Taylor & Francis
Taylor & Francis Group

RESEARCH PAPER

OPEN ACCESS | Check for updates

A compromised developmental trajectory of the infant gut microbiome and metabolome in atopic eczema

Le Duc Huy Ta[a]*, James Chun Yip Chan [b,c]*, Gaik Chin Yap[a], Rikky W. Purbojati [d], Daniela I. Drautz-Moses [d], Yanqing Michelle Koh[d], Carina Jing Xuan Tay[a], Chiung-Hui Huang[a], Dorinda Yan Qin Kioh[e], Jia Yun Woon[e], Elizabeth Huiwen Tham [a,f], Evelyn Xiu Ling Loo[g], Lynette P.C. Shek[a,f], Neerja Karnani[g], Anne Goh[h], Hugo P. S. Van Bever[a,f], Oon Hoe Teoh [h], Yiong Huak Chan[i], Christophe Lay[a,j], Jan Knol[k,l], Fabian Yap[h], Kok Hian Tan[h], Yap-Seng Chong[g,m], Keith M. Godfrey[n], Staffan Kjelleberg[d], Stephan C. Schuster[d], Eric Chun Yong Chan[c,e,g], and Bee Wah Lee[a]

[a]Department of Paediatrics, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore; [b]Skin Research Institute of Singapore, A*STAR, Singapore; [c]Innovations in Food & Chemical Safety Programme, A*STAR; [d]Singapore Centre For Environmental Life Sciences Engineering (SCELSE), Nanyang Technological University, Singapore; [e]Department of Pharmacy, Faculty of Science, National University of Singapore, Singapore; [f]Khoo Teck Puat-National University Children's Medical Institute, National University Health System, Singapore; [g]Singapore Institute for Clinical Sciences (SICS), Agency for Science, Technology and Research (A*STAR), Singapore, Singapore; [h]Department of Paediatrics, KK Women's and Children's Hospital, Singapore; [i]Biostatistics Unit, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore; [j]Danone Nutricia Research, Singapore; [k]Danone Nutricia Research, Utrecht, The Netherlands; [l]Laboratory of Microbiology, Wageningen University, Wageningen, The Netherlands; [m]Department of Obstetrics & Gynaecology, National University of Singapore, Singapore; [n]MRC Lifecourse Epidemiology Unit and NIHR Southampton Biomedical Research Centre, University of Southampton and University Hospital Southampton NHS Foundation Trust, Southampton, UK

ABSTRACT

Evidence is accumulating that the establishment of the gut microbiome in early life influences the development of atopic eczema. In this longitudinal study, we used integrated multi-omics analyses to infer functional mechanisms by which the microbiome modulates atopic eczema risk. We measured the functionality of the gut microbiome and metabolome of 63 infants between ages 3 weeks and 12 months with well-defined eczema cases and controls in a sub-cohort from the Growing Up in Singapore Toward healthy Outcomes (GUSTO) mother-offspring cohort. At 3 weeks, the microbiome and metabolome of allergen-sensitized atopic eczema infants were characterized by an enrichment of *Escherichia coli* and *Klebsiella pneumoniae*, associated with increased stool D-glucose concentration and increased gene expression of associated virulence factors. A delayed colonization by beneficial *Bacteroides fragilis* and subsequent delayed accumulation of butyrate and propionate producers after 3 months was also observed. Here, we describe an aberrant developmental trajectory of the gut microbiome and stool metabolome in allergen sensitized atopic eczema infants.

The infographic describes an impaired developmental trajectory of the gut microbiome and metabolome in allergen-sensitized atopic eczema (AE) infants and infer its contribution in modulating allergy risk in the Singaporean mother-offspring GUSTO cohort. The key microbial signature of AE is characterized by (1) an enrichment of *Escherichia coli* and *Klebsiella pneumoniae* which are associated with accumulation of pre-glycolysis intermediates (D-glucose) via the trehalose metabolic pathway, increased gene expression of associated virulence factors (invasin, adhesin, flagellin and lipopolysaccharides) by utilizing ATP from oxidative phosphorylation and delayed production of butyrate and propionate, (2) depletion of *Bacteroides fragilis* which resulted in lower expression of immunostimulatory bacterial cell envelope structure and folate (vitamin B9) biosynthesis pathway, and (3) accompanied depletion of bacterial groups with the ability to derive butyrate and propionate through direct or indirect pathways which collectively resulted in reduced glycolysis, butyrate and propionate biosynthesis.

CONTACT Bee Wah Lee ✉ paeleebw@nus.edu.sg; Eric Chun Yong Chan ✉ eric.chan@nus.edu.sg 🖅 Department of Paediatrics, National University of Singapore (NUS) Tahir Foundation Building (MD1), Level 15, 12 Science Drive 2, Singapore, Singapore 117549
*These authors contributed equally to this work.
ⓘ Supplemental data for this article can be accessed on the publisher's website.
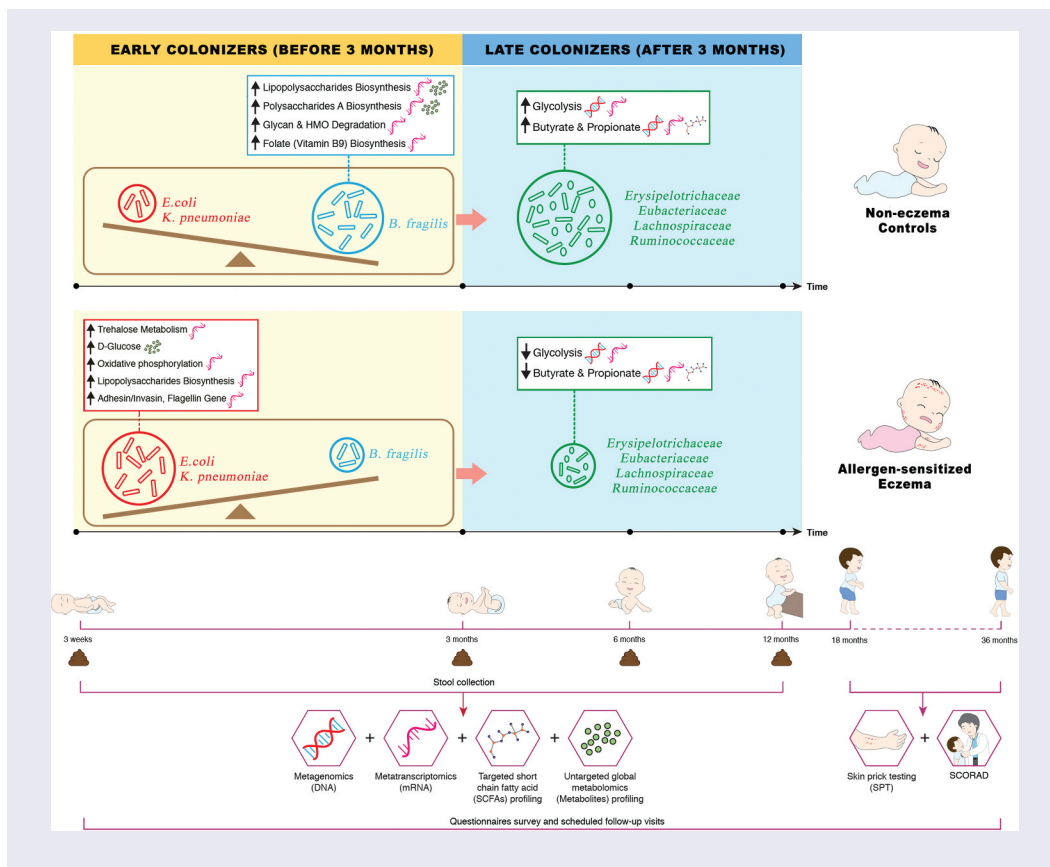
Introduction

Atopic eczema, a chronic inflammatory skin disorder, is one of the most common childhood non-communicable diseases, affecting up to 20% of children and is one of the earliest manifestations of the atopic march.[1] The pathogenesis of eczema has been attributed to skin barrier dysfunctions, immune dysregulation as well as environmental-host-microbial interactions.[2]

Environmental factors and modern lifestyle trends have been pinpointed to indirectly contribute to allergy risk through modulation of the gut microbiome (collective genomes and gene products of the microbiota within the human gut).[3,4,5,6] Evidence is accumulating that the stool metabolome (global collection of small-molecule metabolites found within a stool sample) provides a functional readout of microbial activity and has direct influence on the systemic metabolome and the host through epigenetic modifications as a possible mechanism.[7] Short-chain fatty acids (SCFAs), a source of energy for colonocytes, liver and muscle cells,[8] are the major metabolic end products of anaerobic bacterial fermentation in the gut. SCFAs regulate immune function and modulate host metabolism through specific receptor-mediated regulatory pathways.

In support, several birth cohort studies from varied populations in Europe, Asia, and Canada have demonstrated that lower stool SCFAs levels or butyrate producers in infants influence the development of atopic diseases. Roduit et al. demonstrated in the European PASTURE cohort that infants with the highest levels of stool butyrate and propionate (≥95th percentile) at 1 year of age had significantly less atopic sensitization, diagnosed food allergy, allergic rhinitis and asthma between 3 and 6 years of age.[9] In the mainly Asian PATCH cohort, infants with eczema by 18 months displayed a decreased acquisition of lactate-utilizing butyrate producers (*Eubacterium* and *Anaerostipes*) manifesting with increased stool lactate and decreased butyrate levels in the first 26 weeks of life.[10] In a Korean case study, Song et al. found that stool samples from patients with eczema had decreased levels of butyrate and propionate, associated with reduced number of butyrate and propionate

producer – *Faecalibacterium prausnitzii*.[11] The Canadian CHILD birth cohort showed that lower stool acetate levels in infants was associated with childhood atopic wheeze.[12] In a subsequent publication, this cohort also demonstrated a low butyrate production accompanied by reduced expression of microbial genes encoding key enzymes for butyrate biosynthesis in 3-month old infants with allergen sensitization.[13] These reports suggested the lower production of SCFAs in allergic infants is partly due to the depletion of their corresponding SCFAs producers and this signature in allergic infants is common across different parts of the world.

Using multi-omics approach (metagenomics, metatranscriptomics and metabolomics), we leveraged on a clinically well-characterized mother-offspring cohort, to monitor and compare the gut microbiome and metabolome of infants with early onset eczema (up till 36 months of age) and controls in the first year of life. Here, we hypothesize that infants with atopic eczema and concomitant allergen sensitization present an altered gut microbiome and metabolome in the first 3 weeks of life that led to a subsequent depletion of butyrate and propionate producers. These early events were characterized by an enrichment of *Escherichia coli* and *Klebsiella pneumoniae* accompanied by increased gene expression of virulence factors (invasin, adhesin, flagellin and lipopolysaccharides). These observations were accompanied by a delayed colonization by beneficial *Bacteroides fragilis*, and subsequent depletion of butyrate and propionate producers at ages 6 and 12 months. For the first time, our findings illuminates comprehensive multi-omics insights by unraveling the putative microbial pathways underlying the susceptibility to atopic eczema in early life.

Results

Participant characteristics

The overall demographic, allergy risk factors and clinical characteristics of the 3 clinical groups (allergen sensitized atopic eczema, AE; non-allergen sensitized atopic eczema, NAE; and non-eczema controls) are summarized in Table 1. Only the parental history of eczema differed between groups. The mean age of onset of eczema symptoms in the AE group (3.7 months, range 1 to

18 months) was earlier than that of NAE group (6.4 months, range 1.5 to 18 months). AE participants had a higher mean SCORAD compared to NAE participants at 18 months (23.7 ± 14.0 vs 15.6 ± 5.0) and 36 months (30.1 ± 21.5 vs 17.2 ± 6.9) although this did not reach statistical significance. In terms of allergic comorbidities, food allergy was observed only in the AE participants ($p = .024$). The AE group also had more wheeze and rhinitis compared to controls ($p < .05$) (Table 2). Taking into account of missing samples at different timepoints, there were no significant differences in demographic and clinical confounders between groups for the remaining samples (data not shown).

In this study, we used a "reverse engineering" approach to deconvolute the multi-omics dataset by first delving into the metabolome between the clinical groups. Subsequently, the role of other dominant bacterial species at the early timepoints were then investigated to deduce the potential underlying mechanism involved in the pathogenesis of atopic eczema. The aim was to investigate and unravel the putative microbiome and metabolome risk phenotype of atopic eczema.

Depletion of SCFAs (acetate/propionate/butyrate), butyrate/propionate-producing bacterial families/ species and related functional gene expression in AE group

Of the nine varieties of stool SCFAs analyzed, acetate, propionate and butyrate were the most abundant SCFA among all clinical groups across all four timepoints. They accounted for about 98% of the total SCFAs analyzed (data not shown). Compared to controls, lower absolute concentrations of acetate, butyrate and propionate levels were observed in AE (adj $p < .05$), but not in NAE subjects at all time points. Taking into account of longitudinal mean difference, only the reduction in butyrate and propionate levels was significantly associated with AE compared to controls after adjustment for confounders (adj $p < .05$) (Figure 1). Longitudinal analysis of other SCFAs did not show significant differences between groups (Supplementary Figure 1).

Based on Spearman correlation and functional gene carriage as described in methods, five major butyrate- and propionate-producing families and their 13

Table 1. Demographic and clinical profiles of the three clinical groups.

	Control		All Eczema (Non Allergen Sensitized Atopic Eczema & Allergen Sensitized Atopic Eczema)		Non Allergen Sensitized Atopic Eczema (NAE)		Allergen Sensitized Atopic Eczema (AE)	
	N = 30		N = 33		N = 14		N = 19	
	n	%	n	%	n	%	n	%
DEMOGRAPHIC								
Male gender	12	(40.00)	18	(54.55)	6	(42.86)	12	(63.16)
Presence of sibling	8	(26.67)	14	(42.42)	8	(57.14)	6	(31.58)
Cesarean delivery	10	(33.33)	13	(39.39)	4	(28.57)	9	(47.37)
Pre-term gestation (<37 week)	0	(0.00)	3	(9.09)	1	(7.14)	2	(10.53)
Maternal antibiotics during pregnancy	8	(26.67)	5	(15.15)	2	(14.29)	3	(15.79)
Maternal antibiotics during labor	8	(27.59)	10	(30.30)	7	(50.00)	3	(15.79)
Postnatal antibiotics within first year	6	(20.00)	8	(24.24)	3	(21.43)	5	(26.32)
Ethnicity								
Chinese	13	(43.33)	20	(60.61)	8	(57.14)	12	(63.16)
Malay	9	(30.00)	10	(30.30)	5	(35.71)	5	(26.32)
Indian	8	(26.67)	3	(9.09)	1	(7.14)	2	(10.53)
Infant care attendance within first year	1	(5.88)	3	(9.09)	2	(22.22)	1	(7.14)
Feeding history in 1st 6 months								
Exclusive breastfeeding	2	(6.67)	3	(9.09)	0	(0.00)	3	(15.79)
Predominantly breastfeeding	1	(3.33)	8	(24.24)	3	(21.43)	5	(26.32)
Partial breastfeeding	11	(36.67)	9	(27.27)	2	(14.29)	7	(36.84)
Exclusively formula	16	(53.33)	13	(29.29)	9	(64.29)	4	(21.05)
CLINICAL								
Maternal history								
Any of atopic diseases	4	(13.33)	13	(39.39)*	7	(50.00)*	6	(31.58)
Rhinitis	2	(6.67)	5	(15.15)	3	(21.43)	2	(10.53)
Eczema	1	(3.33)	9	(27.27)*	4	(28.57)*	5	(26.32)*
Asthma	4	(13.33)	5	(15.20)	2	(14.29)	3	(15.79)
Paternal History								
Any of atopic diseases	6	(20.00)	11	(33.33)	3	(21.43)	8	(42.11)
Rhinitis	1	(3.33)	6	(18.18)	3	(21.43)	3	(15.79)
Eczema	1	(3.33)	5	(15.15)	0	(0.00)	5	(26.32)*
Asthma	4	(13.33)	4	(12.12)	1	(7.14)	3	(15.79)

*Significant at $P < .05$ compared with the healthy group in univariate analysis.

**Table 2.** Eczema characteristics of the three clinical groups.

	Control		All Eczema (Non Allergen Sensitized Atopic Eczema & Allergen Sensitized Atopic Eczema)		Non Allergen Sensitized Atopic Eczema (NAE)		Allergen Sensitized Atopic Eczema (AE)	
	N = 30		N = 33		N = 14		N = 19	
	n	%	n	%	n	%	n	%
Age of onset of eczema								
Before 3rd month			12	(40.00)	3	(27.27)	9	(47.37)
Between 3rd and 6th month			11	(36.67)	3	(27.27)	8	(42.11)
Between 6th and 12th month			4	(13.33)	3	(27.27)	1	(5.26)
Between 12th and 18th month			3	(10.00)	2	(18.18)	1	(5.26)
Mean ± SD (Range)			4.7 ± 4.6 (1.0–18.0)		6.4 ± 5.6 (1.5–18.0)		3.7 ± 3.7 (1.0–18.0)	
SCORAD Index Score†								
At 18 Months – Mean ± SD (Range)			19.7 ± 11.0 (7.1–46.1)		15.6 ± 5.0 (10.7–20.3)		23.7 ± 14.0 (7.1–46.1)	
At 36 Months – Mean ± SD (Range)			25.0 ± 17.8 (7.1–69.4)		17.2 ± 6.9 (7.1–22.7)		30.1 ± 21.5 (12.8–69.4)	
Other allergic co-morbiditides								
Wheezing	0	(0.00)	6	(18.75)*	2	(14.29)	4	(22.22)*
Rhinitis	3	(10.00)	20	(60.61)*	9	(64.29)*	11	(57.89)*
Food allergy	0	(0.00)	6	(18.75)*	0	(0.00)	6	(33.33)*

†SCORing of Atopic Dermatitis severity assessments (SCORAD) were performed by trained physicians at 18 and 36 months for evaluation of eczema severity.
*Significant at $P < .05$ compared with the healthy group in univariate analysis.

corresponding species were shortlisted and analyzed according to their abundance (read counts) (Figure 2). They included the following: (1) "Early colonizers" from week 3 till month 3 – *Bacteroidaceae* and (2) "Late colonizers" after month 3 – *Erysipelotrichaceae, Eubacteriaceae, Lachnospiraceae* and *Ruminoco*
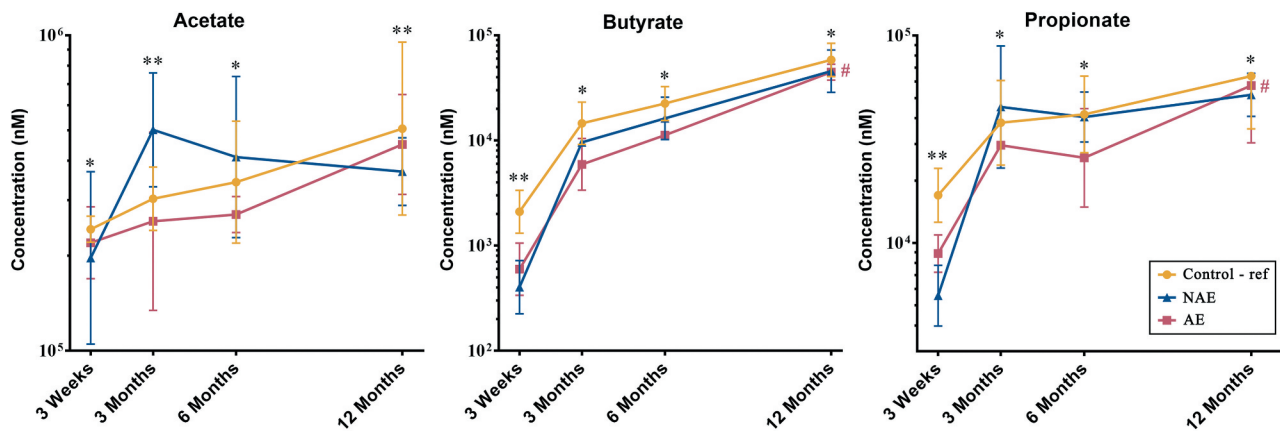
Figure 1. Longitudinal comparison of 3 major stool SCFAs by 3 clinical outcomes. Data are presented as geometric mean and geometric standard deviation range of absolute concentration (nM) in log-scale. Linear mixed-model and general linear model were used to assess difference of abundance between the eczema (AE/NAE) and control (reference) groups adjusting for confounders (gender, birth order, mode of delivery, breastfeeding till 6 months, antibiotics at labor and family of atopic history). List of SCFAs and the comparison of abundance between eczema (NAE or AE) and control (reference) at individual timepoints and longitudinal mean differences are shown in Supplementary Table 6a. *Significant difference at adj $P < .05$ between AE only and control group at specific time point. **Significant difference at adj $P < .05$ between both AE and NAE compared to control groups at specific time point. # (red) annotates significant longitudinal difference between AE and control at adj $P < .05$. Sample size at 3 weeks – Control (n = 13) vs NAE (n = 5) vs AE (n = 5); 3 months – Control (n = 16) vs NAE (n = 11) vs AE (n = 10); 6 months – Control (n = 27) vs NAE (n = 9) vs AE (n = 14) and 12 months – Control (n = 26) vs NAE (n = 8) vs AE (n = 18).

ccaceae (Supplementary Table 1a). Between ages 3 week and 12 months, normal succession of butyrate and propionate producers belonging to *Bacteroidaceae, Eubacteriaceae, Lachnospiraceae* and *Ruminococcaceae* families were disrupted in both AE and NAE compared to controls (adj $p < .05$) except for *Erysipelotrichaceae* of which only longitudinal difference was observed between AE and controls (adj $p < .05$) (Figure 2). The pairwise comparision between clinical groups for the 13 dominant bacterial species belonging to these 5 families were shown in Supplementary Table 1b. Of note, longitudinal difference between AE and controls for *Bacteroides fragilis, Anaerostipes caccae, Blautia producta, Blautia wexlerae, Lachnospiraceae bacterium, Tyzzerella nexilis, Eubacterium limosum, Eubacterium hallii, Eubacterium ramulus, Faecalibacterium prausnitzii, Ruminococcus gnavus, Ruminococcus sp. JC304* and *Erysipelatoclostridium ramosum* (adj <0.05) was observed. These 13 dominant species were collectively lower in abundance in AE, but not in NAE, compared to controls (adj $p < .05$) (Supplementary Figure 2).

Furthermore, the functional gene expression patterns of all 13 bacterial species contributing to the glycolysis-associated, butyrate and propionate metabolic pathways were then analyzed (Supplementary Table 2). When compared to controls, a lower sum of the abundance read counts of gene carriage (metagenomics) and functional gene transcripts (metatranscriptomics) belonging to these metabolic pathways was observed in AE (adj $p < .05$), but not NAE (Figure 3 and Supplementary Table 3, 4a, 4b). Specific functional genes from 13 species that were differentially expressed (adj $p < .05$) between AE and controls were identified and the pathways of the key genes involved are illustrated in Figure 4a. Several key carbohydrate metabolism genes from the glycolysis-associated, butyrate, acetate and propionate metabolic pathways were significantly reduced in AE (adj $p < .05$) (Supplementary Table 5a).

Accumulation of sugar metabolites from glycolysis-associated pathway and depletion of intermediates in butyrate and propionate biosynthesis pathway in AE infants

Data from untargeted stool metabolite profiling was analyzed to evaluate the intermediate metabolites which promote butyrate and propionate production and determine the specific metabolic pathways
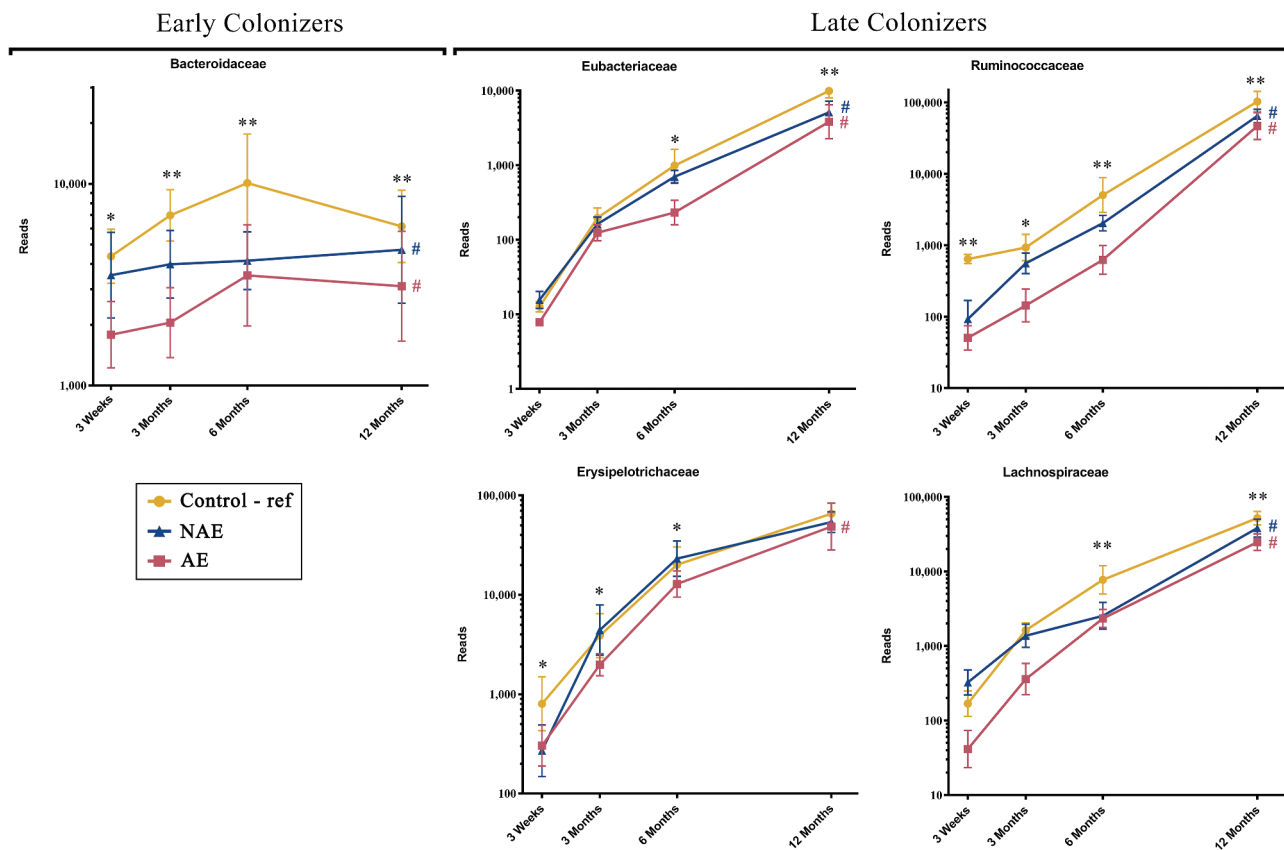
Figure 2. Longitudinal maturation of major butyrate-producing, propionate-producing families in the first year of life. Bacterial community were divided into 2 groups: "Early Colonizers" consisting of major bacterial family and their corresponding predominant species that are dominant from week 3 till month 6: *Bacteroidaceae* (*Bacteroides fragilis*) and "Late Colonizers" consisting of major bacterial families and their corresponding predominant species that are dominant after month 6: *Eubacteriaceae* (*Eubacterium ramulus, Eubacterium limosum, Eubacterium hallii*), *Erysipelotrichaceae* (*Erysipelatoclostridium ramosum*), *Lachnospiraceae* (*Anaerostipes caccae, Blautia wexlerae, Blautia producta, Lachnospiraceae bacterium, Tyzzerella nexilis*) and *Ruminococcaceae* (*Faecalibacterium prausnitzii, Ruminococcus gnavus, Ruminococcus sp. JC304*). Data are presented as geometric mean and geometric standard deviation range of reads in log-scale. Different y-axis scales between graph were used to highlight the difference of microbiota between groups. Linear mixed-model and general linear model were used to assess difference of abundance between the eczema (AE/NAE) and control (reference) groups adjusting for confounders (gender, birth order, mode of delivery, breastfeeding till 6 months, antibiotics at labor and family of atopic history). Comparison of key colonizers at individual timepoints and longitudinal mean differences between eczema (NAE or AE) and control (reference) were listed in Supplementary Table 1b. *Significant difference at adj $p < .05$ between AE only and control group at specific time point. **Significant difference at adj $p < .05$ between both AE and NAE compared to control groups at specific time point. # (blue) annotates significant longitudinal difference between NAE and control at adj $P < .05$. # (red) annotates significant longitudinal difference between AE and control at adj $P < .05$. Sample size at 3 weeks – Control (n = 13) vs NAE (n = 5) vs AE (n = 5); 3 months – Control (n = 16) vs NAE (n = 11) vs AE (n = 10); 6 months – Control (n = 27) vs NAE (n = 9) vs AE (n = 14) and 12 months – Control (n = 26) vs NAE (n = 8) vs AE (n = 18).

responsible for the differences in stool butyrate and propionate levels between AE and controls. The physiologically relevant global metabolites were grouped based on the KEGG database into "glycolysis-associated metabolites", "butyrate intermediates" and "propionate intermediates" as described in methods. Longitudinal analysis showed that compared to controls, AE infants had significantly lower levels of intermediate stool metabolites in the butyrate biosynthesis pathway (predominantly hydroxybutyrate and

dihydroxybutanoate) (adj $p < .05$); propionate biosynthesis pathway (predominantly propanoic acid, 2-oxo-3-(trimethylsilyl)-, trimethylsilyl ester) (adj $p < .05$) and a corresponding increase in accumulation of metabolites belonging to the pre-glycolysis pathway (predominantly D-glucose, D-galactose, lactose and lactate) (adj $p < .05$) in the first 6 months (Figure 5 and Supplementary Table 6b). These data suggest that the production of butyrate and propionate intermediate metabolites is impaired in AE infants.
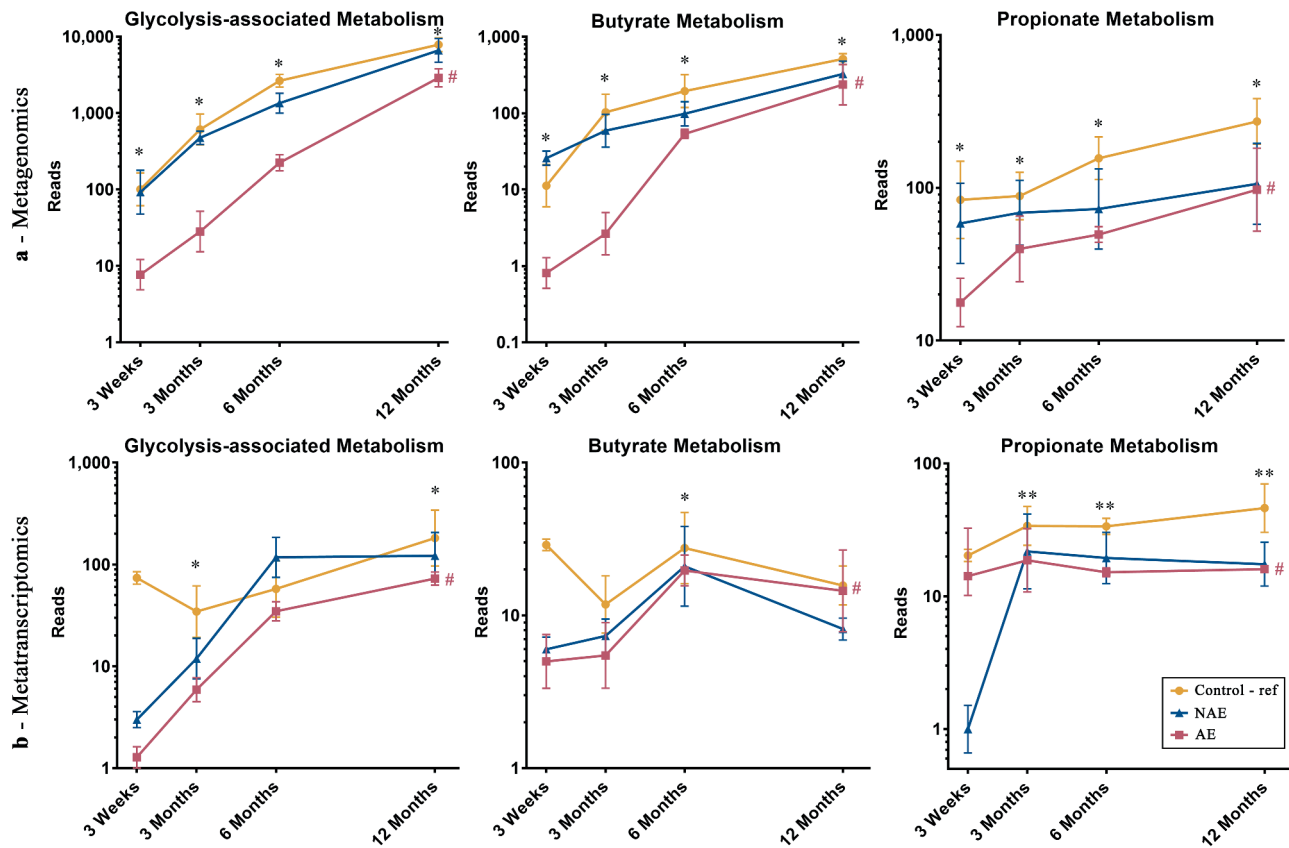
Figure 3. Longitudinal distribution of functional genes belonging to glycolysis, butyrate and propionate biosynthesis pathway of major butyrate-producing and propionate-producing bacterial species from (a) metagenomics and (b) metatranscriptomics. All relevant functional genes carried by shortlisted species were grouped into "Glycolysis-associated Metabolism", "Butyrate Metabolism" and "Propionate metabolism" as described in methods (list of functional genes in Supplementary Table 5). The cumulative read numbers of all genes belonging to the same metabolism pathway were then pooled and plotted as a line plot. Data are presented as geometric mean and geometric standard deviation range of cumulative read number in log-scale. Different y-axis scales between graphs were used to highlight the difference of read counts between groups. Linear mixed-model and general linear model were used to assess difference of abundance between the eczema (AE/NAE) and control (reference) groups adjusting for confounders (gender, birth order, mode of delivery, breastfeeding till 6 months, antibiotics at labor and family of atopic history). Abundance data at individual timepoints are listed in Supplementary Table 4a and 4b. *Significant difference at $P < .05$ between AE only and control group at specific time point. **Significant difference at $P < .05$ between both AE and NAE compared to control groups at specific time point. # (red) annotates significant longitudinal difference between AE and control at adj $P < .05$. Sample size for (a) at 3 weeks – Control (n = 13) vs NAE (n = 5) vs AE (n = 5); 3 months – Control (n = 16) vs NAE (n = 11) vs AE (n = 10); 6 months – Control (n = 27) vs NAE (n = 9) vs AE (n = 14) and 12 months – Control (n = 26) vs NAE (n = 8) vs AE (n = 18). Sample size for (b) at 3 weeks – Control (n = 6) vs NAE (n = 2) vs AE (n = 2); 3 months – Control (n = 11) vs NAE (n = 3) vs AE (n = 7); 6 months – Control (n = 17) vs NAE (n = 6) vs AE (n = 11) and 12 months – Control (n = 13) vs NAE (n = 3) vs AE (n = 11).

Consequently, intermediate metabolites in the pre-glycolysis pathway accumulated in AE, but not NAE subjects, when compared against controls.

Early deviation in the developmental trajectory of the gut microbiome of infants with atopic eczema – *Enterobacteriaceae/Bacteroidaceae* ratio

We next explored if other microbial groups contributed to the microbial signature associated with AE.

Unsupervised analysis of the microbiome and meta-bolome (SCFAs) datasets as a whole indicated that the maturation of the microbiome/metabolome was driven by age from 3 weeks toward 12 months (Supplementary Figure 3). The Shannon species diversity in all three groups increased over time, though there were no significant differences observed between clinical groups (Supplementary Figure 4). At the family level, we observed that members of *Bifidobacteriaceae, Enterobacteriaceae,*
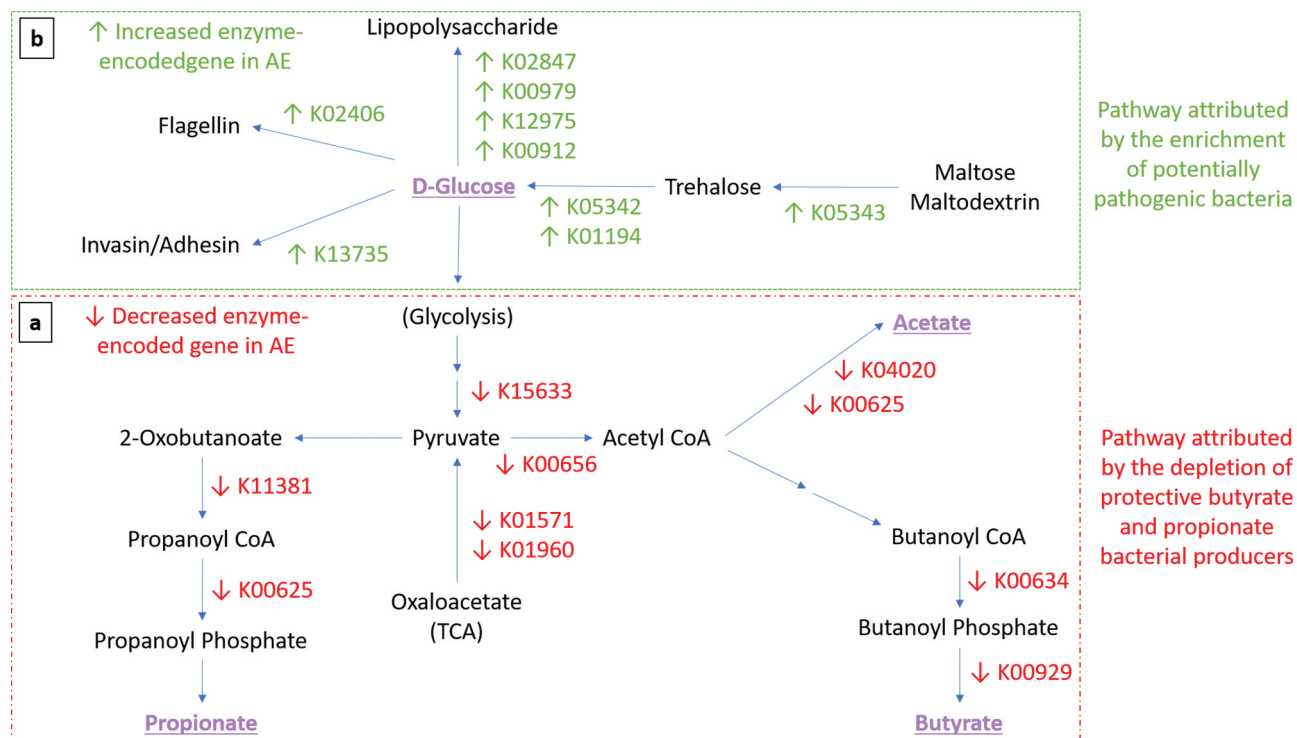
Figure 4. Expression of functional gene profile for carbohydrates metabolism in AE group. Arrows indicate significant differential expression of functional gene in AE compared to control group ($p < .05$). (a-Red Box) Perturbation of carbohydrate metabolism attributed by the depletion of butyrate- and propionate-producing bacteria. Several key carbohydrate metabolism genes were significantly reduced in AE (adj $p < .05$) from the glycolysis pathway (K15633 2,3-bisphosphoglycerate-independent phosphoglycerate mutase, K01571 oxaloacetate decarboxylase and K01960 pyruvate carboxylase subunit B); butyrate biosynthesis pathway (K00634 phosphate butyryltransferase, K00929 butyrate kinase); acetate biosynthesis pathway (K04020 phosphotransacetylase, K00625 phosphate acetyltransferase) and propionate biosynthesis pathway (K01734 methylglyoxal synthase, K00005 glycerol dehydrogenase, K01699 propanediol dehydratase, K13922 propionaldehyde dehydrogenase, K11381 2-oxoisovalerate dehydrogenase, K00625 phosphate acetyltransferase). (b-Green Box) Pathway attributed by the enrichment of potentially pathogenic bacteria. In contrast, several key genes were found to be enriched in AE compared to controls (adj $p < .05$) in the trehalose metabolism pathway (K05343 maltose alpha-D-glucosyltransferase/alpha-amylase, K05342 alpha,alpha-trehalose phosphorylase and K01194 alpha,alpha-trehalase); lipopolysaccharide (LPS) biosynthesis (K02847 O-antigen ligase, K00979 3-deoxy-manno-octulosonate cytidylyltransferase, K12975 KDO II ethanol-amine-phosphotransferase and K00912 tetra-acyl-disaccharide 4'-kinase) and virulence components biosynthesis pathways (K13735 adhesin/invasin, K02406 flagellin). List of metatranscriptomics functional annotation that are significantly different between AE and controls is shown in Supplementary Table 5.

Bacteroidaceae and *Lachnospiraceae* were predominant across the clinical groups (Supplementary Figure 5).

Pairwise group comparisons indicated that there was a higher proportion of *Enterobacteriaceae* and a lower abundance of *Bacteroidaceae* in AE compared to controls at 3 weeks and 3 months (adj $p < .05$). *Lachnospiraceae* was only significantly higher in controls than AE at 6, 12 months (adj $p < .05$) (Supplementary Table 1a), whilst there was no statistically significant difference in *Bifidobacteriaceae* between groups (data not shown). We demonstrated in AE, an inverse relationship between *Bacteroidaceae* (*B. fragilis*) and *Enterobacteriaceae* (*E. coli* and *K. pneumoniae*) (Figure 6), manifesting as high ratios

of *Enterobacteriaceae*/*Bacteroidaceae*, *E. coli*/ *B. fragilis* and *K. pneumoniae*/*B. fragilis* that peaked at week 3 and gradually decreased almost to zero at month 12 (Supplementary Figure 6).

Enrichment of *Escherichia coli*, *Klebsiella pneumoniae* and increased carbohydrate metabolism and virulence gene expression in AE

Compared to controls, AE presented an enrichment of *E. coli* and *K. pneumoniae* in the first 3 weeks of life (adj $p < .05$) (Figure 6). Of these, *E. coli* and *K. pneumoniae* predominated before 6 months of age. Compared to controls, pairwise comparison of carbohydrate gene profiles from
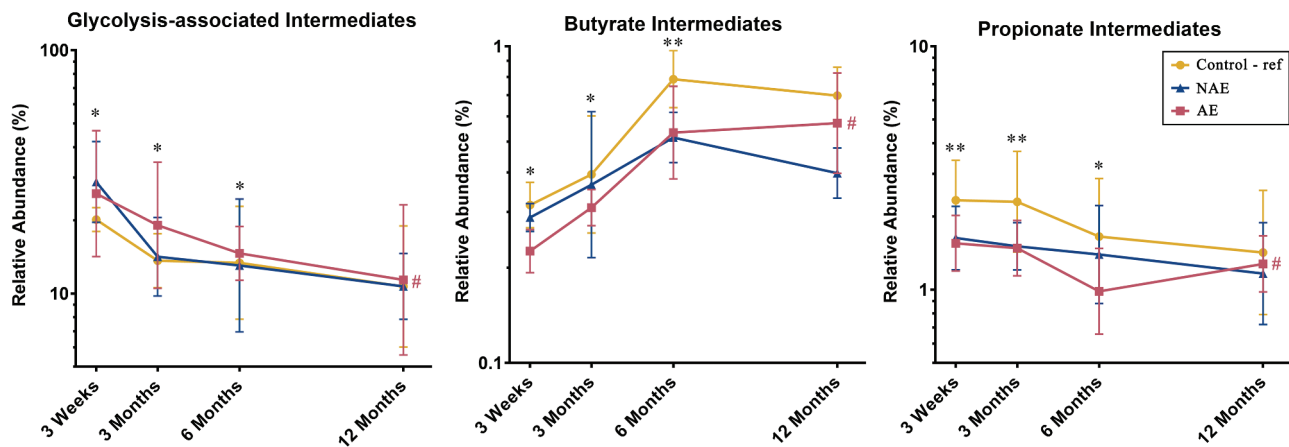
Figure 5. Distribution of intermediate metabolites belonging to the glycolysis-associated pathway, butyrate and propionate biosynthesis pathways (list of intermediate metabolites in Supplementary Table 10). All physiologically and metabolically relevant metabolites were grouped into "Glycolysis-associated intermediates", "Butyrate intermediates" or "Propionate intermediates metabolites" based on the KEGG database. Data are presented as geometric mean and geometric standard deviation range of relative abundance (%) in log-scale. Linear mixed-model and general linear model were used to assess difference of abundance between the eczema (AE/NAE) and control (reference) groups adjusting for confounders (gender, birth order, mode of delivery, breastfeeding till 6 months, antibiotics at labor and family of atopic history). List of predominant metabolites and the comparsion of abundance between eczema (NAE or AE) and control (reference) at individual timepoints and longitudinal mean differences are shown in Supplementary Table 6b. *Significant difference at adj $P < .05$ between AE only and control group at specific time point. **Significant difference at adj $P < .05$ between both AE and NAE compared to control groups at specific time point. # (red) annotates significant longitudinal difference between AE and control at adj $P < .05$. Sample size at 3 weeks – Control (n = 13) vs NAE (n = 5) vs AE (n = 5); 3 months – Control (n = 16) vs NAE (n = 11) vs AE (n = 10); 6 months – Control (n = 27) vs NAE (n = 9) vs AE (n = 14) and 12 months – Control (n = 26) vs NAE (n = 8) vs AE (n = 18).

E. coli and *K. pneumoniae* in AE revealed an increase in several key genes in the trehalose metabolic pathway, which yields from hydrolyzed trehalose D-glucose as substrates for glycolysis (Figure 4b and Supplementary Table 5b). These results are in line with the observed accumulation of pre-glycolysis intermediates such as D-glucose in AE subjects (Figure 5).

In addition, the expression of genes responsible for bacterial LPS, virulence components biosynthesis (adhesin/invasion and flagellin) and oxidative phosphorylation were also increased in AE (Supplementary Table 5b). Further correlation analysis revealed that the increase in D-glucose positively correlated with the increased expression of several virulence genes coding for adhesin/invasin (R = 0.761, $p < .05$), flagellin (R = 0.752, $p < .05$) and LPS biosynthesis (R = 0.805, $p < .05$) in both bacterial species.

Bacteroides fragilis, an early life colonizer depleted in AE infants

In comparison to AE, control infants presented an enrichment of *B. fragilis* from first 3 weeks of life

onwards (Figure 6). Transcriptomics-based functional gene analysis of *B. fragilis* revealed that the depletion of *B. fragilis* in AE was associated with lower gene read counts involving glycan degradation, lipopolysaccharides (LPS) biosynthesis, polysaccharide A (PSA) capsule biosynthesis, membrane transport and folate biosynthesis (adj $p < .05$) (Supplementary Table 5). Further correlation analysis between transcriptomic functional gene and metabolome showed a strong positive correlation between the abundance of *B. fragilis* and LPS fatty acid components namely pentadecanoic acid (R = 0.673, $p < .05$).

Sensitivity Analysis

In lieu of the missing stool data in this cohort at various timepoints, linear mixed modeling was repeatedly performed under two extreme scenarios as described in methods for the two key SCFAs (butyrate and propionate) and three key bacterial species (*B. fragilis, E. coli* and *K. pneumoniae*). For the "best" scenarios, the longitudinal trend of all variables between AE and controls remained significant (adj $p < .05$). For the "worst" scenario, all

Early Colonizers
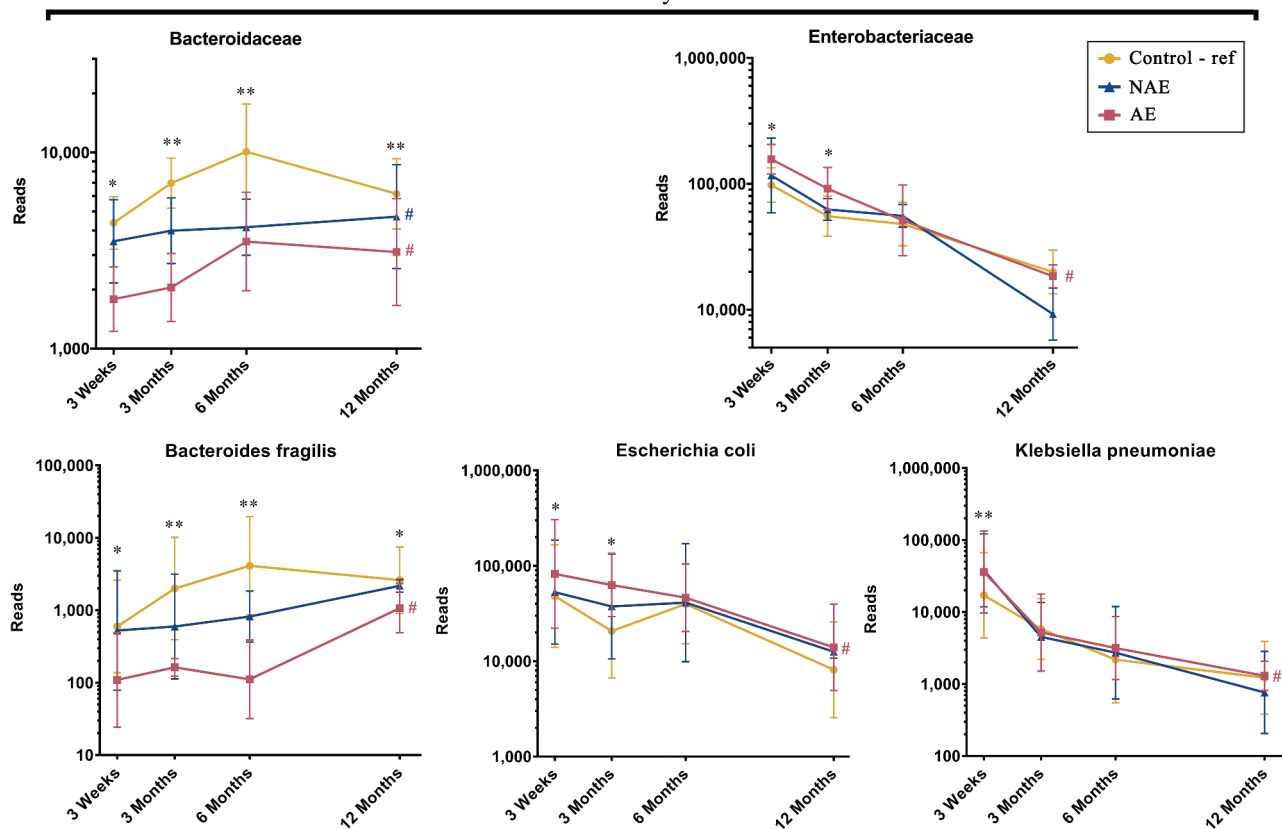
Figure 6. Longitudinal maturation of three early colonizers namely *Bacteroides fragilis* (*Bacteroidaceae* family), *Escherichia coli* (*Enterobacteriaceae* family) and *Klebsiella pneumoniae* (*Enterobacteriaceae* family). Data are presented as geometric mean and geometric standard deviation range of reads in log-scale. Different y-axis scales between graph were used to highlight the difference of microbiota between groups. Linear mixed-model and general linear model were used to assess difference of abundance between the eczema (AE/NAE) and control (reference) groups adjusting for confounders (gender, birth order, mode of delivery, breastfeeding till 6 months, antibiotics at labor and family of atopic history). Comparison of key colonizers at individual timepoints and longitudinal mean differences between eczema (NAE or AE) and control (reference) were listed in Supplementary Table 1b. *Significant difference at adj *p* < .05 between AE only and control group at specific time point. **Significant difference at adj *p* < .05 between both AE and NAE compared to control groups at specific time point. # (blue) annotates significant longitudinal difference between NAE and control at adj *P* < .05. # (red) annotates significant longitudinal difference between AE and control at adj *P* < .05. Sample size at 3 weeks – Control (n = 13) vs NAE (n = 5) vs AE (n = 5); 3 months – Control (n = 16) vs NAE (n = 11) vs AE (n = 10); 6 months – Control (n = 27) vs NAE (n = 9) vs AE (n = 14) and 12 months – Control (n = 26) vs NAE (n = 8) vs AE (n = 18).

variables were significant (adj *p* < .05) except for propionate concentration (adj *p* = .062).

Discussion

This study leveraged an integrated, multi-omics approach to deconvolute the gut microbiome-metabolome cross-talk in infants with atopic eczema in the first year of life. The longitudinal microbial analysis uncovered a pattern of successive microbial colonization at a taxonomic and functional level in atopic eczema. Most of the differences observed in microbial composition, functional gene expression and metabolomes between clinical groups commenced as early as 3 weeks, which in most cases preceded the onset of eczema, suggesting that the early life gut microbiome is likely to influence the risk of developing AE. The abnormalities in the stool microbiome and metabolome were only significant in eczema infants with concomitant allergen sensitization by the age of 36 months (AE), while the data of the non-allergen sensitized atopic eczema group (NAE) was intermediate between AE and controls (Figures 1,2,3,5). We hypothesize that the larger the extent of microbiome dysbiosis, the greater its influence on allergen sensitization and the development of allergic co-morbidities (rhinitis and

wheeze) (Table 1). At the study population level, the maturation of the microbiome and metabolome was driven by age indicating a synchronized process that likely have a role in gut maturation and immunity development.[14] Conversely, a disruption of this process may have a negative impact.

Birth to age 3 months: Over-representation of *E. coli* and *K. pneumoniae* and delayed colonization by *B. fragilis* in AE

In this study, we uncovered an impaired developmental trajectory of the gut microbiome/metabolome in infants with AE. The microbial deviation in AE infants was observed as early as 3 weeks of life and was characterized by a delayed colonization by *Bacteroidaceae family* (*B. fragilis*) with concomitant enrichment of members of the *Enterobacteriaceae* family (*E. coli* and *K. pneumoniae*) (Figure 6), resulting in a high ratio of *Enterobacteriaceae/Bacteroidaceae* in AE compared to controls (Supplementary Figure 6). In support, a high *Enterobacteriaceae/Bacteroidaceae* ratio has been reported in infants with food allergy.[15,16] Contrary to other reports that have found reduced *Bifidobacteriaceae* in infants who develop eczema,[17–19] our study did not find any association of this family at both structural and functional gene level.

We hypothesize that the early colonization by *E. coli* and *K. pneumoniae* in the first weeks of life would subsequently impact the successive establishment of other microbial species, leading to the delayed colonization by butyrate- and propionate-producing bacteria observed in AE. In support, high abundance of *E. coli* was linked to eczema in both longitudinal and cross-sectional studies,[20,21] and early onset (mean age of 2 months) eczema with high total serum IgE levels.[22] A study showed that adherent-invasive *E. coli* colonization alters the intestinal commensal microbiome acquisition in a way that increases its pro-inflammatory potential.[23] At the functional level, we identified multi-omic signatures indicating that in AE infants the presence of *E. coli* and *K. pneumoniae* was associated with the expression of genes involved in the metabolism of trehalose and this was translated to an increased detection of D-glucose in stool samples (Figure 4b). It has

been shown that the accumulation of trehalose from dietary maltodextrin/maltose is an exclusive response to stress conditions of *E. coli* and *K. pneumoniae* before undergoing hydrolysis to D-glucose as an alternative energy source.[24,25] We also observed that the elevations of stool D-glucose was associated with the expression of genes encoding for virulence factors (LPS, adhesin, invasion and flagellin) suggesting that D-glucose could be diverted for the biosynthesis of those virulence factors by utilizing ATP from oxidative phosphorylation. Adhesin, invasin and flagellin have been shown to increase the ability of *E. coli* and *K. pneumoniae* to colonize mucin and epithelial surfaces,[26,27,28] allowing them to compete with *B. fragilis* for colonization in niches which normally favor *B. fragilis* colonization.[29] D-glucose is also a well-known key component of LPS core oligosaccharides in *E. coli* and *K. pneumoniae*.[30,31] Accumulation of stool D-glucose has been reported previously, with stool samples profiled in 3-month old infants with eczema being higher compared to controls.[32] LPS, found in the outer membrane of Gram-negative *Enterobacteriaceae*, is also known to contribute to atopic risk. Binding of LPS to CD14 and TLR4 receptors leads to activation of MyD88-dependent NF-κB and IRF3, culminating in production of pro-inflammatory cytokines.[33,34]

Conversely, in control infants, the presence of *B. fragilis* indicates its protective role against the development of atopic eczema. *B. fragilis* is transmitted vertically from the maternal microbiota, colonizes the infant gut as early as day three and has the ability to consume Human Milk Oligosaccharides (HMOs).[35,36,37,38] Current evidence supports breastfeeding, and when not possible formula supplemented with prebiotics and/or probiotics as a nutritional solution may prevent the primary development of eczema by supporting a healthy gut microbiome.[39,40,41] Studies on the use of prebiotics to specifically support the colonization and enrichment of beneficial *B. fragilis* are needed to demonstrate its potential therapeutic application for eczema prevention. Here, we postulate that the acquisition of *B. fragilis* during birth allows the modulation of a gut environment that would promote the establishment of butyrate and propionate producers. The reduced abundance of *Bacteroides* genus and *B. fragilis* within the first year

of life in infants with AE are consistent with other reports.[18,42] A low abundance of *B. fragilis* in early life has also been associated with Cesarean delivery, a known risk factor for eczema development.[43] Cesarean delivery was, however, not significantly associated with eczema in our study as mode of delivery was similar in the cases and controls [Control: 10 cases (33.33%) vs Eczema: 13 cases (39.39%)] (Table 1). A study on the larger GUSTO cohort (1077 pregnant women) also supports this observation.[44]

Besides members of the *Enterobacteriaceae*, LPS is also a common feature of *B. fragilis*. In support, we observed a positive association between abundance of *B. fragilis*, LPS biosynthesis gene and LPS component (pentadecanoic acid). Several studies have shown differences in LPS immunogenicity between bacterial species, whereby LPS from *Bacteroides* species promote immune tolerance; while those from *E. coli* induce inflammation due to their structurally distinct LPS structures.[45,46] Those data suggest that *Bacteroides* LPS could potentially inhibit the immunostimulatory activity of *E. coli* LPS. Hence, increased *E. coli* and *K. pneumoniae* LPS in early life may thus be pro-inflammatory and hence adversely influence the development of immune tolerance.

Additionally, increased expression of essential genes by *B. fragilis* involved in PSA biosynthesis was associated with controls. Studies reported that *B. fragilis* exert anti-inflammatory properties through its PSA which induces IL-10 and IFN-γ, restoring Th1/Th2 balance in the host.[47] Furthermore, *B. fragilis* carries glycan degradation genes which were shown to be reduced in AE. These enzymes metabolize HMOs consumption due to their structural similarities with glycan.[48] *B. fragilis* metabolizes HMOs to precursor substrates for butyrate and propionate production, which serve as nutrients for colonocytes and other gut epithelial cells.[49] Moreover, other relevant functional genes of *B. fragilis* involved in membrane transport and folate biosynthesis were reduced in AE. Membrane transports, especially ATP-binding cassette (ABC) transporters have previously been shown to be immunogenic.[50] Folate (vitamin B9) is known to be an important nutrient for DNA synthesis, replication, and repair as well as the

maintenance of regulatory T cells.[51,52] Thus, the presence of this early colonizer might be an important determinant for normal immune maturation in early life.

Beyond age 3 months: Delayed colonization by butyrate and propionate producers and consequent reduction of SCFAs production

Between ages 3 and 12 months, the normal succession of butyrate and propionate producers belonging to *Erysipelotrichaceae*, *Eubacteriaceae*, *Lachnospiraceae* and especially *Ruminococcaceae* families were disrupted in AE compared to controls (Figure 2). Other studies substantiate our observations, with *Erysipelatoclostridium* higher in abundance in healthy compared with allergic infants.[53] A previous report observed that eczema severity correlated inversely with the abundance of *Eubacteriaceae*, *Lachnospiraceae* and *Ruminococcaceae* in 6-month-old infants.[54] High-risk infants who developed eczema by 18 months likewise demonstrated delayed acquisition by *Eubacterium* and *Anaerostipes* species and correspondingly reduced butyrate and propionate levels at 26 weeks of age, compared to healthy controls.[10] *F. prausnitzii* (*Ruminococcaceae* family) is a major component of the healthy human microbiome and is depleted in atopic children.[55] The Korean COCOA birth cohort study showed that the stool of six-month-old infants with eczema was depleted of *R. gnavus* (*Ruminococcaceae* family) and *L. bacterium* (*Lachnospiraceae* family).[56]

Consequently, the impact of the reduced abundance of these SFCA producers in AE participants was reflected in the accumulation of pre-glycolysis intermediates as well as lower expression of functional genes and transcripts involved in butyrate and propionate biosynthesis (Figures 1, 5). These strengthen the notion that the observed accumulation of stool intermediates derived from the pre-glycolysis pathway and, low butyrate and propionate levels are related to diminished glycolysis, butyrate and propionate biosynthesis capabilities by corresponding SCFAs producers. These finding corroborate with those of several recent cohort studies.[9,10,12,13] High dietary fiber or live biotherapeutic intervention could regulate protective pathways in the gastrointestinal tract and promote oral

tolerance through the modulation of the gut microbiome.[57] Our findings indicated that the delayed production of SCFA starts to occur around the weaning period. A recent preclinical study related to the weaning reaction, a key period that allows the establishment of the SCFA producers. Failing to orchestrate this microbial programming event led to pathological inflammatory disorders.[58]

The potential limitation of this study is the small sample size at early timepoint (3 weeks) which preclude us from investigating the effect of other early-life confounding factors such as diet or lifestyle factors on these findings. To overcome the possibility of a type 1 error from small sample size, statistical power was increased by bootstrapping method whereas sensitivity test was performed under 2 extreme scenarios to assess the influence of missing samples. The statistical significance for key variables remained unchanged. However, caution should also be exercised in interpreting the findings as they are not necessarily representative of AE in other geographical settings and replication is required. Moreover, as with previous cohort studies,[9–12] SCFAs were quantified based on per mg of wet stool due to the high volatility of the SCFAs. Differential stool consistency between samples may however be a potential confounding factor. It is proposed that stool consistency should be taken into account in the design of future studies. For example, the dry weight of an equal aliquot of each stool sample could be incorporated to standardize measurement. Despite the small cohort size, our study is the first integrated multi-omics study in the field of allergy including atopic eczema. This strategy elucidates the cross-talk between host-microbiota and their metabolites, and therefore provides comprehensive multi-omics insights into unraveling the microbial pathways underlying the pathogenesis of eczema in early life.

In summary, we described key early life gut microbiome and metabolomic signatures that are associated with AE in a Singaporean mother-offspring cohort. Taken together, our data depict a disrupted developmental trajectory of the gut microbiome that commences in the first weeks of life. We identified microbiome and metabolome risk phenotypes for AE that occurs in the first three months of life. A compromised microbiome in the first 100 days of life has been

identified as a risk factor for allergy.[10,12,59] Our findings provide key directions for future mechanistic studies and targets for intervention strategies to optimize the influential role of the maturing gut microbiome on the programming of the immune system for the prevention of eczema in early life.

Methods

Subjects recruitment and sample collection

GUSTO is an ongoing mother-offspring cohort study in Singapore which recruited 1237 pregnant women from the general population irrespective of atopy and followed their offspring prospectively. The primary objective of GUSTO is to identify and evaluate the role of risk factors and determinants influencing body composition and metabolism during early development, and their influence on infants' health in later childhood. The study methodology has been described in detail previously.[60] Standardized questionnaires, including the modified ISAAC questionnaire previously validated in other epidemiological studies for assessment of allergic outcomes,[61,62,63] were administered by trained interviewers at birth, 3 weeks, 3, 6, 9, 12, 15, 18, 24 and 36 months. Information on demographics, family history of allergy, social and lifestyle factors, such as diet and breastfeeding, were collected. Outcomes were classified as absent when the answers for all visits up to 36 months were "no". Family history of allergy was defined as positive if the mother, father or an older sibling ever had atopic eczema, asthma or allergic rhinitis. A physician's diagnosis of eczema in the child was determined by a positive answer to the written question: "Has your child ever been diagnosed with eczema?". SCORAD (SCORing of Atopic Dermatitis (AD) severity) assessments were performed by trained physicians at 18 and 36 months for evaluation of eczema severity.

Allergen sensitization was assessed through skin prick testing (SPT) to aeroallergens (house dust mites *Dermatophagoides pteronyssinus*, *Dermatophagoides farinae*, and *Blomia tropicalis*) and to food allergens (egg, peanut and cow's milk) at the 18- and 36-month visits. These are the most common aeroallergens and food allergens in

sensitized Singaporean children.[64] All skin prick extracts were obtained from Greer Laboratories (Lenoir, NC, USA), except for *B. tropicalis*, which was obtained from our in-house laboratory. *B. tropicalis* extract was prepared as previously described.[65] A wheal of at least 3 mm was defined as a positive SPT and a child was considered as SPT-positive (allergen-sensitized) if any one or more of the individual tests were positive with a positive reaction to histamine (positive control) and negative reaction to saline (negative control).

Other potential allergic co-morbidities (wheezing, rhinitis) were also assessed at 18 and 36 months. Rhinitis was defined as having symptoms of sneezing, runny and/or blocked nose that lasted for at least four weeks in single or multiple episodes (each episode lasting at least two weeks). This definition was based on the Allergic Rhinitis and its Impact on Asthma (ARIA) guidelines.[66] Wheezing was defined as the presence of wheeze symptoms (noisy breathing with a high-pitched, whistling sound heard from the chest, not the mouth) and with the use of nebulizer. Food allergy was defined as a convincing history of an IgE mediated reaction to a food allergen and a positive SPT to the specified food allergen.

A sub-cohort of 63 participants who had clinical data from birth up to the 36-month time-point, and stool samples collected at 3 weeks, 3, 6 and 12 months, were analyzed in this study. Thirty-three participants with physician-diagnosed eczema were further classified into non-allergen sensitized atopic eczema (NAE) (n = 14) and allergen-sensitized atopic eczema (AE) (n = 19). Atopic eczema cases were selected with similar characteristics as 30 non-eczema controls by age, mode of delivery, breastfeeding pattern till 6 months, use of antibiotic at labor and postnatal antibiotics to minimize potential selection bias. (Supplementary Figure 7). Ethics approval was obtained from the Domain Specific Review Board of Singapore, National Healthcare Group and the Centralized Institutional Review Board of SingHealth (DSRB D/09/021 and CIRB 2009/280/D). Informed written consent was obtained from all subjects.

For metagenomics sequencing and metabolomics profiling, stool samples were collected by the parents using sterile feces containers and stored at −20°C at home. For metatranscriptomic

sequencing, stool samples were additionally collected directly into containers with RNAlater Stabilization Solution (Thermo Fisher Scientific, Waltham, Massachusetts) and stored at 4°C at home. Both samples were then transported to the lab in cold chain within 20 hours of sample collection for processing. After processing, sample were stored at −80°C until further analysis. A total of 162 stool samples were collected from subjects of the three clinical groups over four timepoints which were all then subjected to SCFAs and untargetted metabolomic profiling and metagenomics sequencing. Excluding samples with low RNA yield and low read counts, 91 stool samples eventually underwent metatranscriptomic sequencing (technique summary in Supplementary Table 7).

Profiling of Stool SCFAs using LC/MS/MS and Untargeted Stool Metabolite Profiling by GC/TOFMS

For targeted analysis of SCFAs, details of chemicals and reagents sources, sample and calibration standards preparation were performed as reported previously which was shown to able to detect a total of 9 SCFAs including 6 common SCFAs (acetate, propionate, butyrate, isobutyrate, isovalerate and caproate) and 3 uncommon SCFAs (valerate, 2-methylbutyrate and 4-methylvalerate).[67] For global profiling of stool metabolites, samples were randomized then processed according to our in-house protocol. Briefly, lyophilized stool sample (80 mg) was ultrasonicated with 1 mL of ice-cold extraction solvent (methanol:water [8:2]) containing 1 μg/mL d_{27}-myristic acid as an internal standard at 4°C in a bath ultrasonicator (Elma Transsonic 460, Germany) for 30 min, and vortex-mixed for 2 min. The samples were then centrifuged at 18,000 g for 20 min at 4°C, and 0.5 mL of the supernatant was extracted carefully followed by drying at 50°C under a gentle stream of nitrogen gas (Turbovap LV, Caliper Life Sciences, Hopkinton, MA, USA). 100 μL of toluene (kept anhydrous with sodium sulfate) was added to the dry residue and evaporated completely again at 50°C under nitrogen gas to remove traces of water. The dried metabolic extract was then oximated with 50 μL of MOX (20 mg/mL) at 60°C for 2 h. Following centrifugation, 100 μL of MSTFA with 1% TMCS was added and the mixture was

incubated at 60°C for 45 min to form trimethylsilyl (TMS) derivatives. Finally, 100 μL of the TMS derivatives was transferred into a GC vial and subjected to gas chromatography/time-of-flight mass spectrometry (GC/TOFMS) analysis. 50 μL of each stool extract prepared during extraction of lyophilized stool were also pooled to prepare quality control (QC) samples.

Analysis was performed on a Pegasus GC/TOFMS (LecoCorp, St. Joseph, MI, USA) coupled to an Agilent 7890 gas chromatograph (Agilent Technologies, Santa Clara, California, USA). A DB-1 (30 m × 250 μm i.d.) fused silica capillary column (Agilent J&W Scientific, Folsom, CA), with 0.25 μm film thickness, was used with open split interface. Helium was employed as the carrier gas at a constant flow rate of 1.5 mL/min and the injector split ratio was adjusted to 1:10. An injection volume of 1 μL was used. The injector, transfer line and ion source temperatures were maintained at 220, 280 and 250°C, respectively. The oven temperature was programmed at 70°C for 0.2 min, increased at 9°C/min to 270°C where it was sustained for 5 min, and further increased at 40°C/min to 310°C where it was held for 11 min. The mass spectrometer was operated in the electron impact ionization mode at 70 eV. Data were acquired in full scan mode from m/z 40 to 600 with an acquisition rate of 20 spectra per second. To detect retention time shifts and facilitate Kovats retention index (RI) calculation, a standard alkane series mixture (C7-C30) was injected before every batch during sample analysis. This was accompanied by an injection of one pooled QC after every injection of 10 samples.

Gas chromatograms were subjected to baseline correction, smoothing, noise reduction, deconvolution, library matching and area calculation using ChromaTOF software (version 4.41, LecoCorp). Each peak was then manually checked for proper peak alignment and integration. Metabolites were assigned putative identities by comparing both the RI and the mass spectra with those available in the National Institute of Standards and Technology (NIST) library, Fiehn library and/or our in-house spectral library. The peak areas were subjected to quality control-based robust LOESS signal correction.[34] Subsequently, the peak areas were normalized by the lyophilized weight of each stool sample. Thereafter, metabolites with coefficient of variation greater than 30% were removed from the dataset, and total area normalization was performed. List of all global metabolites identified are shown in Supplementary Table 8. All raw metabolomics data are available at Metabolomics Workbench data repository (http://www.metabolomicsworkbench.org),[68] (Project Digital Object Identifier: 10.21228/M8N104; Project ID: PR000983; Study ID: ST001431, ST001432).

Nucleic acid extraction and metagenomics sequencing

Approximately 100–150 mg of stool material from four serial stool samples (collected at week 3, months 3, 6 and 12) were used to extract genomic DNA with the ZR Fecal DNA MicroPrep Kit (Zymo Research, USA) according to the manufacturer's instructions. For each stool sample, a sequencing library was first constructed using Illumina's Truseq Nano DNA Library Preparation Kit (Illumina, San Diego, USA). The samples were sheared on a Covaris E220 to ~450 bp, following the manufacturer's recommendation, and uniquely tagged with one of Illumina's TruSeq HT DNA barcode combinations to enable sample pooling for sequencing.

The finished libraries were quantitated using Invitrogen's Picogreen assay and the average library size was determined on Bioanalyzer 2100, using a DNA 7500 chip (Agilent). Library concentrations were then normalized to 4 nM and validated by qPCR on a ViiA-7 real-time thermocycler (Applied Biosystems), using qPCR primers recommended in Illumina's qPCR protocol, and Illumina's PhiX control library as standard. The libraries were then pooled at equimolar concentrations and sequenced on an Illumina HiSeq2500 sequencer in rapid mode at a read-length of 250 bp paired-end.

For metatranscriptomic sequencing, approximately 100–150 mg of stool material was used to extract total RNA with the Quick-RNA Fecal/Soil Microprep Kit (Zymo Research, USA) according to the manufacturer's instructions and including the optional in-column DNase I digestion step to eliminate any traces of leftover DNA. For each stool sample, a sequencing library was first constructed using the Truseq Stranded mRNA Library Preparation Kit (Illumina, San Diego, USA) with the following modifications:

The oligo-dT mRNA purification step was omitted and instead, 200ng of total RNA was directly added to the Elution2-Frag-Prime step. The PCR amplification step, which selectively enriches for library fragments that have adapters ligated at both ends, was performed according to the manufacturer's recommendation but the number of amplification cycles was reduced to 12 instead of 15. Each library was uniquely tagged with one of Illumina's TruSeq HT RNA barcode combinations to enable sample pooling for sequencing.

Finished libraries were quantitated using Promega's QuantiFluor dsDNA assay and the average library size was determined on an Agilent Tapestation 4200. Library concentrations were then normalized to 4 nM and validated by qPCR on a QuantStudio-3 real-time PCR system (Applied Biosystems), using the Kapa library quantification kit for Illumina platforms (Kapa Biosystems). The libraries were then pooled at equimolar concentrations and sequenced on an Illumina HiSeq2500 sequencer in rapid mode at a read-length of 100bp paired-end.

Bioinformatics

The metagenomics and metatranscriptomics Illumina reads were adapter-trimmed and quality-trimmed using cutadapt-1.8.1 with parameter of "-q 20 – trim-n – minimum-length 30 – match-read-wildcards".[69]

The metagenomics dataset was mapped against the hg19 human reference genome with bowtie2 2.1.0 using "–very-sensitive-local" preset as its sensitivity parameter.[70] Any reads that can't be confidently mapped against the human genome, which were generated by " – un-conc" switch, were separated and processed as the non-host reads. The non-host reads were then aligned against NCBI non-redundant protein database (downloaded on 3 January 2016) with Diamond version 0.8.5.[71] Based on the resulting alignments, the microbial taxonomical classification was determined using the Lowest-Common-Ancestor (LCA) algorithm implemented in MEGAN6 version 6.4.19 (parameters: maxmatches = 25 minscore = 100 minsupport = 100).[72] In addition, the functional classification was also determined with MEGAN6 against its internal KEGG database (gi2keggMarch2016.bin).

The metatranscriptomics dataset was mapped against GRCh38 reference genome using Hisat2 2.1.0 with default parameters.[73] The unmapped reads, as generated with the " – un-conc" option, were separated into putative ribosomal RNA and non-ribosomal RNA reads with sortMeRNA 2.1.[74] The non-ribosomal reads were then assembled into transcript contigs with Trinity 2.6.5 and the samples' abundance were estimated using the Salmon 0.9.1 package. To annotate the function of the transcripts, the contigs were aligned to NCBI non-redundant protein database (downloaded on 3 January 2016) with Diamond version 0.85.[71] The microbial taxonomical classification was determined using the Lowest-Common-Ancestor (LCA) algorithm implemented in MEGAN6 version 6.4.19 (parameters: maxmatches = 25 minscore = 100 minsupport = 100). In addition, the functional classification was also determined with MEGAN6 against its internal KEGG database (gi2keggMarch2016.bin). Metagenomics abundance data was normalized against the smallest sample library size, while the metatranscriptomics abundance data was normalized using trimmed mean of M values (TMM) method implemented by Trinity.[75] Both metagenomics and metatranscriptomics normalized taxonomical abundance and KEGG classification data was then cross-referenced to produce the table of microbial composition and their gene expression profile. The Shannon Diversity Index were calculated for each sample at genus level. All raw microbiome data are available at NCBI (https://www.ncbi.nlm.nih.gov) (Accession Number PRJNA642723).

Statistical analysis

R Package "vegan" in the R version 3.4.1 within R studio version 1.0.136 was used to plot Principal Coordinates Analysis (PCoA) based on Bray-Curtis distance to visualize the separation of microbiome and metabolites profiles between groups.

All subsequent statistical analysis were carried out using IBM SPSS version 24.0 (IBM SPSS Statistics, Armonk, NY). Univariate analysis (Pearson chi-square test) was used to study demographic, lifestyle, and clinical factors between groups. Data (SCFA absolute concentrations, relative abundance of global metabolites, bacterial

taxonomic metagenomics read counts and metagenomics/transcriptomics functional genes read counts) were subjected to log-transformation in order to normalize the data prior to subsequent analysis. The selection of confounding factors for subsequent analysis adjustment were based on two criteria: (1) known confounding factors that have been reported by other studies and (2) any other potential factors that show significant unequal distribution between groups by Chi-square test ($p < .05$). Exclusion of factors was based on the "backward exclusion" approach in linear mixed models analysis. Briefly, all relevant factors were first all included into the model for adjustment. Factors were then excluded if their p-value were > 0.9, indicating minimal or no effect on the model. Linear mixed model was used to evaluate the longitudinal mean difference of the normalized abundance data and assess the trend significance of the trajectories of eczema group (NAE/AE) compared to controls group (i.e., at the selected four timepoints of week 3, months 3, 6 and 12) among the three clinical groups, adjusted for seven potential confounders (i.e., gender, birth order, mode of delivery, breastfeeding till 6 months, antibiotics during labor, maternal and paternal atopic history). Three variance-covariance structures (compound symmetry, 1st order autoregressive and unstructured) were used for linear mixed model, and the selection of covariance structure was based on Akaike's Information Criterion (AIC) and Schwarz's Bayesian Criterion (BIC). General linear models were performed to compare the normalized abundance data between groups at each timepoint with Bonferroni correction for pairwise comparisons adjusting for baseline values and the 7 mentioned potential confounders. To increase statistical power in view of the relative small sample size, linear mixed model with bootstrapping (1000 iterations) was also performed to evaluate the trend significance of the trajectories. Spearman correlation was performed to assess the strength and direction (positive or negative) of the relationship between variables (SCFA absolute concentrations, relative abundance of global metabolites, bacterial taxonomic metagenomics read counts and metagenomics/transcriptomics functional genes read counts).

For identification of major butyrate- and propionate-producing bacterial families and species, the following criteria were used: (1) Demonstration of a positive Spearman correlation between the bacterial abundance in read count and absolute concentration of corresponding SCFAs ($p < .05$, Spearman correlation value $\rho \geq 0.65$) and (2) Carriage of at least one genomic/transcriptomics functional gene from the butyrate and/or propionate metabolic pathway. (Supplementary Table 9A,9B).

All physiologically relevant global metabolites were grouped based on the KEGG database according to three main carbohydrate metabolic pathways: (1) "Glycolysis-associated intermediates" involving the upstream biosynthesis pathways of both butyrate and propionate which includes pre-glycolysis (starch and sucrose metabolism), glycolysis, para-glycolysis (citrate cycle – TCA, pentose phosphate metabolism) and post-glycolysis (pyruvate metabolism); (2) "Butyrate intermediates" and (3) "Propionate intermediates" (Supplementary Table 10).

All relevant functional genes carried by short-listed species were grouped into "Glycolysis-associated Metabolism", "Butyrate Metabolism" and "Propionate metabolism". Based on KEGG database, "Glycolysis-associated Metabolism" includes functional gene from Glycolysis pathway (ko00010 Glycolysis/Gluconeogenesis), Pre-glycolysis pathway (ko00500 Starch and sucrose metabolism), Para-glycolysis pathway (ko00020 Citrate cycle – TCA cycle, ko00030 Pentose phosphate pathway) and Post-glycolysis pathway (ko00620 Pyruvate metabolism). "Butyrate Metabolism" includes gene from ko00650 Butanoate metabolism. "Propionate Metabolism" includes gene from ko00640 Propanoate metabolism (Supplementary Table 5).

Geometric means and standard deviation range (SD) were used to represent abundance data in log-scale. Abundance data were plotted in GraphPad Prism 7.0 (GraphPad Software, La Jolla, Calif). All statistical significance tests and confidence intervals (CIs) were 2-sided and set at a P value of less than 0.05. A sensitivity analysis involving two extreme "best" and "worst" cases analysis was conducted to test the stability of our model to account for missing data. In "best" scenarios, the missing value(s) were replaced with the highest values of data from other timepoints within that subject. Whereas in

"worst" scenarios, the missing value(s) were replaced with the lowest value. The linear mixed model was then performed using the dataset generated under two extreme "best" and "worst" scenarios while adjusting for the same 7 confounding factors.

Disclosure of potential conflicts of interest

No potential conflicts of interest were disclosed.

ORCID

James Chun Yip Chan http://orcid.org/0000-0003-3162-8389
Rikky W. Purbojati http://orcid.org/0000-0002-2790-4056
Daniela I. Drautz-Moses http://orcid.org/0000-0002-3739-0917
Elizabeth Huiwen Tham http://orcid.org/0000-0003-1037-6143
Oon Hoe Teoh http://orcid.org/0000-0002-5100-1543

List of abbreviation

NAE	Non allergen-sensitized atopic eczema
AE	Allergen-sensitized atopic eczema
HMOs	Human milk oligosaccharides
GUSTO	Growing Up in Singapore Towards healthy Outcomes
SCORAD	SCORing of Atopic Dermatitis (AD) severity
SCFAs	Short-chain fatty acids
KEGG	Kyoto Encyclopedia of Genes and Genomes database
HMDB	Human Metabolome Database
SPT	Skin prick testing
LPS	Lipopolysaccharides
PSA	Polysaccharide A

References

1. Flohr C, Mann J. New insights into the epidemiology of childhood atopic dermatitis. Allergy. 2014;69(1):3–16. doi:10.1111/all.12270.
2. Leung DY, Guttman-Yassky E. Deciphering the complexities of atopic dermatitis: shifting paradigms in treatment approaches. J Allergy Clin Immunol. 2014;134(4):769–779. doi:10.1016/j.jaci.2014.08.008.
3. Blaser MJ, Falkow S. What are the consequences of the disappearing human microbiota? Nat Rev Microbiol. 2009;7(12):887–894. doi:10.1038/nrmicro2245.
4. Mancabelli L, Milani C, Lugli GA, Turroni F, Ferrario C, van Sinderen D, Ventura M. Meta-analysis of the human gut microbiome from urbanized and pre-agricultural populations. Environ Microbiol. 2017;19(4):1379–1390. doi:10.1111/1462-2920.13692.
5. Thompson AL, Monteagudo-Mera A, Cadenas MB, Lampl ML, Azcarate-Peril MA. Milk- and solid-feeding practices and daycare attendance are associated with differences in bacterial diversity, predominant communities, and metabolic and immune function of the infant gut microbiome. Front Cell Infect Microbiol. 2015;5:3. doi:10.3389/fcimb.2015.00003.
6. Tamburini S, Shen N, Wu HC, Clemente JC. The microbiome in early life: implications for health outcomes. Nat Med. 2016;22(7):713–722. doi:10.1038/nm.4142.
7. Berni Canani R, Di Costanzo M, Leone L. The epigenetic effects of butyrate: potential therapeutic implications for clinical practice. Clin Epigenetics. 2012;4(1):4. doi:10.1186/1868-7083-4-4.
8. Morrison DJ, Preston T. Formation of short chain fatty acids by the gut microbiota and their impact on human metabolism. Gut Microbes. 2016;7(3):189–200. doi:10.1080/19490976.2015.1134082.
9. Roduit C, Frei R, Ferstl R, Loeliger S, Westermann P, Rhyner C, Schiavi E, Barcik W, Rodriguez-Perez N,

Wawrzyniak M, et al. High levels of butyrate and propionate in early life are associated with protection against atopy. Allergy. 2019;74(4):799–809. doi:10.1111/all.13660.

10. Wopereis H, Sim K, Shaw A, Warner JO, Knol J, Kroll JS. Intestinal microbiota in infants at high risk for allergy: effects of prebiotics and role in eczema development. J Allergy Clin Immunol. 2018;141 (4):1334–1342.e5. doi:10.1016/j.jaci.2017.05.054.

11. Song H, Yoo Y, Hwang J, Na YC, Kim HS. Faecalibacterium prausnitzii subspecies-level dysbiosis in the human gut microbiome underlying atopic dermatitis. J Allergy Clin Immunol. 2016;137 (3):852–860. doi:10.1016/j.jaci.2015.08.021.

12. Arrieta MC, Stiemsma LT, Dimitriu PA, Thorson L, Russell S, Yurist-Doutsch S, Kuzeljevic B, Gold MJ, Britton HM, Lefebvre DL, et al. Early infancy microbial and metabolic alterations affect risk of childhood asthma. Sci Transl Med. 2015;7(307):307ra152. doi:10.1126/scitranslmed.aab2271.

13. Cait A, Cardenas E, Dimitriu PA, Amenyogbe N, Dai D, Cait J, Sbihi H, Stiemsma L, Subbarao P, Mandhane PJ, et al. Reduced genetic potential for butyrate fermentation in the gut microbiome of infants who develop allergic sensitization. J Allergy Clin Immunol. 2019;144(6):1638–1647 e3. doi:10.1016/j.jaci.2019.06.029.

14. Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, Magris M, Hidalgo G, Baldassano RN, Anokhin AP, et al. Human gut microbiome viewed across age and geography. Nature. 2012;486(7402):222–227. doi:10.1038/nature11053.

15. Azad MB, Konya T, Guttman DS, Field CJ, Sears MR, HayGlass KT, Mandhane PJ, Turvey SE, Subbarao P, Becker AB, et al. Infant gut microbiota and food sensitization: associations in the first year of life. Clin Expl Allergy. 2015;45(3):632–643. doi:10.1111/cea.12487.

16. Smith PK, Masilamani M, Li XM, Sampson HA. The false alarm hypothesis: food allergy is associated with high dietary advanced glycation end-products and proglycating dietary sugars that mimic alarmins. J Allergy Clin Immunol. 2017;139(2):429–437. doi:10.1016/j.jaci.2016.05.040.

17. Lee SY, Lee E, Park YM, Hong SJ. Microbiome in the gut-skin axis in atopic dermatitis. Allergy Asthma Immunol Res. 2018;10(4):354–362. doi:10.4168/aair.2018.10.4.354.

18. Zheng H, Liang H, Wang Y, Miao M, Shi T, Yang F, Liu E, Yuan W, Ji ZS, Li DK. Altered gut microbiota composition associated with eczema in infants. PLoS One. 2016;11(11):e0166026. doi:10.1371/journal.pone.0166026.

19. Reddel S, Del Chierico F, Quagliariello A, Giancristoforo S, Vernocchi P, Russo A, Fiocchi A, Rossi P, Putignani L, El Hachem M. Gut microbiota profile in children affected by atopic dermatitis and evaluation of intestinal persistence of a probiotic mixture. Sci Rep. 2019;9(1):4996. doi:10.1038/s41598-019-41149-6.

20. Pei-Ying Hong BWL, Marion A, Shek LPC, Yap GC, Chua KY, Liu W-T. Comparative analysis of fecal microbiota in infants with and without eczema. PLoS One. 2010;5(4):e9964. doi:10.1371/journal.pone.0009964.

21. Penders J, Thijs C, van den Brandt PA, Kummeling I, Snijders B, Stelma F, Adams H, van Ree R, Stobberingh EE. Gut microbiota composition and development of atopic manifestations in infancy: the KOALA birth cohort study. Gut. 2007;56(5):661–667. doi:10.1136/gut.2006.100164.

22. Kirjavainen PV, Arvola T, Salminen SJ, Isolauri E. Aberrant composition of gut microbiota of allergic infants: a target of bifidobacterial therapy at weaning? Gut. 2002;51(1):51–55. doi:10.1136/gut.51.1.51.

23. Chassaing B, Koren O, Carvalho FA, Ley RE, Gewirtz AT. AIEC pathobiont instigates chronic colitis in susceptible hosts by altering microbiota composition. Gut. 2014;63 (7):1069–1080. doi:10.1136/gutjnl-2013-304909.

24. Purvis JE, Yomano LP, Ingram LO. Enhanced trehalose production improves growth of Escherichia coli under osmotic stress. Appl Environ Microbiol. 2005;71 (7):3761–3769. doi:10.1128/AEM.71.7.3761-3769.2005.

25. Ruhal R, Kataria R, Choudhury B. Trends in bacterial trehalose metabolism and significant nodes of metabolic pathway in the direction of trehalose accumulation. Microb Biotechnol. 2013;6(5):493–502. doi:10.1111/1751-7915.12029.

26. Tieng V, Le Bouguenec C, Du Merle L, Bertheau P, Desreumaux P, Janin A, Charron D, Toubert A. Binding of escherichia coli adhesin AfaE to CD55 triggers cell-surface expression of the MHC class I-related molecule MICA. Proc Natl Acad Sci U S A. 2002;99 (5):2977–2982. doi:10.1073/pnas.032668099.

27. Cieza RJ, Hu J, Ross BN, Sbrana E, Torres AG. The IbeA invasin of adherent-invasive Escherichia coli mediates interaction with intestinal epithelia and macrophages. Infect Immun. 2015;83(5):1904–1918. doi:10.1128/IAI.03003-14.

28. Khan MA, Bouzari S, Ma C, Rosenberger CM, Bergstrom KS, Gibson DL, Steiner TS, Vallance BA. Flagellin-dependent and -independent inflammatory responses following infection by enteropathogenic Escherichia coli and Citrobacter rodentium. Infect Immun. 2008;76(4):1410–1422. doi:10.1128/IAI.01141-07.

29. Huang JY, Lee SM, Mazmanian SK. The human commensal Bacteroides fragilis binds intestinal mucin. Anaerobe. 2011;17(4):137–141. doi:10.1016/j.anaerobe.2011.05.017.

30. Shu HY, Fung CP, Liu YM, Wu KM, Chen YT, Li LH, Liu TT, Kirby R, Tsai SF. Genetic diversity of capsular polysaccharide biosynthesis in Klebsiella pneumoniae clinical isolates. Microbiology. 2009;155(Pt 12):4170–4183. doi:10.1099/mic.0.029017-0.

31. Wang X, Quinn PJ. Lipopolysaccharide: biosynthetic pathway and structure modification. Prog Lipid Res. 2010;49(2):97–107. doi:10.1016/j.plipres.2009.06.002.

32. Kim HK, Rutten NB, Besseling-van der Vaart I, Niers LE, Choi YH, Rijkers GT, van Hemert S. Probiotic supplementation influences faecal short chain fatty acids in infants at high risk for eczema. Benef Microbes. 2015;6(6):783–790. doi:10.3920/BM2015.0056.

33. Park BS, Lee JO. Recognition of lipopolysaccharide pattern by TLR4 complexes. Exp Mol Med. 2013;45:e66. doi:10.1038/emm.2013.97.

34. Simpson A, Martinez FD. The role of lipopolysaccharide in the development of atopy in humans. Clin Expl Allergy. 2010;40(2):209–223. doi:10.1111/j.1365-2222.2009.03391.x.

35. Wampach L, Heintz-Buschart A, Fritz JV, Ramiro-Garcia J, Habier J, Herold M, Narayanasamy S, Kaysen A, Hogan AH, Bindl L, et al. Birth mode is associated with earliest strain-conferred gut microbiome functions and immunostimulatory potential. Nat Commun. 2018;9 (1):5091. doi:10.1038/s41467-018-07631-x.

36. Marcobal A, Barboza M, Froehlich JW, Block DE, German JB, Lebrilla CB, Mills DA. Consumption of human milk oligosaccharides by gut-related microbes. J Agric Food Chem. 2010;58(9):5334–5340. doi:10.1021/jf9044205.

37. Jakobsson HE, Abrahamsson TR, Jenmalm MC, Harris K, Quince C, Jernberg C, Bjorksten B, Engstrand L, Andersson AF. Decreased gut microbiota diversity, delayed Bacteroidetes colonisation and reduced Th1 responses in infants delivered by caesarean section. Gut. 2014;63(4):559–566. doi:10.1136/gutjnl-2012-303249.

38. Martin R, Makino H, Cetinyurek Yavuz A, Ben-Amor K, Roelofs M, Ishikawa E, Kubota H, Swinkels S, Sakai T, Oishi K, et al. Early-life events, including mode of delivery and type of feeding, siblings and gender, shape the developing gut microbiota. PLoS One. 2016;11(6):e0158498. doi:10.1371/journal.pone.0158498.

39. Szari S, Quinn JA. Supporting a Healthy Microbiome for the Primary Prevention of Eczema. Clin Rev Allergy Immunol. 2019;57(2):286–293. doi:10.1007/s12016-019-08758-5.

40. Fiocchi A, Pawankar R, Cuello-Garcia C, Ahn K, Al-Hammadi S, Agarwal A, Beyer K, Burks W, Canonica GW, Ebisawa M, et al. world allergy organization-McMaster University guidelines for allergic disease prevention (GLAD-P): probiotics. World Allergy Organ J. 2015;8(1):4. doi:10.1186/s40413-015-0055-2.

41. Cuello-Garcia CA, Fiocchi A, Pawankar R, Yepes-Nunez JJ, Morgano GP, Zhang Y, Ahn K, Al-Hammadi S, Agarwal A, Gandhi S, et al. World allergy organization-McMaster university guidelines for allergic disease prevention (GLAD-P): prebiotics. World Allergy Organ J. 2016;9:10. doi:10.1186/s40413-016-0102-7.

42. Abrahamsson TR, Jakobsson HE, Andersson AF, Bjorksten B, Engstrand L, Jenmalm MC. Low diversity of the gut microbiota in infants with atopic eczema. J Allergy Clin Immunol. 2012;129(2):434–40, 440 e1-2. doi:10.1016/j.jaci.2011.10.025.

43. Stewart CJ, Ajami NJ, O'Brien JL, Hutchinson DS, Smith DP, Wong MC, Ross MC, Lloyd RE, Doddapaneni H, Metcalf GA, et al. Temporal development of the gut microbiome in early childhood from the TEDDY study. Nature. 2018;562(7728):583–588. doi:10.1038/s41586-018-0617-x.

44. Loo EXL, Sim JZT, Loy SL, Goh A, Chan YH, Tan KH, Yap F, Gluckman PD, Godfrey KM, Van Bever H, et al. Associations between caesarean delivery and allergic outcomes: results from the GUSTO study. Ann Allergy Asthma Iimmunol. 2017;118(5):636–638. doi:10.1016/j.anai.2017.02.021.

45. d'Hennezel E, Abubucker S, Murphy LO, Cullen TW. Total lipopolysaccharide from the human gut microbiome silences toll-like receptor signaling. mSystems. 2017;2(6). doi:10.1128/mSystems.00046-17.

46. Evrard B, Balestrino D, Dosgilbert A, Bouya-Gachancard JL, Charbonnel N, Forestier C, Tridon A. Roles of capsule and lipopolysaccharide O antigen in interactions of human monocyte-derived dendritic cells and Klebsiella pneumoniae. Infect Immun. 2010;78 (1):210–219. doi:10.1128/IAI.00864-09.

47. Troy EB, Kasper DL. Beneficial effects of Bacteroides fragilis polysaccharides on the immune system. Front Biosci (Landmark Ed). 2010;15:25–34. doi:10.2741/3603.

48. Koropatkin NM, Cameron EA, Martens EC. How glycan metabolism shapes the human gut microbiota. Nat Rev Microbiol. 2012;10(5):323–335. doi:10.1038/nrmicro2746.

49. Yu ZT, Chen C, Newburg DS. Utilization of major fucosylated and sialylated human milk oligosaccharides by isolated human gut microbes. Glycobiology. 2013;23 (11):1281–1292. doi:10.1093/glycob/cwt065.

50. Garmory HS, Titball RW. ATP-binding cassette transporters are targets for the development of antibacterial vaccines and therapies. Infect Immun. 2004;72 (12):6757–6763. doi:10.1128/IAI.72.12.6757-6763.2004.

51. Kunisawa J, Hashimoto E, Ishikawa I, Kiyono H. A pivotal role of vitamin B9 in the maintenance of regulatory T cells in vitro and in vivo. PLoS One. 2012;7(2):e32094. doi:10.1371/journal.pone.0032094.

52. Stover PJ. Physiology of folate and vitamin B12 in health and disease. Nutr Rev. 2004;62(6 Pt 2):S3–12. discussion S13. doi:10.1111/j.1753-4887.2004.tb00070.x.

53. Shen X, Wang M, Zhang X, He M, Li M, Cheng G, Wan C, He F. Dynamic construction of gut microbiota may influence allergic diseases of infants in Southwest China. BMC Microbiol. 2019;19(1):123. doi:10.1186/s12866-019-1489-4.

54. Nylund L, Nermes M, Isolauri E, Salminen S, de Vos WM, Satokari R. Severity of atopic disease inversely correlates with intestinal microbiota diversity and butyrate-producing bacteria. Allergy. 2015;70 (2):241–244. doi:10.1111/all.12549.

55. Candela M, Rampelli S, Turroni S, Severgnini M, Consolandi C, De Bellis G, Masetti R, Ricci G, Pession A, Brigidi P. Unbalance of intestinal microbiota in atopic children. BMC Microbiol. 2012;12:95. doi:10.1186/1471-2180-12-95.

56. Lee MJ, Kang MJ, Lee SY, Lee E, Kim K, Won S, Suh DI, Kim KW, Sheen YH, Ahn K, et al. Perturbations of gut microbiome genes in infants with atopic dermatitis according to feeding type. J Allergy Clin Immunol. 2018;141(4):1310–1319. doi:10.1016/j.jaci.2017.11.045.

57. Tan J, McKenzie C, Vuillermin PJ, Goverse G, Vinuesa CG, Mebius RE, Macia L, Mackay CR. Dietary fiber and bacterial SCFA enhance oral tolerance and protect against food allergy through diverse cellular pathways. Cell Rep. 2016;15(12):2809–2824. doi:10.1016/j.celrep.2016.05.047.

58. Al Nabhani Z, Dulauroy S, Marques R, Cousu C, Al Bounny S, Dejardin F, Sparwasser T, Berard M, Cerf-Bensussan N, Eberl G. A weaning reaction to microbiota is required for resistance to immunopathologies in the adult. Immunity. 2019;50(5):1276–1288 e5. doi:10.1016/j.immuni.2019.02.014.

59. Low JSY, Soh SE, Lee YK, Kwek KYC, Holbrook JD, Van der Beek EM, Shek LP, Goh AEN, Teoh OH, Godfrey KM, et al. Ratio of Klebsiella/Bifidobacterium in early life correlates with later development of paediatric allergy. Benef Microbes. 2017;8(5):681–695. doi:10.3920/BM2017.0020.

60. Soh SE, Lee SS, Hoon SW, Tan MY, Goh A, Lee BW, Shek LP, Teoh OH, Kwek K, Saw SM, et al. The methodology of the GUSTO cohort study: a novel approach in studying pediatric allergy. Asia Pac Allergy. 2012;2(2):144–148. doi:10.5415/apallergy.2012.2.2.144.

61. Bunyavanich S, Rifas-Shiman SL, Platts-Mills TA, Workman L, Sordillo JE, Camargo CA Jr., Gillman MW, Gold DR, Litonjua AA. Peanut, milk, and wheat intake during pregnancy is associated with reduced allergy and asthma in children. J Allergy Clin Immunol. 2014;133(5):1373–1382. doi:10.1016/j.jaci.2013.11.040.

62. Erkkola M, Kaila M, Nwaru BI, Kronberg-Kippila C, Ahonen S, Nevalainen J, Veijola R, Pekkanen J, Ilonen J, Simell O, et al. Maternal vitamin D intake during pregnancy is inversely associated with asthma and allergic rhinitis in 5-year-old children. Clin Expl Allergy. 2009;39(6):875–882. doi:10.1111/j.1365-2222.2009.03234.x.

63. Floistrup H, Swartz J, Bergstrom A, Alm JS, Scheynius A, van Hage M, Waser M, Braun-Fahrlander C, Schram-Bijkerk D, Huber M, et al. Allergic disease and sensitization in Steiner school children. J Allergy Clin Immunol. 2006;117(1):59–66. doi:10.1016/j.jaci.2005.09.039.

64. Khoo J, Shek LP, Khor ES, Wang DY, Lee BW. Pattern of sensitization to common environmental allergens amongst atopic Singapore children in the first 3 years of life. Asian Pac J Allergy Immunol. 2001;19:225–229.

65. Yi FC, Chew FT, Jimenez S, Chua KY, Lee BW. Culture of blomia tropicalis and IgE immunoblot characterization of its allergenicity. Asian Pac J Allergy Immunol. 1999;17:189–194.

66. Bousquet J, Khaltaev N, Cruz AA, Denburg J, Fokkens WJ, Togias A, Zuberbier T, Baena-Cagnani CE, Canonica GW, van Weel C, et al. Allergic rhinitis and its impact on asthma (ARIA) 2008 update (in collaboration with the world health organization, GA(2) LEN and AllerGen). Allergy. 2008;63:8–160.

67. Chan JC, Kioh DY, Yap GC, Lee BW, Chan EC. A novel LCMSMS method for quantitative measurement of short-chain fatty acids in human stool derivatized with (12)C- and (13)C-labelled aniline. J Pharm Biomed Anal. 2017;138:43–53. doi:10.1016/j.jpba.2017.01.044.

68. Sud M, Fahy E, Cotter D, Azam K, Vadivelu I, Burant C, Edison A, Fiehn O, Higashi R, Nair KS, et al. Metabolomics workbench: an international repository for metabolomics data and metadata, metabolite standards, protocols, tutorials and training, and analysis tools. Nucleic Acids Res. 2016;44(D1):D463–70. doi:10.1093/nar/gkv1042.

69. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. 2011;17:10–12. doi:10.14806/ej.17.1.200.

70. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9(4):357–359. doi:10.1038/nmeth.1923.

71. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nat Methods. 2015;12(1):59–60. doi:10.1038/nmeth.3176.

72. Huson DH, Mitra S, Ruscheweyh HJ, Weber N, Schuster SC. Integrative analysis of environmental sequences using MEGAN4. Genome Res. 2011;21(9):1552–1560. doi:10.1101/gr.120618.111.

73. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015;12(4):357–360. doi:10.1038/nmeth.3317.

74. Kopylova E, Noe L, Touzet H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics. 2012;28(24):3211–3217. doi:10.1093/bioinformatics/bts611.

75. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc. 2013;8(8):1494–1512. doi:10.1038/nprot.2013.084.