [166-206]
12 months	364 [345-383]	365.5 [348-383]	363.5 [345-378]	365 [345-377]	363.5 [348-383]
Breastfeeding (yes/no)					
baseline	72 / 0	20 / 0	52 / 0	20 / 0	52 / 0
6 months	71 / 1	20 / 0	51 / 1	20 / 0	51 / 1
12 months	58 / 14	17 / 3	41 / 11	18 / 2	40 / 12
Formula Feeding (yes/no)					
baseline	0 / 72	0 / 20	0 / 52	0 / 20	0 / 52
6 months	7 / 65	2 / 18	5 / 47	1 / 19	6 / 46
12 months	23 / 49	6 / 14	17 / 35	4 / 16	19 / 33
Milk feeding§ (BF / FF / MMF)					
baseline	72 / 0 / 0	20 / 0 / 0	52 / 0 / 0	20 / 0 / 0	52 / 0 / 0
6 months	65 / 1 / 6	18 / 0 / 2	47 / 1 / 4	19 / 0 / 1	46 / 1 / 5
12 months	49 / 12 / 11	12 / 2 / 4	35 / 10 / 7	48 / 2 / 2	33 / 10 / 9
Complementary Feeding (yes/no)					
baseline	0 / 72	0 / 20	0 / 52	0 / 20	0 / 52
6 months	55 / 17	13 / 7	42 / 10	15 / 5	40 / 12
12 months	72 / 0	20 / 0	52 / 0	20 / 0	52 / 0

†The two subjects who had developed IgE-mediated food/respiratory allergy were also diagnosed with IgE-mediated skin allergy
‡Even though the numbers for allergy and delivery mode are the same (20 / 52), the infants in the four groups are different. More specifically 39 non-allergic and 13 allergic subjects were delivered vaginally; while 13 were non-allergic and delivered via a C-section; 7 were allergic and delivered via a C-section.
§BF – breastfed, infants receiving breastmilk and no formula milk; FF – formula-fed, infants receiving infant formula milk and not breastmilk, MMF – mixed milk-fed – infants receiving breastmilk and formula milk

Age had a strong effect on the metabolome, with LMM analysis identifying 99 metabolites that significantly changed within the first 6 months of life and 92 metabolites in the second half of the first year (**Figure 1A**). B vitamins and derivatives, AAs and derivatives, BAs, nucleobases, nucleosides and derivatives, SCFAs, and phenolic acids increased significantly throughout the whole first year, between baseline and 6m or between 6m and 12m. Among those the primary BAs, CA and CDCA, increased in the first six months, glyco-conjugated BAs in the latter six months, and secondary BAs during either or both halves of the year (**Figure 1A**).

Host-derived tryptophan metabolites also increased with age, whereas the microbial aromatic AA metabolites followed varying time trends. Aromatic lactic acids (PLA, ILA, 4-OH-PLA#) increased until 6m. Then, while PLA remained unchanged, ILA and 4-OH-PLA# decreased. The tryptophan-derived indoxyl sulfuric acid and phenylalanine-derived PAGIn and hippuric acid also declined after 6m. Meanwhile, the acetic aromatic acids 4-OH-PAA# and IAA increased after 6m, with IAA decreasing before 6m (**Figure 1A**).

LCFAs declined through the first year, except for ALA#, LA, and mead acid which declined significantly only until 6m. Even though an overall decline in acylcarnitines was observed after 6m, before 6m the long-chain acylcarnitines increased, whereas the short- and medium-chain acylcarnitines remained unchanged (**Figure 1A**).

The metabolites that could not be analyzed using LMMs were assessed using a Mann-Whitney U test (**Figure S1A**). Consistent with LMM findings, acylcarnitines decreased, whereas AAs and derivatives; B vitamins and derivatives; nucleobases, nucleosides and derivatives; SCFAs; phenolic acids; and host tryptophan metabolites increased over time. The acetic and propionic aromatic acids, PAA and 4-OH-PPA, also increased whereas IPA levels remained stable despite the decline in its missingness with age (missingness of 91.7%, 79.2%, 34.7% for baseline, 6m, 12m respectively). TUDCA and secondary BAs also increased, particularly in the second half of the first year (**Figure S1A**). Although LCA and DCA did not pass QC, visual inspection suggested a rise, especially in some subjects at 12m (**Figure S3**).

3.3 Dietary changes were associated with fecal metabolome alterations

Infant diets evolved during the first year (**Table 1**), where at baseline (<16 weeks), all infants were breastfed, at 6m 90% infants were breastfed, 8.3% mixed-milk-fed (breastfed and formula-fed), and 1.4% formula-fed, while at 12m, 68% were breastfed, 15% mixed-milk-fed, and 17% formula-fed. Meanwhile, complementary feeding had started for 76% of the participants by 6m and for all by 12m.

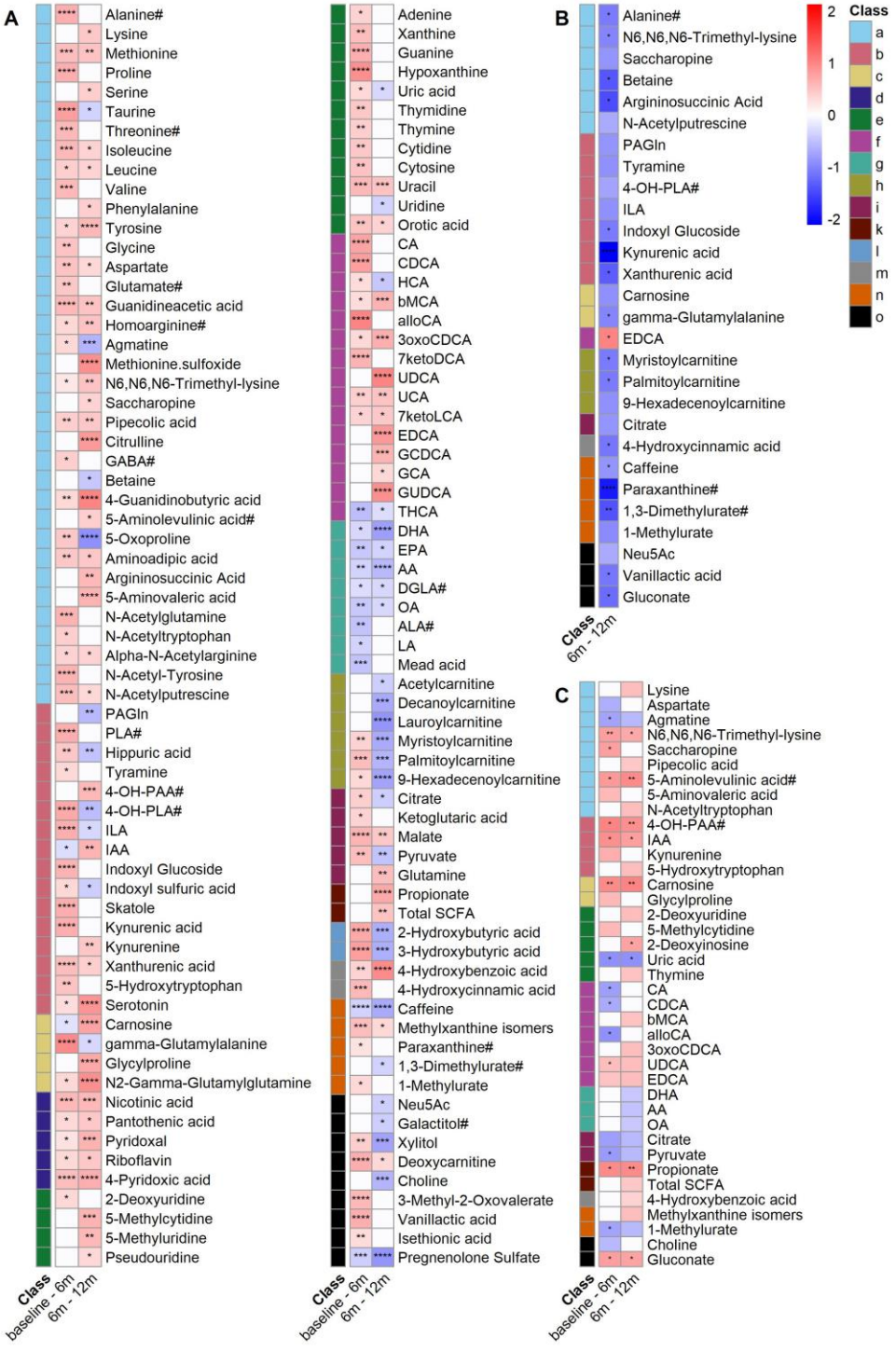
A
Alanine#
Lysine
Methionine
Proline
Serine
Taurine
Threonine#
Isoleucine
Leucine
Valine
Phenylalanine
Tyrosine
Glycine
Aspartate
Glutamate#
Guanidineacetic acid
Homoarginine#
Agmatine
Methionine.sulfoxide
N6,N6,N6-Trimethyl-lysine
Saccharopine
Pipecolic acid
Citrulline
GABA#
Betaine
4-Guanidinobutyric acid
5-Aminolevulinic acid#
5-Oxoproline
Aminoadipic acid
Argininosuccinic Acid
5-Aminovaleric acid
N-Acetylglutamine
N-Acetyltryptophan
Alpha-N-Acetylarginine
N-Acetyl-Tyrosine
N-Acetylputrescine
PAGln
PLA#
Hippuric acid
Tyramine
4-OH-PAA#
4-OH-PLA#
ILA
IAA
Indoxyl Glucoside
Indoxyl sulfuric acid
Skatole
Kynurenic acid
Kynurenine
Xanthurenic acid
5-Hydroxytryptophan
Serotonin
Carnosine
gamma-Glutamylalanine
Glycylproline
N2-Gamma-Glutamylglutamine
Nicotinic acid
Pantothenic acid
Pyridoxal
Riboflavin
4-Pyridoxic acid
2-Deoxyuridine
5-Methylcytidine
5-Methyluridine
Pseudouridine
Class
baseline - 6m
6m - 12m
Adenine
Xanthine
Guanine
Hypoxanthine
Uric acid
Thymidine
Thymine
Cytidine
Cytosine
Uracil
Uridine
Orotic acid
CA
CDCA
HCA
bMCA
alloCA
3oxoCDCA
7ketoDCA
UDCA
UCA
7ketoLCA
EDCA
GCDCA
GCA
GUDCA
THCA
DHA
EPA
AA
DGLA#
OA
ALA#
LA
Mead acid
Acetylcarnitine
Decanoylcarnitine
Lauroylcarnitine
Myristoylcarnitine
Palmitoylcarnitine
9-Hexadecenoylcarnitine
Citrate
Ketoglutaric acid
Malate
Pyruvate
Glutamine
Propionate
Total SCFA
2-Hydroxybutyric acid
3-Hydroxybutyric acid
4-Hydroxybenzoic acid
4-Hydroxycinnamic acid
Caffeine
Methylxanthine isomers
Paraxanthine#
1,3-Dimethylurate#
1-Methylurate
Neu5Ac
Galactitol#
Xylitol
Deoxycarnitine
Choline
3-Methyl-2-Oxovalerate
Vanillactic acid
Isethionic acid
Pregnenolone Sulfate
Class
baseline - 6m
6m - 12m
B
Alanine#
N6,N6,N6-Trimethyl-lysine
Saccharopine
Betaine
Argininosuccinic Acid
N-Acetylputrescine
PAGln
Tyramine
4-OH-PLA#
ILA
Indoxyl Glucoside
Kynurenic acid
Xanthurenic acid
Carnosine
gamma-Glutamylalanine
EDCA
Myristoylcarnitine
Palmitoylcarnitine
9-Hexadecenoylcarnitine
Citrate
4-Hydroxycinnamic acid
Caffeine
Paraxanthine#
1,3-Dimethylurate#
1-Methylurate
Neu5Ac
Vanillactic acid
Gluconate
Class
6m - 12m
2
1
0
-1
-2
Class
a
b
c
d
e
f
g
h
i
k
l
m
n
o
C
Lysine
Aspartate
Agmatine
N6,N6,N6-Trimethyl-lysine
Saccharopine
Pipecolic acid
5-Aminolevulinic acid#
5-Aminovaleric acid
N-Acetyltryptophan
4-OH-PAA#
IAA
Kynurenine
5-Hydroxytryptophan
Carnosine
Glycylproline
2-Deoxyuridine
5-Methylcytidine
2-Deoxyinosine
Uric acid
Thymine
CA
CDCA
bMCA
alloCA
3oxoCDCA
UDCA
EDCA
DHA
AA
OA
Citrate
Pyruvate
Propionate
Total SCFA
4-Hydroxybenzoic acid
Methylxanthine isomers
1-Methylurate
Choline
Gluconate
Class
baseline - 6m
6m - 12m

Figure 1. *Fecal metabolome alterations associated with age (**A**), cessation of breastfeeding (**B**), introduction of complementary feeding (**C**) between baseline and 6m and/or 6m and 12m, assessed using LMM. Colors represent the model coefficient: positive (red), negative (blue), P>0.05 (white). In (**A**) a positive coefficient represents an increase of the metabolite between the visits; in (**B**) a positive coefficient represents an increase of the metabolite associated with cessation of breastfeeding between 6m and 12m; while in (**C**) a positive coefficient represents an increase of the metabolite with introduction of complementary feeding between baseline and 6m or between 6m and 12m. Class annotation: **a** - AAs and derivatives; **b** aromatic AAs metabolites; **c** - dipeptides and tripeptides; **d** - B vitamins and derivatives; **e** - nucleobases, nucleosides and derivatives; **f** – BAs; g – LCFAs; **h** – carnitines; **j** – energy metabolites; **k** – SCFAs; **l** - hydroxy acids and derivatives; **m** - phenolic acids; **n** – xanthines; **o** – other. Asterisks indicate statistical significance: Q < 0.1 (*), Q < 0.01 (**), Q < 0.001 (***), Q < 0.0001 (****). The "#" in the metabolite names indicates that the metabolite coeluted with another target metabolite. All abbreviations and coeluting metabolites can be found in Table S1.*

Initiation of formula-feeding had a minor effect on the metabolome (LMM, **Table S4**). It was associated with lower levels of B vitamins i.e. pyridoxal, pantothenic acid, nicotinic acid as well as thymine, 2-deoxyuridine, 2-deoxyinosine but with higher guanosine# and allantoin until 6m. However, following multiple testing correction, only the association of thymine remained significant.

Complementary feeding was associated with significantly higher propionate, carnosine and aromatic acetic acids 4-OH-PAA# and IAA but lower uric acid and pyruvate levels among others (**Figure 1C**). Until 6m, complementary feeding was also negatively associated with the primary BAs CA, CDCA, alloCA but positively with the secondary BA UDCA and syringic acid. The latter was detected only after the introduction of complementary food except for one infant (**Figure S2**).

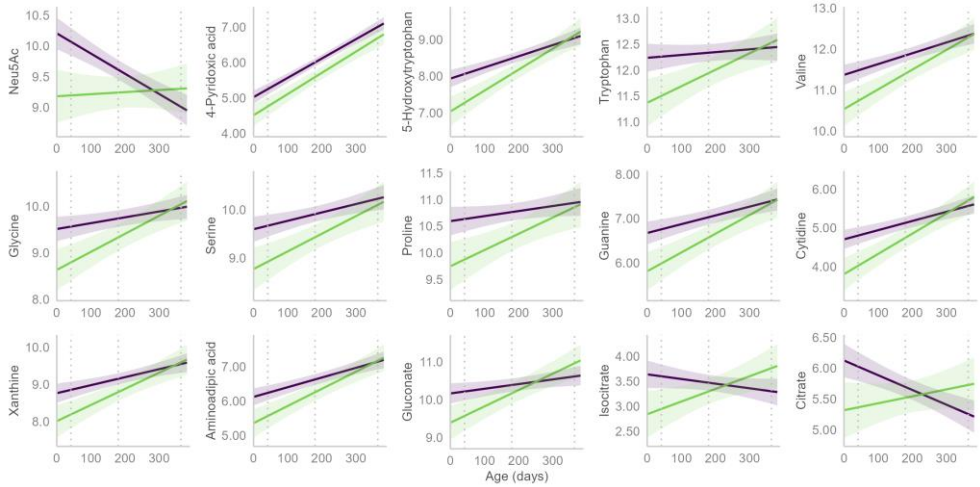

Figure 2. *Scaled abundance levels of the metabolites that significantly differed between infants delivered vaginally (purple, solid) and via C-section (green, solid) group as a function of age, based on LMM analysis. The shaded areas represent the 95% confidence intervals, while the dotted grey lines represent the median age at each visit (baseline, 6m, 12m). All abbreviations and coeluting metabolites can be found in Table S1.*

Meanwhile, the cessation of breastfeeding was associated with higher EDCA but lower long-chain acylcarnitines, caffeine and metabolites, Neu5Ac, gamma-glutamylalanine, and 4-hydroxycinnamic acid. The tryptophan and tyrosine metabolites kynurenic acid, indoxyl glucoside, xanthurenic acid (Q<0.1), ILA, tyramine, and 4-OH-PLA# (0.01<P<0.05, Q>0.1) were also negatively associated with cessation of breastfeeding (**Figure 1B**).

The effect of feeding practices for metabolites that could not be analyzed using LMM analysis, were assessed using the Mann-Whitney U test (**Figure S1B-C**). Butyrate, secondary BAs, and phenolic acids were higher in the complementary-fed versus non-complementary-fed infants at 6m (Q<0.1) and breastfed versus non-breastfed subjects at 12m (P<0.05). N2,N2-dimethylguanosine and 2-octenoylcarnitine were respectively higher and lower in the complementary-fed versus non-complementary-fed infants, whereas the tryptophan metabolite IPA was higher in the non-breastfed versus breastfed infants.

3.4 Delivery mode affected the fecal metabolome up to 6 months of age

TEMPO enrolled infants delivered vaginally and via a C-section, allowing an investigation into the effect of the delivery mode on the metabolome. Fifteen metabolites, including Neu5Ac, AAs and derivatives, pyrimidine and purine derivatives, and carboxylic acids, were significantly lower in the C-section compared to the vaginal group at baseline (**Figure 2**). For all, the group differences decreased with age until the groups completely overlapped at 12m. Notably, Neu5Ac levels remained stable over time in the C-section group, while they declined in the vaginal group. In contrast, proline and tryptophan were stable in the vaginal group but increased over time in the C-section group. Citrate and isocitrate also followed opposing trends, decreasing in the vaginal group while increasing in the C-section group.

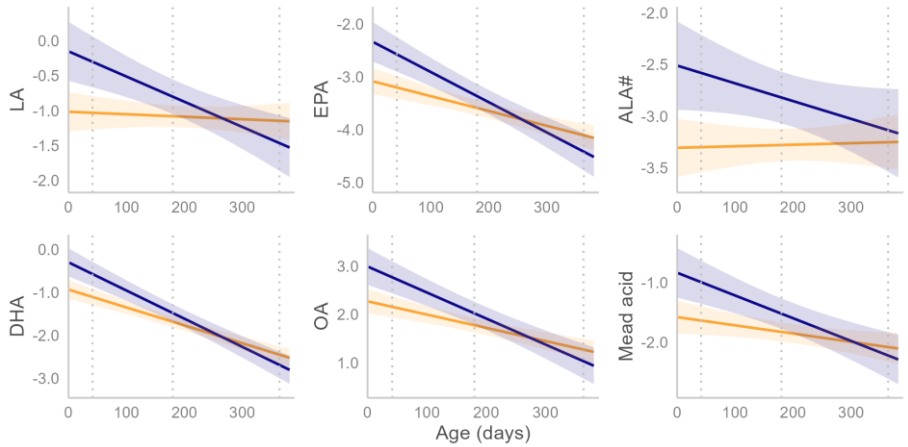

Figure 3. *Scaled relative abundance levels of LCFAs as a function of age in allergic (blue, solid) and non-allergic (orange, solid) groups, based on LMM analysis. The shaded areas represent the 95% confidence intervals, while the dotted grey lines represent the median age at each visit (baseline, 6m, 12m). The "#" in the metabolite names indicates that the metabolite coeluted with another target metabolite. All abbreviations and coeluting metabolites can be found in Table S1.*

3.5 Higher LCFA levels in infants who developed allergy

LMMs were used to examine the longitudinal metabolite alterations with age between the infants who developed allergies during the first year of life and those who did not. At baseline, no participants were allergic and allergies developed between 69 and 299 days (median age 126.5 days). A few LCFAs, namely LA, EPA, ALA#, DHA, OA, and mead acid, were found to be significantly higher at baseline in the allergic compared to the non-allergic group. However, the group separation disappeared over time and the groups overlapped at 6m and 12m (**Figure 3**).

3.6 Lower Bifidobacterium spp. in infants prior to allergy development

FISH was applied to quantify the *Bifidobacterium* spp. which are characteristic GM members in breastfed infants and ER/CC, which is primarily composed of *Lachnospiraceae* species and is more common in adults.[22] The analysis showed that the *Bifidobacterium* spp. levels were significantly lower at 12m compared to baseline and 6m, whereas the opposite was the case for ER/CC ($P < 0.05$, $Q < 0.1$, **Figure 4A-B**). ER/CC was also significantly lower in infants that were still breastfed versus non-breastfed infants ($P < 0.05$, $Q < 0.1$, **Figure 4D**) and those not receiving formula versus those that did at 12m ($P < 0.05$, $Q > 0.1$, **Figure S4**). Complementary feeding and delivery mode were not associated with significant differences in the examined taxa (**Figure S4**). The baseline *Bifidobacterium* spp. levels of the infants who developed allergy by 12m were lower than those that did not ($P < 0.05$, $Q > 0.01$, **Figure 4C**). A follow-up Spearman correlation analysis showed no evidence of a correlation between reduced *Bifidobacterium* spp. levels and elevated LCFAs levels (**Figure S5**).

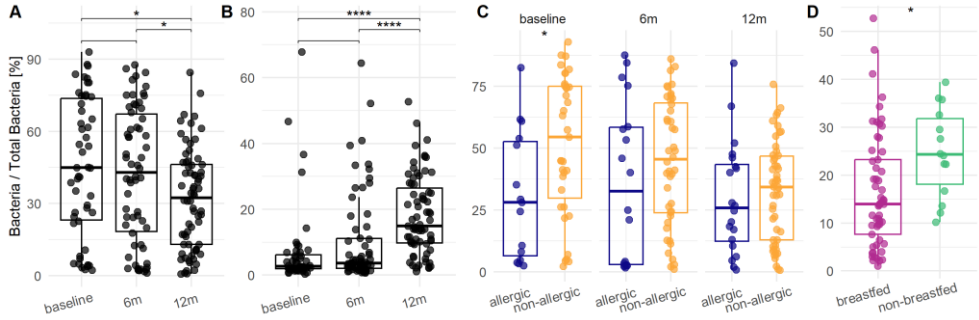

Figure 4. *The levels of A) Bifidobacterium spp. between visits; B) ER/CC between visits; C) Bifidobacterium spp. between the allergic (blue) and non-allergic (orange) infants at each visit; D) ER/CC between breastfed (pink) and non-breasted (green) at 12m as proportion of the total bacteria. Statistical analysis was performed using Mann-Whitney test: P < 0.05 (*), P < 0.01 (**), P < 0.001 (***), P < 0.0001 (****). Number of measurements per group and visit: Bifidobacterium spp.: n = [50, 62, 70]; ER/CC: n = [48, 60, 71] for baseline, 6m and 12m, respectively.*

4. Discussion

In this study, healthy breastfed infants at risk of allergy were followed throughout their first year. During this period, alteration in the fecal metabolome and key microbiome members

(*Bifidobacterium* spp., ER/CC) were examined in relation to age, feeding practices, mode of delivery, and allergy development. Strong age-associated alterations were observed including an overall increase in AAs and derivatives, BAs, nucleobases, nucleosides and derivatives, B vitamins and derivatives, SCFAs, and phenolic acids, along with a decrease in LCFAs and acylcarnitines. Feeding practices, specifically the intake of complementary food and cessation of breastfeeding were significantly associated with changes to the metabolome. Delivery mode had a pronounced impact on the metabolome with distinct differences observed mainly at baseline, some of which persisted until 6m. Meanwhile, infants who developed allergy had significantly lower *Bifidobacterium* spp. and higher LCFA levels at baseline.

These strong age-associated metabolome changes align with previous studies examining the fecal metabolome in early life.[23,24] These pronounced shifts are expected, given the rapid physical growth[2] and GM diversification associated with the transition from milk to solid food during this period.[25] In our cohort, the diversification is evident by the significant decline in bifidobacteria and the increase in the adult-like ER/CC at 12m as well as higher ER/CC in non-breastfed versus breastfed infants at 12m.

The shift to a diet richer in fiber and protein, and the resulting GM diversification, is also clearly reflected at the metabolomic level. The decline in pyruvate after 6m following an initial increase, and its negative association with complementary feeding, likely reflects its conversion to downstream metabolites as the GM diversifies.[26] The increase in fiber intake was also evident by the rise of butyrate and propionate after 6m and their positive association with complementary feeding and cessation of breastfeeding, respectively, in agreement with Tsukuda *et al.*[27] That aligns well with the observed increase in the well-known butyrate producers within ER/CC.[28] The observed temporal increase in phenolic acids, alongside associations with breastfeeding and complementary feeding, is consistent with their diverse origins, including plants,[29] breastmilk,[30] and microbial flavonoid and tyrosine transformation.[29] Meanwhile, the higher levels of carnosine[31] and N2,N2-dimethylguanosine[32] in complementary-fed infants may indicate meat consumption.

A shift from a bifidobacteria-rich to a more adult-like microbiome was also evident by the change in microbial aromatic AA metabolites. As expected, the aromatic lactic acids ILA, 4-OH-PLA#, and PLA#, known to be produced by infant-type bifidobacterial species,[33] increased until 6m. Subsequently PLA# levels remained unchanged, whereas those of ILA and 4-OH-PLA# declined and were lower in infants who received no breastmilk at 12m, supporting Sillner *et al.*'s findings.[34] The observations align with the bifidobacterial decline at 12m. The aromatic acetic and propionic acids IAA, 4-OH-PAA, PAA, and 4-OH-PPA increased after 6m and were positively associated with complementary feeding, likely reflecting increased microbial protein degradation.[26] In contrast, IPA did not rise with age, however, it was detected in more infants at 12m compared to 6m and was positively associated with cessation of breastfeeding, suggesting that IPA producers are more common GM members at 12m. The phenylalanine-derived PAGIn and tryptophan-derived

indoxyl sulfuric acid, declined with the introduction of complementary feeding contrary to their expected increase.[26] These metabolites are of particular interest due to their known detrimental effects on health in adults and remain understudied in early life.[26] Though B vitamins can be obtained from the diet, including breastmilk,[35] their temporal rise is also likely attributed to microbial production as multiple GM members are well-established B vitamin producers.[36]

As anticipated, the abundance and diversity of secondary BAs increased with age and the two drivers of GM diversification: introduction to complementary foods and the cessation of breastfeeding. Similar to Sillner *et al.*[34] we report on less-studied secondary BAs (7ketoLCA, 3oxoCDCA, 7,12oxoLCA, 7oxoDCA, 3oxoCA, UCA) in infancy, along with almost complete absence of LCA and DCA until 12m.[34] The latter aligns with the observed increase in ER/CC well-known for its high 7α-dehydroxylating activity required for their production.[37] Meanwhile, the rise in glyco-BAs after 6m likely reflects the reduction in particularly effective glyco-BAs deconjugators bifidobacteria.[38]

The decline in acylcarnitines after 6m and their positive association with breastfeeding, along with the negative association of LCFAs with age suggest increasing reliance on beta oxidation. Production of conjugated linoleic and linolenic acid isomers by bifidobacteria[39] possibly also contributes to the decline in LA and ALA# before 6m, a period characterized by bifidobacterial dominance.

Multiple studies have shown strong fecal metabolome differences between breastfed and formula-fed infants.[23,34,40,41] However, unlike these studies, our cohort consisted of infants breastfed for at least 16 weeks, with formula-feeding often initiated alongside breastfeeding, mainly after the introduction of complementary feeding. The minor significant associations observed with formula feeding in this cohort, agree with He *et al.*,[40] who reported convergence of the metabolome profiles between breast-fed and formula-fed infants following complementary feeding.

Despite its known importance in shaping the GM,[42] delivery mode was not associated with microbiome differences in this cohort. It did, however, affect the metabolome, especially at baseline and up to 6m. Earlier studies reported no metabolome changes despite shifts in the microbiome composition,[43] or significant alterations that differ from our findings and between each other.[24,44] These discrepancies may reflect ethnic or age-related cohort differences. We found the HMO building-block Neu5Ac to be significantly higher in the vaginal compared to the C-section group and maternal ppBMI to be associated with C-section. Since ppBMI has been linked to HMO composition,[45] the difference in Neu5Ac may be attributed to variations in breastmilk composition. Another possibility is that vaginally-delivered infants' guts are richer in taxa like *Bacteroides* capable of cleaving sialic acids HMO residues.[46]

Unexpectedly, no metabolome differences were observed between the allergic groups at 6m and 12m, despite the emergence of allergic symptoms in this period. Feeding changes

and the resulting GM shifts may have masked these differences. Infants who developed allergy during the study did, however, have significantly higher baseline levels of the LCFAs, mainly polyunsaturated fatty acids (PUFAs), including n-6 LA and n-3 EPA, ALA#, DHA. Although elevated plasma n-3 and n-6 PUFA levels have also been reported in children with food allergy,[47] lower n-3 PUFA levels are generally associated with increased allergy risk.[48,49] Along with the higher LCFA levels, we observed lower *Bifidobacterium* spp. in the allergic group. The absence of significant correlation between the two, however, suggests that the lower LCFA levels in the allergic infants are unlikely to be due to bifidobacteria. Instead, the difference may be due to variations in mother's breastmilk composition,[50] microbial transformation,[39] or differences in intestinal absorption.

Our study has several limitations, including the small sample size, the wide age range at baseline, and the infrequent sampling. The limited sample size, especially in the allergic group, prevented a separate analysis of the different allergy types. To enhance metabolomic interpretation and clarify the GM–allergy link, future research should consider whole microbiome dynamics rather than focusing solely on specific taxa and prioritize integrated microbiome–metabolome analyses. Meanwhile, examining the circulating metabolome and breastmilk compositional analysis are of interest to respectively understand the plausible link between LCFAs and allergy and aid the interpretation of the delivery mode findings.

5. Conclusion

This study offers valuable new insights into the longitudinal fecal metabolome development in infancy, a critical period with lasting implications for immune system development. Our findings reveal substantial metabolomic shifts with age likely due to changes to the host metabolism, diet, and the GM. Notably, we show that C-section is significantly associated with fecal metabolome alterations up to 6m, though the health implications of these changes require further investigation. This study showed that low *Bifidobacterium* spp. and LCFAs precede allergy, indicating a temporal association that suggests a direction for follow-up studies on their potential role in allergy development.

Author contributions

M.V.S: Conceptualization, Investigation, Methodology, Formal Analysis, Visualization, Data curation, Writing – Original Draft Preparation; **P.Z.**: Conceptualization, Investigation, Methodology, Writing – Review & Editing; **A.K.**: Conceptualization, Supervision, Writing – Review & Editing; **The TEMPO study team**: Resources; **H.W.**: Conceptualization, Writing – review and editing **C.B.**: Conceptualization, Funding acquisition, Writing – review and editing; **A.C.H.**: Conceptualization, Supervision, Writing – Review & Editing; T.H.: Conceptualization, Supervision, Funding acquisition, Writing – review & editing.

Acknowledgments

Pascal Mass is greatly appreciated for his invaluable assistance in metabolomics data pre-processing. We thank all the study investigators for their contribution in data and sample collection in the TEMPO study, namely: Mazin Alhakim, László Barkai, Csaba Bartha, Ildikó Batta, Viktor Bauer, Shira Benor, Kirsten Beyer, Elena Bradatan, Katrina Cathie, Chong Chan Poh, An-Chyi Chen, Shih-Ming Chu, Elisa Civardi, Ronit Confino-Cohen, Maria Couce Pico, Daniel Drazan, Jitka Fabianova, Allessandro Giovanni Fiocchi, Montserrat Garriga, Francisco Giménez Sánchez, Anne Goh Eng Neo, Monique Gorissen, Martin Gregora, Ludmila Grossmanova, Zuzana Havlicekova, Stephen Hughes, Jose Hurtado, Natalia Klocanova, Éva Kovács, Silvia Labovska, István Laki, Anja Lange, Yu Lung Lau, Ting F. Leung, Danica Mankova, Nofar Marcus, Louise J. Michaelis, Zuzana Nagyova, López Eduardo Narbona, Antonio Nieto, Lee Noimark, Daniela Olexova, Miroslava Ondrejkova, Nikolaos G. Papadopoulos, Stefaan Peeters, Paola Roggero, Renata Ruzkova, Miguel Sáenz de Pipaón, Ignacio Salamanca de la Cueva, Vered Schichter-Konfino, Beata Sediva, Eduardo Shahar, Pavol Simurka, Sylva Skalova, Françoise Smets, László Somorjai, Zev Sthoeger, Zbynek Stranak, Edina Stunya, Erzsebet Szakos, Ron van Beek, Vivienne van de Walle, Hans van Goudoever, Yvan Vandenplas, Mirko Zibolen.

Conflict of interest statement

Harm Wopereis is an employee of Danone Research & Innovation. The project is part of a partnership programme between NWO-TTW and Danone Research & Innovation. The other authors declare that they have no known conflicts of interest.

Financial support

This study was part of the EARLYFIT project (Partnership programme NWO Domain AES-Danone Research & Innovation), funded by the Dutch Research Council (NWO) and Danone Research & Innovation (project number: 16490). Pingping Zhu Would like to acknowledge the China Scholarship Council (CSC, No. 201906240049). A.C.H and T.H. are supported by the Dutch Research Council (NWO) funded Netherlands X-omics Initiative (project number 184.034.019).

Data availability

The metabolomics data of this study are submitted in MetaboLights at www.ebi.ac.uk/metabolights/ with reference number MTBLS8954, along with limited clinical metadata. Additional individual-level metadata, even pseudonymized, are sensitive and are protected by the GDPR and not publicly available. Reasonable data sharing requests based on data processing and material transfer agreements can be made to Danone Research & Innovation (www.danoneresearch.com/).

References

1. Wu, Z. A. & Wang, H. X. A Systematic Review of the Interaction Between Gut Microbiota and Host Health from a Symbiotic Perspective. SN Compr. Clin. Med. 1, 224–235 (2019).
2. Yatsunenko, T. et al. Human gut microbiome viewed across age and geography. Nature 486, 222–227 (2012).
3. Simon, A. K., Hollander, G. A. & McMichael, A. Evolution of the immune system in humans from infancy to old age. Proceedings of the Royal Society B: Biological Sciences 282, 20143085 (2015).
4. Gensollen, T., Iyer, S. S., Kasper, D. L. & Blumberg, R. S. How colonization by microbiota in early life shapes the immune system. Science 352, 539–544 (2016).
5. Sarkar, A., Yoo, J. Y., Valeria Ozorio Dutra, S., Morgan, K. H. & Groer, M. The Association between Early-Life Gut Microbiota and Long-Term Health and Diseases. Journal of Clinical Medicine 10, 459 (2021).
6. Fazlollahi, M. et al. Early-life gut microbiome and egg allergy. Allergy 73, 1515–1524 (2018).
7. Savova, M. V. et al. Current insights into cow's milk allergy in children: Microbiome, metabolome, and immune response—A systematic review. Pediatric Allergy and Immunology 35, e14084 (2024).
8. Yang, L., Fu, J. & Zhou, Y. Research Progress in Atopic March. Front. Immunol. 11, (2020).
9. Milani, C. et al. The First Microbial Colonizers of the Human Gut: Composition, Activities, and Health Implications of the Infant Gut Microbiota. Microbiology and Molecular Biology Reviews 81, 10.1128/mmbr.00036-17 (2017).
10. Ríos-Covian, D., Langella, P. & Martín, R. From Short- to Long-Term Effects of C-Section Delivery on Microbiome Establishment and Host Health. Microorganisms 9, 2122 (2021).
11. Hoskinson, C. et al. Antibiotics taken within the first year of life are linked to infant gut microbiome disruption and elevated atopic dermatitis risk. Journal of Allergy and Clinical Immunology 154, 131–142 (2024).
12. Nuzzi, G., Di Cicco, M. E. & Peroni, D. G. Breastfeeding and Allergic Diseases: What's New? Children (Basel) 8, 330 (2021).
13. Garwolińska, D., Namieśnik, J., Kot-Wasik, A. & Hewelt-Belka, W. Chemistry of Human Breast Milk—A Comprehensive Review of the Composition and Role of Milk Metabolites in Child Development. J. Agric. Food Chem. 66, 11881–11896 (2018).
14. Parigi, S. M., Eldh, M., Larssen, P., Gabrielsson, S. & Villablanca, E. J. Breast Milk and Solid Food Shaping Intestinal Immunity. Front. Immunol. 6, (2015).
15. Lin, C. et al. Intestinal 'Infant-Type' Bifidobacteria Mediate Immune System Development in the First 1000 Days of Life. Nutrients 14, 1498 (2022).
16. Krautkramer, K. A., Fan, J. & Bäckhed, F. Gut microbial metabolites as multi-kingdom intermediates. Nat Rev Microbiol 19, 77–94 (2021).
17. Papadopoulos, N. G. et al. Mixed Milk Feeding: A New Approach to Describe Feeding Patterns in the First Year of Life Based on Individual Participant Data from Two Randomised Controlled Trials. Nutrients 14, 2190 (2022).
18. Sim, K. et al. Improved Detection of Bifidobacteria with Optimised 16S rRNA-Gene Based Pyrosequencing. PLoS ONE 7, e32543 (2012).
19. Zhu, P. et al. Exploring the Fecal Metabolome in Infants With Cow's Milk Allergy: The Distinct Impacts of Cow's Milk Protein Tolerance Acquisition and of Synbiotic Supplementation. Mol Nutr Food Res 69, e202400583 (2025).
20. Peet, M. van der et al. mzQuality: A tool for quality monitoring and reporting of targeted mass spectrometry measurements. 2025.01.22.633547 Preprint at https://doi.org/10.1101/2025.01.22.633547 (2025).

21. Wopereis, H. et al. Intestinal microbiota in infants at high risk for allergy: Effects of prebiotics and role in eczema development. Journal of Allergy and Clinical Immunology 141, 1334-1342.e5 (2018).
22. Mueller, S. et al. Differences in Fecal Microbiota in Different European Study Populations in Relation to Age, Gender, and Country: a Cross-Sectional Study. Applied and Environmental Microbiology 72, 1027–1033 (2006).
23. Holzhausen, E. A. et al. Longitudinal profiles of the fecal metabolome during the first 2 years of life. Sci Rep 13, 1886 (2023).
24. Ouyang, R. et al. Maturation of the gut metabolome during the first year of life in humans. Gut Microbes 15, 2231596 (2023).
25. Laursen, M. F. et al. Infant Gut Microbiota Development Is Driven by Transition to Family Foods Independent of Maternal Obesity. mSphere 1, 10.1128/msphere.00069-15 (2016).
26. Roager, H. M., Stanton, C. & Hall, L. J. Microbial metabolites as modulators of the infant gut microbiome and host-microbial interactions in early life. Gut Microbes 15, 2192151 (2023).
27. Tsukuda, N. et al. Key bacterial taxa and metabolic pathways affecting gut short-chain fatty acid profiles in early life. The ISME Journal 15, 2574–2590 (2021).
28. Barcenilla, A. et al. Phylogenetic Relationships of Butyrate-Producing Bacteria from the Human Gut. Applied and Environmental Microbiology 66, 1654–1661 (2000).
29. Kiriyama, Y., Tokumaru, H., Sadamoto, H., Kobayashi, S. & Nochi, H. Effects of Phenolic Acids Produced from Food-Derived Flavonoids and Amino Acids by the Gut Microbiota on Health and Disease. Molecules 29, 5102 (2024).
30. Fiecke, C., Knox, N., Andres, A., Ferruzzi, M. G. & Kay, C. Polyphenol metabolites in human milk: Potential role in support of healthy infant development, a narrative review. 2025.02.04.25321667 Preprint at https://doi.org/10.1101/2025.02.04.25321667 (2025).
31. Mitry, P. et al. Plasma concentrations of anserine, carnosine and pi-methylhistidine as biomarkers of habitual meat consumption. Eur J Clin Nutr 73, 692–702 (2019).
32. Wishart, D. S. et al. HMDB 5.0: the Human Metabolome Database for 2022. Nucleic Acids Research 50, D622–D631 (2022).
33. Laursen, M. F. et al. Bifidobacterium species associated with breastfeeding produce aromatic lactic acids in the infant gut. Nat Microbiol 6, 1367–1382 (2021).
34. Sillner, N. et al. Longitudinal Profiles of Dietary and Microbial Metabolites in Formula- and Breastfed Infants. Front. Mol. Biosci. 8, (2021).
35. Qiao, W. et al. A cohort study of vitamins contents in human milk from maternal-infant factors. Front. Nutr. 9, (2022).
36. Wan, Z. et al. Intermediate role of gut microbiota in vitamin B nutrition and its influences on human health. Front. Nutr. 9, (2022).
37. Ridlon, J. M., Kang, D. J., Hylemon, P. B. & Bajaj, J. S. Bile Acids and the Gut Microbiome. Curr Opin Gastroenterol 30, 332–338 (2014).
38. Kim, G.-B., Yi, S.-H. & Lee, B. H. Purification and Characterization of Three Different Types of Bile Salt Hydrolases from Bifidobacterium Strains. Journal of Dairy Science 87, 258–266 (2004).
39. Gorissen, L. et al. Production of conjugated linoleic acid and conjugated linolenic acid isomers by Bifidobacterium species. Appl Microbiol Biotechnol 87, 2257–2266 (2010).
40. He, X. et al. Fecal microbiome and metabolome of infants fed bovine MFGM supplemented formula or standard formula with breast-fed infants as reference: a randomized controlled trial. Sci Rep 9, 11589 (2019).
41. Chalifour, B. et al. The potential role of early life feeding patterns in shaping the infant fecal metabolome: implications for neurodevelopmental outcomes. npj Metab Health Dis 1, 2 (2023).
42. Rutayisire, E., Huang, K., Liu, Y. & Tao, F. The mode of delivery affects the diversity and colonization pattern of the gut microbiota during the first year of infants' life: a systematic review. BMC Gastroenterol 16, 86 (2016).

43. Hoen, A. G. et al. Association of Cesarean Delivery and Formula Supplementation with the Stool Metabolome of 6-Week-Old Infants. Metabolites 11, 702 (2021).
44. Li, N. et al. Distinct gut microbiota and metabolite profiles induced by delivery mode in healthy Chinese infants. Journal of Proteomics 232, 104071 (2021).
45. Han, S. M. et al. Maternal and Infant Factors Influencing Human Milk Oligosaccharide Composition: Beyond Maternal Genetics. The Journal of Nutrition 151, 1383–1393 (2021).
46. Kijner, S., Ennis, D., Shmorak, S., Florentin, A. & Yassour, M. CRISPR-Cas-based identification of a sialylated human milk oligosaccharides utilization cluster in the infant gut commensal Bacteroides dorei. Nat Commun 15, 105 (2024).
47. Crestani, E., Benamar, M., Phipatanakul, W., Rachid, R. & Chatila, T. A. Age-specific Metabolomic profiles in children with food allergy. Clinical Immunology 261, 109928 (2024).
48. Jonsson, K. et al. Serum fatty acids in infants, reflecting family fish consumption, were inversely associated with allergy development but not related to farm residence. Acta Paediatrica 105, 1462–1471 (2016).
49. Lee-Sarwar, K. et al. Dietary and Plasma Polyunsaturated Fatty Acids Are Inversely Associated with Asthma and Atopy in Early Childhood. The Journal of Allergy and Clinical Immunology: In Practice 7, 529-538.e8 (2019).
50. Bobiński, R. & Bobińska, J. Fatty acids of human milk – a review. International Journal for Vitamin and Nutrition Research (2020).

Supplementary Materials

Sample collection and storage procedure

Fecal samples were collected at home and immediately stored in freezers, then transferred on ice to the participant hospitals and stored at −80◦C until transfer to Danone Research & Innovation (Utrecht, The Netherlands) for wet sample aliquoting and fluorescence *in situ* hybridization (FISH), SCFAs and lactic acid analysis. Sample aliquots for LC-MS metabolomics analysis were transferred on dry ice to Leiden University and stored at −80◦C until analysis.

Data analysis details

For the 162 metabolites with missingness <20%, missing values were imputed using quantile regression imputation of left-censored data method.[1] Two LMMs were built to investigate the effect of age and diet on the fecal metabolome with age (days), breastfeeding (yes/no), formula feeding (yes/no), and complementary feeding (yes/no) used as fixed effects and subject ID as a random effect. The first model investigated the baseline and 6m measurements (*Metabolite ~ age + complementary feeding + formula feeding + (1|ID)*), whereas the second the 6m and 12m measurements (*Metabolite ~ age + complementary feeding + formula feeding + breastfeeding + (1|ID)*). Reference levels for the variables were the median age at baseline, "no" for complementary and formula feeding, and "yes" for breastfeeding. Breastfeeding was included only in the second model since a single subject was non-breastfed at 6m (**Table 1**). As infants' diet changed during the first year (**Table 1**), the feeding practices (breastfeeding, formula feeding, complementary feeding) at each visit were considered. Because of the change in feeding practices and due to the choice of reference levels, the breastfeeding, complementary feeding and formula feeding coefficients are interpreted respectively as cessation of breastfeeding, introduction of complementary feeding and introduction of formula feeding. To study the effect of delivery mode on the metabolome a LMM with age (days), delivery mode (Vaginal/C-section), and their interaction as fixed effects was used (*Metabolite ~ age + delivery mode + age:delivery mode + (1|ID)*). Reference levels for the variables were the median age at baseline and "Vaginal" delivery. Similarly, the LMM for allergy had age, allergy status at 12 months (allergic, non-allergic) and their interaction as fixed effects (*Metabolite ~ age + allergy status + age:allergy status + (1|ID)*). Median age at baseline and the non-allergic group were used as a reference. A separate LMM was then constructed after stratifying the allergy group by allergy type (IgE-mediated, non-IgE–mediated, and non-allergic), with the non-allergic group as the reference.

Supplementary References

1. Wei, R. et al. Missing Value Imputation Approach for Mass Spectrometry-based Metabolomics Data. Sci Rep 8, 663 (2018).

Table S1. Analyte names and abbreviations

Metabolite(s)	Abbreviation
1,3-Dimethylurate/1,7-Dimethylurate	1,3-Dimethylurate#
12-keto Chenodeoxycholic Acid	12ketoCDCA
12-keto Lithocholic Acid	12ketoLCA
p-Hydroxymandelate/3,4-Dihydroxyphenylacetic acid	3,4-Dihydroxyphenylacetic acid#
3-Hydroxyhippurate/2-Hydroxyhippurate/4-hydroxyhippurate	3-Hydroxyhippurate#
3-Methylhistidine/1-Methylhistidine	3-Methylhistidine#
3-Oxocholic Acid	3oxoCA
3-oxochenodeoxycholic acid	3oxoCDCA
3-oxo Deoxycholic Acid	3oxoDCA
Mandelic acid/4-Hydroxyphenylacetic acid	4-OH-PAA#
Dihydrocaffeic acid / 3-hydroxy-3-(3-hydroxyphenyl)propanoic acid/Hydroxyphenyllactic acid (4-OH-PLA/DHCA/HPHPA)	4-OH-PLA#
4-Hydroxyphenylpropionic acid	4-OH-PPA
5-Aminolevulinic acid/4-Hydroxyproline	5-Aminolevulinic acid#
7,12-Diketolithocholic Acid	7.12oxoLCA
7-Ketodeoxycholic acid	7ketoDCA
7-ketolithocholic acid	7ketoLCA
Arachidonic acid	AA
Adenosine/Deoxyguanosine	Adenosine#
Alpha-Linolenic acid/Gamma-Linolenic acid	ALA#
Alanine/beta-Alanine/Sarcosine	Alanine#
Allocholic acid	alloCA
Allolithocholic Acid	alloLCA
beta-Muricholic Acid	bMCA
Butyrylcarnitine/Isobutyrylcarnitine	Butyrylcarnitine#
Cholic acid	CA
Chenodeoxycholic acid	CDCA
Creatine/3-Guanidinopropanoate	Creatine#
Deoxycholic acid	DCA
Dihomo-alpha-linolenic acid; eicosatrienoic acid/Dihomo-gamma-linolenic acid (DGLA)	DGLA#
Docosahexaenoic acid	DHA
Docosapentaenoic acid	DPA
3-Epideoxycholic Acid	EDCA
Eicosapentaenoic acid	EPA
Alpha-aminobutyric acid/Gamma-aminobutyric acid	GABA#
Galactitol/Mannitol/Sorbitol	Galactitol#
Glycocholic acid	GCA
Glycochenodeoxycholic acid	GCDCA
Glycodeoxycholic acid	GDCA
Glycohyocholic Acid	GHCA
Glycohydeoxycholic acid	GHDCA
Glycolithocholic acid	GLCA
O-Acetylserine/Glutamic acid	Glutamate#
Guanosine/8-Hydroxy-2-deoxyguanosine	Guanosine#
Glycoursodeoxycholic acid	GUDCA

Metabolite(s)	Abbreviation
Hyocholic acid	HCA
Hyodeoxycholic acid	HDCA
Hyodeoxycholic acid	HDCA
Targinine/Homoarginine	Homoarginine#
Indoleacetic acid	IAA
Indolelactic acid	ILA
Indolepropionic acid	IPA
Isolithocholic Acid	isoLCA
Isoursodeoxycholic Acid	isoUDCA
Linoleic acid	LA
Lithocholic acid	LCA
Murideoxycholic acid	MDCA
3-Methylxanthine/1-Methylxanthine/7-Methylxanthine	Methylxanthine isomers
N-Acetylneuraminic Acid	Neu5Ac
Oleic acid	OA
Phenylacetylglutamine	PAGln
Paraxanthine/Theophylline	Paraxanthine#
Phenyllactic acid/3-(3-Hydroxyphenyl)propanoic acid	PLA#
Taurocholic acid	TCA
Taurochenodesoxycholic acid	TCDCA
Taurodeoxycholic acid	TDCA
Taurohyocholic acid	THCA
Taurohyodeoxycholic Acid	THDCA
Threonine/Homoserine	Threonine#
Taurolithocholic acid	TLCA
Taurolithocholic acid 3-sulfate	TLCA-3S
Trimethylamine N-oxide	TMAO
Total branched-chain short chain fatty acids	Total BSCFA
Total short-chain fatty acids	Total SCFA
Tauroursodeoxycholic Acid	TUDCA
Ursocholic acid	UCA
Ursodeoxycholic acid	UDCA

Table S2. Assessment of association between categorical clinical variables and either allergy status and delivery mode using Fisher's exact test. Only significant associations are displayed (P > 0.05)

Clinical Variables	Groups	P
dermatitis and eczema by 12 months	Allergic - Non allergic	2.8E-12
dermatitis and eczema by 6 months	Allergic - Non allergic	1.7E-09
rashes, eruptions and exanthems (unclassified cause) by 6 months	formula fed - not formula fed at 6 m	1.6E-04
rashes, eruptions and exanthems (unclassified cause) by 12 months	formula fed - not formula fed at 6 m	1.6E-04
mother's alcohol consumption prepregnancy	complementary-fed - not-complementary-fed at 6m	4.2E-03
mineral supplements by 12 months	Vaginal - C-section	4.7E-03
country	formula fed - not formula fed at 6 m	4.8E-03
ear infections by 12 months	Allergic - Non allergic	5.4E-03

Clinical Variables	Groups	P
analgesics by 6 months	complementary-fed - not-complementary-fed at 6m	1.0E-02
ophthalmological and otological preparations by 6 months	Vaginal - C-section	1.1E-02
analgesics by 12 months	formula fed - not formula fed at 6 m	1.2E-02
upper respiratory tract signs and symptoms by 12 months	Vaginal - C-section	1.5E-02
pets at home	complementary-fed - not-complementary-fed at 6m	1.6E-02
country	Allergic - Non allergic	1.8E-02
otologicals by 12 months	Allergic - Non allergic	1.9E-02
emollients and protectives by 6 months	Allergic - Non allergic	1.9E-02
emollients and protectives by 12 months	Allergic - Non allergic	1.9E-02
pets at home	breastfed - non breastfed at 12 m	1.9E-02
general signs and symptoms (unclassified cause) by 12 months	breastfed - non breastfed at 12 m	2.1E-02
dyspeptic signs and symptoms by 6 months	formula fed - not formula fed at 6 m	2.3E-02
dyspeptic signs and symptoms by 12 months	formula fed - not formula fed at 6 m	2.3E-02
rashes, eruptions and exanthems (unclassified cause) by 6 months	formula fed - not formula fed at 12 m	3.3E-02
rashes, eruptions and exanthems (unclassified cause) by 12 months	formula fed - not formula fed at 12 m	3.3E-02
non site specific injuries (unclassified cause) by 12 months	breastfed - non breastfed at 12 m	3.6E-02
ophthalmological and otological preparations by 12 months	Vaginal - C-section	3.7E-02
antihistamines for systemic use by 12 months	formula fed - not formula fed at 12 m	3.8E-02
daycare	complementary-fed - not-complementary-fed at 6m	3.9E-02
analgesics by 6 months	formula fed - not formula fed at 6 m	3.9E-02
antiinflammatory and antirheumatic products by 12 months	formula fed - not formula fed at 6 m	4.0E-02
antihemorrhagics by 6 months	breastfed - non breastfed at 12 m	4.0E-02
antihemorrhagics 12 months	breastfed - non breastfed at 12 m	4.0E-02
pets at home	Vaginal - C-section	4.1E-02
mother's alcohol consumption during lactation	complementary-fed - not-complementary-fed at 6m	4.9E-02
ophthalmological and otological preparations by 12 months	formula fed - not formula fed at 12 m	5.0E-02
nasal preparations by 6 months	breastfed - non breastfed at 12 m	5.7E-02

Table S3. Assessment of association between numeric clinical variables and either allergy status and delivery mode using Mann–Whitney U test. Only significant associations are displayed (P < 0.05)

Clinical Variable	Groups	P
gestational age	Vaginal - C-section	2.79E-03
maternal pre-pregnancy BMI	Vaginal - C-section	4.49E-02
number of vaccines by 6 months	Allergic - Non allergic	5.42E-02
age at baseline	Allergic - Non allergic	5.67E-02
number of vaccines by 12 months	Allergic - Non allergic	5.88E-02
birth length	breastfed - non breastfed at 12 m	7.44E-03
birth head circumference	breastfed - non breastfed at 12 m	2.26E-02
number of vaccines by 6 months	complementary-fed - not-complementary-fed at 6m	3.10E-02
number of vaccines by 12 months	complementary-fed - not-complementary-fed at 6m	3.54E-02

Table S4. Results of the LMM analysis examining the effect of age, breastfeeding, complementary feeding, and formula feeding. The table displays only the fixed-effect coefficients for formula feeding; coefficients for the other predictors are not shown and only estimates significant prior multiple testing correction (P < 0.05). Information on LMM used and all abbreviations can be found in Supplementary materials.

Metabolite	Model	Estimate	$P_{unadjusted}$	$P_{adjusted}$ (Q)
Thymine	baseline – 6m	-1.29E+00	3.19E-04	5.16E-02
2-Deoxyinosine	baseline – 6m	-9.42E-01	1.15E-02	4.08E-01
2-Deoxyuridine	baseline – 6m	-9.58E-01	8.88E-03	4.08E-01
Pantothenic acid	baseline – 6m	-9.06E-01	1.26E-02	4.08E-01
Pyridoxal	baseline – 6m	-9.61E-01	7.74E-03	4.08E-01
Nicotinic acid	baseline – 6m	-8.48E-01	1.60E-02	4.33E-01
Allantoin	baseline – 6m	7.88E-01	3.12E-02	4.77E-01
Guanosine#	baseline – 6m	8.01E-01	2.31E-02	4.77E-01
Propionylcarnitine	baseline – 6m	9.13E-01	3.24E-02	4.77E-01
Saccharopine	baseline – 6m	-7.06E-01	3.15E-02	4.77E-01
Uracil	baseline – 6m	-8.23E-01	2.61E-02	4.77E-01
Glutamate#	baseline – 6m	-8.03E-01	4.34E-02	5.86E-01

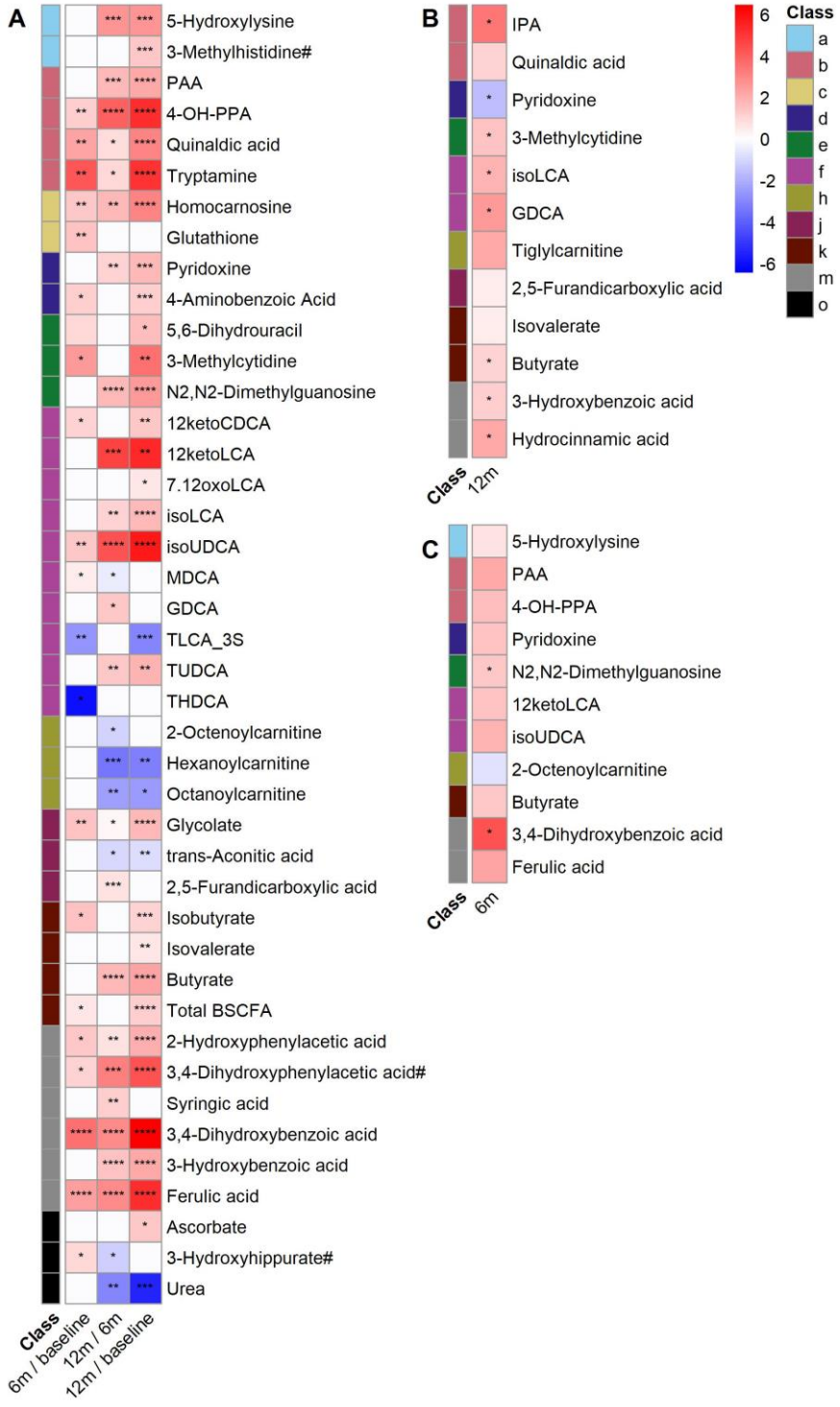
A
5-Hydroxylysine
3-Methylhistidine#
PAA
4-OH-PPA
Quinaldic acid
Tryptamine
Homocarnosine
Glutathione
Pyridoxine
4-Aminobenzoic Acid
5,6-Dihydrouracil
3-Methylcytidine
N2,N2-Dimethylguanosine
12ketoCDCA
12ketoLCA
7.12oxoLCA
isoLCA
isoUDCA
MDCA
GDCA
TLCA_3S
TUDCA
THDCA
2-Octenoylcarnitine
Hexanoylcarnitine
Octanoylcarnitine
Glycolate
trans-Aconitic acid
2,5-Furandicarboxylic acid
Isobutyrate
Isovalerate
Butyrate
Total BSCFA
2-Hydroxyphenylacetic acid
3,4-Dihydroxyphenylacetic acid#
Syringic acid
3,4-Dihydroxybenzoic acid
3-Hydroxybenzoic acid
Ferulic acid
Ascorbate
3-Hydroxyhippurate#
Urea
Class
6m / baseline
12m / 6m
12m / baseline
B
IPA
Quinaldic acid
Pyridoxine
3-Methylcytidine
isoLCA
GDCA
Tiglylcarnitine
2,5-Furandicarboxylic acid
Isovalerate
Butyrate
3-Hydroxybenzoic acid
Hydrocinnamic acid
Class
12m
6
4
2
0
-2
-4
-6
Class
a
b
c
d
e
f
h
j
k
m
o
C
5-Hydroxylysine
PAA
4-OH-PPA
Pyridoxine
N2,N2-Dimethylguanosine
12ketoLCA
isoUDCA
2-Octenoylcarnitine
Butyrate
3,4-Dihydroxybenzoic acid
Ferulic acid
Class
6m

Figure S1. *Comparison of the fold changes of metabolites A) between the visits, B) breastfeeding groups at 12m, C) complementary feeding at 6m. The comparison was performed using paired Mann-Whitney U test on the metabolites excluded for the LMM analysis due to high missingness. Colors represent the fold change: above 0 (red), below 0 (blue), p>0.05 (white). Asterisks indicate statistical significance: Q < 0.1 (*), Q < 0.01 (**), Q < 0.001 (***), Q > 0.001 (****). The "#" in the metabolite names indicates that the metabolite has coeluted with another target metabolite. Information on that and all abbreviations can be found in* ***Table S1***. *Class annotations: a - AAs and derivatives; b aromatic AAs metabolites; c - dipeptides and tripeptides; d - B vitamins and derivatives; e - nucleobases, nucleosides and derivatives; f – BAs; h – carnitines; j - organic acids; k – SCFAs; m - phenolic acids; o – other. The "#" in the metabolite names indicates that the metabolite coeluted with another target metabolite. All abbreviations and coeluting metabolites can be found in* ***Table S1***.

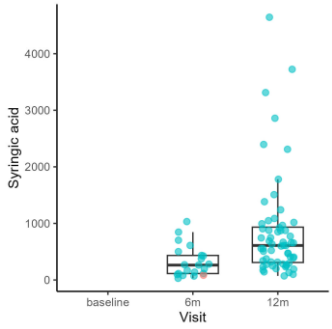

Figure S2. *Boxplots showing the distribution of syringic acid at the three visits (baseline, 6m, 12m). Individual data points are jittered and colored according to the complementary feeding status: blue – complementary-fed and orange: not complementary fed*

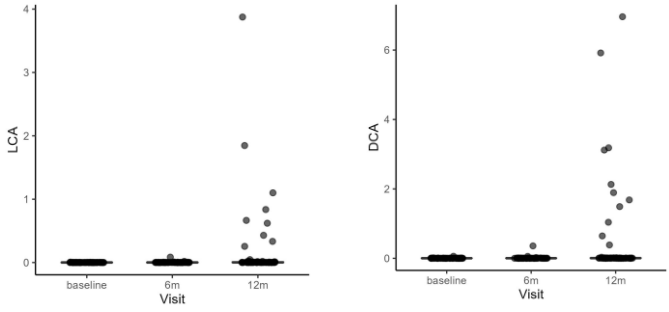

Figure S3. *Boxplots showing the distribution of LCA and DCA between the visits (baseline, 6m, 12m). Individual data points are jittered*

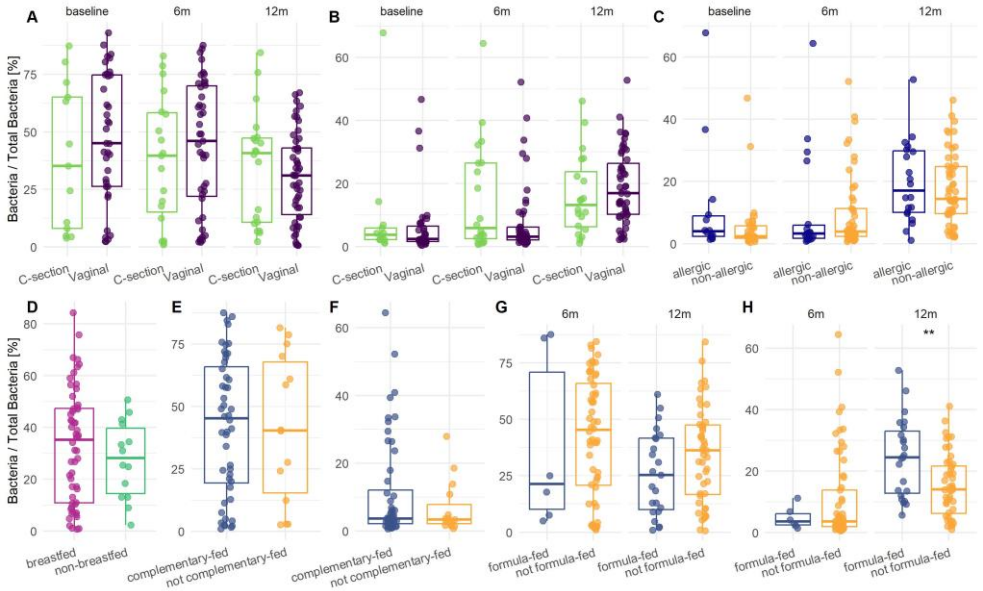

Figure S4. *FISH quantified levels of* ***A)*** *Bifidobacterium spp. between vaginal (purple) and C-section (green) deliveries;* ***B)*** *ER/CC between vaginal (purple) and C-section (green) deliveries;* ***C)*** *ER/CC between allergic (blue) and non-allergic (orange) infants);* ***D)*** *Bifidobacterium spp. between breastfed (pink) and non-breasted (green) infants at 12m;* ***E)*** *Bifidobacterium spp. between complementary-fed (blue) and not complementary-fed (orange) infants at 6m;* ***F)*** *ER/CC between complementary-fed (blue) and not complementary-fed (orange) infants at 6m;* ***G)*** *Bifidobacterium spp. between formula-fed (blue) and not formula-fed (orange) infants;* ***H)*** *ER/CC between formula-fed (blue) and not formula-fed (orange) infants. Statistical analysis was performed using Mann-Whitney test: P < 0.05 (*), P < 0.01 (**), P < 0.001 (***), P < 0.0001 (****). Number of measurements per group and visit: Bifidobacterium spp.: n = [50, 62, 70]; ER/CC: n = [48, 60, 71] for baseline, 6m and 12m, respectively.*

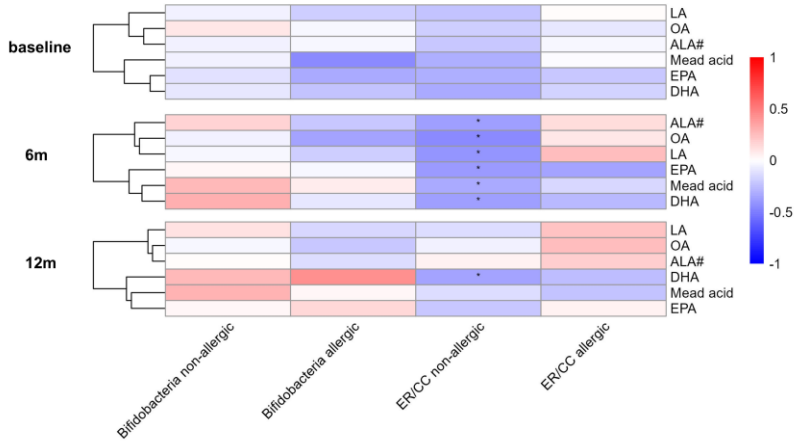

Figure S5. *Spearman correlation between the long-chain fatty acids found to be significantly higher at baseline in the allergic infants and Bifidobacterium spp. and ER/CC at baseline, 6m, 12m, Q < 0.1 (*). The "#" in the metabolite names indicates that the metabolite coeluted with another target metabolite. All abbreviations and coeluting metabolites can be found in* ***Table S1****.*