KØBENHAVNS UNIVERSITET
FACULTY OF SCIENCE

COPSAC
COPENHAGEN PROSPECTIVE STUDIES
ON ASTHMA IN CHILDHOOD

PhD Thesis

Jie Jiang

Cesarean section, gut microbiome, and asthma risk

Supervisor: Jakob Stokholm
Date: December 19, 2025

Title	Cesarean section, gut microbiome, and asthma risk
Author	Jie Jiang, Msc
Principle Supervisor	Jakob Stokholm, MD, PhD, Professor
Co-supervisor	Jonathan Thorsen, MD, PhD, Associate Professor
Co-supervisor	Casper Sahl Poulsen, PhD, Postdoc
Affiliation	COPSAC, Copenhagen Prospective Studies on Asthma in Childhood, Copenhagen University Hospital - Herlev and Gentofte, Copenhagen, Denmark
Assessment committee	Bekzod Khakimov, Associate Professor (Chair) University of Copenhagen John Penders, PhD, Professor Maastricht University Medical Centre, Netherlands Carsten Eriksen, PhD, Tenure-track scientist Technical University of Denmark

List of papers

Paper I: Key bacteria and environmental factors associated with the restoration of the infant gut microbiome after cesarean section

Jie Jiang, Casper Sahl Poulsen, Ulrika Boulund, Shiraz Shah, Urvish Trivedi, Madhumita Bhattacharyya, Avidan U. Neumann, Darlene L.Y. Dai, Charisse Petersen, Courtney Hoskinson, Theo J. Moraes, Piushkumar J. Mandhane, Elinor Simons, Meghan B. Azad, Padmaja Subbarao, Klaus Bønnelykke, Bo Chawes, Stuart E Turvey, Søren J. Sørensen*, Jonathan Thorsen*, Jakob Stokholm*
Accepted in *Nature Commuinications*, December 2025

Paper II: Early Gut Microbiome Composition and Functional Pathways Linking Cesarean section to Asthma Susceptibility

Jie Jiang, Shiraz Shah, Lili Yang, Kaare Tranæs, Casper Sahl Poulsen, Ulrika Boulund, Trine Zachariasen, Xuanji Li, Urvish Trivedi, Klaus Bønnelykke, Bo Chawes, Søren J. Sørensen, Jonathan Thorsen*, Jakob Stokholm*
Manuscript in preparation.

Paper III: Microbially-derived tryptophan metabolites in the infant gut mediate the increased risk of asthma following cesarean section

Jie Jiang, Darlene L.Y. Dai, Casper Sahl Poulsen, Kaare Tranæs, Tingting Wang, Charisse Petersen, Courtney Hoskinson, Theo J. Moraes, Piushkumar J. Mandhane, Elinor Simons, Meghan B. Azad, Padmaja Subbarao, Bo Chawes, Klaus Bønnelykke, Jonathan Thorsen*, Stuart E Turvey*, Jakob Stokholm*
Submitted manuscript.

Table of Contents

Acknowledgement

I have mentally drafted my acknowlegement on my first day in Denmark: thanking Jakob for offering me this opportunity and for renting me his daughter’s apartment; thanking Yang and Chao for picking me up at 6:00 a.m. at the airport. And the list grew longer and longer as time went by.

But before that, I want to take a moment to thank myself and my parents. The last few months were tough, as expected. I did not regret it even one, as expected. Three years age, I made a big decision: I quit my job - a pretty decent job, I would say - left my family, my comfort zone behind, and came to Denmark to build a life here and pursue a PhD. At the mutual alignment meeting, Jakob, Jonathan, and I talked about why we choose to do research. That was where I learned the COPSAC spirit: “What’s in it for the child”. It is a very good reason, and it has been part of my motivation ever since. But my first reason was simpler - I wanted to keep the intellectual stimulation. I wanted to nurture my creativity and maintain my ability to be curious. It turned out I made the right decision, and I have embraced everything that came with it, the passion, the happiness, the frustration, and even the periods of depression.

这一切都离不开我父母的支持. 我们是相对来说比较传统的家庭. 传统但有爱. 跟朋友聊起我的爸妈, 他们会说你爸妈很开明, 能够同意你一个人出国读博. 我却不这么想. 我能感受到肯定有那么些个夜晚, 两个人聊起这件事, 还是会叹气, 还是会想她跑那么远干什么, 呆在我们身边多好, 生病了还能有个照应, 不像现在, 连吃个韭菜饺子都是奢侈. 即使如此, 结尾总是随她去吧, 她大了, 该是自己做决定的时候了. 与开不开明无关, 只是两个拿小孩没办法了的父母罢了. 这三年来, 我感受到了我的成长, 也感受到了我爸妈的成长. 我们的对话开始变得平等, 直抒胸臆. 我感受到来自家的能量. 衷心感谢我的父母, 愿意倾听我的想法, 愿意同我一起成长. 我们都要照顾好自己, 你们托举我到我现在的地方, 接下来轮到我带你们来看一看我眼中的风景.

Now I thank Jakob. I don’t need to tell you about his expertise or how professional he is. He is already a professor. What I want to thank Jakob for is being such a great group leader and the best supervisor I could have hoped for. One day, during my first few weeks at COPSAC, Jonathan told me in passing that he looks up to Jakob a lot. I don’t remember the context or the rest of the conversation, but I remember the sincerity in Jonathan’s eyes. It means a lot when an employee genuinely looks up to the senior in the group, it says something about leadership, and it says something about the atmosphere. Now I see it too, and I completely agree with Jonathan. Beyond showing me how to do good research, Jakob taught me how to make presentations and even took the time to explain Danish jokes. When I once said I might work during my vacation, he looked genuinely surprised and asked why. I didn’t know how to answer that question, so I ended up not working at all during my vacation. The microbiome office is a special place because of our group leader (and, of course, because we are in it). When Jakob moved upstairs, we cried a bit. (This is a joke.) In our group, I feel safe asking “stupid” questions, asking for help, and I feel happy when I become capable of helping others. Thank you, Jakob, for creating that environment. And honestly, how you manage to support a family, be a professor, write so many beautiful R functions, and publish so many great scientific articles is still beyond me. I burned out after writing one PhD thesis.

Jonathan is the anchor of our office. Trine once said that Jonathan has this “adult vibe” that people naturally want his second opinion on anything. I think I know why. Jonathan takes initiative. The first defense I attended was Rikke’s, held at a hospital about a 1.5-hour ride from COPSAC. Jonathan suggested we carpool, so he made a Google Sheet and sent an email to get it organized. COPSAC people like chili sauce; he said we should have a chili-sauce club, so he sent another email and created the CCCC club. He is full of knowledge, and he doesn’t mind sharing it with us. He started the yearly statistics sessions, going through basic analyses. And he does it every year. I learn a lot by talking with Jonathan and listening to him, and I learn even more simply by observing him turn an idea into action. And he plays instruments, guitar, bass, and piano. I started learning piano, which was inspired by him.

Casper is my stat buddy - the best stat buddy ever. Well, I’ve only had one stat buddy, but still. I started our very first stat-buddy meeting by asking, “What’s a phyloseq object?” At the second meeting, I said I didn’t like phyloseq because I didn’t understand it. So Casper started over and walked me through the structure of phyloseq. At the third meeting, I asked if we could go through the utilities again. It becomes a “stupid question” when you ask it the third time. Casper didn’t have to explain it three times, he could have said, “Just read the documentation.” But he explained it three times. I had very little experience with R when I started my PhD. I was eager to learn, but I didn’t know how. Casper was the one who guided me patiently, step by step, with hands-on exercises. Now I’ve become someone else’s stat buddy, and I’m proud to tell them that I’m teaching what I learned from mine.

That’s my supervision team. I also appreciate everyone else in our office. Shiraz, who always brings fresh air to the room, and contagious enthusiasm when there’s good science or autoencoders. Michael, my fellow PhD friend, good luck with your submission in a few weeks. Kaare and Anton, both brilliant at physiology and math, and my first go-to people when the seniors aren’t around. Ulrika, thank you for being part of our team, I admire how calm, peaceful, and professional you are. We are such different people brought together by research and science, and I feel honored to do this work alongside you.

And I especially want to thank Connie. When I’m not talking about science, I talk about everything else with Connie. She recommended great places to buy hiking shoes, Korean drama series, the best bakery for rye bread, and the right snacks for the right mood. The microbiome office took me in because we do research, COPSAC took me in because Connie took me in.

Last but not least, my Chinese friends. 罗洋和超超, 婷姐和玩玩师兄,感谢有你们在丹麦. 我真的去你们家吃了太多顿饭了. 在来丹麦之前我建立了全方位的防备, 因为刷到了太多被自己人骗了的案例. 我给自己的信条是, 对每一个遇见的中国人都先信任, 直到对方失去了我的信任为止. 这是一个非常极端且矛盾的状态, 要么是乌托邦, 要么是全盘崩溃, 没有中间值. 你们的存在给我建立了一

个乌托邦, 让我觉得在我的小圈子里我是安全的. 与此同时, 我希望我也给你们带来了一些能量. 以后你们的朋友我也终于是有能力请你们吃饭了, 我非常开心, 非常非常开心. 罗交和吴老师, 我觉得罗洋去另一个单位转了一圈最大的收获就是给他儿子找到了这么好的干妈, 给我们找到了一个非常好的新朋友, 你的真诚大方和善良带给我很多触动, 我希望你也能感受到我的真诚. 丽丽和卓冰, 我同甘苦共患难的朋友, 读博好苦, 好在我们有伴儿, 能互相加油打劲. 还有我远在国内的朋友们, 感谢你们还愿意跟我保持联系, 海内存知己, 天涯共比邻. 王玲, 感谢你第一个来丹麦找我玩, 时不时去探望我的父母, 最近我们聊得很少, 但我相信再见面, 什么都没有变. 史俊杰, 第二个来丹麦找我玩的朋友, 去了最北的地方, 感受了零下三十度的芬兰, 聊了很多, 发现各自都在自己的人生里成长了很多. 淑怡, 谁能想到前同事能成为互相支持互相陪伴的好朋友呢. 从我出国前你给准备的记事本, 项链, 便利贴, 毛毯甚至书包, 到出国后一次不落的节日祝福和时不时的关心和问候, 我只觉得何德何能. 曾珍和慧敏. 第三批来找我玩的朋友. 曾珍真的是 E 人典范了, 主动联系朋友的楷模. 十年不联系, 出国前愣是加上了好友带我吃了很多无锡好吃的餐馆. 张玲, 不夸张的说你出国日本是激励我出国读博的事迹之一, 现在你也来到了丹麦, 遇到了很好的老板, 正在做非常有意义的课题, 感谢有你, 我们来日方长. 最后的最后, 烁烁, 我永远都会记得我毕业时出国前你对我的鼓励, 我一句想试试, 你把所有能给的资料全给我发来了, 数十条 60 秒的语音给我讲你当时的申请经验. 这三年里我们聊得不多, 但是每一次都熟稔得像是我们一直在一个地方一样. 很高兴你找到了自己的幸福, 等着我去找你玩. 还有很多渐行渐远的朋友, 我在这里暂且按下不表, 但不代表我不记得. 我感谢我们一起走过的时光, 每一段. 我希望大家都顺利, 都能过自己想要的生活, 都能成为自己想成为的人.

最后, 感谢国家留学基金委员会, 因为有你们, 我才能有这么好的机会出国学习. 我想用毛主席的语录选段结束我的致谢, 这段话在我灰心的时候给了我很大的力量: 不要被敌人的气势汹汹所吓到, 不要被尚能忍耐的困难所沮丧, 不要为一时的挫折而灰心. 道路是曲折的, 前途是光明的. 黑暗即将过去, 光明就在眼前. 有利的条件和主动的恢复产生于再坚持一下的努力之中!

姜洁 Jie Jiang

哥本哈根 2025 年冬

Copenhagen, December 2025

Summary

Early life represents a critical window for immune development, during which the gut microbiome plays a central role in shaping long-term health. Delivery by cesarean section (CS) profoundly alters the initial establishment of the infant gut microbiota and has been consistently associated with increased risks of immune-mediated diseases, including asthma. However, the mechanisms linking early microbial perturbations to later disease remain incompletely understood, and potential modifiable factors supporting microbiome restoration have not been clearly identified. This PhD thesis aimed to characterize the development of the infant gut microbiome following CS delivery, identify key ecological and environmental drivers of its restoration, and elucidate the microbial and metabolic pathways that may mediate the association between CS delivery and asthma risk.

The PhD thesis is based on data from the Copenhagen Prospective Studies on Asthma in Childhood 2010 (COPSAC$_{2010}$) cohort of 700 mother-child pairs, recruited during pregnancy week 24 and monitored through childhood with frequent detailed clinical visits to the COPSAC study clinic. A validation and replication study was based on data from a Canadian birth cohort (the CHILD study), which enrolled 3,405 subjects since pregnancy from 4 largely urban study centers across Canada (Vancouver, Edmonton, Winnipeg, and Toronto) from 2008 to 2012. The primary endpoints for both cohorts are the development of asthma, which was prospectively diagnosed using pre-defined quantitative criteria in both cohorts.

The thesis contains three scientific papers.

In **paper I**, we investigated the ecological and environmental factors that facilitate the restoration of the gut microbiome in CS-born infants. Based on a 1-year restoration score defined in our previous work, we quantified how closely the 1-year microbial community of each infant resembled that of vaginally delivered infants. Through 16S rRNA gene sequencing at 1 week and 1 month in the COPSAC cohort, we found that the abundance of *Bifidobacterium longum*, *Bacteroides fragilis*, and *Bacteroides vulgatus* at 1 week predicted a higher degree of restoration at 1 year. The presence of older siblings emerged as a strong environmental factor promoting microbiome restoration, mediated through increased colonization with restoration-associated bacteria and inhibition of *Clostridium perfringens*. The restoration score, its determinants, and its relationship with later asthma were successfully replicated in the independent CHILD cohort (16S rRNA), highlighting the robustness of these findings. Together, this work identifies specific microbial and environmental targets for early interventions to mitigate the long-term health risks associated with CS delivery.

In **paper II**, we expanded our previous findings based on 16S rRNA data to metagenomics data in the COPSAC$_{2010}$ cohort. We profiled not only the taxonomic composition but also the functional development of the gut microbiome at 1 month and 1 year of age using metagenomics sequencing, and investigated how CS affects the gut microbiome composition and its pathways. At 1 month, CS-born infants showed very distinct taxonomic and functional compositions. By 1 year, taxonomic profiles remained distinguishable by delivery mode,

whereas pathway profiles largely converged. An increased asthma risk was only associated with 1-year CS microbial scores but not with CS pathway scores (1-year CS microbial score: Generalized Estimating Equation (GEE) Odds ratio (OR) per SD=1.21, 95% confidence interval [1.01,1.45], P=0.03; Adjusted GEE OR=1.23, [1.02,1.48], P=0.03). In conclusion, we confirmed the association between the persistent perturbed microbial composition and asthma risk, but we found no association between the CS-altered pathways and asthma risk, expanding our understanding of the influence of CS delivery on the gut microbiome.

In **paper III**, we leveraged metagenomic and gut metabolomic data from the CHILD cohort to identify metabolic alterations following CS. We characterized the CS microbial signatures at 3 months and 1 year of age using machine learning models. The association between CS and the gut microbiota composition was more pronounced at 3 months (area under the curve (AUC)=0.78) and attenuated to age 1 year (AUC=0.60). An increased asthma risk was only associated with 1-year CS microbial scores (Odds ratio 1.41[1.12,1.80], P<0.01; adjusted Odds ratio 1.30[1.03,1.68], P=0.03), which is in line with previous work. Extending this, we identified metabolic imbalances associated with 1-year cesarean section microbial scores, marked by elevated tryptophan metabolites in the stool metabolome at 1 year of age. Our results replicate the mediating role of the 1-year cesarean section microbial signature in the association between cesarean section and asthma risk, and provide new insight into possible underlying mechanisms of host-microbe interaction following cesarean section.

Collectively, these findings shed light on the relevance of the gut microbiome for asthma in childhood based on different birth cohorts and data layers. In the future, strategies for restoring the gut microbiota may help prevent and alleviate asthma in childhood.

Abbreviations

COPSAC - COpenhagen Prospective Studies on Asthma in Childhood

COPSAC_{2010} - COPSAC unselected pregnancy and birth cohort

CHILD - The Canadian Healthy Infant Longitudinal Development (CHILD) birth cohort study

CS – Cesarean Section

sPLS - sparse partial least squares (sPLS) model

GABA – Gamma-aminobutyric acid

LPS – Lipopolysaccharide

ROS – Reactive Oxygen Species

SCFA – Short-chain fatty acid(s)

T6SS – Type VI secretion system

Resumé

Det tidlige liv udgør et kritisk vindue for immunsystemets udvikling, hvor tarmmikrobiomet spiller en central rolle for barnets senere sundhed. Fødsel ved kejsersnit (CS) ændrer markant den initiale etablering af spædbarnets tarmmikrobiota og er gentagne gange blevet associeret med øget risiko for immunmedierede sygdomme, herunder astma. Mekanismerne, der forbinder tidlige mikrobielle forstyrrelser med senere sygdom, er imidlertid endnu kun delvist forstået, og potentielt modificerbare faktorer, der kan understøtte genopretning af mikrobiomet, er ikke klart identificeret. Formålet med denne ph.d.-afhandling var at karakterisere udviklingen af spædbarnets tarmmikrobiom efter fødsel ved kejsersnit, identificere centrale økologiske og miljømæssige drivkræfter for dets genopretning samt at belyse de mikrobielle og metaboliske pathways, der kan mediere sammenhængen mellem fødsel ved kejsersnit og astmarisiko.

Ph.d.-afhandlingen er baseret på data fra fødselskohorten Copenhagen Prospective Studies on Asthma in Childhood 2010 (COPSAC2010) bestående af 700 mor-barn-par, inkluderet i graviditetsuge 24 og fulgt gennem barndommen med hyppige, detaljerede kliniske besøg i COPSAC-studieklinikken. Et validerings- og replikationsstudie var baseret på data fra en canadisk fødselskohorte (CHILD-studiet), som inkluderede 3.405 deltagere fra graviditeten ved 4 overvejende urbane studiecentre i Canada (Vancouver, Edmonton, Winnipeg og Toronto) i perioden 2008–2012. Det primære endepunkt i begge kohorter er udvikling af astma, som blev prospektivt diagnosticeret ud fra prædefinerede kvantitative kriterier i begge kohorter.

Afhandlingen består af tre videnskabelige artikler.

I **artikel I** undersøgte vi de økologiske og miljømæssige faktorer, der faciliterer genopretningen af tarmmikrobiomet hos børn født ved kejsersnit. Med udgangspunkt i en 1-års “restoration score” (genopretningsscore) defineret i vores tidligere arbejde kvantificerede vi, hvor tæt den mikrobielle sammensætning ved 1 års-alderen hos det enkelte barn lignede den hos vaginalt fødte børn. Ved hjælp af 16S rRNA-gensekventering ved 1 uge og 1 måned i COPSAC-kohorten fandt vi, at en højere forekomst af *Bifidobacterium longum*, *Bacteroides fragilis* og *Bacteroides vulgatus* ved 1 uge forudsagde en højere grad af genopretning ved 1 år. Tilstedeværelsen af ældre søskende fremkom som en stærk miljøfaktor, der fremmer genopretning af mikrobiomet, medieret gennem øget kolonisering med genopretnings-associerede bakterier og hæmning af *Clostridium perfringens*. Genopretningsscoren, dens determinanter og dens relation til senere astma blev succesfuldt replikeret i den uafhængige CHILD-kohorte (16S rRNA), hvilket understreger robustheden af disse fund. Samlet identificerer dette arbejde specifikke mikrobielle og miljømæssige mål for tidlige interventioner, der kan reducere de langsigtede helbredsrisici forbundet med kejsersnit.

I **artikel II** udvidede vi vores tidligere fund baseret på 16S rRNA-data til metagenomiske data i COPSAC2010-kohorten. Vi beskrev både den taksonomiske og funktionelle udvikling af tarmmikrobiomet ved 1 måned og 1 års-alderen ved hjælp af metagenomsekventering og undersøgte, hvordan kejsersnit påvirker tarmmikrobiomets sammensætning og dets pathways.

Ved 1 måned havde børn født ved kejsersnit meget distinkte taksonomiske og funktionelle profiler. Ved 1 år var de taksonomiske profiler fortsat adskillelige efter fødselsmåde, mens pathway-profiler i høj grad var konvergeret. En øget astmarisiko var kun associeret med 1-års kejsersnit-mikrobielle scores, men ikke med kejsersnit-pathway-scores (1-års kejsersnit-mikrobiel score: Generalized Estimating Equation (GEE) odds ratio (OR) pr. SD = 1,21, 95 % konfidensinterval [1,01; 1,45], P = 0,03; justeret GEE OR = 1,23 [1,02; 1,48], P = 0,03). Sammenfattende bekræftede vi sammenhængen mellem en vedvarende ændret mikrobiel sammensætning og astmarisiko, men fandt ingen association mellem de kejsersnit-ændrede pathways og astmarisiko, hvilket udvider vores forståelse af kejsersnits indflydelse på tarmmikrobiomet.

I **artikel III** udnyttede vi metagenomiske data og data om tarmmetabolomet fra CHILD-kohorten til at identificere metaboliske ændringer efter kejsersnit. Vi karakteriserede kejsersnitsrelaterede mikrobielle signaturer ved 3 måneder og 1 års-alderen ved hjælp af machine learning modeller. Sammenhængen mellem kejsersnit og tarmmikrobiotaens sammensætning var mere udtalt ved 3 måneder (arealet under kurven (AUC) = 0,78) og var attenueret ved 1 års-alderen (AUC = 0,60). En øget astmarisiko var kun associeret med 1-års kejsersnit-mikrobielle scores (odds ratio 1,41 [1,12; 1,80], P < 0,01; justeret odds ratio 1,30 [1,03; 1,68], P = 0,03), hvilket er i overensstemmelse med tidligere arbejde. Herudover identificerede vi metaboliske ubalancer associeret med 1-års kejsersnitsrelaterede mikrobielle scores, karakteriseret ved forhøjede tryptofanmetabolitter i fæcesmetabolomet ved 1 års-alderen. Vores resultater replikerer den medierende rolle af den 1-årige kejsersnitsrelaterede mikrobielle signatur i sammenhængen mellem kejsersnit og astmarisiko og giver ny indsigt i mulige underliggende mekanismer for vært–mikrobe-interaktion efter kejsersnit.

Samlet kaster disse fund lys over betydningen af tarmmikrobiomet for astma i barndommen på tværs af forskellige fødselskohorter og datalag. I fremtiden kan strategier til genopretning af tarmmikrobiotaen potentielt bidrage til at forebygge og afhjælpe astma i barndommen.

Introduction

Cesarean section

Cesarean section (CS) derives from the Roman legal code, the *lex Caesarea*, which prescribed that a baby should be cut from its mother's womb if she dies before giving birth[1]. The initial purpose was essentially to retrieve the infant from a dead or dying mother, either in the vain hope of saving the baby's life, or as commonly required by religious edicts, so the infant might be buried separately from the mother[2]. The first historically recorded case of a mother and a baby surviving CS was in 1500 in Siegersausen, Switzerland, although the accuracy of this record is questionable. The mother lived and subsequently gave birth to five other children by vaginal deliveries including twins[2]. For centuries, CS procedure remained crude, involving vertical incisions in the maternal abdominal wall and uterus, and rarely any attempt to save the mother[3]. In the late 19th century, the introduction of uterine suturing by Max Sänger and the lower uterine segment transverse incision by Ferdinand Kehrer were critical milestones, reducing severe bleeding and infections[4]. There was a dictum, "Once a cesarean, always a cesarean", which was debated[5]. One argument is that subsequent deliveries can take place naturally provided that the operator has avoided sepsis and has repaired uterine wound properly[6]; however, neonatal complications are significantly higher in failed vaginal birth after cesarean[7]. CS was then further improved with the development of the aseptic technique, anesthesia, and surgical tools[3,4]. It was the introduction of antibiotics after the 1940s that dramatically reduced maternal and neonatal mortality from post-operative infections[8,9].

Today, CS is performed both for strict medical indications and, increasingly, as a personal or elective choice[4]. It is generally categorized into elective and emergency procedures based on their timing and circumstances. An elective procedure is planned before labor begins, while an emergency procedure is performed after labor has begun. The principal circumstances that require an emergency CS are maternal (eclampsia and severe pre-eclampsia), obstetrical (uterine rupture, hemorrhage), and fetal (umbilical cord prolapse, severe fetal brady-arrhythmias)[10]. Reasons for an elective CS include both medical indications mentioned above and maternal request. CS on maternal request refers to a primary pre-labor CS performed in the absence of fetal and maternal medical indications, driven by fears, beliefs, and experiences[11–13]. When medically necessary, a CS can effectively prevent maternal and newborn mortality, but when the rate of CS goes above 10%, there's no evidence that mortality rates improve[14]. In fact, CS increases the risk of some pregnancy complications in subsequent pregnancies, including placenta accreta, placenta praevia, placental abruption, and stillbirth[15]. For the child, there's a growing body of research showing that CS is associated with elevated risks of asthma and obesity[16,17]. Therefore, vaginal delivery is recommended by The World Health Organization if CS is not medically necessary[18].

Asthma

Asthma is one of the most common chronic non-communicable diseases that affects over 260 million people and is responsible for over 450,000 deaths each year worldwide[19]. Substantial progress has been made in the understanding and management of patients with asthma in recent years, yet this heterogeneous disease still poses an unacceptable health burden globally[20]. Multiple genetic and environmental factors cause asthma, and possibly their interaction, across the life course[21]. Genetic factors play a substantial but incomplete role in asthma susceptibility, accounting for 40%-70% liability to asthma[22]. Children with one asthmatic parent have 2-3 times increased risk, while those with both parents affected have up to 15-fold increased risk[23]. Recent evidence also reveals that males appear to have a higher risk of asthma than females, but this sex-specific modulation of genetic risk is offset by hormonal and environmental factors that lead to female predominance in adult asthma prevalence[24].

There has been significant development and research into asthma phenotypes, but none has been widely accepted[25]. The broad categories of phenotypes are grouped into three: phenotypes defined by clinical or physiological criteria (Severity-defined, Exacerbation-prone, defined by chronic restriction, Treatment-resistent, Defined by age at onset), phenotypes related to environmental triggers (Aspirin or on-steroidal anti-inflammatory drugs, Environmental allergens, Occupational allergens or irritants, Menses, Exercise), and phenotypes defined by their pathobiology(Eosinophilic, Neutrophilic, Pauci-granulocytic)[26]. In addition to phenotypes, the concept of endotype has been proposed, referring to a subtype of the disease defined functionally and pathologically by a molecular mechanism or by treatment response[27]. According to age-based classification, asthma is divided into childhood asthma and adult asthma. The differences between childhood and adult asthma reflect variations in immune development, hormonal influences, environmental exposures, and disease phenotypes/endotypes[26,28,29]. In this project, we focused on childhood asthma.

Childhood asthma

Phenotypes and endotypes

Different approaches have been used to disentangle asthma heterogeneity, resulting in multiple phenotypes and endotypes, which can be classified into clinical phenotypes, inflammatory endotypes, and molecular ("omic" based) endotypes[30].

The initial classification of childhood asthma phenotypes emerged from observations of wheezing patterns in preschool children, leading to two primary clinical phenotypes: Episodic Viral Wheeze[31] and Multiple-Trigger Wheeze[32]. Among school-age asthmatic children, phenotypes have been differentiated based on allergic sensitization status (Atopy-based classification), including allergic asthma and nonallergic asthma[33]. There's also classification based on responsiveness to standard therapy: Easy-to-Control asthma and Difficult-to-Control asthma[34,35]. Advanced statistical clustering analyses have identified reproducible asthma phenotypes in children based on combinations of atopic burden, airway obstruction, and exacerbation history. Five distinct clusters have been characterized in pediatric populations: Nonallergic with preserved lung function and minimal exacerbations; Atopic with low obstruction and intermediate exacerbations; Atopic with high obstruction and intermediate exacerbations; Moderately atopic with high obstruction and frequent exacerbations; and Highly atopic with high obstruction and frequent exacerbations[36]. The identification of inflammatory endotypes has advanced beyond clinical features to characterize underlying immune mechanisms. Four distinct inflammatory endotypes have been consistently identified across pediatric asthma populations, including Type 2-High[37] and Type 2-Low[38] Asthma. There are still many more phenotypes and endotypes, including obesity-related asthma[39] and genetic and molecular endotypes (e.g., the 17q21 chromosomal region)[40], which underscore the complexity of this disease and emphasize the need for more integrative, multi-omic approaches to unravel its underlying biology.

Risk factors

Childhood asthma is a complex disease arising from the intricate interplay of genetic predisposition, environmental exposures, and developmental factors. Genetic risk factors, such as family history and specific genetic loci, are among the most significant[41,42]. However, genetic susceptibility alone does not predict asthma development. Early-life environmental exposures can interact with genes in asthma pathogenesis. Respiratory viral infections represent significant risk factors for asthma inception, particularly in the context of genetic predisposition and allergic sensitization, such as infections caused by rhinovirus and respiratory syncytial virus[43–45]. Early-life exposures, including air pollution exposure[46,47], tobacco smoke exposure[48,49], indoor allergens[50,51] were reported to increase the risk of childhood asthma, as well as multiple perinatal factors, such as gestational diabetes[52], antibiotics exposure to the child[53] and to the mother[54,55], and CS[56–58].

CS and childhood asthma

CS is reported to increase the risk of asthma in multiple large-scale studies[56]. In a landmark Finnish birth cohort study, the cumulative incidence of asthma at age seven years was significantly higher in children born by CS (4.2%) compared to those delivered vaginally (3.3%), with an adjusted odds ratio of 1.21 (95% Cdonfidence Interval 1.08-1.36)[59]. In our previous work done on the COPSAC2000 cohort, the adjusted hazard ratio for asthma was increased by CS relative to vaginal birth 2.18 (1.27-3.73), and these findings were replicated in the registry data[60]. Also, CS before rupture of membranes carries a significantly higher risk of asthma than CS after rupture of membranes[60,61]. However, there appears to be significant heterogeneity across studies in different regions[62]. A case–control study from Saudi Arabia found no independent association between cesarean section delivery and childhood asthma after adjusting for confounders, and also reported no link between delivery mode and asthma control[63]. A recent meta-analysis of thirty-five cohort studies also concluded that CS was a risk factor for asthma in children from Europe, North America, and Oceania but not in the Asian population[64]. More longitudinal studies are needed to clarify the relationship between CS and childhood asthma across different regions. There is a lot of work exploring the mechanisms of childhood asthma. Possible mechanisms include epigenetic reprogramming of the developing immune system[65,66] and broader innate immune alterations such as changes in dendritic cell programming, Toll-like receptor signaling, and trained immunity[67,68]. When it comes to the influence of CS on childhood asthma, the gut-lung axis has raised broad interest. Researchers are now investigating how the local microbiota influences immunity at distal sites, in particular how the gut microbiota influences lungs[69]. The gut-lung axis represents a bidirectional communication pathway between the gastrointestinal tract and respiratory system, through which microbial communities and their metabolites shape immune responses and influence respiratory health. Ever since it was first proposed in 1991 by Pugin et al.[70], this axis has emerged as a critical concept for understanding how early-life gut microbial perturbations contribute to the development of childhood asthma[71].

Gut environment

Understanding the gut environment is crucial for comprehending how delivery mode affects early-life gut microbiome development and subsequent health outcomes.

pH level

The physicochemical environment of the gastrointestinal tract—particularly its pH profile—plays a central role in shaping early microbial colonization, nutrient metabolism, and mucosal immune development[72]. Along the intestinal tract, pH follows a gradient: values rise from mildly acidic (6.1) in the duodenum to near neutrality in the distal small intestine (7.1-7.5), then drop in the proximal colon (cecum: 6.0, Rectum: 7.0), where microbial fermentation generates short-chain fatty acids. Fecal pH is a measurement of stool acidity, which can reflect the pH of the lower colon and rectum. Breastfed infants typically have lower fecal pH due to the predominance of *Bifidobacterium*. Over the past century, the pH in healthy breastfed infants has increased from 5.0 to 6.5 in association with reduction in *Bifidobacterium*[73]. These changes in pH and microbiome overtime are likely multifactorial in nature and may be related to increased rates of CS, use of infant formulas, and antibiotic exposure[73].

Oxygen level

The neonatal gut environment is distinct from a mature adult gut environment. Notably, oxygen traces can be detected in the neonatal gut immediately after birth, whereas the adult gut is almost entirely hypoxic[74]. This initial oxygen presence creates an environment where the first colonizers—typically facultative anaerobes such as Lactobacillaceae and Enterobacteriaceae—can thrive[75]. Oxygen consumption occurs through both microbial and host-mediated processes. Facultative anaerobes are the primary oxygen consumers in early life, such as *Escherichia coli* and other Enterobacteriaceae, which consume oxygen through aerobic respiration[76]. Colonocyte metabolism and chemical oxidation are different pathways of oxygen consumption. During homeostasis, colonocyte metabolism is directed toward oxidative phosphorylation, thereby increasing epithelial oxygen consumption. Notably, β-oxidation of short-chain fatty acids (SCFAs) accounts for approximately 70% of oxygen consumption in colonocytes[77]. In germ-free animals, the gut lumen maintains anaerobic conditions through chemical oxygen-consuming reactions, particularly lipid oxidation. Oxidation of unsaturated phospholipids and fatty acids in the luminal contents can consume oxygen over several hours, though at substantially slower rates than microbial respiration. This mechanism may be particularly relevant during initial colonization in neonates when bacterial biomass is low[78]. The most luminal aspect of the healthy colon exists at a pO2 of less than 10 mmHg[79], any oxygen present (from blood supply, ingestion) is continuously consumed by microbial and colonocyte consumption and chemical reactions[80].

Mucus layer

Besides the longitudinal pH and oxygen gradient from duodenum to distal colon, there's also a steep transversal gradient existing across the intestinal wall because of the presence of a mucus layer[81,82]. The mucus was initially thought to play solely a protective role by shielding the epithelium from the shear and mechanical stresses imposed by feces. Nowadays, it is recognized that mucus plays additional roles in gut homeostasis[83]. The colon possesses the most well-developed and organized mucus system, featuring a distinct two-layered structure. The inner mucus layer is a dense, stratified, and firmly adherent sheet of highly cross-linked MUC2 polymer that remains attached to goblet cells and is essentially impermeable to bacteria, creating a sterile zone at the epithelial surface[84]. In contrast, the outer mucus layer is a loosely organized, easily detachable, and highly expanded matrix formed as MUC2 undergoes proteolytic cleavage. This outer layer is permeable and serves as the primary niche for commensal microbes, with its thickness and structure varying with microbial degradation activity[85].

A pH gradient exists across the mucus layer from the epithelium to the lumen in the colon, with mean cecal, proximal, and distal colonic mucus pH ranging between 7.3 and 7.5, while the corresponding intraluminal contents have pH values around 6.8[86]. In addition, a dramatic radial oxygen gradient exists from the highly oxygenated submucosa through the epithelium and mucus layer to the anaerobic lumen[79]. From blood vessels to epithelium, the oxygen concentration drops from approximately 3% O_2 (~23 mmHg pO_2) to < 1% O_2 (~7.6 mmHg). The basolateral surface of the epithelium (adjacent to blood vessels) is well oxygenated, while the apical surface of the epithelium (facing the lumen) is kept hypoxic due to high epithelial oxygen consumption[80,81,87]. Across the mucus layers, the inner mucus layer is adjacent to the oxygen-consuming epithelium and thus experiences low but detectable oxygen concentrations (~approximately 15 mmHg (~210 nM))[88], while the outer mucus layer experiences even lower oxygen levels. These gradients are essential for maintaining the appropriate spatial separation between the microbiota and the epithelium while providing ecological niches for beneficial commensals.

CS and gut environment

Being born by CS doesn't change the basic physical structure/anatomy of the gut compared with being born vaginally. However, delivery by CS results in increased intestinal permeability possibly due to the reduction of the expression of tight junction protein related genes[89]. Besides, the initial seeding of the facultative anaerobes in the CS born infants' gut, could have an impact on creasting a different gut environment compared to vaginally born infants in terms of pH, oxygen level.

Although there aren't direct comparative measurements of fecal pH specifically contrasting infants born by CS against those born vaginally, increased CS delivery since the 1980s is recognized as a factor that limits the natural fecal-oral transfer of *Bifidobacterium* from mother to infant typically associated with vaginal delivery, which is reflected in a higher fecal pH[73]. This was also associated with meta-omics signatures characteristic of a microbiome adapted to a more oxygen-rich gut environment, enriched with genes associated with reactive oxygen species metabolism and lipopolysaccharide biosynthesis, and depleted in genes involved in the metabolism of milk carbohydrates[90]. An existing hypothesis suggests that when strict anaerobes are depleted, the resulting lack of SCFAs as an energy source for colonocytes prevents those cells from consuming oxygen via β-oxidation, leading to increased epithelial oxygenation that drives the aerobic expansion of Enterobacteriaceae[90].

First colonizers

Before discussing the differences in the first colonizers, it is essential to establish the current scientific understanding of the prenatal environment. The healthy human fetus develops in a sterile environment, a conclusion now supported by a comprehensive scientific consensus published in Nature in 2023[91]. This was challenged when several studies reported detecting bacterial DNA in the placenta, amniotic fluid, and meconium, suggesting the existence of a "fetal microbiome"[92,93]. However, a comprehensive international consensus involving 46 leading experts in reproductive biology, microbiome science, and immunology unanimously refuted the concept of a fetal microbiome[94]. The team concluded that bacterial signals detected in fetal tissues were due to contamination of samples drawn from the womb during vaginal delivery, clinical procedures, or laboratory analysis[91]. This means that newborns acquire their first microbes at birth.

During vaginal delivery, the newborn acquires microbiota mainly from the maternal fecal microbiomes via vertical transmission[95]. In vaginally delivered infants without antibiotic exposure, approximately 72% of the gut microbial population at 3-7 days after birth was transmitted from their mothers[96]. Another study found that mothers' stool microbiota accounted for 22.1% of overall microbial abundance in the neonate's gut[97]. Advanced metagenomic strain-level analysis demonstrates that specific maternal gut strains are transmitted to infants and remain transcriptionally active in the infant gut[98]. The most frequently transmitted bacterial species from maternal feces to vaginally delivered infants include facultative anaerobes (*Escherichia coli*, *Enterococcus* species, and *Veillonella* species) and obligate anaerobes (*Bacteroides* species, *Bifidobacterium* species, and *Clostridiales* species)[99,100]. The maternal vaginal microbiome undergoes significant compositional changes during pregnancy, becoming more stable and dominated by specific *Lactobacillus* species[101,102]. However, the transmission of *Lactobacillus* to the infant gut is limited, and the relative abundance of ASVs shared with maternal vaginal swabs was only 0% at day 28 in both vaginally- and cesarean-born neonates[97]. Factors affecting vertical transmission include delivery mode, antibiotic exposure, and spontaneous rupture of membranes[103].

Where do they colonize

An ecological niche is the match of a species to a specific environmental condition, including resource utilization (specific nutrients, carbon source, and growth substrates a bacterium can use), environmental conditions (pH ranges, oxygen tension, temperature, spatial location (lumen or mucus layer)), and ecological role (the functional contribution the bacterium makes to the community)[104]. When discussing bacterial competition for colonization niches, we're referring to the ecological processes by which bacteria compete for limited resources, physical space, and environmental conditions.

Competition for resources

Recent research demonstrates that colonization resistance is achieved through nutrient blocking—where the early-established microbiota removes or monopolizes key nutrients, preventing pathogen colonization[105]. Among the early colonizers, there’s nutrient competition, like proline competition among *Escherichia coli*[106], and Human milk oligosaccharide (HMO) utilization among *Bifidobacterium* and other species[107]. Among closely related bacteria that may occupy the same general niche, they can coexist through resource partitioning - using different subsets of available resources[108–110], for example, different *Bacteroides* species in the infant gut have evolved distinct glycoside hydrolase profiles that allow them to degrade and utilize different carbohydrate substrates, reducing direct competition[111]. Interestingly and importantly, there’s a priority effect, and arrival order matters. Priority effects occur when the order of arrival of bacteria into the gut determines which species becomes established, regardless of their inherent competitive ability[112,113]. At the end of the competition, one will inevitably outcompete the other, according to the competitive exclusion principle[114,115].

Physical space

Beyond nutrient competition, physical space and spatial location represent distinct niches. In the colon, the transverse oxygen gradient leads to aetolerant bacteria like Akkermansia that occupy the outer mucus layer, accessing oxygen and specific MUC2 glycans[116,117], while strict anaerobes dominate the luminal and deeper mucosal niches. The longitudinal pH gradient can partition bacteria into distinct niches, with *Pseudomonas aeruginosa* preferring acidic niches and *Staphylococcus aureus* preferring neutral pH[118].

Bacteroicins and Antimicrobial Competition

There are other antimicrobial mechanisms of bacterial competition, including bacteriophages and bacteriocins. *Enterococcus faecalis V583* harbors a composite bacteriophage, ϕV1/7, composed of two prophage elements (prophage 1 and prophage 7), which can infect and lyse closely related *Enterococcus faecalis strains* that lack the prophage[119]. Bacteriocins are the products of the secondary metabolism of many bacterial genera that must display antimicrobial activity[120]. They are ribosomally produced proteins or peptides that display either inhibitory or lytic activity against bacterial cells, whether of the same genus as the producing bacteria or a closely related genus.

Functional profile

Bacteria in the gut microbiome contribute a wide array of essential functions that shape community structure, maintain host health, and enable ecological resilience. These functions include growth-related functions that support bacterial survival and replication, functions that contribute to the community, and functions involved in the host interaction that regulate metabolism, immunity, and physiology. Lipopolysaccharide (LPS) is a major component of the outer membrane of Gram-negative bacteria, consisting of lipid A (endotoxin), core oligosaccharide, and O-antigen polysaccharide. LPS biosynthesis is essential for bacterial cell wall integrity, survival, and host interaction[121,122]. Peptidoglycan (PGN) forms a multi-layer reticular macromolecular structure that protects bacterial cells against stress and maintains tolerance of the external environment. PGN is composed of long polysaccharide chains with repetitive unit structure: N-acetylglucosamine (NAG) and N-acetylmuramic acid (NAM) linked by β-1,4 glycosidic bonds, cross-linked by short peptides[123]. There are also functions involving membrane biogenesis and phospholipid synthesis. Bacterial membranes consist of phospholipids, lipoproteins, and membrane proteins, all of which are essential for cellular compartmentalization, nutrient transport, energy generation, and signal transduction[124,125]. Metatranscriptomic studies reveal that other functions dominate bacterial gene expression in the healthy gut, including ribosomal protein function and translation, amino acid synthesis, energy production and conversion, nucleotide metabolism, and cofactor biosynthesis[126,127].

Nutrients

One of the main carbon and nitrogen sources of the gut microbiota in infants is breastmilk. Human breast milk contains approximately 87-88% water and 12-13% solid components: carbohydrates (60-70 g/L, including lactose (50-70 g/L), HMOs (12-24 g/L), and monosaccharide like glucose and galactose in small amounts), proteins (8-16 g/L, including whey proteins, caseins, and functional proteins (enzymes, growth factors, and immunoglobulins)), fats (35-40 g/L, including triglycerides, essential fatty acids, phospholipids and cholesterol) , and other bioactive components (immunoglobulins, lactoferrin, vitamins and minerals, and living bacteria). The majority of breastmilk nutrients are digested and absorbed in the stomach and small intestine, with only specific components reaching the colon, which were HMOs, small fraction of proteins (secretory IgA, lactoferrin, milk mucins and glycoproteins), small amounts of lipids, and finally living bacteria[128].

Carbon source

HMOs are a family of glycans found in breastmilk with over 200 identified structures. They are indigestible by infants but metabolized by certain gut bacteria[129], including *Bifidobacterium* and *Bacteroides* species[111,130]. *Bifidobacterium longum* subsp. *infantis* is the most efficient HMO degrader, possessing genes for importing and intracellularly degrading a broad range of fucosylated and sialylated HMOs. This species transports HMOs with degree of polymerization ≤7 into cells and hydrolyzes them using dedicated glycosyl hydrolases[131]. *Bifidobacterium bifidum* employs an extracellular degradation strategy, using membrane-anchored glycosidases to partially degrade HMOs outside the cell before importing oligosaccharides and monosaccharides[132]. Other HMO-degrading species include *Bifidobacterium breve* and *Bifidobacterium longum* subsp. *longum*, though their capacity is more limited and often confined to specific structures like lacto-N-tetraose[133]. Beyond bifidobacteria, *Bacteroides* species can also efficiently utilize HMOs via polysaccharide utilization loci (PULs)[134,135]. *Akkermansia muciniphila* strains also demonstrate HMO-utilization capacity, co-opting their mucin-degradation machinery[136].

Some members of the microbiota can also utilize mucin O-glycans as a carbon source. To degrade these host glycans, the bacteria express multiple carbohydrate-active enzymes

(CAZymes), such as glycoside hydrolases, sulfatases, and esterases, which act on specific linkages[137]. *Akkermansia muciniphila* binds to mucins via LacNAc-containing O-glycans after sialic acid removal and degrades mucin glycans using sialidases, fucosidases, and various glycoside hydrolases[116]. *Bacteroides* species possess PULs encoding mucin-degrading enzymes and utilize extracellular hydrolysis followed by oligosaccharide import[138]. *Ruminococcus torques* serves as a keystone mucin degrader, releasing oligosaccharides used by *Bacteroides thetaiotaomicron*[139,140].

Other bacteria, unable to metabolize complex glycans, employ sophisticated cross-feeding strategies to obtain carbon, such as acetate and lactate cross-feeding, monosaccharide sharing, and additional cross-feeding metabolites (including formate, succinate, pyruvate, fucose, and 1,2-propanediol released during glycan fermentation)[141,142]. During the weaning transition, this metabolic network adapts as dietary substrates shift from milk-based to solid foods[143]. Complementary feeding introduces complex polysaccharides, selecting for different bacterial populations with diverse carbohydrate-active enzymes. Species unable to degrade these fibers rely on cross-feeding relationships established during the milk-feeding period, ensuring ecosystem continuity during dietary transitions[144].

Nitrogen source

Nitrogen is an essential nutrient for bacterial growth, supporting core biosynthetic processes that are fundamental to microbial survival and proliferation[145]. The nitrogen source in an infant's gut includes protein and non-protein nitrogen, mainly from breastmilk. Amino acids and peptides represent major nitrogen sources[146,147], and urea constitutes a significant nitrogen source, particularly in breastfed infants[148,149], ammonia (NH_4^+/NH_3) serves as the most readily assimilable nitrogen source and supports the fastest bacterial growth rates in vitro[150].

Bacteria employ diverse, species-specific strategies to acquire nitrogen from available sources in the infant gut, including core ammonia assimilation pathways (the glutamine synthetase-glutamate synthase (GS-GOGAT) pathway and the glutamate dehydrogenase (GDH) pathway)[151,152], urea utilization strategy[148], and direct amino acid uptake and catabolism[146].

CS and first colonizers

First study looking into the impact of CS on the infant gut microbiota was published in 1999 by Grönlund et al., using culture-based observation, which showed that the primary gut flora in CS infants is disturbed for up to 6 months after birth[153]. With the development of modern sequencing technique, Dominguez-Bello et al. proved that CS infants were colonized by skin and hospital environment bacteria (*Staphylococcus*, *Corynebacterium*, *Propionibacterium*) using multiplexed 16S rRNA gene pyrosequencing[154]. Then recent large-scale studies like the TEDDY study[155] and Shao et al.[155], refined the developmental phases of the gut microbiome following CS, and brought interest of the lack of *Bacteroides*. Nowadays, decrease in *Bifidobacterium* and *Bacteroides* in the gut microbiome community following CS has been proved in longitudinal studies[156,157].

Host-interaction

There are several possible mechanisms of host-interaction of the gut microbiota: metabolites, Microbe-Associated Molecular Patterns (MAMPs) and Pattern Recognition Receptors (PRRs), and other mechanisms such as Bacterial Extracellular Vesicles (BEVs)[158].

Metabolites

Recent work has uncovered a central role for bacterial metabolites in orchestrating the host immune response[159]. Major metabolites groups that have been studied include Short-Chain Fatty Acids (SCFAs, acetate, propionate, and butyrate), tryptophan metabolites, and secondary bile acids.

SCFAs are detected through G-protein coupled receptors (GPR) and epigenetic mechanisms[160]. Acetate activates FFAR2 (GPR43), propionate preferentially binds FFAR3 (GPR41), while butyrate engages both FFAR3 and GPR109a[161]. All three SCFAs inhibit histone deacetylases (HDACs), leading to altered gene expression, enhanced regulatory T cell differentiation, and anti-inflammatory effects[162].

Breast milk is the only source of the essential amino acid tryptophan in breast-fed infants[163]. Tryptophan is mainly metabolized through four interconnected routes: the kynurenine pathway, the serotonin pathway, multiple indole pathways, and minor alternative branches (e.g. tryptamine)[164]. Roughly 90–95% of tryptophan enters the host kynurenine pathway (host dominant), a smaller fraction enters the serotonin/tryptamine pathway (host- and microbiota-modulated), and about 5–10% is converted by gut microbes into indoles and related derivatives (microbiota dominant)[165]. Tryptophan metabolites primarily signal through the aryl hydrocarbon receptor (AhR), a ligand-activated transcription factor crucial for intestinal immunity. Indole, indole-3-acetic acid (IAA), indole-3-lactic acid (ILA), and indole-3-propionic acid (IPA) all activate AhR, promoting IL-22 production, enhancing barrier integrity, and inducing anti-inflammatory programs. IPA additionally activates the pregnane X receptor (PXR), contributing to neuroprotection and metabolic regulation[166].

Other microbial metabolites include secondary bile acids[167], vitamins[168], neurotransmitters like GABA and serotonin (5-HT)[169], branched-chain amino acids[170], gaseous metabolites including hydrogen sulfide (H_2S)[171] and nitric oxide (NO)[172,173].

Microbe-associated molecular patterns

Specific pattern recognition receptors recognize microbial-associated molecular patterns[174]. Gut microbial LPS is thought to be one of the most potent activators of innate immune signaling and an essential mediator of the microbiome's influence on host physiology by the activation of Toll-like receptor 4[121]. Peptidoglycan (PGN) fragments are detected by multiple receptors depending on structure, such as Nucleotide oligomerization domain (NOD)-like receptors (NLRs)[175,176]. Flagellin is recognized by Toll-like receptor 5 (TLR5) expressed on intestinal epithelial cells and dendritic cells[177,178].

Other mechanisms

Additional host interaction mechanisms operate through diverse molecular mediators; for example, quorum-sensing molecules enable bacteria-to-bacteria and bacteria-to-host communication[179]. Bacterial extracellular vesicles (BEVs) carry complex cargo, including proteins, nucleic acids, metabolites, and MAMPs that engage multiple host receptors and cell types[180,181].

CS and host-interaction

CS and feeding induced particular differences in the gut metabolome since birth, such as the deficit in the acylcarnitines which was associated with the presence of bacteria such as *Bacteroides vulgatus* and *Parabacteroides merdae*[182]. This suggests the potential influence of delivery mode on fatty acid oxidation[183]. Besides, the altered initial seeding of the bacteria may lead to a potential disruption of the "bacterial-epithelial crosstalk" essential for establishing a functional immune phenotype[184]. Research suggests that only vaginal delivery promotes the production of various cytokines implicated in neonatal immunity[185]. For instance, the prominent commensal bacterium Bacteroides fragilis directs the development of Foxp3+ regulatory T cells (Tregs) that induce anti-inflammatory cytokine production, thereby promoting mucosal tolerance[185].

Restoration

We are losing microbes.

“...as terrible as these resistant pathogens are, the loss of diversity within our microbiome is far more pernicious. Its loss changes development itself, affecting our metabolism, immunity, and cognition. I have called this process the ‘disappearing microbiota’.” - *Blaser, Martin. Missing Microbes: How Killing Bacteria Creates Modern Plagues*

The human gut microbiome was historically much more diverse—evidence from ancient samples shows that modern industrial populations have lost many unique microbial species compared to ancestral populations and non-industrialized societies[186]. This decline is linked to shifts in diet, sanitation, antibiotic use, and the extinction of many bacterial taxa across generations[187]. Consequences of these microbial losses extend beyond gut health: higher colonic pH and reduced microbial diversity in early life are linked to increased risks of childhood asthma, allergies, obesity, and multiple immune-mediated diseases. Restoring the infant microbiome, especially in CS-born infants, has been one of the most attractive solutions to restore beneficial bacterial populations and functional metabolites.

Restoration of the CS gut microbiome

Reviews on early-life microbiome and CS-related dysbiosis list probiotics and prebiotics as established strategies to beneficially modulate infant gut composition and function[188]. The emerging “microbiota transplantation”, including maternal vaginal microbiota seeding and maternal fecal microbiota transplantation (FMT), is an experimental method aiming at mimicking natural vertical transmission. Randomized trials and systematic reviews indicate that pre-, pro-, and synbiotic supplementation in CS-born infants can increase bifidobacterial dominance, alter early colonization patterns, and shift community structure toward that seen in vaginally born infants, though usually without fully restoring all depleted taxa (notably Bacteroides)[189]. A recent RCT of a synbiotic formula (GOS/FOS plus Bifidobacterium breve M-16V) reported dominant and diverse bifidobacterial communities and, in CS infants, restoration of Parabacteroides by 17 weeks and *Bacteroides* spp. by 12 months to levels closer to vaginally born controls, while the prebiotic-only formula had weaker effects[190]. On the other hand, A randomized placebo-controlled trial of oral administration of maternal vaginal microbiota to CS infants found no detectable effect on gut microbiome composition or function at 1 and 3 months, with persistent low Bacteroides and rare maternal vaginal strain engraftment[191]. Other work, including observational and interventional studies using topical swabbing, reports partial restoration of some taxa (e.g. increases in Lactobacillus or enrichment of selected genera) and increased maternal–infant microbial sharing[192], but concludes that vaginal seeding does not fully replicate the vaginally born gut profile, particularly for Bacteroides, and its benefits and safety remain uncertain[193]. A proof-of-concept trial in 7 CS-born infants receiving maternal FMT in breast milk showed that their microbiota trajectory rapidly shifted away from the CS profile and closely resembled that of vaginally born infants, including restoration of Bacteroidales and Bifidobacteriales and increased diversity, with no reported short-term adverse events[194]. Maternal FMT was reported to have the most dramatic effect on microbiota composition, shifting it entirely toward that of vaginally born infants[195].

Several environmental factors have been reported to promote maturation of the gut microbiome, including breastfeeding[196], a rural living environment[197], and having older siblings at home[198]. Breastfeeding is recognized as one of the most influential drivers of gut microbiome composition during infancy, which provides a dynamic source of nutrition that delivers live microbes, immunoglobulins, and bioactive compounds essential for shaping the infant gut microbiome[199]. Besides, Infants from rural areas often exhibit a more diverse gut microbiome and early colonization of Bacteroides, which promotes further microbiome maturation[200]. Lastly, infants with older siblings tend to have a more mature gut microbiota by the age of one year, and this maturity is linked to a higher diversity of gut bacteria, which potentially mediates the protective effects of siblings in relation to allergies[201].

Aim and objectives

The overall objectives of this thesis were to elucidate how CS delivery influences the early-life gut microbiome development and its core functional capacity, and to investigate how persistent CS-associated microbiome perturbations contribute to the risk of childhood asthma. This research sought to provide insight into the restoration process of a perturbed gut microbiome from early life and to evaluate potential pathways by which the CS-associated microbiome is linked with later risk of asthma, with the ultimate perspective of reducing the asthma risk in infants born by CS.

Paper I - Key bacteria and environmental factors associated with the restoration of the infant gut microbiome after cesarean section.
The objective of this paper was to describe early-life environmental and microbial predictors and their interplay in the 1-year gut microbiome restoration.

Paper II - Early gut microbiome composition and functional pathways linking cesarean birth to asthma susceptibility.
The objective of this paper was to profile the taxonomic and functional development of the gut microbiome at 1 month and 1 year of age, to investigate how CS affects gut microbiome composition and its pathways, and to associate CS microbial and pathway signatures with asthma risk.

Paper III - Microbially-derived tryptophan metabolites in the infant gut mediate the increased risk of asthma following cesarean section.
The objective of this paper was to replicate and expand the association between prolonged CS-induced perturbation of the gut microbiome and the development of childhood asthma; furthermore, to associate CS microbial signatures with metabolomic profiles.

Materials and Methods

In this section, I will present an overview of the two cohorts included in the three papers comprising this thesis: the COPSAC$_{2010}$ prospective mother-child cohort and the CHILD birth cohort. We analyzed 16S rRNA gene sequencing data, shotgun metagenomic sequencing data, and metabolomics data. I will describe the study population, asthma diagnosis, fecal sample collection and processing, sequencing procedures, and bioinformatics workflow used for taxonomic and functional annotation, followed by the statistical methods common to all three papers. For further methodological details, please refer to each individual paper.

The COPSAC$_{2010}$ cohort

Study population

The COPSAC$_{2010}$ cohort is a population-based mother-child cohort of 700 children and their families, recruited in week 24 of pregnancy and followed prospectively at the COPSAC research unit during the years 2008-2011[202]. The children were followed by COPSAC study physicians and nurses collecting all biosamples, clinical measurements and diagnoses during clinical visits scheduled at 1 week, 1, 3, 6, 12, 18, 24, 30, and 36 months, thereafter yearly until the age of 6 and again at age 8 and 10 years. The study was conducted in accordance with the guiding principles of the Declaration of Helsinki and was approved by the Local Ethics Committee (H-B-2008-093) and the Danish Data Protection Agency (2015-41-3696). Both parents gave oral and written informed consent before enrolment.

Asthma diagnosis

In the COPSAC$_{2010}$ cohort[202], all children were examined in the clinic at episodes of troublesome lower airway symptoms, if symptoms were significant and the parents found the child needed medical attention. 'Recurrent wheeze' was diagnosed based on (1) recurrent episodes of troublesome lower airway symptoms recorded in the daily dairy cards as 5 episodes within 6 months, each episode lasting at least 3 consecutive days; (2) multi-trigger wheeze[203] including wheezing, breathlessness, and recurrent cough outside common cold; and (3) need for intermittent rescue use of inhaled β2-agonist. An asthma diagnosis was given if a child fulfilled the criteria for recurrent wheeze, showed a response to a 3-month course of inhaled corticosteroids, and had a relapse of symptoms when stopping treatment[204].

Study outcome

Papers I and **II** are based on data from the COPSAC$_{2010}$ cohort, but with different outcomes and sequencing layers. Paper I is a follow-up study of our prior study[157], investigating the gut microbiome development and asthma risk after birth by CS in the COPSAC$_{2010}$ mother-child cohort. In **Paper I**, we focused on the association between early environmental factors and fecal microbiota composition at 1 week and 1 month, using a 1-year restoration score as the outcome measure. This restoration score was first introduced in our previous paper as a CS score, calculated by constructing a cross-validated sparse partial least squares (sPLS) model (described below) on gut microbial composition at 1 year of age predicting delivery mode (vaginal/CS)[157]. The model identified the gut microbial composition at 1 year of age most associated with CS delivery. By reversing the CS microbial score, we define this restoration score as our outcome measure to characterize the restoration of 1-year gut microbiome. Thus, a higher restoration score means a child's microbiome resembles being born vaginally, while a low restoration score means that a child retains a CS-like gut microbiota composition. This score was used as either a continuous score or dichotomized (above and below median value). In **Paper II**, the outcome is persistent wheeze/asthma from 1 to 10 years of age[205,206], we focused on the CS-associated gut microbiome and pathway signatures and their associations with the asthma risk up to 10 years of age.

Fecal samples collection and sequencing

In **Paper I**, all children with fecal samples collected and characterized by 16S rRNA sequencing (requiring at least 2000 reads) at any of the three time points of 1 week (n = 552), 1 month (n = 607), and 1 year (n = 625) were included in the analyses. Fecal samples were collected either at the research clinic or by the parents at home using detailed instructions. Each sample arrived at the laboratory within 24 hours and was mixed on arrival with 1 ml of 10% (v/v) glycerol broth (SSI, Copenhagen, Denmark) and frozen at −80°C. DNA was extracted using the PowerMag Soil DNA Isolation Kit (MO-BIO Laboratories, Inc., Carlsberg, CA, USA) on an epMotion 5075 (Eppendorf), amplified using a two-step polymerase chain reaction (PCR) with 515F and 806R primers flanking the V4 region of 16S rRNA gene, and sequenced using the v2 kit (paired-end 250–base pair reads) on the MiSeq platform (Illumina Inc., San Diego, CA). A full description of the laboratory workflow has been described previously[207].

Raw fastq files were demultiplexed using the MiSeq controller software prior to downstream analysis. As described in previous study[208], the primers and adaptors in sequencing reads were removed using Cutadapt[209]. The determination of amplicon sequence variants (ASVs) was performed on QIIME2 Core 2020.11 platform[210] using Amplicon Denoising Algorithm 2 (DADA2) analysis pipelines[211]. The resulting ASV sequences were annotated using the AnnotIEM[212] pipeline (v.1.3), which combines sequence alignment against four databases: EzBioCloud[213] (r. 2018-05), NCBI[214] (v. refseq 202), RDP[215] (v.11.5), and Silva[216] (v. 138SSU) followed by a high confidence selection of best probable annotation. Genus annotations were correctly annotated for all genera in the mock community, and 15/20 was annotated to the correct species. Therefore, species annotations should be considered putative.

In **Paper II**, children with fecal samples collected at 1 month (n = 600) and 1 year of age (n = 660) and characterized by metagenomics sequencing were included in the analyses. The storage and transportation was the same as described above. DNA was extracted using the PowerMag Soil DNA Isolation Kit (Qiagen) and NucleoSpin Stool Kit (Macherey-Nagel). No differences were found between the different DNA extraction kits after comparison. Before library preparation, the DNA was quantified by Tecan Infinite F Nano+ Plate Reader using Quant-iT dsDNA BR Assay Kit. The enzymatic fragmentation of DNA and library construction was conducted by Tecan DreamPrep NGS using Celero EZ DNA-seq Core Module Kit. The fragmented DNA was amplified using polymerase chain reaction (PCR). Short and large DNA fragments were removed using double-sided magnetic bead size selection (AMPure XP, Beckman Coulter). Adapter sequences from Celero 96-Plex Adaptor Plate were added to each sample during library construction. The final concentration for each library was quantified by Tecan Infinite F Nano+ Plate Reader using NuQuant NGS Library Quantification Module and Qubit. The final fragment distribution is evaluated using a Fragment Analyzer 5200 (Agilent). Qubit and TapeStation were used to determine the concentration of the final library before sequencing. The library was sequenced using 2 × 150 bp paired-end sequencing on an Illumina NovaSeq 6000 platform (Illumina, San Diego, CA, USA). The sequencing depth and batch were adjusted for in our statistical analysis.

Quality control of raw FASTQ files was performed using KneadData (v. 0.6.1) to remove low-quality bases and reads derived from the host genome as follows: Using Trimmomatic (v. 0.36), the reads were quality trimmed by removing Nextera adapters, leading and trailing bases with a Phred score below 20, and trailing bases in which the Phred score over a window of size 4 drops below 20. Trimmed reads shorter than 100 bases were discarded as low-quality reads. Reads that mapped to the human reference genome GRCh38 (with Bowtie2 v. 0.2.3.2 using default settings)[217] were also discarded. Read pairs in which both reads passed filtering were retained; these were classified as high-quality non-host (HQNH) reads.

Taxonomic composition of metagenomes was profiled using MetaPhlAn v4.1[218] on quality-controlled reads. MetaPhlAn was run with default parameters using the species-level genomic bin (SGB) markergene database mpa_vJun23_CHOCOPhlAnSGB_202403 (ChocoPhlAn SGB database vJun23), yielding relative abundances of microbial taxa. For descriptive analysis, taxa were retained if their prevalence >5%; For differential abundance analyses, relative abundances were log-transformed internally by the DA.lli2 function after filtering by prevalence >5% and mean relative abundance >0.01%; For sPLS model training and testing, relative abundances were center-log-ratio (CLR) transformed after filtering by prevalence >5% and mean relative abundance >0.01% and applied a pseudocount of min(relative abundance)/2.

Microbial functional profiling was performed using HUMAnN v3.9[219]. HUMAnN's prescreen step (prescreen threshold 0.01) used the MetaPhlAn v4.1 taxonomic profiles to build a sample-specific nucleotide database from the HUMAnN ChocoPhlAn pangenome collection (v201901_v31). Reads were first aligned to this database with Bowtie2 v2.2 using the *--very-sensitive* preset. Reads not mapped at the nucleotide level were searched in translated mode against the UniRef90 protein database using DIAMOND v2.0.15 (*--top 1 --outfmt 6*, e-value ≤ 1.0, minimum amino-acid identity 80%, and minimum subject and query coverage of 50% and 90%, respectively). Gene family abundances were mapped to MetaCyc pathways (MetaCyc v24), and pathway abundances were inferred using MinPath with gap-filing enabled (MinPath on, gap fill on, Xipe off). HUMAnN outputs for gene families and pathways were then normalized to copies per million (CPM) using *humann_renorm_table (units = cpm)* and merged across samples with *humann_join_tables*, so that each feature's abundances reflects its estimated copy number per million total microbial copies in each sample. For descriptive analyses, pathways were retained if their prevalence >5%; For differential abundance analyses, relative abundances were log-transformed internally by the DA.lli2 function after filtering by prevalence >5% ; For sPLS model training and testing, relative abundances were CLR transformed after filtering by prevalence >5% and applied a pseudocount of min(relative abundance)/2.

The CHILD cohort

Study population

In **Paper I** and **Paper III**, we based our results on 16S rRNA and metagenomics data in the CHILD cohort, respectively.
The CHILD study is a prospective longitudinal birth cohort study, which enrolled 3,405 subjects since pregnancy from 4 largely urban study centers across Canada (Vancouver, Edmonton, Winnipeg, and Toronto) from 2008 to 2012[220]. All children with fecal samples collected and characterized by 16s rRNA sequencing at around 1 year visit (9 months to 12 months, n = 325) were included in the **Paper I**. A subsample of 1,479 children having shotgun metagenomic data processed from fecal samples collected at the 3-month visit (n=1,422) and the 12-month visit (n=1,426) were included in the **Paper III**. Questionnaires related to environmental exposures, psychosocial stresses, nutrition and general health were administered at recruitment, prenatally, at 3, 6, 12, 18, 24, 30 months, and at 3, 4, and 5 years.

Asthma diagnosis

In the CHILD study, childhood asthma was diagnosed (as Yes/Possible/No) by an expert study physician at the clinical assessment at the age of 5 years based on published approach[220]. For **Paper I** and **Paper II**, children were considered to have asthma only if the response was ‘Yes’ and the asthma phenotype was defined as comparing children with asthma at 5 years versus children without asthma at 5 years, children diagnosed as “possible” were excluded.

Study outcome

In **Paper I**, the outcome of the validation study on the CHILD cohort was *CHILD 1-year restoration score*. The validation study was conducted by applying the sPLS model (described below) from our previous study[157] to the CHILD 1-year gut microbiome. In the previous work, the sPLS model was trained on 1-year gut microbiome in the COPSAC$_{2010}$ cohort and selected 5 genera to predict if the child was born by CS or not (1/0). Those taxa were: *Bacteroides*, *Family Enterobacteriaceae*, *Sutterella*, *Escherichia/Shigella*, and *Dysgonomonas*. We found the same or the closest annotation in the CHILD cohort as the input of the model: *Bacteroides*, *Family Enterobacteriaceae*, *Sutterella*, *Escherichia/Shigella*, and *Dysgonomonas*. The preprocessing of the abundance was the same as described in the original work, where we agglomerated the rank into genus level, log transformed the relative abundance, using half the lowest nonzero value as a pseudocount. We then used the predict() function from mixOmics package to apply the model on these five genera, and output the CHILD 1-year restoration score. We also tested the association between this CHILD 1-year restoration score and asthma risk at 5 years of age in the Paper I to validate the same association as observed in the COPSAC study. In the **Paper III**, the main outcome is the asthma risk at 5 years of age, we focused on the association between the CS-associated gut microbiome and asthma risk, and investigated the role of gut metabolites in the association between the CS-associated gut microbiome and asthma risk.

Fecal samples collection and sequencing

The fecal sample collection and sequencing in the CHILD study were comparable to the COPSAC cohort. In the **Paper I**, a soiled diaper was provided on the same day for infant stool collection. Samples were refrigerated at home for up to 24 hours before being collected and processed by study staff. An additional infant stool sample was provided at the 1-year clinical assessment. DNA was extracted from fecal samples using the commercial kits (Qiagen Mo Bio PowerSoil) optimized for the Thermofisher KingFisher® robot. The V4 hypervariable region of the 16S rRNA gene of fecal DNA was amplified by PCR using universal bacterial primers (V4-515f: V4-806r). Pooled PCR amplicons were subjected to paired-end sequencing on the Illumina MiSeq platform. Using VSEARCH and Deblur[221] within the QIIME2 pipeline[210], forward and reverse demultiplexed reads were assembled for a final length of 247 bp (unassembled sequences were discarded) and filtered against the GREENGENES reference database ((v13·8)[222,223]. Taxonomic classification of the resulting unique amplicon sequence variants (ASVs) was achieved using a naïve Bayes classifier trained on reference reads extracted from the reference database at 97% sequence similarity.

In **Paper III**, metagenomic sequencing was performed. Fecal sample collection, storage and transportaion were described as above. Shotgun metagenomic sequencing data with an average depth of 5 million reads per sample was generated by Diversigen (Minneapolis, MN, USA) from fecal samples. DNA was extracted from faecal samples using the MO Bio PowerSoil Pro with bead beating in 0.1mm glass bead plates. The amount of high-quality input DNA was measured using Quant-iT Picogreen. Libraries were then prepared and sequenced on an Illumina NextSeq using single-end 1 x 150 reads. Low quality (Q-Score<30) and length (<50) sequences were removed, and adapter sequences trimmed. Host and low-quality reads were removed, and only samples with at least 100,000 remaining reads or more were retained for downstream analysis. Shotgun metagenomic reads were mapped using the bioBakery 3 pipeline to identify taxonomic (species and strain level) and functional features within each sample. The taxonomic annotation was obtained by mapping reads to a customized MetaPhlAn database (https://github.com/biobakery/MetaPhlAn/wiki/MetaPhlAn-4)[224], and agglomerated to the species level in the analyses.

NMR and LC-MS/MS metabolite quantification

In **Paper III**, we investigated the role of gut metabolome in the association between the CS-associated gut microbiome and asthma risk. Metabolic profiles were created from the same sequenced stool samples at The Metabolomics Innovation Center (TMIC) in Edmonton, Alberta using two separate assays. Targeted nuclear magnetic resonance (NMR) analysis of 31 metabolites was performed across 62 batches. Targeted liquid chromatography with tandem mass spectrometry (LC-MS/MS) analysis of 590 metabolites was performed TMIC's Microbiome Metabolism (MEGA) assay across 27 batches. NMR and LC-MS/MS precision were confirmed to be <5 and < 10% coefficients of variability (CV), respectively. Additionally, overlapping metabolites detected by both methods were cross-checked to confirm the accuracies of the reported concentrations values. Detailed methods of both NMR and LC-MS/MS analyses were described in previous work[225]. The superclass information of metabolites were retrieved from HMDB (Human Metabolome Database) using R package hmdbQuery (v1.28.0).

Statistical analysis

Within-sample and between-sample diversity

In **Paper I**, **II**, and **III**, we quantified within-sample microbial diversity (alpha-diversity) using two commonly applied metrics: Observed richness and Shannon diversity. Richness reflects the number of distinct taxa observed in a sample, and the Shannon diversity index describes both richness and evenness of relative abundances within a sample[226]. In **Paper I**, we calculated the observed richness and Shannon diversity at 1 week and 1 month of age and assessed associations between early-life alpha diversity and 1-year restoration scores using linear regression. In **Paper II**, we calculated observed richness for both bacterial species and functional pathways and compared richness across delivery modes and relevant subgroups. In **Paper III**, we compared observed richness and Shannon diversity between delivery modes and subgroups in the CHILD cohort.

We then assessed between-sample diversity (beta-diversity) by calculating pairwise distances between samples based on compositional dissimilarity. As distance metrics can be defined in multiple ways, we used weighted UniFrac distance[227] in Paper I and Bray-Curtis dissimilarity in Paper II and III. Differences in community composition were visualized using Principal Coordinate Analysis (PCoA). In **Paper I**, we tested compositional differences between infants with high versus low 1-year restoration scores (above/below the median). In **Paper II**, we examined compositional differences based on both gut microbiome and microbial core pathways profiles. In **Paper III**, we tested the compositional differences between delivery modes and relevant subgroups.

Differential abundance analysis

We used differential abundance analysis to identify microbial taxa, functional pathways, or metabolites whose relative abundances differed systematically between exposures or outcomes of interest. We used different R packages in papers, but the underlying analytical framework was consistent. Relative abundances were log-transformed to improve normality, and associations were assessed using linear regression models with relevant covariate adjustment. In **Paper I**, differential abundance analysis was used to identity bacterial species at 1 week and 1 month of age that were associated with 1-year restoration scores. In **Paper II**, we identified bacterial species and microbial core pathways associated with delivery mode. In **Paper III**, we identified bacterial species associated with delivery mode and further assessed associations between gut metabolites and 1-year CS microbial scores.

Sparse Partial Least Squares (sPLS) model

Different from the differential abundance analysis, which tests the associations individually, Partial Least Squares (PLS) is a multivariable approach used to explore and explain the relationship between two datasets by calculating latent components that maximise covariance[228]. This method is designed to reduce the high dimensionality of data and perform simultaneous variable selection. In sparse PLS (sPLS), lasso penalisation is applied on the loading vectors to identify the most important variables[229]. In All three papers, we used sPLS models to identify microbial species or microbial core pathways that were most strongly associated with outcomes of interest. To optimize model complexity and limit overfitting, we selected the optimum number of components and the number of input variables using 11-repeated 10-fold cross-validation of the correlation statistic for continuous scores or area under the ROC curve (AUC) for binary outcomes. The median-performing repeat was used to improve stability. Predictions were retrieved as outcome-associated scores to perform further analyses.

In **Paper I**, we performed sPLS models on gut microbiome and environmental factors separately and jointly to predict the 1-year restoration scores, to reveal microbial taxa and environmental factors most descriptive of 1-year restoration score. We used spearman correlation between the predictions (1-week bacterial scores) and 1-year restoration scores to assess the model performance. 1-week bacterial scores were further used in the following analysis.

In **Paper II**, sPLS models were performed on the gut microbiome and microbial core pathways at both 1 month and 1 year of age predicting delivery mode (vaginal/CS) to reveal microbial taxa and core pathways most associated with CS. AUC was used to assess the model performance. Predictions from these four models were further associated with asthma risk up to 10 years of age, which were 1-month CS microbial scores, 1-month CS core pathway scores, 1-year CS microbial scores, and 1-year CS core pathway scores.

In **Paper III**, sPLS models were performed on the gut microbiome at 3 months and 1 year of age in the CHILD cohort to predict delivery mode (vaginal/CS) to reveal microbial taxa most descriptive of CS. 3-month and 1-year CS microbial scores were retrieved to associated with asthma risk at 5 years of age. 1-year CS microbial scores were further associated with gut metabolites to investigate the metabolomics profile associated with CS-associated gut microbiome.

Associations with asthma risk

In **Paper I** and **Paper III**, we used logistic regression to assess associations between exposures and binary outcomes, such as asthma (yes/no). Logistic regression models the log-odds of the outcome as a linear function of one or more predictors for estimating adjusted associations when controlling for potential confounders.

In Paper II, generalized estimating equitions (GEE) models were used to analyze asthma prevalence assessed repeatedly over time within the same child. GEE extends generalized linear models to clustered or longitudinal data by accounting for within-subject correlation through a working correlation structure, yielding population-averaged effect estimates with robust standard errors.

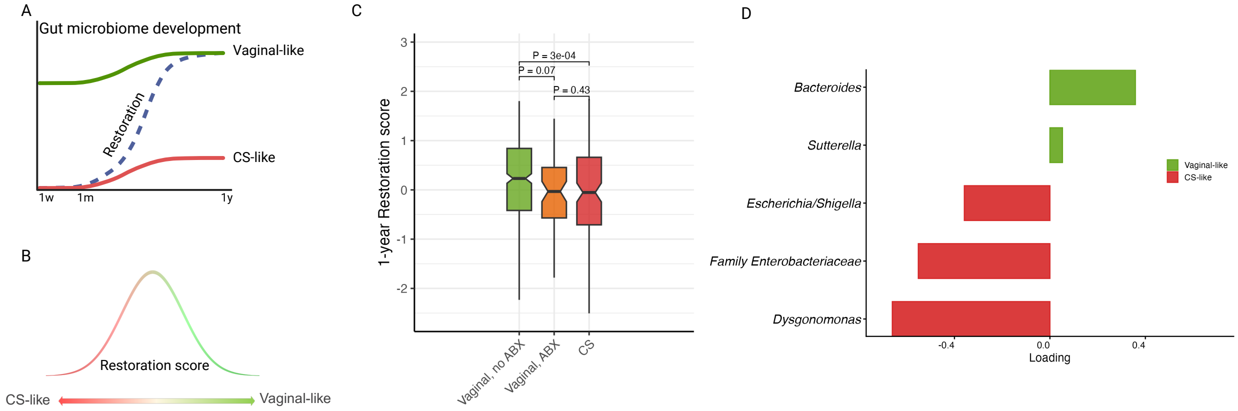
Results are reported as odds rations (ORs) with 95% confidence intervals. The OR is the ratio of the odds of the outcome in the exposed group to the odds in the reference group. An OR>1 indicates higher odds of the outcome associated with the exposure, an OR<1 indicates lower odds, and an OR=1 indicates no association; confidence intervals that exclude 1 suggest statistical evidence of an association at the chosen significance level.

Results

In this section, I will summarize the results obtained in each of the three papers included in the thesis, including their main figures and tables. For further details, please refer to the full papers, which are included as appendices.

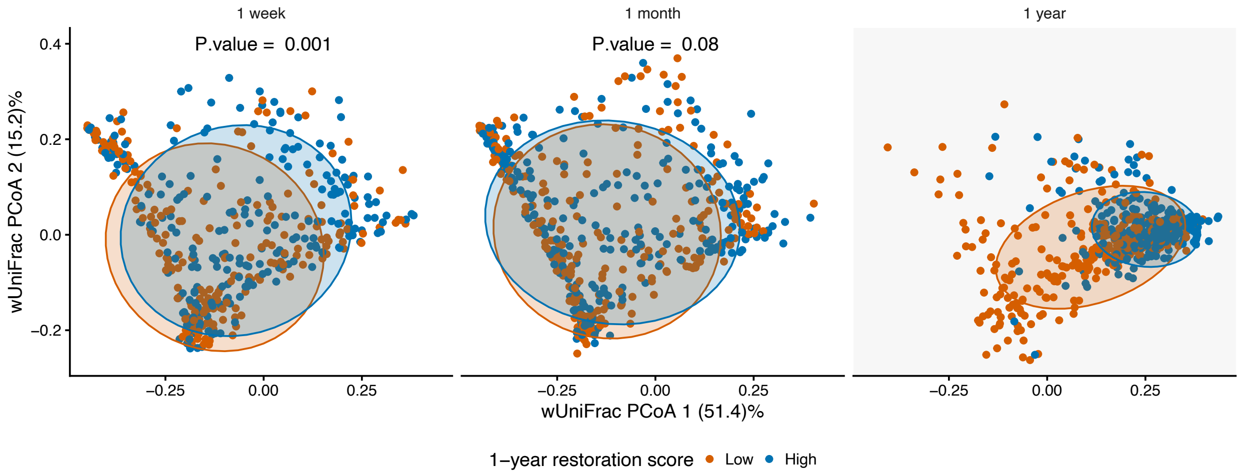
Paper I: Key bacteria and environmental factors associated with the restoration of the infant gut microbiome after cesarean section

In the first paper, we examined the early gut microbiome at 1 week and 1 month, and environmental factors as predictors for the 1-year restoration scores. Based on the prior work, we developed a restoration score quantifying the degree to which the 1-year gut microbiome resembled that of vaginally delivered infants - a higher score indicates that a child's 1-year gut microbiome resembles that of vaginally delivered infants (**Paper I, Fig 1**). Compared to vaginal delivery, children born by CS had a significant lower restoration score (Estimate -0.32, 95% Confidence Interval (CI) [-0.51,-0.13], P=8e-04). In the subgroup analysis (**Paper I, Fig 1C**), children born by CS had a lower restoration score compared to vaginal delivery without antibiotics exposure (-0.36[-0.55,-0.16], P=3e-04). Notably, children born by vaginal delivery whose mother received antibiotics at birth also had a lower restoration score compared to those without antibiotics exposure (-0.23[-0.48, 0.02], P=0.07), and with a score comparable to the CS-born infants (0.13[-0.19,0.45], P=0.43). There was no detectable differences observed between the planned and emergency procedures (0.18[-0.20,0.57], P=0.35).



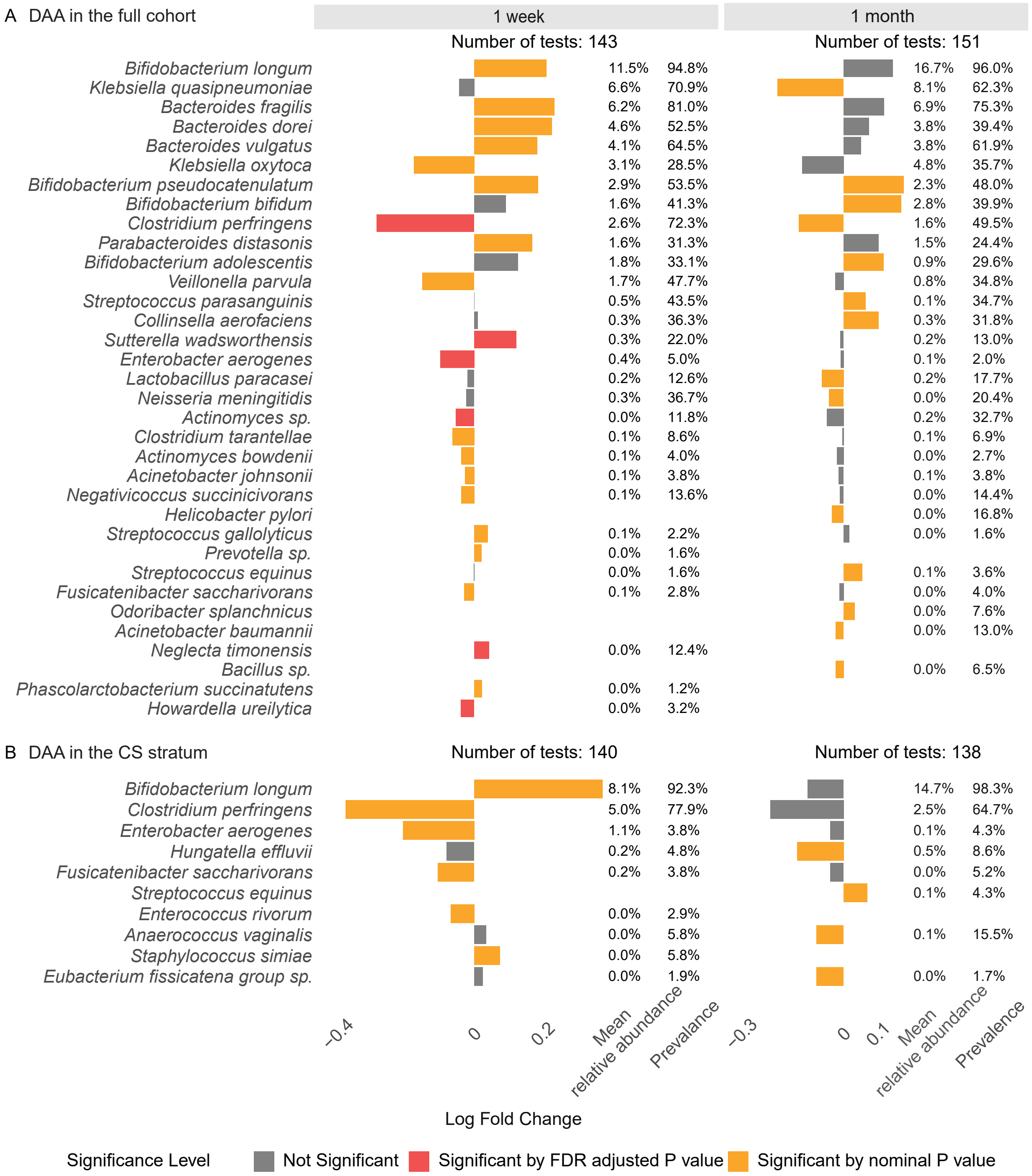
Paper I, Fig 1: Conceptual design of a restoration score. (A) illustrates the hypothesis that a CS-perturbed microbiome composition can be influenced by environmental exposures or ecological interactions and be restored towards a vaginal-like microbiome composition. (B) shows that the restoration score in this study describes the status of 1-year gut microbiome, a higher restoration score means a child's microbiome resembles being born vaginally, while a low restoration score means that a child's microbiome resembles being born by CS[157] (See also Methods). (C) Boxplot of 1-year restoration score according to different groups, where red represents CS delivery, orange represents vaginal delivery with antibiotic treatment, and green vaginal delivery without antibiotic treatment, indicating that while CS born children have lower scores at 1 year of age, vaginally born children could possibly have a CS-like gut microbiome due to the perturbations like antibiotics treatment at birth. P values were derived from linear models comparing the restoration scores among these three groups. (D) Bacteria contributing to the 1-year restoration score. Loadings are derived from the sPLS model trained on the COPSAC2010 1-year gut microbiome in the previous work[157]. Negative/Positive loading corresponds to lower/higher genus abundance in vaginally delivered children compared to CS born children.

We first examined the beta diversity of the gut microbiome at 1 week and 1 month to obtain an overview of the association between overall composition of the early-life gut microbiome and 1-year restoration score. The gut microbiota composition differed between children with high and low 1-year restoration scores (above vs below median values) (**Paper I, Fig 2**) using PERMANOVA analysis. From the PCoA plots ordinated across all three timepoints, children with low 1-year restoration scores appeared to have different compositions compared to those with high 1-year restoration scores. The compositional difference of having a high vs low restoration score was most obvious at age 1-week (F=4.9, R2=0.9%, P=0.001) compared to age 1-month (F=2.2, R2=0.4%, P=0.08).



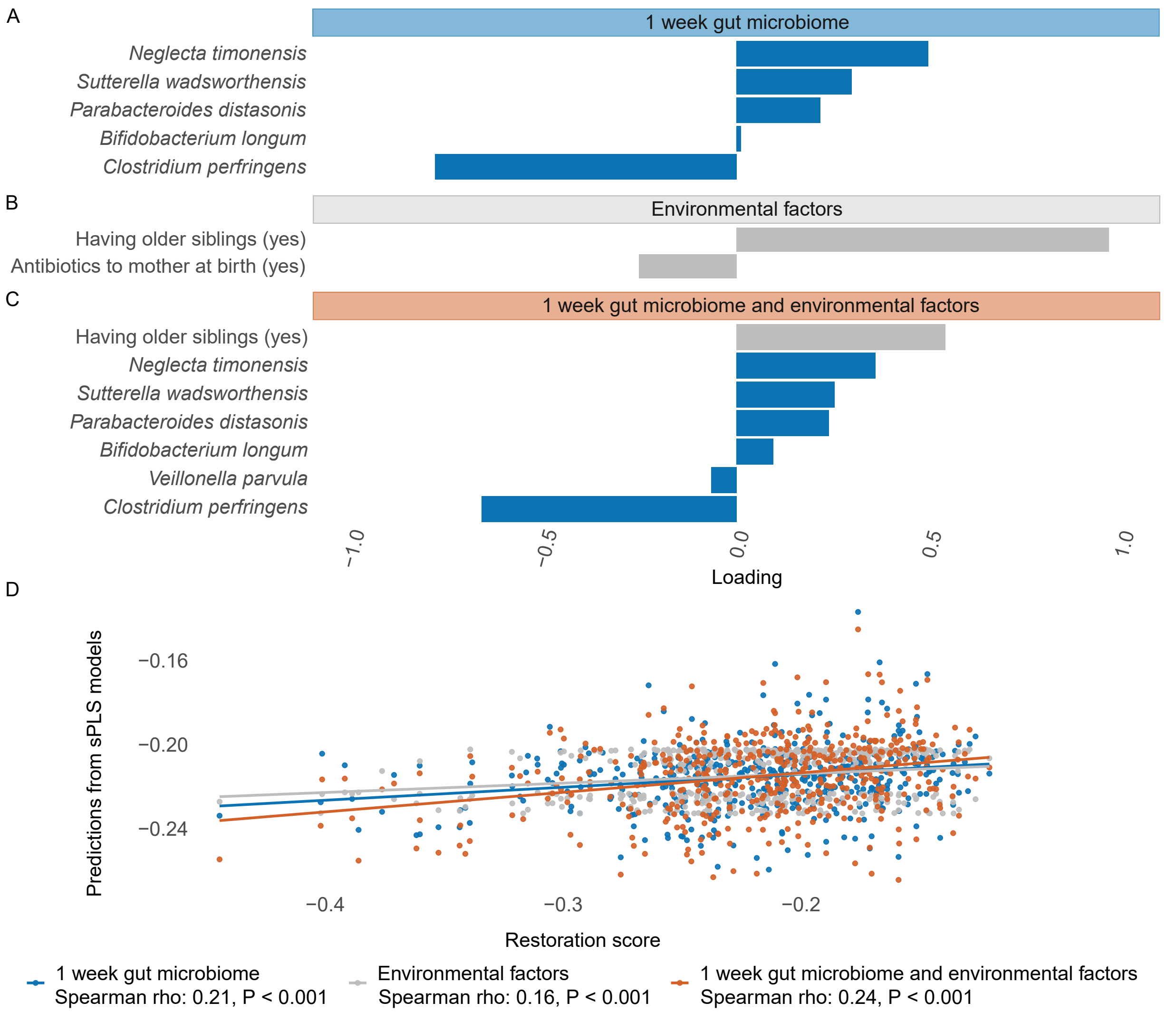
Paper I, Fig 2: Comparison of the fecal microbiota between children having a low or high 1-year restoration score in the full cohort. Weighted UniFrac distances were used as input in principal coordinate analysis (PCoA) plots and colored according to children having high restoration scores (above median) (orange, n = 246, 285, 312 at 1 week, 1 month and 1 year) and low restoration scores (below median) (blue, n = 253, 269, 312 at 1 week, 1 month and 1 year). Gray shading at the 1-year time point marks when the restoration score was derived, hence groups are different by design. PERMANOVA was used to compare group differences. Each dot represents one fecal sample. Ellipses represent 1 standard deviation (1 SD), encompassing approximately 68% of the data points, assuming a bivariate normal distribution.

We next investigated individual taxa at 1 week and 1 month to determine which members of the microbiota were main drivers of the observed compositional differences using the LIMMA model (**Paper I, Fig 3**). At 1 week of age, children with higher restoration scores had higher relative abundances of *Sutterella wadsworthensis* and *Neglecta timonensis* (**Paper I, Fig 3A**, FDR adjusted *P*<0.05). On the other hand, we found that *Clostridium perfringens* at 1 week was negatively associated with the restoration score and was the most differentially abundant species found, followed by *Enterobacter aerogenes* and *Actinomyces sp.* (FDR adjusted *P<0.05*). Compared to the gut microbiome at 1 week of age, the 1-month gut microbiome was less associated with the 1-year restoration score. We observed associations between the relative abundances of species at 1 month and restoration scores in the same direction as at 1 week, but they were not FDR significant. In the CS stratum, higher relative abundances of Bifidobacterium longum and Staphylococcus simiae at 1 week and only Streptococcus equinus at 1 month were associated with higher restoration score at 1 year (**Paper I, Fig 3B**, P<0.05), but none were FDR significant.



Paper I, Fig 3: Differential abundance analysis on the species associated with 1-year restoration score. In the full cohort (A) and in the CS stratum (B). Species with prevalence of at least 0.1% and relative abundance of more than 0.01% of the total were eligible. The analysis in the full cohort was adjusted for delivery mode. The species represented by the red bars were significant after FDR-adjustment, while species represented by yellow bars were nominally significant, and species represented by grey bars were not significant. A positive log fold change value indicates an increase in the abundance of the species as the 1-year restoration score increases, while a negative fold change indicates a decrease.

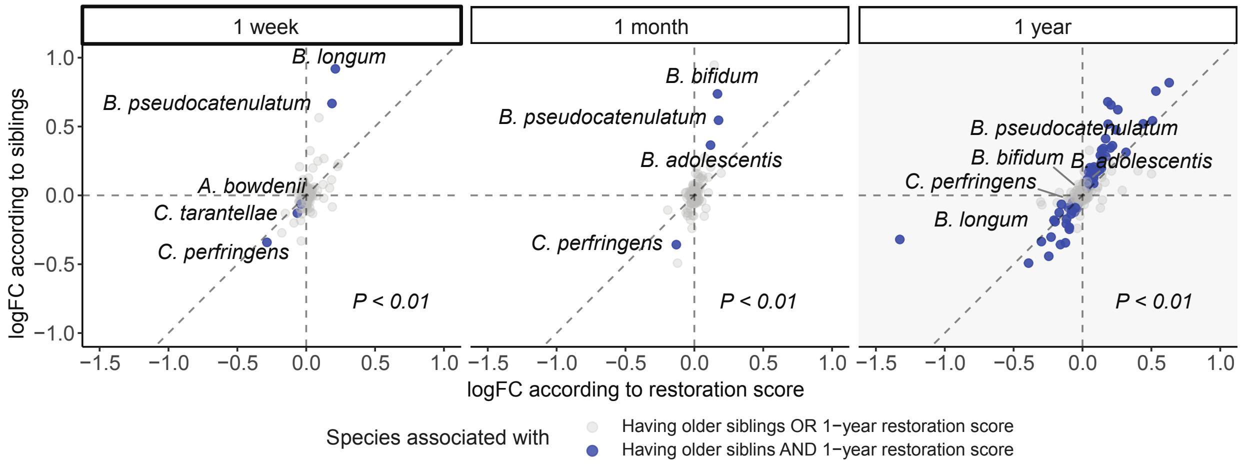
We then used multivariable models to identify species and environmental factors jointly associated with the 1-year restoration score (**Paper I, Fig 4**). We employed three sparse Partial Least Squares (sPLS) models on only the 1-week gut microbiome, only environmental factors, and combined 1-week gut microbiome and environmental factors as input features, respectively. These models were evaluated based on the correlations between their cross-validated predictions and 1-year restoration scores (model on only gut microbiome, Spearman rho 0.21, P<0.001; model on only environmental factors, 0.16, *P*<0.001; model on combined gut microbiome and environmental factors, 0.24, P<0.001). The gut microbiome at 1 week was enriched for *Bifidobacterium longum*, *Parabacteroides distasonis*, *Sutterella wadsworthensis*, and *Neglecta timonensis* in children with higher restoration scores at 1 year, while *Clostridium* perfringens was depleted. Of these, *Bifidobacterium longum* was the most abundant and prevalent (see **Paper 1, Fig 3**), while *Parabacteroides distasonis* and *Sutterella wadsworthensis* were less so. *Neglecta timonensis* was only present in few samples (prevalence 12.4%), but had a strong association with the 1-year restoration score. *Clostridium perfringens* contributed the most to the model with a strong negative loading towards the restoration score. In the model with environmental factors, only maternal antibiotics at birth and having older siblings had negative and positive loadings with the 1-year restoration score, respectively. The model on combined 1-week gut microbiome and environmental factors showed a better performance than each individual model, predicting the 1-year restoration score with 7 variables selected, including the same five species selected by the model on only the gut microbiome, and *Veillonella parvula* and having older siblings.



Paper I, Fig 4: Sparse partial least squares (sPLS) models on restoration score using the full cohort at 1 week of age. (A), (B), and (C) show loadings from models on gut microbiome at 1 week of age (143 species, 499 samples), environmental factors (39 factors, 466 samples), and combined factors (182 variables, 466 samples), respectively. Loadings represent the contribution of each variable to the sPLS models. Negative loadings indicate associations with lower restoration scores, while positive loadings indicate associations with higher restoration scores.
(D) Spearman correlation between cross-validated predictions from the above three sPLS models and the 1-year restoration score. The correlation coefficient and P value are shown in the legend. Blue for the sPLS model on gut microbiome, grey for the sPLS model on environmental factors and orange for the sPLS model on combined gut microbiome and environmental factors.

These results suggest that having older siblings may promote the restoration of the gut microbiome, we then investigated whether the species associated with having older siblings were also associated with the 1-year restoration score, by comparing the logFC of the species abundances between two outcomes.

A Species associated with having older siblings and 1-year restoration score



B Number of species associated with having older siblins and 1-year restoration score

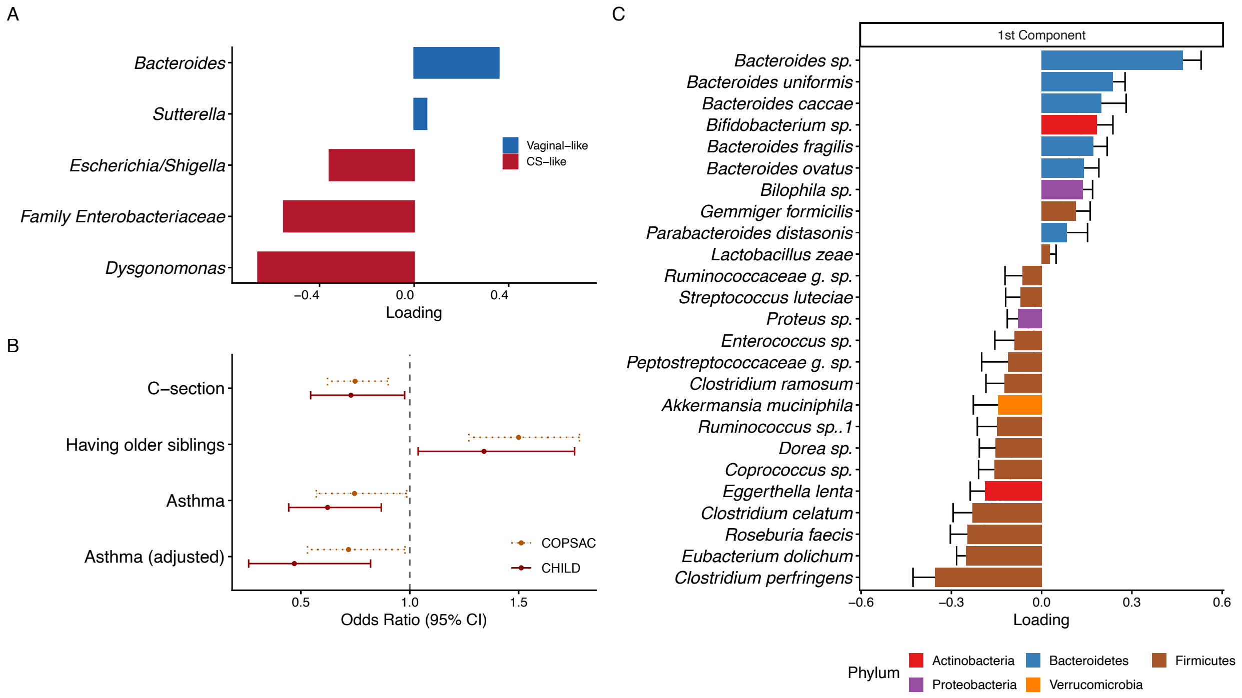
Paper 1, Fig 5: Differential abundance analysis reveals subset of species associated with having older siblings and restoration in the full cohort. (A) Scatter plot comparing differential abundance analyses of having older siblings and the restoration score. Each point represents a species, those associated with both factors are highlighted in blue. Labeled species were significant at 1 week and 1 month. A positive log fold change value in the first quadrant (upper right) indicates an increase in the abundance of certain bacteria associated with a higher 1-year restoration score and having older siblings at home, while a negative fold change in the third quadrant (bottom left) indicates a decrease in the abundance of certain bacteria associated with a lower 1-year restoration score and not having older siblings. P values were derived from permutation tests examining whether the direction of the associations with both predictors is concordant by chance, a P value smaller than 0.05 indicates the concordant direction is unlikely under the null hypothesis. (B)Venn diagram indicating the number of differentially abundant species according to having older siblings and according to the 1-year restoration score. P values were derived from permutation tests examining whether the number of taxa associated with both predictors is greater than expected by chance, a P value smaller than 0.05 indicates the overlap is unlikely under the null hypothesis.

At 1 week, 85 species showed a similar logFC association for both having older siblings and the 1-year restoration score (**Paper 1, Fig 5A**, 85 points, grey and blue, were in the first and third quadrants at 1 week). Among them, 5 were significantly associated with both outcomes (**Paper I, Fig.5B**). There was a higher abundance of *Bifidobacterium longum* and *Bifidobacterium pseudocatenulatum* in the children with siblings and a high restoration score; and lower abundance of *Actinomyces bowdenii*, *Clostridium tarantellae*, and *Clostridium perfringens* in the children without siblings and with a low restoration score. At 1 month of age, 4 species were significantly associated with both outcomes in the same direction. At 1 year of age, the time point where the restoration score was defined, 63 species showed significant consistent positive or negative association with both outcomes. To determine whether the number of taxa that were significant for both predictors was greater than expected by chance, and whether the direction of their associations was concordant by chance, we calculated empirical p-values based on permutation tests. Both of the empirical p-values were below 0.05, indicating that the overlap and concordant direction are highly unlikely by chance.

Having identified older siblings as a potential protective factor that may facilitate the restoration of a CS-perturbed gut microbiome, we next considered these factors together: We performed a mediation analysis to investigate the possible mediating role of the 1-week gut microbiome between having older siblings and the restoration of the gut microbiome by 1 year. Here, we used the cross-validated predictions from the sPLS model of the 1-week gut microbiome to predict the 1 year restoration score and expressed it as a 1-week microbial score representing the degree to which a child's 1-week microbiome looked like it would be restored by 1 year of age. After adjusting for delivery, the results showed a significant mediation effect of older siblings on the 1-year restoration score through the 1-week microbial score (*P*=0.004), but also a strong direct effect of older siblings on the 1-year restoration score (*P*<0.001). The indirect pathway accounted for a portion of 11.2% of the total effect. This suggests that older siblings may contribute to the restoration partially through influencing the very early gut microbiome.

We next sought to validate our findings in an independent dataset. We applied the sPLS model trained on the COPSAC$_{2010}$ cohort on the CHILD cohort's 1-year samples (n=325) to create a restoration score (**Paper I, Fig 6A**). This 1-year restoration score was, like in COPSAC$_{2010,}$ negatively associated with delivery by CS and positively associated with older siblings in the CHILD cohort (**Paper I, Fig 6B**). Additionally, the CHILD 1-year restoration score was associated with reduced asthma risk at 5 years in both univariate and adjusted logistic regression models (OR 0.63[0.44, 0.87], *P*=0.005, aOR 0.47 [0.26, 0.82], *P*=0.0084, **Paper I, Fig 6B**). This replicates the results in our previous study[17] and is consistent with the hypothesis that an appropriate restoration of the gut microbiota could mitigate the increased asthma risk associated with gut microbial changes due to CS delivery.

To identify species at early time points associated with the 1-year restoration score, we trained an sPLS model on the gut microbiome at 3 months in the CHILD cohort, to predict the 1-year restoration score(**Paper I, Fig 6C**). *Clostridium perfringens* was selected in the model with the highest negative loading as we saw in the COPSAC$_{2010}$ cohort, while *Bacteroides sp.* had the highest positive loading in the model, followed by *Bacteroides uniforms*, *Bacteroides caccae*, *Bifidobacterium sp.*, *Bacteroides fragilis*, and *Bacteroides ovatus*. Having older siblings was positively associated with the 1-year restoration score, which supports the idea that older siblings promote the restoration of a CS-perturbed gut microbiome(**Paper I, Fig 6B**).



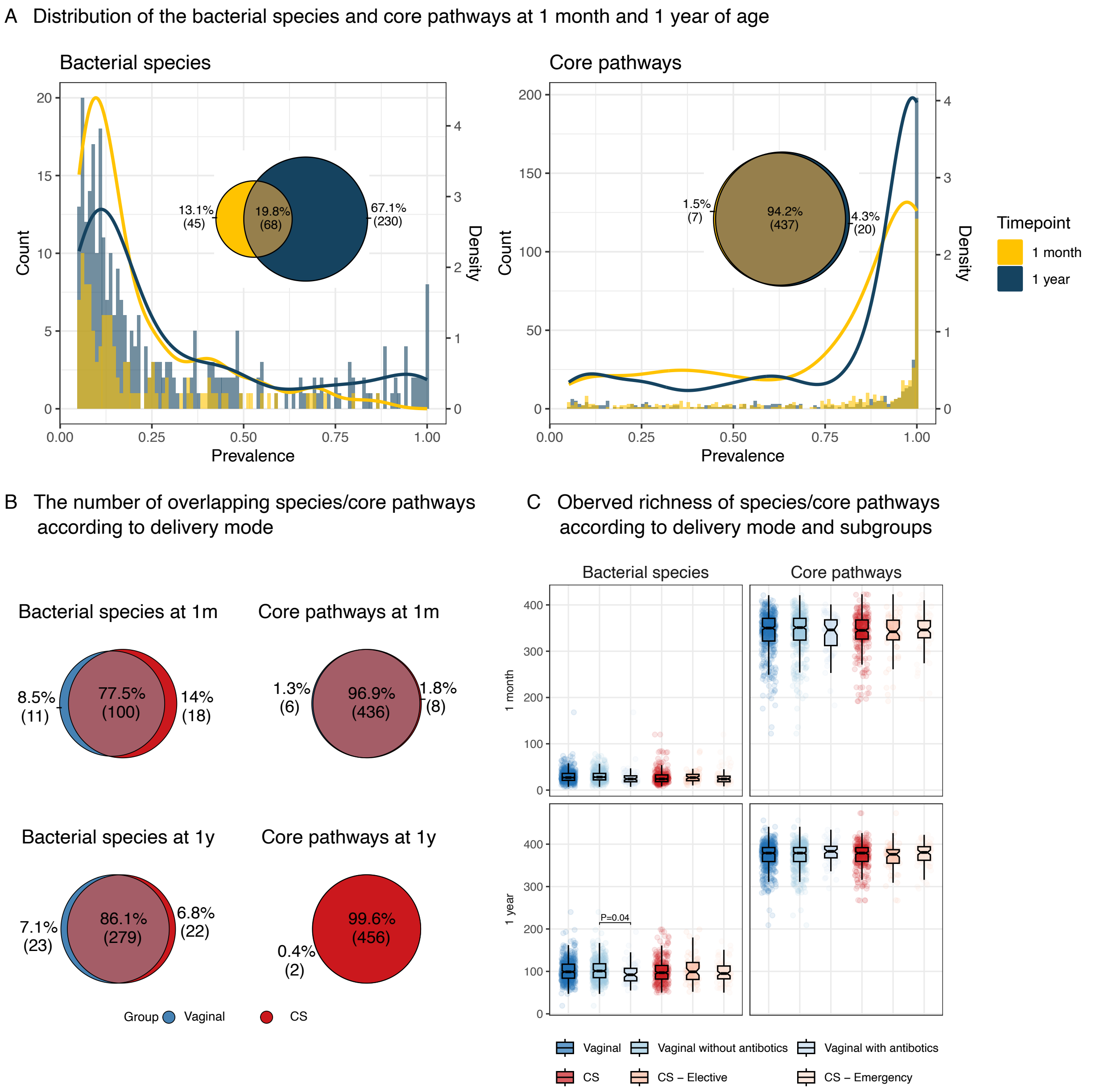
Paper 1, Fig 6 Validation of the COPSAC$_{2010}$ 1-year restoration score in the independent CHILD cohort. (A) Loadings are derived from the sPLS model trained on the COPSAC$_{2010}$ 1-year gut microbiome to output a restoration score, the annotations are derived from the CHILD cohort 16s rRNA data. (B) Univariate logistic regression models on the 1-year restoration score versus CS, having older siblings and asthma, and the adjusted odds ratio from multivariable logistic regression model on asthma, in both cohorts (dashed line: COPSAC, solid line: CHILD). The asthma diagnosis was at 5 years in the CHILD study, while the asthma diagnosis was at 6 years of age in the COPSAC cohort. The covariates adjusted in the models were: gestational age, hospitalization after birth, antibiotics exposure to children at 1 year, having older siblings, family asthma history, gender, race, birth season, (and study center for the CHILD study). (C) sPLS model on gut microbiome at 3-month in the CHILD cohort predicting 1-year restoration score, positive loadings indicated positive association with higher restoration score, while negative loadings indicated negative associations with higher restoration score.

Paper II: Early Gut Microbiome Composition and Functional Pathways Linking Cesarean section to Asthma Susceptibility

In this second paper, we profiled gut microbiome taxonomic and core functional composition at 1 month and 1 year of life in infants from the COPSAC_{2010} cohort born by CS or vaginal delivery. We investigated the effect of delivery modes on both the gut microbiome taxonomic and core functional compositions, and examined the role of bacterial species and core pathways in the association between delivery mode and asthma risk up to 10 years of age.

We first characterized the development of the taxonomic composition and core functional profile at 1 month and 1 year of age by describing the presence and absence of bacterial species and functional pathways, and quantifying the extent to which these features were shared between the two time points (**Paper II, Fig 1**). After filtering by prevalence above 5% across samples within each time point, we identified 113 and 298 bacterial species at 1 month and 1 year of age, respectively; 343 species in total. For functional profiling, reads were mapped to pathways using HUMAnN 3[219], yielding 444 and 457 pathways with prevalence above 5% at 1 month and 1 year of age, respectively; 464 pathways in total. These common pathways accounted for 6% and 5% of the total read-level abundance (Copies Per Million, CPM) at 1 month and 1 year, respectively. Unmapped reads and unintegrated genes (ie. genes not assigned to a core pathway) were excluded from all downstream analyses (1 month: 12% unmapped, 82% unintegrated of total CPM; 1 year: 15% unmapped, 80% unintegrated). Thus, our pathway abundance dataset, henceforth referred to as core pathways, focuses on broadly shared, well-annotated functions and represents a conservative view of the core functional capacity of the infant gut microbiome, likely understating the contribution of rare or less well characterized pathways.

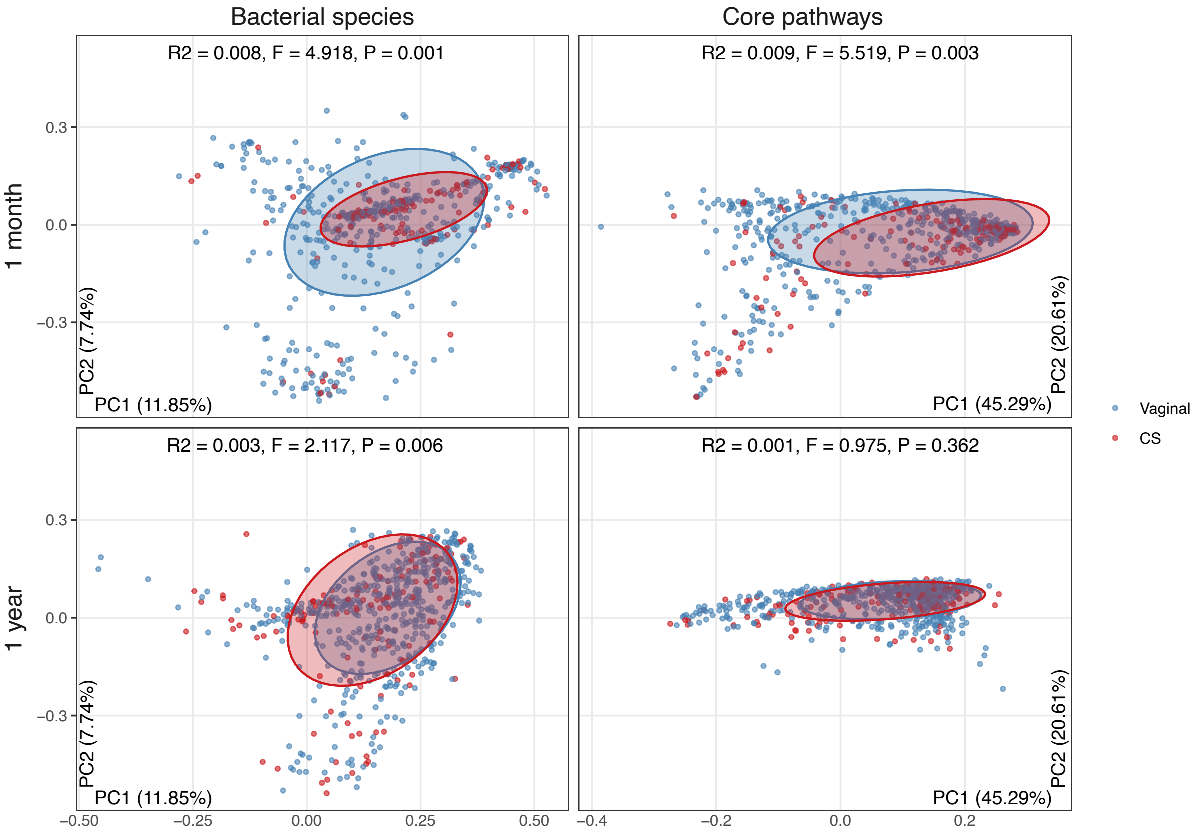
In general, most bacterial species had low prevalence at 1 month and only slightly higher at 1 year of age (**Paper II, Fig 1A** Bacterial species). Among all 343 detected species, 45 (13.1%) were unique to the 1-month time point, such as *Staphylococcus epidermidis* (prevalence 68%), *Streptococcus vestibularis*(49%), and *Staphylococcus aureus*(45%); while 229 (67.1%) were unique to the 1-year time point, such as *Faecalibacterium prausnitzii*(100%), *Anaerostipes hadrus*(98%), and *Ruminococcus bromii*(96%); and 69 (19.8%) were shared between time points, such as *Streptococcus salivarius*(prevalence at 1m/1y: 87%/80%), *Bifidobacterium longum*(80%/100%), and *Veillonella dispar*(72%/83%). In contrast, most pathways were present in all samples at 1 month and 1 year of age (**Paper II, Fig 1A** Microbial pathways). Among all 464 core pathways, 437 (94%) pathways were shared between time points, 7(2%) and 20(4%) pathways were unique to the 1-month and 1-year time point, respectively.



Paper II, Fig 1: Prevalence of the bacterial species and pathways. (A) Prevalence and the number of overlapping species/pathways at 1 month and 1 year of age (prevalence > 5% across samples within each time point). (B) The number of overlapping species/pathways between vaginal and CS group at 1 month and 1 year of age (prevalence > 5% across samples within each stratum at each time point). (C)Observed richness of species/pathways according to delivery mode and subgroups. The differences were tested using linear regression, adjusted for sequencing depth and batch. P values above 0.05 are not shown.

We then investigated bacterial species and pathways unique to the vaginal or CS group at 1 month and 1 year of age after filtering by prevalence above 5% across samples within each stratum at each time point ((Vaginal/CS at 1m: N=465/135; at 1y: N=517/143, **Paper II, Fig 1B**). At 1 month of age, 100 (77.5%) species were shared between the two groups, while at 1 year this increased to 86.1%. At 1 month of age, 436 core pathways (96.9%) were present in both groups, while at 1 year this increased to 99.6%. We then tested differences in observed bacterial species richness and core pathway richness using linear regression (**Paper II, Fig 1C**). The only significant difference was a lower bacterial species richness in infants born vaginally whose mothers received antibiotics at birth compared with infants born vaginally whose mothers did not receive antibiotics at 1 year of age. No differences were detected in core pathway richness.

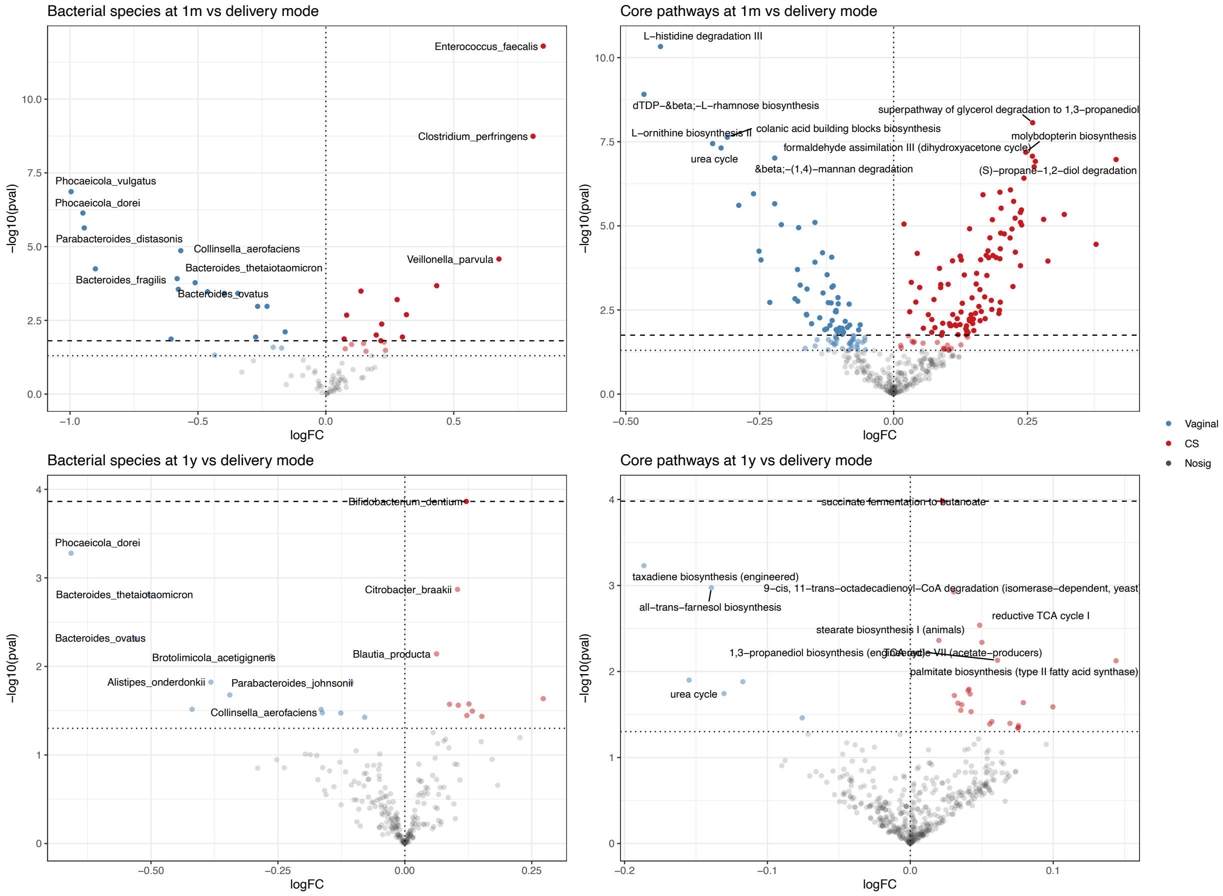
Associations between delivery mode vs. overall gut microbial composition and core pathway composition (beta diversity) were evaluated using Bray-Curtis distances (**Paper II, Fig. 2**). At 1 month of age, we observed compositional differences in both bacterial species and core pathways according to birth mode; the species difference was more pronounced. At 1 year of age, differences in species composition were attenuated but remained significant, whereas no significant differences were observed in core pathway composition between the vaginal and CS groups.



Paper II, Fig 2: Comparison of beta diversity of bacterial species and core pathways in relation to delivery mode. PERMANOVA was used to compare group differences. Each dot represents one fecal sample. Ellipses represent 1 standard deviation (1 SD), encompassing approximately 68% of the data points, assuming a bivariate normal distribution.

We then investigated individual bacterial species and core pathways associated with delivery mode (**Paper II, Fig 3**). Children born vaginally, whose mothers received antibiotics were excluded and bacterial species were filtered by prevalence above 5% and abundance above 0.01% within each time point, and core pathways by prevalence above 5% within each time point, yielding 90 species and 443 pathways at 1 month of age, and 229 species and 457 core pathways at 1 year of age.

At 1 month of age, 29 species and 159 core pathways were significantly different between infants born by CS and vaginal delivery (FDR adj P<0.05, **Paper II, Fig 3**). CS was associated with higher relative abundances of 13 species, including *Enterococcus faecalis*, *Clostridium perfringens*, and *Veillonella parvula*, and 103 core pathways including superpathway of glycerol degradation to 1,3-propanediol, molybdopterin biosynthesis, and formaldehyde assimilation III (dihydroxyacetone cycle). Conversely, vaginal delivery was associated with higher relative abundances of 16 species, including *Phocaeicola vulgatus*, *Phocaeicola dorei*, and *Parabacteroides distasonis*, and 56 core pathways, including L-histidine degradation III, dTDP-beta-L-rhamnose biosynthesis, and colanic acid building blocks biosynthesis. At 1 year of age, only *Bifidobacterium dentium* and succinate fermentation to butanoate was found to be individually associated with CS (*Bifidobacterium dentium* FDR adj P=0.03, succinate fermentation to butanoate FDR adj P=0.05).



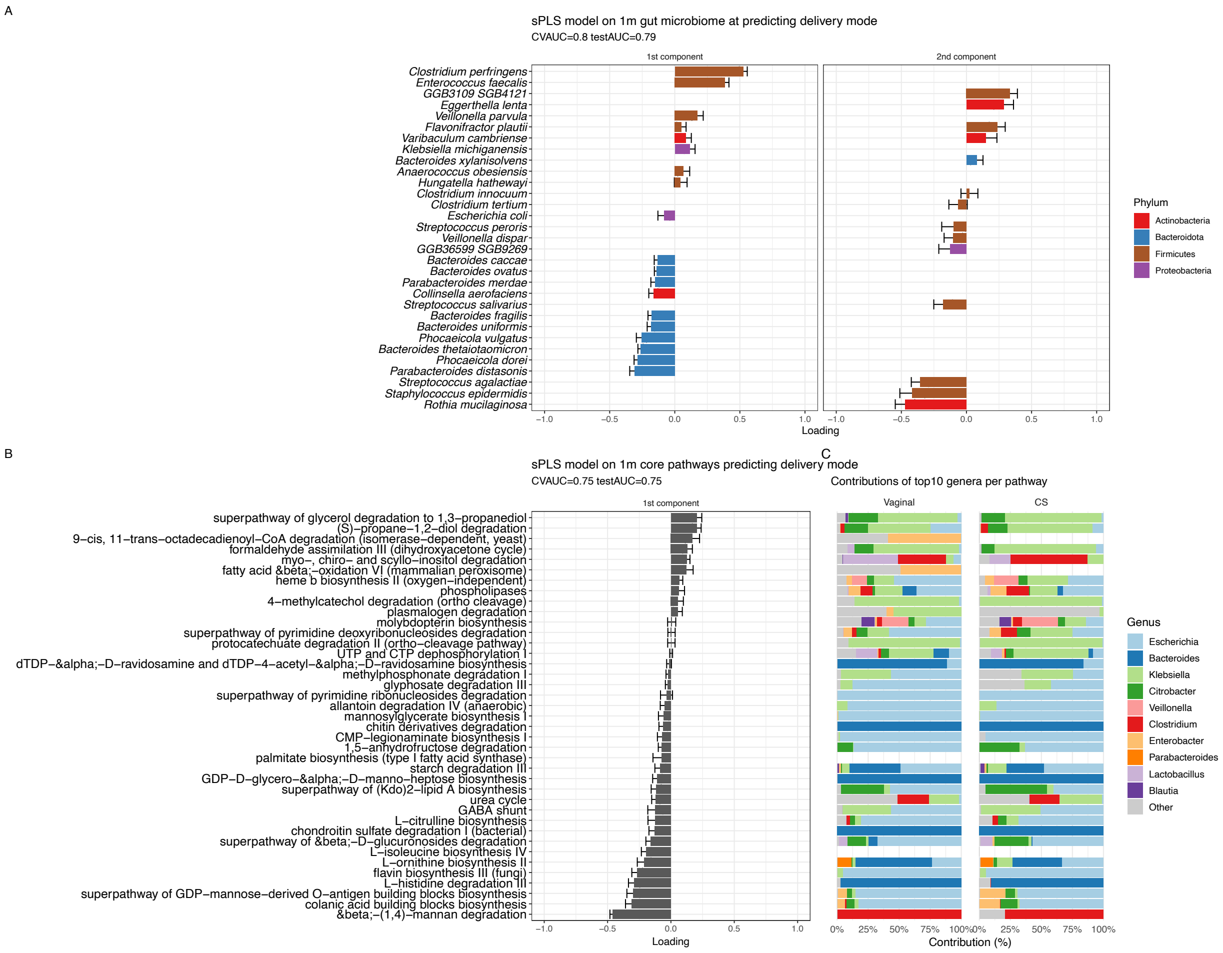
Paper II, Fig 3: Differential abundance analysis of bacterial species and core pathways according to delivery mode. Species with prevalence of at least 5% and relative abundance of more than 0.01% of the total were eligible. Core pathways with prevalence of at least 5% were eligible. All analyses were adjusted for sequencing depth and batch. Red dots represent species/pathways more abundant in the CS group, blue dots represent those more abundant in the vaginal without antibiotics group. The dotted line represents nominal P<0.05, and the dashed line represents FDR-adjusted P<0.05.

We next constructed cross-validated sparse partial least squares (sPLS) models to identify jointly correlated species and core pathway abundances associated with CS delivery against the vaginal without antibiotics group at 1 month and 1 year of age (1 month, N=532; 1 year, N=589). These models identified the sets of gut microbial species and core pathways at each time point most closely associated with CS using variable selection via L1-norm penalization. We extracted cross-validated prediction scores from these models and used them as CS microbial scores.

A 2-component model captured 31 different species in the 1-month CS microbial profile (**Paper II, Fig 4A**). The first component mainly included species positively correlated with CS, such as *Clostridium perfringens*, *Enterococcus faecalis*, and *Veillonella parvula*, as well as the vaginal birth associated *Parabacteroides distasonis*, *Phocaeicola dorei*, and *Bacteroides thetaiotaomicron*.

In parallel, a 1-component model selected 39 core pathways to best predict delivery mode at 1 month (**Paper II, Fig 4B**). Pathways positively associated with CS included superpathway of glycerol degradation to 1,3-propanediol, (S)-propane-1,2-diol degradation, 9-cis, 11-trans-octadecadienoyl-CoA degradation (isomerase-dependent, yeast). The abundances of these pathways were contributed to by *Klebsiella*, *Citrobacter*, and *Clostridium* (**Paper II, Fig 4C**). In contrast, pathways associated with vaginal delivery included beta-(1,4)-mannan degradation, colanic acid building blocks biosynthesis, and superpathway of GDP-mannose-derived O-antigen building blocks biosynthesis, which were mostly contributed by *Escherichia*, *Bacteroides*, *Klebsiella*, and *Clostridium* (**Paper II, Fig 4C**). The composition of species involved in each pathway was similar in infants born by vaginal delivery and those born by CS, suggesting that the differences in pathway abundances may be driven by the differential abundance of bacterial species rather than by changes in overall microbiota composition.

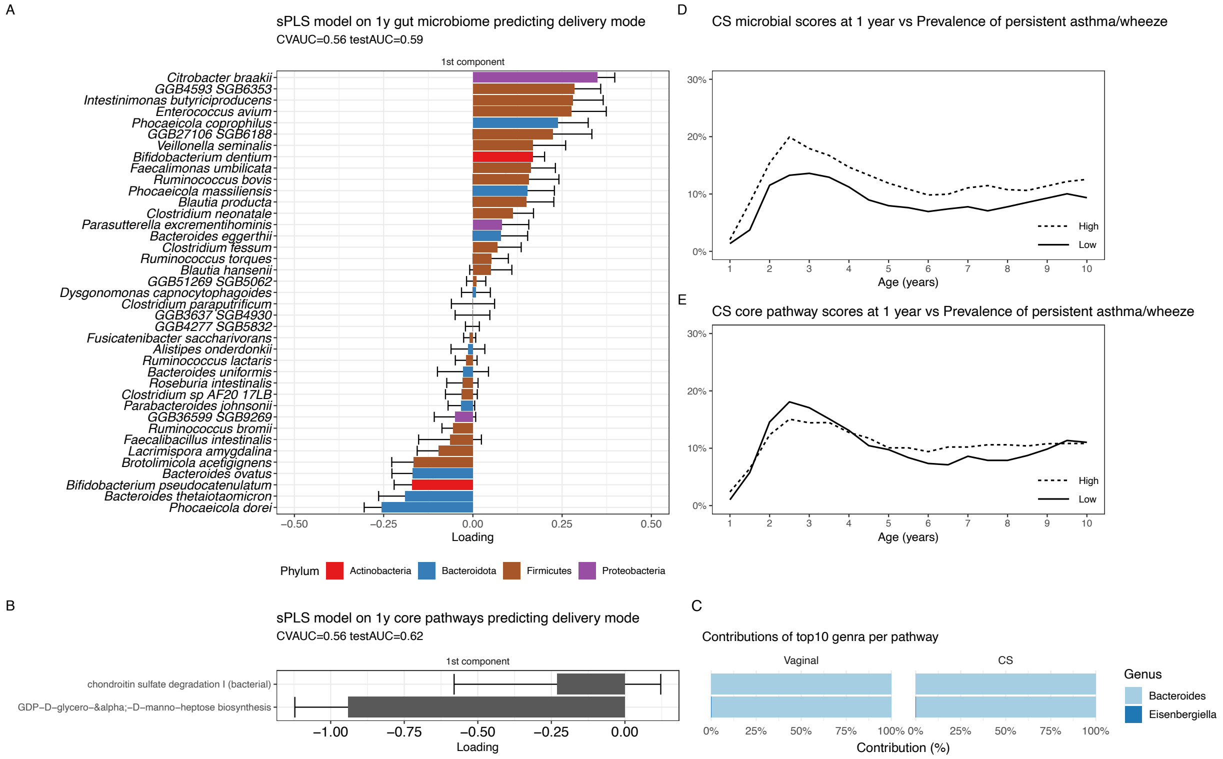
The model predictions were utilized as 1-month CS microbial and core pathway scores, and tested for associations with asthma from 1 to 10 years of age. There was no detectable association between these 1-month scores and asthma risk (P>0.05).



Paper II, Fig 4: sPLS model predicting delivery mode at 1 month of age based on the (A) gut microbiome and (B,C) microbial core pathways. Sparse partial least squares (sPLS) models were used to distinguish (A) fecal microbial composition and (B) microbial core pathway composition between children born by CS and vaginal delivery without antibiotics. Receiver-operator area under the curve (AUC_{ROC}) values calculated from repeated 10-fold cross-validation of the models (CVAUC) and in the test set (25% samples, testAUC) are shown. A random prediction yields a score of 0.5, with a perfect prediction being 1. Loadings (model contributions) for each species are ranked by magnitude and colored by phylum. Negative/positive loadings correspond to lower/higher abundance in CS born infants compared to vaginally delivered infants. (C) Bacterial contributions to each pathway, stratified by delivery mode. Contribution (%) represents the proportion of each species' relative abundance divided by the total relative abundance of all species contributing to that pathway within each stratum. (Note: unclassified contributions were excluded.)

At 1 year of age, a 1-component model included 39 species in the 1-year CS microbial score (**Paper II, Fig 5A**). The 3 main contributors positively correlated with CS were *Citrobacter braakii*, *GGB4593 SGB6353*, and *Intestinimonas butyriciproducens*; the 3 main contributors negatively correlated with CS were *Phocaeicola dorei*, *Bacteroides thetaiotaomicron*, and *Bifidobacterium pseudocatenulatum*. In contrast, only two pathways were selected in a similar model (**Paper II, Fig 5B**), which were chondroitin sulfate degradation I (bacterial) and GDP-D-glycero-alpha-D-manno-heptose biosynthesis. In line with the alpha and beta diversity results, this may indicate that the overall functional (core pathway) capacity of the gut microbiome was nearly identical between infants born vaginally and by CS.

When associating 1-year CS microbial scores and 1-year CS core pathway scores with asthma risk to 10 years of age, we found higher microbial scores to be associated with higher asthma risk over time (Generalized Estimating Equation (GEE) Odds ratio (OR) per SD=1.21, 95% confidence interval [1.01,1.45], P=0.03; Adjusted GEE OR=1.23, [1.02,1.48], P=0.03) (**Paper II, Fig.5D**), while the core pathway score was not associated with asthma risk (GEE OR=1.08, [0.88,1.32], P=0.46; Adjusted GEE OR=1.07, [0.86,1.32], P=0.56) (**Paper II, Fig.5E**).

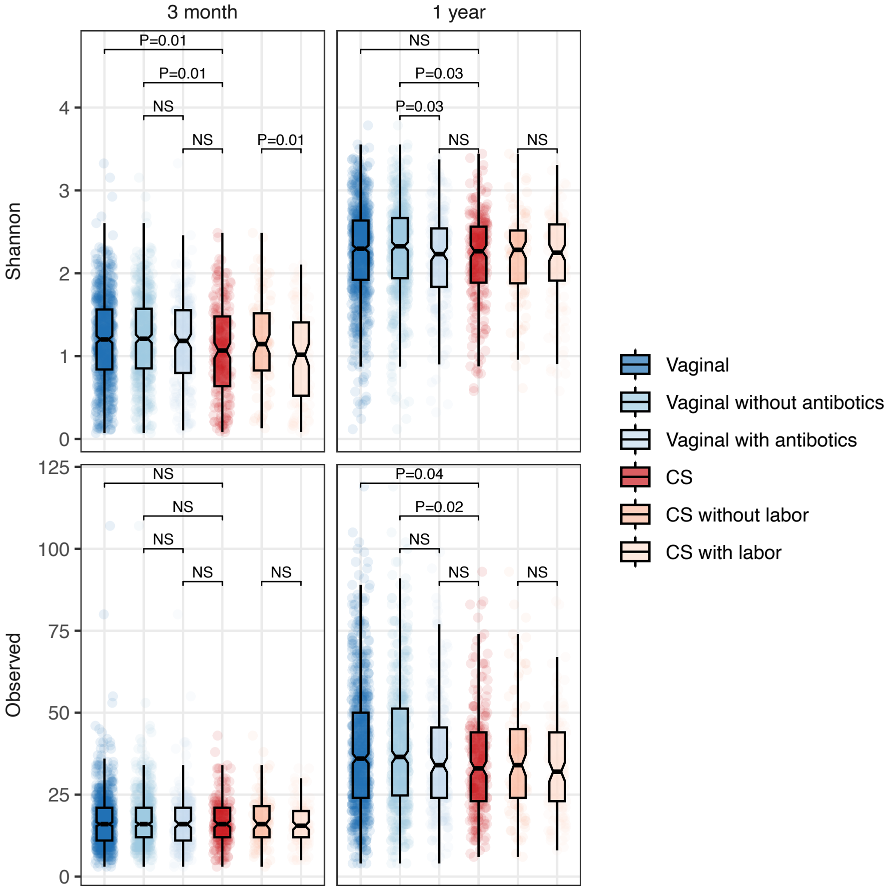


Paper II, Fig 5: sPLS model predicting delivery mode at 1 year of age based on the (A) gut microbiome and (B,C) core pathways. Sparse partial least squares (sPLS) models were used to distinguish (A) fecal microbial composition and (B) core pathway composition between children born by CS and vaginal delivery without antibiotics. Receiver-operator area under the curve (AUCROC) values calculated from repeated 10-fold cross-validation of the models and in the test set are shown. A random prediction yields a score of 0.5, with a perfect prediction being 1. Loadings (model contributions) for each species are ranked by magnitude and colored by phylum. Negative/positive loadings correspond to lower/higher abundance in CS born infants compared to vaginally delivered infants. (C)Bacterial contributions to each pathway, stratified by delivery mode. Contribution (%) represents the proportion of each species' relative abundance divided by the total relative abundance of all species contributing to that pathway within each stratum. (Note: unclassified contributions were excluded.) (D) and (E) The scores were dichotomized into high/low (above/below median) to visualize the prevalence of persistent asthma/wheeze, P values in the main text were measured using continuous scores, covariates adjusted in the model GEE models included: exclusive breastfeeding at 1 month, any breastfeeding at 1 year, hospitalization after birth, gestational age, maternal bmi, maternal asthma history, household income at 1 year, passive smoke exposure, maternal age, paternal age, the presence of older siblings, sex, race, birth season, sequencing depth, and batch.

Paper III: Microbially-derived tryptophan metabolites in the infant gut mediate the increased risk of asthma following cesarean section

In **Paper I**, we validated our model in the CHILD cohort by directly applying the model trained in the COPSAC$_{2010}$ cohort to 16S rRNA data from CHILD. In **Paper III**, we instead used shotgun metagenomic data and trained the model within the CHILD cohort to replicate the full methodology. Specifically, we assessed whether delivery mode was associated with gut microbiome composition at 3 months and 1 year of age in CHILD, and then test whether the resulting CS microbial scores were associated with asthma risk at 5 years of age.

We first looked into the alpha and beta diversity and investigated the effect of delivery mode on the overall diversity. At 3 months of age, children born by CS had a lower Shannon diversity compared to those born vaginally, but the observed richness was not significantly different (**Paper III, Fig 1**). In the subgroup analysis, Vaginal without antibiotics showed higher Shannon diversity than CS-born infants, whereas Vaginal with antibiotics showed Shannon diversity similar to that of CS-born infants. At 1 year of age the Vaginal without antibiotics group continued to have a higher diversity than CS-born infants, while the Vaginal with antibiotics group showed no difference from the CS-born infants, suggesting that the antibiotic exposure to the mother at birth affects the development of infants' gut microbiome during the first year of life. CS without labor had higher Shannon diversity than CS with labor at 3 months of age. At 1 year of age, we found no difference in the observed richness and Shannon diversity between birth modalities (**Paper III, Fig 1**).



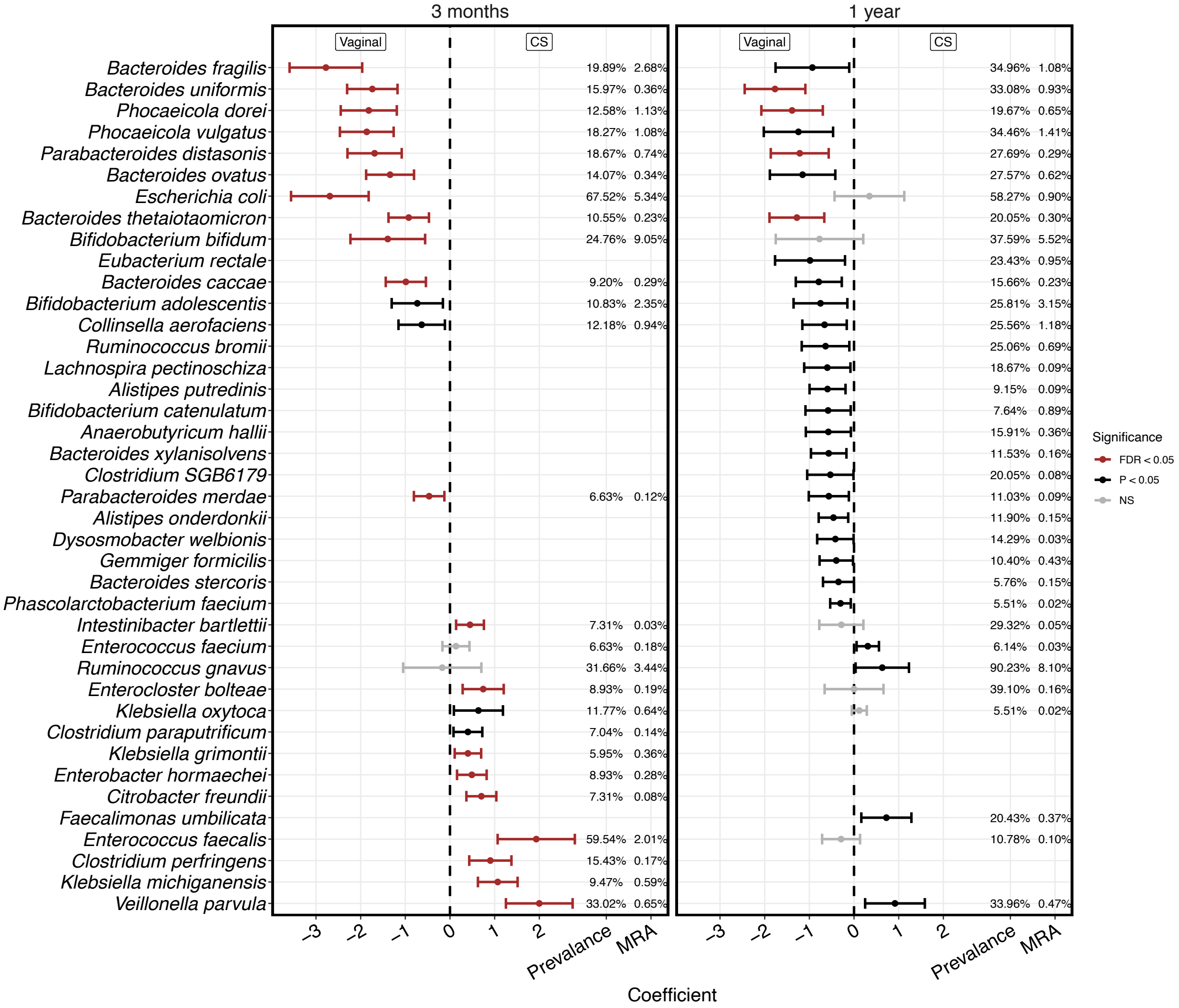
Paper III, Fig 1: Observed and Shannon diversity of the gut microbiome at 3 months and 1 year of age. Linear regression was employed for statistical comparisons at 3 months of age and 1 year of age, adjusting for processing period, exact age, and the sample collection site

Destinct compositional (Bray-Curtis distance) differences were observed between CS and vaginally born infants at 3 months, but not at 1 year of age (**Paper III, Table 1**). In contrast to the alpha diversity metrics, differences were also observed between vaginally born infants with and without antibiotic exposure at both 3 months and 1 year of age. Similar to the alpha diversity, vaginally born infant with antibiotics exposure were compositionally similar to CS-born infants, suggesting that intrapartum antibiotics may lead to a more “CS-like” microbial composition at 1 year of age. There was no detectable difference between CS with labor or without labor at 3 months or at 1 year of age. We performed the same analysis in Paper II (refer to attached **Paper II, Table S9**), and the compositional differences among delivery modes and subgroups were comparable between two independent cohorts.

Paper III, Table 1: Birth mode effects on beta diversity calculated based on Bray-curtis distances

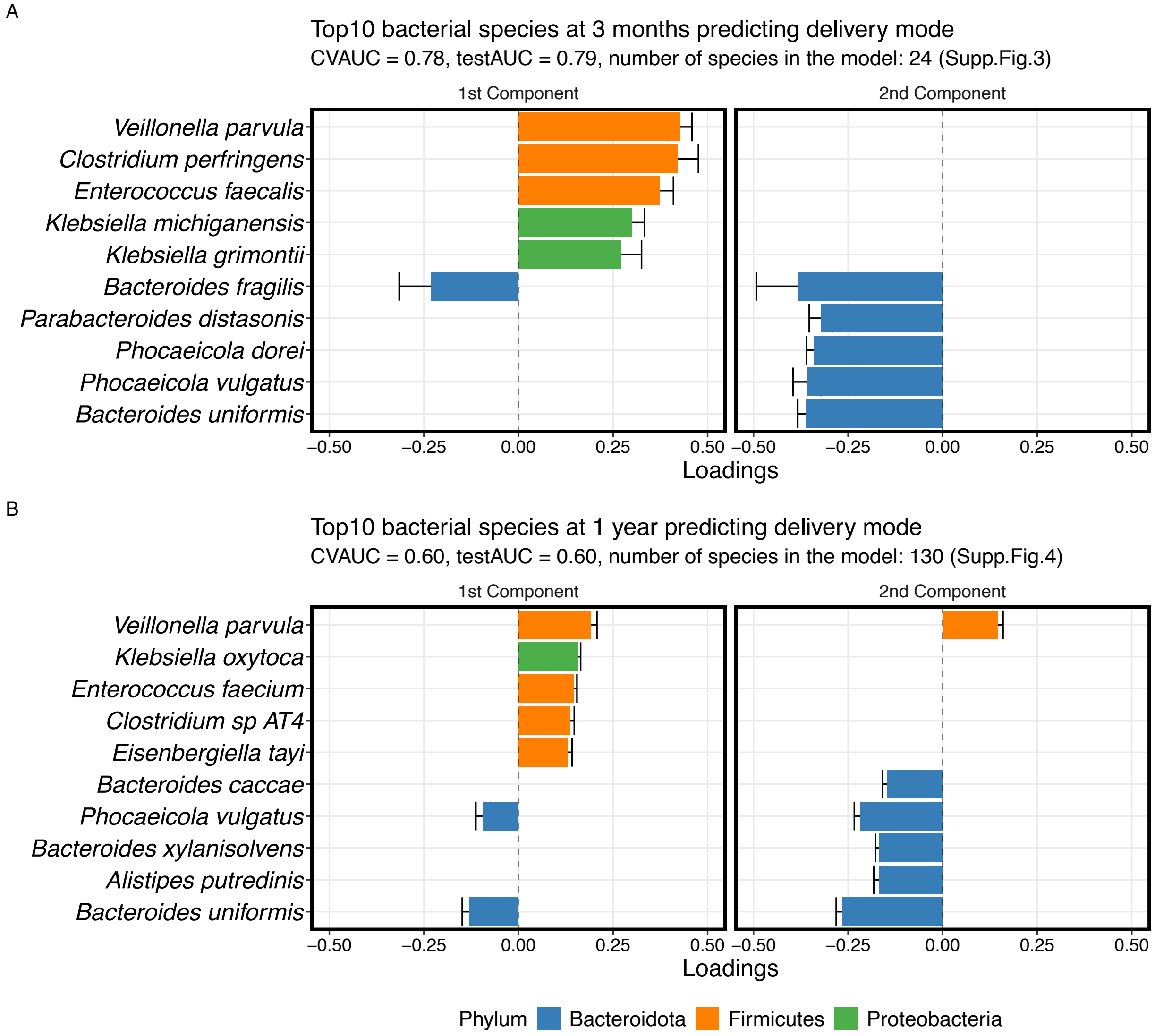
Group	Visit	R2	F statistic	P value	Sample size
CS vs Vaginal	3 month	0.006	5.685	**0.001**	956
	1 year	0.002	2.16	**0.005**	1029
CS vs Vaginal with antibiotics	3 month	0.007	2.835	**0.004**	425
	1 year	0.002	0.884	0.608	461
CS vs Vaginal without antibiotics	3 month	0.009	6.582	**0.001**	737
	1 year	0.003	2.771	**0.001**	794
Vaginal with antibiotics vs Vaginal without antibiotics	3 month	0.004	3.26	**0.001**	746
	1 year	0.002	1.986	**0.006**	795
CS with labor vs Vaginal	3 month	0.005	4.265	**0.001**	861
	1 year	0.002	2.266	**0.006**	923
CS without labor vs Vaginal	3 month	0.004	3.691	**0.001**	842
	1 year	0.001	1.333	0.126	904
CS with labor vs CS without labor	3 month	0.008	1.622	0.063	209
	1 year	0.006	1.458	0.066	231

To further evaluate the temporal microbial differences after CS delivery, differential abundance analyses were performed, comparing CS with Vaginal without antibiotics (**Paper III, Fig 2**). At 3 months of age, 20 (29%) of the 69 tested species were significantly different after FDR adjustment and at 1 year 4 (3%) of the 130 tested species were significantly different. This analysis mirrors the differential abundance analysis used in the **Paper II**, and we observed largely consistent delivery mode-associated species across the two cohorts. For example, at 3 month of age, *Clostridium perfringens*, *Veillonella parvula*, and *Enterococcus faecalis* were among the species most strongly associated with CS, consistent with findings in the COPSAC$_{2010}$ cohort (**Paper II, Fig 3**). At both 3 months and 1 year of age, children delivered by CS had significantly lower abundances of *Bacteroides uniformis*, *Phocaeicola dorei*, *Parabacteroides distasonis*, and *Bacteroides thetaiotaomicron*. Collectively, these results suggest a highly generablizabe effect of CS delivery on the infant gut microbiome at the bacterial species level.



Paper III, Fig 2: MaAslin2 forest plot of differential abundances of species in the gut microbiome according to birth mode, adjusting for processing period, exact age, and the sample collection site. Species significantly associated with CS at either 3 months or 1 year of age are shown here. Positive coefficients indicate a higher abundance of bacteria in the CS group than in the vaginal group. The prevalence and Mean Relative Abundance (MRA) of the bacteria across all samples are shown for each included species.

We then constructed cross-validated sparse partial least squares (sPLS) models to identify jointly correlated species associated with CS delivery vs the Vaginal without antibiotics group at 3 months and 1 year of age (3 months, N=739; 1 year, N=798, **Paper III, Fig 3**). These models identified the gut microbial composition at each time point most closely associated with CS and determined a CS microbial signature at both 3 months and 1 year. At 3 months of age, the model could predict the delivery mode based solely on the gut microbial composition, but less so at 1 year of age (cross-validated area under the curve (CVAUC) at 3 months: 0.78, **Paper III, Fig 3A**, also refer to the attached **Paper III Supp. Fig 3**; 1 year: 0.60, **Paper III, Fig 3B**, also refer to the attached **Paper III Supp. Fig 4**). At 3 months of age, a 2-component model captured 35 different species from 4 phyla: Actinobacteria, Bacteroidota, Firmicutes, and Proteobacteria in the CS microbial profile. The first component mainly included species positively correlated with CS, such as *Veillonella parvula*, *Clostridium perfringens*, *and Enterococcus Faecalis* and two species negatively correlated with CS, which were *Escherichia coli* and *Bacteroides fragilis* (refer to the attached **Paper III Supp. Fig 3**). The second component was driven by species negatively correlated with CS, with the biggest contributor being *Bacteroides fragilis*, followed by *Bacteroides uniformis*, *Phocaeicola vulgatus*, *Phocaeicola dorei*, *Parabacteroides distasonis*, and other species with lesser contributions. Most of these species were chosen in the 1-month model in Paper II to predict CS as well, see **Paper II, Fig 4A**.

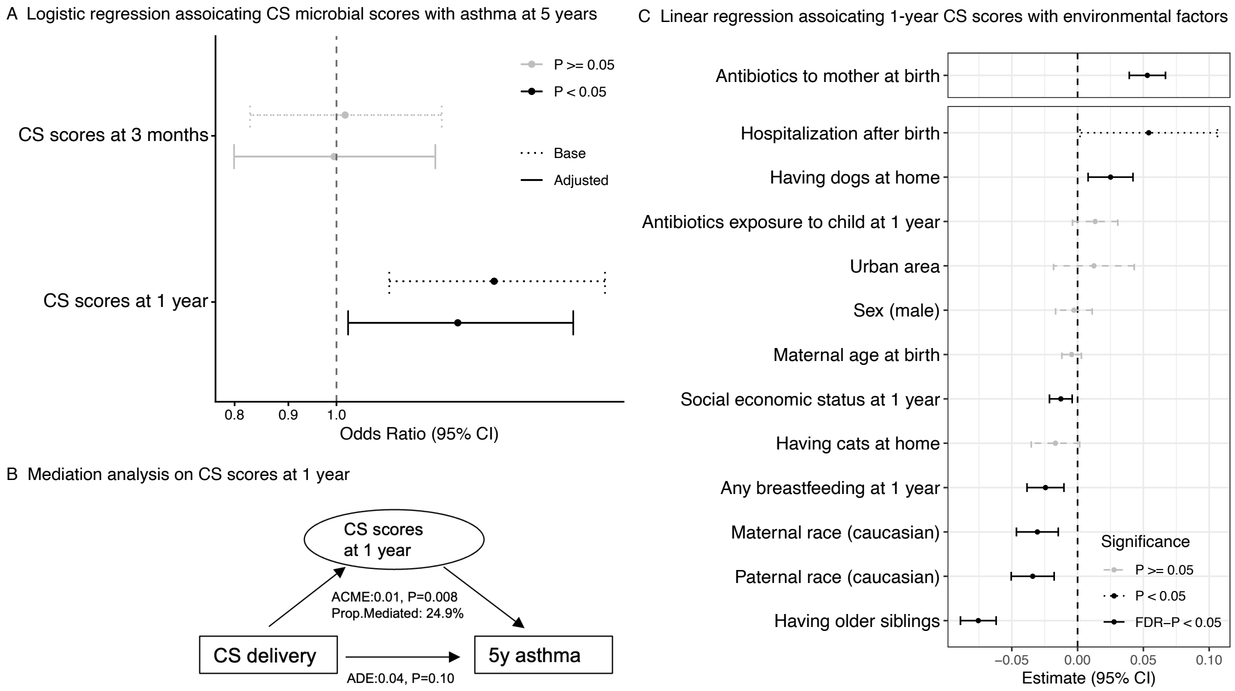


Paper III, Fig. 3: Top10 bacterial species contributing to the sPLS models predicting birth mode. sPLS models were trained to distinguish fecal microbial composition between children born by CS and vaginal delivery at A) 3 months of age and B) 1 year of age. Top 10 drivers were shown here. Area under the curve (AUC) values, calculated from repeated 10-fold cross-validation of the models and on the test set, are shown. Loading (model contribution) for each species is colored by phylum. Negative/positive loading corresponds to lower/higher abundance in CS born infants compared to vaginally delivered infants.

At 1 year of age, a 2-component model included 130 species from phyla Actinobacteria, Bacillota, Bacteroidota, Firmicutes, Proteobacteria, and Verrucomicrobia (**Paper III, Fig 3B**). The 5 main contributors positively correlated with CS were *Veillonella parvula*, *Escherichia coli*, *Streptococcus lutetiensis*, *Faecalimonas umbilicata*, and *GGB9469 SGB14862*. The 5 main contributors negatively correlated with CS were *Bacteroides uniformis*, *Phocaeicola vulgatus*, *Phocaeicola dorei*, *Bacteroides thetaiotaomicron*, and *Faecalibacterium prausnitzii*. Again, this is comparable to the model in Paper II, see **Paper II, Fig 5A**.

We then extracted the CS microbial scores at 3 months and 1 year of age by applying the models on all samples (larger age window, 3 months, N=1,355; 1 year, N=1,356, **Paper III, Fig 4**), but keeping the model loadings fixed. We then examined whether the CS microbial scores derived from the sPLS models were associated with the risk of asthma at age 5 in the entire cohort. Only the 1-year CS microbial score was associated with a higher risk of asthma at age 5 in both base and covariate-adjusted logistic regression models (Base OR 1.41 per SD [1.12,1.80], P<0.01; Covariate-adjusted OR (aOR) 1.30 [1.03,1.68], P=0.03, N cases/controls = 101/1,048, **Paper III, Fig 4A**). When performing a combined model including both 1-year CS microbial score and delivery mode in relation to 5-year asthma, we still observed associations for the microbial score (Base OR 1.35 [1.08,1.74], P=0.01; aOR 1.27 [0.99,1.64], P=0.06), but not for CS delivery (OR 1.50 [0.94, 2.36], P=0.08; aOR 1.34 [0.82, 2.15], P=0.24), suggesting that the score might mediate the CS association. A mediation analysis revealed an average causal mediation effect (ACME) to be positive (ACME 0.01 [0.003,0.02], P=0.008, **Paper III, Fig 4B**), suggesting that part of the association of birth mode on asthma risk may be mediated by the 1-year gut microbial scores. The proportion of the total effect mediated was approximately 25% (Prop. mediated 0.25 [0.04, 1.32], P=0.03).

To investigate the possible effect of environmental factors on the development of CS-assoicated microbial composition, we tested the assoications between several exposures and 1-year CS scores using linear regression (**Paper III, Fig 4C**). Antibiotics to mother at birth was tested in the vaginal stratum, and was significantly associated with higher 1-year CS scores, indicating a "CS-like" microbial composition. In contrast, having older siblings was most associated with lower 1-year CS scores, followed by maternal and paternal race, and any breastfeeding at 1 year of age. Consistent with findings from **Paper I** and **Paper II**, having older siblings showed the strongest association with reduced CS microbial scores at 1 year of age, suggesting a protective effect in restoring a perturbed gut microbiome.



Paper III, Fig 4: Associations between CS scores, asthma at age 5, and environmental factors. A) Forest plot of logistic regression models evaluating the association of CS scores with asthma at 5 years. The base models were adjusted for processing period, exact age of infants, and the sample collection site, and the covariate-adjusted models were adjusted for covariates including gestational age, hospitalization after birth, birth weight, sex, race (father and mother), any breastfeeding by 1 year of age, exclusive breastfeeding ever, antibiotics exposure to infants by 1 year of age, paternal asthma history, maternal asthma history, birth season, processing period, exact age of infants, and the sample collection site. B) Mediation analysis on CS score at 1 year adjusted for processing period, exact age, and the sample collection site. C) Linear regression models associating 1-year CS scores with environmental factors, adjusted for processing period, exact age, and the sample collection site. Antibiotics to mother at birth was tested in the vaginal stratum, while other factors were tested in the full cohort.

We then investigated the link between the CS microbial scores and gut metabolites quantified by nuclear magnetic resonance (NMR) and liquid chromatography with tandem mass spectrometry (LC-MS/MS) (**Paper III, Fig 5**). Out of 244 metabolites measured in the fecal samples at the concurrent timepoint (3 months, N=414; 1 year, N=1,243), 52 and 141 metabolites were found to be associated with 3-month and 1-year CS scores, respectively. These metabolites belonged to 7 classes including: organic acids and derivatives (3-month score: 27; 1-year score: 57), lipids and lipid-like molecules (4;42), organoheterocyclic compounds (4;12), benzenoids (4;9), organic nitrogen compounds (7;8), organic oxygen compounds (3;10), and nucleosides, nucleotides, and analogues (3;3).

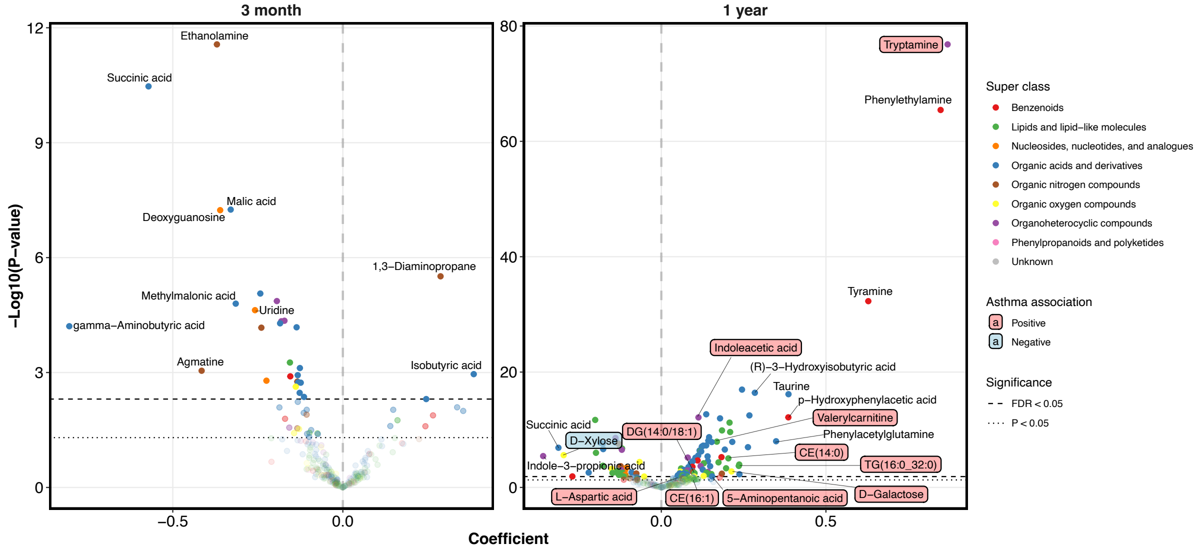
At 3 months of age, most metabolites were negatively associated with the 3-month CS microbial score, including ethanolamine, succinic acid, and malic acid, while 1,3-diaminopropane and isobutyric acid were positively associated with the 3-month CS microbial score (FDR adj. $P<0.05$, **Paper III, Fig 5A**).

At 1 year of age, most metabolites were positively associated with CS microbial score, such as tryptamine, followed by phenylethylamine, tyramine, and taurine, while succinic acid, indole-3-

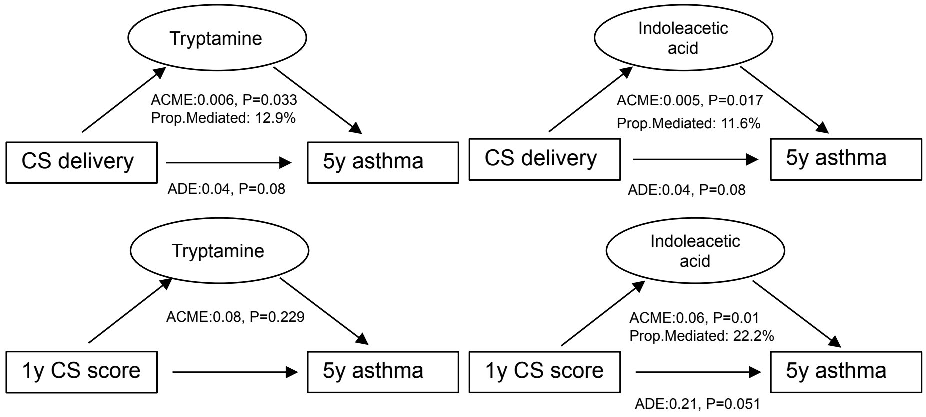
propionic acid, and D-xylose were negatively associated with the CS microbial score (**Paper III, Fig 5A**). It is worth noting that the associations between metabolites and the 3-month CS microbial scores were mostly in the same direction as the associations between metabolites and 1-year CS microbial score (Spearman Rho=0.40, P<0.01). In other words, those metabolites of which the levels decreased with a CS-associated gut microbiome at 3 months of age most likely kept being at a lower level at 1 year of age as well.

Furthermore, in the 1-year stool metabolome, we found that 11 of the CS microbial score-associated metabolites were also significantly associated with the risk of asthma at age 5 in a directionally consistent manner (**Paper III, Fig 5A**). For example, D-xylose, negatively associated with the 1-year CS microbial score, was found to be negatively associated with asthma risk (OR 0.89 [0.81,0.99], P=0.03); Tryptamine, positively associated with the 1-year CS microbial scores, was positively associated with asthma risk (OR 1.17 [1.03,1.35], P=0.02). This indicates that the link between a CS-perturbed gut microbiome at 1 year of age and later asthma risk could be mediated through an altered gut metabolomic profile.

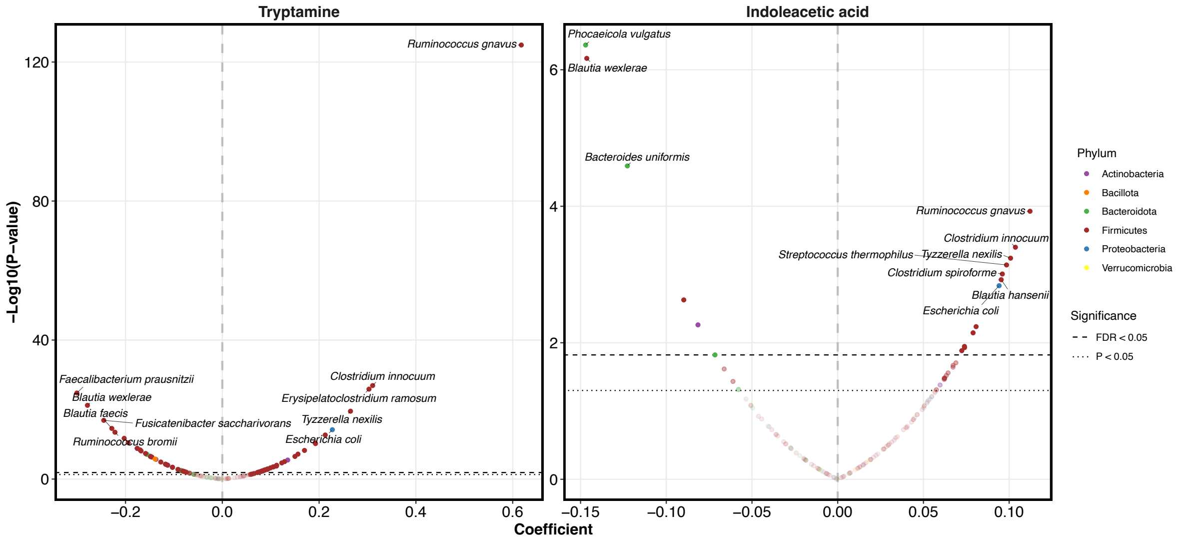
A Stool metabolites associated with CS microbial scores and asthma risk



B Mediation analysis on Tryptamine and Indoleacetic acid



C 1-year bacterial species assoicated with Tryptamine and Indoleacetic acid



Paper III, Fig 5: Metabolites associated with CS microbial scores at 3 months and 1 year of age. A) MaAslin2 volcano plot of fecal metabolites associated with CS microbial scores at 3 months and 1 year of age, adjusting for processing period, exact age, and the sample collection site. Positive coefficients indicate a higher level of metabolites associated with higher CS microbial scores. The top 10 metabolites ranked by the coefficients were labeled as plain text. Those boxed labels indicate metabolites associated with asthma at age 5 according to logistic regression, adjusting for processing period, exact age, and the sample collection site. B) Mediation models for associations between CS delivery or 1-year CS microbial scores and asthma risk at age 5, adjusting for processing period, exact age, and the sample collection site. C) Linear models associating Tryptamine and Indoleacetic acid with abundances of 1-year gut microbiome. Positive coefficients indicate a higher abundance of the bacteria associated with higher levels of Tryptamine or Indoleacetic acid in the stool at 1 year of age. The top 10 species ranked by the coefficients were labeled as plain text.

To further clarify this relationship, we then tested the mediating role of these 11 metabolites in the association between CS delivery and asthma risk at age 5, and found that tryptamine and indoleacetic acid (a tryptophan metabolite) had a positive ACME in the mediation analysis (Tryptamine, ACME 0.006 [0.0004,0.01], P=0.03; Indoleacetic acid, ACME 0.005 [0.0005,0.01], P=0.02, **Paper III, Fig 5B**). However, these mediation effects were not significant after adjusting for 1-year CS scores, indicating that the microbial scores and metabolites reflected each other in CS-related asthma risk. Therefore, we next tested the mediating role of the metabolites in the association between the 1-year CS microbial scores and asthma risk. Indoleacetic acid and three other metabolites showed a positive ACME (Indoleacetic acid, ACME 0.06 [0.01, 0.11], P=0.01; **Paper III, Fig 5B**). After adjusting for birth mode, it was still significant (ACME 0.05 [0.01, 0.10], P=0.01), which indicates that Indoleacetic acid may also be important for vaginally born infants with high CS scores.

Using linear regression, we tested associations between the abundances of 133 gut bacterial species at 1 year and the concentrations of tryptamine and indoleacetic acid. After FDR correction, 78 species (59%) were significantly associated with tryptamine and 11 species (8%) with indoleacetic acid (**Paper III, Fig 5C**). Species associated with tryptamine belong to *Blautia* (7), *Streptococcus* (6), *Bifidobacterium* (5), *Roseburia* (4), *Mediterraneibacter* (4) and 36 other genera (52). *Blautia faecis* showed the strongest negative association with tryptamine levels, while *Ruminococcus gnavus* was significantly associated with higher tryptamine levels. Metabolites associated with indoleacetic acid belong to *Blautia* (2), *Erysipelatoclostridium* (2), *Bacteroides* (1), *Escherichia* (1), *Mediterraneibacter* (1), and 4 other genera (4). *Phocaeicola vulgatus* showed the strongest negative association with Indoleacetic acid levels, followed by *Blautia wexlerae*, while *Clostridium spiroforme* was associated with higher indoleacetic acid levels, followed by *Ruminococcus gnavus*.

Discussion

In this section, I will discuss the combined results from the three included papers. The results are more thoroughly discussed within each individual paper.

This thesis examines the association between CS delivery and asthma risk, and how this relationship may be mediated by the gut microbiome, its functional capacity, and corresponding gut metabolites. Prior work in the COPSAC_{2010} cohort established a mediating role of a CS-associated gut microbiome signature in the association between CS delivery and asthma risk[157]. In **Paper I**, we followed this work and transformed the 1-year CS microbial score into a "restoration score" and, using data from 1 week and 1 month of age, identified early microbial predictors and environmental factors associated with microbiome restoration; these findings were subsequently validated in the CHILD cohort. In **Paper II**, we extended the prior work in **Paper I** using shotgun metagenomic data in the COPSAC_{2010} cohort and incorporated a functional profiling layer, there by refining our understanding of the effects of delivery mode on gut microbiome composition and microbial core pathways. In **Paper III**, we replicated the microbiome findings and further expanded the framework by integrating stool metabolomics in the CHILD cohort. Across all three papers, we observed consistent delivery mode-associated differences in the early gut microbiome that persist to at least 1 year of age, and we confirmed associations between the 1-year CS microbial score and asthma risk up to 5, 6, and 10 years of age. Moreover, the presence of older siblings showed a consistent protective association across all three papers, suggesting a potential role in gut restoration and thereby reducing the risk of developing asthma.

The effect of CS on the gut microbiome was most pronounced at early time points, and attenuated by 1 year of age. This pattern likely reflects differences in the initial seeding, and the overrepresentation of specific taxa in CS-born infants may alter the gut environment and thereby exert lasting influence on the developmental trajectory of the perturbed gut microbiome. Being born by CS doesn't change the basic physical structure/anatomy of the gut compared with being born vaginally. However, delivery by CS was reported to be associated with increased intestinal permeability possibly due to the reduction of the expression of tight junction protein related genes[89]. Increased permeability is considered a normal finding in the neonatal intestine[230], but a defective intestinal barrier and an altered microbiome composition are associated with intestinal disorders[231]. The first colonizers of CS born infants are mainly from members of the maternal skin microbiota[232] or the hospital environment[233]. This difference in the initial seeding was associated with a different gut environment after birth. Vaginally delivered infants develop a more acidic gut environment as early as day 3 after birth characterized by an enrichment of strict anaerobes. However, CS born infants presented an enrichment of Enterobacteriaceae, which was associated with meta-omics signatures characteristic of a microbiome adapted to a more oxygen-rich gut environment, enriched with genes associated with Reactive Oxygen Species (ROS) metabolism[90]. In our **Paper II** and **III**, we observed enriched *Enterococcus faecalis*, *Clostridium perfringens*, and *Klebsiella* at early time points following CS in both cohorts, according to the literature, these bacterial species are all able to produce ROS.

Enterococcus faecalis is able to respire aerobically when grown in the presence of hemin[234], and detoxify ROS[235]. Hemin is the oxidized form of heme, which occurs when heme is exposed to oxygen or other oxidizing agents. *Enterococcus faecalis* itself does not synthesize heme, but we observed heme b biosynthesis II (oxygen-independent) to be enriched in the CS-assoicated microbial functional profile, contributed by *Escherichia*, Citrobacter, *Klebsiella*, *Veillonella*, and *Enterobacter* in **Paper II**. In the presence of oxygen, *Enterococcus faecalis* can produce superoxide anions, hydrogen peroxide, and hydroxy radicals[236,237]. This capability to produce ROS gives this species a competitive advantage in terms of its ability to survive and to be more virulent toward the host[237–239].

Klebsiella michiganensis is a member of *Klebsiella oxytoca* species complex[240], of which some strains, particularly those with a type VI secretion system (T6SS), can generate ROS as a weapon against competitor bacteria[241,242]. Interestingly, phospholipase was identified as T6SS effector[243], which we observed in our study, enriched in the CS microbial functional profile, contributed mainly by *Escherichia*, *Klebsiella*, *Clostridium*, and *Enterobacter*. Bacterial phospholipases are classified into phospholipase A (PLA), PLB, PLC, and PLD by their cleavage site on a phospholipid molecule[244]. They have a range of effects *in vivo* and *in vitro* from minor alterations in cell membrane composition and function to lethality at low concentrations[243,245,246].

Another well-known phospholipase producer is *Clostridium perfringens*, a Gram-positive anaerobic bacterium that can form spores that are crucial during transmission[247], which we observed to be highly abundant in the CS born infants in Lay and **Paper III**. In **Paper I**, its abundance at 1 week was associated with 1-year perturbation status, indicating its long-lasting effect on the development of the gut microbiome. In fact, it can survive in the presence of oxygen, and/or low concentrations of ROS like superoxide or hydroxyl radicals by an adaptive response involving antioxidant enzymes[248,249]. As an aero-tolerant anaerobe, it could survive through aerobic environments such as on surfaces in hospital wards[250], which is probably where this spore former passes on to in-hospital neonates. It has been linked to various intestinal diseases in both humans and animals, particularly the necrotizing enterocolitis in neonates[251,252]. Of all the toxins *Clostridium perfringens* is able to produce[253], PLC (*Clostridium perfringens* alpha-toxin) is reported to participate in cellular signaling and regulation by virtue of its ability to hydrolyze membrane phospholipids into di-acyl-glycerol (DAG) and inositol triphosphate (IP3), which further causes the activation of other signaling pathways involved in various processes, including immune response[254,255]. At high concentrations, alpha-toxin causes membrane disruption and cytolysis[256]; at low concentrations, it leads to unregulated generation of secondary messengers - ROS[257–259]. When it comes to the host-interaction in the gut, we have to think about the gut barrier - mucus layer. There’s a protective mucus layer on the intestinal surface, especially in the colon, that separates the epithelial cells from the large population of commensal bacteria[260]. Mucins are the main structural components of mucus and play an integral and multifaceted role in the interaction between microbes and epithelial surfaces[261]. In the colon, the mucus layer differentiates into inner and outer layers due to the secretion of mucins by abundant goblet cells. When the outer layer allows bacteria to colonize (adhesion, growth, and reproduction), the inner layer, with a smaller pore size and higher viscosity under normal physiological conditions, does not permit bacterial passage[262]. However, smaller objects, for example, 50 kDa proteins, are able to pass through this layer[263]. Alpha toxin of *Clostridium perfringens* is 42.5 kDa[264].

On the other hand, vaginal delivery is assoicated with more anaerobes, especially from *Parabacteroides*, *Bacteroides*, and *Phocaeicola*, observed in both **Paper II** and **Paper III**. In **Paper I**, the abundance of *Bifidobacterium* at 1 week and 1 month was associated with increased restoration scores at 1 year of age, indicating a protective role. Different from CS-assoicated bacteria, most vaginal delivery assoicated bacteria are gram-negative, of which the outer membrane is made of lipopolysaccharide (LPS)[265]. LPS is a large glycolipid composed of three structural domains: lipid A, the core oligosaccharide, and the O antigen[266]. The biosynthetic pathway of LPS is a complex process that begins with the synthesis of Kdo2-lipid A in the cytoplasm and on the inner leaflet of the inner membrane, followed by elongation of the O-polysaccharide (O-antigen) chain in the periplasm. The completed LPS is then transported to the outer membrane for insertion[267]. We then looked into the functional profiles in **Paper II,** in this study, the enriched pathways of core LPS (superpathway of (Kdo)2-lipid A biosynthesis, GDP-D-glycero-α-D-manno-heptose biosynthesis, palmitate biosynthesis (type I fatty acid synthase)), O-antigen/other envelope polysaccharides (superpathway of GDP-mannose-derived O-antigen building blocks biosynthesis, CMP-legionaminate biosynthesis I, colanic acid building blocks biosynthesis) reflected a microbial composition rich in gram-negative bacteria.

The beta-(1,4)-mannan degradation pathway was upregulated in vaginal-associated core pathways as reported in other studies[268]. This pathway was reported to be increased in infants with a Bacteroidaceae-dominated gut microbiome[268]. The genome of *Bacteroides fragilis* was found to contain an operon encoding putative mannan metabolizing enzymes[269], and this operon is conserved in the genomes of other *Bacteroides* species, as well as species of *Parabacteroides*[270]. According to the literature, we believe this pathway was contributed by those vaginal delivery-associated bacteria. It’s also reported that mannose can inhibit bacterial adhesion[271], which may indicate the competition for colonization niche in the gut.

Besides, the enriched GABA shunt could reflect a distinct cross-feeding strategy in the gut microbiota of vaginally delivered infants. Gamma-Aminobutyric acid (GABA) shunt is a biochemical pathway that creates a “closed loop” to produce and conserve GABA. It bypasses two steps of the TCA cycle, converting alpha-ketoglutarate to succinate, and is linked to the mitochondrial electron transport chain[272]. The primary function of the GABA shunt is the synthesis and maintenance of GABA, although it also contributes to energy production by providing NADH and succinate to the cell[273]. As a major neurotransmitter in the brain, GABA plays a crucial role in behavior, cognition, and the body’s stress response[274]; while succinate is a key cross-feeding metabolite[275,276]. Interestingly, we found lower levels of succinate associated with CS-associated microbiome at 3 months in the CHILD cohort in **Paper III**, which was consistent with the depleted abundance of this pathway in the CS-associated microbiome. In **Paper II**, GABA shunt was mainly contributed by *Klebsiella* and *Escherichia*, but it was also proved to be actively expressed by *Bacteroides* and *Parabacteroides* in a transcriptome analysis[277], which is consistent with the higher abundance of these genera in the vaginally delivered infants. As we found the abundance of *Bacteroides* substantially decreased while *Veillonella parvula* (a succinate consumer) was enriched in the gut microbiome of CS-born infants in both **Paper II** and **Paper III**, it is expected that the production of succinate was lower and consumption higher than in vaginally-born infants. *Bacteroides*-derived succinate is reported to favor the colonization by anaerobes[275,276]. Consequently, this low concentration of succinate in the gut environment can affect microbiota composition, as well as microbial functional capacity. This pathway profile may also reflect the production of short-chain fatty acids (SCFA), as evidenced by the enriched GABA shunt, starch degradation III, chondroitin sulfate degradation I (bacterial), superpathway of β-D-glucuronosides degradation, chitin derivatives degradation, and 1,5-anhydrofructose degradation. SCFA have been reported to increase mucus secretion[278,279], which may have beneficial effects on gut health by supporting mucosal barrier integrity.

Although 1-month CS microbial scores were not associated with asthma risk, the CS microbial composition at early time points may shape the gut microenvironment favoring a divergent developmental trajectory compared with the vaginal-delivery-associated microbial composition. It’s important to mention that *Enterococcus faecalis* and *Clostridium perfringens* existed in vaginally born infants as well, just at lower abundances than in CS born infants. However, *Enterococcus faecalis* may respire anaerobically and the sporulation of *Clostridium perfringens* can be inhibited by *Bacteroides* and short-chain fatty acids (SCFA)[280]. In vaginally born infants, cross-feeding interactions and SCFA production could promote a hypoxic and acidic environment beneficial for a stable transition from facultative anaerobes to obligate anaerobes, which was reduced or absent in the CS-associated gut microbiomes[281]. Besides, these results appear to point towards a hypothesis that CS-born infants might be exposed to increased generation of ROS in the local intestine due to the colonization strategy of opportunists, which may lead to oxidative stress - an imbalance between ROS and the body's ability to detoxify them with antioxidants. This microbial profile is also consistent with multiple studies reporting that CS born infants have a gut microbiome with increased gut redox potential, resulting from a delayed establishment of strict anaerobes[282]. Oxidative stress has been recognized as both a trigger and an endpoint for several events, including inflammation with impairment of pulmonary function and prolonged lung damage[283]. However, CS born infants don’t have an “unhealthy” gut, and oxidative stress on the colon epithelium from gut bacteria is most likey low level, and the inflammation, if any, is low grade, chronic and subject to modulation by diet (breastmilk) and changing flora. Counterintuitively, reduced inflammatory responses in CS-born infants have been characterized as a “hypo-inflammatory state” with functional consequences for immune competence[284–286]. In contrast, vaginally born infants might have elevated systemic inflammation due to the lipopolysaccharide (LPS) produced by gram-negative bacteria enriched in the gut[287]. Thus, a more oxidizing gut environment in CS-born infants may shape early microbiome assembly while providing less of the immediate epithelial innate immune stimulation reported to occur after vaginal delivery[232]. Consistent with this, we observed sustained depletion of *Phocaeicola dorei*, *Bacteroides theraiotaomicron*, and *Bacteroides ovatus* in the CS born infants all the way from 1 month to 1 year of age, while *Citrobacter braakii*, *Veillonella parvula*, and

Ruminococcus gnavus were found to be more abundant. This compositional pattern was then linked to asthma risk up to 10 years of age.

Across all three papers, we observed that only 1-year microbial scores were associated with asthma risk even though the CS-signature was more pronounced in the gut microbiome at early time points. We speculate that this is the consequence of both the delay of immune training and specific taxa overrepresented at 1 year of age. As mentioned above, the vaginal-associated microbial functional profile is reported to be associated with elevated inflammation due to the LPS biosynthesis[287]. In the same study, an increase in the number of Gram-negative bacteria in the CS born infants at 1 year of age was reported, resulting in similar levels of TNF-alpha as observed in the vaginally born infants at 1 year of age[287]. In **Paper II**, *Citrobacter braakii* was the largest contributor to the 1-year CS microbial score. *Citrobacter braakii*, an anaerobic, gram-negative bacterium, has been isolated from the environment, food, and humans[288,289]. It was reported to have proinflammatory and cytotoxic role in gastric epithelial cells, and has been associated with acute mucosal inflammation in the intestine, respiratory tract, and urinary tract[289]. Surprisingly, T6SS was described in this species[289], which is also characterized in the *Klebsiella* mentioned above to generate ROS. A positive relationship between a group of bacteria including *Citrobacter braakii* and ROS overgeneration was also detected *in vitro* study[290], supporting that this bacterium is a ROS producer in the gut. In **Paper II** and **Paper III**, we observed higher abundance of *Ruminococcus gnavus* associated with 1-year CS-associated microbial scores. *Ruminococcus gnavus* is confirmed to be able to degrade mucin by the utilization of terminal mucin glycans[291], and *Ruminococcus gnavus*-produced tryptophan metabolites increases anion-dependent proximal colon secretion[292], and can influence inflammation as a ligand to the aryl hydrocarbon receptor[293,294]. In **Paper III**, we observed that tryptamine was the most pronounced metabolite positively associated with high CS microbial scores. Indoleacetic acid, another AhR ligand which is also a downstream product of tryptamine, was found to mediate the association between CS delivery and 1-year CS microbial signature and asthma risk at age 5. These two tryptophan metabolites were all positively associated with the abundance of *Ruminococcus gnavus* in **Paper II**.

In **Paper I**, we hypothesized that an appropriate restoration of the gut microbiota could mitigate the increased asthma risk associated with gut microbial changes due to CS delivery. We tested the association between different environmental factors and the 1-year restoration score. Having older siblings has been associated with increased gut microbial diversity during early childhood[295]. This association was seen in **Paper II** and **III** as well, demonstrated by the negative association between the presence of older siblings and CS microbial scores across all time points. Key taxa who appeared as partial mediators of this association between the presence of older siblings and restoration of the perturbed gut microbiome included the enriched *Bifidobacterium longum*, *Bifidobacterium pseudocatenulatum* and *Bifidobacterium adolescentis*. The mechanism of the presence of older siblings promoting these *Bifidobacterium* is not yet elucidated, however, *Bifidobacterium* was reported to be associated with increased SCFA production[296]. This may promote an acidic environment in the gut, and facilitate the consumption of oxygen by the β-oxidation of colocytes[77], which then shape the gut microbiome shifting towards obligate anaerobes[297]. Among SCFAs, butyrate is the preferred energy source for the colon epithelial cells[298]. *Bifidobacterium* is not a byturate-producer, but it has been reported to be involved in the cross-feeding interactions with byturate-producing colon bacteria, such as *Faecalibacterium prausnitzii*, *Anaerostipes*, and *Eubacterium*[299]. In our study, the abundance of *Bifidobacterium pseudocatenulatum* at early time points was found to be associated with higher 1-year restoration scores in **Paper I**, and negatively associated with 1-year CS microbial scores in **Paper II**.

Across the three papers, this thesis has several key strengths. First, we leverage deeply phenotyped prospective birth cohorts with clinical follow-up and physician-diagnosed asthma based on predefined criteria. Second, we combined longitudinal sampling with complementary microbiome data layers including 16S rRNA profiling in two independent cohorts for replication and shotgun metagenomic sequencing to resolve species-level composition and core functional potential, further extended with stool metagenomics to generate mechanistic hypotheses about microbe-derived metabolites. Most importantly, the consistent finding across cohorts confirming the link between the persistent CS-associated microbial signatures at 1 year of age and asthma risk at 5, 6, and up to 10 years of age demonstrate the robustness and generalizability of this work. With the species-level resolution, the identification of modifiable ecological and environmental factors further motivates the translational potential of strategies aimed at restoring a perturbed gut microbiome, and ultimately reducing asthma risk in CS-born infants.

We also acknowledge several limitations. As these are observational studies, the analyses only support associations rather than causality, and residual or unmeasured confounding cannot be fully excluded. For example, in **Paper I** and **II**, higher CS microbial scores were not exclusive to CS-born infants, it can also be observed among vaginally delivered infants whose mothers received antibiotics at birth or those exposed to other birth complications. This is important when interpreting the score as a marker of microbial perturbation status rather than delivery mode *per se*. Methodologically, 16S rRNA sequencing in **Paper I** limits taxonomic resolution. In **Paper II**, interpretation of functional findings was constrained by incomplete functional annotation. A substantial fraction of metagenomic reads map to genes not assigned to curated core pathways, meaning that potentially relevant non-core or species-specific functions may not have been captured in this study. Finally, in **Paper III**, stool metabolomic measurements could not distinguish bacterially produced from host-derived or absorbed metabolites, nor localize metabolite production within the gut. Future work will benefit from integrative multi-omics approaches and experimental validation to test proposed mechanisms, including gut environment alterations (pH, oxygen level, redox potential), ecological interactions underlying microbiome restoration and persistence, and the role of microbe-derived metabolites (tryptophan metabolites) in asthma susceptibility.

Conclusion

This PhD thesis set out to elucidate how CS delivery influences early-life gut microbiome development, its functional capacity, and to determine the role of persistent CS-associated microbial signatures in the association with the risk of childhood asthma.

Collectively, our findings demonstrated that CS delivery affects the gut microbiome from initial seeding through subsequent developmental trajectories, with compositional difference persisting until 1 year of age, despite substantial convergence in core functional capacity. Importantly, it was the prolonged microbial perturbation, and possibly specific taxa, rather than core functions of the microbiome, that consistently associated with an increased risk of childhood asthma. This association was mediated by the elevated tryptophan metabolites which co-occurred with a higher abundance of *Ruminococcus gnavus*. Finally, the presence of older siblings at birth was associated with the restoration of a perturbed gut microbiome, mediated in part by the enrichment of *Bifidobacterium*.

References

1. Todman, D. A history of caesarean section: From ancient world to the modern era. *Aust. N. Z. J. Obstet. Gynaecol.* **47**, 357–361 (2007).
2. Cesarean Section - A Brief History: Part 1. https://www.nlm.nih.gov/exhibition/cesarean/part1.html.
3. “Cesarean Section -- A Brief History” (1993), by Jane Eliot Sewell | Embryo Project Encyclopedia. https://embryo.asu.edu/pages/cesarean-section-brief-history-1993-jane-eliot-sewell.
4. Tostado Fernandez, F. J. Caesarean section: history of a surgical procedure that has always been with us. *Eur. Gynecol. Obstet.* **6**, 121 (2024).
5. Flamm, B. L. Once a cesarean, always a controversy. *Obstet. Gynecol.* **90**, 312–315 (1997).
6. Allen, F. R. W. K. Cæsarean Section. *Indian Med. Gaz.* **80**, 406–412 (1945).
7. Asgarian, A., Rahmati, N., Nasiri, F. & Mohammadbeigi, A. The Failure Rate, Related Factors, and Neonate Complications of Vaginal Delivery after Cesarean Section. *Iran. J. Nurs. Midwifery Res.* **25**, 65–70 (2019).
8. Hesseltine, H. C. & Thelen, C. Sulfonamides as a prophylactic agent in conjunction with cesarean section. *Am. J. Obstet. Gynecol.* **52**, 813–816 (1946).
9. Cesarean Section - A Brief History: Part 3. https://www.nlm.nih.gov/exhibition/cesarean/part3.html.
10. Gosset, M., Ilenko, A., Bouyou, J. & Renevier, B. Emergency caesarean section. *J. Visc. Surg.* **154**, 47–50 (2017).
11. Jaan, F., Qureshi, A. & Reddy, S. S. M. Beyond Medical Indications: Maternal Reasons for Elective Cesarean Sections. *Eur. J. Cardiovasc. Med.* **15**, 600–604 (2025).
12. Shaterian, N., Alsadat Rahnemaei, F., Ghavidel, N. & Abdi, F. Elective cesarean section on maternal request without indication: reasons for it, and its advantages and disadvantages. *Cent. Eur. J. Nurs. Midwifery* **12**, 458–469 (2021).
13. Jenabi, E., Khazaei, S., Bashirian, S., Aghababaei, S. & Matinnia, N. Reasons for elective cesarean section on maternal request: a systematic review. *J. Matern. Fetal Neonatal Med.* **33**, 3867–3872 (2020).
14. Betran, A. *et al.* WHO Statement on Caesarean Section Rates. *Bjog* **123**, 667–670 (2016).
15. Keag, O. E., Norman, J. E. & Stock, S. J. Long-term risks and benefits associated with cesarean delivery for mother, baby, and subsequent pregnancies: Systematic review and meta-analysis. *PLOS Med.* **15**, e1002494 (2018).
16. Darabi, B., Rahmati, S., HafeziAhmadi, M. R., Badfar, G. & Azami, M. The association between caesarean section and childhood asthma: an updated systematic review and meta-analysis. *Allergy Asthma Clin. Immunol.* **15**, 62 (2019).
17. Zhou, C. *et al.* Risk of early childhood overweight/obesity following cesarean section: A systematic review and meta-analysis. *Int. J. Gynecol. Obstet.* **n/a**,.
18. WHO statement on caesarean section rates. https://www.who.int/publications/i/item/WHO-RHR-15.02.
19. World Asthma Day 2025. *Global Initiative for Asthma - GINA* https://ginasthma.org/wad-2025/.

20. Asthma 2025. https://www.thelancet.com/series-do/asthma-2025.
21. Koppelman, G. H. *et al.* Genetic and environmental risk factors for asthma: towards prevention. *Lancet Respir. Med.* **13**, 1011–1025 (2025).
22. El-Husseini, Z. W., Gosens, R., Dekker, F. & Koppelman, G. H. The genetics of asthma and the promise of genomics-guided drug target discovery. *Lancet Respir. Med.* **8**, 1045–1056 (2020).
23. Melén, E. *et al.* Asthma Inception: Epidemiologic Risk Factors and Natural History Across the Life Course. *Am. J. Respir. Crit. Care Med.* **210**, 737–754 (2024).
24. Naeem, A. & Silveyra, P. Sex Differences in Paediatric and Adult Asthma. *Eur. Med. J. Chelmsf. Engl.* **4**, 27–35 (2019).
25. Wenzel, S. E. Asthma phenotypes: the evolution from clinical to molecular approaches. *Nat. Med.* **18**, 716–725 (2012).
26. Wenzel, S. E. Asthma: defining of the persistent adult phenotypes. *The Lancet* **368**, 804–813 (2006).
27. Anderson, G. P. Endotyping asthma: new insights into key pathogenic mechanisms in a complex, heterogeneous disease. *The Lancet* **372**, 1107–1119 (2008).
28. Trivedi, M. & Denton, E. Asthma in Children and Adults—What Are the Differences and What Can They Tell us About Asthma? *Front. Pediatr.* **7**, (2019).
29. Phenotypic Differences between Pediatric and Adult Asthma. https://www.atsjournals.org/doi/epdf/10.1513/pats.200906-046DP?role=tab doi:10.1513/pats.200906-046DP.
30. Foppiano, F. & Schaub, B. Childhood asthma phenotypes and endotypes: a glance into the mosaic. *Mol. Cell. Pediatr.* **10**, 9 (2023).
31. Ng, M. C. W. & How, C. H. Recurrent wheeze and cough in young children: is it asthma? *Singapore Med. J.* **55**, 236–241 (2014).
32. Raaymakers, M. J. A. *et al.* Episodic viral wheeze and multiple-trigger wheeze in preschool children are neither distinct nor constant patterns. A prospective multicenter cohort study in secondary care. *Pediatr. Pulmonol.* **54**, 1439–1446 (2019).
33. Gerday, S. *et al.* Revisiting differences between atopic and non-atopic asthmatics: When age is shaping airway inflammatory profile. *World Allergy Organ. J.* **15**, 100655 (2022).
34. Le, A. V. & Simon, R. A. The Difficult-to-Control Asthmatic: A Systematic Approach. *Allergy Asthma Clin. Immunol.* **2**, 109 (2006).
35. Brown, K. R. *et al.* Endotypes of difficult-to-control asthma in inner-city African American children. *PLOS ONE* **12**, e0180778 (2017).
36. Howrylak, J. A. *et al.* Classification of childhood asthma phenotypes and long-term clinical responses to inhaled anti-inflammatory medications. *J. Allergy Clin. Immunol.* **133**, 1289-1300.e12 (2014).
37. Skov, F. R. *et al.* Type 2-high airway inflammation in childhood asthma distinguishes a more severe phenotype. *Pediatr. Allergy Immunol.* **36**, e70032 (2025).
38. Mishra, P. E., Melén, E., Koppelman, G. H. & Celedón, J. C. T2-low Asthma in School Aged Children: The “Elephant in the Room”. *Lancet Respir. Med.* **11**, 1044–1045 (2023).
39. Fainardi, V., Passadore, L., Labate, M., Pisi, G. & Esposito, S. An Overview of the Obese-Asthma Phenotype in Children. *Int. J. Environ. Res. Public. Health* **19**, 636 (2022).

40. Golebski, K. *et al.* Childhood asthma in the new omics era: challenges and perspectives. *Curr. Opin. Allergy Clin. Immunol.* **20**, 155–161 (2020).
41. Forsström, V. *et al.* Association of Asthma Risk Alleles With Acute Respiratory Tract Infections and Wheezing Illnesses in Young Children. *J. Infect. Dis.* **228**, 990–998 (2023).
42. Ntontsi, P., Photiades, A., Zervas, E., Xanthou, G. & Samitas, K. Genetics and Epigenetics in Asthma. *Int. J. Mol. Sci.* **22**, 2412 (2021).
43. Ahanchian, H., Jones, C. M., Chen, Y. & Sly, P. D. Respiratory viral infections in children with asthma: do they matter and can we prevent them? *BMC Pediatr.* **12**, 147 (2012).
44. Nguyen, V. T. *et al.* The usual suspects: Respiratory syncytial virus and rhinovirus drive asthma development and exacerbation. *Ann. Allergy. Asthma. Immunol.* **135**, 261–267 (2025).
45. Agache, I. *et al.* The Impact of Rhinovirus, Syncytial Respiratory Virus and Helminth Infection on the Risk of New-Onset Asthma and Other Allergic Conditions—A Systematic Review for the EAACI Guidelines on Environmental Science for Allergic Diseases and Asthma. *Allergy* **80**, 1878–1898 (2025).
46. Zanobetti, A. *et al.* Early-Life Exposure to Air Pollution and Childhood Asthma Cumulative Incidence in the ECHO CREW Consortium. *JAMA Netw. Open* **7**, e240535 (2024).
47. Melaram, R. Early life exposures of childhood asthma and allergies—an epidemiologic perspective. *Front. Allergy* **5**, (2024).
48. Sunde, R. B. *et al.* Prenatal tobacco exposure and risk of asthma and allergy outcomes in childhood. *Eur. Respir. J.* **59**, 2100453 (2022).
49. Farber, H. J., Batsell, R. R., Silveira, E. A., Calhoun, R. T. & Giardino, A. P. The Impact of Tobacco Smoke Exposure on Childhood Asthma in a Medicaid Managed Care Plan. *Chest* **149**, 721–728 (2016).
50. Akar-Ghibril, N. & Phipatanakul, W. The Indoor Environment and Childhood Asthma. *Curr. Allergy Asthma Rep.* **20**, 43 (2020).
51. Gold, D. R. Environmental tobacco smoke, indoor allergens, and childhood asthma. *Environ. Health Perspect.* **108 Suppl 4**, 643–651 (2000).
52. Huang, X., Huang, Z., Zhang, J. & Jiang, Y. Maternal gestational diabetes mellitus and the childhood asthma in offspring: a meta-analysis. *Ital. J. Pediatr.* **49**, 139 (2023).
53. Zou, Z., Liu, W., Huang, C., Sun, C. & Zhang, J. First-Year Antibiotics Exposure in Relation to Childhood Asthma, Allergies, and Airway Illnesses. *Int. J. Environ. Res. Public. Health* **17**, 5700 (2020).
54. Loewen, K., Monchka, B., Mahmud, S. M., Jong, G. 't & Azad, M. B. Prenatal antibiotic exposure and childhood asthma: a population-based study. *Eur. Respir. J.* **52**, (2018).
55. Lamont, R. F., Møller Luef, B. & Stener Jørgensen, J. Childhood inflammatory and metabolic disease following exposure to antibiotics in pregnancy, antenatally, intrapartum and neonatally. *F1000Research* **9**, F1000 Faculty Rev-144 (2020).
56. Zhou, W. & Tang, J. Prevalence and risk factors for childhood asthma: a systematic review and meta-analysis. *BMC Pediatr.* **25**, 50 (2025).
57. Chu, S. *et al.* Cesarean section without medical indication and risk of childhood asthma, and attenuation by breastfeeding. *PLoS ONE* **12**, e0184920 (2017).
58. Al Yassen, A., Al-Asadi, J. & Khalaf, S. The role of Caesarean section in childhood asthma. *Malays. Fam. Physician Off. J. Acad. Fam. Physicians Malays.* **14**, 10–17 (2019).

59. Kero, J. *et al.* Mode of Delivery and Asthma – Is There a Connection? *Pediatr. Res.* **52**, 6–11 (2002).
60. Sevelsted, A., Stokholm, J. & Bisgaard, H. Risk of Asthma from Cesarean Delivery Depends on Membrane Rupture. *J. Pediatr.* **171**, 38-42.e4 (2016).
61. Almqvist, C., Cnattingius, S., Lichtenstein, P. & Lundholm, C. The impact of birth mode of delivery on childhood asthma and allergic diseases–a sibling study. *Clin. Exp. Allergy* **42**, 1369–1376 (2012).
62. Wypych-Ślusarska, A. *et al.* Caesarean delivery and risk of childhood asthma development: meta-analysis. *Adv. Dermatol. Allergol. Dermatol. Alergol.* **38**, 819–826 (2021).
63. Boker, F., Alzahrani, A., Alsaeed, A., Alzhrani, M. & Albar, R. Cesarean Section and Development of Childhood Bronchial Asthma: Is There A Risk? *Open Access Maced. J. Med. Sci.* **7**, 347–351 (2019).
64. Zhong, Z. *et al.* Association of cesarean section with asthma in children/adolescents: a systematic review and meta-analysis based on cohort studies. *BMC Pediatr.* **23**, 571 (2023).
65. Liu, C. *et al.* Immunomodulatory roles of butyrate in asthma: mechanisms and therapeutic potentials. *Front. Immunol.* **16**, 1639606 (2025).
66. Li, Y. *et al.* Epigenomic and proteomic analyses provide insights into early-life immune regulation and asthma development in infants. *Nat. Commun.* **16**, 3556 (2025).
67. Idiiatullina, E. & Parker, D. Trained immunity in the lung. *eLife* **14**, e104918.
68. Li, Y. *et al.* Plasmacytoid dendritic cells alleviate allergic asthma via airway epithelial cell–dependent thymosin β4 expression. *J. Allergy Clin. Immunol.* **156**, 171–185 (2025).
69. Budden, K. F. *et al.* Emerging pathogenic links between microbiota and the gut–lung axis. *Nat. Rev. Microbiol.* **15**, 55–63 (2017).
70. Pugin, J. & Chevrolet, J. C. [The intestine-liver-lung axis in septic syndrome]. *Schweiz. Med. Wochenschr.* **121**, 1538–1544 (1991).
71. Song, X. *et al.* Gut-lung axis and asthma: A historical review on mechanism and future perspective. *Clin. Transl. Allergy* **14**, e12356 (2024).
72. Yamamura, R., Inoue, K. Y., Nishino, K. & Yamasaki, S. Intestinal and fecal pH in human health. *Front. Microbiomes* **2**, (2023).
73. Henrick, B. M. *et al.* Elevated Fecal pH Indicates a Profound Change in the Breastfed Infant Gut Microbiome Due to Reduction of Bifidobacterium over the Past Century. *mSphere* **3**, 10.1128/msphere.00041-18 (2018).
74. Penders, J. *et al.* Factors Influencing the Composition of the Intestinal Microbiota in Early Infancy. *Pediatrics* **118**, 511–521 (2006).
75. Sanidad, K. Z. & Zeng, M. Y. Neonatal gut microbiome and immunity. *Curr. Opin. Microbiol.* **56**, 30–37 (2020).
76. Podlesny, D. & Fricke, W. F. Strain inheritance and neonatal gut microbiota development: A meta-analysis. *Int. J. Med. Microbiol.* **311**, 151483 (2021).
77. Rath, E. & Haller, D. Intestinal epithelial cell metabolism at the interface of microbial dysbiosis and tissue injury. *Mucosal Immunol.* **15**, 595–604 (2022).
78. Friedman, E. S. *et al.* Microbes vs. chemistry in the origin of the anaerobic gut lumen. *Proc. Natl. Acad. Sci.* **115**, 4170–4175 (2018).

79. Albenberg, L. *et al.* Correlation Between Intraluminal Oxygen Gradient and Radial Partitioning of Intestinal Microbiota in Humans and Mice. *Gastroenterology* **147**, 1055-1063.e8 (2014).
80. Glover, L. E., Lee, J. S. & Colgan, S. P. Oxygen metabolism and barrier regulation in the intestinal mucosa. *J. Clin. Invest.* **126**, 3680–3688.
81. Walaas, G. A. *et al.* Physiological hypoxia improves growth and functional differentiation of human intestinal epithelial organoids. *Front. Immunol.* **14**, (2023).
82. Chikina, A. & Matic Vignjevic, D. At the right time in the right place: How do luminal gradients position the microbiota along the gut? *Cells Dev.* **168**, 203712 (2021).
83. Herath, M., Hosie, S., Bornstein, J. C., Franks, A. E. & Hill-Yardin, E. L. The Role of the Gastrointestinal Mucus System in Intestinal Homeostasis: Implications for Neurological Disorders. *Front. Cell. Infect. Microbiol.* **10**, 248 (2020).
84. Johansson, M. E. V. *et al.* The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. *Proc. Natl. Acad. Sci.* **105**, 15064–15069 (2008).
85. Johansson, M. E. V., Larsson, J. M. H. & Hansson, G. C. The two mucus layers of colon are organized by the MUC2 mucin, whereas the outer layer is a legislator of host–microbial interactions. *Proc. Natl. Acad. Sci.* **108**, 4659–4665 (2011).
86. Barmpatsalou, V. *et al.* Physiological properties, composition and structural profiling of porcine gastrointestinal mucus. *Eur. J. Pharm. Biopharm.* **169**, 156–167 (2021).
87. Ward, J. B. J., Keely, S. J. & Keely, S. J. Oxygen in the regulation of intestinal epithelial transport. *J. Physiol.* **592**, 2473–2489 (2014).
88. Ouwerkerk, J. P. *et al.* Adaptation of Akkermansia muciniphila to the Oxic-Anoxic Interface of the Mucus Layer. *Appl. Environ. Microbiol.* **82**, 6983–6993 (2016).
89. Silva, A. P. V. *et al.* Targeting the Microbiota Reverses C-Section-Induced Effects on Intestinal Permeability, Microbiota Composition, and Amygdala Gene Expression in the Mouse. *Neurogastroenterol. Motil.* **37**, e70107 (2025).
90. Lay, C. *et al.* A synbiotic intervention modulates meta-omics signatures of gut redox potential and acidity in elective caesarean born infants. *BMC Microbiol.* **21**, 191 (2021).
91. Kennedy, K. M. *et al.* Questioning the fetal microbiome illustrates pitfalls of low-biomass microbial studies. *Nature* **613**, 639–649 (2023).
92. Collado, M. C., Rautava, S., Aakko, J., Isolauri, E. & Salminen, S. Human gut colonisation may be initiated in utero by distinct microbial communities in the placenta and amniotic fluid. *Sci. Rep.* **6**, 23129 (2016).
93. Walter, J. & Hornef, M. W. A philosophical perspective on the prenatal in utero microbiome debate. *Microbiome* **9**, 5 (2021).
94. Betkowski, B. New research refutes claims that humans are colonized by bacteria before birth. https://www.ualberta.ca/en/folio/2023/01/new-research-refutes-claims-that-humans-are-colonized-by-bacteria-before-birth.html.
95. Mitchell, C. M. *et al.* Delivery Mode Affects Stability of Early Infant Gut Microbiota. *Cell Rep. Med.* **1**, (2020).
96. Li, W. *et al.* Vertical Transmission of Gut Microbiome and Antimicrobial Resistance Genes in Infants Exposed to Antibiotics at Birth. *J. Infect. Dis.* **224**, 1236–1246 (2021).

97. Ronde, E. *et al.* The influence of maternal gut and vaginal microbiota on gastrointestinal colonization of neonates born vaginally and per caesarean section. *BMC Pregnancy Childbirth* **25**, 254 (2025).
98. Studying Vertical Microbiome Transmission from Mothers to Infants by Strain-Level Metagenomic Profiling. https://journals.asm.org/doi/epdf/10.1128/msystems.00164-16 doi:10.1128/mSystems.00164-16.
99. Browne, H. P., Shao, Y. & Lawley, T. D. Mother–infant transmission of human microbiota. *Curr. Opin. Microbiol.* **69**, 102173 (2022).
100. Wang, S. *et al.* Metagenomic analysis of mother-infant gut microbiome reveals global distinct and shared microbial signatures. *Gut Microbes* **13**, 1911571 (2021).
101. Ozcan, G. *et al.* The vaginal microbiome composition during pregnancy in a region compromising different ethnic origins. *Arch. Gynecol. Obstet.* **310**, 369–375 (2024).
102. The Vaginal Microbiota of Pregnant Women Varies with Gestational Age, Maternal Age, and Parity. https://journals.asm.org/doi/epdf/10.1128/spectrum.03429-22 doi:10.1128/spectrum.03429-22.
103. Feehily, C. *et al.* Detailed mapping of Bifidobacterium strain transmission from mother to infant via a dual culture-based and metagenomic approach. *Nat. Commun.* **14**, 3015 (2023).
104. Ecological niche. *Wikipedia* (2025).
105. Spragge, F. *et al.* Microbiome diversity protects against pathogens by nutrient blocking. *Science* **382**, eadj3502 (2023).
106. Ducarmon, Q. R. *et al.* Gut Microbiota and Colonization Resistance against Bacterial Enteric Infection. *Microbiol. Mol. Biol. Rev. MMBR* **83**, e00007-19 (2019).
107. Shao, Y. *et al.* Primary succession of Bifidobacteria drives pathogen resistance in neonatal microbiota assembly. *Nat. Microbiol.* **9**, 2570–2582 (2024).
108. Michalska-Smith, M. *et al.* Network structure of resource use and niche overlap within the endophytic microbiome. *ISME J.* **16**, 435–446 (2022).
109. Yang, C. *et al.* Life history strategies complement niche partitioning to support the coexistence of closely related *Gilliamella* species in the bee gut. *ISME J.* **19**, wraf016 (2025).
110. Brochet, S. *et al.* Niche partitioning facilitates coexistence of closely related honey bee gut bacteria. *eLife* **10**, e68583 (2021).
111. Marcobal, A. *et al.* Bacteroides in the Infant Gut Consume Milk Oligosaccharides via Mucus-Utilization Pathways. *Cell Host Microbe* **10**, 507–514 (2011).
112. Segura Munoz, R. R. *et al.* Experimental evaluation of ecological principles to understand and modulate the outcome of bacterial strain competition in gut microbiomes. *ISME J.* **16**, 1594–1604 (2022).
113. Sprockett, D., Fukami, T. & Relman, D. A. Role of priority effects in the early-life assembly of the gut microbiota. *Nat. Rev. Gastroenterol. Hepatol.* **15**, 197–205 (2018).
114. Pocheville, A. The Ecological Niche: History and Recent Controversies. in *Handbook of Evolutionary Thinking in the Sciences* (eds Heams, T., Huneman, P., Lecointre, G. & Silberstein, M.) 547–586 (Springer Netherlands, Dordrecht, 2015). doi:10.1007/978-94-017-9014-7_26.
115. Competitive exclusion principle. *Wikipedia* (2025).

116. Elzinga, J. *et al.* Binding of Akkermansia muciniphila to mucin is O-glycan specific. *Nat. Commun.* **15**, 4582 (2024).
117. Qu, S. *et al.* Excessive consumption of mucin by over-colonized Akkermansia muciniphila promotes intestinal barrier damage during malignant intestinal environment. *Front. Microbiol.* **14**, 1111911 (2023).
118. Quinn, R. A. *et al.* Niche partitioning of a pathogenic microbiome driven by chemical gradients. *Sci. Adv.* **4**, eaau1908 (2018).
119. Duerkop, B. A., Clements, C. V., Rollins, D., Rodrigues, J. L. M. & Hooper, L. V. A composite bacteriophage alters colonization by an intestinal commensal bacterium. *Proc. Natl. Acad. Sci.* **109**, 17621–17626 (2012).
120. Solis-Balandra, M. A. & Sanchez-Salas, J. L. Classification and Multi-Functional Use of Bacteriocins in Health, Biotechnology, and Food Industry. *Antibiotics* **13**, 666 (2024).
121. d'Hennezel, E., Abubucker, S., Murphy, L. O. & Cullen, T. W. Total Lipopolysaccharide from the Human Gut Microbiome Silences Toll-Like Receptor Signaling. *mSystems* **2**, e00046-17 (2017).
122. Yin, R. *et al.* Immunogenic molecules associated with gut bacterial cell walls: chemical structures, immune-modulating functions, and mechanisms. *Protein Cell* **14**, 776–785 (2023).
123. Irazoki, O., Hernandez, S. B. & Cava, F. Peptidoglycan Muropeptides: Release, Perception, and Functions as Signaling Molecules. *Front. Microbiol.* **10**, (2019).
124. Wu, Q. Y., Liu, B. C., Ruan, X. Z. & Ma, K. L. Intestinal microbiota-derived membrane vesicles and their role in chronic kidney disease. *Biochim. Biophys. Acta BBA - Mol. Basis Dis.* **1868**, 166478 (2022).
125. Abraham, L., Raise, A., Beney, L., Lapaquette, P. & Rieu, A. Membrane vesicles produced by next-generation probiotics from the gut as innovative tools for human health. *Gut Microbes* **17**, 2552344 (2025).
126. Gosalbes, M. J. *et al.* Metatranscriptomic Approach to Analyze the Functional Human Gut Microbiota. *PLOS ONE* **6**, e17447 (2011).
127. Forster, S. C. *et al.* A human gut bacterial genome and culture collection for improved metagenomic analyses. *Nat. Biotechnol.* **37**, 186–192 (2019).
128. Boudry, G. *et al.* The Relationship Between Breast Milk Components and the Infant Gut Microbiota. *Front. Nutr.* **8**, (2021).
129. Kijner, S., Kolodny, O. & Yassour, M. Human milk oligosaccharides and the infant gut microbiome from an eco-evolutionary perspective. *Curr. Opin. Microbiol.* **68**, 102156 (2022).
130. Laursen, M. F. & Roager, H. M. Human milk oligosaccharides modify the strength of priority effects in the Bifidobacterium community assembly during infancy. *ISME J.* **17**, 2452–2457 (2023).
131. Underwood, M. A., German, J. B., Lebrilla, C. B. & Mills, D. A. Bifidobacterium longum subspecies infantis: champion colonizer of the infant gut. *Pediatr. Res.* **77**, 229–235 (2015).
132. Arzamasov, A. A. & Osterman, A. L. Milk glycan metabolism by intestinal bifidobacteria: insights from comparative genomics. *Crit. Rev. Biochem. Mol. Biol.* **57**, 562–584 (2022).

133. Wichmann, A. Biological effects of combinations of structurally diverse human milk oligosaccharides. *Front. Pediatr.* **12**, 1439612 (2024).
134. Grondin, J. M., Tamura, K., Déjean, G., Abbott, D. W. & Brumer, H. Polysaccharide Utilization Loci: Fueling Microbial Communities. *J. Bacteriol.* **199**, 10.1128/jb.00860-16 (2017).
135. Salli, K. *et al.* Selective Utilization of the Human Milk Oligosaccharides 2′-Fucosyllactose, 3-Fucosyllactose, and Difucosyllactose by Various Probiotic and Pathogenic Bacteria. *J. Agric. Food Chem.* **69**, 170–182 (2021).
136. Luna, E. *et al.* Utilization Efficiency of Human Milk Oligosaccharides by Human-Associated Akkermansia Is Strain Dependent. *Appl. Environ. Microbiol.* **88**, e01487-21 (2022).
137. Raba, G. & Luis, A. S. Mucin utilization by gut microbiota: recent advances on characterization of key enzymes. *Essays Biochem.* **67**, 345–353 (2023).
138. Crouch, L. I. *et al.* Prominent members of the human gut microbiota express endo-acting O-glycanases to initiate mucin breakdown. *Nat. Commun.* **11**, 4017 (2020).
139. Schaus, S. R. *et al. Ruminococcus torques* is a keystone degrader of intestinal mucin glycoprotein, releasing oligosaccharides used by *Bacteroides thetaiotaomicron*. *mBio* **15**, e00039-24 (2024).
140. Berkhout, M. D. *et al.* Mucin-driven ecological interactions in an in vitro synthetic community of human gut microbes. *Glycobiology* **34**, cwae085 (2024).
141. Rowland, I. *et al.* Gut microbiota functions: metabolism of nutrients and other food components. *Eur. J. Nutr.* **57**, 1–24 (2018).
142. Cheng, C. C. *et al.* Ecological Importance of Cross-Feeding of the Intermediate Metabolite 1,2-Propanediol between Bacterial Gut Symbionts. *Appl. Environ. Microbiol.* **86**, e00190-20 (2020).
143. Flores, J. N., Lubin, J.-B. & Silverman, M. A. The case for microbial intervention at weaning. *Gut Microbes* **16**, 2414798 (2024).
144. Lalli, M. K. *et al.* Associations between dietary fibers and gut microbiome composition in the EDIA longitudinal infant cohort. *Am. J. Clin. Nutr.* **121**, 83–99 (2025).
145. Harling, L. C., Hecht, A. L., Meyer, F. & Wu, G. D. Revisiting nitrogen assimilation strategies in the mammalian gut: lessons from Enterobacteriaceae as pathobiont models and a challenge to the limitation paradigm. *Arch. Microbiol.* **207**, 203 (2025).
146. Doranga, S. & Conway, T. Nitrogen assimilation by E. coli in the mammalian intestine. *mBio* **15**, e00025-24 (2024).
147. Bertin, Y. *et al.* Enterohaemorrhagic Escherichia coli gains a competitive advantage by using ethanolamine as a nitrogen source in the bovine intestinal content. *Environ. Microbiol.* **13**, 365–377 (2011).
148. You, X., Rani, A., Özcan, E., Lyu, Y. & Sela, D. A. Bifidobacterium longum subsp. infantis utilizes human milk urea to recycle nitrogen within the infant gut microbiome. *Gut Microbes* **15**, 2192546 (2023).
149. Schimmel, P., Kleinjans, L., Bongers, R. S., Knol, J. & Belzer, C. Breast milk urea as a nitrogen source for urease positive Bifidobacterium infantis. *FEMS Microbiol. Ecol.* **97**, fiab019 (2021).

150. Wang, J., Yan, D., Dixon, R. & Wang, Y.-P. Deciphering the Principles of Bacterial Nitrogen Dietary Preferences: a Strategy for Nutrient Containment. *mBio* **7**, e00792-16 (2016).
151. Gao, Y. *et al.* Ammonia-Assimilating Bacteria Promote Wheat (Triticum aestivum) Growth and Nitrogen Utilization. *Microorganisms* **13**, 43 (2024).
152. Harper, C. J., Hayward, D., Kidd, M., Wiid, I. & van Helden, P. Glutamate dehydrogenase and glutamine synthetase are regulated in response to nitrogen availability in Myocbacterium smegmatis. *BMC Microbiol.* **10**, 138 (2010).
153. Fecal Microflora in Healthy Infants Born by Different Methods of Delivery: Permanent Changes in Intestinal Flora After Cesarean Delivery - Grölund - 1999 - Journal of Pediatric Gastroenterology and Nutrition - Wiley Online Library. https://onlinelibrary-wiley-com.kb-ku.idm.oclc.org/doi/10.1002/j.1536-4801.1999.tb01997.x.
154. Dominguez-Bello, M. G. *et al.* Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc. Natl. Acad. Sci.* **107**, 11971–11975 (2010).
155. Stewart, C. J. *et al.* Temporal development of the gut microbiome in early childhood from the TEDDY study. *Nature* **562**, 583–588 (2018).
156. Mueller, N. T., Differding, M. K., Østbye, T., Hoyo, C. & Benjamin-Neelon, S. E. Cesarean delivery is associated with decreased beneficial gut microbes and increased pathobionts and butyrate excretion: a longitudinal study over the first year of life. *BJOG Int. J. Obstet. Gynaecol.* **128**, 1293–1303 (2021).
157. Stokholm, J. *et al.* Delivery mode and gut microbial changes correlate with an increased risk of childhood asthma. *Sci. Transl. Med.* **12**, eaax9929 (2020).
158. Sanam, M. *et al.* Bridging two worlds: Host Microbiota crosstalk in health and dysregulation. *Innate Immun.* **31**, 17534259251392993 (2025).
159. Levy, M., Thaiss, C. A. & Elinav, E. Metabolites: messengers between the microbiota and the immune system. *Genes Dev.* **30**, 1589–1597 (2016).
160. Kalyanaraman, B., Cheng, G. & Hardy, M. Gut microbiome, short-chain fatty acids, alpha-synuclein, neuroinflammation, and ROS/RNS: Relevance to Parkinson's disease and therapeutic implications. *Redox Biol.* **71**, 103092 (2024).
161. Liu, H. *et al.* Butyrate: A Double-Edged Sword for Health? *Adv. Nutr.* **9**, 21–29 (2018).
162. Kopczyńska, J. & Kowalczyk, M. The potential of short-chain fatty acid epigenetic regulation in chronic low-grade inflammation and obesity. *Front. Immunol.* **15**, 1380476 (2024).
163. O'Rourke, L. *et al.* Tryptophan metabolic profile in term and preterm breast milk: implications for health. *J. Nutr. Sci.* **7**, e13 (2018).
164. Gao, K., Mu, C., Farzi, A. & Zhu, W. Tryptophan Metabolism: A Link Between the Gut Microbiota and Brain. *Adv. Nutr.* **11**, 709–723 (2020).
165. Liaqat, H., Parveen, A. & Kim, S. Y. Neuroprotective Natural Products' Regulatory Effects on Depression via Gut–Brain Axis Targeting Tryptophan. *Nutrients* **14**, 3270 (2022).
166. Owe-Larsson, M. *et al.* Microbiota-Derived Tryptophan Metabolite Indole-3-Propionic Acid-Emerging Role in Neuroprotection. *Molecules* **30**, 3628 (2025).
167. Larabi, A. B., Masson, H. L. P. & Bäumler, A. J. Bile acids as modulators of gut microbiota composition and function. *Gut Microbes* **15**, 2172671.

168. Tarracchini, C. *et al.* Exploring the vitamin biosynthesis landscape of the human gut microbiota. *mSystems* **9**, e00929-24 (2024).
169. Loh, J. S. *et al.* Microbiota–gut–brain axis and its therapeutic applications in neurodegenerative diseases. *Signal Transduct. Target. Ther.* **9**, 37 (2024).
170. Yang, W. & Cong, Y. Gut microbiota-derived metabolites in the regulation of host immune responses and immune-related inflammatory diseases. *Cell. Mol. Immunol.* **18**, 866–877 (2021).
171. Hayes, J. A. *et al.* Engineered bacteria titrate hydrogen sulfide and induce concentration-dependent effects on host in a gut microphysiological system. 2023.05.16.538950 Preprint at https://doi.org/10.1101/2023.05.16.538950 (2023).
172. Bryan, N. S. & Lefer, D. J. Update on Gaseous Signaling Molecules Nitric Oxide and Hydrogen Sulfide: Strategies to Capture their Functional Activity for Human Therapeutics. *Mol. Pharmacol.* **96**, 109–114 (2019).
173. Wu, D., Hu, Q. & Zhu, D. An Update on Hydrogen Sulfide and Nitric Oxide Interactions in the Cardiovascular System. *Oxid. Med. Cell. Longev.* **2018**, 4579140 (2018).
174. Li, D. & Wu, M. Pattern recognition receptors in health and diseases. *Signal Transduct. Target. Ther.* **6**, 291 (2021).
175. Iyer, K., Erkert, L. & Becker, C. Know your neighbors: microbial recognition at the intestinal barrier and its implications for gut homeostasis and inflammatory bowel disease. *Front. Cell Dev. Biol.* **11**, 1228283 (2023).
176. Bersch, K. L. *et al.* Bacterial Peptidoglycan Fragments Differentially Regulate Innate Immune Signaling. *ACS Cent. Sci.* **7**, 688–696 (2021).
177. Rhee, S. H. *et al.* Pathophysiological role of Toll-like receptor 5 engagement by bacterial flagellin in colonic inflammation. *Proc. Natl. Acad. Sci.* **102**, 13610–13615 (2005).
178. Vijay-Kumar, M. & Gewirtz, A. T. Flagellin: key target of mucosal innate immunity. *Mucosal Immunol.* **2**, 197–205 (2009).
179. Su, Y. & Ding, T. Targeting microbial quorum sensing: the next frontier to hinder bacterial driven gastrointestinal infections. *Gut Microbes* **15**, 2252780.
180. Mottawea, W. *et al.* Multi-level analysis of gut microbiome extracellular vesicles-host interaction reveals a connection to gut-brain axis signaling. *Microbiol. Spectr.* **13**, e01368-24 (2024).
181. Qian, D. *et al.* Bacterial extracellular vesicles for gut microbiome–host communication and drug development. *Acta Pharm. Sin. B* **15**, 1816–1840 (2025).
182. Ouyang, R. *et al.* Maturation of the gut metabolome during the first year of life in humans. *Gut Microbes* **15**, 2231596.
183. Vidarsdottir, H. *et al.* Mode of delivery was associated with transient changes in the metabolomic profile of neonates. *Acta Paediatr.* **110**, 2110–2118 (2021).
184. Adeyeye, T. E. *et al.* Effects on Neonatal Immunoglobulin Concentrations by Infant Mode of Delivery in the Upstate KIDS Study (2008–2010). *Am. J. Reprod. Immunol. N. Y. N 1989* **89**, e13688 (2023).
185. Neu, J. & Rushing, J. Cesarean versus Vaginal Delivery: Long term infant outcomes and the Hygiene Hypothesis. *Clin. Perinatol.* **38**, 321–331 (2011).
186. Wibowo, M. C. *et al.* Reconstruction of ancient microbial genomes from the human gut. *Nature* **594**, 234–239 (2021).

187. Blaser, M. J. The Past and Future Biology of the Human Microbiome in an Age of Extinctions. *Cell* **172**, 1173–1177 (2018).
188. Nunez, H. *et al.* Early life gut microbiome and its impact on childhood health and chronic conditions. *Gut Microbes* **17**, 2463567 (2025).
189. Gong, Y. *et al.* Effect of Probiotic Supplementation on the Gut Microbiota Composition of Infants Delivered by Cesarean Section: An Exploratory, Randomized, Open-label, Parallel-controlled Trial. *Curr. Microbiol.* **80**, 341 (2023).
190. Wang, Y. *et al.* Restoration of gut microbiota with a specific synbiotic-containing infant formula in healthy Chinese infants born by cesarean section. *Eur. J. Clin. Nutr.* **79**, 567–575 (2025).
191. Wilson, B. C. *et al.* Oral administration of maternal vaginal microbes at birth to restore gut microbiome development in infants born by caesarean section: A pilot randomised placebo-controlled trial. *EBioMedicine* **69**, 103443 (2021).
192. Zhou, L. *et al.* Effects of vaginal microbiota transfer on the neurodevelopment and microbiome of cesarean-born infants: A blinded randomized controlled trial. *Cell Host Microbe* **31**, 1232-1247.e5 (2023).
193. Wang, X. *et al.* Impact of vaginal seeding on the gut microbiome of infants born via cesarean section: A systematic review. *J. Infect.* **89**, 106348 (2024).
194. Song, S. J. *et al.* Naturalization of the microbiota developmental trajectory of Cesarean-born neonates after vaginal seeding. *Med* **2**, 951-964.e5 (2021).
195. Korpela, K. & de Vos, W. M. Infant gut microbiota restoration: state of the art. *Gut Microbes* **14**, 2118811 (2022).
196. Davis, E. C. *et al.* Gut microbiome and breast-feeding: Implications for early immune development. *J. Allergy Clin. Immunol.* **150**, 523–534 (2022).
197. Yang, Z. *et al.* Rural environment reduces allergic inflammation by modulating the gut microbiota. *Gut Microbes* **14**, 2125733 (2022).
198. Gao, Y. *et al.* Gut microbiota maturity mediates the protective effect of siblings on food allergy. *J. Allergy Clin. Immunol.* **152**, 667–675 (2023).
199. Ames, S. R., Lotoski, L. C. & Azad, M. B. Comparing early life nutritional sources and human milk feeding practices: personalized and dynamic nutrition supports infant gut microbiome development and immune system maturation. *Gut Microbes* **15**, 2190305 (2023).
200. Anumula, S., Nalla, K., Pandala, P., Kotha, R. & Harsha, N. Rural Versus Urban Mothers' Microbiome Difference and Its Effect on Neonates: A Systematic Review. *Cureus* **16**, e55607.
201. Laursen, M. F. *et al.* Having older siblings is associated with gut microbiota development during early childhood. *BMC Microbiol.* **15**, 154 (2015).
202. Bisgaard, H. *et al.* Deep phenotyping of the unselected COPSAC2010 birth cohort study. *Clin. Exp. Allergy* **43**, 1384–1394 (2013).
203. Brand, P. L. P. *et al.* Definition, assessment and treatment of wheezing disorders in preschool children: an evidence-based approach. *Eur. Respir. J.* **32**, 1096–1110 (2008).
204. Bisgaard, H. *et al.* Childhood Asthma after Bacterial Colonization of the Airway in Neonates. *N. Engl. J. Med.* **357**, 1487–1495 (2007).

205. Bisgaard, H. *et al.* Fish Oil-Derived Fatty Acids in Pregnancy and Wheeze and Asthma in Offspring. *N. Engl. J. Med.* **375**, 2530–2539 (2016).
206. Chawes, B. L. *et al.* Effect of Vitamin D_3 Supplementation During Pregnancy on Risk of Persistent Wheeze in the Offspring: A Randomized Clinical Trial. *JAMA* **315**, 353 (2016).
207. Stokholm, J. *et al.* Maturation of the gut microbiome and risk of asthma in childhood. *Nat. Commun.* **9**, 141 (2018).
208. Thorsen, J. *et al.* The airway microbiota of neonates colonized with asthma-associated pathogenic bacteria. *Nat. Commun.* **14**, 6668 (2023).
209. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).
210. Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
211. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
212. Bhattacharyya, M. *et al.* AnnotIEM: novel tool for microbiome species-level annotation of 16S gene based microbial sequencing. 2024.10.31.621263 Preprint at https://doi.org/10.1101/2024.10.31.621263 (2024).
213. Yoon, S.-H. *et al.* Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int. J. Syst. Evol. Microbiol.* **67**, 1613–1617 (2017).
214. Federhen, S. The NCBI Taxonomy database. *Nucleic Acids Res.* **40**, D136–D143 (2012).
215. Cole, J. R. *et al.* The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res.* **37**, D141-145 (2009).
216. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590-596 (2013).
217. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
218. Blanco-Míguez, A. *et al.* Extending and improving metagenomic taxonomic profiling with uncharacterized species using MetaPhlAn 4. *Nat. Biotechnol.* **41**, 1633–1644 (2023).
219. Beghini, F. *et al.* Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *eLife* **10**, e65088 (2021).
220. Subbarao, P. *et al.* The Canadian Healthy Infant Longitudinal Development (CHILD) Study: examining developmental origins of allergy and asthma: Table 1. *Thorax* **70**, 998–1000 (2015).
221. Amir, A. *et al.* Deblur Rapidly Resolves Single-Nucleotide Community Sequence Patterns. *mSystems* **2**, 10.1128/msystems.00191-16 (2017).
222. DeSantis, T. Z. *et al.* Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. *Appl. Environ. Microbiol.* **72**, 5069–5072 (2006).
223. McDonald, D. *et al.* An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J.* **6**, 610–618 (2012).
224. Ennis, D., Shmorak, S., Jantscher-Krenn, E. & Yassour, M. Longitudinal quantification of Bifidobacterium longum subsp. infantis reveals late colonization in the infant gut independent of maternal milk HMO composition. *Nat. Commun.* **15**, 894 (2024).

225. Hoskinson, C. *et al.* Delayed gut microbiota maturation in the first year of life is a hallmark of pediatric allergic disease. *Nat. Commun.* **14**, 4785 (2023).
226. Shannon, C. E. A Mathematical Theory of Communication.
227. Lozupone, C. A., Hamady, M., Kelley, S. T. & Knight, R. Quantitative and Qualitative β Diversity Measures Lead to Different Insights into Factors That Structure Microbial Communities. *Appl. Environ. Microbiol.* **73**, 1576–1585 (2007).
228. Rohart, F., Gautier, B., Singh, A. & Cao, K.-A. L. mixOmics: An R package for ‘omics feature selection and multiple data integration. *PLOS Comput. Biol.* **13**, e1005752 (2017).
229. Thorsen, J. Implementation and utility functions for using the mixOmics package in the caret framework: github.com/jonathanth/mixOmicsCaret. (2017).
230. Weaver, L. T., Laker, M. F. & Nelson, R. Intestinal permeability in the newborn. *Arch. Dis. Child.* **59**, 236–241 (1984).
231. Hanning, N. *et al.* Intestinal barrier dysfunction in irritable bowel syndrome: a systematic review. *Ther. Adv. Gastroenterol.* **14**, 1756284821993586 (2021).
232. Zhang, C. *et al.* The Effects of Delivery Mode on the Gut Microbiota and Health: State of Art. *Front. Microbiol.* **12**, 724449 (2021).
233. Shin, H. *et al.* The first microbial environment of infants born by C-section: the operating room microbes. *Microbiome* **3**, 59 (2015).
234. Winstedt, L., Frankenberg, L., Hederstedt, L. & von Wachenfeldt, C. Enterococcus faecalis V583 Contains a Cytochrome bd-Type Respiratory Oxidase. *J. Bacteriol.* **182**, 3863–3866 (2000).
235. Nishikawa, T. *et al.* Effect of Nitric Oxide on the Oxygen Metabolism and Growth of E. faecalis. *J. Clin. Biochem. Nutr.* **44**, 178–184 (2009).
236. Huycke, M. M. & Moore, D. R. In vivo production of hydroxyl radical by enterococcus faecalis colonizing the intestinal tract using aromatic hydroxylation. *Free Radic. Biol. Med.* **33**, 818–826 (2002).
237. Huycke, M. M., Abrams, V. & Moore, D. R. Enterococcus faecalis produces extracellular superoxide and hydrogen peroxide that damages colonic epithelial cell DNA. *Carcinogenesis* **23**, 529–536 (2002).
238. Riboulet, E. *et al.* Relationships between Oxidative Stress Response and Virulence in Enterococcus faecalis. *J. Mol. Microbiol. Biotechnol.* **13**, 140–146 (2007).
239. Portela, C. A. F., Smart, K. F., Tumanov, S., Cook, G. M. & Villas-Bôas, S. G. Global Metabolic Response of Enterococcus faecalis to Oxygen. *J. Bacteriol.* **196**, 2012–2022 (2014).
240. McCartney, A. L. & Hoyles, L. Host interactions of bioactive molecules produced by Klebsiella spp. *Microbiota Host* **3**, (2025).
241. Storey, D. *et al.* Klebsiella pneumoniae type VI secretion system-mediated microbial competition is PhoPQ controlled and reactive oxygen species dependent. *PLOS Pathog.* **16**, e1007969 (2020).
242. Wong Fok Lung, T. *et al. Klebsiella pneumoniae* induces host metabolic stress that promotes tolerance to pulmonary infection. *Cell Metab.* **34**, 761-774.e9 (2022).
243. de Carvalho, M. de M. L. *et al.* Dual Role for Pld1 in Klebsiella pneumoniae Virulence: Transcriptomics and Proteomics Provide Insights into Direct and Indirect Effects. *J. Proteome Res.* **24**, 2874–2884.

244. Flores-Díaz, M., Monturiol-Gross, L., Naylor, C., Alape-Girón, A. & Flieger, A. Bacterial Sphingomyelinases and Phospholipases as Virulence Factors. *Microbiol. Mol. Biol. Rev.* **80**, 597–628 (2016).
245. Lery, L. M. S. *et al.* Comparative analysis of Klebsiella pneumoniae genomes identifies a phospholipase D family protein as a novel virulence factor. *BMC Biol.* **12**, 41 (2014).
246. Songer, J. G. Bacterial phospholipases and their role in virulence. *Trends Microbiol.* **5**, 156–161 (1997).
247. Camargo, A., Ramírez, J. D., Kiu, R., Hall, L. J. & Muñoz, M. Unveiling the pathogenic mechanisms of Clostridium perfringens toxins and virulence factors. *Emerg. Microbes Infect.* **13**, 2341968 (2024).
248. Briolat, V. & Reysset, G. Identification of the Clostridium perfringens Genes Involved in the Adaptive Response to Oxidative Stress. *J. Bacteriol.* **184**, 2333–2343 (2002).
249. Jean, D., Briolat, V. & Reysset, G. Oxidative stress response in Clostridium perfringens. *Microbiology* **150**, 1649–1659 (2004).
250. Machida, Y. *et al.* [An outbreak of enterocolitis due to Clostridium perfringens in a hospital for the severely disabled]. *Kansenshogaku Zasshi* **63**, 410–416 (1989).
251. Kiu, R. & Hall, L. J. An update on the human and animal enteric pathogen Clostridium perfringens. *Emerg. Microbes Infect.* **7**, 141 (2018).
252. Kiu, R. *et al.* Particular genomic and virulence traits associated with preterm infant-derived toxigenic Clostridium perfringens strains. *Nat. Microbiol.* **8**, 1160–1175 (2023).
253. Rood, J. I. *et al.* Expansion of the *Clostridium perfringens* toxin-based typing scheme. *Anaerobe* **53**, 5–10 (2018).
254. Singh, V. *et al.* Phospholipase C: underrated players in microbial infections. *Front. Cell. Infect. Microbiol.* **13**, 1089374 (2023).
255. Sakurai, J., Nagahama, M. & Oda, M. Clostridium perfringens Alpha-Toxin: Characterization and Mode of Action. *J. Biochem. (Tokyo)* **136**, 569–574 (2004).
256. Flores-Díaz, M., Thelestam, M., Clark, G. C., Titball, R. W. & Alape-Girón, A. Effects of *Clostridium perfringens* phospholipase C in mammalian cells. *Anaerobe* **10**, 115–123 (2004).
257. Oda, M. *et al.* Signal transduction mechanism involved in Clostridium perfringens alpha-toxin-induced superoxide anion generation in rabbit neutrophils. *Infect. Immun.* **74**, 2876–2886 (2006).
258. Monturiol-Gross, L. *et al.* Reactive Oxygen Species and the MEK/ERK Pathway Are Involved in the Toxicity of Clostridium perfringens α-Toxin, a Prototype Bacterial Phospholipase C. *J. Infect. Dis.* **206**, 1218–1226 (2012).
259. Monturiol-Gross, L., Flores-Díaz, M., Pineda-Padilla, M. J., Castro-Castro, A. C. & Alape-Giron, A. Clostridium perfringens Phospholipase C Induced ROS Production and Cytotoxicity Require PKC, MEK1 and NFκB Activation. *PLOS ONE* **9**, e86475 (2014).
260. Luis, A. S. & Hansson, G. C. Intestinal mucus and their glycans: A habitat for thriving microbiota. *Cell Host Microbe* **31**, 1087–1100 (2023).
261. Kim, Y. S. & Ho, S. B. Intestinal Goblet Cells and Mucins in Health and Disease: Recent Insights and Progress. *Curr. Gastroenterol. Rep.* **12**, 319–330 (2010).
262. Lin, Q. *et al.* A Review of the Mechanisms of Bacterial Colonization of the Mammal Gut. *Microorganisms* **12**, 1026 (2024).

263. Pelaseyed, T. *et al.* The mucus and mucins of the goblet cells and enterocytes provide the first defense line of the gastrointestinal tract and interact with the immune system. *Immunol. Rev.* **260**, 8–20 (2014).
264. Schoepe, H. *et al.* Naturally Occurring Clostridium perfringens Nontoxic Alpha-Toxin Variant as a Potential Vaccine Candidate against Alpha-Toxin-Associated Diseases. *Infect. Immun.* **69**, 7194–7196 (2001).
265. Bertani, B. & Ruiz, N. Function and Biogenesis of Lipopolysaccharides. *EcoSal Plus* **8**, (2018).
266. Di Lorenzo, F., De Castro, C., Silipo, A. & Molinaro, A. Lipopolysaccharide structures of Gram-negative populations in the gut microbiota and effects on host interactions. *FEMS Microbiol. Rev.* **43**, 257–272 (2019).
267. Sperandeo, P., Martorana, A. M. & Polissi, A. Lipopolysaccharide Biosynthesis and Transport to the Outer Membrane of Gram-Negative Bacteria. in *Bacterial Cell Walls and Membranes* (ed. Kuhn, A.) 9–37 (Springer International Publishing, Cham, 2019). doi:10.1007/978-3-030-18768-2_2.
268. Leech, S. M. *et al.* Delivery mode is a larger determinant of infant gut microbiome composition at 6 weeks than exposure to peripartum antibiotics. *Microb. Genomics* **10**, 001269 (2024).
269. Kawaguchi, K. *et al.* The mannobiose-forming exo-mannanase involved in a new mannan catabolic pathway in Bacteroides fragilis. *Arch. Microbiol.* **196**, 17–23 (2014).
270. Senoura, T. *et al.* New microbial mannan catabolic pathway that involves a novel mannosylglucose phosphorylase. *Biochem. Biophys. Res. Commun.* **408**, 701–706 (2011).
271. Wagenlehner, F., Lorenz, H., Ewald, O. & Gerke, P. Why d-Mannose May Be as Efficient as Antibiotics in the Treatment of Acute Uncomplicated Lower Urinary Tract Infections—Preliminary Considerations and Conclusions from a Non-Interventional Study. *Antibiotics* **11**, 314 (2022).
272. Sarasa, S. B. *et al.* A Brief Review on the Non-protein Amino Acid, Gamma-amino Butyric Acid (GABA): Its Production and Role in Microbes. *Curr. Microbiol.* **77**, 534–544 (2020).
273. Zhang, Q. *et al.* Insights and progress on the biosynthesis, metabolism, and physiological functions of gamma-aminobutyric acid (GABA): a review. *PeerJ* **12**, e18712 (2024).
274. Petroff, O. A. C. Book Review: GABA and Glutamate in the Human Brain. *The Neuroscientist* **8**, 562–573 (2002).
275. Fernández-Veledo, S. & Vendrell, J. Gut microbiota-derived succinate: Friend or foe in human metabolic diseases? *Rev. Endocr. Metab. Disord.* **20**, 439–447 (2019).
276. Rotstein, O. D., Nasmith, P. E. & Grinstein, S. The Bacteroides by-product succinic acid inhibits neutrophil respiratory burst by reducing intracellular pH. *Infect. Immun.* **55**, 864–870 (1987).
277. Strandwitz, P. *et al.* GABA Modulating Bacteria of the Human Gut Microbiota. *Nat. Microbiol.* **4**, 396–403 (2019).
278. Shimotoyodome, A., Meguro, S., Hase, T., Tokimitsu, I. & Sakata, T. Short chain fatty acids but not lactate or succinate stimulate mucus release in the rat colon. *Comp. Biochem. Physiol. A. Mol. Integr. Physiol.* **125**, 525–531 (2000).
279. Willemsen, L. E. M., Koetsier, M. A., van Deventer, S. J. H. & van Tol, E. A. F. Short chain fatty acids stimulate epithelial mucin 2 expression through differential effects on

prostaglandin E1 and E2 production by intestinal myofibroblasts. *Gut* **52**, 1442–1447 (2003).

280. Wrigley, D. M. Inhibition of *Clostridium perfringens* sporulation by *Bacteroides fragilis* and short-chain fatty acids. *Anaerobe* **10**, 295–300 (2004).
281. Bui, T. N. Y., Paul, A., Guleria, S., O’Sullivan, J. M. & Toldi, G. Short-chain fatty acids—a key link between the gut microbiome and T-lymphocytes in neonates? *Pediatr. Res.* 1–9 (2025) doi:10.1038/s41390-025-04075-0.
282. Flores-Lopez, R., Maasakkers, C. M., Lambers, T. T., Nauta, A. & Morán-Ramos, S. Cesarean section and socioeconomic status are associated with Bifidobacterium abundance and intestinal physiology: observations from a cross-sectional study in formula-fed infants. *BMC Microbiol.* **25**, 780 (2025).
283. Cannavò, L. *et al.* Oxidative Stress and Respiratory Diseases in Preterm Newborns. *Int. J. Mol. Sci.* **22**, 12504 (2021).
284. Liao, S.-L. *et al.* Caesarean Section is associated with reduced perinatal cytokine response, increased risk of bacterial colonization in the airway, and infantile wheezing. *Sci. Rep.* **7**, 9053 (2017).
285. Kiilerich, P. *et al.* Delivery Modality Affect Neonatal Levels of Inflammation, Stress, and Growth Factors. *Front. Pediatr.* **9**, 709765 (2021).
286. Thompson, A. L. Caesarean delivery, immune function, and inflammation in early life among Ecuadorian infants and young children. *J. Dev. Orig. Health Dis.* **10**, 555–562 (2019).
287. Busi, S. B. *et al.* Persistence of birth mode-dependent effects on gut microbiome composition, immune system stimulation and antimicrobial resistance during the first year of life. *ISME Commun.* **1**, 8 (2021).
288. Pasquali, F. *et al.* Citrobacter braakii Isolated from Salami and Soft Cheese: An Emerging Food Safety Hazard? *Foods* **14**, 1887 (2025).
289. Yu, M. *et al.* Characterization of cytotoxic Citrobacter braakii isolated from human stomach. *FEBS Open Bio* **14**, 487–497 (2024).
290. Lenický, M. *et al.* Identification of Bacterial Profiles and Their Interactions with Selected Quality, Oxidative, and Immunological Parameters of Turkey Semen. *Anim. Open Access J. MDPI* **11**, 1771 (2021).
291. Crost, E. H. *et al.* The mucin-degradation strategy of Ruminococcus gnavus: The importance of intramolecular trans-sialidases. *Gut Microbes* **7**, 302–312 (2016).
292. Bhattarai, Y. *et al.* Gut Microbiota-Produced Tryptamine Activates an Epithelial G-Protein-Coupled Receptor to Increase Colonic Secretion. *Cell Host Microbe* **23**, 775-785.e5 (2018).
293. Cheng, Y. *et al.* Aryl Hydrocarbon Receptor Activity of Tryptophan Metabolites in Young Adult Mouse Colonocytes. *Drug Metab. Dispos.* **43**, 1536–1543 (2015).
294. Zelante, T. *et al.* Tryptophan catabolites from microbiota engage aryl hydrocarbon receptor and balance mucosal reactivity via interleukin-22. *Immunity* **39**, 372–385 (2013).
295. Laursen, M. F. *et al.* Having older siblings is associated with gut microbiota development during early childhood. *BMC Microbiol.* **15**, 154 (2015).

296. De Bruyn, F., James, K., Cottenet, G., Dominick, M. & Katja, J. Combining Bifidobacterium longum subsp. infantis and human milk oligosaccharides synergistically increases short chain fatty acid production ex vivo. *Commun. Biol.* **7**, 943 (2024).
297. Litvak, Y., Byndloss, M. X. & Bäumler, A. J. Colonocyte metabolism shapes the gut microbiota. *Science* **362**, eaat9076 (2018).
298. Salvi, P. S. & Cowles, R. A. Butyrate and the Intestinal Epithelium: Modulation of Proliferation and Inflammation in Homeostasis and Disease. *Cells* **10**, 1775 (2021).
299. Rivière, A., Selak, M., Lantin, D. R., Leroy, F. & De Vuyst, L. Bifidobacteria and Butyrate-Producing Colon Bacteria: Importance and Strategies for Their Stimulation in the Human Gut. *Front. Microbiol.* **7**, (2016).

Appendix - Paper I: Key bacteria and environmental factors associated with the restoration of the infant gut microbiome after cesarean section

Jie Jiang[1,2], Casper Sahl Poulsen[1], Ulrika Boulund[1], Shiraz Shah[1], Urvish Trivedi[1,5], Madhumita Bhattacharyya[4], Avidan U. Neumann[4], Darlene L.Y. Dai[7], Charisse Petersen[7], Courtney Hoskinson[7,8], Theo J. Moraes[9], Piushkumar J. Mandhane[10,11], Elinor Simons[12], Meghan B. Azad[12,13], Padmaja Subbarao[9,14,15], Klaus Bønnelykke[1,3], Bo Chawes[1,3], Stuart E Turvey[7], Søren J. Sørensen*[5], Jonathan Thorsen*[1,3], Jakob Stokholm*[1,2,6]

[1]COPSAC, Copenhagen Prospective Studies on Asthma in Childhood, Copenhagen University Hospital - Herlev and Gentofte, Copenhagen, Denmark
[2]Department of Food Science, Faculty of Science, University of Copenhagen, Frederiksberg C, Denmark
[3]Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark
[4]Institute of Environmental Medicine and Integrative Health, Faculty of Medicine, University of Augsburg, Germany
[5]Department of Biology, Faculty of Science, University of Copenhagen, Copenhagen, Denmark
[6]Department of Pediatrics, Slagelse Hospital, Slagelse, Denmark
[7]Department of Pediatrics, BC Children's Hospital, University of British Columbia; Vancouver, BC, Canada.
[8]Department of Microbiology and Immunology, University of British Columbia; Vancouver, British Columbia, Canada.
[9]Department of Pediatrics, The Hospital for Sick Children, Toronto, Canada
[10]Department of Pediatrics, University of Alberta, Edmonton, Canada
[11]Department of Medicine, Faculty of Medicine and Health Sciences, UCSI University, Kuala Lumpur, Malaysia
[12]Section of Allergy and Immunology, Department of Pediatrics and Child Health, University of Manitoba, Winnipeg, MB, Canada
[13]Manitoba Interdisciplinary Lactation Centre (MILC), Children's Hospital Research Institute of Manitoba, Winnipeg, MB, Canada
[14]Department of Medicine, McMaster University, Hamilton, ON, Canada
[15]Dalla Lana School of Public Health, University of Toronto, Toronto, Canada

*Joint senior authors

Corresponding author:

Professor Jakob Stokholm, MD, PhD
E-mail: stokholm@copsac.com
Website: www.copsac.com

Abstract

Long-term gut microbiome perturbation following Cesarean section (CS) delivery has been associated with an increased risk of developing childhood asthma. Whether such CS-associated microbiome composition can be modulated by environmental exposures or ecological interactions, and thereby mitigate disease risk, is unclear. In the COPSAC$_{2010}$ birth cohort (N = 700), we longitudinally profiled the gut microbiome in infants at 1 week, 1 month, and 1 year of age using 16S rRNA gene sequencing. We developed a restoration score quantifying the degree to which the 1-year gut microbiome resembled that of vaginally delivered infants. Predictions of a restoration score were discovered in the very early 1-week gut microbiome: higher microbial diversity and higher relative abundances of specific taxa, including *Bifidobacterium longum*, *Bacteroides fragilis*, and *Bacteroides vulgatus*. In addition, having older siblings was linked to a higher restoration score, mediated by increased abundances of restoration-associated bacteria, suggesting a role for shared household microbiota in shaping early microbial trajectories. The restoration score, including association with delivery mode, older siblings and later asthma as well as early bacterial drivers, was successfully replicated in the independent Canadian birth cohort, CHILD. This validation strengthens the evidence that specific early-life microbial signatures and sibling exposure may support microbiome restoration and confer protective effects in CS-born infants. These insights highlight the potential targets for early interventions aimed at mitigating long-term health risks after CS delivery.

Keywords:
Cesarean Section(CS); gut microbiome; perturbation; maturation; restoration; siblings; COPSAC Cohort Study; CHILD Cohort Study

Introduction

A Cesarean section (CS) can be a life-saving intervention when medically indicated, but can potentially also lead to short-term and long-term health consequences for women and their children[1]. CS is one of the most important factors determining an infant's developing gut microbiome, and in turn, possibly the development and differentiation of their immune system[2]. Such perturbation of the initial infant gut colonization may be involved in the increased risk for diseases in childhood and later in life observed in CS-born individuals[3,4].

The healthy fetus is considered sterile[5] and the newborn child is colonized by the first microbes during and immediately after birth[6–8]. When comparing the gut microbiota composition, infants born by CS have gut microbiomes that resemble those found on the mother's skin surface, while vaginally born infants harbor gut microbiomes more similar to the adult fecal and vaginal microbiome[9,10]. In babies delivered by CS, previous studies have reported disrupted colonization of maternal *Bacteroides* strains, and more prevalent colonization by opportunistic pathogens associated with the hospital environment. A low-*Bacteroides* profile in the infant gut microbiome is considered a hallmark of the microbiome perturbation in CS-born infants[11,12]. However, low-*Bacteroides* profiles can also be found in vaginally delivered infants[13–15], which are usually used as a reference for a healthy infant gut microbiome. Such early microbiomes of vaginally delivered infants, which resembles the CS-perturbed profiles, could be caused by other perturbers such as intrapartum antibiotics[16,17].

Microbiota maturation can be described by age-dependent successional stages: a "mature" microbiota contains certain taxa that are common for that child's age group, while an "immature" or delayed microbiota resembles that of a younger child[18]. The microbiome composition of infants matures through three distinct, conserved stages of development, with the genus predominance shifting from *Escherichia* over *Bifidobacterium* to *Bacteroides*[19]. This ecological succession eventually stabilizes during childhood regardless of early perturbations such as birth by CS. However, such early perturbation may have long-lasting functional effects on immune development. Gut microbiota assembly and immune system development are intimately linked in early life[20,21]. The establishment of immune tolerance by introducing microbial antigens happens in the first few weeks of life[22,23]. Delayed maturation caused by early perturbation might result in a loss of tolerance and a pro-inflammatory response, which has been associated with increased risk of food allergy[24] and asthma[25].

The dynamics of the early life gut microbiota has been a key research area[26,27], the modifiability provides a potential to promote health throughout life[28]. Potential approaches include pre- and probiotics[29], vaginal seeding[30,31] and fecal microbiota transplantation (FMT)[32]. Maternal FMT was reported to have the most dramatic effect on microbiota composition, shifting the composition fully to that of the vaginally born infants[8]. However, a CS-perturbed gut microbiota may also naturally recover, a process which may be influenced by environmental and intrinsic ecological factors in the microbiome.

Several environmental factors have been reported to promote maturation of the gut microbiome, including breastfeeding[33], a rural living environment[34], and having older siblings at home[35].

Breastfeeding is recognized as one of the most influential drivers of gut microbiome composition during infancy, which provides a dynamic source of nutrition that delivers live microbes, immunoglobulins, and bioactive compounds essential for shaping the infant gut microbiome[36]. Besides, Infants from rural areas often exhibit a more diverse gut microbiome and early colonization of Bacteroides, which promotes further microbiome maturation[37,38]. Lastly, infants with older siblings tend to have a more mature gut microbiota by the age of one year, and this maturity is linked to a higher diversity of gut bacteria, which potentially mediates the protective effects of siblings in relation to allergies[35,39].

Previously, we described the link between CS and asthma risk which may be partially mediated via a prolonged perturbation of the infant gut microbiome in the 700 children from the Copenhagen Prospective Studies on Asthma in Childhood$_{2010}$ (COPSAC$_{2010}$) prospective birth cohort[17]. Only children who retained a CS-like gut microbiome composition at 1 year had increased risk of later asthma independent of the magnitude of their initial perturbation. Conversely, children whose gut microbiota was restored by 1 year, resembling those born vaginally, had a similar, lower, risk of asthma. In this follow-up study, we describe early-life environmental and ecological predictors of such restoration in the 1-year gut microbiome with the aim to enhance our understanding of the natural microbiome restoration process in the context of CS-associated risk of asthma.

Results

Characteristics of the COPSAC$_{2010}$ cohort

Information on maternal and child characteristics were obtained during scheduled visits to the COPSAC clinic. We observed differences in gestational age, parental age, maternal BMI, breastfeeding duration, and hospitalization after birth according to delivery mode (Table S1). We derived a 1-year restoration score based on previous work[17], which is found to be negatively associated with asthma risk at 6 years of age. This restoration score is to characterize the restoration of the 1-year gut microbiome - a higher score indicates that a child's 1-year gut microbiome resembles that of vaginally delivered infants (Fig.1A, Fig.1B). Compared to vaginal delivery, children born by CS had a significant lower restoration score (Estimate -0.32, 95% Confidence Interval (CI) [-0.51,-0.13], P=8e-04). In the subgroup analysis (Fig.1C), children born by CS had a lower restoration score compared to vaginal delivery without antibiotics exposure (-0.36[-0.55,-0.16], P=3e-04). Notably, children born by vaginal delivery whose mother received antibiotics at birth also had a lower restoration score compared to those without antibiotics exposure (-0.23[-0.48, 0.02], P=0.07), and with a score comparable to the CS-born infants (0.13[-0.19,0.45], P=0.43). There was no detectable differences observed between the planned and emergency procedures (0.18[-0.20,0.57], P=0.35).

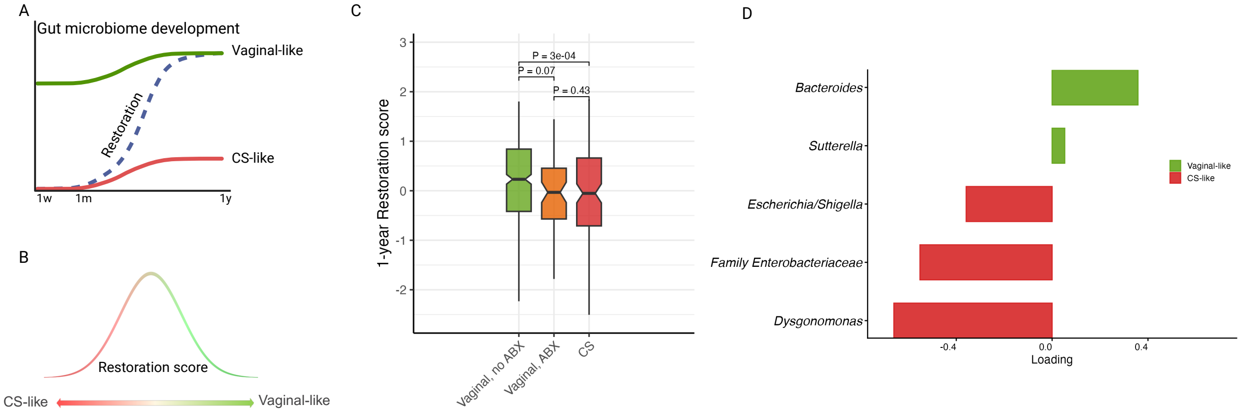


Fig.1. Conceptual design of a restoration score. (A) illustrates the hypothesis that a CS-perturbed microbiome composition can be influenced by environmental exposures or ecological interactions and be restored towards a vaginal-like microbiome composition. (B) shows that the restoration score in this study describes the status of 1-year gut microbiome, a higher restoration score means a child's microbiome resembles being born vaginally, while a low restoration score means that a child's microbiome resembles being born by CS[17] (See also Methods). (C) Boxplot of 1-year restoration score according to different groups, where red represents CS delivery, orange represents vaginal delivery with antibiotic treatment, and green vaginal delivery without antibiotic treatment, indicating that while CS born children have lower scores at 1 year of age, vaginally born children could possibly have a CS-like gut microbiome due to the perturbations like antibiotics treatment at birth. P values were derived from linear models comparing the restoration scores among these three groups. (D) Bacteria contributing to the 1-year restoration score. Loadings are derived from the sPLS model trained on the COPSAC$_{2010}$ 1-year gut microbiome in the previous work[17]. Negative/Positive loading corresponds to lower/higher genus abundance in vaginally delivered children compared to CS born children.

Gut microbiome diversity at 1 week associates with 1-year restoration score

Children with higher 1-year restoration scores had a higher gut microbial α-diversity at early time points (Table S2). At age 1 week, higher Shannon and Faith's Phylogenetic Diversity (PD) diversity were observed in children with higher 1-year restoration score (Shannon P=0.02, PD P=0.03). At 1 month of age, the same directionality was observed, but it was not significant for any of the indices. In the CS stratum (n=151), higher PD at 1 week of age was associated (P=0.04) with higher 1-year restoration score. For other measures, similar estimates were observed but they were not statistically significant (Table S2).

The gut microbiota composition differed between children with high and low 1-year restoration scores (above vs below median values) (Fig.2) using PERMANOVA analysis. From the PCoA plots ordinated across all three timepoints, children with low 1-year restoration scores appeared to have different compositions compared to those with high 1-year restoration scores. The compositional difference of having a high vs low restoration score was most obvious at age 1-week (F=4.9, R^2=0.9%, P=0.001) compared to age 1-month (F=2.2, R^2=0.4%, P=0.08).In the CS stratum, there were no detectable differences in microbial composition between high and low 1-year restoration score groups at 1 week and 1 month (Fig.S1).

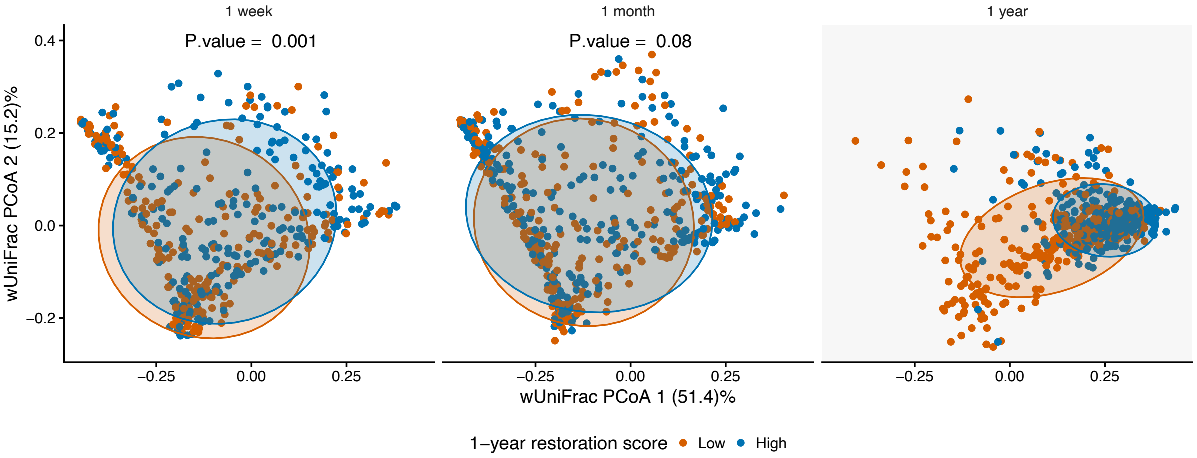


Fig.2.Comparison of the fecal microbiota between children having a low or high 1-year restoration score in the full cohort. Weighted UniFrac distances were used as input in principal coordinate analysis (PCoA) plots and colored according to children having high restoration scores (above median) (orange, n = 246, 285, 312 at 1 week, 1 month and 1 year) and low restoration scores (below median) (blue, n = 253, 269, 312 at 1 week, 1 month and 1 year). Gray shading at the 1-year time point marks when the restoration score was derived, hence groups are different by design. PERMANOVA was used to compare group differences. Each dot represents one fecal sample. Ellipses represent 1 standard deviation (1 SD), encompassing approximately 68% of the data points, assuming a bivariate normal distribution.

Early gut taxa associate with 1-year restoration score

We next investigated individual taxa at 1 week and 1 month to determine which members of the microbiota were main drivers of the observed compositional differences using the LIMMA model (Fig. 3). At 1 week of age, children with higher restoration scores had higher relative abundances of *Sutterella wadsworthensis* and *Neglecta timonensis* (Fig. 3A, FDR adjusted

P<0.05). On the other hand, we found that *Clostridium perfringens* at 1 week was negatively associated with the restoration score and was the most differentially abundant species found, followed by *Enterobacter aerogenes* and *Actinomyces sp.* (FDR adjusted *P<0.05*). Compared to the gut microbiome at 1 week of age, the 1-month gut microbiome was less associated with the 1-year restoration score. We observed associations between the relative abundances of species at 1 month and restoration scores in the same direction as at 1 week, but they were not FDR significant.

In the CS stratum, higher relative abundances of *Bifidobacterium longum* and *Staphylococcus simiae* at 1 week and only *Streptococcus equinus* at 1 month were associated with higher restoration score at 1 year (Fig.3B, *P*<0.05), but none were FDR significant.

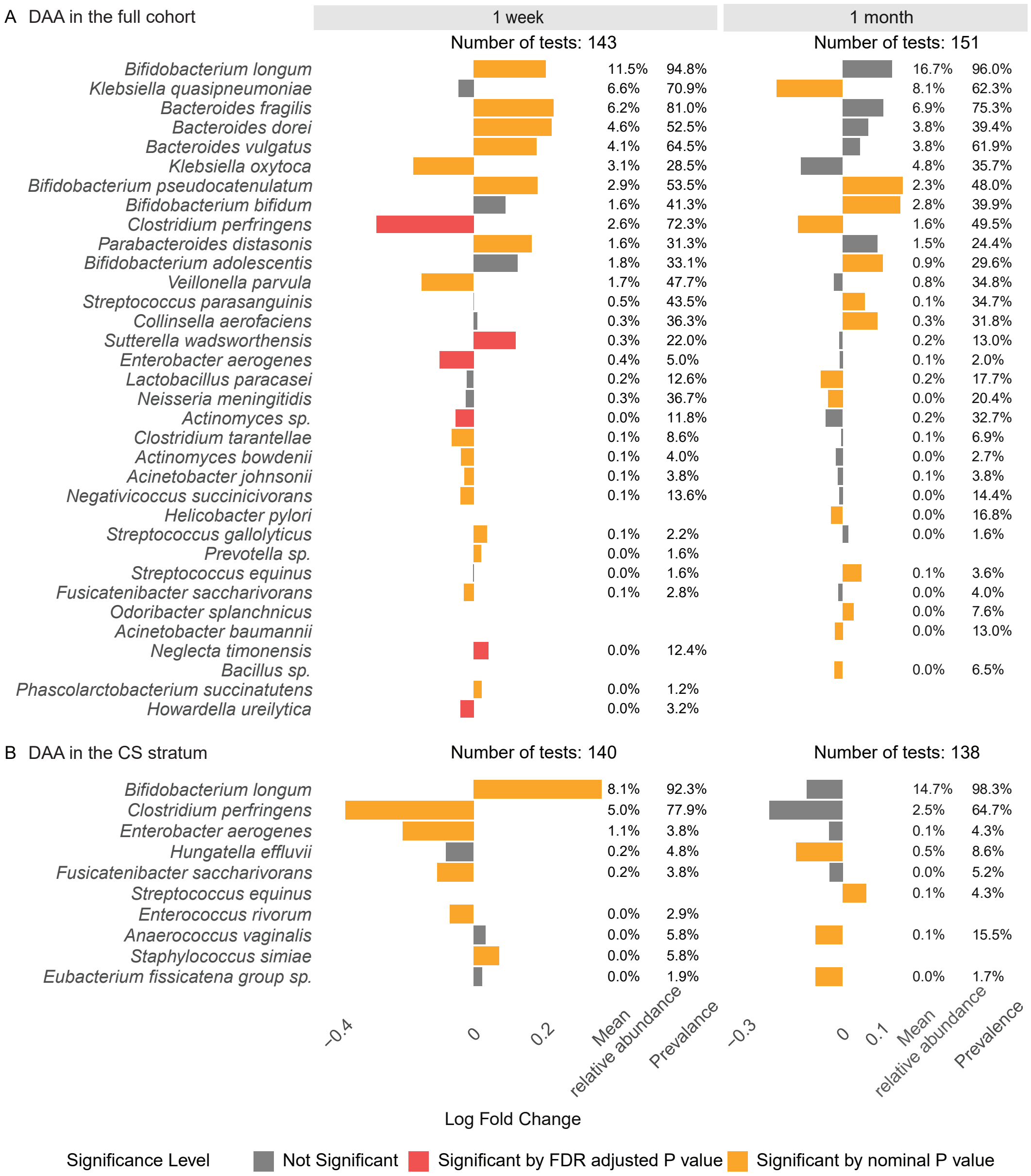


Fig.3 Differential abundance analysis on the species associated with 1-year restoration score. In the full cohort (A) and in the CS stratum (B). Species with prevalence of at least 0.1% and relative abundance of more than 0.01%

of the total were eligible. The analysis in the full cohort was adjusted for delivery mode. The species represented by the red bars were significant after FDR-adjustment, while species represented by yellow bars were nominally significant, and species represented by grey bars were not significant. A positive log fold change value indicates an increase in the abundance of the species as the 1-year restoration score increases, while a negative fold change indicates a decrease.

Environmental factors associate with 1-year restoration score

We next examined the relationship between early life environmental factors and the 1-year restoration score (Table S3). We found that having older siblings was associated with a higher restoration score in the full cohort (Full cohort, 0.36 [0.21;0.51], FDR adjusted *P*<0.01). Interestingly, the age of the youngest older sibling was negatively associated with the restoration score; i.e. the closer in age the older sibling was the higher the restoration score was at 1 year (Full cohort, -0.14 [-0.23;-0.05], FDR adjusted *P*=0.04).

To identify whether specific factors would influence the restoration process differently for children born by CS, we performed this subgroup analysis (Table S3). In the CS stratum, having older siblings, having cats, and being born in a rural area were positively associated with the restoration score while being breastfed for more than 6 months was negatively associated with the restoration score, but none were FDR significant.

Multivariable models on the early life gut microbiome and environmental factors predict the 1-year restoration score

Multivariable analyses were used to identify species and environmental factors jointly associated with the 1-year restoration score in the full cohort (Fig.4). We employed three sparse Partial Least Squares (sPLS) models on only the 1-week gut microbiome, only environmental factors, and combined 1-week gut microbiome and environmental factors as input features, respectively. These models were evaluated based on the correlations between their cross-validated predictions and 1-year restoration scores (model on only gut microbiome, Spearman rho 0.21, P<0.001; model on only environmental factors, 0.16, *P*<0.001; model on combined gut microbiome and environmental factors, 0.24, P<0.001). The gut microbiome at 1 week was enriched for *Bifidobacterium longum*, *Parabacteroides distasonis*, *Sutterella wadsworthensis*, and *Neglecta timonensis* in children with higher restoration scores at 1 year, while *Clostridium* perfringens was depleted. Of these, *Bifidobacterium longum* was the most abundant and prevalent (see Fig 3), while *Parabacteroides distasonis* and *Sutterella wadsworthensis* were less so. *Neglecta timonensis* was only present in few samples (prevalence 12.4%), but had a strong association with the 1-year restoration score. *Clostridium perfringens* contributed the most to the model with a strong negative loading towards the restoration score. In the model with environmental factors, only maternal antibiotics at birth and having older siblings had negative and positive loadings with the 1-year restoration score, respectively. The model on combined 1-week gut microbiome and environmental factors showed a better performance than each individual model, predicting the 1-year restoration score with 7 variables selected, including the same five species selected by the model on only the gut microbiome, and *Veillonella parvula* and having older siblings.

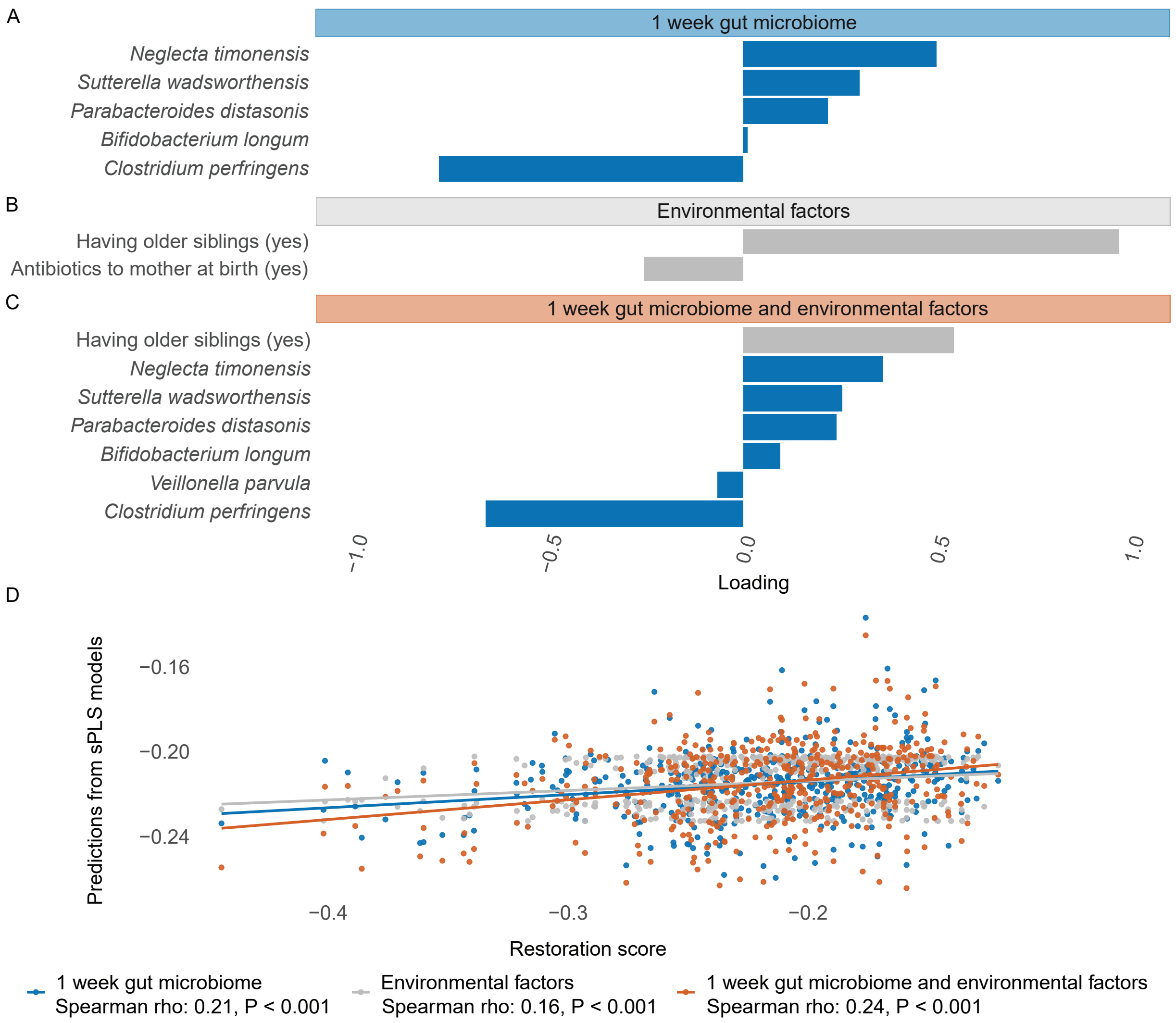


Fig.4 Sparse partial least squares (sPLS) models on restoration score using the full cohort at 1 week of age. (A), (B), and (C) show loadings from models on gut microbiome at 1 week of age (143 species, 499 samples), environmental factors (39 factors, 466 samples), and combined factors (182 variables, 466 samples), respectively. Loadings represent the contribution of each variable to the sPLS models. Negative loadings indicate associations with lower restoration scores, while positive loadings indicate associations with higher restoration scores.
(D) Spearman correlation between cross-validated predictions from the above three sPLS models and the 1-year restoration score. The correlation coefficient and P value are shown in the legend. Blue for the sPLS model on gut microbiome, grey for the sPLS model on environmental factors and orange for the sPLS model on combined gut microbiome and environmental factors.

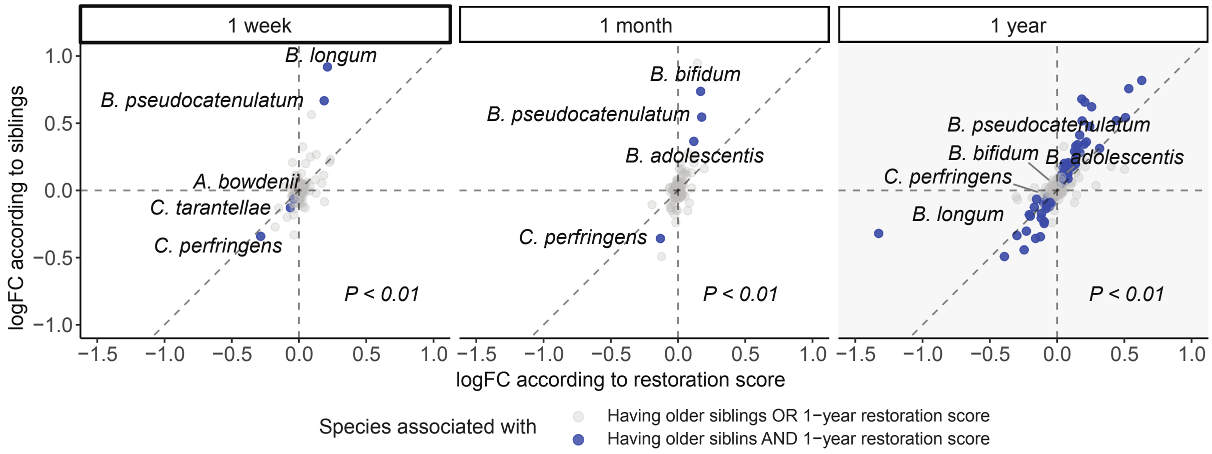
At 1 month of age, the models had moderate performance on predicting restoration scores (Fig.S2, model on gut microbiome, Spearman rho 0.10, *P*=0.023; model on environmental factors, 0.15, *P*<0.001; model on combined gut microbiome and environmental factors, 0.16, *P*<0.001). The combined model's cross-validated predictions had a higher correlation with 1-year restoration scores than each individual model, but the inclusion of the gut microbiome didn't improve the model much compared to the model on only environmental factors. In this combined model, 3 environmental factors and 14 species were selected (Fig.S2). Of all the selected variables, having older siblings and *Clostridium perfringens* were also selected in the 1-week combined model with the same directionality.

In the CS stratum, the comparatively best model was on the combined 1 week gut microbiome and environmental factors, the Spearman rho between the cross-validated predictions and the restorations scores was 0.19, *P*=0.064 (Fig.S3). However, the models on the 1 month gut microbiome and environmental factors in the CS stratum were not able to predict the restoration scores, with Spearman rho correlations around 0. The overall better performance of models on combined gut microbiome and environmental factors suggested that both might contribute to the restoration process of the perturbed gut microbiome.

Having older siblings facilitates CS restoration through gut microbiome development

We further investigated whether the species associated with having siblings were also associated with the 1-year restoration score, by comparing the logFC of the species abundances between the two outcomes(Fig.3, Fig.5, Table S4).

A Species associated with having older siblings and 1-year restoration score



B Number of species associated with having older siblins and 1-year restoration score

Fig.5 Differential abundance analysis reveals subset of species associated with having older siblings and restoration in the full cohort. (A) Scatter plot comparing differential abundance analyses of having older siblings and the restoration score. Each point represents a species, those associated with both factors are highlighted in blue. Labeled species were significant at 1 week and 1 month. A positive log fold change value in the first quadrant (upper

right) indicates an increase in the abundance of certain bacteria associated with a higher 1-year restoration score and having older siblings at home, while a negative fold change in the third quadrant (bottom left) indicates a decrease in the abundance of certain bacteria associated with a lower 1-year restoration score and not having older siblings. P values were derived from permutation tests examining whether the direction of the associations with both predictors is concordant by chance, a P value smaller than 0.05 indicates the concordant direction is unlikely under the null hypothesis. (B)Venn diagram indicating the number of differentially abundant species according to having older siblings and according to the 1-year restoration score. P values were derived from permutation tests examining whether the number of taxa associated with both predictors is greater than expected by chance, a P value smaller than 0.05 indicates the overlap is unlikely under the null hypothesis.

At 1 week, 85 species showed a similar logFC association for both having older siblings and the 1-year restoration score (Fig.5A, 85 points, grey and blue, were in the first and third quadrants at 1 week). Among them, 5 were significantly associated with both outcomes (Fig.5B). There was a higher abundance of *Bifidobacterium longum* and *Bifidobacterium pseudocatenulatum* in the children with siblings and a high restoration score; and lower abundance of *Actinomyces bowdenii*, *Clostridium tarantellae*, and *Clostridium perfringens* in the children without siblings and with a low restoration score. At 1 month of age, 4 species were significantly associated with both outcomes in the same direction. At 1 year of age, the time point where the restoration score was defined, 63 species showed significant consistent positive or negative association with both outcomes. To determine whether the number of taxa that were significant for both predictors was greater than expected by chance, and whether the direction of their associations was concordant by chance, we calculated empirical p-values based on permutation tests. Both of the empirical p-values were below 0.05, indicating that the overlap and concordant direction are highly unlikely by chance.

In the CS stratum, 87 species, and 73 species showed consistent associations with having older siblings and the restoration score, at 1 week and 1 month time point, respectively, but less significant compared to the associations in the full cohort (Fig.S4, Table S5). Nevertheless, higher abundance of *Bifidobacterium longum* was still found positively associated with having older siblings and higher restoration score. We also found the same directionality at 1 year of age in the CS stratum, which further suggests a positive role of having older siblings on the restoration of the gut microbiome by such very early influences. The number of taxa associated with both predictors and concordant direction in the CS stratum was significant in the permutation test.

Having identified older siblings as a potential protective factor that may facilitate the restoration of a CS-perturbed gut microbiome, we next considered these factors together: We performed a mediation analysis to investigate the possible mediating role of the 1-week gut microbiome between having older siblings and the restoration of the gut microbiome by 1 year. Here, we used the cross-validated predictions from the sPLS model of the 1-week gut microbiome to predict the 1 year restoration score and expressed it as a 1-week microbial score representing the degree to which a child’s 1-week microbiome looked like it would be restored by 1 year of age (Fig.S5). After adjusting for delivery, the results showed a significant mediation effect of older siblings on the 1-year restoration score through the 1-week microbial score (*P*=0.004), but also a strong direct effect of older siblings on the 1-year restoration score (*P*<0.001). The indirect pathway accounted for a portion of 11.2% of the total effect. This suggests that older siblings may contribute to the restoration partially through influencing the very early gut microbiome.

Early gut taxa and environmental factors associated with restoration scores in the vaginal stratum

While the restoration score was conceived in the context of CS delivery, we also wanted to investigate its dynamics in vaginally delivered children, where low restoration scores would indicate that other factors than CS may have contributed to having a CS-like composition at 1 year. We separately investigated the early gut microbiome associated with the 1-year restoration scores in the vaginal stratum (Fig. S6), and perinatal events during vaginal delivery in addition to the previously described prenatal and postnatal factors (Table S8). At 1 week of age, children with higher 1-year restoration scores had higher relative abundances of *Sutterella wadsworthensis* and *Neglecta timonensis* as we saw in the full cohort (FDR adjusted $P<0.05$). In contrast, *Veillonella parvula*, *Actinomyces sp.*, and *Howardella ureilytica* were associated with lower restoration scores at 1 year of age in the vaginal stratum (FDR adjusted $P<0.05$). Having older siblings was significantly associated with higher 1-year restoration scores as we saw in the CS stratum (0.02[0.01,0.03], FDR adjusted $P<0.05$). In contrast, we found antibiotics exposure to mothers and children at birth and meconium-stained amniotic fluid exposure to be negatively associated with the 1-year restoration score, however not FDR significant.

Validation of the 1-year restoration score on the CHILD cohort

We next sought to validate our findings in an independent dataset. The CHILD study is the largest prospective longitudinal birth cohort study in Canada[40]. We observed similar differences in gestational age, parental age, maternal BMI, and hospitalization after birth according to delivery mode in the CHILD study. On the other hand, compared to the COPSAC cohort, CS-born children had higher antibiotics exposure at 3 months and shorter breastfeeding duration in the CHILD study (Table S1).

We applied the sPLS model trained on the COPSAC$_{2010}$ cohort[17] on the CHILD cohort's 1-year samples (n=325) to create a restoration score (Fig. 6A). This 1-year restoration score was, like in COPSAC$_{2010,}$ negatively associated with delivery by CS and positively associated with older siblings in the CHILD cohort (Fig. 6B). Additionally, the CHILD 1-year restoration score was associated with reduced asthma risk at 5 years in both univariate and adjusted logistic regression models (OR 0.63[0.44, 0.87], *P*=0.005, aOR 0.47 [0.26, 0.82], *P*=0.0084, Fig. 6B). This replicates the results in our previous study[17] and is consistent with the hypothesis that an appropriate restoration of the gut microbiota could mitigate the increased asthma risk associated with gut microbial changes due to CS delivery.

To identify species at early time points associated with the 1-year restoration score, we trained an sPLS model on the gut microbiome at 3 months in the CHILD cohort, to predict the 1-year restoration score(Fig. 6C). *Clostridium perfringens* was selected in the model with the highest negative loading as we saw in the COPSAC$_{2010}$ cohort, while *Bacteroides sp.* had the highest positive loading in the model, followed by *Bacteroides uniforms*, *Bacteroides caccae*, *Bifidobacterium sp.*, *Bacteroides fragilis*, and *Bacteroides ovatus*. Having older siblings was positively associated with the 1-year restoration score, which supports the idea that older siblings promote the restoration of a CS-perturbed gut microbiome(Fig. 6B).

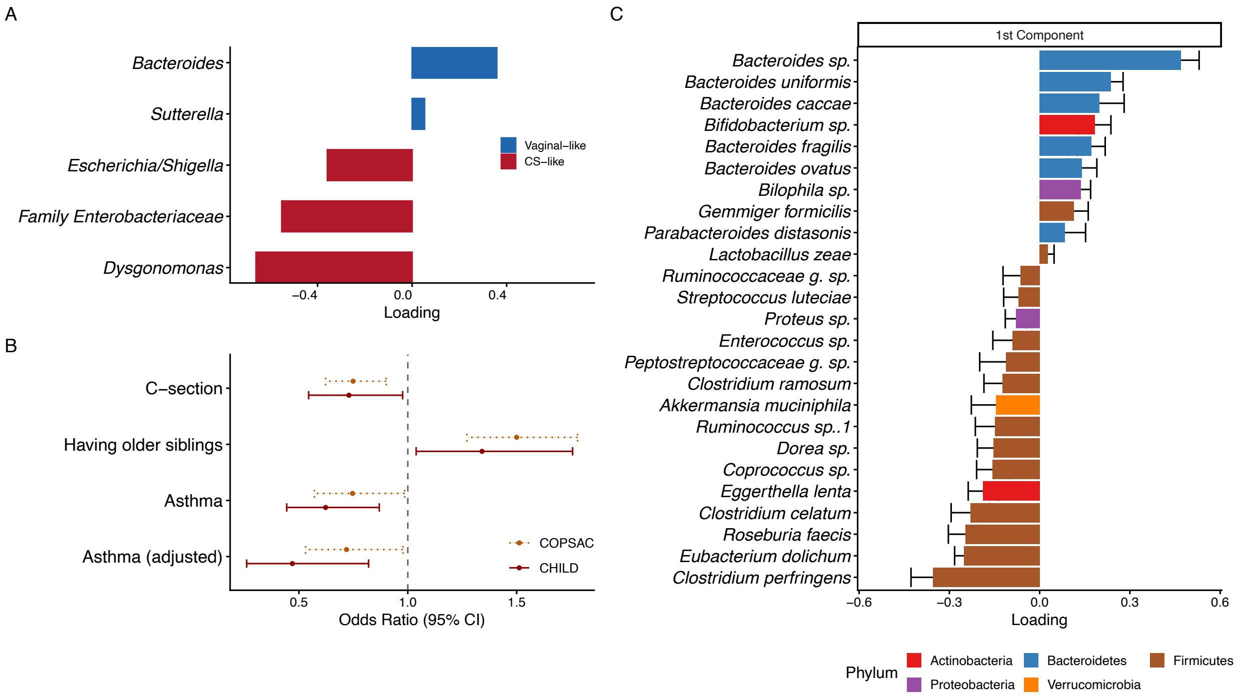


Fig.6 Validation of the COPSAC$_{2010}$ 1-year restoration score in the independent CHILD cohort. (A) Loadings are derived from the sPLS model trained on the COPSAC$_{2010}$ 1-year gut microbiome to output a restoration score, the annotations are derived from the CHILD cohort 16s rRNA data. (B) Univariate logistic regression models on the 1-year restoration score versus CS, having older siblings and asthma, and the adjusted odds ratio from multivariable logistic regression model on asthma, in both cohorts (dashed line: COPSAC, solid line: CHILD). The asthma diagnosis was at 5 years in the CHILD study, while the asthma diagnosis was at 6 years of age in the COPSAC cohort. The covariates adjusted in the models were: gestational age, hospitalization after birth, antibiotics exposure to children at 1 year, having older siblings, family asthma history, gender, race, birth season, (and study center for the CHILD study). (C) sPLS model on gut microbiome at 3-month in the CHILD cohort predicting 1-year restoration score, positive loadings indicated positive association with higher restoration score, while negative loadings indicated negative associations with higher restoration score.

Discussion

In this study, we investigated the early-life gut microbiome and environmental factors in relation to restoration of the gut microbiome in the first year of life in a prospective birth cohort. A successful microbiome restoration to a vaginal-born-like composition during this period seems to alleviate the otherwise increased risk of asthma in CS-born children. We found that a higher abundance of several bacteria including members from *Bacteroides* and *Bifidobacterium* was associated with increased restoration of the gut microbiome after CS by 1 year of age. Moreover, having older siblings improved restoration in part mediated via the 1-week gut microbiota. The restoration score, including association with delivery mode, older siblings and later asthma as well as early bacterial drivers, was successfully replicated in the independent CHILD cohort.

During infancy the diet, host genetics, environment and medical interventions determine the establishment and progression of the intestinal microbiota, which interacts with and trains the developing immune system[41]. We found that *Bacteroides* and *Bifidobacterium* species were among the microbiome features at 1 week of age that were associated with a higher restoration score at 1 year. Their abundance drives delivery mode-dependent infant gut microbiota developmental trajectories[15]. By occupying ecological niches within the gut, *Bacteroides* can protect against pathogenic bacteria[42], in part via the production of a unique capsular polysaccharide known as polysaccharide A[43–45].

In the first year of life *Bifidobacterium* has been linked with lower diversity, but its dominance helps establish a protective gut environment[46,47]. However, the timing of having a high *Bifidobacterium* abundance seems to matter. While we observed positive associations with higher abundance early, we also found a negative association between the 1-year *Bifidobacterium longum* abundance and the restoration score (Table S5, Table S6). Some studies indicate that *Bifidobacterium longum* subsp. *infantis* may colonize the gut later in the breastfeeding period, evidenced by a detectable abundance starting only at 10 weeks of age[48] or later[49]. In addition, the timing of infant colonization with *Bifidobacterium longum* subsp. *infantis* is consistent with horizontal transmission of this subspecies, rather than the vertical transmission previously reported for other *Bifidobacterium* species[50]. In the CS stratum, we observed a negative association between breastfeeding for longer than 6 months with the restoration score, suggesting that breastfeeding until 6 months may decelerate restoration, perhaps through a persistent high *Bifidobacterium* colonization. In a Swedish cohort, *Bifidobacterium longum* subsp. *infantis* reached its peak prevalence at 17 weeks followed by a gradual decline in both prevalence and relative abundance by 52 weeks[48,51]. This dominance of *Bifidobacterium* competes against *Bacteroides* and other bacteria contributing to the restoration, which could be further associated with a persisting CS-like gut microbiome composition.

We also found other species associated with restoration score, but there is limited research relating them to the restoration. *Neglecta timonensis* is phylogenetically close to *Clostridium sporosphaeroides* and *Clostridium leptum*[52], which are producers of short-chain fatty acids such as butyrate, acetate and propionate from fermentation of fibers, which might play an immunomodulatory role in disease such as asthma[53,54]. *Sutterella wadsworthensis* has been isolated from a variety of clinical specimens throughout the human gastrointestinal tract, in both healthy individuals and patients with gastrointestinal diseases like inflammatory bowel disease (IBD)[55,56]. Few studies describe the role of *Sutterella wadsworthensis*. However, an FMT study showed an enrichment of *Sutterella wadsworthensis* and other species and increased levels of heme and lipopolysaccharide biosynthesis in IBD patients who did not achieve remission after FMT[57]. *Parabacteroides distasonis* has been studied for its potential probiotic properties, with studies suggesting it can produce pentadecnoic acid[58–60].

Conversely, the relative abundance of *Clostridium perfringens* was at all time points negatively associated with the restoration score at 1 year of age in both the full cohort and in the CS stratum, suggesting a negative role in the restoration from a perturbed gut microbiome composition. *Clostridium perfringens* is a common colonizer of the infant gut microbiome, and can colonize the infant very early[61]. While being part of the normal gut microbiota, overgrowth of specific strains can disrupt the microbial diversity. In particular, strains of *Clostridium perfringens* harboring a gene encoding the toxin perfringolysin O (*pfoA)* have been associated with necrotizing enterocolitis (NEC) in preterm infants when present in high abundance[62]. It is reported that CS born infants exhibited higher carriage of toxigenic *Clostridium perfringens,* and infants carrying toxigenic *Clostridium perfringens* had lower levels of *Bacteroides*, *Bifidobacterium*, and *Lactobacillus* groups[63].

We also tested the association between different environmental factors and the 1-year restoration score. Having older siblings has been associated with increased gut microbial diversity during early childhood[39,64]. We have previously reported that having a sibling captures microbial signatures that associate with protection against asthma at 6 years of age, and additionally that presence of siblings promotes a more mature gut microbiome composition at 1 year, which in turn also associates with protection from asthma[65] and allergy[35]. This suggests a long-lasting effect of having older siblings on the gut microbiome and disease risk. In line with a previous study in COPSAC[17], we found a positive association between older siblings and the restoration score, and an association with key gut microbial taxa who appeared as partial mediators of this association, including the enriched *Bifidobacterium longum*, *Bifidobacterium pseudocatenulatum* and *Bifidobacterium adolescentis* and the depleted *Actinomyces bowdenii*, *Clostridium tarantellae*, and *Clostridium perfringens*. Siblings share many strains, which might be due to direct contact between the siblings or due to parallel colonization from a common source, such as the mother or the environment[66]. The transfer of beneficial microorganisms from older siblings could help in the development of the infant's immune system, providing protection against allergic and autoimmune disorders[67]. Furthermore, we found that this protective effect started as early as 1 week of age, and lasted at least till 1 year of age.

In previous studies, breastfeeding contributed to shaping a specific microbial composition in early life, which had the potential to influence infant immune development[33]. We therefore hypothesized that breastfeeding could also promote the restoration of a CS-perturbed gut microbiome. However, we did not observe this association, probably due to the imbalanced groups as the vast majority of children were breastfed at 1 week (exclusive breastfeeding 90%, N=625) and 1 month (exclusive breastfeeding 79%, N=548) of age, resulting in a small overall variation. Lastly, being born in a rural area and having cats at home were found to be positively associated with the 1-year restoration score only in the CS stratum, which suggests a positive role in restoring a CS-perturbed gut microbiome.

We were able to perform an external validation of our restoration score in an independent cohort, which greatly strengthens the interpretability and robustness of this metric. Here, we applied the previously developed model - trained solely on COPSAC_{2010} 1-year gut microbiome[17] - to the CHILD 1-year gut microbiome. Without any retraining or recalibration, the restoration score produced in the CHILD dataset not only recapitulated the association with delivery mode seen in COPSAC_{2010} cohort, but also demonstrated biological relevance by showing significant association with lower asthma risk at 5 years.

The predictive power of the restoration score in this separate cohort indicates that it can distinguish microbiome development patterns in a biologically consistent way, independent of cohort-specific features and geographical context. This supports our interpretation that the score represents a continuum of microbiota restoration - from perturbed (CS-like) to normalized (vaginal-like). It also affirms that the score captures biologically meaningful variation in vaginally born infants, which we interpret not as a flaw, but as an opportunity to infer the variation in CS born infants, since we included those born by vaginal delivery and received antibiotics at birth in the previous work, and this perturbation could cause a “CS-like” microbiome in vaginally born infants[17]. In the current study, we observed meconium-stained amniotic fluid exposure to be negatively associated with restoration scores in the vaginal stratum.

By showing that infants with higher restoration scores exhibit early microbiome profiles and taxa compositions more typical of vaginally born infants, we believe that these taxa (e.g. *Bacteroides fragilis* and *Bifidobacterium longum*) could contribute to the restoration of a perturbed gut microbiome. On the other hand, we found early *Clostridium perfringens* negatively associated with restoration in both cohorts, which could add on the knowledge of the development of a perturbed gut microbiome. Last but not least, we also validated the positive association between having older siblings and restoration score, which is also consistent with the hygiene hypothesis[68].

There are some limitations to our study. Due to the limitation of 16S rRNA sequencing, our analysis was not able to delineate the composition of subspecies in *Bifidobacterium longum*, which would require metagenomic data or strain isolation. In the beta diversity analyses of the early time points in relation to later restoration score, the variance explained was modest, which is a common observation in microbiome studies due to the high dimensionality and dynamic nature of the infant gut microbiome. This likely reflects the rapid ecological changes occurring during the first year of life, where many taxa are competing in this developing niche. Given the observational design of the study, we only inferred associations between gut microbiome at early and late timepoints and environmental factors. Many of our single ASV results were not significant after FDR correction, which is important for interpretation. We chose to report them here so that others may attempt replication. Furthermore, our analysis approach included multivariable analyses that do not require FDR correction but lack the inference on individual taxa.

Conclusion

The early-life gut microbiome and environmental factors are important for the restoration of a CS-perturbed gut microbiome. In the future, this can be key to further reduce long-term health risks associated with a perturbed gut microbiome such as childhood asthma. Beneficial bacteria like *Bifidobacterium longum* appear to play an important role in restoration, and having older siblings is associated with the restoration process. These findings underscore the potential for early interventions to improve health outcomes in CS-born infants.

Methods

The COPSAC$_{2010}$ cohort

Ethics

The study was conducted in accordance with the guiding principles of the Declaration of Helsinki and was approved by the Local Ethics Committee (H-B-2008-093) and the Danish Data Protection Agency (2015-41-3696). Both parents gave oral and written informed consent before enrolment.

Study population

The COPSAC$_{2010}$ cohort is a population-based mother-child cohort of 700 children and their families, recruited in pregnancy and followed prospectively at the COPSAC research unit. The children were followed by COPSAC study physicians and nurses collecting all biosamples, clinical measurements and diagnoses during clinical visits scheduled at 1 week, 1, 3, 6, 12, 18, 24, 30, and 36 months, thereafter yearly until the age of 6 and again at age 8 and 10 years.

Study endpoints

This study is a follow-up of our prior study[17], investigating the gut microbiome development and asthma risk after birth by CS in the COPSAC$_{2010}$ mother-child cohort.
In the present study, we focused on the association between early environmental factors and fecal microbiota composition at 1 week and 1 month, using a 1-year restoration score as the outcome measure. This restoration score was first introduced in our previous paper as a CS score, calculated by constructing a cross-validated sparse partial least squares (sPLS) model on gut microbial composition at 1 year of age predicting delivery mode (vaginal/CS)[17]. The model identified the gut microbial composition at 1 year of age most associated with CS delivery. By reversing the CS microbial score, we define this restoration score as our outcome measure to characterize the restoration of 1-year gut microbiome. Thus, a higher restoration score means a child's microbiome resembles being born vaginally, while a low restoration score means that a child retains a CS-like gut microbiota composition. This score was used as either a continuous score or dichotomized (above and below median value).

Fecal sample collection and sequencing

All children with fecal samples collected and characterized by 16S rRNA sequencing (requiring at least 2000 reads) at any of the three time points of 1 week (n = 552), 1 month (n = 607), and 1 year (n = 625) were included in the analyses. Fecal samples were collected either at the research clinic or by the parents at home using detailed instructions. Each sample arrived at the laboratory within 24 hours and was mixed on arrival with 1 ml of 10% (v/v) glycerol broth (SSI, Copenhagen, Denmark) and frozen at −80°C. DNA was extracted using the PowerMag Soil DNA Isolation Kit (MO-BIO Laboratories, Inc., Carlsberg, CA, USA) on an epMotion 5075 (Eppendorf), amplified using a two-step polymerase chain reaction (PCR) with 515F and 806R primers flanking the V4 region of 16S rRNA gene, and sequenced using the v2 kit (paired-end

250–base pair reads) on the MiSeq platform (Illumina Inc., San Diego, CA). A full description of the laboratory workflow has been described previously[65].

Bioinformatic processing

Raw fastq files were demultiplexed using the MiSeq controller software prior to downstream analysis. As described in previous study[69], the primers and adaptors in sequencing reads were removed using Cutadapt[70]. The determination of amplicon sequence variants (ASVs) was performed on QIIME2 Core 2020.11 platform[71] using Amplicon Denoising Algorithm 2 (DADA2) analysis pipelines[72]. The resulting ASV sequences were annotated using the AnnotIEM[73] pipeline (v.1.3), which combines sequence alignment against four databases: EzBioCloud[74] (r. 2018-05), NCBI[75] (v. refseq 202), RDP[76] (v.11.5), and Silva[77] (v. 138SSU) followed by a high confidence selection of best probable annotation[73]. Genus annotations were correctly annotated for all genera in the mock community, and 15/20 was annotated to the correct species. Therefore, species annotations should be considered putative.

Maternal and child characteristics

Information on antibiotics in pregnancy, antibiotics to mother and child at birth, antibiotics to children during the first year of life, gestational age, age when starting daycare, maternal age, maternal BMI, having older siblings at home, age of the youngest sibling, physician-diagnosed asthma, rhinitis, and dermatitis in the mother and father, paternal age, duration of exclusive and total breastfeeding period, hospitalization after birth, any furred animal at home, cats at home, dogs at home, and birth area was obtained during the scheduled visits to the research clinic. Delivery mode was encoded as “vaginal” and “CS” (including planned and emergency CS). Information on intrapartum antibiotics was validated against birth records from hospitals and information on antibiotic use during pregnancy and childhood was validated against national registries[78]. Information on perinatal exposures (asphyxia including meconium-stained amniotic fluid, vacuum-assisted delivery, and prelabor membrane-rupture) was obtained from national registries[78]. Living environment was based on home address at birth using the CORINE satellite-based land cover database[79].

The CHILD cohort

Ethics

Ethical approval for the CHILD Cohort Study, including the oversight of the CHILD biological samples and the CHILD database (CHILDdb), was obtained from the local Research Ethics Board of each study site: the University of British Columbia, the University of Alberta, the University of Manitoba, the Hospital for Sick Children and McMaster University.

Study population

CHILD study is a prospective longitudinal birth cohort study, which enrolled 3,405 subjects since pregnancy from 4 largely urban study centers across Canada (Vancouver, Edmonton, Winnipeg, and Toronto) from 2008 to 2012[40]. All children with fecal samples collected and characterized by 16s rRNA sequencing at around 1 year visit (9 months to 12 months, n = 325) were included in the analyses (Table S7). Questionnaires related to environmental exposures, psychosocial stresses, nutrition and general health were administered at recruitment, prenatally, at 3, 6, 12, 18, 24, 30 months, and at 3, 4, and 5 years. Covariates used in this study were: delivery mode, gestational age, hospitalization after birth, gender, antibiotics exposure to children at 1 year, family asthma history, race (mother and father), birth season and study center.

Study endpoints

Childhood asthma was diagnosed (as Yes/Possible/No) by an expert study physician at the clinical assessment at the age of 5 years based on published approach[40]. For this study, children were considered to have asthma only if the response was ‘Yes’ and the asthma phenotype was defined as comparing children with asthma at 5 years versus children without asthma at 5 years, children diagnosed as “possible” were excluded.

Fecal sample collection and sequencing

Sequencing data generation for infant stool microbiota has been previously described[80]. Briefly, a soiled diaper was provided on the same day for infant stool collection. Samples were refrigerated at home for up to 24 hours before being collected and processed by study staff. An additional infant stool sample was provided at the 1-year clinical assessment. DNA was extracted from fecal samples using the commercial kits (Qiagen Mo Bio PowerSoil) optimized for the Thermofisher KingFisher® robot. The V4 hypervariable region of the 16S rRNA gene of fecal DNA was amplified by PCR using universal bacterial primers (V4-515f: V4-806r).

Bioinformatic processing

Pooled PCR amplicons were subjected to paired-end sequencing on the Illumina MiSeq platform. Using VSEARCH and Deblur[81] within the QIIME2 pipeline[71], forward and reverse demultiplexed reads were assembled for a final length of 247 bp (unassembled sequences were discarded) and filtered against the GREENGENES reference database ((v13·8)[82,83]. Taxonomic classification of the resulting unique amplicon sequence variants (ASVs) was achieved using a naïve Bayes classifier trained on reference reads extracted from the reference database at 97% sequence similarity.

Statistics and data analysis

Chao 1 index, Shannon diversity and Faith’s Phylogenetic Diversity (PD) were used as measures of the within-sample diversity (α-diversity). Chao 1 index and Shannon diversity were calculated using estimate_richness() from package phyloseq v1.48.0, PD was calculated using pd() from package picante v1.8.2. Linear regression was used for analyzing simple associations between α-diversity and the 1-year restoration score. A p-value of 0.05 was considered statistically significant. The between-sample diversity metrics (β-diversity) were computed as weighted UniFrac distances[84]. Differences in β-diversity were visualized with principal coordinates analysis (PCoA) plots, where the restoration score was dichotomized into high and low groups (above/below the median), and tested for inference using permutational multivariable analysis of variance (PERMANOVA; adonis2 from the package vegan v2.6-8 with 999 permutations). A p-value of 0.05 was considered statistically significant.

At species level, we filtered species based on the threshold of mean relative abundance > 0.01%. Differential abundance analysis against 1-year restoration score at species level was performed using LIMMA[85] (Linear Models for Microarray Data, DA.lli2 from DAtest[86] package), which fits a linear model to compositional data (log-transformed) for each taxa, multiple testing was controlled using the False Discovery Rate (FDR) adjustment[87] within the function. The associations between the 1-year restoration score and environmental factors were tested using linear regression, multiple testing was controlled using the False Discovery Rate (FDR) adjustment. We performed supervised sparse partial least squares (sPLS) regression models on fecal microbiome and environmental factors separately and jointly using package mixOmics v6.28.0. This multivariable model is designed to reduce the high dimensionality of data and perform simultaneous variable selection. By using sPLS, we were to reveal microbial taxa and environmental factors most descriptive of 1-year restoration score. We $\log_{10}$-transformed counts data, using the lowest non-zero value as a pseudocount, and then calibrated the data for sequence depth by adjusting the number of reads and the log-transformed number of reads in the linear model. We selected the optimum number of input variables using 11-repeated 10-fold cross-validation of the correlation statistic to avoid overfitting and chose the median repeat for stability. The performance of the model was evaluated by assessing the correlation between predictions from the model and true restoration score. To evaluate the mediation role of the 1-week gut microbiome in the association between having older siblings and the 1-year restoration score, we applied mediation analysis using R package mediation (v4.5.1). A p-value of 0.05 was considered statistically significant. We performed a permutation test to determine whether the number of taxa that are significant for both the 1-year restoration score and older siblings' status was greater than expected by chance, and whether the direction of their associations was concordant by chance. Using 10,000 permutations, we randomly shuffled the two predictor vectors in the data at each time point, recalculated differential abundance for each predictor (adjusting for reads and delivery mode), and merged the results by species. For each permutation, we counted how many species were significant for both predictors and computed the Spearman correlation between their log-fold-changes. The empirical p-values were calculated as P_perm_rho = (sum(rho_perm >= rho_obs) +1) / (10,000 + 1) and P_perm_n = (sum(n_perm >= n_obs) +1) / (10,000 + 1) for the direction and the number of associations, respectively. Analyses were conducted in the full cohort and within the stratum of children born by CS throughout and for vaginal born children for specific analyses. All analyses were conducted in R v4.4.0.

The validation study was conducted by applying the sPLS model from our previous study[17] to the CHILD 1-year gut microbiome. In the previous work, the sPLS model was trained on 1-year gut microbiome in the COPSAC_{2010} cohort and selected 5 genera to predict if the child was born by CS or not (1/0). Those taxa were: *Bacteroides*, Family *Enterobacteriaceae*, *Sutterella*, *Escherichia/Shigella*, and *Dysgonomonas*. We found the same or the closest annotation in the CHILD cohort as the input of the model: *Bacteroides*, Family *Enterobacteriaceae*, *Sutterella*, *Escherichia/Shigella*, and *Dysgonomonas*. The preprocessing of the abundance was the same as described in the original work, where we agglomerated the rank into genus level, log transformed the relative abundance, using half the lowest nonzero value as a pseudocount. We then used the predict() function from mixOmics package to apply the model on these five genera, and output the *CHILD 1-year restoration score*. To test the association between the CHILD 1-year restoration score and asthma endpoint, we performed both univariate logistic regression and multivariable logistic regression. The covariates adjusted in the multivariable logistic regression models were: gestational age, hospitalization after birth (days), gender, antibiotics exposure to child at 1 year, maternal asthma history, paternal asthma history, maternal race, paternal race, birth season, and study center. The performance of the model was evaluated by associating the CHILD 1-year restoration score with the outcome of the model - delivery mode - using the univariate logistic regression. And the validation of the association between having older siblings and restoration score was tested using univariate logistic regression as well.

Data availability

All data obtained from the COPSAC participants associated with this study are present in the paper or the Supplementary Materials. The 16S rRNA gene sequencing data are deposited at the Sequence Read Archive with the accession no. PRJNA417357. Participant-level personally identifiable data are protected under the Danish Data Protection Act and European Regulation 2016/679 of the European Parliament and of the Council (GDPR) that prohibit distribution even in pseudo-anonymized form. However, participant-level data can be made available under a data transfer agreement as part of a collaboration.

The informed consent obtained from the CHILD participants, in addition to the CHILD Inter-Institutional Agreement (IIA) which has been executed between the five Canadian institutions responsible for the study, govern the sharing of CHILD data. The accession numbers for the 16S rRNA gene sequence data reported in this paper are BioProject accession (NCBI): PRJNA481046. Data described in the manuscript are available by registration to the CHILD database (https://childstudy.ca/childdb/) and the submission of a formal request. All reasonable requests will be accommodated. More information about data access for the CHILD Cohort Study can be found at https://childstudy.ca/for-researchers/data-access/. Researchers interested in collaborating on a project and accessing CHILD Cohort Study data should contact child@mcmaster.ca.

References

1. Sandall, J. *et al.* Short-term and long-term effects of caesarean section on the health of women and children. *The Lancet* **392**, 1349–1357 (2018).
2. Hill, C. J. *et al.* Evolution of gut microbiota composition from birth to 24 weeks in the INFANTMET Cohort. *Microbiome* **5**, 4 (2017).
3. Shao, Y. *et al.* Stunted microbiota and opportunistic pathogen colonization in caesarean-section birth. *Nature* **574**, 117–121 (2019).
4. Tamburini, S., Shen, N., Wu, H. C. & Clemente, J. C. The microbiome in early life: implications for health outcomes. *Nat. Med.* **22**, 713–722 (2016).
5. Kennedy, K. M. *et al.* Questioning the fetal microbiome illustrates pitfalls of low-biomass microbial studies. *Nature* **613**, 639–649 (2023).
6. Blaser, M. J. *et al.* Lessons learned from the prenatal microbiome controversy. *Microbiome* **9**, 8 (2021).
7. Walter, J. & Hornef, M. W. A philosophical perspective on the prenatal in utero microbiome debate. *Microbiome* **9**, 5 (2021).
8. Korpela, K. & de Vos, W. M. Infant gut microbiota restoration: state of the art. *Gut Microbes* **14**, 2118811 (2022).
9. 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).
10. Bogaert, D. *et al.* Mother-to-infant microbiota transmission and infant microbiota development across multiple body sites. *Cell Host Microbe* **31**, 447-460.e6 (2023).
11. Martin, R. *et al.* Early-Life Events, Including Mode of Delivery and Type of Feeding, Siblings and Gender, Shape the Developing Gut Microbiota. *PLoS ONE* **11**, e0158498 (2016).
12. Milani, C. *et al.* Exploring Vertical Transmission of Bifidobacteria from Mother to Child. *Appl. Environ. Microbiol.* **81**, 7078–7087 (2015).
13. Shaterian, N., Abdi, F., Ghavidel, N. & Alidost, F. Role of cesarean section in the development of neonatal gut microbiota: A systematic review. *Open Med.* **16**, 624–639 (2021).
14. Yassour, M. *et al.* Natural history of the infant gut microbiome and impact of antibiotic treatments on strain-level diversity and stability. *Sci. Transl. Med.* **8**, 343ra81 (2016).
15. Matharu, D. *et al.* Bacteroides abundance drives birth mode dependent infant gut microbiota developmental trajectories. *Front. Microbiol.* **13**, (2022).
16. Stokholm, J. *et al.* Cesarean section changes neonatal gut colonization. *J. Allergy Clin. Immunol.* **0**, (2016).
17. Stokholm, J. *et al.* Delivery mode and gut microbial changes correlate with an increased risk of childhood asthma. *Sci. Transl. Med.* **12**, eaax9929 (2020).
18. Bokulich, N. A. *et al.* Antibiotics, birth mode, and diet shape microbiome maturation during early life. *Sci. Transl. Med.* **8**, 343ra82 (2016).
19. Beller, L. *et al.* Successional Stages in Infant Gut Microbiota Maturation. *mBio* **12**, e01857-21 (2021).
20. 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).

21. Alcazar, C. G.-M. *et al.* The association between early-life gut microbiota and childhood respiratory diseases: a systematic review. *Lancet Microbe* **3**, e867–e880 (2022).
22. Browne, H. P., Shao, Y. & Lawley, T. D. Mother–infant transmission of human microbiota. *Curr. Opin. Microbiol.* **69**, 102173 (2022).
23. Knoop, K. A. *et al.* Microbial antigen encounter during a preweaning interval is critical for tolerance to gut bacteria. *Sci. Immunol.* **2**, eaao1314 (2017).
24. Yang, X. *et al.* The prevalence of food allergy in cesarean-born children aged 0–3 years: A systematic review and meta-analysis of cohort studies. *Front. Pediatr.* **10**, 1044954 (2023).
25. Słabuszewska-Jóźwiak, A., Szymański, J. K., Ciebiera, M., Sarecka-Hujar, B. & Jakiel, G. Pediatrics Consequences of Caesarean Section—A Systematic Review and Meta-Analysis. *Int. J. Environ. Res. Public. Health* **17**, 8031 (2020).
26. Roswall, J. *et al.* Developmental trajectory of the healthy human gut microbiota during the first 5 years of life. *Cell Host Microbe* **29**, 765-776.e3 (2021).
27. Stewart, C. J. *et al.* Temporal development of the gut microbiome in early childhood from the TEDDY study. *Nature* **562**, 583–588 (2018).
28. Pantazi, A. C. *et al.* Development of Gut Microbiota in the First 1000 Days after Birth and Potential Interventions. *Nutrients* **15**, 3647 (2023).
29. Kuitunen, M. *et al.* Probiotics prevent IgE-associated allergy until age 5 years in cesarean-delivered children but not in the total cohort. *J. Allergy Clin. Immunol.* **123**, 335–341 (2009).
30. Dominguez-Bello, M. G. *et al.* Partial restoration of the microbiota of cesarean-born infants via vaginal microbial transfer. *Nat. Med.* **22**, 250–253 (2016).
31. Wilson, B. C. *et al.* Oral administration of maternal vaginal microbes at birth to restore gut microbiome development in infants born by caesarean section: A pilot randomised placebo-controlled trial. *eBioMedicine* **69**, (2021).
32. Korpela, K. *et al.* Maternal Fecal Microbiota Transplantation in Cesarean-Born Infants Rapidly Restores Normal Gut Microbial Development: A Proof-of-Concept Study. *Cell* **183**, 324-334.e5 (2020).
33. Davis, E. C. *et al.* Gut microbiome and breast-feeding: Implications for early immune development. *J. Allergy Clin. Immunol.* **150**, 523–534 (2022).
34. Yang, Z. *et al.* Rural environment reduces allergic inflammation by modulating the gut microbiota. *Gut Microbes* **14**, 2125733 (2022).
35. Gao, Y. *et al.* Gut microbiota maturity mediates the protective effect of siblings on food allergy. *J. Allergy Clin. Immunol.* **152**, 667–675 (2023).
36. Ames, S. R., Lotoski ,Larisa C. & and Azad, M. B. Comparing early life nutritional sources and human milk feeding practices: personalized and dynamic nutrition supports infant gut microbiome development and immune system maturation. *Gut Microbes* **15**, 2190305 (2023).
37. Anumula, S., Nalla, K., Pandala, P., Kotha, R. & Harsha, N. Rural Versus Urban Mothers' Microbiome Difference and Its Effect on Neonates: A Systematic Review. *Cureus* **16**, e55607.
38. Derrien, M. *et al.* Gut microbiome function and composition in infants from rural Kenya and association with human milk oligosaccharides. *Gut Microbes* **15**, 2178793.

39. Laursen, M. F. *et al.* Having older siblings is associated with gut microbiota development during early childhood. *BMC Microbiol.* **15**, 154 (2015).
40. Subbarao, P. *et al.* The Canadian Healthy Infant Longitudinal Development (CHILD) Study: examining developmental origins of allergy and asthma. *Thorax* **70**, 998–1000 (2015).
41. Donald, K. & Finlay, B. B. Early-life interactions between the microbiota and immune system: impact on immune system development and atopic disease. *Nat. Rev. Immunol.* **23**, 735–748 (2023).
42. Gregory, K. E., LaPlante, R. D., Shan, G., Kumar, D. V. & Gregas, M. Mode of Birth Influences Preterm Infant Intestinal Colonization with Bacteroides Over the Early Neonatal Period. *Adv. Neonatal Care Off. J. Natl. Assoc. Neonatal Nurses* **15**, 386–393 (2015).
43. Troy, E. B. & Kasper, D. L. Beneficial effects of Bacteroides fragilis polysaccharides on the immune system. *Front. Biosci. J. Virtual Libr.* **15**, 25–34 (2010).
44. Eribo, O. A., du Plessis, N. & Chegou, N. N. The Intestinal Commensal, Bacteroides fragilis, Modulates Host Responses to Viral Infection and Therapy: Lessons for Exploration during Mycobacterium tuberculosis Infection. *Infect. Immun.* **90**, e00321-21.
45. Zafar, H. & Saier, M. H. Gut Bacteroides species in health and disease. *Gut Microbes* **13**, 1848158.
46. Laursen, M. F. *et al.* Bifidobacterium species associated with breastfeeding produce aromatic lactic acids in the infant gut. *Nat. Microbiol.* **6**, 1367–1382 (2021).
47. Stuivenberg, G. A., Burton, J. P., Bron, P. A. & Reid, G. Why Are Bifidobacteria Important for Infants? *Microorganisms* **10**, 278 (2022).
48. Ennis, D., Shmorak, S., Jantscher-Krenn, E. & Yassour, M. Longitudinal quantification of Bifidobacterium longum subsp. infantis reveals late colonization in the infant gut independent of maternal milk HMO composition. *Nat. Commun.* **15**, 894 (2024).
49. Vatanen, T. *et al.* Genomic variation and strain-specific functional adaptation in the human gut microbiome during early life. *Nat. Microbiol.* **4**, 470–479 (2019).
50. Taft, D. H. *et al.* Bifidobacterium Species Colonization in Infancy: A Global Cross-Sectional Comparison by Population History of Breastfeeding. *Nutrients* **14**, 1423 (2022).
51. Bäckhed, F. *et al.* Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life. *Cell Host Microbe* **17**, 690–703 (2015).
52. Bessis, S., Ndongo, S., Lagier, J.-C., Raoult, D. & Fournier, P.-E. 'Neglecta timonensis' gen. nov., sp. nov., a new human-associated species. *New Microbes New Infect.* **13**, 13–14 (2016).
53. Huang, F. *et al.* Early-Life Exposure to Clostridium leptum Causes Pulmonary Immunosuppression. *PLoS ONE* **10**, e0141717 (2015).
54. Cibis, K. G., Gneipel, A. & König, H. Isolation of acetic, propionic and butyric acid-forming bacteria from biogas plants. *J. Biotechnol.* **220**, 51–63 (2016).
55. Kirk, K. F., Andersen, K. L., Tarpgaard, I. H. & Nielsen, H. L. Three cases of *Sutterella wadsworthensis* bacteremia secondary to abdominal infections. *Anaerobe* **72**, 102460 (2021).
56. Mukhopadhya, I. *et al.* A comprehensive evaluation of colonic mucosal isolates of Sutterella wadsworthensis from inflammatory bowel disease. *PloS One* **6**, e27076 (2011).

57. Paramsothy, S. *et al.* Specific Bacteria and Metabolites Associated With Response to Fecal Microbiota Transplantation in Patients With Ulcerative Colitis. *Gastroenterology* **156**, 1440-1454.e2 (2019).
58. Girdhar, K. *et al.* The Impact of Parabacteroides distasonis Colonization on Hosts' Microbiome, Metabolome, Immune Responses, and Diabetes Onset. 2024.06.13.598927 Preprint at https://doi.org/10.1101/2024.06.13.598927 (2024).
59. Wei, W. *et al.* Parabacteroides distasonis uses dietary inulin to suppress NASH via its metabolite pentadecanoic acid. *Nat. Microbiol.* **8**, 1534–1548 (2023).
60. Wu, W. K. K. Parabacteroides distasonis: an emerging probiotic? *Gut* **72**, 1635–1636 (2023).
61. Huang, K. *et al.* Molecular characterization of Clostridium perfringens isolates from a tertiary children's hospital in Guangzhou, China, establishing an association between bacterial colonization and food allergies in infants. *Gut Pathog.* **15**, 47 (2023).
62. Kiu, R. *et al.* Particular genomic and virulence traits associated with preterm infant-derived toxigenic Clostridium perfringens strains. *Nat. Microbiol.* **8**, 1160–1175 (2023).
63. Nagpal, R. *et al.* Gut dysbiosis following C-section instigates higher colonisation of toxigenic Clostridium perfringens in infants. *Benef. Microbes* **8**, 353–365 (2017).
64. Christensen, E. D. *et al.* The developing airway and gut microbiota in early life is influenced by age of older siblings. *Microbiome* **10**, 106 (2022).
65. Stokholm, J. *et al.* Maturation of the gut microbiome and risk of asthma in childhood. *Nat. Commun.* **9**, 141 (2018).
66. Enav, H., Bäckhed, F. & Ley, R. E. The developing infant gut microbiome: A strain-level view. *Cell Host Microbe* **30**, 627–638 (2022).
67. Rook, G. A. W. The old friends hypothesis: evolution, immunoregulation and essential microbial inputs. *Front. Allergy* **4**, (2023).
68. Strachan, D. P. Hay fever, hygiene, and household size. *BMJ* **299**, 1259–1260 (1989).
69. Thorsen, J. *et al.* The airway microbiota of neonates colonized with asthma-associated pathogenic bacteria. *Nat. Commun.* **14**, 6668 (2023).
70. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).
71. Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
72. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
73. Bhattacharyya, M. *et al.* AnnotIEM: A Novel algorithm for Species level annotation of 16S gene based microbial OTUs. *F1000Research* **8**, (2019).
74. Yoon, S.-H. *et al.* Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int. J. Syst. Evol. Microbiol.* **67**, 1613–1617 (2017).
75. Federhen, S. The NCBI Taxonomy database. *Nucleic Acids Res.* **40**, D136–D143 (2012).
76. Cole, J. R. *et al.* The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res.* **37**, D141-145 (2009).
77. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590-596 (2013).

78. Stokholm, J. *et al.* Prevalence and Predictors of Antibiotic Administration during Pregnancy and Birth. *PLoS ONE* **8**, e82932 (2013).
79. Lehtimäki, J. *et al.* Urbanized microbiota in infants, immune constitution, and later risk of atopic diseases. *J. Allergy Clin. Immunol.* **148**, 234–243 (2021).
80. Patrick, D. M. *et al.* Decreasing antibiotic use, the gut microbiota, and asthma incidence in children: evidence from population-based and prospective cohort studies. *Lancet Respir. Med.* **8**, 1094–1105 (2020).
81. Amir, A. *et al.* Deblur Rapidly Resolves Single-Nucleotide Community Sequence Patterns. *mSystems* **2**, 10.1128/msystems.00191-16 (2017).
82. DeSantis, T. Z. *et al.* Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. *Appl. Environ. Microbiol.* **72**, 5069–5072 (2006).
83. McDonald, D. *et al.* An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J.* **6**, 610–618 (2012).
84. Lozupone, C. A., Hamady, M., Kelley, S. T. & Knight, R. Quantitative and Qualitative β Diversity Measures Lead to Different Insights into Factors That Structure Microbial Communities. *Appl. Environ. Microbiol.* **73**, 1576–1585 (2007).
85. Smyth, G. K. limma: Linear Models for Microarray Data. in *Bioinformatics and Computational Biology Solutions Using R and Bioconductor* (eds Gentleman, R., Carey, V. J., Huber, W., Irizarry, R. A. & Dudoit, S.) 397–420 (Springer, New York, NY, 2005). doi:10.1007/0-387-29362-0_23.
86. Russel, J. *et al.* DAtest: a framework for choosing differential abundance or expression method. 241802 Preprint at https://doi.org/10.1101/241802 (2018).
87. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B Methodol.* **57**, 289–300 (1995).

Supplementary tables and figures

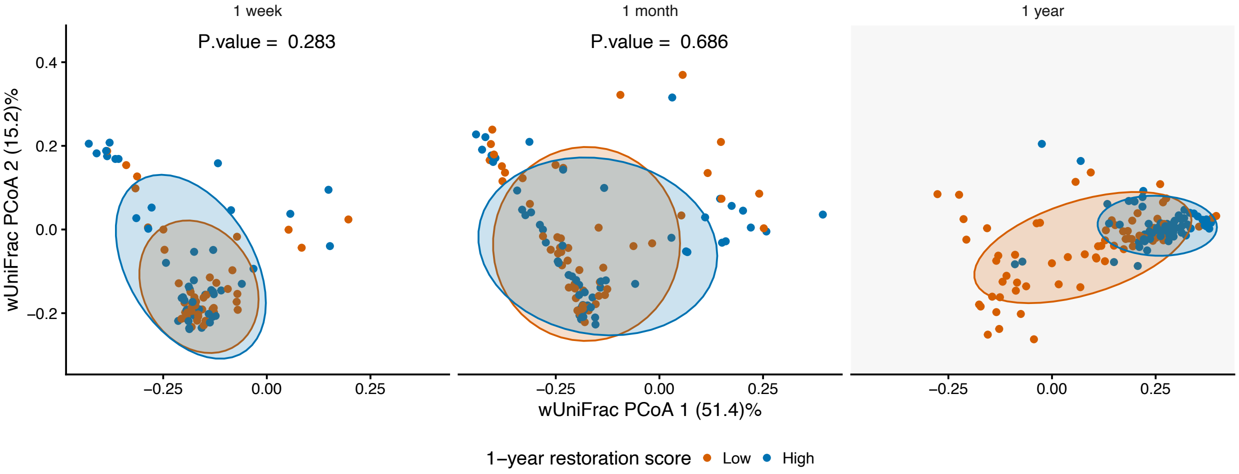


Fig.S1. Comparison of beta diversity in fecal samples during the first year of life according to 1-year restoration score in the CS stratum. Principal coordinate analysis (PCoA) plots of wUniFranc distances for fecal microbial beta diversity in the first week, first month, and first year of life for children having high restoration score (orange, n = 41, 48, 53 at 1 week, 1 month and 1 year) and low restoration score(blue, n = 63, 68, 80 at 1 week, 1 month and 1 year). Gray shading at the 1-year time point marks when the restoration score was derived. P values were calculated by Adonis PERMANOVA tests. Each dot represents one fecal sample. Ellipses encompass 55% of the data points assuming a bivariate normal distribution.

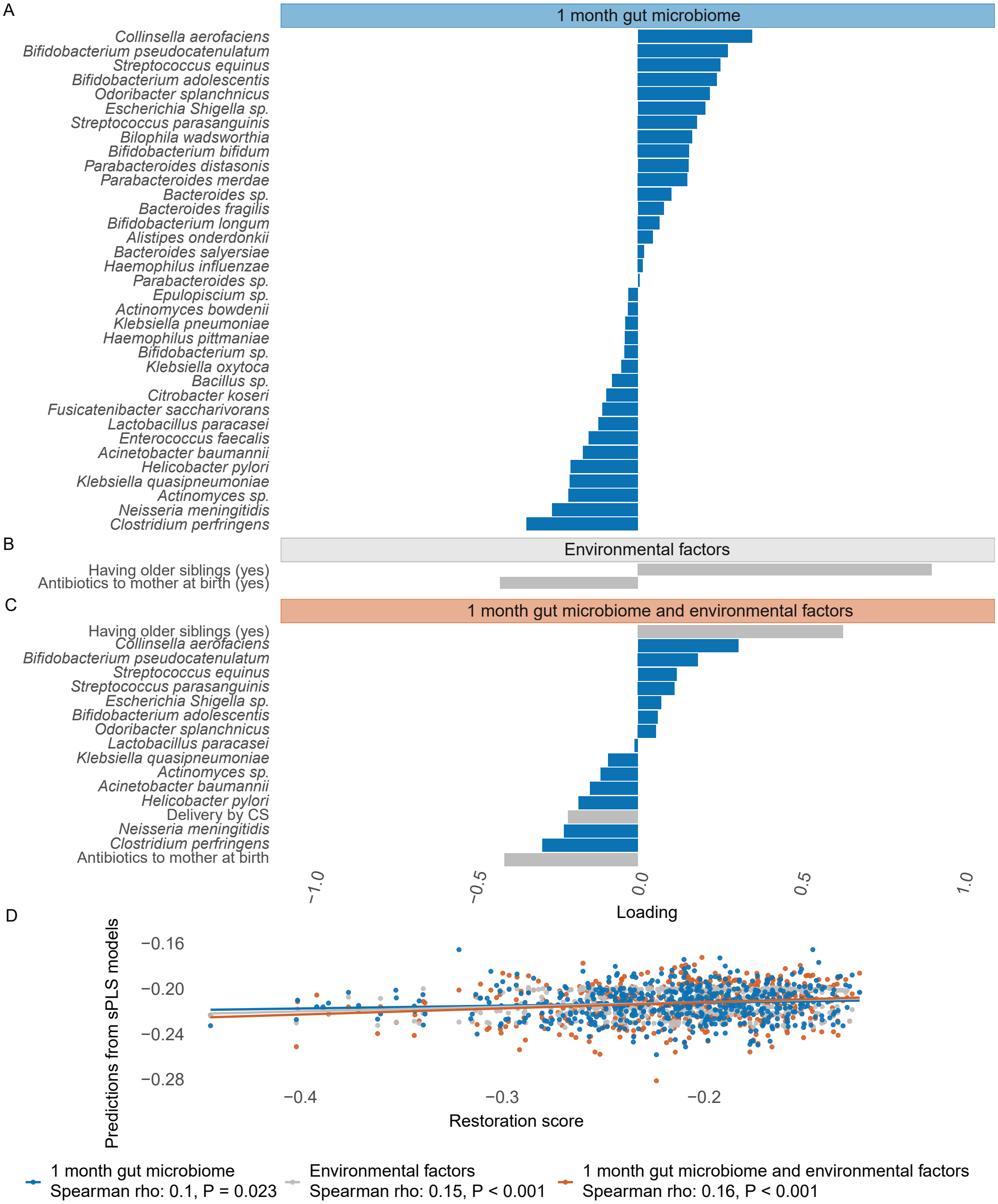


Fig.S2 Sparse partial least squares (sPLS) models according to restoration score in the full cohort at 1 month of age. (A), (B), and (C) show loadings from models on gut microbiome at 1 month of age (152 species, 554 samples), environmental factors (39 factors, 512 samples), and combined factors (191 variables, 512 samples), respectively. Loadings represent the contribution of each variable to the sPLS models. Negative loadings indicate associations with lower restoration scores, while positive loadings indicate associations with higher restoration scores.

(D) Spearman correlation between predictions from the above three sPLS models and the 1-year restoration score. The correlation coefficient and P value are shown in the legend. Blue for the sPLS model on gut microbiome, grey for the sPLS model on environmental factors and orange for the sPLS model on combined gut microbiome and environmental factors.

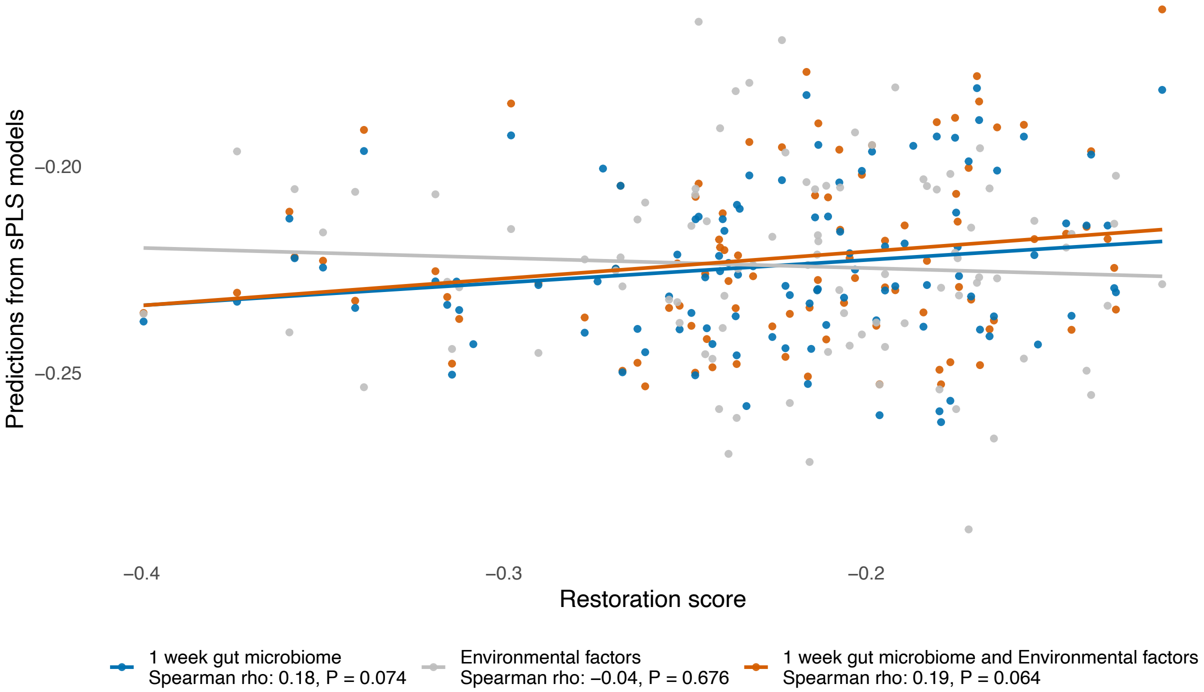
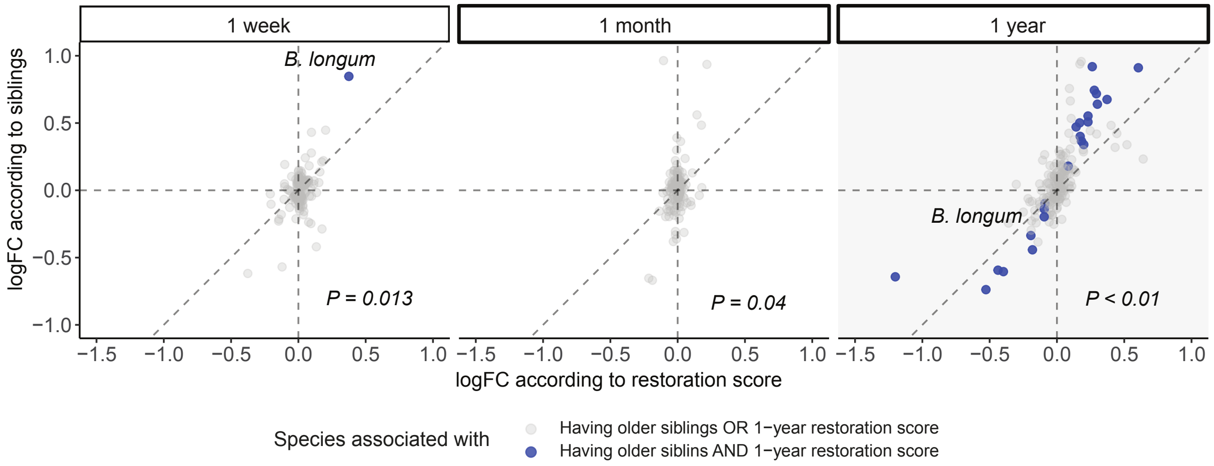


Fig.S3. Spearman correlation between predictions from three sPLS models predicting 1-year restoration score in the CS stratum at 1 week. The correlation coefficient and P value are shown in the legend. Blue for the sPLS model on gut microbiome (140 species, 104 samples), grey for the sPLS model on environmental factor (36 factors, 93 samples) and orange for the sPLS model on combined gut microbiome and environmental factors (176 variables, 93 samples).

A Species associated with having older siblings and 1-year restoration score



B Number of species associated with having older siblins and 1-year restoration score

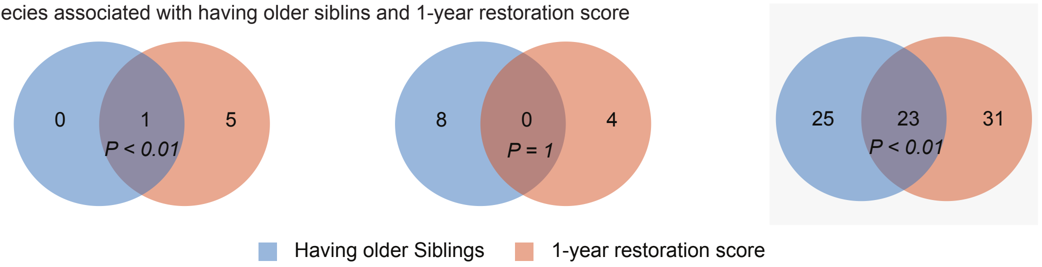


Fig.S4 Differential abundance analysis reveals subset of species associated with having older siblings and restoration in the CS stratum. (A) Scatter plot comparing differential abundance analyses of having older siblings and the restoration score. Each point represents a species, those associated with both factors are highlighted in blue. Labeled species were significant at 1 week and 1 month. A positive log fold change value in the first quadrant (upper right) indicates an increase in the abundance of certain bacteria associated with a higher 1-year restoration score and having older siblings at home, while a negative fold change in the third quadrant (bottom left) indicates a decrease in the abundance of certain bacteria associated with a lower 1-year restoration score and not having older siblings. P values were derived from permutation tests examining whether the direction of the associations with both predictors is concordant by chance, a P value smaller than 0.05 indicates the concordant direction is unlikely under the null hypothesis. (B)Venn diagram indicating the number of differentially abundant species according to having older siblings and according to the 1-year restoration score. P values were derived from permutation tests examining whether the number of taxa associated with both predictors is greater than expected by chance, a P value smaller than 0.05 indicates the overlap is unlikely under the null.

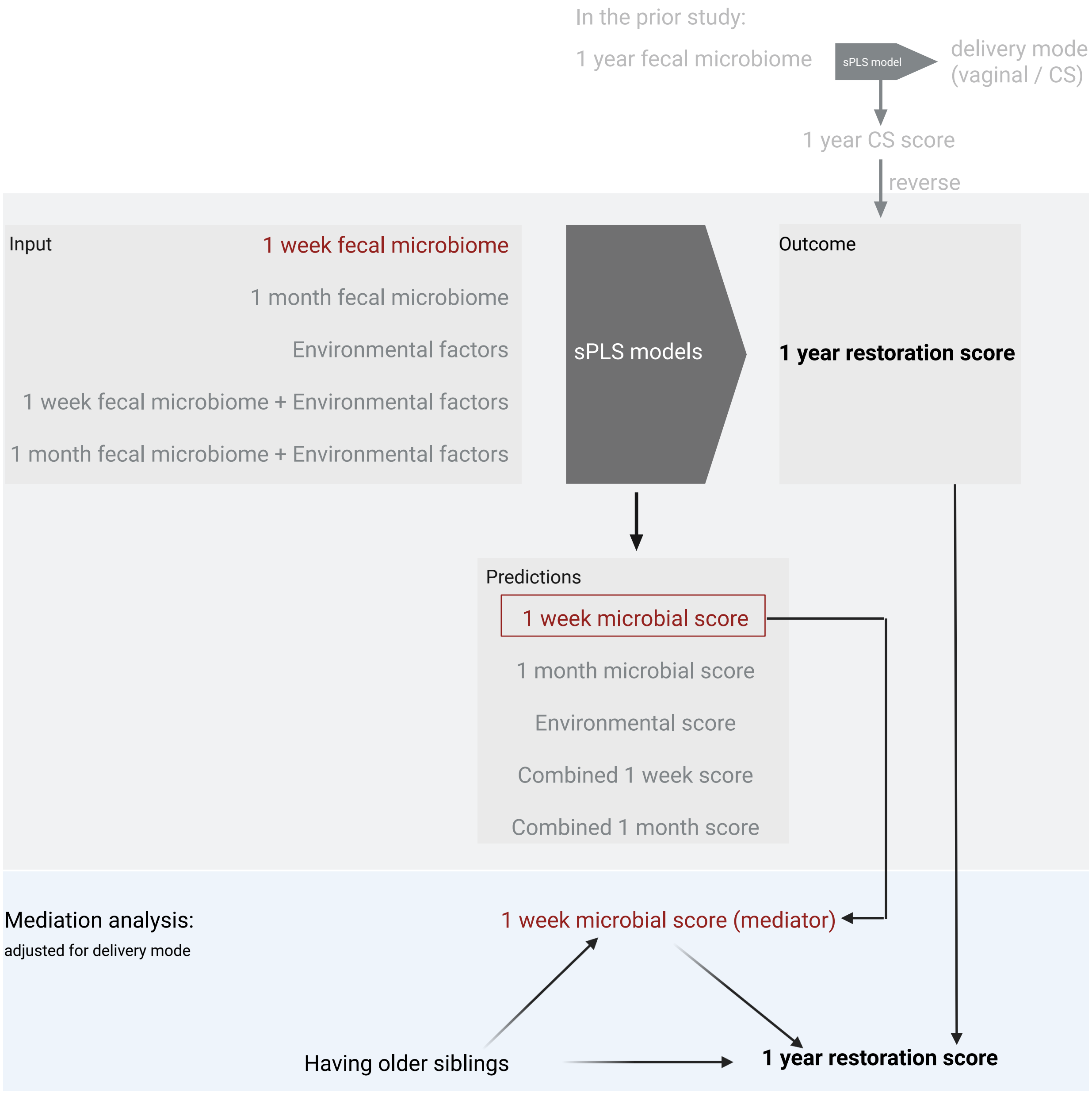


Fig.S5 Flowchart illustrating the overview of the mediation analyses.

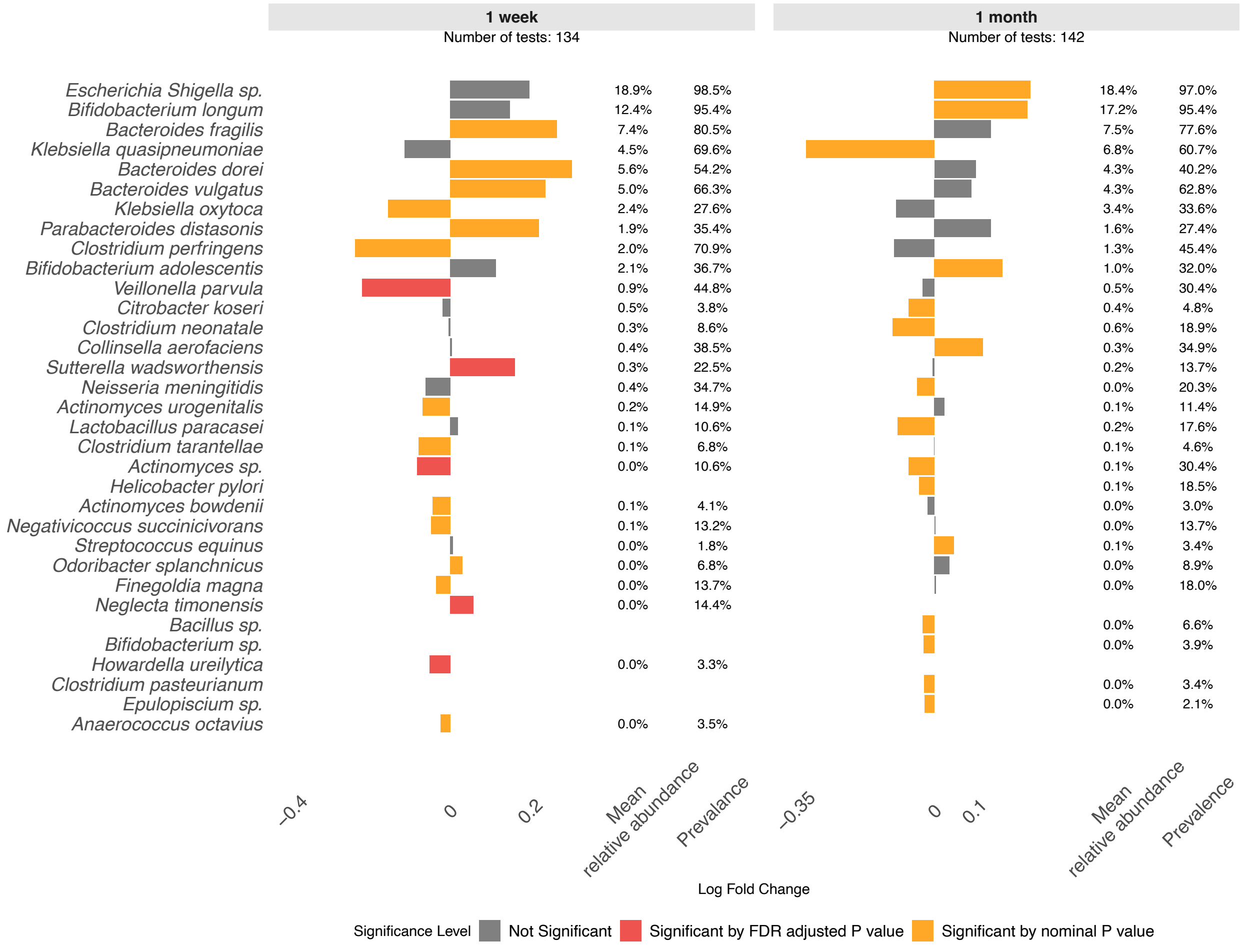


Fig.S6 Differential abundance analysis on the species associated with 1-year restoration score in the vaginal stratum. Species with prevalence of at least 0.1% and relative abundance of more than 0.01% of the total were eligible. The species represented by the red bars were significant after FDR-adjustment, while species represented by yellow bars were nominally significant, and species represented by grey bars were not significant. A positive log fold change value indicates an increase in the abundance of the species as the 1-year restoration score increases, while a negative fold change indicates a decrease.

Appendix - Paper II: Early Gut Microbiome Composition and Functional Pathways Linking Cesarean section to Asthma Susceptibility

Jie Jiang[1,2], Shiraz Shah[1], Lili Yang[3], Kaare Tranæs[1], Casper Sahl Poulsen[1], Ulrika Boulund[1], Trine Zachariasen[3], Xuanji Li[3], Urvish Trivedi[1,3], Klaus Bønnelykke[1,4], Bo Chawes[1,4], Søren J. Sørensen[3], Jonathan Thorsen*[1,4], Jakob Stokholm*[1,2]

[1]COPSAC, Copenhagen Prospective Studies on Asthma in Childhood, Copenhagen University Hospital - Herlev and Gentofte, Copenhagen, Denmark

[2]Department of Food Science, Faculty of Science, University of Copenhagen, Frederiksberg C, Denmark

[3]Department of Biology, Faculty of Science, University of Copenhagen, Copenhagen, Denmark

[4]Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

[4]Institute of Environmental Medicine and Integrative Health, Faculty of Medicine, University of Augsburg, Germany

*Joint senior authors

Corresponding author:

Professor Jakob Stokholm, MD, PhD
E-mail: stokholm@copsac.com
Website: www.copsac.com

Abstract

Cesarean section (CS) has been linked to increased asthma risk, and a mediating role through a prolonged microbial perturbation has been suggested. However, the underlying mechanisms remain unclear, including whether specific microbial species or their functions are involved. Here, we profiled gut microbiome taxonomic and core functional composition at 1 month and 1 year of age in 689 infants from the COPSAC$_{2010}$ cohort born by CS or vaginal delivery. At 1 month, CS-born infants showed very distinct taxonomic and functional compositions. Using machine learning models, we quantified CS-associated microbial and core pathway scores. By 1 year, taxonomic profiles remained distinguishable by delivery mode, whereas core pathway profiles largely converged. We associated these scores with risk of later asthma, prospectively assessed up to 10 years of age. Persistence of a CS-associated microbial score at 1 year characterized by *Citrobacter braakii*, *Veillonella parvula* and *Ruminococcus gnavus*, was associated with increased asthma risk, while core pathway scores were not. Our findings suggest that the persistence of specific gut bacteria, rather than core functional capacity, may be critical for CS-associated asthma risk mediated via the developing gut microbiome.

Keywords:
Cesarean Section(CS); gut microbiome; function; microbial pathway; asthma; COPSAC Cohort Study

Introduction

Cesarean section (CS) delivery has become increasingly common worldwide, rising from approximately 6.7% of births in 1990 to 21.1% by 2018, and are projectedto reach 28.5% by 2030[1–3]. This trend has prompted a major public health concern regarding the long-tem consequences for offspring health[4–7]. A link between CS and increased asthma risk has been extensively demonstrated[8,9] and is increasingly understood to be mediated by a perturbed gut microbiome following CS[10]. The first year(s) of life is believed to represent a “critical window of opportunity” during which the gut microbiota helps shape immune maturationand the establishment of long-term immune tolerance[11]. CS delivery alters the process of bacterial colonization by maternal vaginal and fecal bacteria, resulting in delayed colonization of taxa such as *Bacteroides* and *Bifidobacterium*[12], and an enrichment of Enterobacteriaceae[13]. In our previous study, a prolonged microbial perturbation after CS was associated with a higher risk of asthma up to 6 years of age[14]. However, as our analyses were based on 16S rRNA sequencing, it was unclear whether these associations were mediated by the presence and/or abundance of specific bacterial taxa or their functional capacity.

With the introduction of metagenomic sequencing, studies are increasingly exploring the functional profile of the gut microbiome, shifting the general focus from “who is there” to “what are they doing”. It has been reported that CS can alter the gut microbiome’s metabolic capabilities[15,16]. Lipopolysaccharide (LPS) biosynthesis has been found to be enriched in vaginally born infants, which in turn is associated with elevated levels of tumour necrosis factor (TNF-α) and interleukin 18 (IL-18) in plasma at three days after birth[15,16]. In meconium samples and samples from early infancy, the gut microbiome after vaginal delivery shows enriched amino acid biosynthesis and degradation, while the microbiome after CS delivery shows enrichment of nucleoside and nucleotide degradation and acetyl-CoA biosynthesis[17,18]. The specific functional capacity of the gut microbiome affected by CS and whether these functional alterations might mechanistically contribute to asthma development have not been studied. The lack of coherent, longitudinal data on functional microbiome capacity represents a major barrier to establishing causal relationships between CS-associated microbiota profiles and asthma risk.

In this study, we profiled the taxonomic and functional development of the gut microbiome at 1 month and 1 year of age in children from the Copenhagen Prospective Studies on Asthma in Childhood 2010 (COPSAC_{2010}) mother-child cohort using shotgun metagenomics. We investigated how CS was associated with the gut microbiome composition and its core functional pathways. We then associated CS microbial and CS core pathway scores with later asthma risk, environmental factors, and immune profiles, aiming to provide insights into early host-microbial interactions that may underlie the elevated asthma risk observed in CS-delivered children.

Results

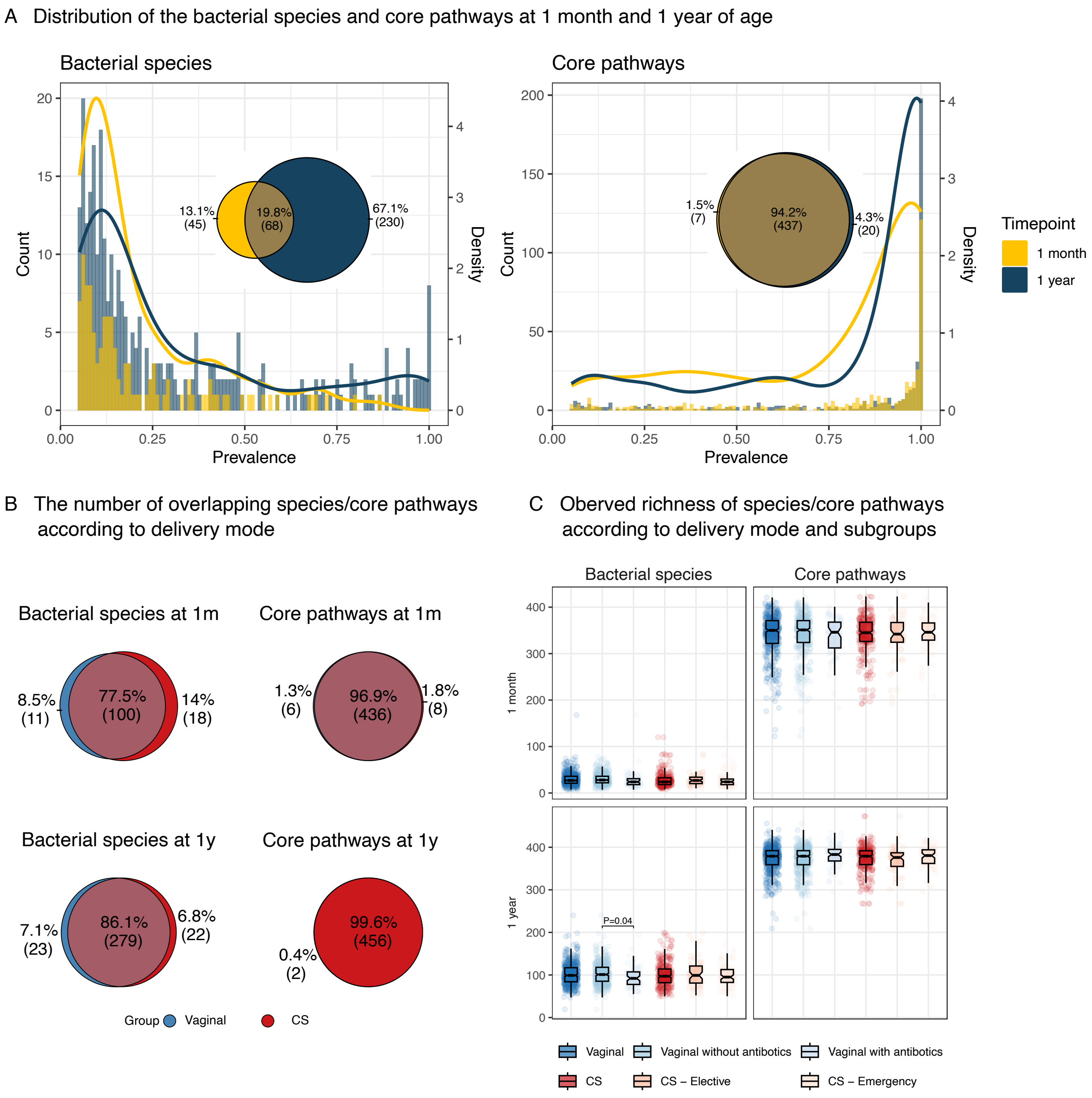
Delivery-related characteristics of the $COPSAC_{2010}$ cohort

This study includes data from 689 children (98% of 700 children) from the prospective $COPSAC_{2010}$ mother-child cohort, of which 150 children were born by CS. Characteristics of children born by CS have been described in previous work[14]. In $COPSAC_{2010}$ cohort, CS delivery was associated with lower gestational age, higher maternal age, higher maternal pre-pregnancy body mass index, increased rate of hospitalization of the child after birth, and a decreased rate of exclusive breastfeeding. All mothers giving birth by CS received intrapartum antibiotics. By contrast, only 13.2% of mothers who gave birth by vaginal delivery recieved intrapartum antibiotics (Table S1). Of the 539 women who gave birth vaginally, 22 (4.3%) were diagnosed with preeclampsia, 128 (25.2%) underwent induction of labor and 52 (7.8%) delivered by vacuum extraction. Prelabor rupture of membranes occurred in 40 (7.9%) women: 35 at >= 37 weeks gestation and 5 as preterm prelabor rupture of membranes. Asphyxia was recorded in 106 (20.8) infants, of which 26 were exposed to meconium-stained amniotic fluid (Table S2).

Distribution of the gut microbiome and its core pathways at 1 month and 1 year

We first profiled the gut microbiome and its functional pathways in infants at 1 month (N=600) and 1 year (N=660) of age (Fig.1). After filtering out bacteria with a prevalence below 5% across all samples within each time point, we identified 113 and 298 bacterial species at 1 month and 1 year of age, respectively, for a total of 343 species. For functional profiling, sequencing reads were mapped to pathways using HUMAnN 3[19], yielding 444 and 457 pathways with a prevalence above 5% at 1 month and 1 year of age, respectively; 464 pathways in total. These core pathways accounted for 6% and 5% of the total read-level abundance (Copies Per Million, CPM) at 1 month and 1 year, respectively. Unmapped reads and unintegrated genes (ie. genes not assigned to a core pathway) were excluded from all downstream analyses (1 month: 12% unmapped, 82% unintegrated of total CPM; 1 year: 15% unmapped, 80% unintegrated). Thus, our pathway abundance dataset, henceforth referred to as core pathways, focuses on broadly shared, well-annotated functions and represents a conservative view of the core functional capacity of the infant gut microbiome, likely understating the contribution of rare or less well characterized pathways.

In general, most bacterial species had low prevalence across samples at 1 month and only slightly higher at 1 year of age (Fig.1A Bacterial species, Table S3). Among all 343 detected species, 45 (13.1%) were unique to the 1-month time point, such as *Staphylococcus epidermidis* (prevalence 68%), *Streptococcus vestibularis* (49%), and *Staphylococcus aureus* (45%). At the 1-year time point, 229 (67.1%) were unique including *Faecalibacterium prausnitzii* (100% prevalence), *Anaerostipes hadrus* (98%), and *Ruminococcus bromii* (96%); and 69 species (19.8%) were present at both time points, including *Streptococcus salivarius* (prevalence at 1 month/1 year: 87%/80%), *Bifidobacterium longum* (80%/100%), and *Veillonella dispar* (72%/83%). In contrast, most pathways were present in all samples at 1 month and 1 year of age (Fig.1A Core pathways, Table S4). Among all 464 core pathways, 437 (94%) pathways were shared between time points, 7(2%) and 20(4%) pathways were unique to the 1-month and 1-year time point, respectively.


A Distribution of the bacterial species and core pathways at 1 month and 1 year of age

B The number of overlapping species/core pathways according to delivery mode

C Oberved richness of species/core pathways according to delivery mode and subgroups

Fig.1 Prevalence of the bacterial species and core pathways. (A) Prevalence and the number of overlapping species/pathways at 1 month and 1 year of age (prevalence > 5% across samples within each time point). (B) The number of shared species/pathways between the vaginal and CS group at 1 month and 1 year of age (prevalence > 5% across samples within each stratum at each time point). (C)Observed richness of species/pathways according to delivery mode and antibiotics use. The differences were tested using linear regression, adjusted for sequencing depth and batch. P values above 0.05 are not shown.

We then investigated how many bacterial species and metabolic pathways were shared or unique to the vaginal and CS groups at 1 month and 1 year of age after prevalence-filtering ((Vaginal/CS at 1 month: N=465/135; at 1 year: N=517/143, Fig.1B, Table S5-8). The number of species shared between the groups increased from 100 (77.5% of all species) at 1 month to 279

(86.1%) at 1 year. At 1 month of age, 436 core pathways (96.9% of total core pathways) were shared between groups, which increased to 456(99.6%) at 1 year. We then tested differences in observed bacterial species richness and core pathway richness using linear regression (Fig. 1C). The only significant difference was a lower bacterial species richness at 1 year of age in infants born vaginally whose mothers received intrapartum antibiotics compared with infants born vaginally whose mothers did not receive intrapartum antibiotics. No differences were detected in core pathway richness.

Compositional differences of the gut microbiome and core pathways in relation to delivery mode

Associations between delivery mode, overall gut microbial composition and core pathway composition (beta diversity) were evaluated based on Bray-Curtis dissimilarity metric (PERMANOVA, Fig. 2, Table S9). At 1 month of age, we observed birth mode-dependent compositional differences, mainly in the bacterial species but also in the core pathways. At 1 year of age, differences in species composition were smaller but remained significant, whereas the differences observed in core pathway composition between the vaginal and CS groups were no longer significant.

To evaluate whether some of the compositional differences could be caused by intrapartum antibiotic treatment, which is administered during all CS deliveries, we compared the gut microbial composition between vaginally delivered children born by treated or untreated mothers (PERMANOVA, Table S9). At 1 month, the gut microbial composition of the vaginally delivered children whose mothers were treated with intrapartum antibiotics was not different from that of infants whose mothers were not treated (F=1.31, R2=0.3%, P=0.15). However, differences were detected between these two groups at 1 year of age (F=1.69, R2=0.3%, P=0.03), echoing the findings from the alpha diversity analysis (Fig. 1). Notably, by 1 year of age, the gut microbial composition of vaginally born infants exposed to intrapartum maternal antibiotics did not differ significantly from that of CS-born infants (F=1.26, R2=0.6%, P=0.17). In contrast, at 1 month of age we detected differences in core pathway composition between the CS-born infants and vaginally born infants without antibiotic exposure (F=6.54, R2=1.2%, P=0.003) as well as between vaginally born infants with versus without intrapartum antibiotics exposure (F=3.12, R2=0.7%, P=0.02). No significant differences in core pathway composition were observed between any groups at 1 year of age. Also, no significant differences were detected between elective and emergency CS delivery at any time point for either microbial or core pathway compositions (Table S9).

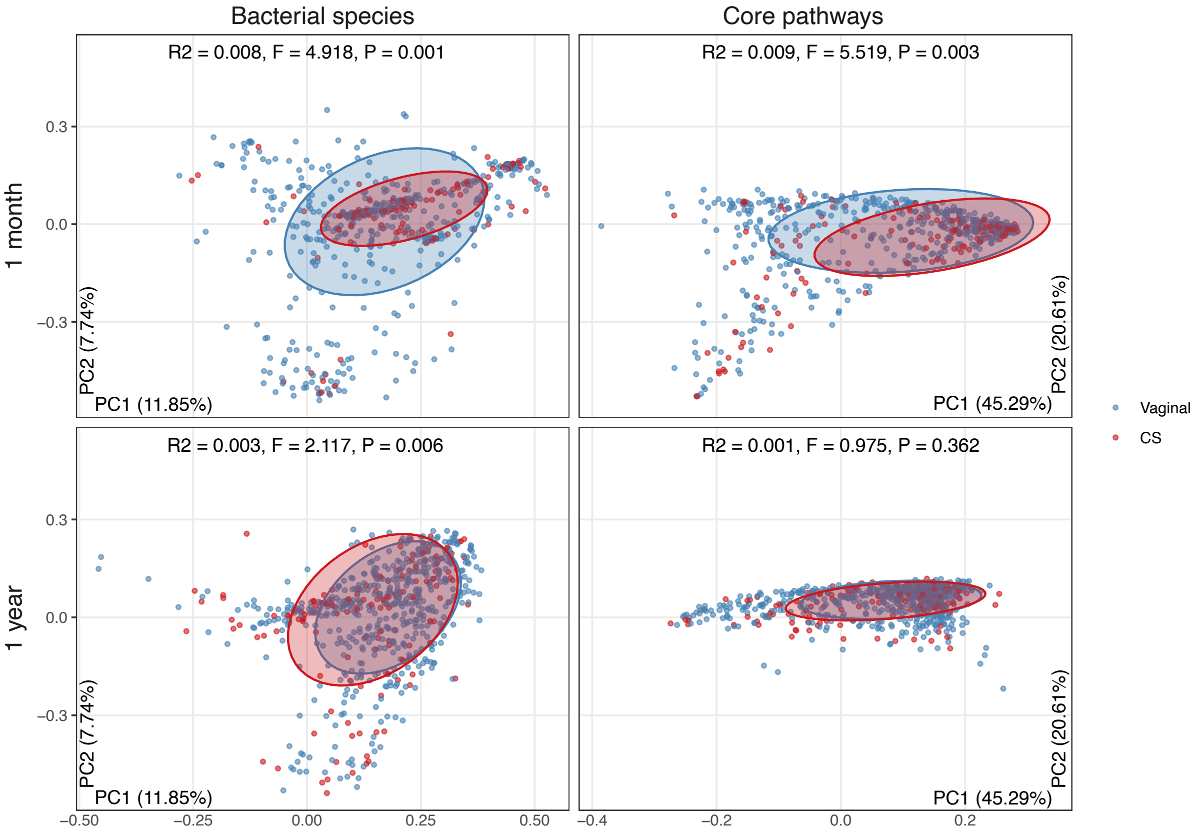


Fig.2 Comparison of beta diversity of bacterial species and core pathways in relation to delivery mode. PERMANOVA was used to compare group differences. Each dot represents a fecal sample. Ellipses represent 1 standard deviation (1 SD), encompassing approximately 68% of the data points, assuming a bivariate normal distribution.

Differential abundance of the gut microbiome and core pathways in relation to delivery mode

We investigated the abundance of individual bacterial species and core pathways associated with delivery mode (Fig.3, Table S10-13). To minimixe confounding by intrapartum antibiotics exposure we excluded vaginally delivered infants whose mothers received intrapartum antibiotics. We retainedbacterial species with >5% prevalence and >0.01% relative abundance within each time point, and core pathways with >5% prevalence., After filtering 90 species and 443 pathways were available at 1 month of age, and 229 species and 457 core pathways at 1 year of age.

At 1 month of age, 29 species and 159 core pathways were significantly different between infants born by CS and by vaginal delivery (FDR adj P<0.05, Fig.3). CS was associated with higher relative abundances of 13 species, including *Enterococcus faecalis*, *Clostridium perfringens*, and *Veillonella parvula*, and 103 core pathways including superpathway of glycerol degradation to 1,3-propanediol, molybdopterin biosynthesis, and formaldehyde assimilation III (dihydroxyacetone cycle). Conversely, vaginal delivery was associated with higher relative abundances of 16 species, including *Phocaeicola vulgatus*, *Phocaeicola dorei*, and *Parabacteroides distasonis*, and 56 core pathways, including L-histidine degradation III, dTDP-beta-L-rhamnose biosynthesis, and colanic acid building blocks biosynthesis. At 1 year of age, only *Bifidobacterium dentium* and succinate fermentation to butanoate was found to be individually associated with CS (*Bifidobacterium dentium* FDR adj P=0.03, succinate fermentation to butanoate FDR adj P=0.05).

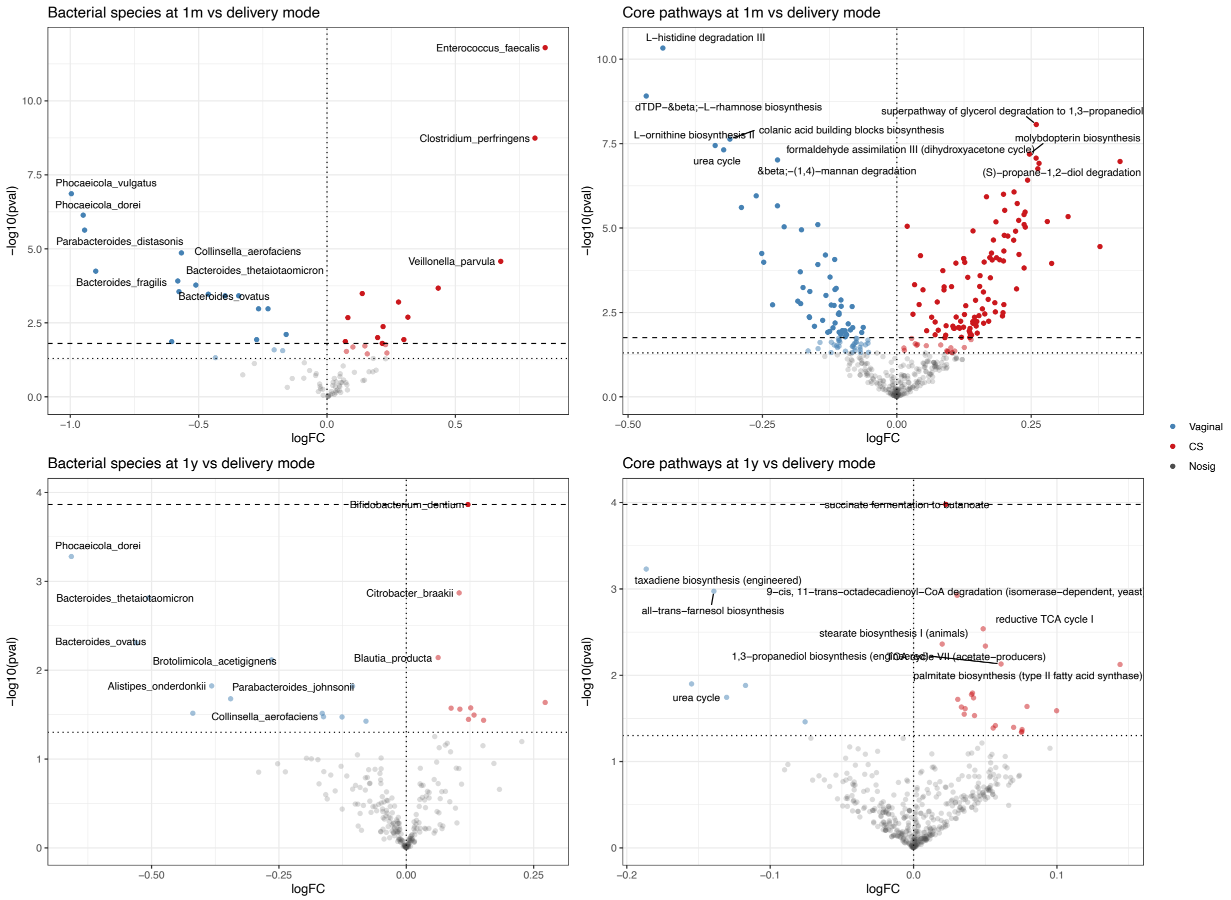


Fig.3 Differential abundance analysis of bacterial species and core pathways according to delivery mode. Species with prevalence of at least 5% and relative abundance of more than 0.01% of the total were eligible. Core pathways with prevalence of at least 5% were eligible. All analyses were adjusted for sequencing depth and batch. Red dots represent species/pathways more abundant in the CS group, blue dots represent those more abundant in the vaginal without antibiotics group. The dotted line represents nominal P<0.05, and the dashed line represents FDR-adjusted P<0.05.

Supervised models of delivery mode and the gut microbiome and core pathways

We next constructed cross-validated sparse partial least squares (sPLS) models to identify jointly correlated species and core pathway abundances associated with CS delivery against the vaginal without antibiotics group at 1 month and 1 year of age (1 month, N=532; 1 year, N=589). These models identified the sets of gut microbial species and core pathways at each time point most closely associated with CS using variable selection via L1-norm penalization. We extracted cross-validated prediction scores from these models and used them as CS microbial scores.

A 2-component model captured 31 different species in the 1-month CS microbial profile (Fig.4A). The first component mainly included species positively correlated with CS, such as *Clostridium perfringens*, *Enterococcus faecalis*, and *Veillonella parvula*, as well as the vaginal birth associated *Parabacteroides distasonis*, *Phocaeicola dorei*, and *Bacteroides thetaiotaomicron*.

In parallel, a 1-component model selected 39 core pathways to best predict delivery mode at 1 month (Fig.4B). Pathways positively associated with CS included superpathway of glycerol degradation to 1,3-propanediol, (S)-propane-1,2-diol degradation, 9-cis, 11-trans-octadecadienoyl-CoA degradation (isomerase-dependent, yeast). The abundances of these pathways were contributed to by *Klebsiella*, *Citrobacter*, and *Clostridium* (Fig.4C). In contrast, pathways associated with vaginal delivery included beta-(1,4)-mannan degradation, colanic acid building blocks biosynthesis, and superpathway of GDP-mannose-derived O-antigen building blocks biosynthesis, which were mostly contributed by *Escherichia*, *Bacteroides*, *Klebsiella*, and *Clostridium* (Fig.4C). The composition of species involved in each pathway was similar in infants born by vaginal delivery and those born by CS, suggesting that the differences in pathway abundances may be driven by the differential abundance of bacterial species rather than by changes in overall microbiota composition.

The model predictions were utilized as 1-month CS microbial and core pathway scores, and tested for associations with asthma from 1 to 10 years of age. There was no detectable association between these 1-month scores and asthma risk (Fig.S1A+B, P>0.05).

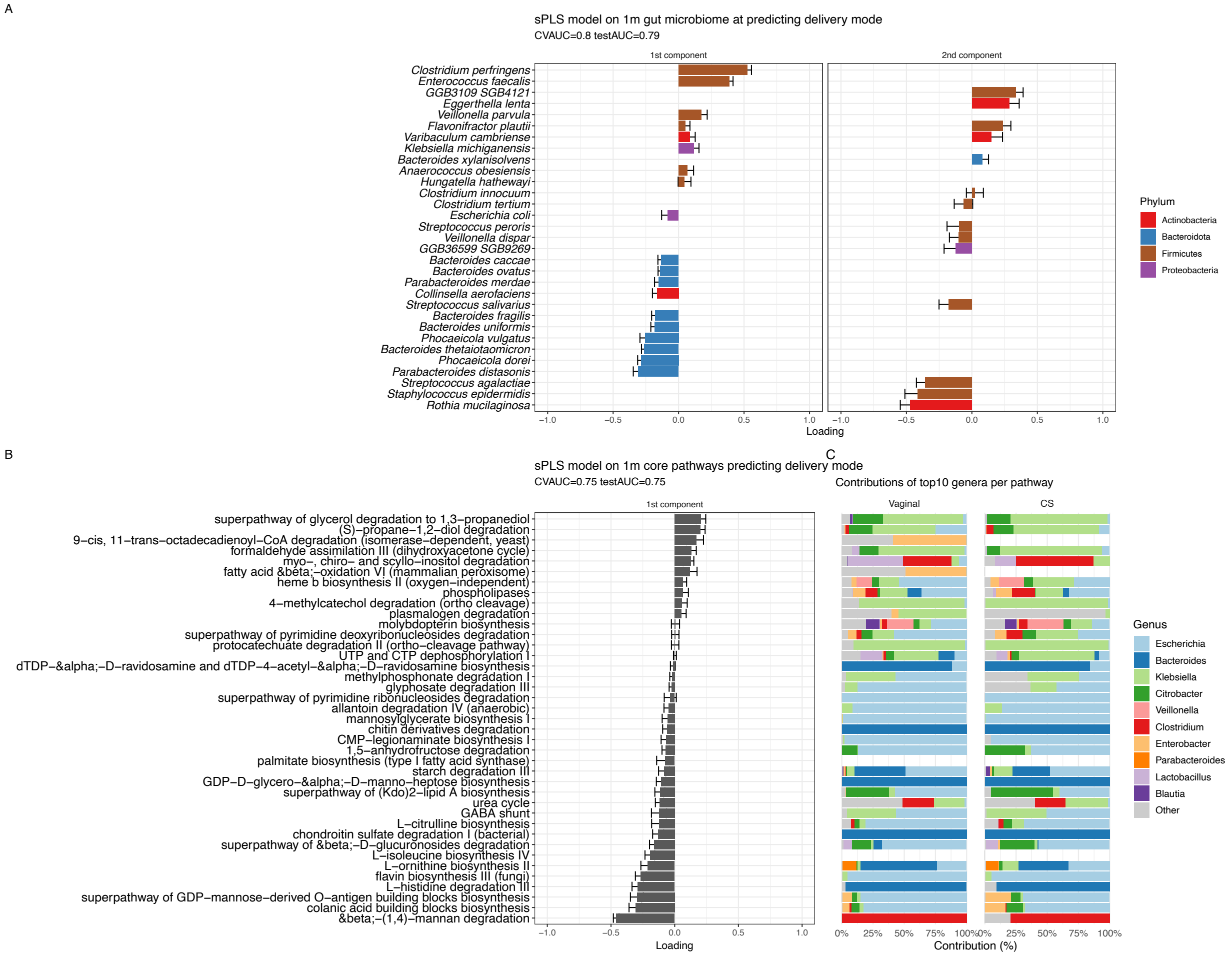


Fig.4 sPLS model predicting delivery mode at 1 month of age based on the (A) gut microbiome and (B,C) microbial core pathways. Sparse partial least squares (sPLS) models were used to distinguish (A) fecal microbial composition and (B) microbial core pathway composition between children born by CS and vaginal delivery without antibiotics. Receiver-operator area under the curve (AUC_{ROC}) values calculated from repeated 10-fold cross-validation of the models (CVAUC) and in the test set (25% samples, testAUC) are shown. A random prediction yields a score of 0.5, with a perfect prediction being 1. Loadings (model contributions) for each species are ranked by magnitude and colored by phylum. Negative/positive loadings correspond to lower/higher abundance in CS born infants compared to vaginally delivered infants. (C) Bacterial contributions to each pathway, stratified by delivery mode. Contribution (%) represents the proportion of each species' relative abundance divided by the total relative abundance of all species contributing to that pathway within each stratum. (Note: unclassified contributions were excluded.)

At 1 year of age, a 1-component model included 39 species in the 1-year CS microbial score (Fig.5A). The 3 main contributors positively correlated with CS were *Citrobacter braakii*, *GGB4593 SGB6353*, and *Intestinimonas butyriciproducens*; the 3 main contributors negatively correlated with CS were *Phocaeicola dorei*, *Bacteroides thetaiotaomicron*, and *Bifidobacterium pseudocatenulatum*. In contrast, only two pathways were selected in a similar model (Fig.5B), which were chondroitin sulfate degradation I (bacterial) and GDP-D-glycero-alpha-D-manno-heptose biosynthesis. In line with the alpha and beta diversity results, this may indicate that the overall functional (core pathway) capacity of the gut microbiome was nearly identical between infants born vaginally and by CS.

When associating 1-year CS microbial scores and 1-year CS core pathway scores with asthma risk to 10 years of age, we found higher microbial scores to be associated with higher asthma risk over time (Generalized Estimating Equation (GEE) Odds ratio (OR) per SD=1.21, 95% confidence interval [1.01,1.45], P=0.03; Adjusted GEE OR=1.23, [1.02,1.48], P=0.03) (Fig.5D), while the core pathway score was not associated with asthma risk (GEE OR=1.08, [0.88,1.32], P=0.46; Adjusted GEE OR=1.07, [0.86,1.32], P=0.56) (Fig.5E).

We next examined whether any species or core pathways not picked out by the models above were also positively associated with the 1-year CS microbial and core pathway scores, which can occur when features capture redundant information and/or their inclusion does not improve model performance. Here, we found substantial differences in both taxonomic and core pathway abundances (Fig.S2): 70 species and 361 core pathways were found to be significantly associated with their respective CS microbial scores at 1 year. Besides *Citrobacter braakii* which was selected in the sPLS model, there were *Veillonella parvula*, *Ruminococcus gnavus*, *Bifidobacterium longum* and superpathway of pyrimidine deoxyribonucleosides degradation, 5-oxo-L-proline metabolism, and heme b biosynthesis II (oxygen-independent) to be strongly associated with higher CS microbial scores. By examining the bacterial species mapped to these pathways, we found that the abundances of these three top pathways were contributed to by *Citrobacter braakii*(Fig.S3).

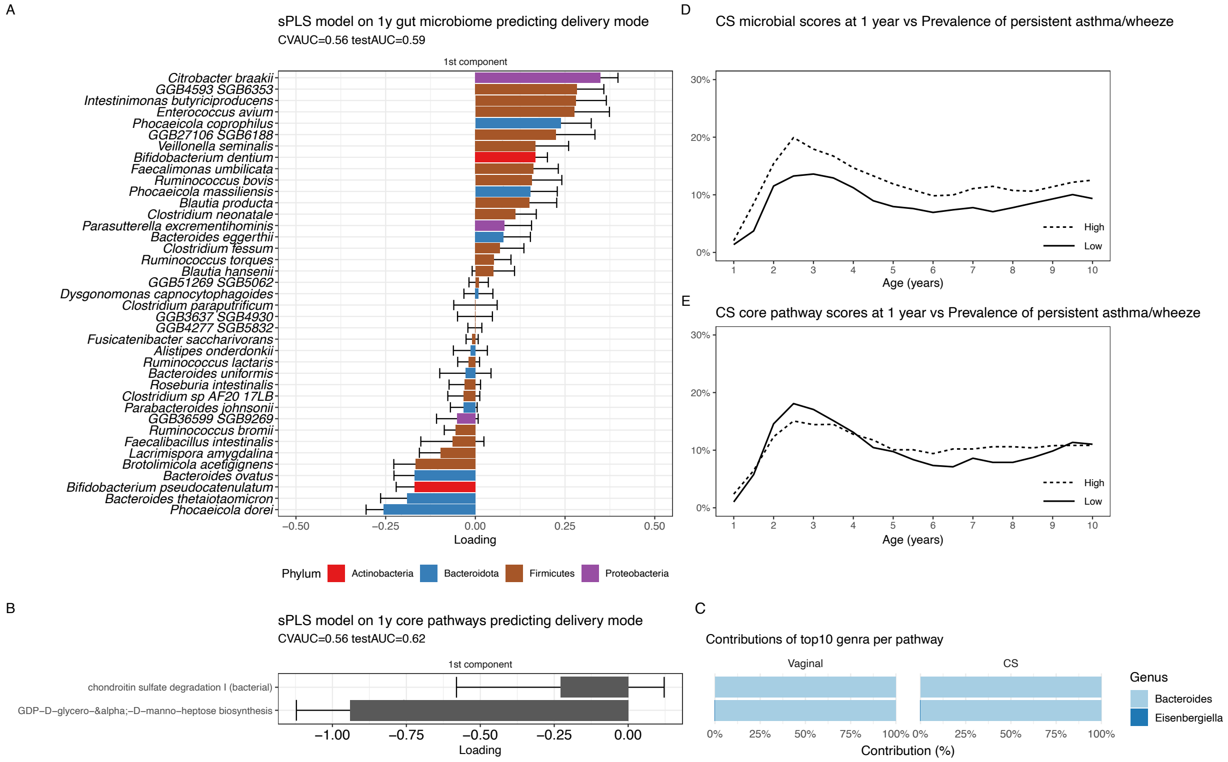


Fig.5 sPLS model predicting delivery mode at 1 year of age based on the (A) gut microbiome and (B,C) core pathways. Sparse partial least squares (sPLS) models were used to distinguish (A) fecal microbial composition and (B) core pathway composition between children born by CS and vaginal delivery without antibiotics. Receiver-operator area under the curve (AUCROC) values calculated from repeated 10-fold cross-validation of the models and in the test set are shown. A random prediction yields a score of 0.5, with a perfect prediction being 1. Loadings (model contributions) for each species are ranked by magnitude and colored by phylum. Negative/positive loadings correspond to lower/higher abundance in CS born infants compared to vaginally delivered infants. (C)Bacterial contributions to each pathway, stratified by delivery mode. Contribution (%) represents the proportion of each species' relative abundance divided by the total relative abundance of all species contributing to that pathway within each stratum. (Note: unclassified contributions were excluded.) (D) and (E) The scores were dichotomized into high/low (above/below median) to visualize the prevalence of persistent asthma/wheeze, P values in the main text were measured using continuous scores, covariates adjusted in the model GEE models included: exclusive breastfeeding at 1 month, any breastfeeding at 1 year, hospitalization after birth, gestational age, maternal bmi, maternal asthma history, household income at 1 year, passive smoke exposure, maternal age, paternal age, the presence of older siblings, sex, race, birth season, sequencing depth, and batch.

Early life exposures and CS microbial and core pathway scores

Next, we investigated which early life factors were associated with the CS scores. All CS scores were strongly correlated with one another and remained so in the vaginal stratum, but not significant in the CS stratum (Fig.S4). Antibiotic exposure to children at birth and hospitalization after birth were positively associated with CS scores, while gestational age, the presence of older siblings, and exclusive breastfeeding at 1 month were negatively associated with CS scores. In addition, maternal fucose secretion (FUT2 gene, rs602662 single nucleotide variant) was associated with higher 1-year CS microbial score and 1-month CS core pathway score, possibly indicating the effect of different compositions of breastmilk on the functional profile of the gut microbiome at 1 month of age, which may contribute to shaping the gut microbiome composition at 1 year of age. Interestingly, the child's secretor status showed the opposite association with 1-month and 1-year CS microbial scores (Fig.6).

Furthermore, perinatal factors and birth complications were evaluated in relation to the CS scores in the vaginal stratum. Birth complications, including asphyxia, assisted delivery, and rupture of membranes, showed positive associations with higher 1-year microbial scores, although none remained significant after FDR correction (Fig.S5).

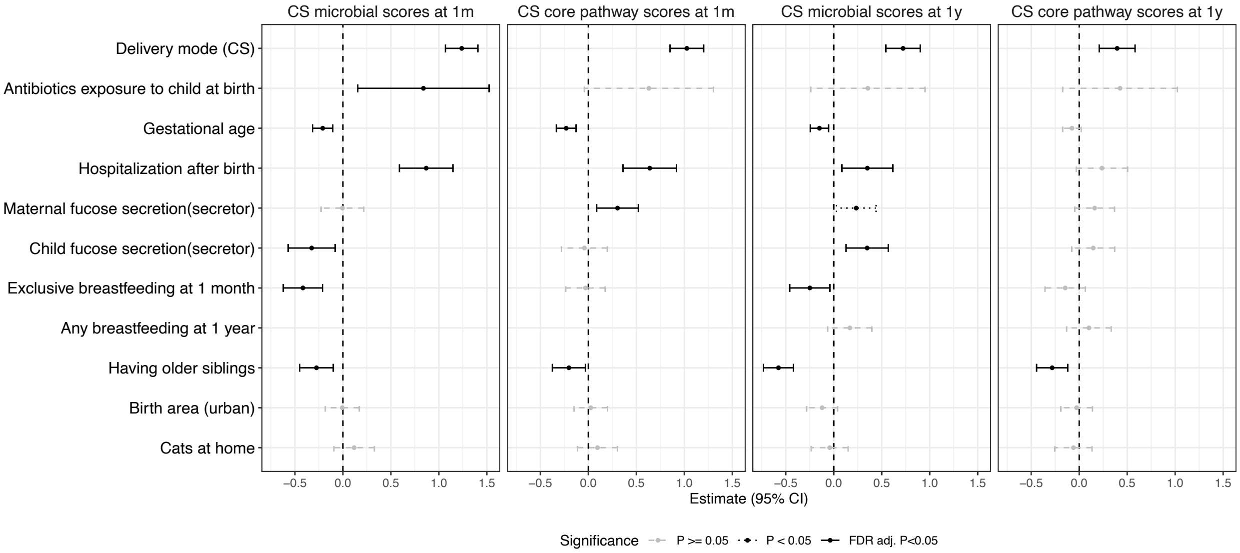


Fig.6 CS scores associated with environmental factors and exposures using linear regression, adjusting for sequencing depth and batch.

Discussion

Our study shows that delivery by CS was associated with substantial differences in both gut microbial and core pathway composition at 1 month of age, while only the gut microbial composition remained distinguishable by delivery mode at age 1 year. Moreover, persistence of a high CS microbial score to 1 year of age was associated with an increased risk of asthma up to 10 years of age, while core pathway scores were not associated with asthma risk at any time point. This suggests that the link between CS and risk of asthma may be driven more by abundances of specific bacterial taxa rather than their core functions.

Major strengths of this study include the longitudinal design with repeated sampling, good sample size, and deep metagenomic resolution of both taxonomic and core functional profiles. A key limitation of our work is the incomplete annotation of microbial functional pathways, especially when stratifying pathway abundances to the species level. More than 80% of the reads mapped to genes not assigned to any core pathways, indicating that while these pathways indeed cover core microbiome functions, many other functions exist which are not described in the HUMAnN pipeline used in this study. This poses a challenge, as species-specific and/or non-core functionality could be important for asthma development. Thus, we can only confidently describe core pathways, and our taxonomic stratification of functions does not capture the entirety of species functional profiles. Furthermore, the association between the CS microbial profiles at age 1-year and later risk of asthma could be affected by unmeasured or residual confounding, however it did survive extensive covariate adjustment. Finally, higher CS microbial scores are not exclusive to children born by CS, and can also be found in vaginally delivered children in part due to other early life exposures.

Our results showed a substantial increase in the gut microbial richness during the first year of life, while the number of core pathways increased only modestly, suggesting pronounced taxonomic turnover but with a relatively stable core functional capacity. This was additionally exemplified in many individual species found only at one of the timepoints, while the two timepoints were almost identical with regards to presence of core functions. Over the first year of life, there's an increase in microbes such as *Blautia*, *Ruminococcus*, and *Faecalibacterium*, replacing those dominant in infancy: *Bacteroides*, *Bifidobacterium*, and *Escherichia coli*[20]. The microbiomes of newborns are enriched in genes required for degrading sugars from the breast milk[21], but some pathways can be used to degrade more complex dietary glycans later in infancy[22–24]. Thus, the overall core functional repertoire already appears to be established by 1 month of age.

Delivery mode is an important driver for early gut microbial acquisition and development, in particular for the first year of life[25]. In our study, we mainly observed presence/absence differences between delivery modes at the taxonomic level, while at the core functional level no unique functions were found for vaginal or CS-born children. Here, only 77.5% and 86.1% of species were shared between delivery modes at 1 month and 1 year, respectively. Furthermore, higher early abundance of *Enterococcus faecalis*, *Clostridium perfringens*, *Veillonella parvula*, and *Klebsiella* was observed in the CS born infants, consistent with environmental and hospital sources[26] as opposed to maternal. This microbial profile is also consistent with multiple studies reporting that CS born infants have a gut microbiome with increased gut redox potential, resulting from a delayed establishment of strict anaerobes[27,28]. CS born infants have lower abundances of *Bacteroides* and *Phocaeicola* early, which was pronounced in our 1-month samples, and for many children these lower abundances persist to age 1 year. It has been reported that even if CS-born children acquire these taxa early, they may fail to maintain them, leading to depleted *Bacteroides* and persistent underrepresentation up to 6-18 months[29]. Previous studies have examined microbial pathway differences between infants born by CS and vaginal delivery and demonstrated "functional convergence despite taxonomic differences"[30–32]. This suggests that even though the microbial communities are different, they do not necessarily differ in their functional potentials[33]. The differences in abundances that we observe for both taxonomy and core functions, among infants born by CS vs vaginal delivery at 1 month of age could be due to the fact that CS-born infants are initially exposed to microbes that are more taxonomically and functionally diverse[34] than the initial taxa in a vaginally delivered child, a difference declining by age 1 year, as seen in Fig 3. Besides CS, we observed a difference in core pathways associated with intrapartum antibiotics at 1 month of age in the vaginal subgroup analysis of beta diversity while the taxonomic composition was comparable, suggesting a specific effect of antibiotics on core pathways.

The difference in core pathways seen at 1 month of age could either taxonomic differences or rather, the identified functions may have actively affected their hosts' fitness and colonization potential in the dynamic niche that is the gut in the first month of life. *Enterococcus faecalis*, *Klebsiella michiganensis*, and *Klebsiella pneumoniae* (Table S10), all enriched in CS-born children, are facultative anaerobes. Their overrepresentation may reflect a more oxygenated gut environment[28]. *Enterococcus faecalis* is able to respire aerobically in the presence of hemin[35], and hemin is the oxidized form of heme, which occurs when heme is exposed to oxygen or other oxidizing agents. *Enterococcus faecalis* itself does not synthesize heme, but we observed higher heme b biosynthesis II (oxygen-independent) capacity in CS-born children. *Klebsiella pneumoniae* and *Klebsiella michiganensis* are members of *Klebsiella oxytoca* species complex[36], of which some strains, particularly those with a type VI secretion system (T6SS), can generate reactive oxygen species (ROS) as a weapon against competitor bacteria[37,38]. Phospholipase, a T6SS effector[39], was enriched in the CS pathway profile at 1 month of age. Phospholipase was contributed to by *Clostridium perfringens* as well. As a Gram-positive anaerobic bacterium that can form spores that are crucial during transmission[40], it could survive through aerobic environments such as on surfaces in hospital wards[41], which is probably where this spore former passes on to in-hospital neonates. It has been linked to pre-term necrotising enterocolitis due to its ability to generate a complement of extracellular toxins and hydrolytic enzymes[42,43]. In contrast, vaginal delivery is associated with more anaerobes, especially from *Parabacteroides*, *Bacteroides*, and *Phocaeicola*, which was also reflected by the enriched pathways of lipopolysaccharide (LPS) biosynthesis. Most vaginal delivery-associated bacteria are gram-negative, of which the outer membrane is made of LPS[44]. In our study, the enriched pathways of core LPS compartments reflected a microbial composition rich in gram-negative bacteria. The beta-(1,4)-mannan degradation pathway was upregulated in vaginal-associated core pathways as reported in other studies[45]. It was mainly contributed to by *Clostridium,* which needs to be interpreted with caution. This pathway was reported to be increased in infants with a Bacteroidaceae-dominated gut microbiome[45]. The genome of *Bacteroides fragilis* was found to contain an operon encoding putative mannan metabolizing enzymes[46], and this operon is conserved in the genomes of other *Bacteroides* species, as well as species of *Parabacteroides*[47].

Although 1-month CS microbial scores were not associated with asthma risk, the CS microbial composition possibly shaped the gut microenvironment favoring a skewed microbiome compared to vaginal-associated micorbial composition. It's important to mention that *Enterococcus faecalis* and *Clostridium perfringens* existed in vaginally born infants as well, just at lower abundances than in CS born infants. However, *Enterococcus faecalis* may respire anaerobically and the sporulation of *Clostridium perfringens* can be inhibited by *Bacteroides* and short-chain fatty acids (SCFA)[48]. In vaginally born infants, cross-feeding interactions and SCFA production could promote a hypoxic and acidic environment beneficial for a stable transition from facultative anaerobes to obligate anaerobes[49], which was reduced or absent in the CS-associated gut microbiomes[50]. Such differences in the gut environment may contribute to a long-lasting divergence in the gut microbial composition up to 1 year of age. Consistent with this, we observed sustained depletion of *Phocaeicola dorei*, *Bacteroides theraiotaomicron*, and *Bacteroides ovatus* in the CS born infants all the way from 1 month to 1 year of age, while *Citrobacter braakii*, *Veillonella parvula*, and *Ruminococcus gnavus* were found to be more abundant. This compositional pattern was then linked to asthma risk up to 10 years of age, indicating the importance of taxanomic differences. In light of the largely converged core pathway profile, it's reasonable to hypothesize that key CS-associated taxa contribute through species-specific traits rather than through differences in shared core pathways.

The risk of asthma has previously been associated with a lower gut microbiome maturity at one-year of age[51–53]. Here, a lower gut maturity was in turn driven mostly by lower abundances of key maturity-associated bacterial taxa, while only few taxa were associated with decreased maturity. The potential mechanism seems to be different, when evaluating the risk of asthma after CS-delivery, where the asthma associations are mainly observed with higher abundances of specific taxa at age one-year, enriched after CS-delivery. As such CS delivery seems to promote long-term persistence of asthma-promoting taxa, while the maturation hypothesis suggests that most taxa in the infant gut promote protective properties, when introduced at an adequate timing.

Citrobacter braakii was the largest contributor to the 1-year CS microbial score. It is an anaerobic, gram-negative bacterium, and has been isolated from the environment, food, and humans[54,55]. It has been reported to have proinflammatory and cytotoxic roles in gastric epithelial cells, and has been associated with acute mucosal inflammation in the intestine, respiratory tract, and urinary tract[55]. Surprisingly, the secretion system T6SS was also described in this species[55] like *Klebsiella* mentioned above. A positive relationship between a group of bacteria including *Citrobacter braakii* and ROS overgeneration was also shown *in vitro*[56], suggesting that this bacterium may be linked to oxidative processes. Several early-life taxa can produce ROS, such as *Entercoccus faecalis*[57,58], *Clostridium perfringens*[59] and *Klebsiella*[37]. This capability to produce ROS may confer a competitive advantage in oxidative environments and has been linked to increased virulence toward the host[58,60,61]potentially inducing local gut inflammation. This interpretation is consistent with the reported enriched genes associated with ROS metabolism in infants born by C-section[28]. However, in our analyses, we didn't identify specific ROS metabolism pathways, so we can only speculate on specific taxa driving this in our dataset. Nevertheless, we identified pathways linked to the CS-microbial score at 1 year, which included superpathway of pyrimidine deoxyribonucleosides degradation, 5-oxo-L-proline metabolism, and heme b biosynthesis II (oxygen-independent). Several additional taxa contributed to 1-year CS microbial score included *Intestinimonas butyriciproducens*, *Enterococcus avium*, and *Phocaeicola coprophilus*. Yet, there are no studies on these bacteria linking to asthma risk. We found *Ruminococcus gnavus* and *Veillonella parvula* more abundant with a higher CS microbial score. Higher *Veillonella parvula* abundance at 6 weeks among CS-born infants has been associated with respiratory infection symptoms[62]. *Ruminococcus gnavus* has been reported to lead to airway hyper-responsiveness and airway inflammation[63,64]. In the gut, *Ruminococcus gnavus* is a mucus-associated taxon that can degrade mucin through utilization of terminal mucin glycans[65]. *Ruminococcus gnavus*-produced tryptophan metabolites increase anion-dependent proximal colon secretion[66], and can influence inflammation as a ligand to the aryl hydrocarbon receptor[67,68]. Taken together, *Citrobacter braakii*, *Veillonella parvula*, and *Ruminococcus gnavus* in the 1-year CS microbial community could be associated with a proinflammatory role in the local gut, which likely contribute to the elevated risk of asthma in the later life. At last, we looked into several environmental factors and found that exclusive breastfeeding at 1 month and the presence of older siblings were associated with a reduced 1-year CS microbial score, indicating a protective role in restoring a perturbed gut microbiome.

Conclusion

Here, we followed infants through the first year of life to characterize how delivery mode shapes gut microbiome development and functional potential, and how these early-life scores relate to later asthma risk. We observed a persistent CS-associated taxonomic score at 1 year, which was associated with increased asthma risk up to 10 years of age. In contrast, core pathways largely converged between groups at 1 year of age, and were not associated with asthma risk at any time point, suggesting that the specific species present in the gut microbiome and how long the perturbation persists may be more critical for asthma risk than their core functional pathway repertoire.

Methods

The COPSAC$_{2010}$ cohort

Ethics

The study was conducted in accordance with the guiding principles of the Declaration of Helsinki and was approved by the Local Ethics Committee (H-B-2008-093) and the Danish Data Protection Agency (2015-41-3696). Both parents gave oral and written informed consent before enrolment.

Study population

The COPSAC$_{2010}$ cohort is a population-based mother-child cohort of 700 children and their families, recruited in pregnancy and followed prospectively at the COPSAC research unit. The children were followed by COPSAC study physicians and nurses collecting all biosamples, clinical measurements and diagnoses during clinical visits scheduled at 1 week, 1, 3, 6, 12, 18, 24, 30, and 36 months, thereafter yearly until the age of 6 and again at age 8 and 10 years. Details have been described in previous work[69].

Study endpoint

Persistent wheeze/asthma from 1 to 10 years of age: the diagnosis required a minimum of 5 episodes of diary-recorded and physician-verified troublesom lung symptoms lasting more than 3 consecutive days within the preceding 6 months or 4 weeks of consecutive symptoms[70,71].

Fecal sample collection

Fecal samples were collected either at the research clinic or by the parents at home using detailed instructions. Each sample arrived at the laboratory was mixed on arrival with 1 ml of 10% (v/v) glycerol broth (SSI, Copenhagen, Denmark) and frozen at −80°C. DNA was extracted using the PowerMag Soil DNA Isolation Kit (Qiagen) and NucleoSpin Stool Kit (Macherey-Nagel). No differences were found between the different DNA extraction kits after comparison. Before library preparation, the DNA was quantified by Tecan Infinite F Nano+ Plate Reader using Quant-iT dsDNA BR Assay Kit. The enzymatic fragmentation of DNA and library construction was conducted by Tecan DreamPrep NGS using Celero EZ DNA-seq Core Module Kit. The fragmented DNA was amplified using polymerase chain reaction (PCR). Short and large DNA fragments were removed using double-sided magnetic bead size selection (AMPure XP, Beckman Coulter). Adapter sequences from Celero 96-Plex Adaptor Plate were added to each sample during library construction. The final concentration for each library was quantified by Tecan Infinite F Nano+ Plate Reader using NuQuant NGS Library Quantification Module and Qubit. The final fragment distribution is evaluated using a Fragment Analyzer 5200 (Agilent). Qubit and TapeStation were used to determine the concentration of the final library before sequencing. The library was sequenced using 2 × 150 bp paired-end sequencing on an Illumina NovaSeq 6000 platform (Illumina, San Diego, CA, USA). The sequencing depth and batch were adjusted for in our statistical analysis.

Gut microbiota characterization and pre-processing

Quality control of raw FASTQ files was performed using KneadData (v. 0.6.1) to remove low-quality bases and reads derived from the host genome as follows: Using Trimmomatic (v. 0.36), the reads were quality trimmed by removing Nextera adapters, leading and trailing bases with a Phred score below 20, and trailing bases in which the Phred score over a window of size 4 drops below 20. Trimmed reads shorter than 100 bases were discarded as low-quality reads. Reads that mapped to the human reference genome GRCh38 (with Bowtie2 v. 0.2.3.2 using default settings)[72] were also discarded. Read pairs in which both reads passed filtering were retained; these were classified as high-quality non-host (HQNH) reads.

Taxonomic composition of metagenomes was profiled using MetaPhlAn v4.1[73] on quality-controlled reads. MetaPhlAn was run with default parameters using the species-level genomic bin (SGB) markergene database mpa_vJun23_CHOCOPhlAnSGB_202403 (ChocoPhlAn SGB database vJun23), yielding relative abundances of microbial taxa. For descriptive analysis, taxa were retained if their prevalence >5%; For differential abundance analyses, relative abundances were log-transformed internally by the DA.lli2 function after filtering by prevalence >5% and mean relative abundance >0.01%; For sPLS model training and testing, relative abundances were center-log-ratio (CLR) transformed after filtering by prevalence >5% and mean relative abundance >0.01% and applied a pseudocount of min(relative abundance)/2.

Microbial functional profiling was performed using HUMAnN v3.9[19]. HUMAnN's prescreen step (prescreen threshold 0.01) used the MetaPhlAn v4.1 taxonomic profiles to build a sample-specific nucleotide database from the HUMAnN ChocoPhlAn pangenome collection (v201901_v31). Reads were first aligned to this database with Bowtie2 v2.2 using the *--very-sensitive* preset. Reads not mapped at the nucleotide level were searched in translated mode against the UniRef90 protein database using DIAMOND v2.0.15 (*--top 1 --outfmt 6*, e-value ≤ 1.0, minimum amino-acid identity 80%, and minimum subject and query coverage of 50% and 90%, respectively). Gene family abundances were mapped to MetaCyc pathways (MetaCyc v24), and pathway abundances were inferred using MinPath with gap-filing enabled (MinPath on, gap fill on, Xipe off). HUMAnN outputs for gene families and pathways were then normalized to copies per million (CPM) using *humann_renorm_table (units = cpm)* and merged across samples with *humann_join_tables*, so that each feature's abundances reflects its estimated copy number per million total microbial copies in each sample. For descriptive analyses, pathways were retained if their prevalence >5%; For differential abundance analyses, relative abundances were log-transformed internally by the DA.lli2 function after filtering by prevalence >5% ; For sPLS model training and testing, relative abundances were CLR transformed after filtering by prevalence >5% and applied a pseudocount of min(relative abundance)/2.

Statistical analysis

Wilcoxon test or Fisher's exact test were used for analyzing simple associations in the baseline characteristics of the cohort according to delivery mode. Observed richness was used as the measure of the within-sample diversity of both taxonomic and pathway compositions. The between-sample diversity metrics were computed as bray-curtis distances. Differences in the observed richness among delivery mode and subgroups were tested using linear regression adjusting for sequencing depth and batch, differences in the beta diversity were visualized with principle coordinates analysis (PCoA) plots and tested for inference using permutational multivariate analysis of variance (PERMANOVA; Adonis2 from the package vegan[74] v2.7.1 with 999 permutations), adjusting for sequencing depth and batch using argument by = "margin". Differential abundance analyses were analyzed using DAtest[75] v2.8.0, adjusting for sequencing depth and batch. sPLS modes were trained against delivery mode for each time point on both taxonomic and pathway compositions. We selected the optimum number of input variables using repeated 10-fold cross-validation of the AUC statistic on the training set (train/test splitting: 0.75/0.25), we evaluated the performance by cross-validated AUC and AUC on the testset. The predicted values of these models were combined into CS microbial and pathway scores, and scaled for analysis. Associations between CS microbial scores and asthma risk were tested using general estimation equations (GEE, geeglm function from package geepack[76] v1.3.13) model to compute the overall odds ratios with 95% CI. Covariates adjusted in the model included: exclusive breastfeeding at 1 month, any breastfeeding at 1 year, hospitalization after birth, gestational age, maternal bmi, maternal asthma history, household income at 1 year, passive smoke exposure, maternal age, paternal age, the presence of older siblings, sex, race, birth season, sequencing depth, and batch. Associations between CS microbial scores and environmental factors were tested using linear regression, adjusting for sequencing depth and batch. A significance level of 0.05 was used in all analyses, and controlled with a false discovery rate of 5% for multiple testing[77]. All data analyses were performed with the statistical software package R v4.5.0, with the package phyloseq[78] v1.52.0 to handle the microbiome and pathway abundance data.

Data availablity

COPSAC sequencing data is available in the Sequence Read Archive (SRA) under accession no. PRJNA715601. Individual-level data is protected under Danish and European law that prohibits publication even in pseudonymized form. However, data can be made available to researchers under a data processing agreement by contacting COPSAC's Data Protection Officer (administration@dbac.dk)

References

1. Betrán, A. P. *et al.* The Increasing Trend in Caesarean Section Rates: Global, Regional and National Estimates: 1990-2014. *PLOS ONE* **11**, e0148343 (2016).
2. Betran, A. P., Ye, J., Moller, A.-B., Souza, J. P. & Zhang, J. Trends and projections of caesarean section rates: global and regional estimates. *BMJ Glob Health* **6**, (2021).
3. Chien, P. Global rising rates of caesarean sections. *BJOG: An International Journal of Obstetrics & Gynaecology* **128**, 781–782 (2021).
4. Chua, W. C., Chen, Y.-L., Yen, C.-F. & Chen, H.-L. Long-term health outcomes of children born by cesarean section: A nationwide population-based retrospective cohort study in Taiwan. *Journal of the Formosan Medical Association* **124**, 1034–1038 (2025).
5. Chavarro, J. E. *et al.* Association of Birth by Cesarean Delivery With Obesity and Type 2 Diabetes Among Adult Women. *JAMA Netw Open* **3**, e202605 (2020).
6. Bernardi, J. R. *et al.* Cesarean delivery and metabolic risk factors in young adults: a Brazilian birth cohort study. *Am J Clin Nutr* **102**, 295–301 (2015).
7. Fritz, J. *et al.* Cesarean birth is associated with lower motor and language development scores during early childhood: a longitudinal analysis of two cohorts. *Sci Rep* **14**, 23438 (2024).
8. Zhou, W. & Tang, J. Prevalence and risk factors for childhood asthma: a systematic review and meta-analysis. *BMC Pediatrics* **25**, 50 (2025).
9. Liu, X. *et al.* Risk of Asthma and Allergies in Children Delivered by Cesarean Section: A Comprehensive Systematic Review. *The Journal of Allergy and Clinical Immunology: In Practice* **12**, 2764–2773 (2024).
10. Song, X. *et al.* Gut-lung axis and asthma: A historical review on mechanism and future perspective. *Clin Transl Allergy* **14**, e12356 (2024).
11. Biagioli, V., Volpedo, G., Riva, A., Mainardi, P. & Striano, P. From Birth to Weaning: A Window of Opportunity for Microbiota. *Nutrients* **16**, 272 (2024).
12. de Koff, E. M. *et al.* Mode of delivery modulates the intestinal microbiota and impacts the response to vaccination. *Nat Commun* **13**, 6638 (2022).
13. Lay, C. *et al.* A synbiotic intervention modulates meta-omics signatures of gut redox potential and acidity in elective caesarean born infants. *BMC Microbiol* **21**, 191 (2021).
14. Stokholm, J. *et al.* Delivery mode and gut microbial changes correlate with an increased risk of childhood asthma. *Sci Transl Med* **12**, (2020).
15. Wampach, L. *et al.* Birth mode is associated with earliest strain-conferred gut microbiome functions and immunostimulatory potential. *Nat Commun* **9**, 5091 (2018).
16. Busi, S. B. *et al.* Persistence of birth mode-dependent effects on gut microbiome composition, immune system stimulation and antimicrobial resistance during the first year of life. *ISME Commun* **1**, 8 (2021).
17. Shi, Y.-C. *et al.* Initial meconium microbiome in Chinese neonates delivered naturally or by cesarean section. *Sci Rep* **8**, 3255 (2018).
18. Reznik, S. E. *et al.* The effect of cesarean delivery on the neonatal gut microbiome in an under-resourced population in the Bronx, NY, USA. *BMC Pediatr* **24**, 450 (2024).
19. Beghini, F. *et al.* Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *eLife* **10**, e65088 (2021).

20. Coker, M. O. *et al.* Infant Feeding Alters the Longitudinal Impact of Birth Mode on the Development of the Gut Microbiota in the First Year of Life. *Front. Microbiol.* **12**, (2021).
21. Bäckhed, F. *et al.* Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life. *Cell Host & Microbe* **17**, 690–703 (2015).
22. Marcobal, A. *et al.* Bacteroides in the Infant Gut Consume Milk Oligosaccharides via Mucus-Utilization Pathways. *Cell Host & Microbe* **10**, 507–514 (2011).
23. Ioannou, A. & Belzer, C. Applicability of synthetic communities of the infant gut microbiota to understand microbial dynamics in an interchanging glycan landscape. *Current Opinion in Microbiology* **88**, 102659 (2025).
24. Stewart, C. J. *et al.* Temporal development of the gut microbiome in early childhood from the TEDDY study. *Nature* **562**, 583–588 (2018).
25. Wang, S. *et al.* Metagenomic analysis of mother-infant gut microbiome reveals global distinct and shared microbial signatures. *Gut Microbes* **13**, 1911571 (2021).
26. Zhang, C. *et al.* The Effects of Delivery Mode on the Gut Microbiota and Health: State of Art. *Front Microbiol* **12**, 724449 (2021).
27. Flores-Lopez, R., Maasakkers, C. M., Lambers, T. T., Nauta, A. & Morán-Ramos, S. Cesarean section and socioeconomic status are associated with Bifidobacterium abundance and intestinal physiology: observations from a cross-sectional study in formula-fed infants. *BMC Microbiol* **25**, 780 (2025).
28. Lay, C. *et al.* A synbiotic intervention modulates meta-omics signatures of gut redox potential and acidity in elective caesarean born infants. *BMC Microbiol* **21**, 191 (2021).
29. Mitchell, C. M. *et al.* Delivery Mode Affects Stability of Early Infant Gut Microbiota. *Cell Rep Med* **1**, 100156 (2020).
30. Heintz-Buschart, A. & Wilmes, P. Human Gut Microbiome: Function Matters. *Trends in Microbiology* **26**, 563–574 (2018).
31. Matharu, D. *et al.* Bacteroides abundance drives birth mode dependent infant gut microbiota developmental trajectories. *Front. Microbiol.* **13**, (2022).
32. Busi, S. B. *et al.* Persistence of birth mode-dependent effects on gut microbiome composition, immune system stimulation and antimicrobial resistance during the first year of life. *ISME COMMUN.* **1**, 8 (2021).
33. Allison, S. D. & Martiny, J. B. H. Resistance, resilience, and redundancy in microbial communities. *Proceedings of the National Academy of Sciences* **105**, 11512–11519 (2008).
34. Guittar, J., Shade, A. & Litchman, E. Trait-based community assembly and succession of the infant gut microbiome. *Nat Commun* **10**, 512 (2019).
35. Winstedt, L., Frankenberg, L., Hederstedt, L. & von Wachenfeldt, C. Enterococcus faecalis V583 Contains a Cytochrome bd-Type Respiratory Oxidase. *Journal of Bacteriology* **182**, 3863–3866 (2000).
36. McCartney, A. L. & Hoyles, L. Host interactions of bioactive molecules produced by Klebsiella spp. *Microbiota and Host* **3**, (2025).
37. Storey, D. *et al.* Klebsiella pneumoniae type VI secretion system-mediated microbial competition is PhoPQ controlled and reactive oxygen species dependent. *PLOS Pathogens* **16**, e1007969 (2020).

38. Wong Fok Lung, T. *et al.* *Klebsiella pneumoniae* induces host metabolic stress that promotes tolerance to pulmonary infection. *Cell Metabolism* **34**, 761-774.e9 (2022).
39. de Carvalho, M. de M. L. *et al.* Dual Role for Pld1 in Klebsiella pneumoniae Virulence: Transcriptomics and Proteomics Provide Insights into Direct and Indirect Effects. *J Proteome Res* **24**, 2874–2884.
40. Camargo, A., Ramírez, J. D., Kiu, R., Hall, L. J. & Muñoz, M. Unveiling the pathogenic mechanisms of Clostridium perfringens toxins and virulence factors. *Emerging Microbes & Infections* **13**, 2341968 (2024).
41. Machida, Y. *et al.* [An outbreak of enterocolitis due to Clostridium perfringens in a hospital for the severely disabled]. *Kansenshogaku Zasshi* **63**, 410–416 (1989).
42. Kiu, R. & Hall, L. J. An update on the human and animal enteric pathogen Clostridium perfringens. *Emerg Microbes Infect* **7**, 141 (2018).
43. Kiu, R. *et al.* Particular genomic and virulence traits associated with preterm infant-derived toxigenic Clostridium perfringens strains. *Nat Microbiol* **8**, 1160–1175 (2023).
44. Bertani, B. & Ruiz, N. Function and Biogenesis of Lipopolysaccharides. *EcoSal Plus* **8**, (2018).
45. Leech, S. M. *et al.* Delivery mode is a larger determinant of infant gut microbiome composition at 6 weeks than exposure to peripartum antibiotics. *Microb Genom* **10**, 001269 (2024).
46. Kawaguchi, K. *et al.* The mannobiose-forming exo-mannanase involved in a new mannan catabolic pathway in Bacteroides fragilis. *Arch Microbiol* **196**, 17–23 (2014).
47. Senoura, T. *et al.* New microbial mannan catabolic pathway that involves a novel mannosylglucose phosphorylase. *Biochemical and Biophysical Research Communications* **408**, 701–706 (2011).
48. Wrigley, D. M. Inhibition of *Clostridium perfringens* sporulation by *Bacteroides fragilis* and short-chain fatty acids. *Anaerobe* **10**, 295–300 (2004).
49. Sanidad, K. Z. & Zeng, M. Y. Neonatal gut microbiome and immunity. *Current Opinion in Microbiology* **56**, 30–37 (2020).
50. Bui, T. N. Y., Paul, A., Guleria, S., O'Sullivan, J. M. & Toldi, G. Short-chain fatty acids—a key link between the gut microbiome and T-lymphocytes in neonates? *Pediatr Res* 1–9 (2025) doi:10.1038/s41390-025-04075-0.
51. Stokholm, J. *et al.* Maturation of the gut microbiome and risk of asthma in childhood. *Nat Commun* **9**, 141 (2018).
52. Hoskinson, C. *et al.* Delayed gut microbiota maturation in the first year of life is a hallmark of pediatric allergic disease. *Nat Commun* **14**, 4785 (2023).
53. Lee-Sarwar, K. *et al.* Association of the gut microbiome and metabolome with wheeze frequency in childhood asthma. *J Allergy Clin Immunol* **150**, 325–336 (2022).
54. Pasquali, F. *et al.* Citrobacter braakii Isolated from Salami and Soft Cheese: An Emerging Food Safety Hazard? *Foods* **14**, 1887 (2025).
55. Yu, M. *et al.* Characterization of cytotoxic Citrobacter braakii isolated from human stomach. *FEBS Open Bio* **14**, 487–497 (2024).
56. Lenický, M. *et al.* Identification of Bacterial Profiles and Their Interactions with Selected Quality, Oxidative, and Immunological Parameters of Turkey Semen. *Animals (Basel)* **11**, 1771 (2021).

57. Huycke, M. M. & Moore, D. R. In vivo production of hydroxyl radical by enterococcus faecalis colonizing the intestinal tract using aromatic hydroxylation. *Free Radical Biology and Medicine* **33**, 818–826 (2002).
58. Huycke, M. M., Abrams, V. & Moore, D. R. Enterococcus faecalis produces extracellular superoxide and hydrogen peroxide that damages colonic epithelial cell DNA. *Carcinogenesis* **23**, 529–536 (2002).
59. Monturiol-Gross, L. *et al.* Reactive Oxygen Species and the MEK/ERK Pathway Are Involved in the Toxicity of Clostridium perfringens α-Toxin, a Prototype Bacterial Phospholipase C. *J Infect Dis* **206**, 1218–1226 (2012).
60. Riboulet, E. *et al.* Relationships between Oxidative Stress Response and Virulence in Enterococcus faecalis. *Journal of Molecular Microbiology and Biotechnology* **13**, 140–146 (2007).
61. Portela, C. A. F., Smart, K. F., Tumanov, S., Cook, G. M. & Villas-Bôas, S. G. Global Metabolic Response of Enterococcus faecalis to Oxygen. *J Bacteriol* **196**, 2012–2022 (2014).
62. Garcia-Mauriño, C. *et al.* Investigation of associations between the neonatal gut microbiota and severe viral lower respiratory tract infections in the first 2 years of life: a birth cohort study with metagenomics. *The Lancet Microbe* **6**, (2025).
63. Crost, E. H., Coletto, E., Bell, A. & Juge, N. Ruminococcus gnavus: friend or foe for human health. *FEMS Microbiol Rev* **47**, fuad014 (2023).
64. Chua, H.-H. *et al.* Intestinal Dysbiosis Featuring Abundance of Ruminococcus gnavus Associates With Allergic Diseases in Infants. *Gastroenterology* **154**, 154–167 (2018).
65. Crost, E. H. *et al.* The mucin-degradation strategy of Ruminococcus gnavus: The importance of intramolecular trans-sialidases. *Gut Microbes* **7**, 302–312 (2016).
66. Bhattarai, Y. *et al.* Gut Microbiota-Produced Tryptamine Activates an Epithelial G-Protein-Coupled Receptor to Increase Colonic Secretion. *Cell Host & Microbe* **23**, 775-785.e5 (2018).
67. Cheng, Y. *et al.* Aryl Hydrocarbon Receptor Activity of Tryptophan Metabolites in Young Adult Mouse Colonocytes. *Drug Metabolism and Disposition* **43**, 1536–1543 (2015).
68. Zelante, T. *et al.* Tryptophan catabolites from microbiota engage aryl hydrocarbon receptor and balance mucosal reactivity via interleukin-22. *Immunity* **39**, 372–385 (2013).
69. Bisgaard, H. The Copenhagen Prospective Study on Asthma in Childhood (COPSAC): design, rationale, and baseline data from a longitudinal birth cohort study. *Ann Allergy Asthma Immunol* **93**, 381–389 (2004).
70. Bisgaard, H. *et al.* Fish Oil-Derived Fatty Acids in Pregnancy and Wheeze and Asthma in Offspring. *N Engl J Med* **375**, 2530–2539 (2016).
71. Chawes, B. L. *et al.* Effect of Vitamin D_3 Supplementation During Pregnancy on Risk of Persistent Wheeze in the Offspring: A Randomized Clinical Trial. *JAMA* **315**, 353 (2016).
72. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357–359 (2012).
73. Blanco-Míguez, A. *et al.* Extending and improving metagenomic taxonomic profiling with uncharacterized species using MetaPhlAn 4. *Nat Biotechnol* **41**, 1633–1644 (2023).
74. Oksanen, J. *et al.* vegan: Community Ecology Package. 2.7-2 https://doi.org/10.32614/CRAN.package.vegan (2001).

75. Russel, J. *et al.* DAtest: a framework for choosing differential abundance or expression method. 241802 Preprint at https://doi.org/10.1101/241802 (2018).
76. Højsgaard, S., Halekoh, U. & Yan, J. The R Package geepack for Generalized Estimating Equations. *Journal of Statistical Software* **15**, 1–11 (2006).
77. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**, 289–300 (1995).
78. McMurdie, P. J. & Holmes, S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE* **8**, e61217 (2013).

Supplementary tables and figures

Table S1 Comparison of beta diversity of bacterial species and microbial pathways in relation to delivery mode and subgroups.

Group	Timepoint	Sample size	**Gut microbiome**			**Microbial pathways**		
			R2	F statistic	P value	R2	F statistic	P value
CS vs Vaginal	1m	600	0.008	4.918	**0.001**	0.009	5.519	**0.003**
	1y	660	0.003	2.117	**0.006**	0.001	0.975	0.362
CS vs Vaginal without antibiotics	**1m**	**532**	0.01	5.293	**0.001**	0.012	6.535	**0.002**
	1y	**589**	0.004	2.31	**0.002**	0.002	1.086	0.323
CS vs Vaginal with antibiotics	1m	201	0.008	1.61	**0.039**	0.002	0.477	0.79
	1y	211	0.006	1.259	0.169	0.003	0.632	0.66
Vaginal with antibiotics vs Vaginal without antibiotics	1m	463	0.003	1.31	0.154	0.007	3.122	**0.024**
	1y	514	0.003	1.692	**0.028**	0.002	0.875	0.444
CS – acute vs Vaginal	1m	542	0.006	3.472	**0.001**	0.009	5.028	**0.002**
	1y	599	0.004	2.248	**0.002**	0.003	1.781	0.108
CS – planned vs Vaginal	1m	523	0.006	3.019	**0.001**	0.004	2.259	0.052
	1y	578	0.002	0.871	0.617	0.001	0.448	0.879
CS – acute vs CS - planned	1m	135	0.009	1.266	0.181	0.014	1.864	0.11
	1y	143	0.005	0.754	0.812	0.007	1.053	0.313

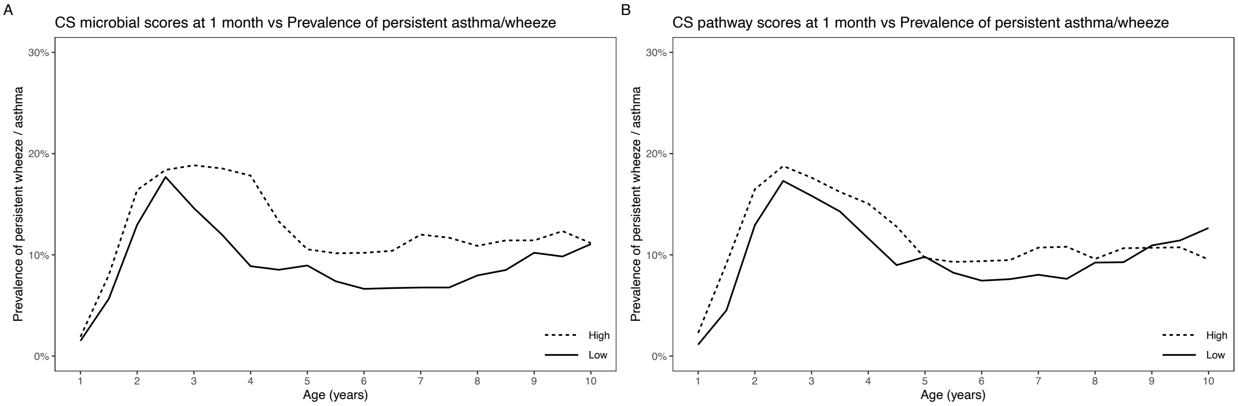


Fig.S1 The distribution of (A) CS microbial scores at 1 month and (B) CS pathway scores and the prevalence of recurrent wheeze, persistent wheeze/asthma up to age 10.

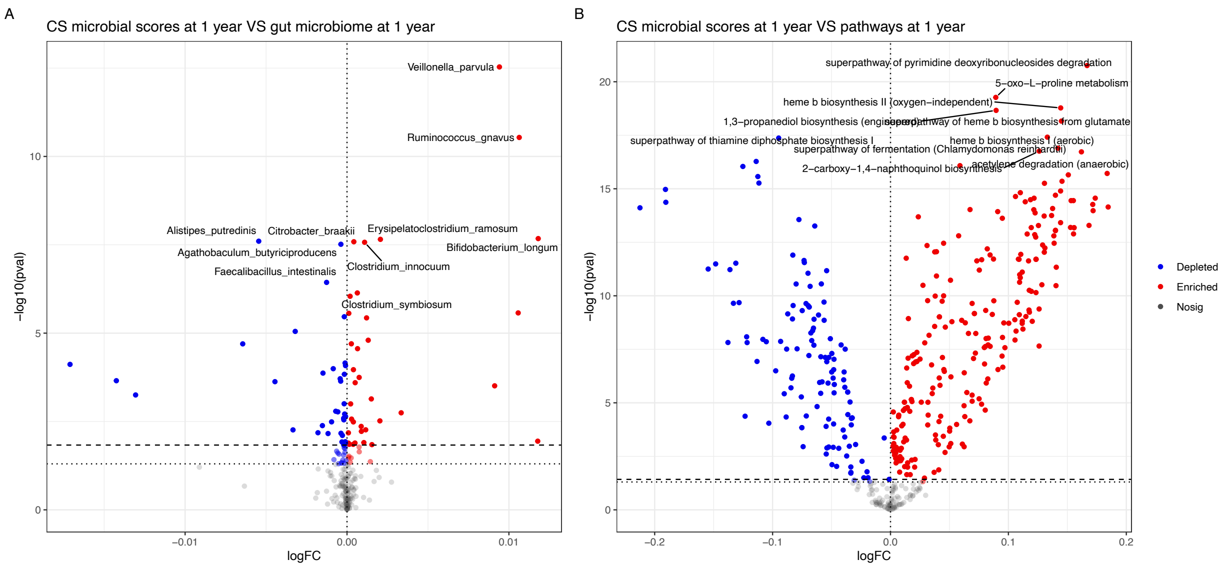


Fig.S2 Differential abundance analysis on the species and pathways according to CS microbial scores at 1 year. Species with prevalence of at least 5% and relative abundance of more than 0.01% of the total were eligible. Pathways with prevalence of at least 5% were eligible. All analyses were adjusted for sequencing depth and batch. Red dots represent species/pathways positively assoicated with CS microbial scores, blue dots represent those negatively associated with CS microbial scores. The dotted line represents nominal P<0.05, and the dashed line represents FDR-adjusted P<0.05.

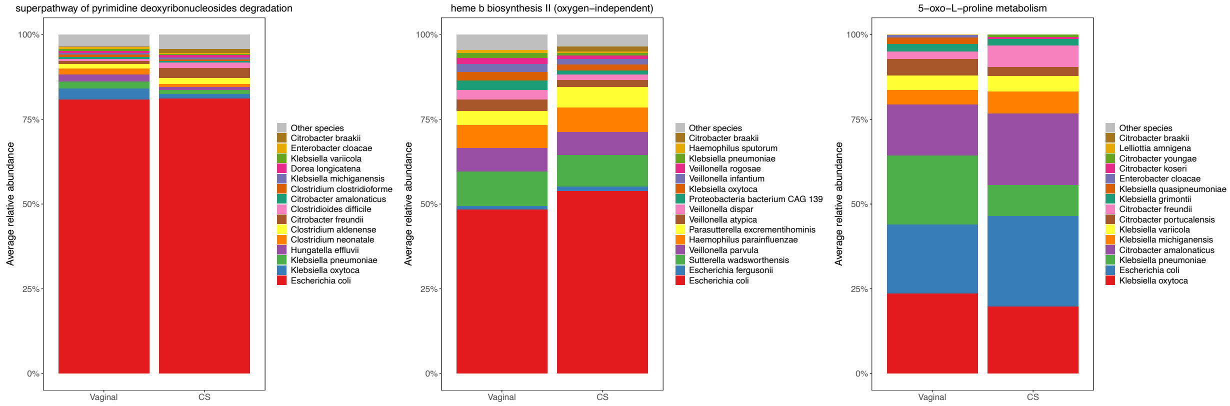


Fig.S3 Distributions of bacterial species to pathways stratified by delivery mode.

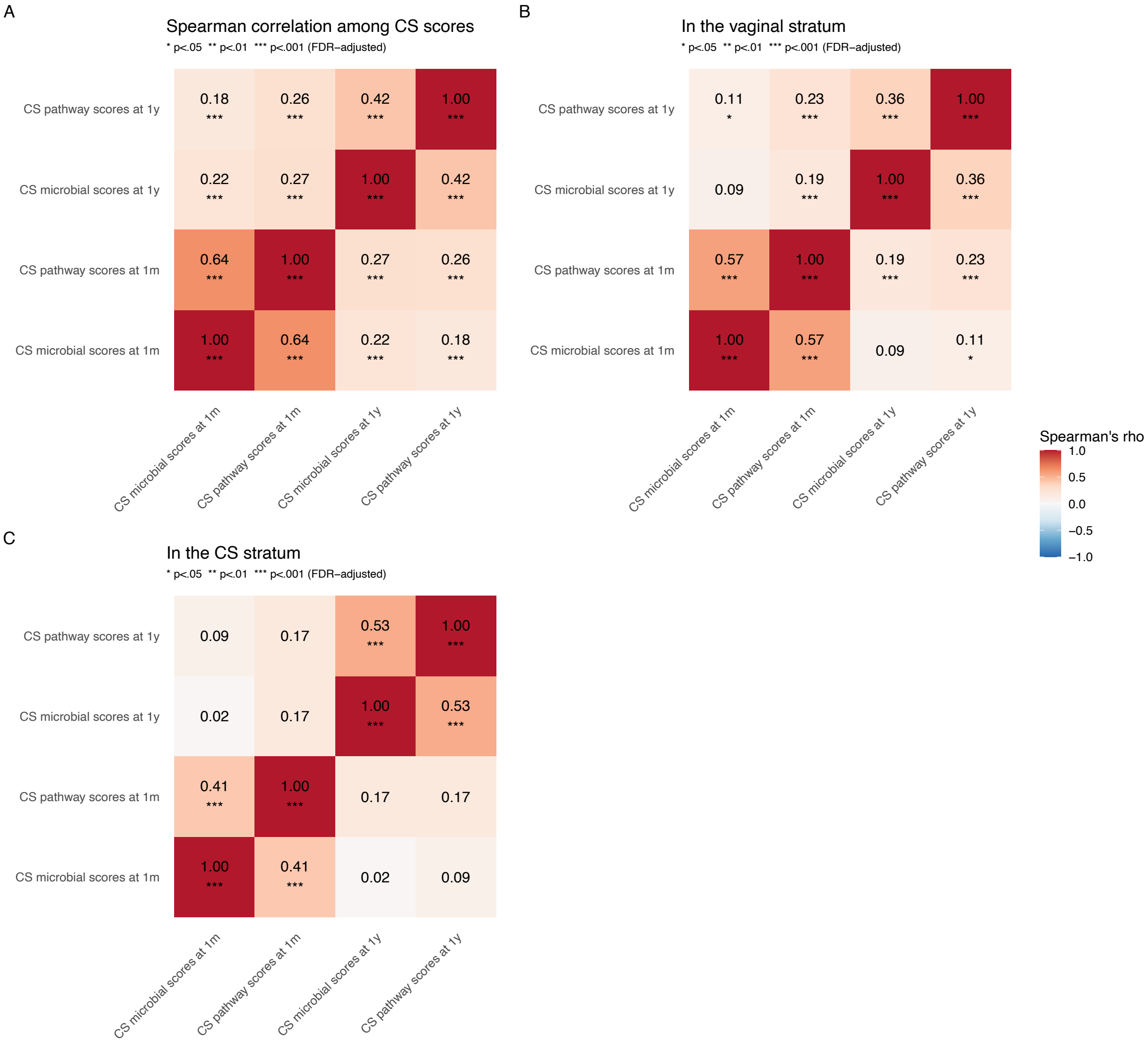


Fig.S4 Spearman correlations among CS scores in A) the full dataset, B) in the vaginal stratum, and C) in the CS stratum.

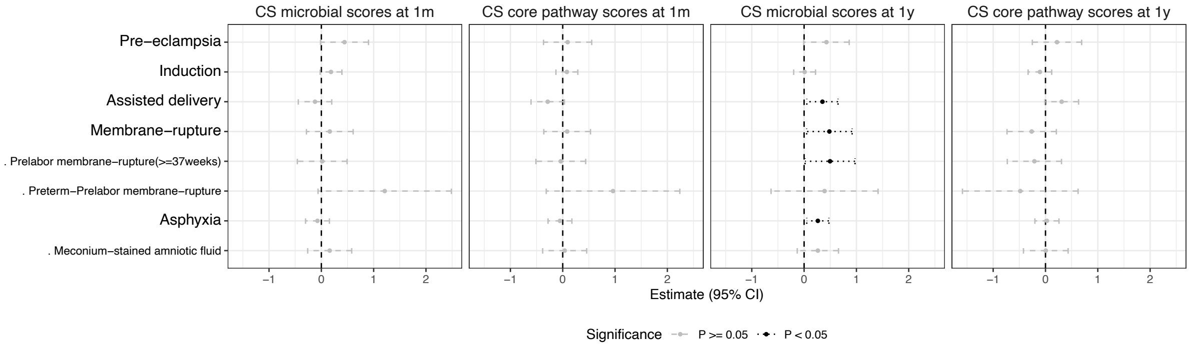


Fig.S5 CS scores associated with perinatal complications using linear regression in the vaginal stratum.

Appendix - Paper III: Microbially-derived tryptophan metabolites in the infant gut mediate the increased risk of asthma following cesarean section

Jie Jiang[1,2], Darlene L.Y. Dai[3], Casper Sahl Poulsen[1], Kaare Tranæs[1], Tingting Wang[1], Charisse Petersen[3], Courtney Hoskinson[3,4], Theo J. Moraes[5], Piushkumar J. Mandhane[6,7], Elinor Simons[8], Meghan B. Azad[8,9], Padmaja Subbarao[5,10,11], Bo Chawes[1,12], Klaus Bønnelykke[1,12], Jonathan Thorsen*[1,12], Stuart E Turvey*[3], Jakob Stokholm*[1,2]

[1]COPSAC, Copenhagen Prospective Studies on Asthma in Childhood, Copenhagen University Hospital - Herlev and Gentofte, Copenhagen, Denmark
[2]Department of Food Science, Faculty of Science, University of Copenhagen, Frederiksberg C, Denmark
[3]Department of Pediatrics, BC Children's Hospital, University of British Columbia; Vancouver, BC, Canada.
[4]Department of Microbiology and Immunology, University of British Columbia; Vancouver, BC, Canada.
[5]Department of Pediatrics, The Hospital for Sick Children, Toronto, Canada
[6]Department of Pediatrics, University of Alberta, Edmonton, Canada
[7]Department of Medicine, Faculty of Medicine and Health Sciences, UCSI University, Kuala Lumpur, Malaysia
[8]Section of Allergy and Immunology, Department of Pediatrics and Child Health, University of Manitoba, Winnipeg, MB, Canada
[9]Manitoba Interdisciplinary Lactation Centre (MILC), Children's Hospital Research Institute of Manitoba, Winnipeg, MB, Canada
[10]Department of Medicine, McMaster University, Hamilton, ON, Canada
[11]Dalla Lana School of Public Health, University of Toronto, Toronto, Canada
[12]Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

*Joint senior authors

Corresponding authors:

Professor Stuart E. Turvey
E-mail: sturvey@bcchr.ca

Professor Jakob Stokholm
E-mail: stokholm@copsac.com

Abstract

Delivery by cesarean section (CS) is a risk factor for childhood asthma and a strong determinant of the early-life gut microbiome. It has been reported that a CS-perturbed gut microbiome plays a mediating role in the association between CS and asthma risk. However, deeper insights into the potential underlying mechanisms are unclear. Here, using deeply phenotyped participants in the Canadian CHILD birth cohort, we characterized the CS microbial composition at 3 months and 1 year of age using machine learning models, and described the perturbation using CS scores. The association between CS and the gut microbiota composition was more pronounced at 3 months (area under the curve (AUC)=0.78) and attenuated at age 1 year (AUC=0.60). An increased asthma risk at age 5 years was associated only with 1-year CS microbial scores (Odds ratio 1.41, P<0.01; adjusted odds ratio 1.30, P=0.03), consistent with our previous work. Extending this, we identified metabolic imbalances associated with 1-year CS microbial scores, marked by elevated tryptophan metabolites in the stool metabolome. These results replicate our previous findings about the mediating role of the 1-year CS microbial signature in the association between CS and asthma risk. Here, we provide new insight into the underlying mechanism of host-microbe interaction following cesarean section.

Keywords:
Cesarean Section(CS); gut microbiome; stool metabolome; perturbation; CHILD Cohort Study

Introduction

When medically justified, cesarean section (CS) can effectively prevent maternal and perinatal mortality and morbidity[1]. However, when the rate of CS goes above 10%, there's no evidence that mortality rates improve[2]. Though efforts have been made to ensure CS is used only when indicated[3], CS rate has reached 43.6% in Southern Europe[4]. CS delivery has been associated with both short- and long-term risks to the child[5], including an altered immune development and a higher risk of asthma and allergy[6].

Asthma is a global health problem, and its prevalence has been increasing, especially among children[7]. A recent meta-analysis of thirty-five cohort studies[8] concluded that CS was a risk factor for asthma in children from Europe, North America, and Oceania (Odds Ratio (OR) = 1.18, P<0.001), but not in the Asian population, which indicates that more longitudinal studies are needed to clarify the relationship between CS and childhood asthma across different regions. This higher risk can extend into late childhood and adolescence up to age 14[9].

The etiology of asthma has not been fully determined, but disruption of the gut microbiome has been studied as a possible explanation for the increased incidence of asthma across many countries[10]. In a Danish birth cohort, we have previously evaluated the gut microbial perturbations after CS delivery[11]. Here, we found that only children who retained a CS-associated gut microbiome all the way to 1 year of age, were at increased asthma risk at 6 years of age. This finding suggested that the elevated risk is likely mediated through persistent CS-associated disruptions to the gut microbiome in early life.

One of the key mechanisms by which the gut microbiota impacts human health is through the production of bioactive metabolites. Studies have shown altered blood metabolomes associated with a higher risk of asthma[12–15]. Specifically, it has been reported that early CS-associated gut microbial colonisation patterns are reflected in the blood metabolome, particularly tryptophan, bile acid and phenylalanine metabolites, which were associated with asthma risk[16]. Recent studies have also shown that fecal metabolites such as butyrate[17], alanine[18], and acetate[19] were altered in children, who later developed asthma[20]. However, it is unknown whether the long-term microbiome profile associated with CS delivery also affects the gut metabolome.

Here, we describe CS-associated microbial compositions at 3 months and 1 year of age in the Canadian Healthy Infant Longitudinal Development (CHILD) cohort. We replicate and expand upon the impact of CS on the composition of the gut microbiome during the first year of life, and further discover metabolomic profiles associated with a CS-perturbed microbiome. Finally, we associate the CS microbial scores with asthma at age 5. This study enhanced our understanding of how the gut microbiome is involved in the increased risk of asthma following CS.

Results

Characteristics of the CHILD study

The CHILD study is a prospective, longitudinal birth cohort of 3,263 children followed from birth (Supp. Table 1)[21]. Of the 3,089 children with complete delivery mode information 78% (2,412) were born by vaginal delivery, and 22% (677) were born by CS delivery. Of all the children with asthma assessment, 6.2% (156, Supp. Table 2) were diagnosed with asthma at 5 years of age. To identify the CS-assoicated microbial composition at 3 months and 1 year of age, we built sparse partial least squares (sPLS) models on a narrow age window (Fig. 1, Supp. Table 3), excluding those whose mother received antibiotics at birth, which resulted in 1,537 samples (3 months, N=739; 1 year, N=798) from 970 (31%) of the CHILD cohort children. The CS scores (predictions from the sPLS models) were obtained by applying models to the large age window, including 2,711 samples (3 months, N=1,355; 1 year, N=1,356) from 1,301(42%) of the CHILD cohort children (Fig.1).

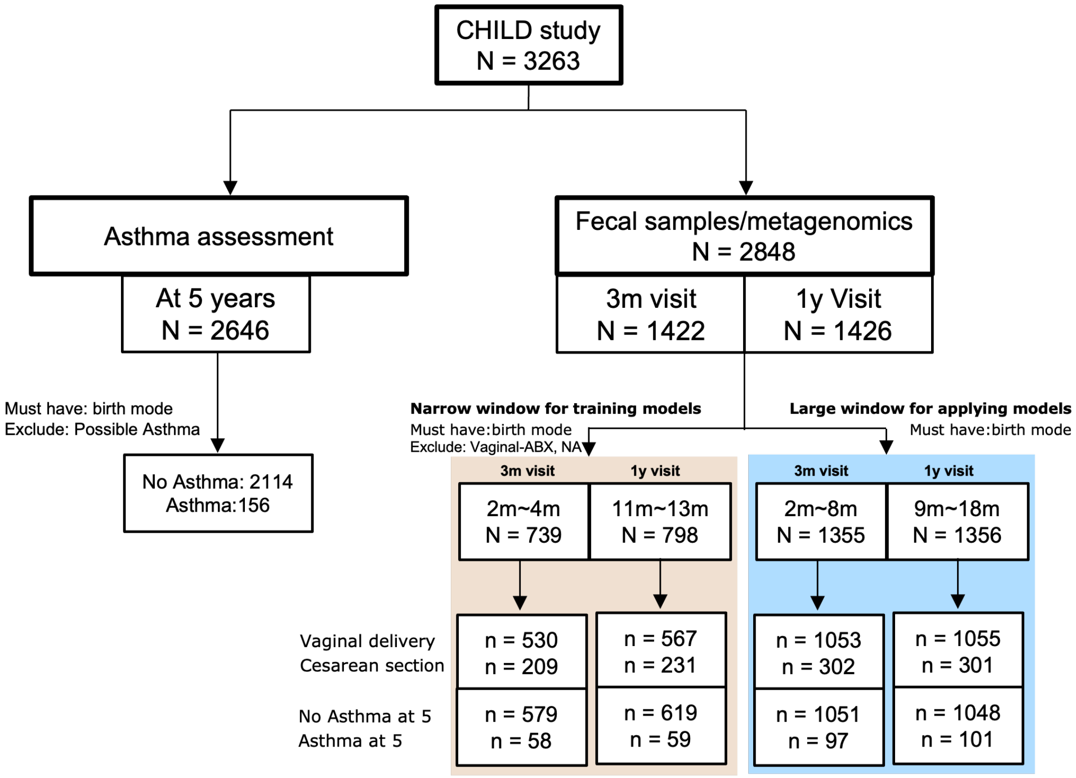


Fig. 1 Flow diagram of sample selection.

Associations between CS and the risk of childhood asthma

CS has been associated with asthma risk at 3 years of age in previous work[22], here, we further investigated the association between CS and the risk of asthma at 5 years of age in the entire cohort and in the subgroups (Fig.2, Supp. Table 4). CS was associated with a higher risk of asthma at 5 years of age (Odds Ratio (OR) 1.66, 95% Confidence Interval (CI) [1.15,2.37], P=0.006). We then stratified the vaginal delivery group by maternal antibiotic exposure at birth to disentangle CS and antibiotic exposure, which is common operational procedure. CS delivery was only significantly associated with asthma when compared with vaginal deliveries without antibiotics (OR 1.76 [1.20,2.56], P=0.003) as compared to the vaginal group who received antibiotics, however there was no significant difference between the two vaginal strata. We also performed a subgroup analysis to test the association of the type of CS, as the vertical transfer of microbes has been suggested to be different and with different influences on the risk of asthma[23]. Compared to vaginal delivery, CS with labor and CS without labor showed comparable risks of asthma (Fig.2).

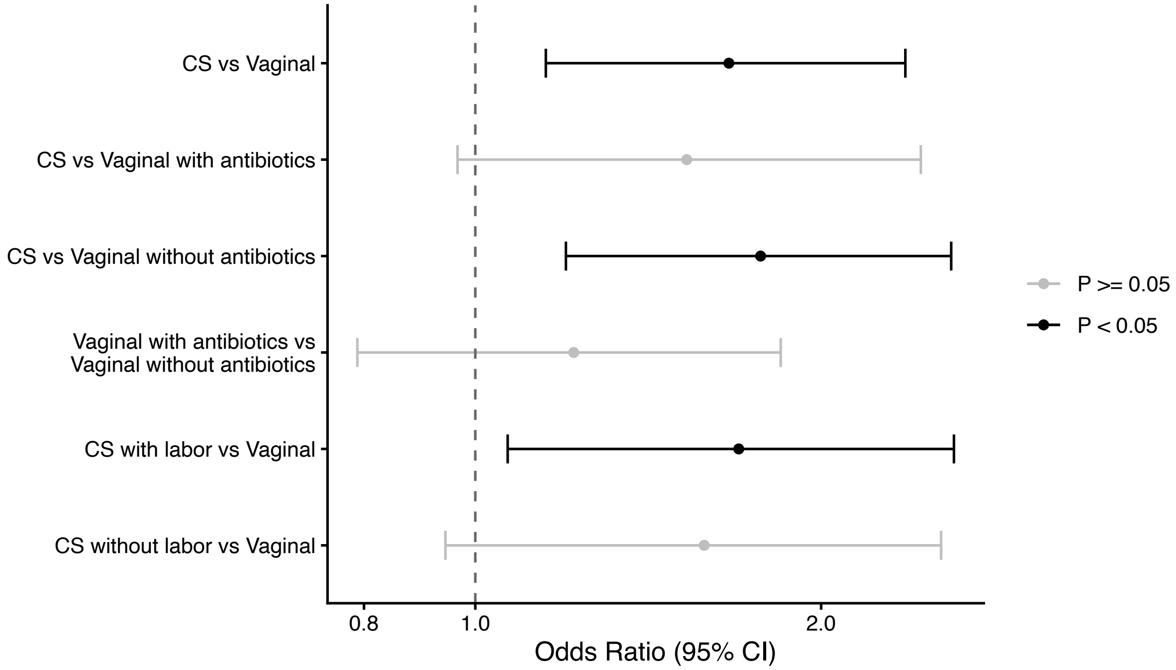


Fig.2 Forest plot of logistic regression models evaluating the associations of CS with asthma at 5 years. The analyses were performed in the crude group: CS vs Vaginal delivery (reference group), and in subgroups: CS vs Vaginal with antibiotics / Vaginal without antibiotics (reference group), CS with labor / CS without labor vs Vaginal (reference group).

Diversity and composition of the gut microbiome in relation to birth mode at early time points

At 3 months of age, children born by CS had a lower Shannon diversity compared to those born vaginally, but the observed richness was not significantly different (Supp. Fig 1). In the subgroup analysis, Vaginal without antibiotics showed higher Shannon diversity than CS-born infants, whereas Vaginal with antibiotics showed Shannon diversity similar to that of CS-born infants. At 1 year of age the Vaginal without antibiotics group continued to have a higher diversity than CS-born infants, while the Vaginal with antibiotics group showed no difference from the CS-born infants, suggesting that the antibiotic exposure to the mother at birth affects the development of infants' gut microbiome during the first year of life. CS without labor had higher Shannon diversity than CS with labor at 3 months of age. At 1 year of age, we found no difference in the observed richness and Shannon diversity between birth modalities (Supp. Fig 1).

Destinct compositional (Bray-Curtis distance) differences were observed between CS and vaginally born infants at 3 months, but not at 1 year of age (Table 1). In contrast to the alpha diversity metrics, differences were also observed between vaginally born infants with and without antibiotic exposure at both 3 months and 1 year of age. . Similar to the alpha diversity, vaginally born infant with antibiotics exposure were compositionally similar to CS-born infants, suggesting that intrapartum antibiotics may lead to a more "CS-like" microbial composition at 1 year of age. There was no detectable difference between CS with labor or without labor at 3 months or at 1 year of age.

Table 1. Birth mode effects on beta diversity calculated based on Bray-curtis distances

Group	Visit	R2	F statistic	P value	Sample size
CS vs Vaginal	3 month	0.006	5.685	**0.001**	956
	1 year	0.002	2.16	**0.005**	1029
CS vs Vaginal with antibiotics	3 month	0.007	2.835	**0.004**	425
	1 year	0.002	0.884	0.608	461
CS vs Vaginal without antibiotics	3 month	0.009	6.582	**0.001**	737
	1 year	0.003	2.771	**0.001**	794
Vaginal with antibiotics vs Vaginal without antibiotics	3 month	0.004	3.26	**0.001**	746
	1 year	0.002	1.986	**0.006**	795
CS with labor vs Vaginal	3 month	0.005	4.265	**0.001**	861
	1 year	0.002	2.266	**0.006**	923
CS without labor vs Vaginal	3 month	0.004	3.691	**0.001**	842
	1 year	0.001	1.333	0.126	904
CS with labor vs CS without labor	3 month	0.008	1.622	0.063	209
	1 year	0.006	1.458	0.066	231

Relative abundances of species associated with CS delivery

We first profiled the distribution of the gut microbiome at 3 months and 1 year of age in the narrow age window (3 months, N=739; 1 year, N=798, Supp. Fig 2). At 3 months of age, the most prevalent species was *Escherichia coli* (prevalence=0.68), followed by *Bifidobacterium longum*, *Streptococcus salivarius*, and *Enterococcus faecalis*. At 1 year of age, *Ruminococcus gnavus* (prevalence = 0.90) was the most prevalent species, followed by *Bifidobacterium longum*, *Eggerthella lenta*, and *Blautia wexlerae*. Among species with prevalence >5% and mean relative abundance >0.01% at both time points, 22 (14%) were unique to the 3-month time point, 83 (55%) were unique to the 1-year time point, and 47 (31%) were shared between time points.

To further evaluate the temporal microbial differences after CS delivery, differential abundance analyses were performed in the narrow age window, comparing CS with Vaginal without antibiotics. At 3 months of age, 20 (29%) of the 69 tested species were significantly different after FDR adjustment and at 1 year 4 (3%) of the 130 tested species were significantly different (Fig. 3). At 3 months of age, especially *Bacteroides fragilis* and *Escherichia coli (*the most prevalent species) were less abundant in the CS delivered children. *Bacteroides fragilis* remained less abundant at 1 year of age, though with less significance, whereas *Escherichia coli abundance* was not different between groups at 1 year of age. *Veillonella parvula* was the species most associated with CS at 3 months of age and at 1 year of age. *Intestinibacter bartlettii* and *Enterococcus faecalis* were significantly more abundant in the CS group at 3 months of age, but not at 1 year of age. *Clostridium perfringens* and *Klebsiella michignensis* were more abundant in the CS group at 3 months, but were not evaluated at 1 year, because of filtering (prevalence >5%, and abundance >0.01%). *Faecalimonas umbilicata* was more abundant in the CS group at 1 year of age, but not evaluated at 3 months of age. At both 3 months and 1 year of age, children delivered by CS had significantly lower abundances of *Bacteroides uniformis*, *Phocaeicola dorei*, *Parabacteroides distasonis*, and *Bacteroides thetaiotaomicron*.

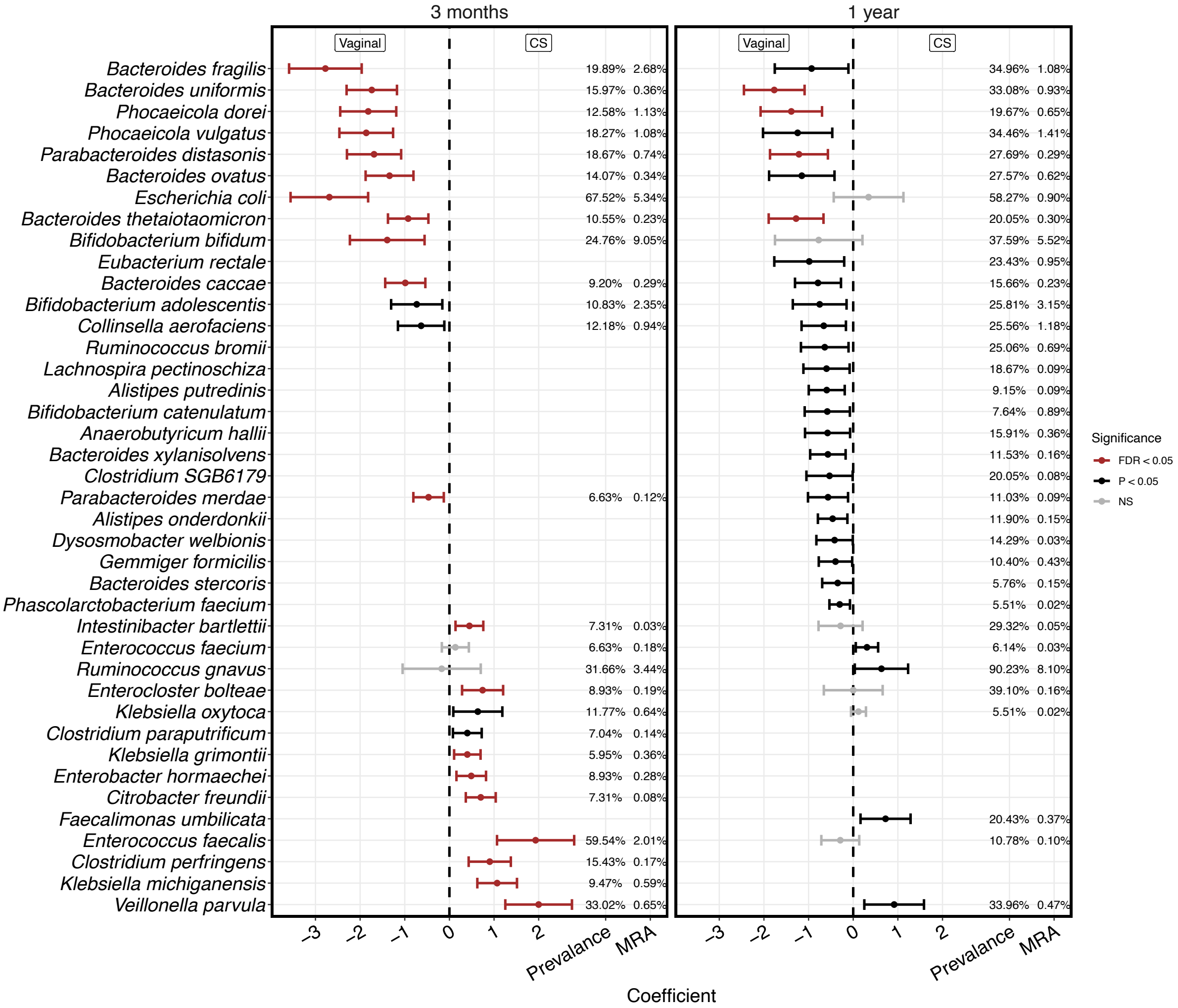


Fig. 3 MaAslin2 forest plot of differential abundances of species in the gut microbiome according to birth mode, adjusting for processing period, exact age, and the sample collection site. Species significantly associated with CS at either 3 months or 1 year of age are shown here. Positive coefficients indicate a higher abundance of bacteria in the CS group than in the vaginal group. The prevalence and Mean Relative Abundance (MRA) of the bacteria across all samples are shown for each included species.

CS microbial signature identified by sPLS models at 3 months and 1 year of age

We constructed cross-validated sparse partial least squares (sPLS) models to identify jointly correlated species associated with CS delivery vs the Vaginal without antibiotics group at 3 months and 1 year of age (3 months, N=739; 1 year, N=798). These models identified the gut microbial composition at each time point most closely associated with CS and determined a CS microbial signature at both 3 months and 1 year. At 3 months of age, the model could predict the delivery mode based solely on the gut microbial composition, but less so at 1 year of age (cross-validated area under the curve (CVAUC) at 3 months: 0.78, Fig. 4A, Supp. Fig 3; 1 year: 0.60, Fig. 4B, Supp. Fig 4). At 3 months of age, a 2-component model captured 35 different species from 4 phyla: Actinobacteria, Bacteroidota, Firmicutes, and Proteobacteria in the CS microbial profile. The first component mainly included species positively correlated with CS, such as *Veillonella parvula*, *Clostridium perfringens*, *and Enterococcus Faecalis* and two species negatively correlated with CS, which were *Escherichia coli* and *Bacteroides fragilis* (Supp. Fig 3). The second component was driven by species negatively correlated with CS, with the biggest contributor being *Bacteroides fragilis*, followed by *Bacteroides uniformis*, *Phocaeicola vulgatus*, *Phocaeicola dorei*, *Parabacteroides distasonis*, and other species with lesser contributions.

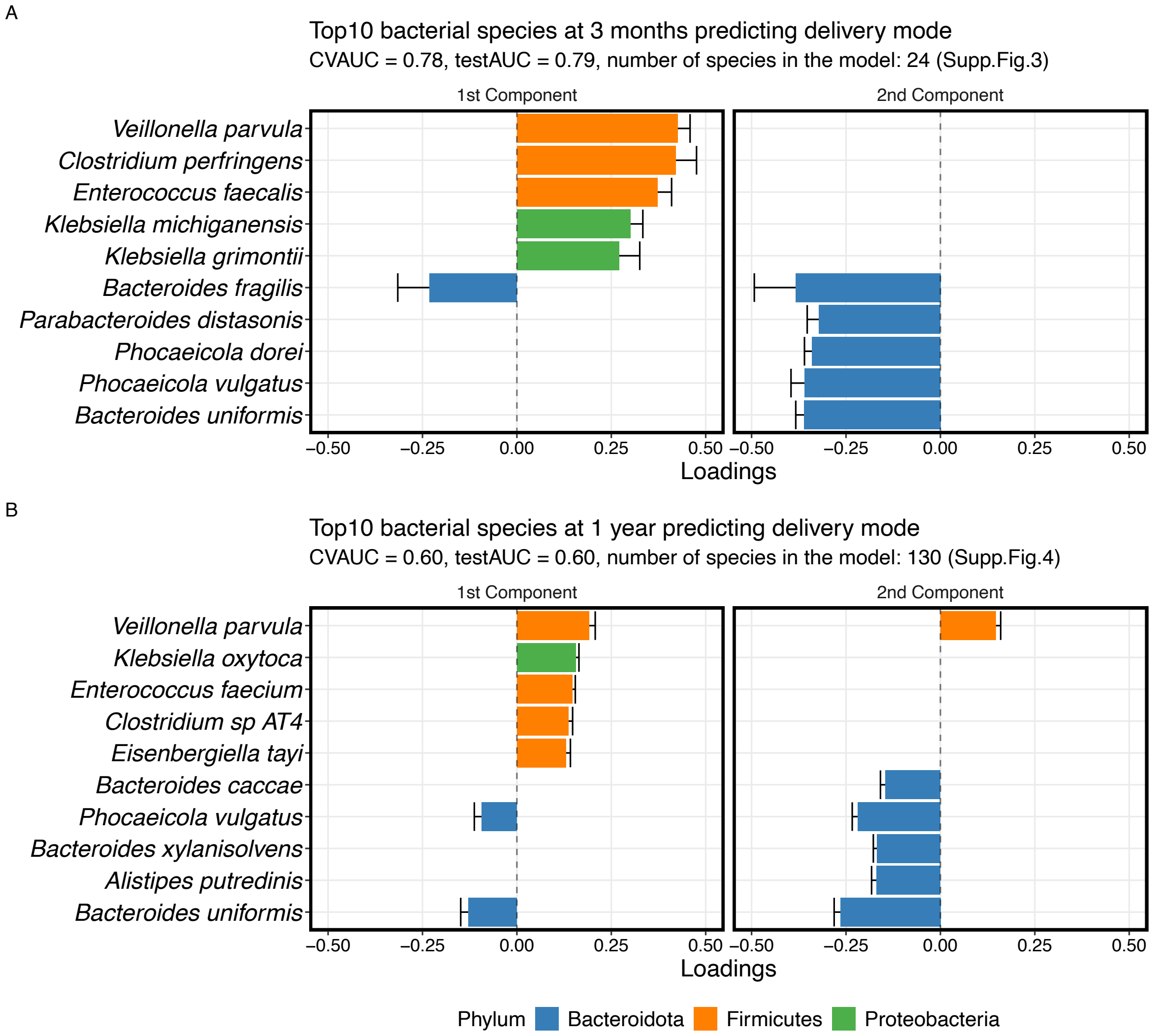


Fig. 4 Top10 bacterial species contributing to the sPLS models predicting birth mode. sPLS models were trained to distinguish fecal microbial composition between children born by CS and vaginal delivery at A) 3 months of age and B) 1 year of age. Top 10 drivers were shown here. Area under the curve (AUC) values, calculated from repeated 10-fold cross-validation of the models and on the test set, are shown. Loading (model contribution) for each species is colored by phylum. Negative/positive loading corresponds to lower/higher abundance in CS born infants compared to vaginally delivered infants.

At 1 year of age, a 2-component model included 130 species from phyla Actinobacteria, Bacillota, Bacteroidota, Firmicutes, Proteobacteria, and Verrucomicrobia (Fig.4B, Supp. Fig 4). The 5 main contributors positively correlated with CS were *Veillonella parvula*, *Escherichia coli*, *Streptococcus lutetiensis*, *Faecalimonas umbilicata*, and *GGB9469 SGB14862*. The 5 main contributors negatively correlated with CS were *Bacteroides uniformis*, *Phocaeicola vulgatus*, *Phocaeicola dorei*, *Bacteroides thetaiotaomicron*, and *Faecalibacterium prausnitzii*.

We then investigated how the performance of the 3-month and 1-year models varied across the age spectrum by applying the models on all samples (larger age window, 3 months, N=1,355; 1 year, N=1,356), but keeping the model loadings fixed. We then calculated the AUC within a sliding age window of the same size as the original windows (3 months N=739; 1 year N=798), and shifted the window by 10 samples for each iteration (Supp. Fig 5). The AUC of the 3-month model peaked at 3 months, but remained above 0.6 at 1 year, indicating the predictive power of the early perturbed gut microbiome for the later timepoint. Surprisingly, the AUC of the 1-year model was greater at 3 months than at 1 year, indicating that the species driving the1-year perturbation were already CS-associated at 3 months, and likely normalized by 1 year of age.

CS microbial signatures and asthma risk at age 5

We then examined whether the CS microbial scores derived from the sPLS models were associated with the risk of asthma at age 5 in the entire cohort (3 months, N=1,355; 1 year, N=1,356). Only the 1-year CS microbial score was associated with a higher risk of asthma at age 5 in both base and covariate-adjusted logistic regression models (Base OR 1.41 per SD [1.12,1.80], P<0.01; Covariate-adjusted OR (aOR) 1.30 [1.03,1.68], P=0.03, N cases/controls = 101/1,048, Fig. 5A). When performing a combined model including both 1-year CS microbial score and delivery mode in relation to 5-year asthma, we still observed associations for the microbial score (Base OR 1.35 [1.08,1.74], P=0.01; aOR 1.27 [0.99,1.64], P=0.06), but not for CS delivery (OR 1.50 [0.94, 2.36], P=0.08; aOR 1.34 [0.82, 2.15], P=0.24), suggesting that the score might mediate the CS association. A mediation analysis revealed an average causal mediation effect (ACME) to be positive (ACME 0.01 [0.003,0.02], P=0.008, Fig. 5B), suggesting that part of the association of birth mode on asthma risk may be mediated by the 1-year gut microbial scores. The proportion of the total effect mediated was approximately 25% (Prop. mediated 0.25 [0.04, 1.32], P=0.03).

To investigate the possible effect of environmental factors on the development of CS-assoicated microbial composition, we tested the assoications between several exposures and 1-year CS scores using linear regression (Fig. 5C). Antibiotics to mother at birth was tested in the vaginal stratum, and was significantly associated with higher 1-year CS scores, indicating a “CS-like” microbial composition. In contrast, having older siblings was most associated with lower 1-year CS scores, followed by maternal and paternal race, and any breastfeeding at 1 year of age.

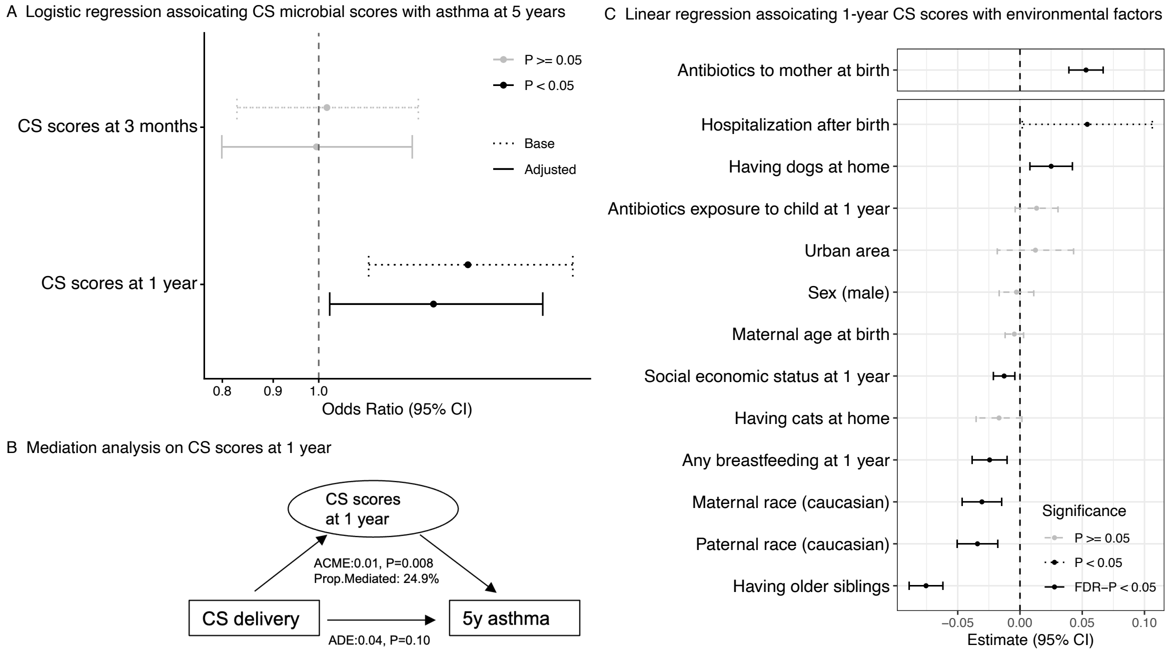


Fig. 5 Associations between CS scores, asthma at age 5, and environmental factors. A) Forest plot of logistic regression models evaluating the association of CS scores with asthma at 5 years. The base models were adjusted for processing period, exact age of infants, and the sample collection site, and the covariate-adjusted models were adjusted for covariates including gestational age, hospitalization after birth, birth weight, sex, race (father and mother), any breastfeeding by 1 year of age, exclusive breastfeeding ever, antibiotics exposure to infants by 1 year of age, paternal asthma history, maternal asthma history, birth season, processing period, exact age of infants, and the

sample collection site. B) Mediation analysis on CS score at 1 year adjusted for processing period, exact age, and the sample collection site. C) Linear regression models associating 1-year CS scores with environmental factors, adjusted for processing period, exact age, and the sample collection site. Antibiotics to mother at birth was tested in the vaginal stratum, while other factors were tested in the full cohort.

CS microbial signatures, gut metabolites, and asthma risk

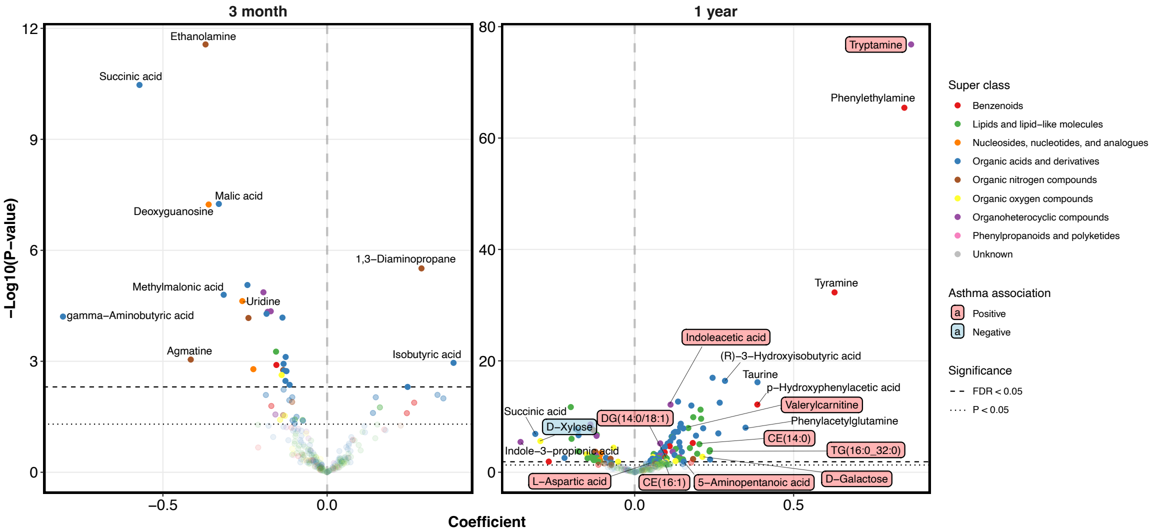
We then investigated the link between the CS microbial scores and gut metabolites quantified by nuclear magnetic resonance (NMR) and liquid chromatography with tandem mass spectrometry (LC-MS/MS). Out of 244 metabolites measured in the fecal samples at the concurrent timepoint (3 months, N=414; 1 year, N=1,243), 52 and 141 metabolites were found to be associated with 3-month and 1-year CS scores, respectively. These metabolites belonged to 7 classes including: organic acids and derivatives (3-month score: 27; 1-year score: 57), lipids and lipid-like molecules (4;42), organoheterocyclic compounds (4;12), benzenoids (4;9), organic nitrogen compounds (7;8), organic oxygen compounds (3;10), and nucleosides, nucleotides, and analogues (3;3) (Fig. 6A, Supp. Fig 6).

At 3 months of age, most metabolites were negatively associated with the 3-month CS microbial score, including ethanolamine, succinic acid, and malic acid, while 1,3-diaminopropane and isobutyric acid were positively associated with the 3-month CS microbial score (FDR adj. P<0.05). According to a pathway enrichment analysis, metabolites negatively associated with the CS microbial scores were found to be involved in pathways including phenylalanine, tyrosine, and tryptophan biosynthesis, glycine, serine, and threonine metabolism, arginine and proline metabolism, and alanine, aspartate, and glutamate metabolism (FDR adj. P<0.05, Supp. Fig 7).

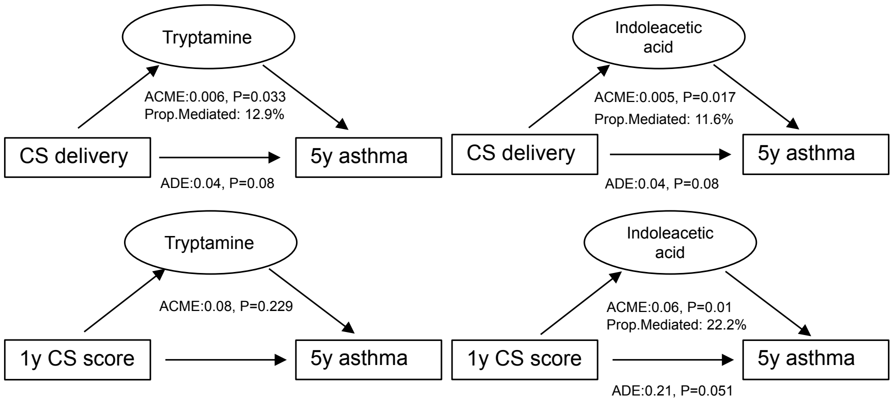
At 1 year of age, most metabolites were positively associated with CS microbial score, such as tryptamine, followed by phenylethylamine, tyramine, and taurine, while succinic acid, indole-3-propionic acid, and D-xylose were negatively associated with the CS microbial score. It is worth noting that the associations between metabolites and the 3-month CS microbial scores were mostly in the same direction as the associations between metabolites and 1-year CS microbial score (Supp. Fig 6, Spearman Rho=0.40, P<0.01). In other words, those metabolites of which the levels decreased with a CS-associated gut microbiome at 3 months of age most likely kept being at a lower level at 1 year of age as well. Metabolites positively associated with the 1-year CS microbial scores were mainly involved in pathways including alanine, aspartate and glutamate metabolism, arginine biosynthesis, glyoxylate and dicarboxylate metabolism, and citrate cycle (TCA cycle) (FDR adj. P<0.05, Supp. Fig 7).

Furthermore, in the 1-year stool metabolome, we found that 11 of the CS microbial score-associated metabolites were also significantly associated with the risk of asthma at age 5 in a directionally consistent manner (Fig. 6A). For example, D-xylose, negatively associated with the 1-year CS microbial score, was found to be negatively associated with asthma risk (OR 0.89 [0.81,0.99], P=0.03); Tryptamine, positively associated with the 1-year CS microbial scores, was positively associated with asthma risk (OR 1.17 [1.03,1.35], P=0.02). This indicates that the link between a CS-perturbed gut microbiome at 1 year of age and later asthma risk could be mediated through an altered gut metabolomic profile.

A Stool metabolites associated with CS microbial scores and asthma risk



B Mediation analysis on Tryptamine and Indoleacetic acid



C 1-year bacterial species assoicated with Tryptamine and Indoleacetic acid

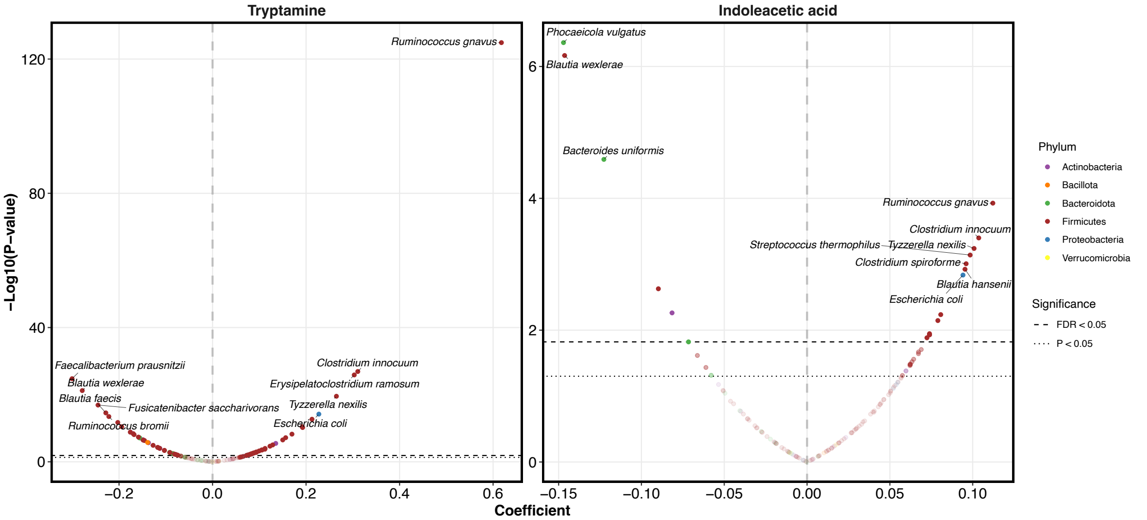


Fig. 6 Metabolites associated with CS microbial scores at 3 months and 1 year of age. A) MaAslin2 volcano plot of fecal metabolites associated with CS microbial scores at 3 months and 1 year of age, adjusting for processing period, exact age, and the sample collection site. Positive coefficients indicate a higher level of metabolites associated with higher CS microbial scores. The top 10 metabolites ranked by the coefficients were labeled as plain text. Those boxed labels indicate metabolites associated with asthma at age 5 according to logistic regression, adjusting for processing period, exact age, and the sample collection site. B) Mediation models for associations between CS delivery or 1-year CS microbial scores and asthma risk at age 5, adjusting for processing period, exact age, and the sample collection site. C) Linear models associating Tryptamine and Indoleacetic acid with abundances of 1-year gut microbiome. Positive coefficients indicate a higher abundance of the bacteria associated with higher levels of Tryptamine or Indoleacetic acid in the stool at 1 year of age. The top 10 species ranked by the coefficients were labeled as plain text.

To further clarify this relationship, we then tested the mediating role of these 11 metabolites in the association between CS delivery and asthma risk at age 5, and found that tryptamine and indoleacetic acid (a tryptophan metabolite) had a positive ACME in the mediation analysis (Tryptamine, ACME 0.006 [0.0004,0.01], P=0.03; Indoleacetic acid, ACME 0.005 [0.0005,0.01], P=0.02, Fig. 6B). However, these mediation effects were not significant after adjusting for 1-year CS scores, indicating that the microbial scores and metabolites reflected each other in CS-related asthma risk. Therefore, we next tested the mediating role of the metabolites in the association between the 1-year CS microbial scores and asthma risk. Indoleacetic acid and three other metabolites showed a positive ACME (Indoleacetic acid, ACME 0.06 [0.01, 0.11], P=0.01; Fig. 6B, Supp table 5). After adjusting for birth mode, it was still significant (ACME 0.05 [0.01, 0.10], P=0.01; Supp table 4), which indicates that Indoleacetic acid may also be important for vaginally born infants with high CS scores.

Using linear regression, we tested associations between the abundances of 133 gut bacterial species at 1 year and the concentrations of tryptamine and indoleacetic acid. After FDR correction, 78 species (59%) were significantly associated with tryptamine and 11 species (8%) with indoleacetic acid (Fig. 6C, Supp. Fig 8). Species associated with tryptamine belong to *Blautia* (7), *Streptococcus* (6), *Bifidobacterium* (5), *Roseburia* (4), *Mediterraneibacter* (4) and 36 other genera (52). *Blautia faecis* showed the strongest negative association with tryptamine levels, while *Ruminococcus gnavus* was significantly associated with higher tryptamine levels. Metabolites associated with indoleacetic acid belong to *Blautia* (2), *Erysipelatoclostridium* (2), *Bacteroides* (1), *Escherichia* (1), *Mediterraneibacter* (1), and 4 other genera (4). *Phocaeicola vulgatus* showed the strongest negative association with Indoleacetic acid levels, followed by *Blautia wexlerae*, while *Clostridium spiroforme* was associated with higher indoleacetic acid levels, followed by *Ruminococcus gnavus*.

Discussion

By leveraging CHILD, a large Canadian prospective birth cohort, we found that a CS-perturbed gut microbiome profile retained until 1 year of age partly mediated the association between CS and an increased risk of developing asthma by age 5, in agreement with previous findings. Moreover, we observed pronounced changes in gut metabolomics associated with the CS-perturbed microbiome. Children with high microbial CS scores at age 1 year tended to have elevated asthma-associated metabolites, which provide new insights into the possible mechanisms behind the increased asthma risk after CS delivery.

The first few months of life are pivotal for establishing the infant gut microbiome and maturation of the immune system. The pioneer bacteria dominating the early gut microbiome are mostly facultative anaerobes acting as oxygen scavengers in the neonatal gut. This early activity facilitates subsequent colonization of beneficial strict anaerobes which later dominate the mature gut microbiome in childhood[24]. However, CS can disrupt the microbial colonization of newborns by limiting exposure to the mother's gut microbiome at birth[25] and instead exposing the infant mainly to the mother's skin microbiota and microbes from the hospital environment[26]. Infants born vaginally develop a hypoxic and acidic gut environment enriched in strict anaerobes by day 3. In contrast, CS-born infants retain a more oxygen-rich gut environment and exhibit microbiomes dominated by Enterobacteriaceae[27]. The oxygenated gut environment can lead to delayed colonization of strict anaerobes such as *Bacteroides* and *Bifidobacterium*, which are crucial for breaking down Human Milk Oligosaccharides[28] (HMOs) and cross-feeding among members of the early gut microbiota[29]. After CS, opportunistic pathogens (including *Enterococcus*, *Enterobacter*, and *Klebsiella* species) may occupy the ecological niche and proliferate[30,31]. In our study, we observed higher abundances of facultative species in the CS-delivered infants at 3 months of age compared to the vaginally delivered, including species from *Klebsiella*, *Enterococcus*, and *Enterobacter*, and the anaerobic bacteria mainly comprised of *Veillonella parvula* and *Clostridium perfringens*. On the other hand, anaerobes from Bacteroidota were significantly depleted in the early gut microbiome of CS-born infants, such as *Bacteroides fragilis*, *Bacteroides uniformis*, and *Phocaeicola dorei*. Furthermore, this depletion of Bacteroidota appeared to last until 1 year of age, which suggests that this disrupted ecological colonization may lead to a lasting disadvantage of beneficial bacteria in the intra-species competition.CS-born infants show reduced colonization with *Escherichia coli*, and that the acquisition was mostly from the hospital milieu instead of from mothers during birth[32]. An antagonistic relationship wasfound antagonistic relationships between *Escherichia coli* and several species from *Klebsiella*, similarly between *Escherichia coli* and *Enterococcus faecalis*[33]. These inter- and intra-species competition dynamics could partially explain the long-lasting perturbations to age 1 year.

The changes in microbial composition after CS were furthermore associated with gut metabolite levels. Most metabolites in the infants' gut at 3 months were negatively associated with the CS-associated microbiome profile. The altered metabolites may, in turn, restrict the colonization of typical early-life bacteria by reducing the availability of key energy substrates.

Succinate, which was reduced in children with a high CS-associated microbiome profile, is a tricarboxylic acid (TCA) cycle intermediary metabolite produced in the mitochondria of host cells, and is also a catabolic metabolite of microbial carbohydrate fermentation. As a common byproduct of some bacteria, especially by primary fermenters such as *Bacteroides*, succinate is also a key cross-feeding metabolite since it can be consumed by secondary fermenters[34]. As we found the abundance of *Bacteroides* substantially decreased while *Veillonella parvula* (a succinate consumer) was enriched in the gut microbiome of CS-born infants, it is expected that the production of succinate was lower and consumption higher than in vaginally-born infants[35]. *Bacteroides*-derived succinate is reported to favor the colonization by anaerobes[34,36]. Consequently, this low concentration of succinate in the gut environment can affect microbiota composition, as well as microbial functional capacity.

Another example of such a microbe-metabolite co-dependency reflecting the colonization of early bacteria is for ethanolamine. Ethanolamine can be utilized as the main source of carbon, nitrogen, and energy under aerobic and anaerobic conditions, but the corresponding microbial genes from conserved operons are absent from almost all species except for *Listeria monocytogenes*, *Clostridium perfringens*, *Salmonella typhimurium*, *Enterococcus faecalis*, and *Escherichia coli*[37,38]. The significant decrease of ethanolamine in the gut of infants with CS-associated microbes could be related to the enrichment of *Clostridium perfringens* and *Enterococcus faecalis*, which then consequently hampered the colonization of *Escherichia coli*[39]. This also agrees with the antagonistic relationships between *Escherichia coli* and *Enterococcus faecalis* previously reported[33].

We previously reported that a prolonged CS-perturbed gut microbiome was associated with later asthma risk[11]. it is reasonable to hypothesize that the affected metabolite profiles observed at 3 months may contribute to both the microbial maturation and restoration of the initial perturbation, as well as the association between the CS-perturbed gut microbiome and asthma risk in later life. However, we didn't observe significant associations between these CS-associated metabolites at 3 months of age and asthma risk. It's worth noting that, consistent with our previous work, only 1-year CS scores mediated the association between the CS delivery and a higher risk of asthma. We therefore hypothesized that the 1-year stool metabolites could partially explain the effect of the 1-year CS microbial signature on the asthma risk.

At 1 year of age, the differences in the metabolites might reflect more of the specific metabolic function of the skewed microbial composition rather than a regulator of the colonization competition observed during the early establishment phase. We observed that tryptamine was the most pronounced metabolite positively associated with high CS microbial scores. Tryptamine is produced by the decarboxylation of tryptophan[40]. It is common in human and rodent feces, and mainly produced by the gut microbiota. *Ruminococcus gnavus* and *Clostridium sporogenes,* in particular, are able to produce tryptophan decarboxylase, the enzyme responsible for decarboxylation of tryptophan to tryptamine[41]. In our study, *Ruminococcus gnavus* was found to be more abundant in CS-born infants at 1 year, and strongly associated with a higher level of tryptamine, suggesting a role of *Ruminococcus gnavus*-produced tryptamine in the association between 1-year CS microbial signature and asthma risk. In fact, the level of tryptamine at 1 year was found to mediate the association between the CS delivery and asthma risk. Tryptamine is a ligand of the aryl hydrocarbon receptor (AhR), which is a key component of the immune response at barrier sites and is thus crucial for intestinal homeostasis by acting on epithelial renewal, barrier integrity, and many immune cell types[42]. The co-activation of immune cell types involved in the asthmatic response by AhR ligands can regulate the outcome of the reaction[43]. Wu et al[44] reported higher concentrations of tryptamine and other tryptophan metabolites in the asthmatic group by analyzing serum metabolites, and concluded that the asthmatic children had an increased indole-AhR pathway.

Indoleacetic acid, another AhR ligand which is also a downstream product of tryptamine, was found to mediate the association between CS delivery and 1-year CS microbial signature and asthma risk at age 5. Several bacteria, such as species from *Bacteroides*, *Clostridium*, and *Escherichia coli* can convert tryptamine to indoleacetic acid[40,42]. In our study, *Ruminococcus gnavus and Clostridium spiroforme* were most positively associated with the level of indoleacetic acid, followed by *Escherichia coli*. Indoleacetic acid has been found to exhibit both pro- and anti-inflammatory properties, depending on concentration, source, and its binding affinity[45]. Although studies showed that the elevated indoleacetic acid levels in the serum could potentially alleviate inflammation in asthma in murine models[46,47], indoleacetic acid also plays a proinflammatory role under certain conditions, largely through its interaction with AhR [45,48]. It's worth noting that our measurements were done before the children developed clinical asthma. More studies are needed to investigate the dynamics between stool metabolites and blood metabolites to make further inferences about the association between microbiota-derived metabolites and asthma.

We further identified molecular pathways based on the CS microbial score-associated metabolites. These molecular pathways could be involved in both host cell and microbial activities. It is under debate whether microbiome pertubations after CS has functional consequences[49]. Some studies report that the functional profiles of the infant gut microbiome are largely stable and resilient to exogenous perturbations, including CS delivery[50], whereas others show up-regulated signaling pathways in the CS born infants, such as tryptophan metabolism[51–53], consistent with our findings. Under limited energy intake, gut bacteria can upregulate TCA cycle genes, and additionally take advantage of the glycoxylate cycle to promote energy production[54]. Recently, high-resolution metabolomics analyses were performed to understand the differences in metabolites in children with asthma, and findings included pathways related to cellular energy and hypoxia, such as the TCA cycle, which possibly reflects the energetic burden due to inflammation and bronchoconstriction[55]. Other amino acid metabolism pathways were reported to be elevated in the asthmatics, including phenylalanine, tyrosine and tryptophan biosynthesis[56,57], glycine, serine, and threonine metabolism[57], alanine, aspartate and glutamate metabolism[58]. We observed metabolites involved in these pathways positively associated with the CS-perturbed microbiome, so we hypothesize that CS alters the gut microbiome, which may subsequently affect the host-microbe interactions via altered metabolites, and further associate with asthma through dysregulated energy metabolism and immune responses. Beyond CS, antibiotic exposure to mother at birth was associated with a similar perturbation in the gut microbiome of vaginally born infants. This makes the restoration of a perturbed gut microbiome important not only to those born by CS, but also to those infants subject to exposures that may cause perturbation in the gut microbiome. On the other hand, we also found having older siblings and any breastfeeding at 1 year of age negatively associated with having a retained CS-perturbed gut microbiome. Further analyses can focus on beneficial microbes these factors interact with, and explore their role in the prevention of asthma risk.

Taken together, we investigated associations between a CS-perturbed gut microbiome through the first year of life and later asthma risk, and evaluated the host-microbe interactions via metabolites. A strength of the study is that we were able to replicate and expand previous findings, using 16S rRNA sequencing in the Danish Copenhagen Prospective Studies on Asthma in Childhood (COPSAC2010) prospective birth cohort[11], with metagenomics, giving a much better resolution at the species level, and gut metabolomics giving credible insights into bacterial metabolism that may affect the child. Using standardized protocols, the CHILD study generated uniform sample collection and phenotyping with physician-diagnosed asthma as the clinical endpoint, enabling robust, generalizable inference. There are some limitations to our study. Given the observational design of the study, it is not possible to directly assess the causal relationship between CS, gut microbiota composition, and asthma. The relationships between the gut microbiome and metabolome are also complicated, and dynamically inter-connected. Measurements in stool do not distinguish between absorbed and bacterially-derived metabolites, or the location of those bacteria. We need more studies with integrative multi-omics methods to further disentangle the network between the microbiome and systemic levels of metabolites, and further explain the microbe-host interactions mediated by the metabolites. Finally, experiments are needed to validate the colonization competition.

Conclusion

In the CHILD cohort, we profiled CS-associated gut microbiome signatures at 3 months and 1 year of age using machine learning, and associated the 1-year CS microbial signature with a higher asthma risk at age 5. We observed metabolic alterations characterized by elevated tryptophan metabolites accompanying higher 1-year CS microbial scores. These findings provide new insight into how CS is be a risk factor for childhood asthma by altering the early gut microbiome and its metabolite production.

Methods

Characteristics of the study population

The CHILD study is a prospective longitudinal birth cohort study, which enrolled 3,405 subjects since pregnancy from 4 largely urban study centers across Canada (Vancouver, Edmonton, Winnipeg, and Toronto) from 2008 to 2012[21]. Our study included 3,263 subjects eligible at birth (Supp. Table 1), who were born at a minimum of 34 weeks of gestation and had no congenital abnormalities. The CHILD study followed children prospectively and collected detailed information on environmental exposures and clinical outcomes using a combination of questionnaires and in-person clinical assessments. Questionnaires related to environmental exposures, psychosocial stressors, nutrition and general health were administered at recruitment, prenatally, at 3, 6, 12, 18, 24, 30 months, and at 3, 4, and 5 years.

Childhood asthma: Childhood asthma was diagnosed (as Yes/Possible/No) by an expert study physician at the clinical assessment at the age of 5 years based on the CHILD Study's published approach, integrating ISAAC-based symptom history, physical examination, allergy skin-prick testing, and age-5 spirometry[21]. For this study, children were considered to have asthma only if the response was 'Yes' and the asthma phenotype was defined as comparing children with asthma (n=165) at 5 years versus children without asthma (n=2,234) at 5 years, children diagnosed as "possible" (n=247) were excluded.

Covariates: Relevant covariates were identified in the crude analyses (Supp. Table 2) associating with asthma risk, including gestational age, hospitalization after birth, birth weight, sex, race (father and mother), any breastfeeding by 1 year of age, exclusive breastfeeding ever, antibiotics exposure to infants by 1 year of age, paternal asthma history, maternal asthma history, birth season, and the sample collection site. These covariates were derived from questionnaires, except that gestational age, hospitalization after birth, birth weight, and sex information were extracted from birth charts. Missing data were handled using imputation. For continuous variables, missing values were imputed with the median. For categorical variables, missing values were replaced with the category "Unknown."

Stool Sample Collection

A subsample of 1,479 children had shotgun metagenomic data processed from fecal samples collected at the 3-month visit (n=1,422) and the 12-month visit (n=1,426). Sample collection and sequencing were performed as previously described[59]. Briefly, stool samples from diapers were collected at a home visit at around 3 months [mean (SD), 3.8 (1.1) months] and a clinic visit at around 1 year [mean (SD), 12.4 (1.3) months]. Samples were briefly stored at 4 °C and then aliquoted into four 2-mL cryovials using a stainless steel depyrogenated spatula and were frozen at −80 °C. CHILD recorded the time between stool collection and long-term storage. This processing time and the age of children at the time of stool sample collection were adjusted for in our statistical analysis. Samples collected for children above 1.5 years of age or samples with a processing time higher than 100 hours were excluded.

Gut Microbiota Characterization and Pre-processing

Shotgun metagenomic sequencing data with an average depth of 5 million reads per sample was generated by Diversigen (Minneapolis, MN, USA) from fecal samples. DNA was extracted from faecal samples using the MO Bio PowerSoil Pro with bead beating in 0.1mm glass bead plates. The amount of high-quality input DNA was measured using Quant-iT Picogreen. Libraries were then prepared and sequenced on an Illumina NextSeq using single-end 1 x 150 reads. Low quality (Q-Score<30) and length (<50) sequences were removed, and adapter sequences trimmed. Host and low-quality reads were removed, and only samples with at least 100,000 remaining reads or more were retained for downstream analysis. Shotgun metagenomic reads were mapped using the bioBakery 3 pipeline[60] to identify taxonomic (species and strain level) and functional features within each sample. The taxonomic annotation was obtained by mapping reads to a customized MetaPhlAn database (https://github.com/biobakery/MetaPhlAn/wiki/MetaPhlAn-4)[61], and agglomerated to the species level in the analyses.

NMR and LC-MS/MS metabolite quantification

Metabolic profiles were created from the same sequenced stool samples at The Metabolomics Innovation Center (TMIC) in Edmonton, Alberta using two separate assays. Targeted nuclear magnetic resonance (NMR) analysis of 31 metabolites was performed across 62 batches. Targeted liquid chromatography with tandem mass spectrometry (LC-MS/MS) analysis of 590 metabolites was performed TMIC's Microbiome Metabolism (MEGA) assay across 27 batches. NMR and LC-MS/MS precision were confirmed to be <5 and < 10% coefficients of variability (CV), respectively. Additionally, overlapping metabolites detected by both methods were cross-checked to confirm the accuracies of the reported concentrations values. Detailed methods of both NMR and LC-MS/MS analyses were described in previous work[62]. The superclass information of metabolites were retrieved from HMDB (Human Metabolome Database) using R package hmdbQuery (v1.28.0).

Statistical Analysis

Data analysis was conducted in R (v4.5.0). Fisher's exact test or Wilcoxon test were used for crude associations regarding birth method and asthma. Logistic regression was used to test the association with asthma risk. Within-sample diversity were assessed using observed richness and shannon diversity. Between-sample diversity were computed as bray-curtis distances (beta diversity). Differences in alpha diversity and beta diversity among groups, adjusting for processing period, exact age, and the sample collection site, were tested using linear regression and using permutational multivariate analysis of variance (PERMANOVA; Adonis from the package vegan (v2.7-1) with 999 permutations, covariates adjusted using by = "margin"), respectively[63]. The phyloseq (v1.52.0) package was used to pre-process the metagenomic taxonomy table[64]. Differential abundance analysis at species level was done using the Maaslin2 (v1.22.0) package with linear regression models (MaAsLin2 function)[65] adjusting for processing period, exact age, and the sample collection site. Models were applied to species detected in at least 5% samples with an abundance above 0.01% (69 species at 3 months of age, 130 species at 1 year of age). MaAsLin2 adds a pseudocount of half the minimum species level detected before the log transformation of relative abundances. P values were corrected using the Benjamini–Hochberg approach and results with FDR <0.05 were considered significant and presented as so. We applied supervised sPLS[66,67] models on the subset of the cohort - (exact age: 3-4 months, N=739; 11-12 months N=798, with Vaginal with antibiotics excluded) - after filtering species (prevalence >5% of children, >0.01% mean relative abundance) and log10 transforming relative abundances, using the lowest nonzero value as a pseudocount. The dataset was split into a training and test set by 75%. The performance was assessed using cross-validated AUC and AUC in the test set, and presented as so. The CS scores in the full dataset were obtained by applying the cross-validated model to the full dataset using the predict function (3-month visit (N=1,355), 12-month visit(N=1,356)). The association between CS scores and metabolomic profiles were tested using MaAsLin2 function[65], and the pathway analyses were done with the online tool - MetaboAnalyst[68]. To evaluate the mediation role of CS-perturbed gut microbiome in the association between CS and asthma risk, we applied mediation analysis using R package mediation (v4.5.1)[69]. The association between gut microbiome and metabolomic profiles was tested using linear regression based on log10 transformed relative abundances of the gut microbiome, adjusted for processing period, exact age, and the sample collection site.

Data availability

The informed consent obtained from the CHILD participants, in addition to the CHILD Inter-Institutional Agreement (IIA) which has been executed between the five Canadian institutions responsible for the study, govern the sharing of CHILD data. Data described in the manuscript are available by registration to the CHILD database (https://childstudy.ca/childdb/) and the submission of a formal request. All reasonable requests will be accommodated. More information about data access for the CHILD Cohort Study can be found at https://childstudy.ca/for-researchers/data-access/. Researchers interested in collaborating on a project and accessing CHILD Cohort Study data should contact child@mcmaster.ca. The shotgun metagenomic data used in this study are available in the NCBI database under BioProject accession code PRJNA838575. The metabolic profile data used in this study are available in the MetaboLights database under accession code MTBLS7919.

References

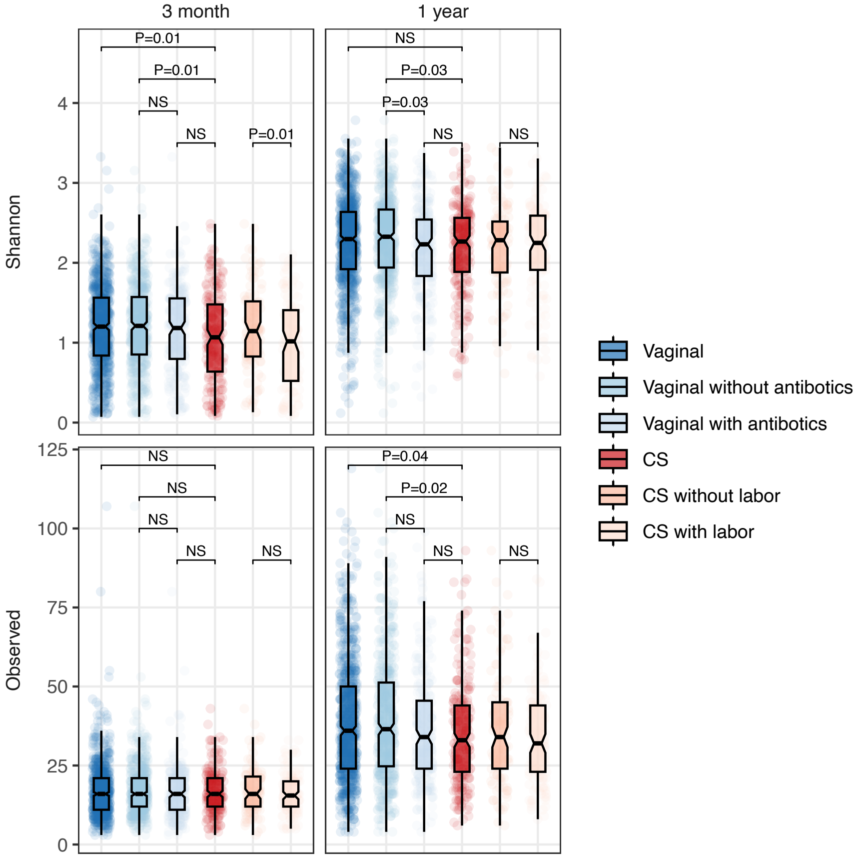
1. WHO Statement on Caesarean Section Rates. Geneva: World Health Organization; 2015 (WHO/RHR/15.02).
2. WHO Statement on caesarean section rates. *Reprod. Health Matters* **23**, 149–150 (2015).
3. Betrán, A. P. *et al.* Interventions to reduce unnecessary caesarean sections in healthy women and babies. *The Lancet* **392**, 1358–1368 (2018).
4. Velebil, P. *et al.* How to influence rising caesarean section rates in Europe? An invited scientific review by European Board and College of Obstetrics and Gynaecology (EBCOG). *Eur. J. Obstet. Gynecol. Reprod. Biol.* **311**, 113979 (2025).
5. World Health Organization. *WHO Recommendations Non-Clinical Interventions to Reduce Unnecessary Caesarean Sections*. (World Health Organization, Geneva, 2018).
6. Sandall, J. *et al.* Short-term and long-term effects of caesarean section on the health of women and children. *The Lancet* **392**, 1349–1357 (2018).
7. 2024 GINA Global Strategy for Asthma Management and Prevention Guideline Summary. *Guideline Central* https://www.guidelinecentral.com/guideline/41774.
8. Zhong, Z. *et al.* Association of cesarean section with asthma in children/adolescents: a systematic review and meta-analysis based on cohort studies. *BMC Pediatr.* **23**, 571 (2023).
9. O'Connor, A., McCarthy, F. P., Kelly, L., Khashan, A. S. & Maher, G. M. Mode of delivery and asthma in childhood and adolescence: Findings from the Millennium Cohort Study. *Clin. Exp. Allergy* **53**, 459–464 (2023).
10. Eder, W., Ege, M. J. & Von Mutius, E. The Asthma Epidemic. *N. Engl. J. Med.* **355**, 2226–2235 (2006).
11. Stokholm, J. *et al.* Delivery mode and gut microbial changes correlate with an increased risk of childhood asthma. *Sci. Transl. Med.* **12**, (2020).
12. Roduit, C. *et al.* High levels of butyrate and propionate in early life are associated with protection against atopy. *Allergy* **74**, 799–809 (2019).
13. Yu, M. *et al.* Aberrant purine metabolism in allergic asthma revealed by plasma metabolomics. *J. Pharm. Biomed. Anal.* **120**, 181–189 (2016).
14. Kelly, R. S. *et al.* Asthma Metabolomics and the Potential for Integrative Omics in Research and the Clinic. *Chest* **151**, 262–277 (2017).
15. Schjødt, M. S., Gürdeniz, G. & Chawes, B. The Metabolomics of Childhood Atopic Diseases: A Comprehensive Pathway-Specific Review. *Metabolites* **10**, 511 (2020).
16. Gürdeniz, G. *et al.* Neonatal metabolome of caesarean section and risk of childhood asthma. *Eur. Respir. J.* **59**, (2022).
17. Zhao, X. *et al.* The role of gut microbiome in the complex relationship between respiratory tract infection and asthma. *Front. Microbiol.* **14**, 1219942 (2023).
18. Chiu, C.-Y. *et al.* Gut microbial-derived butyrate is inversely associated with IgE responses to allergens in childhood asthma. *Pediatr. Allergy Immunol.* **30**, 689–697 (2019).
19. Arrieta, M.-C. *et al.* Early infancy microbial and metabolic alterations affect risk of childhood asthma. *Sci. Transl. Med.* **7**, 307ra152-307ra152 (2015).
20. Zheng, P. *et al.* Gut Microbiome and Metabolomics Profiles of Allergic and Non-Allergic Childhood Asthma. *J. Asthma Allergy* **15**, 419–435 (2022).

21. Subbarao, P. *et al.* The Canadian Healthy Infant Longitudinal Development (CHILD) Study: examining developmental origins of allergy and asthma. *Thorax* **70**, 998–1000 (2015).
22. Klopp, A. *et al.* Modes of Infant Feeding and the Risk of Childhood Asthma: A Prospective Birth Cohort Study. *J. Pediatr.* **190**, 192-199.e2 (2017).
23. Sevelsted, A., Stokholm, J. & Bisgaard, H. Risk of Asthma from Cesarean Delivery Depends on Membrane Rupture. *J. Pediatr.* **171**, 38-42.e4 (2016).
24. Dominguez-Bello, M. G. *et al.* Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc. Natl. Acad. Sci.* **107**, 11971–11975 (2010).
25. Bokulich, N. A. *et al.* Antibiotics, birth mode, and diet shape microbiome maturation during early life. *Sci. Transl. Med.* **8**, 343ra82 (2016).
26. Moore, R. E. & Townsend, S. D. Temporal development of the infant gut microbiome. *Open Biol.* **9**, 190128 (2019).
27. Lay, C. *et al.* A synbiotic intervention modulates meta-omics signatures of gut redox potential and acidity in elective caesarean born infants. *BMC Microbiol.* **21**, 191 (2021).
28. Kijner, S., Cher, A. & Yassour, M. The Infant Gut Commensal Bacteroides dorei Presents a Generalized Transcriptional Response to Various Human Milk Oligosaccharides. *Front. Cell. Infect. Microbiol.* **12**, (2022).
29. Fernandez-Julia, P., Commane, D. M., van Sinderen, D. & Munoz-Munoz, J. Cross-feeding interactions between human gut commensals belonging to the Bacteroides and Bifidobacterium genera when grown on dietary glycans. *Microbiome Res. Rep.* **1**, 12 (2022).
30. Kamada, N., Chen, G. Y., Inohara, N. & Núñez, G. Control of Pathogens and Pathobionts by the Gut Microbiota. *Nat. Immunol.* **14**, 685–690 (2013).
31. Shao, Y. *et al.* Stunted microbiota and opportunistic pathogen colonization in caesarean-section birth. *Nature* **574**, 117–121 (2019).
32. Nowrouzian, F. *et al.* Escherichia coli in Infants' Intestinal Microflora: Colonization Rate, Strain Turnover, and Virulence Gene Carriage. *Pediatr. Res.* **54**, 8–14 (2003).
33. Mäklin, T. *et al.* Strong pathogen competition in neonatal gut colonisation. *Nat. Commun.* **13**, 7417 (2022).
34. Fernández-Veledo, S. & Vendrell, J. Gut microbiota-derived succinate: Friend or foe in human metabolic diseases? *Rev. Endocr. Metab. Disord.* **20**, 439–447 (2019).
35. Leech, S. M. *et al.* Delivery mode is a larger determinant of infant gut microbiome composition at 6 weeks than exposure to peripartum antibiotics. *Microb. Genomics* **10**, 001269 (2024).
36. Rotstein, O. D., Nasmith, P. E. & Grinstein, S. The Bacteroides by-product succinic acid inhibits neutrophil respiratory burst by reducing intracellular pH. *Infect. Immun.* **55**, 864–870 (1987).
37. Korbel, J. O. *et al.* Systematic Association of Genes to Phenotypes by Genome and Literature Mining. *PLOS Biol.* **3**, e134 (2005).
38. Doranga, S. & Conway, T. Nitrogen assimilation by E. coli in the mammalian intestine. *mBio* **15**, e00025-24 (2024).
39. Zhou, J. *et al.* Ethanolamine enhances intestinal functions by altering gut microbiome and mucosal anti-stress capacity in weaned rats. *Br. J. Nutr.* **120**, 241–249 (2018).

40. Gao, J. *et al.* Impact of the Gut Microbiota on Intestinal Immunity Mediated by Tryptophan Metabolism. *Front. Cell. Infect. Microbiol.* **8**, (2018).
41. Bhattarai, Y. *et al.* Gut Microbiota-Produced Tryptamine Activates an Epithelial G-Protein-Coupled Receptor to Increase Colonic Secretion. *Cell Host Microbe* **23**, 775-785.e5 (2018).
42. Agus, A., Planchais, J. & Sokol, H. Gut Microbiota Regulation of Tryptophan Metabolism in Health and Disease. *Cell Host Microbe* **23**, 716–724 (2018).
43. Poulain-Godefroy, O. *et al.* The Aryl Hydrocarbon Receptor in Asthma: Friend or Foe? *Int. J. Mol. Sci.* **21**, 8797 (2020).
44. Wu, H. *et al.* PAH-induced metabolic changes related to inflammation in childhood asthma. *Environ. Sci. Pollut. Res.* **30**, 13739–13754 (2023).
45. Zhong, L., Boopathi, S., Purushothaman, B., Tu, Q. & Zhang, Y. Gut microbiota – indole-3-acetic acid axis in cancer: dual functions, mechanistic insights, and therapeutic potential. *Microbiol. Res.* **300**, 128293 (2025).
46. Li, L. *et al.* Hydrogen-rich water alleviates asthma airway inflammation by modulating tryptophan metabolism and activating aryl hydrocarbon receptor via gut microbiota regulation. *Free Radic. Biol. Med.* **224**, 50–61 (2024).
47. Wang, H. *et al.* Gut Microbiota-Derived Tryptophan Metabolites Alleviate Allergic Asthma Inflammation in Ovalbumin-Induced Mice. *Foods* **13**, 1336 (2024).
48. Vanholder, R. *et al.* Review on uremic toxins: Classification, concentration, and interindividual variability. *Kidney Int.* **63**, 1934–1943 (2003).
49. Torow, N. & Hornef, M. W. The Neonatal Window of Opportunity: Setting the Stage for Life-Long Host-Microbial Interaction and Immune Homeostasis. *J. Immunol.* **198**, 557–563 (2017).
50. Chu, D. M. *et al.* Maturation of the Infant Microbiome Community Structure and Function Across Multiple Body Sites and in Relation to Mode of Delivery. *Nat. Med.* **23**, 314–326 (2017).
51. Lai, C. *et al.* Effect of different delivery modes on intestinal microbiota and immune function of neonates. *Sci. Rep.* **14**, 17452 (2024).
52. Vänni, P. *et al.* Delivery mode and perinatal antibiotics influence the predicted metabolic pathways of the gut microbiome. *Sci. Rep.* **11**, 17483 (2021).
53. Li, N. *et al.* Distinct gut microbiota and metabolite profiles induced by delivery mode in healthy Chinese infants. *J. Proteomics* **232**, 104071 (2021).
54. Yang, Y. *et al.* Caloric Restriction Remodels Energy Metabolic Pathways of Gut Microbiota and Promotes Host Autophagy. 2020.08.16.251215 Preprint at https://doi.org/10.1101/2020.08.16.251215 (2020).
55. Farraia, M. *et al.* Metabolic interactions in asthma. *Eur. Ann. Allergy Clin. Immunol.* **51**, 196 (2019).
56. Kelly, R. S. *et al.* Asthma Metabolomics and the Potential for Integrative Omics in Research and the Clinic. *Chest* **151**, 262–277 (2017).
57. Cottrill, K. A. *et al.* Metabolomics identifies disturbances in arginine, phenylalanine, and glycine metabolism as differentiating features of exacerbating atopic asthma in children. *J. Allergy Clin. Immunol. Glob.* **2**, 100115 (2023).

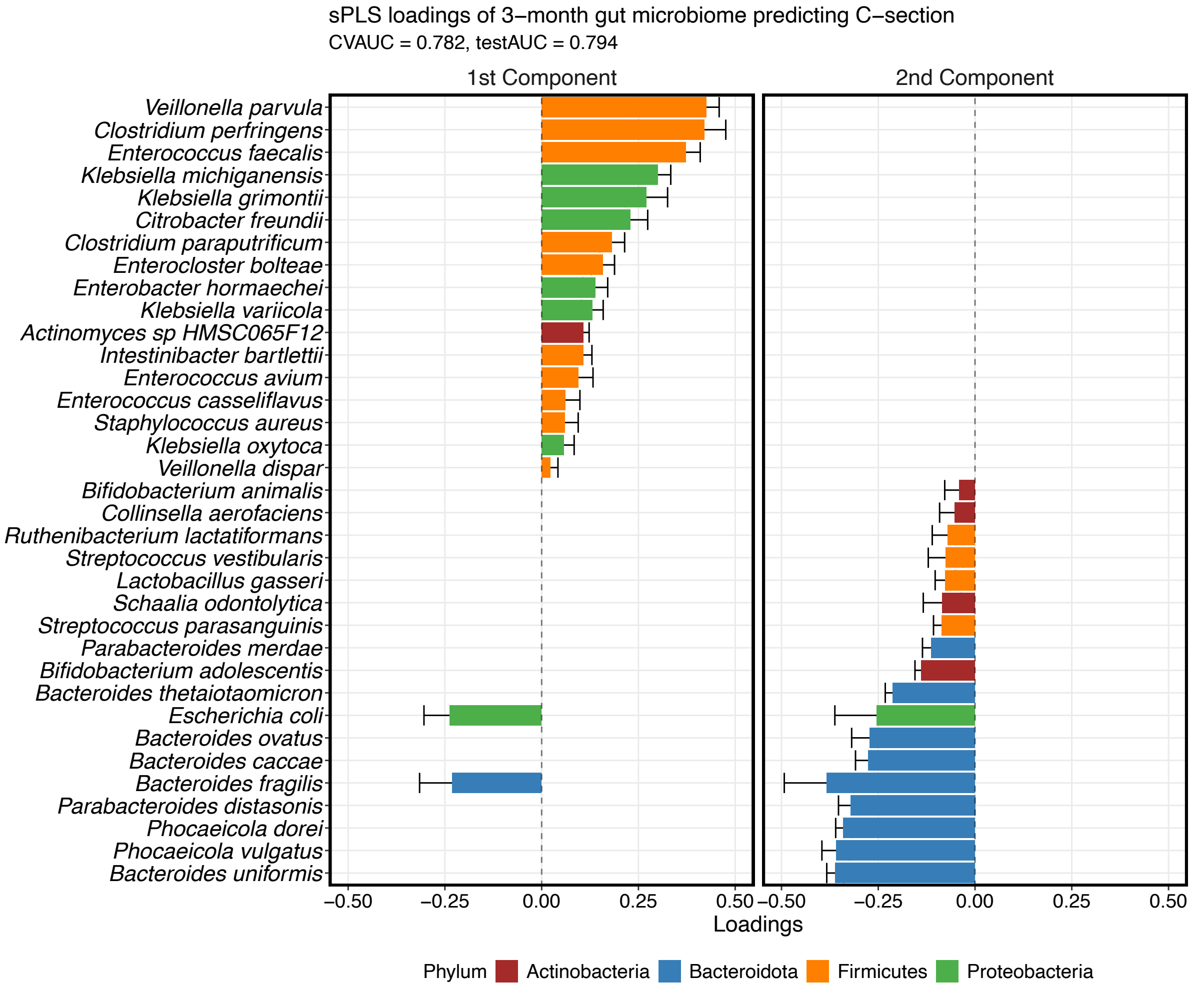
58. Chiu, C.-Y. *et al.* Integrated metabolic and microbial analysis reveals host–microbial interactions in IgE-mediated childhood asthma. *Sci. Rep.* **11**, 23407 (2021).
59. Dai, D. L. Y. *et al.* Breastfeeding enrichment of B. longum subsp. infantis mitigates the effect of antibiotics on the microbiota and childhood asthma risk. *Med* **4**, 92-112.e5 (2023).
60. Beghini, F. *et al.* Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *eLife* **10**, e65088.
61. Ennis, D., Shmorak, S., Jantscher-Krenn, E. & Yassour, M. Longitudinal quantification of Bifidobacterium longum subsp. infantis reveals late colonization in the infant gut independent of maternal milk HMO composition. *Nat. Commun.* **15**, 894 (2024).
62. Hoskinson, C. *et al.* Delayed gut microbiota maturation in the first year of life is a hallmark of pediatric allergic disease. *Nat. Commun.* **14**, 4785 (2023).
63. Jari Oksanen, F. Guillaume Blanchet, Michael Friendly, Roeland Kindt, Pierre Legendre, Dan McGlinn, Peter R. & Minchin, R. B. O'Hara, Gavin L. Simpson, Peter Solymos, M. Henry H. Stevens, Eduard Szoecs and Helene Wagner. vegan: Community Ecology Package. R package version 2.5-6. *2019* https://CRAN.R-project.org/package=vegan.
64. McMurdie, P. J. & Holmes, S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE* **8**, e61217 (2013).
65. Mallick, H. *et al.* Multivariable association discovery in population-scale meta-omics studies. *PLOS Comput. Biol.* **17**, e1009442 (2021).
66. Rohart, F., Gautier, B., Singh, A. & Cao, K.-A. L. mixOmics: an R package for 'omics feature selection and multiple data integration. 108597 Preprint at https://doi.org/10.1101/108597 (2017).
67. Thorsen, J. Implementation and utility functions for using the mixOmics package in the caret framework: github.com/jonathanth/mixOmicsCaret. (2017).
68. Xia, J. & Wishart, D. S. Metabolomic data processing, analysis, and interpretation using MetaboAnalyst. *Curr. Protoc. Bioinforma.* **Chapter 14**, 14.10.1-14.10.48 (2011).
69. Tingley, D., Yamamoto, T., Hirose, K., Keele, L. & Imai, K. mediation: R Package for Causal Mediation Analysis. *J. Stat. Softw.* **59**, 1–38 (2014).

Supplementary figures

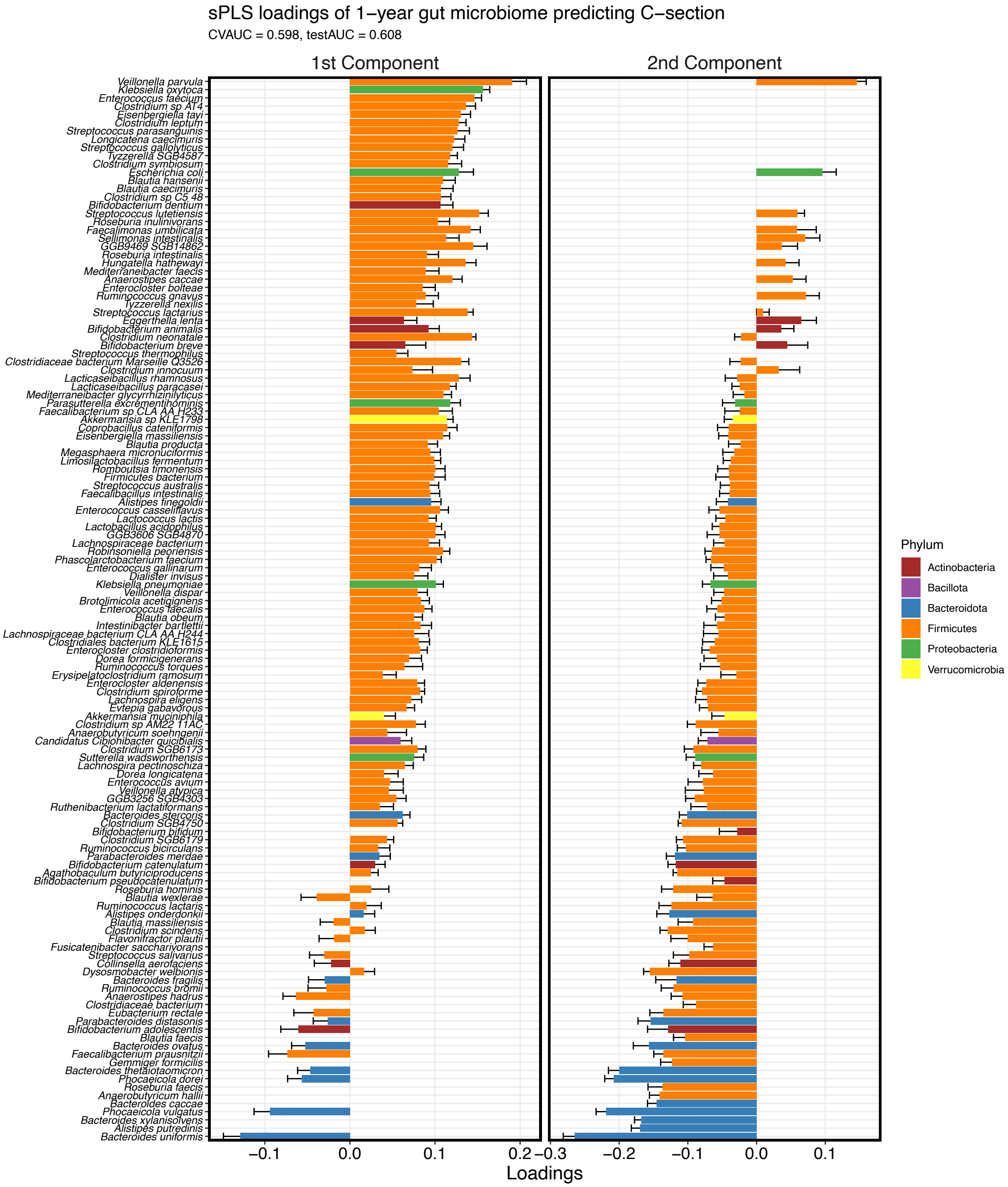


Supplementary Fig.1 Observed and Shannon diversity of the gut microbiome at 3 months and 1 year of age. Linear regression was employed for statistical comparisons at 3 months of age and 1 year of age, adjusting for processing period, exact age, and the sample collection site.

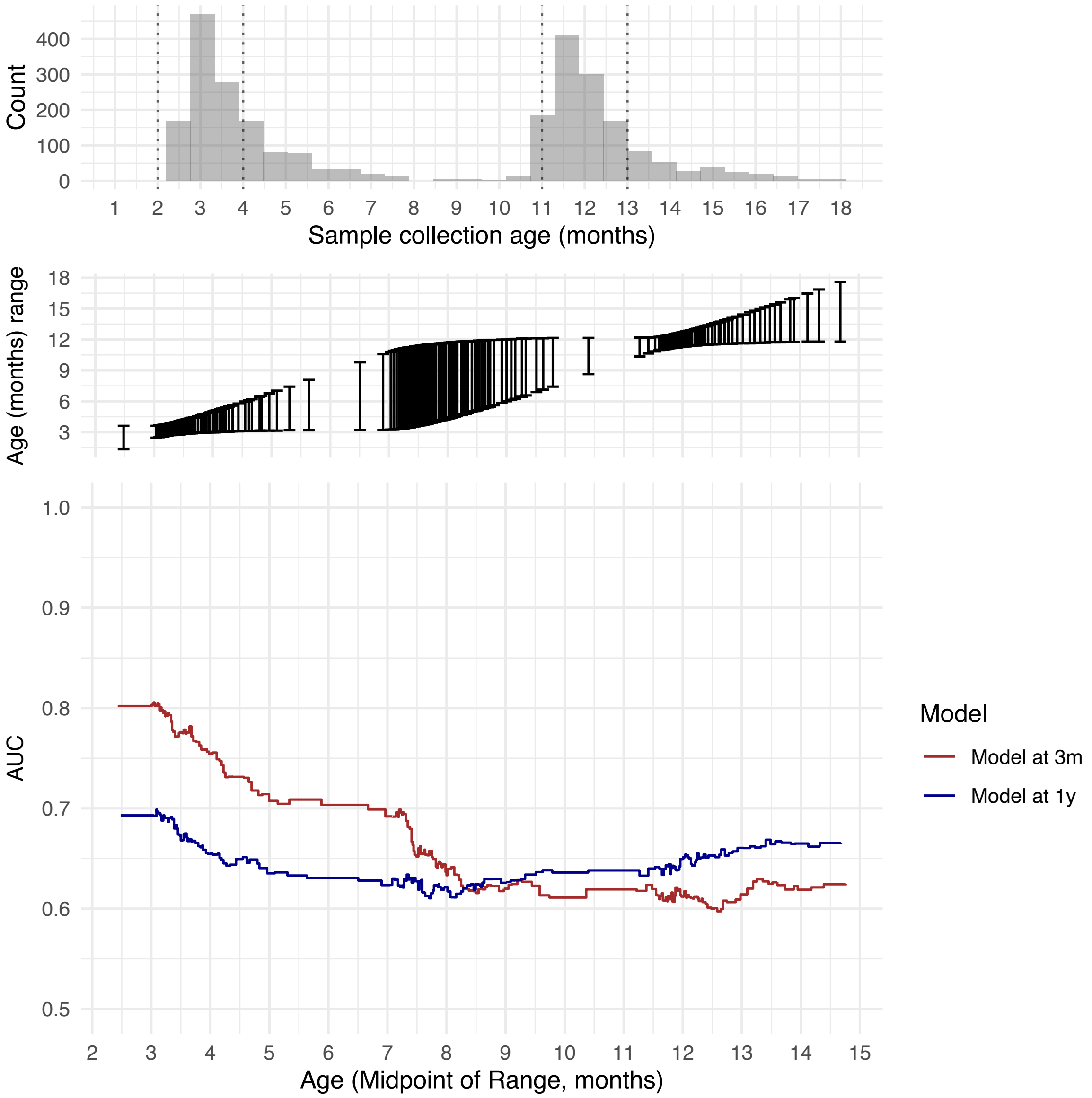
Supplementary Fig. 2 Distribution of the gut microbiome at 3 months and 1 year. A) Prevalence and mean relative abundance of the gut microbiome at 3 months and 1 year. B) Species overlap between 3 months and 1 year (Prevalence > 0.05, mean relative abundance > 0.01%).



Supplementary Fig.3 Fecal microbial compositions assoicated with birth mode at 3 months of age. A sPLS was used to distinguish fecal microbial composition at 3 months of age between children born by CS and vaginal delivery. Area under the curve(AUC) values calculated from the repeated 10-fold cross-validation of the models and on the test set were shown. Loading (model contribution) for each species is ranked by magnitude and colored by phylum. Negative/positive loading corresponds to lower/higher abundance in CS born infants compared to vaginally delivered ones.



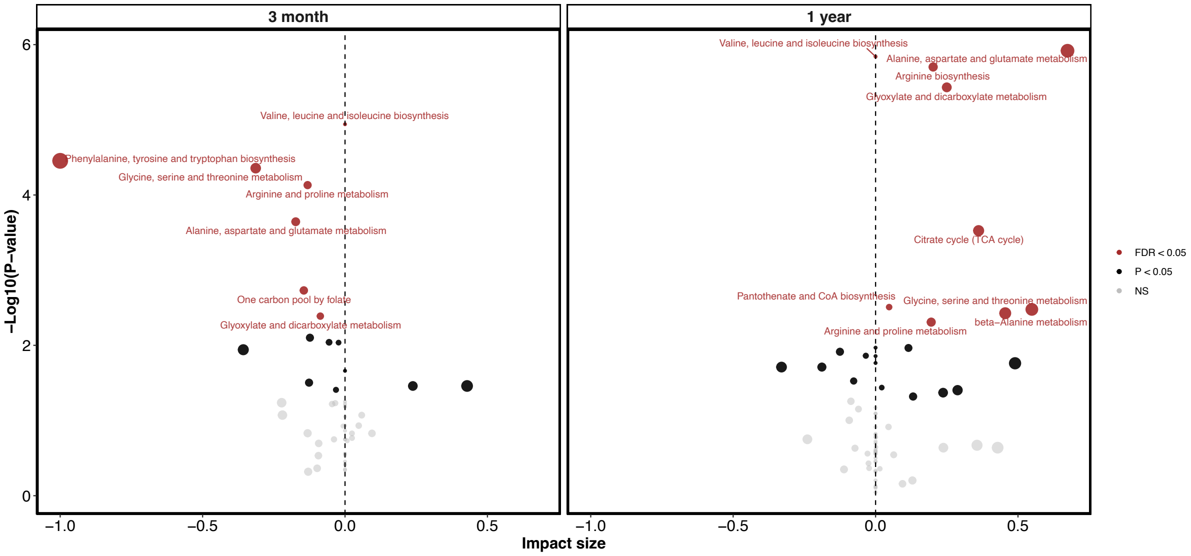
Supplementary Fig.4 Fecal microbial compositions assoicated with birth mode at 1 year of age. A sparse partial least squares (sPLS) was used to distinguish fecal microbial composition at 1 year of age between children born by CS and vaginal delivery. Area under the curve(AUC) values calculated from the repeated 10-fold cross-validation of the models and on the test set were shown. Loading (model contribution) for each species is ranked by the magnitude and colored by phylum. Negative/positive loading corresponds to lower/higher abundance in CS born infants compared to vaginally delivered ones.



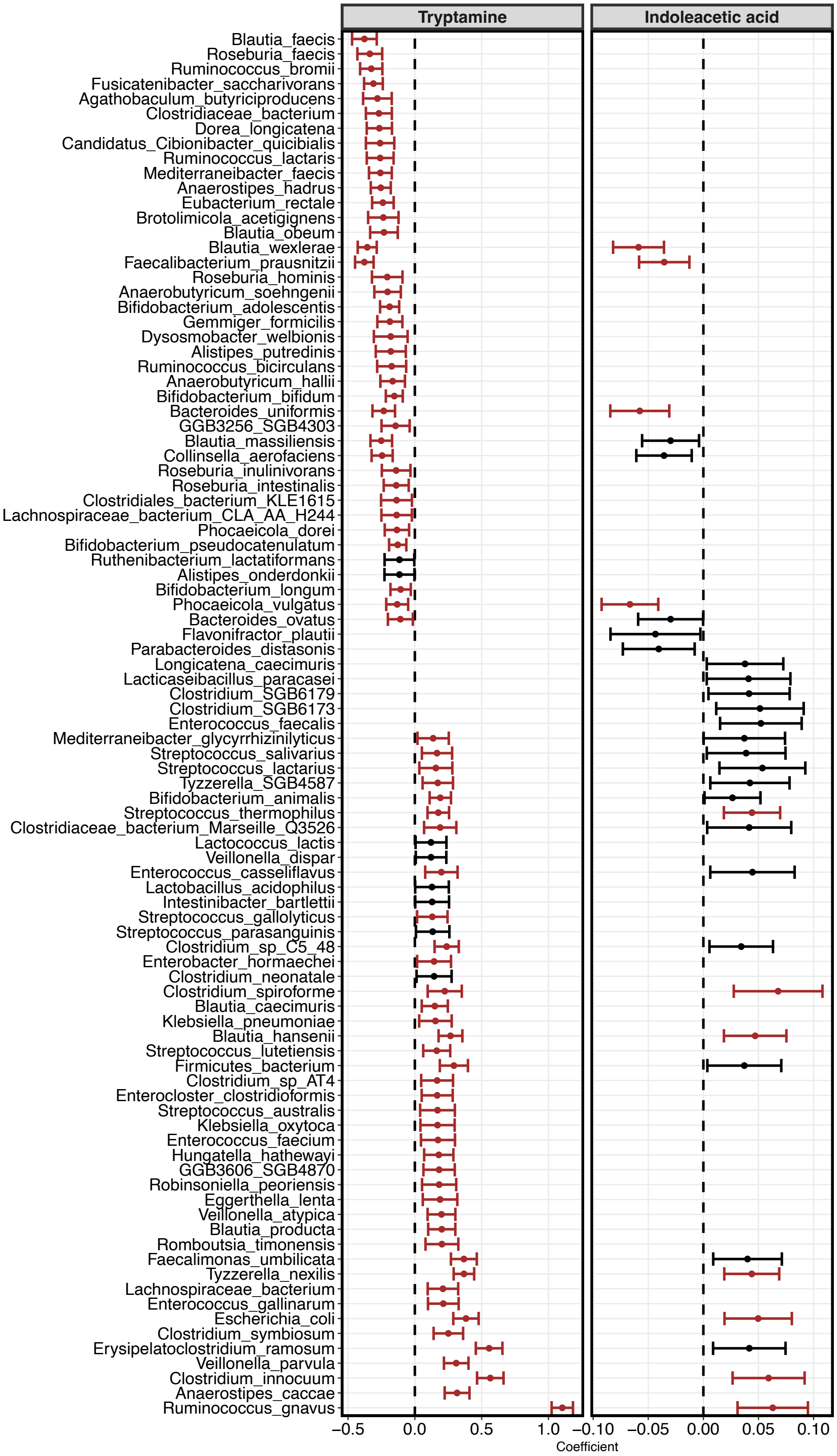
Supplementary Fig.5 The performance of models trained at 3 months and 1 year of age in the whole dataset measured by AUC. The bottom stepline plot illustrates the AUC calculated in windows of samples (739 samples for the 3-month model, and 788 samples for the 1-year model), shifted by 10 samples at each step, based on the predictions from the 3-month and 1-year models for birth mode classification. Each step represents a range of child ages (in months), with the x-axis showing the midpoint of the age range, and the top plot showing the corresponding age range.


3 month
1 year
Succinic acid
gamma−Aminobutyric acid
Malic acid
Ethanolamine
p−Cresol
Deoxyguanosine
Methylmalonic acid
Agmatine
Uridine
Indole−3−propionic acid
Inosine
Hypoxanthine
N−Acetyl−L−aspartic acid
Glutaric acid
Xanthine
Uracil
Spermidine
Nicotinic acid
Ethylmalonic acid
D−Xylose
Acetone
Benzoic acid
N1−Acetylspermidine
DOPA
Dimethylamine
N1,N12−Diacetylspermine
LysoPC(18:1/0:0)
3−Phenylpropyl propanoate
Glyceric acid
LysoPC(16:1/0:0)
Methylhistidine
Pyruvic acid
DG(18:1/18:2)
Trimethylamine
LysoPC(18:2/0:0)
Methionine sulfoxide
L−Tyrosine
Butyric acid
Pimelylcarnitine
Valeric acid
Methylamine
Histamine
Glycerol
Homocitrulline
Methionine
Proline
PC(36:4)
Citrulline
Tartaric acid
Citric acid
L−Alanine
Phenylalanine
DG(18:2/18:2)
Choline
DG(16:0/18:2)
Isoleucine
PC(36:3)
DG(18:1/18:3)
Leucine
Serine
L−Valine
4−Hydroxyphenylpyruvic acid
N−Acetyl−L−tyrosine
Symmetric dimethylarginine
myo−Inositol
GlcCer(d18:1/24:1)
Dimethylglycine
L−Threonine
3−Aminoisobutanoic acid
TG(16:0_36:6)
Spermine
PC(32:0)
CE(17:1)
DG(18:1/20:3)
N−Acetyl−L−methionine
DG(16:0/16:0)
Glycine
L−Tryptophan
Ethanol
N−Acetylisoleucine
GlcCer(d18:1/22:0)
PC(30:0)
N−Acetylvaline
Cer(d18:1/16:0)
alpha−Ketoisovaleric acid
Lysine
PC(36:5)
Isocitric acid
DG(14:0/20:0)
Nudifloramide
Ketoleucine
N−Acetyl−L−alanine
N−Acetylputrescine
CE(16:1)
Serotonin
Threonic acid
DG(16:1/18:1)
N−Acetyl−Leu
Indole
GlcCer(d18:1/16:0)
2−Hydroxy−3−methylbutyric acid
LacCer(d18:1/18:0)
L−Aspartic acid
N−Acetylserine
Isopropyl alcohol
Malonic acid
N−Methyl−D−aspartic acid
TG(16:0_32:0)
CE(20:4)
Indoleacetic acid
L−Arginine
Acetylglycine
Glutamine
N−Acetyl−L−glutamic acid
CE(22:5)
Dopamine
cis−Aconitic acid
Indolelactic acid
CE(20:3)
Acetoin
2−Methylhippuric acid
LacCer(d18:1/22:0)
beta−Alanine
D−Galactose
LacCer(d18:1/24:1)
4−Hydroxybenzoic acid
2−Hydroxyglutarate
DG(14:0/18:1)
CE(22:1)
Oxalacetic acid
L−alpha−Aminobutyric acid
L−Asparagine
Glutamic acid
N−Acetylhistidine
5−Aminopentanoic acid
Valerylcarnitine
CE(20:0)
LacCer(d18:1/16:0)
CE(20:1)
CE(14:0)
CE(22:2)
Butyrylcarnitine
Pyroglutamic acid
2−Hydroxybutyric acid
(S)−3,4−Dihydroxybutyric acid
1,3−Diaminopropane
(R)−3−Hydroxyisobutyric acid
Oxoglutaric acid
Guanidinopropionic acid
Isobutyric acid
Creatine
Taurine
Phenylacetylglutamine
p−Hydroxyphenylacetic acid
Creatinine
Tyramine
Tryptamine
Phenylethylamine
−1.0
−0.5
0.0
0.5
1.0
−1.0
−0.5
0.0
0.5
1.0
Coefficient
Significance
FDR < 0.05
P < 0.05
NS

Supplementary Fig.6 MaAslin2 forest plot of differential abundances of stool metabolites according to CS scores, adjusting for processing period, exact age, and the sample collection site. Positive coefficients indicate a higher abundance of metabolites associated with high CS scores.



Supplementary Fig.7 MetaboAnalyst pathway impact based on representative metabolites associated with CS microbial scores. Positive impact size indicates pathways involved in high-CS-score-associated metabolites, while negative impact size indicates pathways involved in low-CS-score-associated metabolites.


Tryptamine
Indoleacetic acid
Blautia_faecis
Roseburia_faecis
Ruminococcus_bromii
Fusicatenibacter_saccharivorans
Agathobaculum_butyriciproducens
Clostridiaceae_bacterium
Dorea_longicatena
Candidatus_Cibionibacter_quicibialis
Ruminococcus_lactaris
Mediterraneibacter_faecis
Anaerostipes_hadrus
Eubacterium_rectale
Brotolimicola_acetigignens
Blautia_obeum
Blautia_wexlerae
Faecalibacterium_prausnitzii
Roseburia_hominis
Anaerobutyricum_soehngenii
Bifidobacterium_adolescentis
Gemmiger_formicilis
Dysosmobacter_welbionis
Alistipes_putredinis
Ruminococcus_bicirculans
Anaerobutyricum_hallii
Bifidobacterium_bifidum
Bacteroides_uniformis
GGB3256_SGB4303
Blautia_massiliensis
Collinsella_aerofaciens
Roseburia_inulinivorans
Roseburia_intestinalis
Clostridiales_bacterium_KLE1615
Lachnospiraceae_bacterium_CLA_AA_H244
Phocaeicola_dorei
Bifidobacterium_pseudocatenulatum
Ruthenibacterium_lactatiformans
Alistipes_onderdonkii
Bifidobacterium_longum
Phocaeicola_vulgatus
Bacteroides_ovatus
Flavonifractor_plautii
Parabacteroides_distasonis
Longicatena_caecimuris
Lacticaseibacillus_paracasei
Clostridium_SGB6179
Clostridium_SGB6173
Enterococcus_faecalis
Mediterraneibacter_glycyrrhizinilyticus
Streptococcus_salivarius
Streptococcus_lactarius
Tyzzerella_SGB4587
Bifidobacterium_animalis
Streptococcus_thermophilus
Clostridiaceae_bacterium_Marseille_Q3526
Lactococcus_lactis
Veillonella_dispar
Enterococcus_casseliflavus
Lactobacillus_acidophilus
Intestinibacter_bartlettii
Streptococcus_gallolyticus
Streptococcus_parasanguinis
Clostridium_sp_C5_48
Enterobacter_hormaechei
Clostridium_neonatale
Clostridium_spiroforme
Blautia_caecimuris
Klebsiella_pneumoniae
Blautia_hansenii
Streptococcus_lutetiensis
Firmicutes_bacterium
Clostridium_sp_AT4
Enterocloster_clostridioformis
Streptococcus_australis
Klebsiella_oxytoca
Enterococcus_faecium
Hungatella_hathewayi
GGB3606_SGB4870
Robinsoniella_peoriensis
Eggerthella_lenta
Veillonella_atypica
Blautia_producta
Romboutsia_timonensis
Faecalimonas_umbilicata
Tyzzerella_nexilis
Lachnospiraceae_bacterium
Enterococcus_gallinarum
Escherichia_coli
Clostridium_symbiosum
Erysipelatoclostridium_ramosum
Veillonella_parvula
Clostridium_innocuum
Anaerostipes_caccae
Ruminococcus_gnavus
−0.5
0.0
0.5
1.0
−0.10
−0.05
0.00
0.05
0.10
Coefficient

Supplementary Fig.8 Linear models associating Tryptamine and Indoleacetic acid with abundances of 1-year gut microbiome. Positive coefficients indicate a higher abundance of the bacteria associated with higher levels of Tryptamine or Indoleacetic acid in the stool at 1 year of age.