Trajectories of microbiome-derived bile acids in early life – insights into the progression to islet autoimmunity

Santosh Lamichhane[1,2,*], Alex M Dickens[1,3], Tanja Buchacher[1,4], Evita Tianai Lou[5,6], Vincent Charron-Lamoureux[7], Roosa Kattelus[1,4], Pragya Karmacharya[1], Lucas Pinto da Silva[1], Matilda Kråkström[1], Omid Rasool[1,4], Partho Sen[1], Corinn Walker[8], Abubaker Patan[7], Emily C. Gentry[7], Aron Arzoomand[9], Tadepally Lakshmikanth [5,6,9], Jaromir Mikeš [9], Aman Mebrahtu[5,6], Tommi Vatanen[10,11,12,13,14], Manuela Raffatellu[14,15,16], Karsten Zengler[8,15,17], Tuulia Hyötyläinen[18], Ramnik J. Xavier[13], Petter Brodin[5,6,9], Riitta Lahesmaa[1,2,4], Pieter C. Dorrestein[7,15], Mikael Knip[10,19,*], Matej Orešič[1,20,21,*]

[1]Turku Bioscience Centre, University of Turku and Åbo Akademi University, FI-20520 Turku, Finland

[2] Institute of Biomedicine, University of Turku, 20520 Turku, Finland

[3]Department of Chemistry, University of Turku, Turku, 20500, Finland

[4] InFLAMES Research Flagship Center, University of Turku, 20520 Turku, Finland

[5] Medical Research Council Laboratory of Medical Sciences (MRC LMS), Imperial College Hammersmith Campus, London, UK.

[6] Department of Immunology and Inflammation, Imperial College London, W12 0NN, London, UK.

[7] Collaborative Mass Spectrometry Innovation Center, Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, San Diego, CA, USA

[8] Department of Pediatrics, University of California San Diego, La Jolla, California, USA

[9] Department of women's and children's health, Karolinska Institutet, Stockholm, Sweden

[10] Research Program for Clinical and Molecular Metabolism, Faculty of Medicine, University of Helsinki, Finland

[11] Liggins Institute, University of Auckland, Auckland, New Zealand.

[12] Institute of Biotechnology, Helsinki Institute of Life Science, University of Helsinki, Helsinki, Finland.

[13] The Broad Institute of MIT and Harvard, Cambridge, MA, USA

[14] Faculty of Agriculture and Forestry, University of Helsinki, Helsinki, Finland.

[15] Center for Microbiome Innovation, University of California, San Diego, La Jolla, CA, 39 92093, USA

[16] Chiba University-UC San Diego Center for Mucosal Immunology, Allergy, and Vaccines (CU-UCSD cMAV), La Jolla, CA 92093, USA

[17] Department of Bioengineering, University of California, San Diego, La Jolla, CA, 92093, USA

[18] School of Science and Technology, Örebro University, Örebro, Sweden

[19] Department of Pediatrics, Center for Child Health Research, Tampere University Hospital, FI-33520 Tampere, Finland

[20] Department of Life Technologies, University of Turku, Turku, FI-20014 Turku, Finland

[21] School of Medical Sciences, Faculty of Medicine and Health, Örebro University, 702 81 Örebro, Sweden

*Corresponding authors. E-mail addresses: santosh.lamichhane@utu.fi, mikael.knip@helsinki.fi, matej.oresic@utu.fi

Abstract

Recent studies reveal that gut microbes produce diverse bile acid conjugates, termed microbially conjugated bile acids (MCBAs). However, their regulation and health effects remain unclear. Here, we analyzed early-life MCBA patterns and their link to islet autoimmunity. We quantified 110 MCBAs in 303 stool samples collected longitudinally (3–36 months) from children who developed one or more islet autoantibodies and controls who remained autoantibody-negative. Stool MCBAs showed distinct age-dependent trajectories and correlated with gut microbiome composition. Altered levels of ursodeoxycholic and deoxycholic acid conjugates were linked to islet autoimmunity as well as modulated monocyte activation in response to immunostimulatory lipopolysaccharide and Th17/Treg cell balance. These findings suggest MCBAs influence immune development and type 1 diabetes risk.

Introduction

Bile acids (BAs) are cholesterol-derived surfactants produced from cholesterol in the liver, stored in the gall bladder, and released into the small intestine upon meal consumption[1-3]. The primary BAs produced in the hepatocytes consist of cholic acid (CA), chenodeoxycholic acid (CDCA), and their amidates (*i.e.*, either glycine or taurine conjugates)[4]. Gut microbes modify these host-derived primary BAs to a broad range of secondary BAs *via* a series of reactions including deconjugation, dehydroxylation, dehydrogenation, oxidation, and epimerization[5]. Almost 95% of these BAs (both primary and secondary) are recirculated to the liver *via* the enterohepatic circulation across the distal small intestine[6]. The remaining 5% still undergo a wide range of transformations throughout the hind gastrointestinal tract[2,3,7,8]. Traditionally, it was believed that BAs undergo amino acid conjugation only in the liver, mediated by a human enzyme known as bile acid-CoA:amino acid N-acyltransferase[9,10]. However, recent studies demonstrate that in addition to the host, gut microbes can reconjugate BAs *via* bile salt hydrolase N-acyltransferase activity[11,12] . Therefore, hundreds of microbially conjugated amino acid amidates have been identified and termed as novel ‘microbially conjugated bile acids’ (MCBAs)[2,3,7,8].

BAs are known as emulsifiers, facilitating the absorption of lipids in the gut. They are also important signaling molecules and antimicrobial compound that influence human health. BAs act as ligands to activate various receptors such as farnesoid X receptor (FXR), pregnane X receptor (PXR), Takeda G-protein coupled receptor 5 (TGR5), and sphingosine-1-phosphate receptors (S1PRs)[13-15]. These receptors regulate cellular growth and differentiation, host energy metabolism, and modulate the glucose and lipid homeostasis. BAs also play a crucial role in the regulation of the immune system by suppressing proinflammatory cytokine production[16-18]. Dysregulated BA levels are associated with multiple diseases, including type

2 diabetes, inflammatory bowel disease, steatotic liver disease, and various immune-mediated disorders[19].

We previously reported that at an early age, systemic BAs including those from the secondary BA pathways are associated with progression to islet autoimmunity and onset of type 1 diabetes (T1D)[20]. Recently, MCBAs, in particular phenylalanine, tyrosine and leucine conjugated to cholic acid, were also linked with inflammatory bowel disease and cystic fibrosis[7,8]. However, the early-life dynamics of these MCBAs and their potential role in relation to islet autoimmunity and T1D remain unknown.

Herein, in a longitudinal cohort study of young children (ages 3-36 months), we investigate the trajectories of MCBAs during early life including how these microbial BAs are regulated in children who later progressed to islet autoimmunity. The study included children who developed multiple islet autoantibodies (P2Ab) during follow-up and were thus at high risk for progression to T1D later in life[21], the children who developed a single islet autoantibody (P1Ab) but did not progress to T1D during follow-up, and the controls (CTRs), *i.e.*, children who remained islet autoantibody (AAb) negative during the follow-up. We also investigate how the MCBAs modulate the innate and adaptive immune responses, and how they associate with profiles of gut bacterial communities in a prospective series of samples.

Results

Presence of MCBAs in early life

We prospectively analyzed MCBAs in a longitudinal series of stool samples (n = 303) from 74 children across three study groups: P1Ab, P2Ab, and CTR. For each participant, stool samples were collected at up to six time points for the analysis, corresponding to ages 3, 6, 12, 18, 24, or 36 months (**Fig. 1**). At the time of sample collection, none of the children had

been diagnosed with T1D. We analyzed a total of 110 MCBAs, including BAs (Cholic Acid (CA) Chenodeoxycholic Acid (CDCA) Deoxycholic Acid (DCA) Ursodeoxycholic Acid (UDCA) and Hyodeoxycholic Acid (HDCA)) conjugated to 22 different amino acids (L-DOPA, Cys, Pro, Met, Asn, Citrulline, Ala, Ser, Arg, Gln, Trp, Val, Tyr, Thr, Lys, His, Phe, Asp, Leu, Ile, Glu, Ornithine), using a targeted LC-MS/MS assay. Mixed BA standards and a pooled quality control (QC) sample to match retention times and MS/MS spectra were analyzed alongside longitudinal stool extracts, to obtain a level 1 annotation.

Among the 110 MCBAs analyzed, 78 were present in at least one of the 303 fecal samples (**Fig. 2**). Specifically, MCBAs formed from primary BAs, CA and CDCA, were present in the highest number of fecal samples, while those formed from secondary BAs, *i.e.*, DCA, UDCA, and HDCA, were found in fewer samples. BAs conjugated to Pro, Phe, Glu, His, Ile/Leu, Thr, Trp, Asp, and Tyr were detected in a higher number of samples. In most samples, BA amidates were detected more than once, except for ornithine conjugated to UDCA and L-DOPA conjugated to CA, which were detected in only one fecal sample, respectively. Notably, Ile/Leu isomers were not chromatographically separated. Among all the MCBAs, HDCA conjugates were observed in the lowest number of samples.

MCBA trajectories in early life

To explore the influences of various factors on early-life MCBA profiles, we conducted multivariate analyses using linear mixed-effect models in samples where breastfeeding information was available (n = 242). Age, sex, case status (P1Ab, P2Ab, or CTR), and duration of breastfeeding were treated as fixed effects, with random effects within individual samples/subjects were considered. Age emerged as the most significant determinant of MCBA stool profiles compared to sex, case status, and breastfeeding duration. We identified 48 MCBAs that were influenced by age ($p < 0.05$; **Fig. 3**, **Supplementary Table S1**).

Additionally, ten MCBAs varied across case groups, nine showed differences between sexes, and three were associated with breastfeeding duration (**Supplementary Tables S2-S4**). Specifically, primary BA amidates (*i.e.*, CA and CDCA amidates) decreased after the first year of life (**Fig. 3a-b**). Conversely, MCBAs formed from secondary BAs (DCA and UDCA amidates) exhibited an opposing trend, gradually increasing during the first 18 months of life and remaining stable (DCA-conjugated) or slightly decreasing (UDCA conjugates) at the ages of 24 and 36 months (**Fig. 3c-d**). Regarding HDCA conjugates, a mixed trend was observed. HDCA-Asp and HDCA-Tyr decreased, while HDCA-Glu increased with age (**Supplementary Fig. S1**). Here, UDCA-Val, HDCA-Citrulline, CDCA-Phe, CA-His, UDCA-Glu, CA-Ile, CA-Leu, UDCA-Tyr, and UDCA-Phe were affected by sex. Among the 78 detected MCBAs, CDCA-Ala, CA-Pro, and CDCA-Tyr were influenced by breastfeeding status.

MCBAs in the development of islet autoimmunity

Concentrations of ten MCBAs in stool were found to be different in the children who developed single or multiple autoantibodies during follow-up (P1Ab or P2Ab) when compared to the children in the CTR group who were antibody negative ($p < 0.05$; **Fig. 4**, **Supplementary Table S2**). This included two CA conjugates (CA-Citrulline and CA-Cys), three CDCA conjugates (CDCA-Ala, CDCA-Ser, CDCA-Tyr), four DCA conjugates (DCA-Ile, DCA-Leu, DCA-Pro, DCA-Val), and one UDCA conjugate, UDCA-Asn (**Fig. 3a-d**, **Supplementary Table S2**). **Fig. 4** shows the beta coefficients from the linear model. We found that all of the DCA conjugates were lower in the P1Ab group when compared to the CTR, while CDCA-Tyr showed the opposing trend (**Fig. 4a**, **c**, and **d**). Likewise, the BA conjugates UDCA-Asn, CDCA-Ser, CA-Citrulline, and CDCA-Ala remained downregulated in the P2Ab group when compared to the CTR group, with the exception of CA-Cys (**Fig. 4b**,

e and **f**).

We also compared the MCBA stool profiles separately in different age cohorts. There was no persistent trend with respect to MCBA differences between the study groups (P1Ab, P2Ab, or CTR) at individual time points. However, we found that 22 conjugates (12 CA-conjugates, five CDCA-conjugates, four DCA-conjugates, one UDCA-conjugate) remained altered between the P1Ab and CTR groups at either 6, 12, 18, or 36 months of age. Specifically, the CA-conjugates were higher in the P1Ab group (18 months of age), while the DCA-conjugates were generally lower, and other conjugates showed no consistent trend. Additionally, 9 MCBAs were different between the P2Ab and CTR at either 6, 12, or 18 months of age. Notably, UDCA-Asn remained lower in the P2Ab group compared to the CTR group at 12 months of age.

MBCA concentrations in the stool are associated with the specific gut microbes

Gut microbes have the ability to conjugate amino acids to BAs. Thus, our next objective was to determine if the microbial conjugated BA profiles, which differed between the study groups (**Fig. 4**), were associated with bacterial relative abundance in the paired longitudinal metagenomes (n = 110). Using Spearman correlation analysis, we investigated associations between bacterial composition at stain level and ten MCBA profiles that were persistently altered in the islet autoantibody group. Our analysis revealed that deoxycholic acid (DCA) conjugates exhibited the most frequent associations with the paired microbial strains (**Fig. 5**). In total, we identified 56 bacteria that were positively or inversely associated with at least one of the DCA conjugates. Notably, *Eubacterium eligens* remained strongly associated with DCA-Val, DCA-Leu, and DCA-Ile, while *Roseburia intestinalis* showed the strongest positive association with DCA-Pro. Furthermore, we observed that many strains possessing

BA-transforming capabilities were associated with DCA-Val, DCA-Leu, and DCA-Ile. These strains included *Alistipes onderdonkii*, *Alistipes finegoldii*, *Bacteroides finegoldii*, and *Bacteroides xylanisolvens*. Thirty-three of the gut bacterial strains were predominantly positively associated with CDCA conjugates, including *Ruminococcus gnavus*, *Escherichia coli*, *Phascolarctobacterium succinatutens*, *Erysipelotrichaceae bacterium*, *Bifidobacterium breve*, *Veillonella atypica*, and *Lachnospiraceae bacterium* (**Fig. 5** and **Supplementary Table S5**). Meanwhile, *Ruminococcus gnavus*, *Escherichia coli* along with three *Veillonella spp*. were inversely associated with DCA conjugates (**Supplementary Table S5**). We also found that UDCA-Asn was associated with *Prevotella oralis*, *Actinobaculum schaalii*, *Haemophilus parahaemolyticus*, and *Morganella morganii*. Among all conjugates, CA-Cys exhibited the strongest positive association with *Sutterella wadsworthensis* (r = 0.71).

In addition, we leveraged microbeMASST (https://masst.gnps2.org/microbemasst/) to investigate those MCBAs that were persistently altered in the islet autoantibody group and to identify possible microbial producers[22]. We found that CDCA-Ala was predominantly found in cultures of *Clostridium scindens* ATCC 35704 and *Clostridium sordelli AO32*, UDCA-Asn in *Ruminococcus gnavus* ATCC 29149.CA-Citrulline was found in a diverse number of microbial cultures, including *Enterococcus faecium* 513, *Bifidobacterium breve* HPH0326, and 21 others microbial strains (**Supplementary Table S6**). Likewise, DCA-Ile/Leu was detected in *Escherichia coli*, *Bifidobacterium adolescentis* L2-32, *Bifidobacterium angulatum* F16_22, and several other strains (**Supplementary Table S6**). Likewise microbial candidates for CDCA-Ser and DCA-Val are listed in **Supplementary Table S6**.

Based on multi-omics analysis and strain availability, we cultured nine predicted BA metabolizers that were altered in P2Ab group (*Alistipes onderdonkii*, *Ruminococcus bromii*, *Clostridium bartlettii*, *Bacteroides vulgatus*, *Bacteroides wadsworthia*, *Coprococcus comes*,

Parabacteroides distasonis, *Bacteroides intestinalis* and *Eggerthella lenta*) in brain heart infusion media supplemented with conjugated UDCA-Asn, UDCA, and Asn, as well as in control media without supplements (**Supplementary Fig. S2**). After 72 hours of culturing, samples were analyzed using LC-MS/MS and compared to the UDCA-Asn reference. In monocultures of *Clostridium bartlettii*, Bacteroides vulgatus, *Coprococcus comes* had UDCA-Asn deconjugation potential. *Clostridium bartlettii*, and *Coprococcus comes*, Ruminococcus bromii had conjugation potential. In addition, Alistipes onderdonkii and *Eggerthella lenta* supplemented with UDCA-Asn, we observed an increase in the UDCA-Asn signal (**Supplementary Fig. S2**). These findings confirms that specific human gut microorganisms are capable of producing these novel conjugated MCBAs.

MBCAs modulate host immune responses

We further explored the potential of MCBAs to modulate the innate immune system, specifically examining whether the conjugated form of available UDCA and the unconjugated BAs UDCA and DCA exert immunomodulatory properties *in vitro* in response to lipopolysaccharide (LPS) using human whole blood cultures. Whole blood from healthy donors were diluted 1:1 in WB-STIM buffer (Cytodelics AB, Stockholm) and stimulated with a Salmonella derived form of immunostimulatory LPS. We then measured three signaling pathways downstream of the TLR4 receptor in a time-course detecting phosphorylation of p38 (pp38) and ERK (pERK), as well as degradation of IkBa, at the single-cell level by intracellular staining and flow cytometry (*Fig. 6a*). To robustly compare immunomodulatory effects of different BA conjugates across the entire time course we calculated the area under the curve (AUC) for each signaling intermediate and compared theses across pretreatment conditions and used combinations of small molecule inhibitors of ERK and p38 phosphorylation as negative controls. The different UDCA conjugates exerted variable

effects (**Fig. 6b**). Unconjugated UDCA, Asn-UDCA, and Cit-UDCA partially inhibited LPS-induced signaling across all three pathways investigated, while Trp-UDCA, Asp-UDCA, and Glu-UDCA enhanced the LPS-induced signaling (**Fig. 6b**), indicating immunomodulatory potentials of the newly discovered MCBAs. DCA strongly inhibited LPS induced signaling across all three signaling pathways in human blood monocytes and was even more efficient than either of the conjugated UDCAs tested (**Fig. 6c**). Given the abundance of LPS in the human intestine and the previously reported role of LPS exposure early in life in relation to T1D development[23] it is intriguing that BA modulation of LPS-induced responses offers an additional, previously unrecognized layer of regulation for the establishment of healthy immune-microbe interactions early in life.

Next, we sought to investigate whether MCBAs interact with the adaptive immune system. Specifically, we examined the effects of MCBAs during the early *in vitro* differentiation stages of human Th17 and *in vitro* induced Treg (iTreg) cells. Naive CD4+ CD25- T cells were isolated from human cord blood and cultured with MCBAs under conditions promoting Th17 and Treg differentiation (**Fig. 7a**). We screened three MCBAs—UDCA-Asn, CDCA-Tyr and CDCA-Ser, as well as unconjugated UDCA, which we identified as altered in the P2Ab and P1Ab groups compared to the CTR (**Fig. 4**). Our results showed that UDCA-Asn and CDCA-Ser enhanced Th17 cell differentiation, leading to increased IL-17a secretion (**Fig. 7b** and **c**). Conversely, these compounds inhibited the differentiation of iTreg cells, as evidenced by a decrease in Foxp3 expression levels (**Fig. 7f-g**). However, while neither unconjugated UDCA nor CDCA-Tyr affected the iTreg population (**Fig. 7h-i**), both compounds reduced IL-17a secretion, thereby impeding Th17 differentiation (**Fig. 7d-e**).

Discussion

By comprehensive analysis of fecal MCBAs in a longitudinal birth cohort, we demonstrated

that MCBAs display specific age-dependent profile during the first three years of life. Certain BA conjugates, specifically primary BA amidates (CA and CDCA), decreased over time, while the secondary BA amidates (DCA and UDCA) increased during early life, before stabilizing. Previous metabolomics studies suggest that age has a significant impact on the longitudinal trajectories of metabolites. Several studies previously found that systemic lipids, including BAs, show distinct age-related trajectories[20,24]. However, to our knowledge, this is the first study to define the trajectories of MCBAs in the human gut during infancy. Bacteria from the human gut produce bile salt hydrolases (BSHs) that re-amidate BAs to generate MCBAs[11,12]. During the first three years of life, gut microbial maturation is a dynamic process, which can consequently shape the MCBA profiles and thus explain our observations.

The frequencies of MCBA conjugations varied depending on the specific BA type. Notably, the amidates of HDCA and UDCA were the least frequently observed amino acid conjugations. The composition of BAs in humans is markedly influenced by the gut microbiome[2,25]. Additionally, the human gut microbiome composition differs both within and between individuals[26,27], which can influence the amidate conjugations. Recent *in vitro* studies that investigated bacterial species commonly found in the human intestinal tract, have revealed varying abilities among the gut microbial strains to perform these amidate conjugations[5], thus suggesting a complex interplay between microbial diversity and MCBA metabolism within the gastrointestinal ecosystem. Our data also suggest that gut bacteria contribute to the production of these MCBAs.

We found that children at human leukocyte antigen (HLA) - conferred risk for T1D, who later progressed to single or multiple islet autoantibodies in the follow-up period, have a distinct MBCA profile compared to those who remained autoantibody negative. Previously, we demonstrated that at-risk children exhibit persistently altered levels of both host-derived

systemic BAs (including both glycine and taurine conjugates) and host-microbial BA co-metabolism, compared to children who develop at least one single islet autoantibody or remain negative for islet antibodies during follow-up[20].

BAs are recognized as immunoregulatory metabolites[16] and microbially-derived secondary BAs are crucial for the maintenance of immune system homeostasis[28]. Hang *et al.* identified that microbially-derived secondary BAs, including 3-oxoLCA and isoallo LCA, affect host immune responses by directly modulating T-cell differentiation[16]. Moreover, the levels of isoallo LCA were reported to be altered in patients with IBD[28]. Gentry *et al.* discovered that some of these newly discovered MCBAs, particularly CA conjugated to Glu, Ile/Leu, Phe, Thr, Trp, or Tyr, could be associated with IBD. In the latter study, for CDCA-Met, DCA-Met, CDCA-Phe, CDCA-Trp, and CDCA-Tyr, a notable increase in the levels of interferon-γ (IFNγ) was reported, with CDCA-Trp showing a sixfold rise[8]. IFNγ is a cytokine with essential roles in regulating immune homeostasis and inflammatory responses in humans. In line with others, our data suggest that MCBAs control host immune responses by directly influencing *in vitro* cell differentiation process of Th17 and Treg cells. Given the critical roles of Th17 and Treg cells in inflammatory diseases and their close association with gut bacteria[29], our study suggests novel molecular pathways linked to emerging microbiota-derived bioactive compounds that modulate T-cell function.

The sensing of LPS and the degree to which different LPS variants stimulate immune cell responses early in life has previously been associated with T1D in this cohort[23]. Here we show that secondary BAs variably modulate LPS-induced signaling, thus providing an additional layer of regulation conferred by microbiota members. We demonstrate that conjugated forms of UDCA display varying effects on the activation of monocytes, following exposure to immunostimulatory LPS derived from Salmonella *in vitro*. These results indicate

that the presence of immunostimulatory *vs.* immunoregulatory LPS variants in variable proportions, together with the abundance of immunoregulatory BA conjugates, determine the overall potential of LPS to trigger inflammatory responses. This complements the actions of other immunoregulatory metabolites such as indole-3-lactic acid and other tryptophan metabolites, which are also known to modulate inflammatory responses and T cell development in human newborns[30]. Thus, we suggest that disturbances in the gut microbiome-BA axis during early life alters immunomodulation and potentially contribute to the initiation and/or progression of islet autoimmunity.

Taken together, our study presents the first known exploration of MCBAs dynamics in early life, revealing their potential role in shaping islet autoimmunity. Our findings show that MCBA levels in stool are closely linked with the gut microbiome and influence the differentiation of Th17 and Treg immune cells, highlighting MCBAs as important modulators of immune development.

Methods

Study subjects

The DIABIMMUNE study recruited 832 families in Finland (Espoo), Estonia (Tartu), and Russia (Petrozavodsk) with infants carrying HLA alleles which confer risk for islet autoimmunity and T1D. The subjects involved in the current study were chosen from the subset (n = 74) of available samples (matched serum and stool) in the international DIABIMMUNE study. This comprises children who progressed to at least a single AAb (P1Ab, n = 24) or to multiple islet AAb (P2Ab, n = 12), and controls (CTRs, n = 38), *i.e.*, the children who remained islet AAb-negative during the follow-up. Samples were collected longitudinally at 3, 6, 12, 18, 24 and 36 months from each child. Here no prior sample-size

estimation was performed. The study groups were matched for HLA-associated diabetes risk, gender, country and period of birth.

This study was conducted according to the guidelines in the Declaration of Helsinki. The Ethics and Research Committee of the participating Universities and Hospitals approved the study protocol. All families provided written informed consent prior to sample collection. The EDIA study was approved by the Ethical Committee, Tampere University Hospital (approval number R11166).

Quantification of bile acids

The BAs were measured in fecal sample as described previously[20]. All fecal sample were freeze-dried prior to extraction to account for the inconsistency in the fecal water content and dry weight in the stool. The fecal homogenate was prepared using fecal samples and ethanol in 2:1 w/v ratio. The MCBAs were extracted by adding a volume of 10 µL of fecal homogenate to 200 µL of crash solution containing MeOH and 25ppb of internal standard mix: CA-d4, CDCA-d4, UDCA-d4, DCAd4, LCA-d4, TCA-d4, GUDCA-d4, GCDCA-d4, GDCA-d4, GCA-d4, TLCA-d4, TDCA-d4, TCDCA-d4, TUDCA-d4, TCA-d4 and GLCA-d4. The samples were filtered using a 96-Well protein precipitation filter plate (Supelco). The filtrate was transferred to a glass vial and dried with N2 flow at 45°C and resuspended in 50 µL of resuspension solution (methanol:water 4:6). For QC samples, a pooled sample was prepared in the same way.

The analyses were performed on a Sciex 6600 quadrupole time-of-flight mass spectrometer (Sciex, United States) coupled to a Sciex exion LC system. Separation was performed on a ACQUITY PREMIERE HSS T3 (2.1 × 100 mm, 1.8 µm) column, Waters using gradient elution. The eluents were water/methanol (7:3 v/v) with 2mM ammonium acetate (A) and

methanol with 2mM ammonium acetate (B), the flow rate was 0.4 mL/min, the injection volume was 5 µL and the column was kept at 45°C. The gradient program was 0–0.5 min: 5% B, 4 min: 60 % B; 11-16 min; 100% B, 16.1-17 min 5% B. The mass spectrometer was operated in negative mode. Full scan in the range 100-800 m/z was used for all samples, except QC samples where information dependent analysis was performed. The source paramters CUR, GS1, GS2, ISFV and TEM were 25, 40, 40, 4500 and 650 respectively. The accumulation time was 250 ms. The mass was calibrated using sodium formate clusters between every five samples.

Identification of MCBAs was carried out by comparing retention time and accurate mass (mass error <10ppm) to authentic standards, which were synthesized as described previously[8]. Further, for the QC samples, the fragmentation pattern was compared with the authentic standards. Data was processed using SciexOS (3.0). Results were reported as the peak area of the MCBAs normalized with the corresponding internal standard. For quality control, we randomized the order of samples and injected 1) pooled quality control (QC) 2) a blank sample and 3) a known standard every 10 samples. In addition to that the samples were blinded to the person preparing and running the experiments.

Analysis of islet autoantibodies

Four diabetes-associated autoantibodies were analyzed from each serum sample with specific radiobinding assays: insulin autoantibodies (IAA), glutamic acid decarboxylase antibodies (GADA), islet antigen-2 antibodies (IA-2A), and zinc transporter 8 antibodies (ZnT8A) as described previously[20]. The cut-off values for autoantibody positivity were based on the 99th percentile in non-diabetic children and were 2.80 relative units (RUs) for IAA, 5.36 RU for GADA, 0.78 RU for IA-2A and 0.61 RU for ZnT8A.

Gut microbiome analysis by shotgun metagenomics sequencing

Metagenomic shotgun sequencing and data processing were conducted as previously described[20,23,31-33]. Raw metagenomic sequencing data was retrieved from (https://diabimmune.broadinstitute.org/) (NCBI BioProject ID: PRJNA231909). Stool samples (n = 110) were common between the published metagenomics data and the stool BAs measured in the present study. Metagenomic data from the matched samples (n = 110) were considered for further analysis. As stated earlier[20], host genome–contaminated reads and low-quality reads are already removed from the raw sequencing data using kneadData v0.4. Taxonomic microbiome profiles were determined using MetaPhlAn2 using default parameters as described[20].

Bacterial cultures screening

All bacteria cultures were started from glycerol stock and incubated at 37 °C for three days in an anaerobic chamber (10% CO2, 7.5% H2, 82.5% N2) in a filtered BHI medium at a pH adjusted to 7.2 using 5 N NaOH. After the cultures were normalized at OD_{600} = 0.02, 200 µL of bacterial suspension was added in triplicates in a 96-well plate with and without 100 µM of Asn-UDCA and incubated for 72 h at 37 °C. Bacterial cultures were also extracted at the start of the experiment (0 h) to establish a baseline. Following bacterial growth, 200 µL of culture was transferred to a new 2 mL deep-well plate, 600 µL of 50% MeOH/H_2O was added and incubated overnight at 4°C. Samples were centrifuged at 2000 RPM for 10 min and 200 µL was transferred into a deep-well plate and dried overnight in a CentriVap and stored at -80°C until LC-MS/MS analysis.

LC-MS/MS analysis of bacterial culture

The samples were resuspended in 200 µL of 50% MeOH/H_20 with sulfadimethoxine as

internal standard and incubated overnight at -20 °C. Samples were centrifuged at 2000 RPM for 10 min and 150 µL was transferred to a shallow 96 well-plate. A pooled sample was created for quality control. Samples were randomized and analyzed using a Vanquish UHPLC (ultra-high performance liquid chromatography) system coupled to an Orbitrap Exploris 240 mass spectrometer (Thermo Fisher Scientific). The chromatographic separation was achieved by using a Phenomenex polar C18 column (2.6 µm particle size, 100 x 2.1 mm) and the mobile phase consisted of H_2O + 0.1% FA (solvent A) and ACN + 0.1% FA (solvent B). Five microliters of samples were injected and eluted using the following gradient: 0-0.5 min 5% B, 0.5-1.1 min 5% B, 1.1-7.5 min 40% B, 7.5-8.5 min 99% B, 8.5-9.5 min 99% B, 9.5-10 min 5% B, 10-10.5 min 5% B, 10.5-10.75 99%B, 10.75-11.25 99% B, 11.25-12 min 5% B. Data-dependent acquisition (DDA) mode was used to acquire the MS/MS data using positive electrospray ionization (ESI+). Sheath gas was set to 50 L/min, aux gas flow rate was set at 10 L/min, and sweep gas set to 1. The spray voltage was 3.5 kV, ion transfer tube 325 °C, and vaporizer temperature 350 °C. AcquireX Deep Scan method was enabled, and an exclusion list was created by injecting four times the pooled sample. MS/MS scan range was set to 100 – 1000 *m/z*, RF lens (%) was 70, a resolution of 60,000 with 1 microscans, a charge state of 1, the expected peak width was 6 s, and advanced peak determination was enabled. The automatic gain control (AGC) target was set to standard and with a maximum injection time set to auto. Dynamic exclusion was set to custom, and the following parameters were used: Exclude after number of times: 2, if it occurs within 3 s, with a duration of 4 s. Isotope were excluded. Up to 10 scans per MS1 were collected with a resolution at 200 *m/z* of 22,500 with 1 microscans. The isolation window was set to 1 *m/z*. The AGC target was set to custom with 200% as normalized AGC target with a maximum injection time set to Auto. The scan range mode was set to Auto. The collision energies were set to a stepwise increase of 25, 40, and 60 eV.

LC-MS/MS data processing

Thermo RAW files were converted to .mzML using the ProteoWizard MSconvert software. Feature detection and extraction was performed using MZmine 4.4.3. The batch file used for feature extraction (.mzbatch) can be found on the GitHub webpage. The data was imported using MS1 and MS2 detector using the factor of lowest signal of 3 and 2, respectively. Mass detection was also performed using the above parameters. For the chromatogram builder, minimum consecutive scans was set to 4, intensity to 3E4, height to 1.5E5, and 10 ppm for *m/z* tolerance. Smoothing was applied using the Savitzky Golay algorithm before applying the local minimum feature resolver which had the following parameters: chromatographic threshold set to 90%, minimum search range retention time of 0.05 min, minimum ratio of peak top/edge of 2, and a minimum scans of 4. Then, the ^{13}C isotope filter and isotope finder was applied using a *m/z* tolerance of 3 ppm and retention time tolerance of 0.04 min. Features were aligned using the join aligner module at a *m/z* tolerance of 10 ppm and retention time tolerance set to 0.07 min. Feature list rows filter was applied using two samples or 10%. Peak finder was set to an intensity tolerance of 20%, a m/z tolerance of 10 ppm, retention time tolerance of 0.05 min, and three minimum data points before removing duplicate with m/z tolerance of 1.5 ppm and retention time tolerance of 0.04 min. MetaCorrelate and ion identity networking were performed before exporting the final files. GNPS, SIRIUS, and feature information (legacy MZmine 2) modules were used to generate the feature table containing peak areas, .mgf files, and the content of each feature, respectively, which were necessary for downstream analysis.

Data analysis

All tables generated using MZmine software were imported into R 4.4.1 for downstream analysis.

Synthesis of Asn-UDCA

Solid ursodeoxycholic acid (1.27 mmol, 500 mg, 3 eq.) and 5 mL of DMF were added to a 20 mL glass vial with a stir bar. Next, solid EDC (1.27 mmol, 244 mg, 3 eq.) and neat DIPEA (4.25 mmol, 740 µL, 10 eq.) were subsequently added, and the solution was stirred at room temperature. After 15 min, asparagine (845 µmol, 112 mg, 2 eq.), DMAP were added, and the reaction was stirred overnight. The mixture was then concentrated in vacuo and purified by CombiFlash NextGen 300+ using reversed phase column C18 15.5 g Gold at a flow rate13 mL per min with H_2O (Solvent A) and ACN (solvent B) using the following gradient: 0-6 min, 5% B; 6-10 min, 20% B; 10-17 min 20% B; 17-20 min, 30% B; 20-30 min, 30% B; 30-34 min, 40% B; 34-42 min, 40% B; 42-50 min, 80% B; 50-55 min, 80% B. Asn-UDCA eluted at 34 min, 40% B.

LPS stimulation experiment

Whole blood pre-treatment and stimulation ex-vivo for phospho-Flow cytometry analysis

A blood sample was obtained from healthy adult volunteers in BD Vacutainer Heparin Plasma Tubes. The blood sample was mixed in an equal ratio with WB-STIM buffer (Cytodelics AB) at room temperature. The sample was then split into groups: i) non-treated, ii) vehicle control treated with 0.25% dimethylsulfoxide (DMSO; Sigma-Aldrich), iii) treated with each metabolite (50µM conjugated UDCAs, collaborator, Finland), iv) treated with kinase inhibitors Losmapimod (1500nM, p38 inhibitor) and Selumetinib (1000nM, ERK inhibitor), or v) with combinations of both kinase, used as positive control. Metabolite-treated samples were incubated for 3hs, and kinase inhibitors-treated samples were incubated for 45mins at 37 °C and 5% CO2. After incubation, samples were stimulated ex-vivo with LPS

(1ng/ml) derived from Salmonella enterica for 5 time points (3,7,15, 30, and 60 mins) at 37 °C and 5% CO2. Whole blood samples were then fixed, red blood cells lysed, and samples washed using a Whole blood processing kit (Cytodelics AB)[34] according to the manufacturer's instructions. Completely fixed and processed white blood cells (WBCs) (0.5–1 × 10^6 cells per sample) were plated and cryogenically preserved using CRYO#20 buffer (Cytodelics AB).

Detecting intracellular phosphorylated protein using flow cytometry

Cryopreserved cells were quickly thawed at 37˚C and then kept on ice while counted using Cellaca MX (Nexcelom). 0.5- 1 x 10^6 cells were plated to each well in 96 well U bottom plate. Next, the cells were stained overnight at 4 °C with antibodies targeting intracellular antigens (Table 1), washed, and acquired using a Symphony A3 analyzer equipped with HTS system.

$CD4^+CD25^-$ T cell isolation and Th17/iTreg differentiation

Studies with primary human $CD4^+$ T cells were approved by the Finnish Ethics Committee. Oral informed consent was obtained from all donors prior to the onset of the study. Primary human mononuclear cells were isolated from the umbilical cord blood of healthy neonates (Turku University Central Hospital, Turku, Finland) using Ficoll-Paque PLUS (Cytiva, Cat# 17144003) density gradient centrifugation. $CD4^+$ T cells were further enriched using $CD4^+$ Dynal positive selection beads (Invitrogen, Cat# 11331D) followed by $CD25^+$ T cell depletion using the CD25 Microbeads II kit (Miltenyi Biotec, Cat# 130-092-983), according to the manufactures' instructions. Prior to activation, naïve $CD4^+CD25^-$ T cells from different donors, which were highly positive for CD45RA (FITC anti-CD45RA, BD Biosciences Cat# 555488, RRID:AB_395879) and negative for CD45RO (PE anti-CD45RO, BD Biosciences, Cat# 555493, RRID:AB_395884), were characterized by flow cytometry and were pooled.

Th17 cell differentiation was performed as described previously[35,36]. In brief, $CD4^+CD25^-$ T cells were activated with plate-bound anti-CD3 (3.75 µg/mL; Beckman Coulter, Cat# IM1304; RRID:AB_131612) and soluble anti-CD28 (1 µg/mL; Beckman Coulter, Cat# IM1376; RRID:AB_131624) in X-vivo 20 serum-free medium (Lonza), supplemented with L-glutamine (2 mM, Sigma-Aldrich) and antibiotics (50 U/mL penicillin plus 50 µg/mL streptomycin; Sigma-Aldrich). Th17 cells were cultured in the presence of IL6 (20 ng/mL; Roche, Cat# 7270-IL), IL1β (10 ng/mL; R&D Systems, Cat# 201-LB) and TGFβ (10 ng/mL; R&D Systems, Cat# 240-B), in the presence of neutralizing anti-IFNγ (1 µg/mL; R&D Systems, Cat# MAB285; RRID:AB_2123306) and anti-IL4 (1 µg/mL; R&D Systems, Cat# MAB204; RRID:AB_2126745) to block Th1 and Th2 differentiation, respectively.

iTreg cells were cultured as described earlier[36,37], with minor changes. Briefly, $CD4^+CD25^-$ T cells were activated with plate-bound anti-CD3 (2.5 µg/mL, Beckman Coulter, Cat# IM1304; RRID: AB_131612) and soluble anti-CD28 (0.5 µg/mL, Beckman Coulter, Cat# IM1376; RRID: AB_131624) in X-vivo 15 serum-free medium (Lonza), supplemented with L-glutamine (2 mM), penicillin (50 U) and streptomycin (50 µg/ml) (all from Biowest). Treg cell differentiation was induced in the presence of TGF-β (10 ng/mL; R&D Systems) and IL2 (12 ng/mL; R&D Systems).

To study the effect of MCBAs on human Th17 and iTreg cell differentiation, 100 µM of either Asn-UDCA, Ser-CDCA, Tyr-CDCA or UDCA (in DMSO), and DMSO as control, were added to the Th17 and iTreg cell culture media at day 0, and cultured for 72 hours. After differentiation, secreted IL-17a levels were determined from Th17 cell-culture supernatants at 72 hours using the human IL-17a DuoSet ELISA kit (R&D Systems, Cat# DY317-05, DY008). For iTreg, intracellular staining of Foxp3 was performed using the eBioscience™ Foxp3/Transcription Factor Staining Buffer Set (Thermo Scientific, Cat# 00-5523-00),

according to the manufacture's protocol. Cells were stained with PE-conjugated Foxp3 antibody (Thermo Fisher Scientific, clone PCH101, Cat# 12-4776-42, RRID: _AB151878) or corresponding isotype control antibody (Thermo Fisher Scientific Cat# 12-4321-42, RRID: AB_1518773). Cells were incubated with fluorochrome-labelled antibody for 30 min at 4ºC. After staining, the cells were washed twice, resuspended in flow buffer (2% FBS/0.1% Na-azide/PBS) and acquired on BD LSRFortessa (BD Biosciences). The data was analyzed with FlowJo software (FlowJo LLC).

Statistical methods

The metabolites data values were log transformed prior to analysis. The difference in the lipidome and metabolome between the studies groups were compared using a multivariate linear model. For longitudinal samples, linear mixed effects models were regressed with fixed effect (~ sex + case + age + breast feeding duration) and random effect ~ (1 | Subject). For age wise comparisons the metabolites were regressed with various factors such as sex, and disease conditions (*e.g.*, P1Ab *vs*. CTR) using MaAsLin2 package in R (BAs ~ sex + case). To subsequently visualize metabolite level, forest plots from the ggplot2 R package were used.

Data and code availability

Metagenomic sequencing data can be downloaded from https://diabimmune.broadinstitute.org/diabimmune/ (NCBI BioProject ID: PRJNA231909). The targeted bile acid metabolomics datasets generated in this study is available atMassIVE Repository (https://massive.ucsd.edu/ProteoSAFe/static/massive.jsp). MassIVE is a community resource developed by the NIH-funded Center for Computational Mass Spectrometry. The data can be accessed directly at GNPS/MassIVE under the accession

number MSV0000XXXX. Any additional information required to reanalyze the data reported in this work paper is available from the lead contact upon request.

Acknowledgments

This study was supported by the “Inflammation in human early life: targeting impacts on life-course health” (INITIALISE) consortium funded by the Horizon Europe Program of the European Union under Grant Agreement 101094099 (to M.O.), the Research Council of Finland funding (no. 363417 to S.L.), the Research council of Finland project grant (grant no. 333981 to M.O.), and the Novo Nordisk Foundation (grant no. NNF20OC0063971 to T.H. and M.O.). Further support was received from the Swedish Research Council (grant no. 2016-05176 to T.H. and M.O.), Formas (grant no. 2019-00869 to T.H. and M.O.), and the Medical Research Funds, Helsinki University Hospital (to M.K.). Further support was received by the InFLAMES Flagship Programme of the Academy of Finland (grant no. 337530). We further acknowledge support by NICHD, P50HD106463, and NIDDK R01DK136117, U24DK133658 (to P.C.D.). RJX funded by NIH (DK 43351, AI172147) and Center for Microbiome Informatics and Therapeutics. Mass spectrometeric analysis was performed at the Turku Metabolomics Centre with the support of Biocenter Finland. The DIABIMMUNE study was supported by the European Union Seventh Framework Programme FP7/2007-2013 under grant agreement number 202063 (to M.K.), Juvenile Diabetes Research Foundation grants 17-2011-529 and 17-2014-305 to R.L., and The Academy of Finland Centre of Excellence in Molecular Systems Immunology and Physiology Research grant Decision number 250114, 2012–2017 (to M.O. and M.K.).

We thank all voluntary blood donors and personnel of Turku University Hospital, Department of Obstetrics and Gynecology, Maternity Ward (Hospital District of Southwest Finland) for the umbilical cord blood collection. We acknowledge Marjo Hakkarainen (Turku

Bioscience Center, University of Turku) for her excellent technical assistance. We thank to Denys Mavrynsky and Tuomas Karskela for the synthesis of pure UDCA conjugates. We also acknowledge Mauricio Caraballo (Skaggs School of Pharmacy and Pharmaceutical Sciences, UC San Diego) and Nicole Avalon (Scripps Institution of Oceanography) for their assistance with the experimental setup for metabolomics and compound purification.

Author contributions

Conceptualization: S.L. and M.O.; Data curation: M.Kr, P.S., and P.K.; Formal analysis: S.L., P.K., T.B, S.Z., and V.C.L.; Funding acquisition: S.L., T.H., M.Kn, and M.O.; Investigation: S.L., P.K., M.Kn., and M.O.; Methodology: L.P., P.K., M.Kr, A.M.D., E.T.L., V.C.L., C.W., R.K., O.R., and T.H.; Resources: T.V., R.J.X., P.D., E.G., P.B., A.A., J.M., A.M., T.L., M.R., K.Z., R.L., and M.Kn; Supervision: M.Kn and M.O.; Writing – original draft preparation: S.L.; Writing – review and editing: All authors.

Ethics declarations

Competing interests: P.C.D. is an advisor and holds equity in Cybele and Sirenas, is a science advisor and holds equity in bileOmix and is a Scientific co-founder, advisor and holds equity to Ometa, Enveda, and Arome with prior approval by UC-San Diego. PCD also consulted for DSM animal health in 2023. The other authors declare no competing interests.

Supplementary information

Supplementary Figure S1. Trajectory of HDCA-conjugates during early life.

Supplementary Figure S2. Microbial production of UDCA-Asn.

Supplementary Figure S3. Manual gating strategy for identifying the CD14+ monocyte population among whole blood white blood cells analyzed by flow cytometry.

Supplementary Tables S1-S5. p-values, regression coefficients, linear mixed-effect models.

Supplementary Table S6. Output from the microbeMASST.

582 **Table 1.** Fluorescent marker antibodies (surface and intracellular) used in flow cytometry.

Fluorophore	Marker	Catalog number	Antibody dilution, times	Clone	Vendor
BV650	CD14	301835	5000	M5E2	Biolegend
PE-Cy7	p-p38	25-9078-42	5000	4NIT4KK	Invitrogen
PE-DAZZLE	p-ERK	369518	5000	6B8B69	Biolegend
488	IκBα	5743S	200	L35A5	Cell signalling Technologies

583

References

1 Chiang, J. Y. Bile acids: regulation of synthesis. *J Lipid Res* **50**, 1955-1966 (2009). https://doi.org/10.1194/jlr.R900010-JLR200

2 Guzior, D. V. & Quinn, R. A. Review: microbial transformations of human bile acids. *Microbiome* **9**, 140 (2021). https://doi.org/10.1186/s40168-021-01101-1

3 Mohanty, I. *et al.* The changing metabolic landscape of bile acids - keys to metabolism and immune regulation. *Nature reviews. Gastroenterology & hepatology* **21**, 493-516 (2024). https://doi.org/10.1038/s41575-024-00914-3

4 Hofmann, A. F. *et al.* A proposed nomenclature for bile acids. *Journal of lipid research* **33**, 599-604 (1992).

5 Lucas, L. N. *et al.* Dominant Bacterial Phyla from the Human Gut Show Widespread Ability To Transform and Conjugate Bile Acids. *mSystems*, e0080521 (2021). https://doi.org/10.1128/mSystems.00805-21

6 Dawson, P. A. & Karpen, S. J. Intestinal transport and metabolism of bile acids. *Journal of lipid research* **56**, 1085-1099 (2015). https://doi.org/10.1194/jlr.R054114

7 Quinn, R. A. *et al.* Global chemical effects of the microbiome include new bile-acid conjugations. *Nature* **579**, 123-129 (2020). https://doi.org/10.1038/s41586-020-2047-9

8 Gentry, E. C. *et al.* Reverse metabolomics for the discovery of chemical structures from humans. *Nature* **626**, 419-426 (2024). https://doi.org/10.1038/s41586-023-06906-8

9 Killenberg, P. G. & Jordan, J. T. Purification and characterization of bile acid-CoA:amino acid N-acyltransferase from rat liver. *The Journal of biological chemistry* **253**, 1005-1010 (1978).

10 O'Byrne, J., Hunt, M. C., Rai, D. K., Saeki, M. & Alexson, S. E. The human bile acid-CoA:amino acid N-acyltransferase functions in the conjugation of fatty acids to glycine. *The Journal of biological chemistry* **278**, 34237-34244 (2003). https://doi.org/10.1074/jbc.M300987200

11 Rimal, B. *et al.* Bile salt hydrolase catalyses formation of amine-conjugated bile acids. *Nature* **626**, 859-863 (2024). https://doi.org/10.1038/s41586-023-06990-w

12 Guzior, D. V. *et al.* Bile salt hydrolase acyltransferase activity expands bile acid diversity. *Nature* **626**, 852-858 (2024). https://doi.org/10.1038/s41586-024-07017-8

13 Schaap, F. G., Trauner, M. & Jansen, P. L. Bile acid receptors as targets for drug development. *Nature reviews. Gastroenterology & hepatology* **11**, 55-67 (2014). https://doi.org/10.1038/nrgastro.2013.151

14 Makishima, M. *et al.* Identification of a nuclear receptor for bile acids. *Science (New York, N.Y.)* **284**, 1362-1365 (1999). https://doi.org/10.1126/science.284.5418.1362

15 Wang, H., Chen, J., Hollister, K., Sowers, L. C. & Forman, B. M. Endogenous bile acids are ligands for the nuclear receptor FXR/BAR. *Molecular cell* **3**, 543-553 (1999). https://doi.org/10.1016/s1097-2765(00)80348-2

16 Hang, S. *et al.* Bile acid metabolites control T(H)17 and T(reg) cell differentiation. *Nature* **576**, 143-148 (2019). https://doi.org/10.1038/s41586-019-1785-z

17 Paik, D. *et al.* Human gut bacteria produce T(H)17-modulating bile acid metabolites. *Nature* **603**, 907-912 (2022). https://doi.org/10.1038/s41586-022-04480-z

18 Song, X. *et al.* Microbial bile acid metabolites modulate gut RORγ(+) regulatory T cell homeostasis. *Nature* **577**, 410-415 (2020). https://doi.org/10.1038/s41586-019-1865-0

19 Hofmann, A. F. The continuing importance of bile acids in liver and intestinal disease. *Archives of internal medicine* **159**, 2647-2658 (1999). https://doi.org/10.1001/archinte.159.22.2647

20 Lamichhane, S. *et al.* Dysregulation of secondary bile acid metabolism precedes islet autoimmunity and type 1 diabetes. *Cell reports. Medicine* **3**, 100762 (2022). https://doi.org/10.1016/j.xcrm.2022.100762

21 Ziegler, A. G. *et al.* Seroconversion to multiple islet autoantibodies and risk of progression to diabetes in children. *JAMA* **309**, 2473-2479 (2013). https://doi.org/10.1001/jama.2013.6285

22 Zuffa, S. *et al.* microbeMASST: a taxonomically informed mass spectrometry search tool for microbial metabolomics data. *Nature microbiology* **9**, 336-345 (2024). https://doi.org/10.1038/s41564-023-01575-9

23 Vatanen, T. *et al.* Variation in Microbiome LPS Immunogenicity Contributes to Autoimmunity in Humans. *Cell* **165**, 842-853 (2016). https://doi.org/10.1016/j.cell.2016.04.007

24 Lamichhane, S. *et al.* Dynamics of Plasma Lipidome in Progression to Islet Autoimmunity and Type 1 Diabetes - Type 1 Diabetes Prediction and Prevention Study (DIPP). *Sci Rep* **8**, 10635 (2018). https://doi.org/10.1038/s41598-018-28907-8

25 Ridlon, J. M., Kang, D. J., Hylemon, P. B. & Bajaj, J. S. Bile acids and the gut microbiome. *Current opinion in gastroenterology* **30**, 332-338 (2014). https://doi.org/10.1097/mog.0000000000000057

26 Olsson, L. M. *et al.* Dynamics of the normal gut microbiota: A longitudinal one-year population study in Sweden. *Cell host & microbe* **30**, 726-739.e723 (2022). https://doi.org/10.1016/j.chom.2022.03.002

27 Zhu, A., Sunagawa, S., Mende, D. R. & Bork, P. Inter-individual differences in the gene content of human gut bacterial species. *Genome biology* **16**, 82 (2015). https://doi.org/10.1186/s13059-015-0646-9

28 Li, W. *et al.* A bacterial bile acid metabolite modulates T(reg) activity through the nuclear hormone receptor NR4A1. *Cell host & microbe* **29**, 1366-1377.e1369 (2021). https://doi.org/10.1016/j.chom.2021.07.013

29 Nagashima, K. *et al.* Mapping the T cell repertoire to a complex gut bacterial community. *Nature* **621**, 162-170 (2023). https://doi.org/10.1038/s41586-023-06431-8

30 Henrick, B. M. *et al.* Bifidobacteria-mediated immune system imprinting early in life. *Cell* **184**, 3884-3898.e3811 (2021). https://doi.org/10.1016/j.cell.2021.05.030

31 Kostic, A. D. *et al.* The dynamics of the human infant gut microbiome in development and in progression toward type 1 diabetes. *Cell host & microbe* **17**, 260-273 (2015). https://doi.org/10.1016/j.chom.2015.01.001

32 Yassour, M. *et al.* Natural history of the infant gut microbiome and impact of antibiotic treatment on bacterial strain diversity and stability. *Science translational medicine* **8**, 343ra381 (2016). https://doi.org/10.1126/scitranslmed.aad0917

33 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). https://doi.org/10.1038/s41564-018-0321-5

34 Brodin, P., Duffy, D. & Quintana-Murci, L. A Call for Blood-In Human Immunology. *Immunity* **50**, 1335-1336 (2019). https://doi.org/10.1016/j.immuni.2019.05.012

35 Buchacher, T. *et al.* PIM kinases regulate early human Th17 cell differentiation. *Cell Rep* **42**, 113469 (2023). https://doi.org/10.1016/j.celrep.2023.113469

36 Sen, P. *et al.* Quantitative genome-scale metabolic modeling of human CD4(+) T cell differentiation reveals subset-specific regulation of glycosphingolipid pathways. *Cell Rep* **37**, 109973 (2021). https://doi.org/10.1016/j.celrep.2021.109973

37 Andrabi, S. B. A. *et al.* Long noncoding RNA LIRIL2R modulates FOXP3 levels and suppressive function of human CD4(+) regulatory T cells by regulating IL2RA. *Proc Natl Acad Sci U S A* **121**, e2315363121 (2024). https://doi.org/10.1073/pnas.2315363121

Figure legends

Figure 1. Outline of the study. We analyzed microbially conjugated bile acids in a longitudinal series of stool samples collected at 3, 6, 12, 18, 24, and 36 months of age from children at HLA-conferred risk for TD, who later developed (i) multiple islet autoantibodies (P2Ab), (ii) single islet autoantibody (P1Ab), or (iii) remained autoantibody negative (CTRs) during the follow-up.

Figure 2. Presence of microbial conjugated bile acids (MCBAs) in fecal DIABIMMUNE samples. The bar in the color map corresponds to the number of fecal samples in which MCBAs were detected.

Figure 3. Trajectories of MCBA in early life. The loess curve plot of MCBAs over time (3, 6, 12, 18, 24 and 36 months) for CA-conjugates, CDCA-conjugates, DCA-conjugates and UDCA-conjugates.

Figure 4. MCBAs in progression to islet autoimmunity. Forest plot illustrating the coefficient estimate of a linear mixed-effects model for individual MCBAs species, with fixed covariates of Case (CTR *vs.* P1Ab, CTR vs P2AB), Age, Sex, and length of breast-feeding accounting for random effects within individual samples. Filled circles with corresponding confidence intervals represent significant MCBA species. Faded circles depict non-significant species. The loess curve plot of MCBAs over time for significant MCBA species obtained in the linear mixed-effects model. The p-values shown are nominal; adjusted p-values (corrected for multiple comparisons using the Benjamini-Hochberg method) are available in **Supplementary Table S2**.

Figure 5. Cross-correlation between the microbes and stool levels of selected MCBAs. Heatmap showing the correlation coefficients of association between microbes and stool

levels of MCBAs that were found altered in progression to islet autoimmunity, in a subset of the sample with available metagenomics data (n = 110). Red color represents positive correlations, while blue represent negative correlations, as determined by spearman rank correlation.

Figure 6. Effects of secondary bile acids on LPS-induced signaling in human monocytes. a. Time-course analysis of LPS-stimulated phosphorylation of p38 (pp38) and ERK (pERK), as well as degradation of IkBa, markers of canonical NF-kB activation, in human monocytes. Stimulation with LPS induces robust activation of pp38 and pERK and IkBa degradation at 30 minutes. Geometric Mean Fluorescent Intensity (MFI) is measured by flow cytometry, and the average between replicates is plotted in time curves. Quantitative differences among treatments were shown by area under the curve (AUC) analysis. Geometric MFI values were used for AUC calculations (n=6). "Inhibitors" serve as a positive control for pathway inhibition, and "No treatment" represents the baseline LPS response. **b**. UDCA and its conjugates (Trp-UDCA, Asp-UDCA, Glu-UDCA, Asn-UDCA, Cit-UDCA) show limited or no inhibition of these pathways. **c**. DCA markedly inhibits signaling across all three pathways. DCA demonstrates significant suppression of LPS-induced signaling. ANOVA and subsequent pairwise comparisons were performed to analyse the impact of treatments for multiple proteins across time. Statistical significance was determined using FDR-corrected p-values. Adjusted p-values are denoted as follows: ***$p < 0.001$, **$p< 0.01$, and *$p < 0.05$.

Figure 7. Microbial conjugated bile acids modulate Th17 and Treg cell differentiation. **a.** Schematic of the Th17 and iTreg differentiation protocol for primary human naïve $CD4^+CD25^-$ T cells isolated from the umbilical cord blood of healthy neonates. $CD4^+CD25^-$ T cells were activated with anti-CD3/anti-CD28 and differentiated into Th17 or iTreg cells in

735 the presence of corresponding cytokines for 3 days. DMSO control or conjugated bile acids
736 (Asn-UDCA, Ser-CDCA, Tyr-CDCA, and unconjugated UDCA) at 100 µM were added on
737 day 0 of differentiation. **b-e.** IL-17a secretion in the supernatant of Th17 cultures treated with
738 Asn-UDCA (**b**), Ser-CDCA (**c**), Tyr-CDCA (**d**), and UDCA (**e**) was quantified on day 3 of
739 differentiation from four biological replicates using ELISA. **f-i.** Intracellular Foxp3 protein
740 expression in iTregs cultured with Asn-UDCA (**f**), Ser-CDCA (**g**), Tyr-CDCA (**h**), or UDCA
741 (**i**) was assessed on day 3 of differentiation by flow cytometry. Geometric mean fluorescence
742 intensity (MFI) values are shown for four biological replicates. Statistical significance was
743 determined using paired, two-tailed Student's t-test.

744

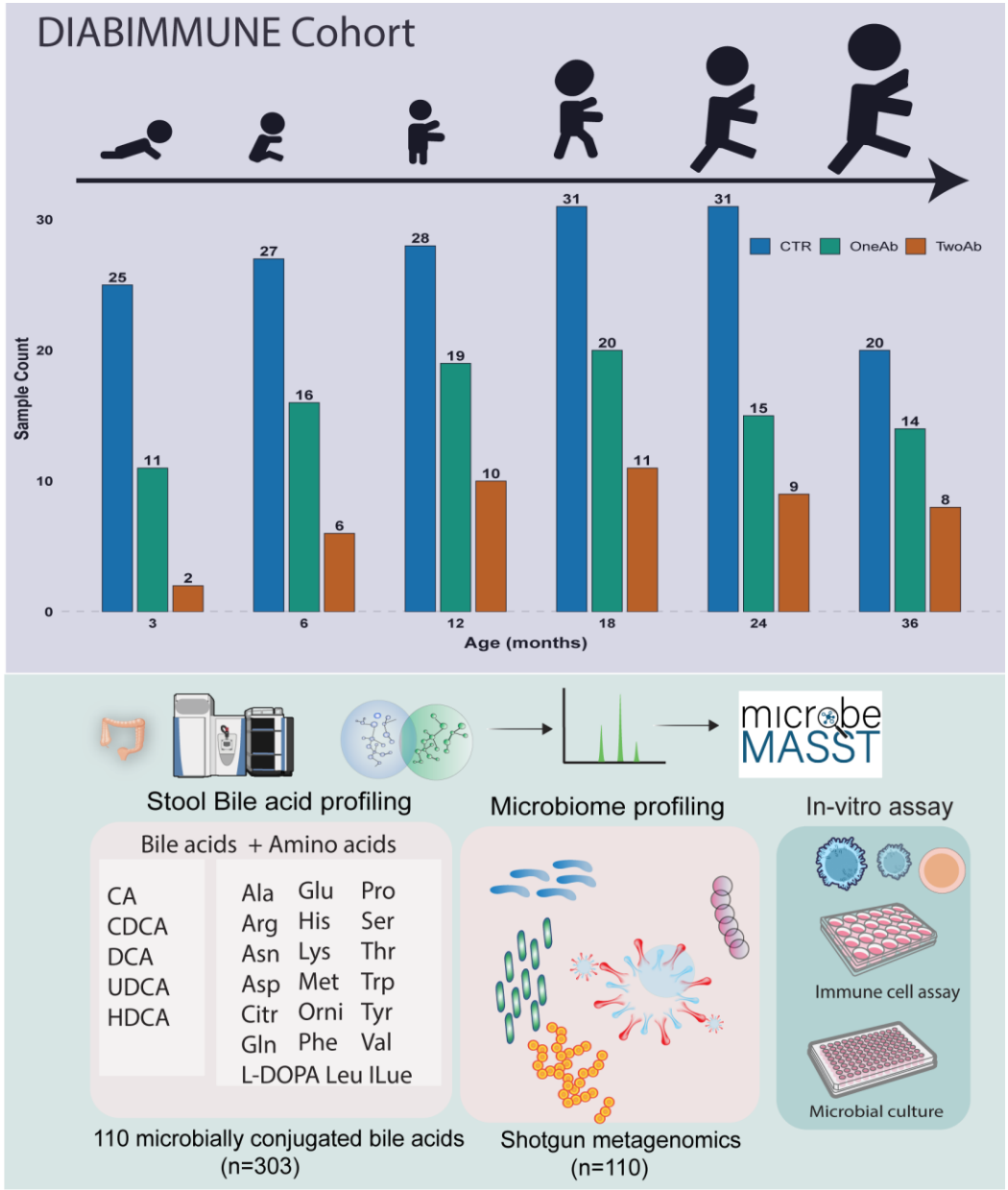

Figure 1. Outline of the study.

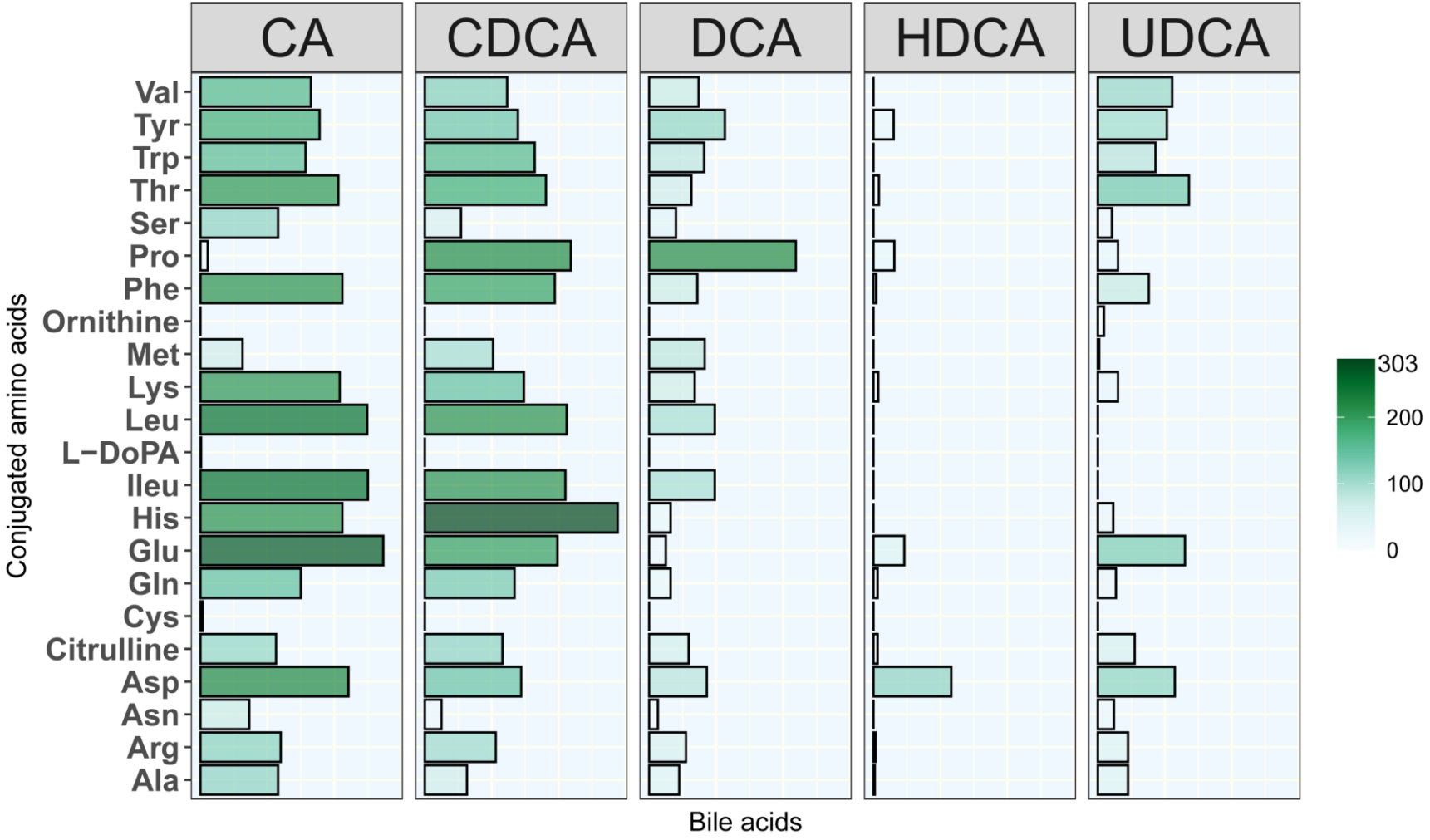

Figure 2. Presence of MCBAs in fecal DIABIMMUNE samples.

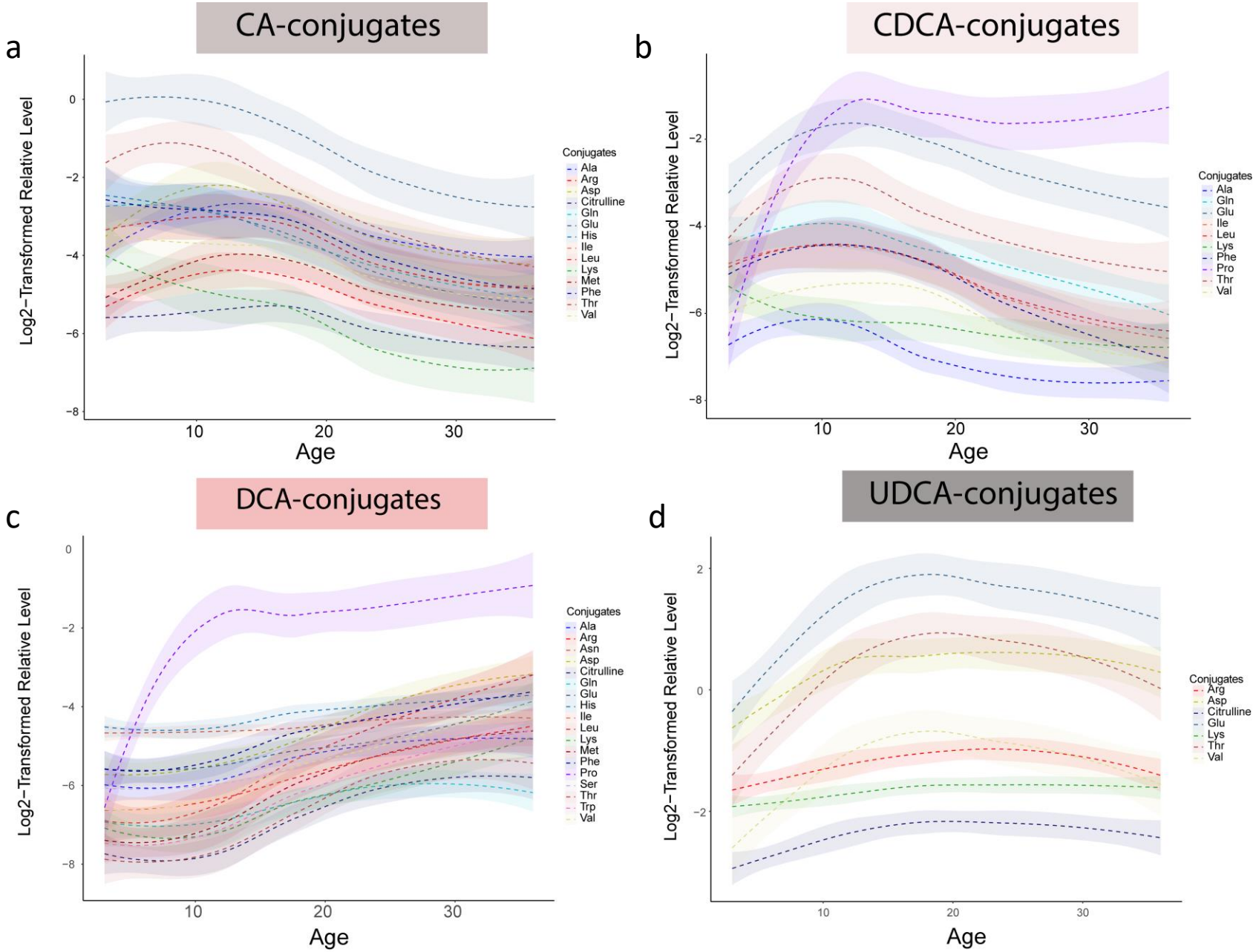

Figure 3. Trajectories of MCBAs in early life.

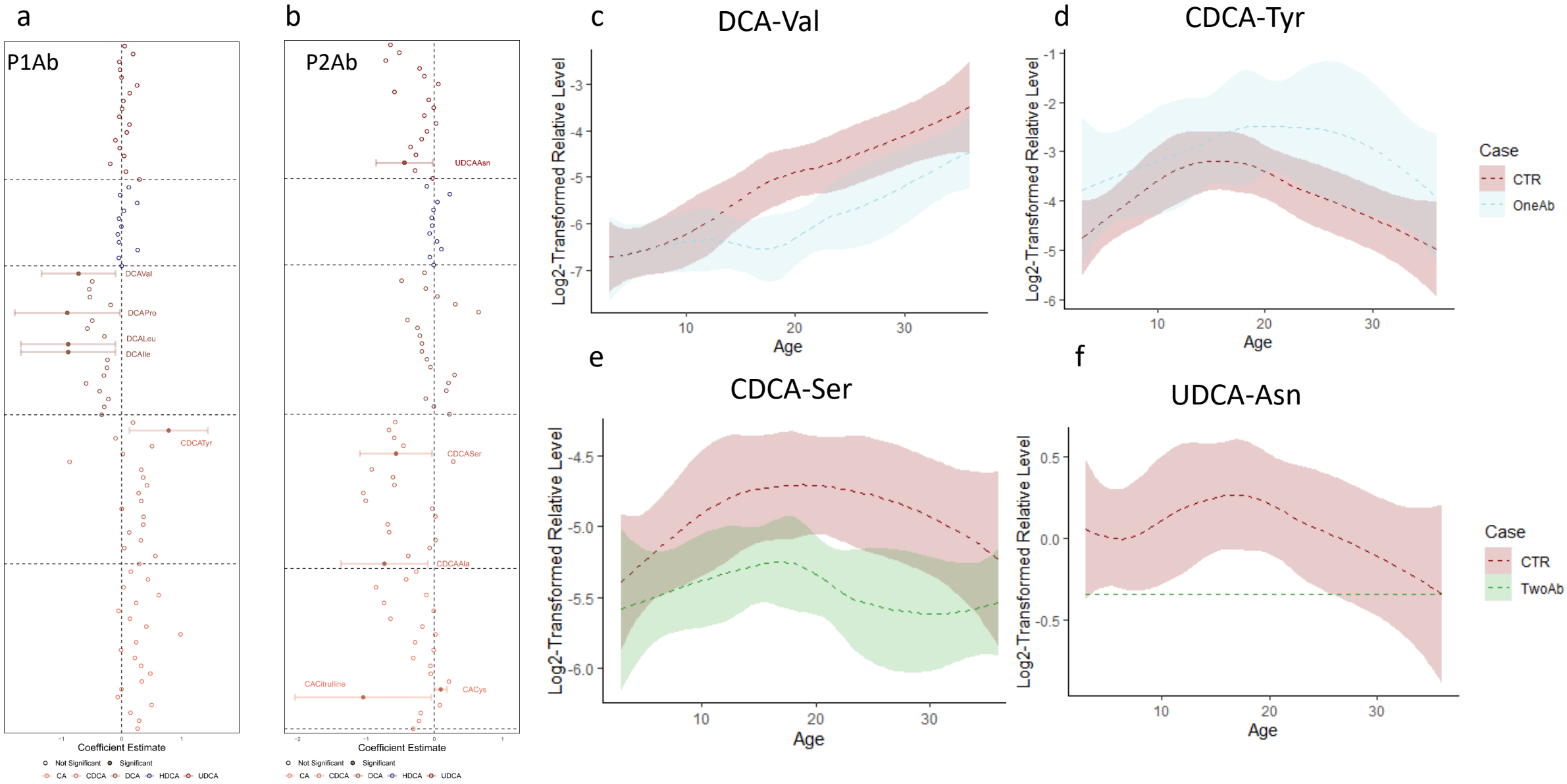

Figure 4. MCBAs in progression to islet autoimmunity.

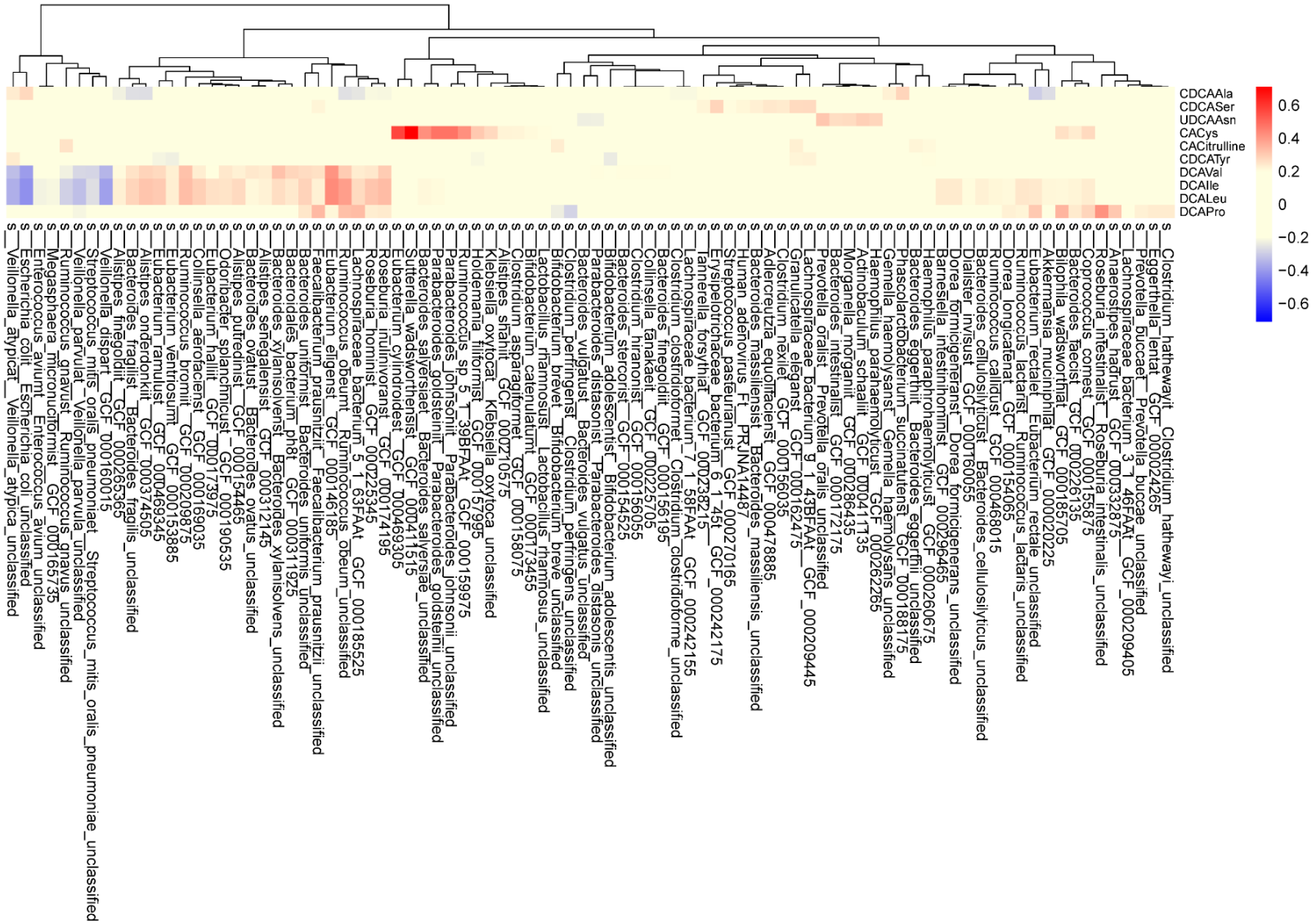

Figure 5. Cross-correlation between the microbes and stool levels of selected MCBAs.

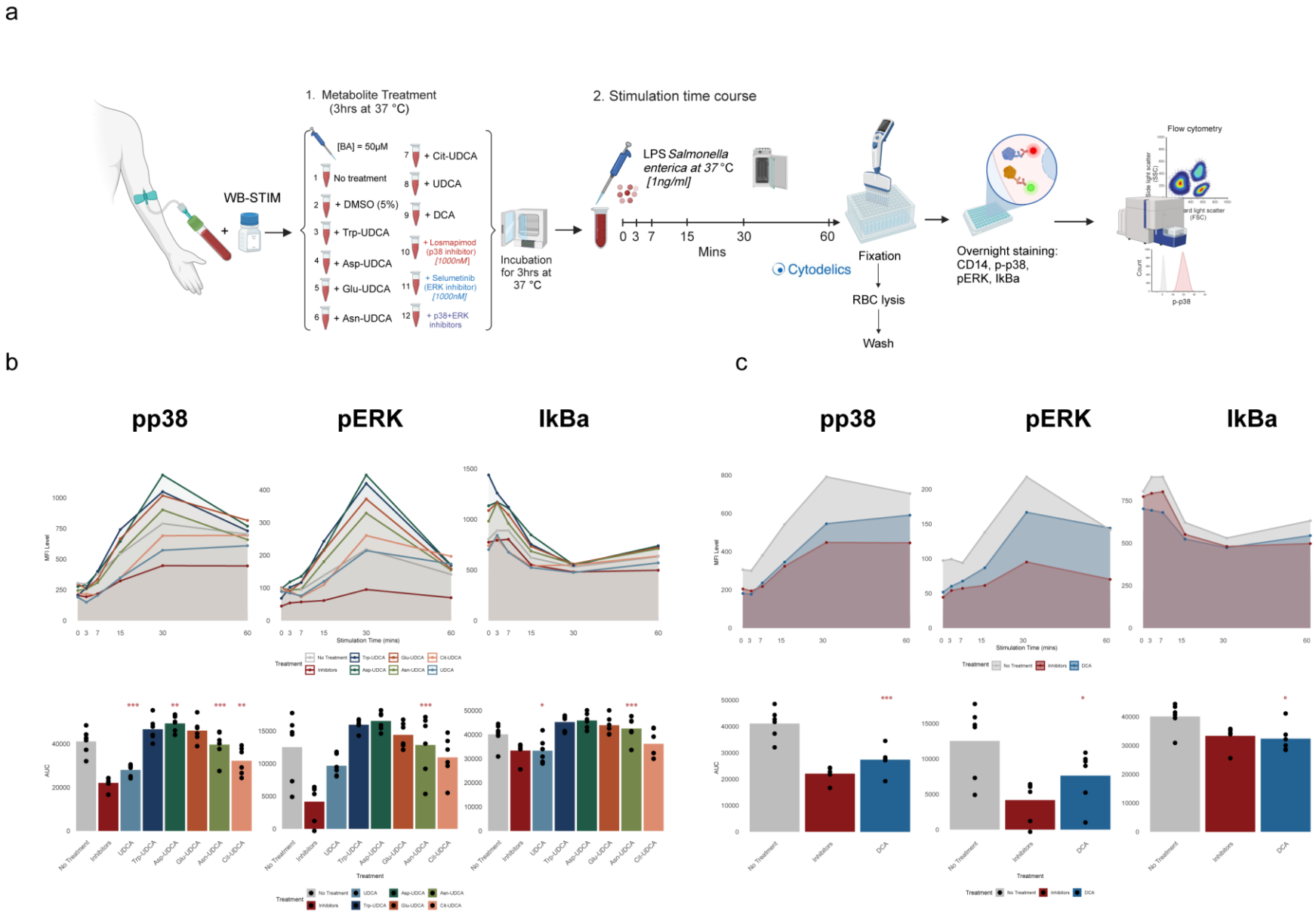

Figure 6. Effects of secondary bile acids on LPS-induced signaling in human monocytes.

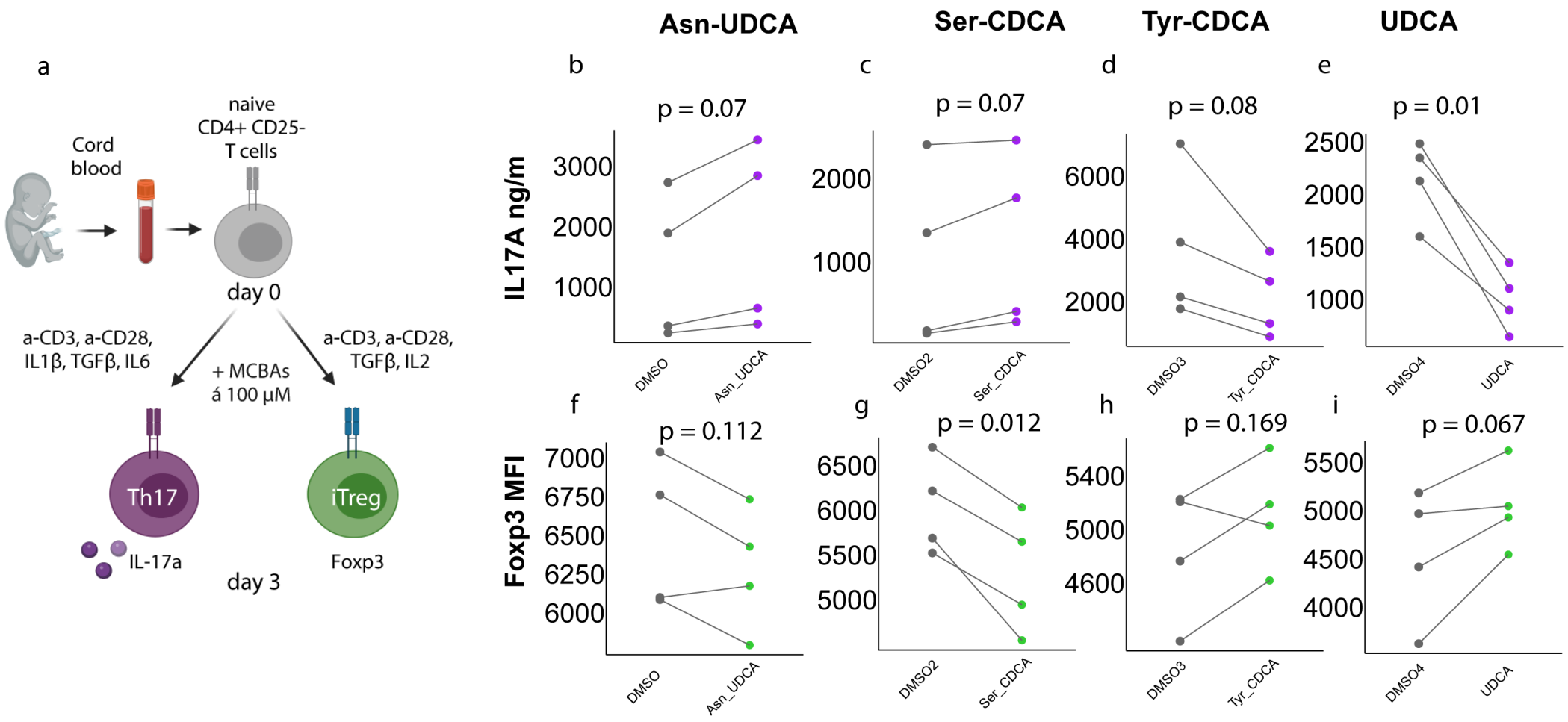

Figure 7. Microbial conjugated bile acids modulate Th17 and Treg cell differentiation.

Supplementary

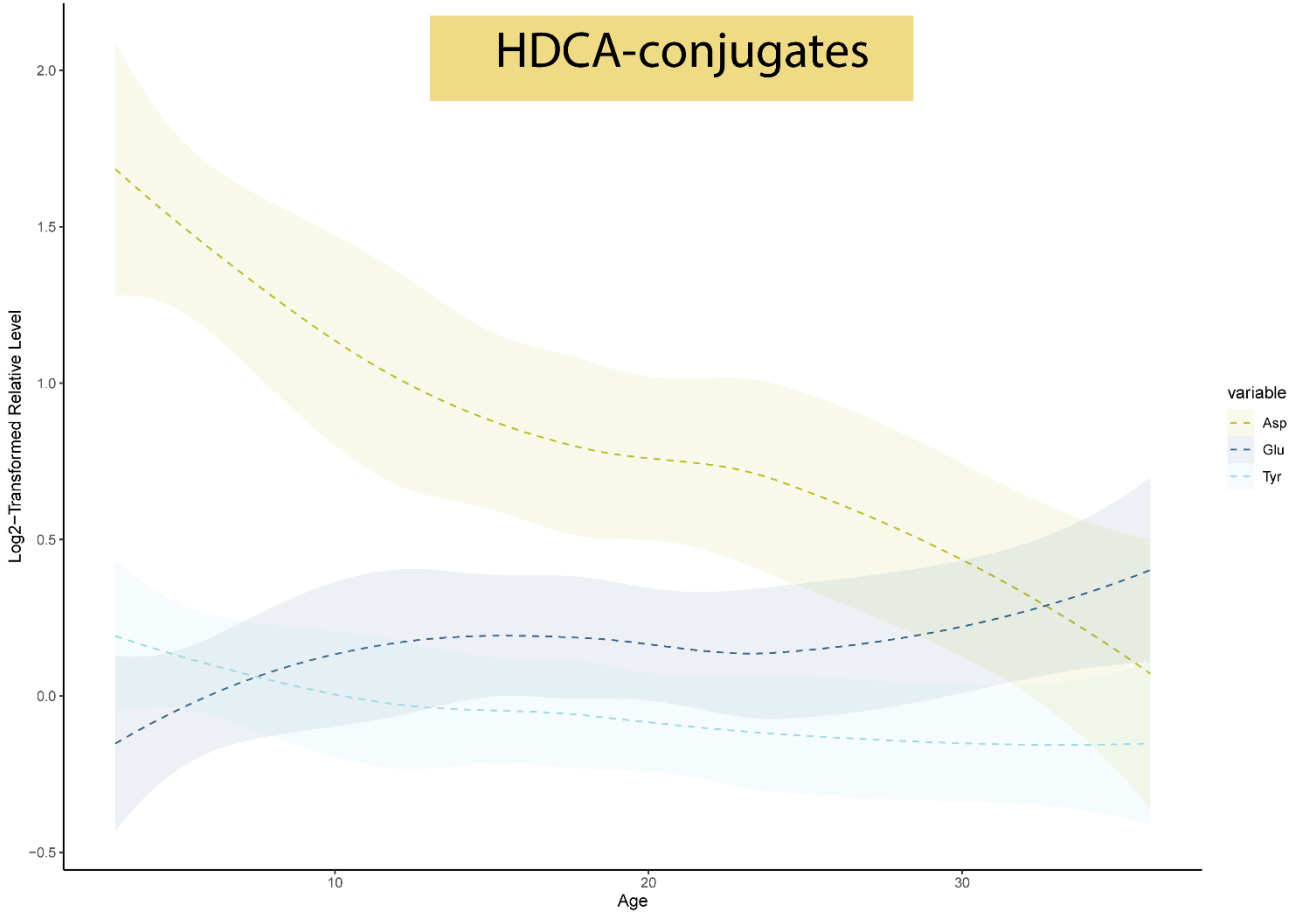

Fig.S1 Trajectories of HDCA conjugated MCBAs in early life.

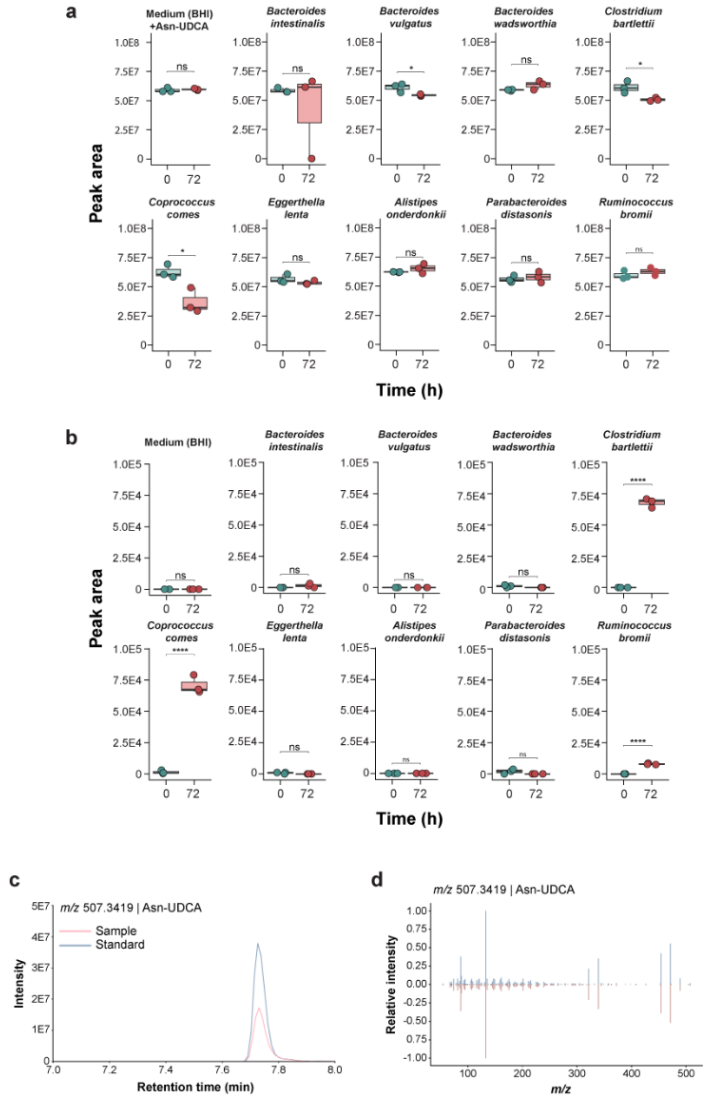

Fig.S2. Microbial production of UDCA-Asn. a. Nine gut microbial strains were cultured in the presence of 200 μM of Asn-UDCA conjugates to assess their deconjugation potential. b. Asn and UDCA were added to the culture medium at 200 μM to evaluate the formation of Asn-UDCA. c. The retention time (min) was compared between the authentic standard and the biological samples. d. Mirror plot displays the MS/MS spectrum similarity between the authentic standard and the biological sample.

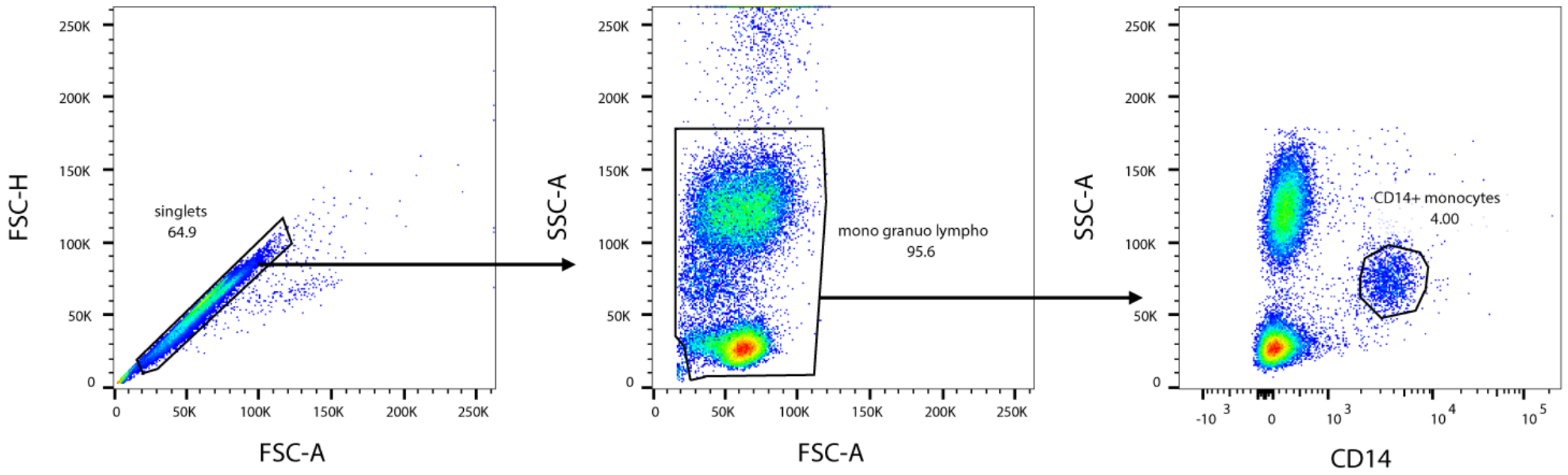

Fig. S3: Manual gating strategy for identifying the CD14+ monocyte population among whole blood white blood cells analyzed by flow cytometry.