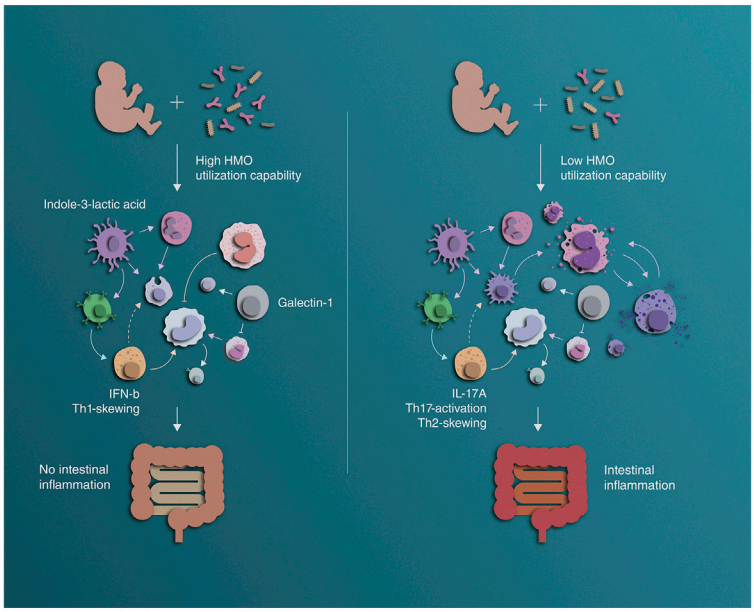
Cell

Bifidobacteria-mediated immune system imprinting early in life

Graphical abstract



Highlights

- An ordered sequence of immune changes after birth driven by microbial interactions
- Lack of gut bifidobacteria and HMO-utilization genes correlates with systemic inflammation
- Feeding *B. infantis* EVC001 upregulates IFNβ and silences intestinal Th2 and Th17
- EVC001-associated indole-3-lactic acid upregulates inhibitory galectin-1 in T cells

Authors

Bethany M. Henrick, Lucie Rodriguez, Tadepally Lakshmikanth, ..., J. Bruce German, Steven A. Frese, Petter Brodin

Correspondence

bhenrick2@unl.edu (B.M.H.), petter.brodin@ki.se (P.B.)

In brief

A lack of bifidobacteria and/or their genes required for the utilization of human milk oligosaccharides from breast milk is associated with systemic inflammation and immune imbalance early in life. Infant supplementation of *Bifidobacterium infantis* EVC001 shows promise in mitigating this by reducing Th2 and Th17 cytokines in the intestine through upregulation of the immunoregulatory factor galectin-1.

Henrick et al., 2021, Cell *184*, 3884–3898
July 22, 2021 © 2021 The Author(s). Published by Elsevier Inc.
https://doi.org/10.1016/j.cell.2021.05.030

CellPress

CellPress
OPEN ACCESS

Cell

Article

Bifidobacteria-mediated immune system imprinting early in life

Bethany M. Henrick,[1,2,*] Lucie Rodriguez,[3] Tadepally Lakshmikanth,[3] Christian Pou,[3] Ewa Henckel,[3,4,5] Aron Arzoomand,[3] Axel Olin,[3] Jun Wang,[3] Jaromir Mikes,[3] Ziyang Tan,[3] Yang Chen,[3] Amy M. Ehrlich,[1] Anna Karin Bernhardsson,[3] Constantin Habimana Mugabo,[3] Ylva Ambrosiani,[4] Anna Gustafsson,[4,5] Stephanie Chew,[1] Heather K. Brown,[1] Johann Prambs,[1] Kajsa Bohlin,[4,5] Ryan D. Mitchell,[1] Mark A. Underwood,[6,7] Jennifer T. Smilowitz,[6,8] J. Bruce German,[6,8] Steven A. Frese,[2,9] and Petter Brodin[3,10,11,*]

[1]Evolve BioSystems, Inc., Davis, CA 95618, USA
[2]Department of Food Science and Technology, University of Nebraska, Lincoln, Lincoln, NE 68588-6205, USA
[3]Science for Life Laboratory, Department of Women's and Children's Health, Karolinska Institutet, 17121 Solna, Sweden
[4]Department of Clinical Science, Intervention and Technology, Karolinska Institutet, 14152 Stockholm, Sweden
[5]Department of Neonatology, Karolinska University Hospital, 14186 Stockholm, Sweden
[6]Foods for Health Institute, University of California, Davis, Davis, CA 95616, USA
[7]Department of Pediatrics, University of California Davis Children's Hospital, Sacramento, CA 95817, USA
[8]Department of Food Science and Technology, University of California, Davis, Davis, CA 95616, USA
[9]Department of Nutrition, University of Nevada, Reno, Reno, NV 89557, USA
[10]Pediatric Rheumatology, Karolinska University Hospital, 17176 Solna, Sweden
[11]Lead contact
*Correspondence: bhenrick2@unl.edu (B.M.H.), petter.brodin@ki.se (P.B.)
https://doi.org/10.1016/j.cell.2021.05.030

SUMMARY

Immune-microbe interactions early in life influence the risk of allergies, asthma, and other inflammatory diseases. Breastfeeding guides healthier immune-microbe relationships by providing nutrients to specialized microbes that in turn benefit the host's immune system. Such bacteria have co-evolved with humans but are now increasingly rare in modern societies. Here we show that a lack of bifidobacteria, and in particular depletion of genes required for human milk oligosaccharide (HMO) utilization from the metagenome, is associated with systemic inflammation and immune dysregulation early in life. In breastfed infants given *Bifidobacterium infantis* EVC001, which expresses all HMO-utilization genes, intestinal T helper 2 (Th2) and Th17 cytokines were silenced and interferon β (IFNβ) was induced. Fecal water from EVC001-supplemented infants contains abundant indolelactate and *B. infantis*-derived indole-3-lactic acid (ILA) upregulated immunoregulatory galectin-1 in Th2 and Th17 cells during polarization, providing a functional link between beneficial microbes and immunoregulation during the first months of life.

INTRODUCTION

Mounting evidence indicates that the composition of the infant gut microbiome is critical to immunological development, particularly during the first 3 months of life, when aberrations in gut microbial composition are most influential in impacting the developing immune system. Indeed, multiple studies have emphasized how early gut microbiome dysbiosis, described as an overabundance of proteobacteria (Shin et al., 2015) and loss of ecosystem function (Duar et al., 2020a), is associated with both acute and chronic immune dysregulation, leading to common conditions such as colic (Rhoads et al., 2018), atopic wheeze and allergy (Arrieta et al., 2015, 2018; Laforest-Lapointe and Arrieta, 2017), and less common but serious immune-mediated disorders such as type 1 diabetes (Vatanen et al., 2016) and Crohn's disease (Hviid et al., 2011). Immune development has been poorly understood in humans due to the difficulty in obtaining samples from infants. Recent developments in systems immunology enable profiling of immune development at the systems level and unraveling of immune cell-regulatory relationships (Davis and Brodin, 2018). Advances in sample processing mean that small-volume samples available from newborn infants are no longer prohibitive and as little as 100 μL of whole blood is sufficient for systems-level immunomonitoring (Olin et al., 2018). We previously unraveled a postnatal adaptation by the newborn immune system to environmental exposures (Olin et al., 2018).

Dysbiosis of the infant gut microbiome is common in modern societies and a likely contributing factor to the increased incidences of immune-mediated disorders (Dominguez-Bello et al., 2019; Mohammadkhah et al., 2018; Sonnenburg and Sonnenburg, 2019). Therefore, there is great interest in identifying microbial factors that can support healthier immune system

 © 2021 The Author(s). Published by Elsevier Inc.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

imprinting and hopefully prevent cases of allergy, autoimmunity, and possibly other conditions involving the immune system (Renz and Skevaki, 2021). Loss of *Bifidobacterium* early in life has been associated with increased risk of developing autoimmunity, as seen in a birth cohort in Finland (Vatanen et al., 2016) and atopic wheeze in another cohort in rural Ecuador (Arrieta et al., 2018). Moreover, observational studies have identified a link between the loss of *Bifidobacterium* in infants and enteric inflammation early in life, but the mechanisms involved are elusive (Henrick et al., 2019; Rhoads et al., 2018).

Human breastmilk contains abundant human milk oligosaccharides (HMOs) that are not digestible by humans as we lack the necessary glucosidases (Sela and Mills, 2010). Instead, the maternal energy spent to create such complex sugars is justified by providing a selective nutritional advantage to "beneficial" microbes specialized in metabolizing HMOs with evolutionarily important functions in the newborn. *Bifidobacterium longum* subspecies (subsp.) *infantis* (*B. infantis*) is one such strain adapted to metabolizing HMOs (LoCascio et al., 2010; Sela et al., 2008; Underwood et al., 2015). *B. infantis* is commonly found in breastfed infants in countries where incidence of immune-mediated disorders is low, such as Bangladesh (Huda et al., 2014) and Malawi (Grześkowiak et al., 2012) but rarely in Europe (Abrahamsson et al., 2014; Avershina et al., 2014; Jost et al., 2012; Roos et al., 2013) and North America (Azad et al., 2013; Casaburi et al., 2021; Lewis et al., 2015). Introducing *B. infantis* has been successfully accomplished in strains such as EVC001 (Evolve BioSystems, Inc.), which is able to stably and persistently colonize and remodel the intestinal microbiome of breastfed infants (Frese et al., 2017), leading to reduced fecal calprotectin, a marker of intestinal inflammation (Henrick et al., 2019).

Here, we extend previous findings by combining longitudinal systems immunology analyses and metagenomic profiling of 208 infants born in Stockholm, Sweden and find that depletion of bifidobacteria and, in particular, HMO-utilization genes from the fecal metagenome is associated with markers of both systemic and intestinal inflammation and immune dysregulation during the first months of life. We also demonstrate a silencing of intestinal inflammation in breastfed infants in California fed *B. infantis* EVC001, a strain harboring all fully functional HMO-utilization genes (Duar et al., 2020a). Fecal water from EVC001-supplemented infants skewed T cell polarization *in vitro* away from a T helper 2 (Th2) state and toward a Th1 state, and indole-3-lactic acid (ILA), 10-fold more abundant in feces from EVC001-treated children versus untreated controls, upregulated immunoregulatory galectin-1 in Th2 and Th17 cells during polarization *in vitro*. This molecular mechanism provides a functional link between beneficial microbes, their metabolites, and immunoregulation during the first critical months of life.

RESULTS

Sequential waves of immune cell expansion during the first months of life

We analyzed longitudinal blood samples (n = 858) from infants (n = 208) born at the Karolinska University Hospital between April 2014 and December 2019 (Olin et al., 2018; Pou et al., 2019). We used mass cytometry and a panel of 44 antibodies (Table S1) targeting activation and differentiation markers across 64 blood immune cell populations (Figure 1A). We also quantified 355 unique plasma proteins using Olink assays (Olink, Uppsala, Sweden) (Lundberg et al., 2011) and together these data elucidated developmental immune system changes postnatally.

When ordering immune cell-frequency measurements by day of sampling, we observed an initial innate response illustrated by the expansion of circulating monocytes that peaked 4–7 days after birth and a transient surge in circulating interferon γ (IFNγ) at days 0–3 as well as elevated levels of circulating interleukin 1RA (IL1RA), a natural inhibitor of the pro-inflammatory IL-1β, likely having a dampening effect on the initial innate response to microbial exposures during the first weeks after birth (Figure 1B). Following the initial monocyte expansion, we observed a gradual increase in the frequency of memory regulatory T cells (Tregs) frequency during the first weeks of life (Figure 1C). We also found a previously unrecognized contraction and subsequent expansion of plasmacytoid dendritic cells (pDCs) after birth (Figure 1C).

From 1 month onward there was a robust increase in circulating $\gamma\delta$T cells, especially within the CD161$^+$ subset of $\gamma\delta$T cells (Figure 1D). These cells are important producers of IL-17A (Maggi et al., 2010), and increased plasma levels of IL-17A were seen during the same first 2-month time window, albeit transiently (Figure 1D). These findings indicate transient innate and adaptive immune responses but also highlight key regulatory mechanisms at play in early and mid infancy. These responses are presumably triggered by colonizing microbes similar to the weaning reaction described in mice (Al Nabhani et al., 2019), but not triggered by weaning and noticeably different in postnatal timing as well as the immune cells and proteins involved.

Expansion of mucosal-specific CD4$^+$ T cells in the blood of newborn children

Previous work has revealed that immune cells primed by antigens at mucosal surfaces circulate and are detectable in peripheral blood. Specifically, memory CD4$^+$ T cells expressing CD38 and lacking the lymphoid tissue homing marker CD62L are mucosal-specific T cells in humans, originate in the intestine, and are identifiable in blood at a frequency of $\sim$4%–8% of total CD4$^+$ T cells (du Pré et al., 2011). Here, we identified this subset of memory CD4$^+$ T cells in the blood of newborn children, but these were more abundant and expanded during the first weeks of life to dominate the circulating memory CD4$^+$ T cell pool (Figure 1E). These cells have downregulated CD45RA and have likely been exposed to antigen, presumably at mucosal surfaces in the intestine, and undergone memory phenotype transition.

To understand these mucosal-specific memory T cells better, we used flow cytometry sorting and bulk mRNA sequencing. We also sorted total memory CD4$^+$ T cells from the same samples and calculated enriched blood transcriptional modules (BTMs) (Li et al., 2014) (Figure 1F). The most enriched hallmark pathways in the mucosal-specific memory CD4$^+$ T cells were type I and type II IFN responses (Figure 1F), and the most upregulated individual genes included complement regulatory factor H (CFH), the cytokine IL-15, important for natural killer (NK) cell homeostasis, as well as the macrophage colony stimulating MA factor CSF1 (M-CSF). Together, these results indicate that mucosal-specific memory CD4$^+$ T cells see antigen after birth, expand in the blood,

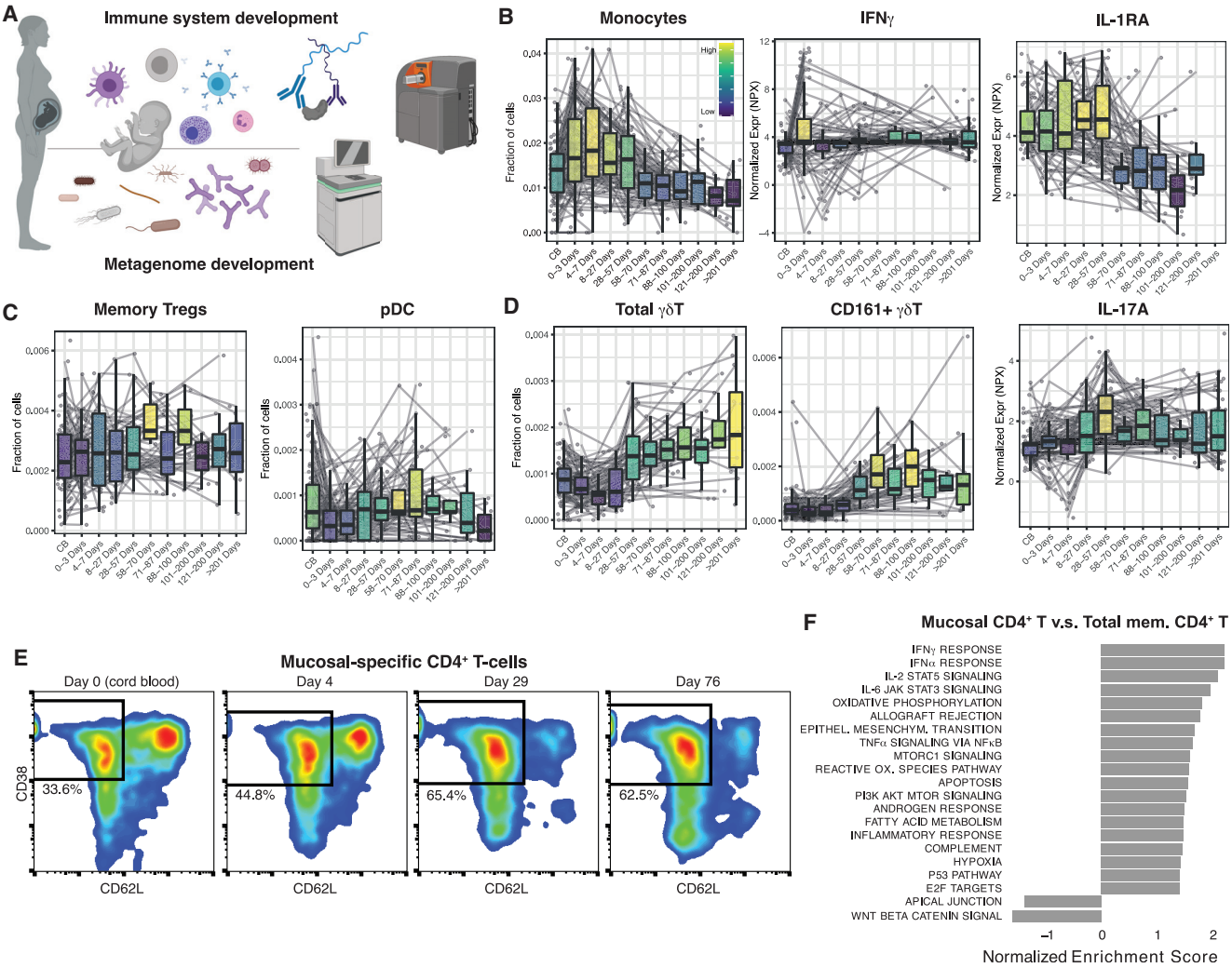


Figure 1. Systems-level analysis of immune development in human newborns

(A) Study overview (antibodies used in mass cytometry listed in Table S1).

(B) Monocyte abundance analyzed by mass cytometry and IFNγ and IL1RA measured by Olink assays in longitudinal blood samples (n = 858) from 208 individual children and binned by sampling day of life. Boxplots are colored by mean rank. CB, cord blood.

(C) Blood mass cytometry analyses of memory Tregs. pDC, plasmacytoid DC.

(D) Blood γδT cell abundance and subset of γδT cells expressing CD161 and plasma IL-17A.

(E) Representative fluorescence-activated cell sorting (FACS) plots of $CD38^{+}CD62L^{-}CD4^{+}$ T cells sorted at postnatal days 0, 4, 29, and 76 from newborn peripheral blood mononuclear cells (PBMCs) and subjected to bulk mRNA sequencing (mRNA-seq).

(F) Gene set enrichment analysis showing the top enriched hallmark pathways in mucosal-specific versus total memory $CD4^{+}$ T cells.

and interact with NK cells and monocytes in the newborn intestinal immune system.

Variable microbiome colonization of the infant gut after birth

To better understand the microbial antigens driving immune responses after birth, we performed shotgun metagenomic sequencing of longitudinal fecal samples (n = 347) from 157 of the 208 infants in this cohort. Gut bacterial composition was highly variable at birth but increasingly converged over time (Figure 2A). At the family level, there was an increase in Bacteroidaceae and Bifidobacteriaceae after birth in the majority of newborns (Figure 2B). Bifidobacteriaceae expanded primarily in breastfed infants without antibiotic exposure (Figure S1). This expansion was observed to involve multiple species of Bifidobacteriaceae, but most frequently *B. longum*, *Bifidobacterium breve*, and *Bifidobacterium bifidum* (Figure 2C). The expansion of bifidobacteria after birth is commonly seen during microbiome colonization according to previous reports (Arrieta et al., 2015, 2018; Vatanen et al., 2016); however, the bifidobacterial expansion is also highly variable across infants with the potential to alter the trajectory of immune system development. Therefore, we decided to study this interaction and the potential consequences in more detail.

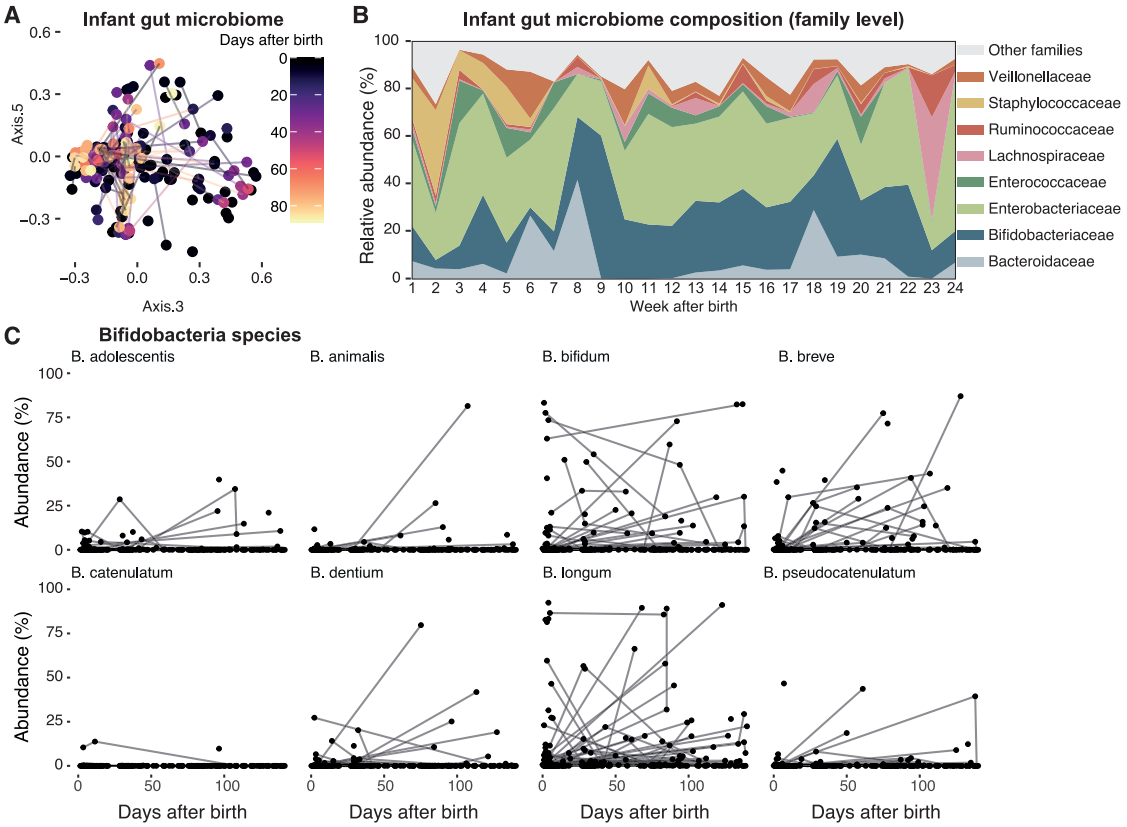


Figure 2. Bifidobacteriaceae expand after birth

(A) Bray-Curtis distance Principal Coordinates Analysis (PCoA) of n = 347 fecal samples collected longitudinally from n = 157 infants.

(B) Mean relative abundance (RA) of gut microbes at the family level across 24 weeks (6 months). Taxa with mean RA <2% are labeled as "other families." Isolated relationship of Bifidobacteriaceae abundance with breastfeeding and antibiotic exposure found in Figure S1.

(C) Species-level abundances within the Bifidobacteriaceae family. Only species detected in at least one sample shown are included.

Immune system states associated with expanded gut bifidobacteria early in life

Given the variable expansion of Bifidobacteriaceae among newborn children, we compared immune system states between infants with abundant Bifidobacteriaceae and those failing to expand such bacteria. Lacking Bifidobacteriaceae was associated with expanded populations of neutrophils, basophils, plasmablasts, and memory CD8$^+$ T cells, indicating both innate and adaptive immune activation (Figure 3A). Mucosal associated invariant T cells (MAIT) are important T cells in the intestinal response to bacterial vitamin B metabolites (Ioannidis et al., 2020) and these were also more abundant in blood from children with a low abundance of Bifidobacteriaceae (Figure 3A). Conversely, children with abundant gut Bifidobacteriaceae had higher frequency of non-classical monocytes, often considered as anti-inflammatory (Narasimhan et al., 2019), as well as antigen-experienced regulatory T cells expressing the CD39 receptor, a highly suppressive Treg subset (Gu et al., 2017) (Figure 3A).

Also, plasma proteins differed between these groups and children lacking Bifidobacteriaceae had elevated levels of tumor necrosis factor α (TNFα) and IL-17A, critical mediators of intestinal inflammation, and also the Th2 cytokines IL-13 and IL-1α, serving as an alarmin released by necrotic cells and synergizing with TNFα in a variety of inflammatory responses (Apte and Voronov, 2008) (Figure 3B). Children with abundant Bifidobacteriaceae had elevated levels of Treg-associated cytokines IL-27 and IL-10 as well as the endogenous IL-1 inhibitor IL1RA, presumably regulating innate IL-1β-mediated responses (Figure 3B). More surprising was the elevated IL-6 in children with abundant Bifidobacteriaceae (Figure 3B).

To better understand the regulatory cell-cell relationships in the newborn immune system, we calculated Spearman correlation matrices and compared infants with high and low abundance of Bifidobacteriaceae (Figure 3C). Such cell-cell correlations indicate co-regulated cell populations and allow for context-dependent perturbations to such relationships to be uncovered (Rodriguez et al., 2020). We found strong positive correlations among naive CD4$^+$ T, naive CD8$^+$ T, and naive regulatory T cells (cluster 1), likely reflecting overall thymic output, and these relationships were comparable among children irrespective of gut Bifidobacteriaceae abundance (Figure 3C). Other modules of co-regulated cell populations were strikingly different between these groups of children. In particular, memory Tregs were inversely correlated with activated CD8$^+$ T cells and pro-inflammatory monocytes (cluster 2) in children with abundant Bifidobacteriaceae, but this relationship was disrupted in infants

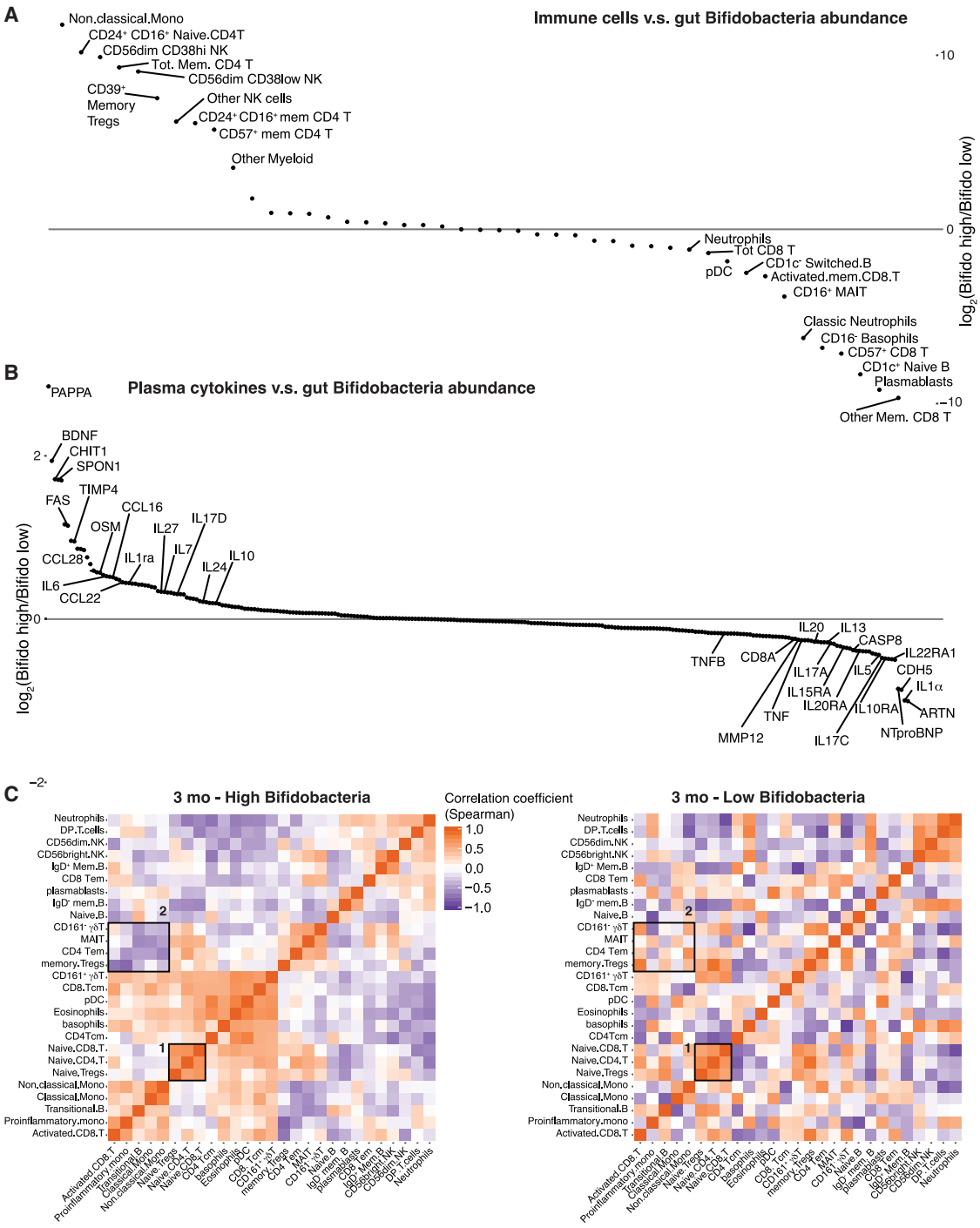


Figure 3. Immune system state in infants with low versus high bifidobacteria

(A) Fold-change immune cell frequencies between 56 and 152 days after birth in infants with high versus low gut Bifidobacteriaceae.

(B) Fold-change plasma protein levels at 3 months of life in infants with high versus low fecal bifidobacteria.

(C) Spearman correlation matrices of immune cell frequencies in the third month of life in children with high versus low fecal bifidobacteria. Black boxes highlight modules of particularly co-regulated immune cell populations. Cluster 1 denotes a thymic output cluster of co-regulated naive T cell subsets, and cluster 2 indicates a regulatory cluster involving memory Tregs, activated T cells, and pro-inflammatory monocytes that differs between children with abundant versus depleted gut bifidobacteria.

lacking such microbes (Figure 3C). We conclude that in infants not colonized by Bifidobacteriaceae or in cases where these microbes fail to expand during the first months of life there is evidence of systemic and intestinal inflammation, increased frequencies of activated immune cells, and reduced levels of regulatory cells indicative of systemic immune dysregulation.

HMO metabolism influences immune system development

Various organic molecules produced by bacteria, including organic acids, phenylalanine, and tryptophan derivatives, have been shown to broadly influence host health (Ehrlich et al., 2020; Fukuda et al., 2011; Laursen et al., 2020). Specifically, *Bifidobacterium*-derived metabolites have been shown to modulate pathogen-induced inflammation via the aryl hydrocarbon receptor (AhR) and NRF-2 pathway (Ehrlich et al., 2020; Meng et al., 2020). To test whether the presence of HMO-utilization genes could explain the immune perturbations in children lacking gut Bifidobacteriaceae, we used the Kyoto Encyclopedia of Genes and Genomes (KEGG) Orthology (KO) database and identified 57 representative key functions necessary to metabolize HMOs (Nguyen et al., 2021; Sela et al., 2008). Next, we assessed the presence of these HMO-utilization genes in fecal metagenomes collected from the infants in our study (Figure 4A). We found that the H5 cluster of HMO-utilization genes was most commonly detected, but the abundances (counts per million; CPM) of HMO-utilization genes were generally low. We then compared relative amounts of the 57 HMO-utilization genes to the 355 plasma proteins measured in blood. Here, we observed significant correlations (Spearman) between a number of plasma proteins, for example, IL-6, TNFα, IL-17A, and IL-13 levels, which were all negatively associated with the presence of HMO-utilization genes, in particular genes within the H5 cluster (Figure 4B; IL-13, $p = 7.79e^{-8}$; IL-6, $p = 1.56e^{-6}$; TNF, $p = 8.85e^{-17}$). Conversely, infants effectively metabolizing HMOs had elevated levels of IL-27 (Figure 4B; $p = 1.36e^{-17}$), a cytokine known to limit Th2- and Th17-type responses in favor of Th1 and regulatory T cell function (Yoshida and Hunter, 2015).

B. infantis EVC001 feeding silences intestinal inflammation early in life

Our data above indicate that HMO-utilization genes expressed by bifidobacteria and other beneficial microbes in breastfed infants correlate with decreased systemic inflammation and a reduction in Th2- and Th17-type responses. Importantly, no isolates from any of the infants in the Swedish cohort expressed all HMO-utilization genes. To assess the beneficial effects of HMO-utilization-gene-expressing microbes, we used an optimized strain of *Bifidobacterium* possessing all HMO-utilization genes. In a second cohort of 60 exclusively breastfed infants in California, we fed approximately half (n = 29) of the newborn children 1.8×10^{10} colony-forming units (CFUs) of *B. longum* subsp. *infantis* EVC001 daily from day 7 to day 28 postnatal and almost half (n = 31) were given no supplementation (Figure 5A). Fecal samples were collected at baseline (day 6) and on day 60 (Figure 5A). As expected, all HMO-utilization genes were abundant in the metagenomes of EVC001-fed but not control children (Figure 5B). Analyses of the microbiome composition showed no significant difference at baseline between these groups; however, by day 60 there was a significant decrease in alpha diversity ($p = 0.0001$; Wilcoxon) in *B. infantis* EVC001-fed infants as compared to controls (Figure S2A).

We then measured fecal cytokine levels in 40 randomly selected newborn children and found no significant differences at baseline but, at day 60, infants fed *B. infantis* EVC001 had reduced levels of IL-13, IL-17A, and regulatory and chemotactic cytokines IL-21, IL-31, IL-33, and MIP3a (Figure 5C; Table S2; $p = 0.015$, 0.0029, 0.00066, 0.007, 0.00011, and 0.0078, respectively). Conversely, there was a significant increase in fecal IFNβ in infants fed *B. infantis* EVC001 (Figure 5C; $p = 0.016$) and this IFNβ concentration correlated with the abundance of Bifidobacteriaceae (Figure S2B). We performed pairwise correlation tests between individual stool taxa (n = 40, day 60) and fecal cytokine concentrations (Spearman, Benjamini-Hochberg false discovery rate [FDR], $\alpha = 0.05$) and identified three taxa, Clostridiaceae, Enterobacteriaceae, and Staphylococcaceae, significantly associated with pro-inflammatory cytokine production. Specifically, Clostridiaceae correlated with IL-17A, IL-21, and IL-33 (Figure 5D; $p = 0.018$, 0.016, and 0.028, respectively) and decreased IFNβ levels (Figure 5D; $p = 0.028$), Enterobacteriaceae correlated with elevated IL-13, IL-17A, IL-21, and IL-33 production (Figure 5D; $p = 0.002$, 0.017, 0.036, and 0.005, respectively) and decreased IFNβ levels (Figure 5D; $p = 0.049$), whereas Staphylococcaceae correlated with higher levels of IL-21 (Figure 5D; $p = 0.028$) and Streptococcaceae correlated with higher levels of IL-23 (Figure 5D; $p = 0.044$). In contrast, only Bifidobacteriaceae was associated with elevated IFNβ (Figure 5D; $p = 0.001$) and negatively correlated with levels of IL-13, IL-17A, IL-21, and IL-33 (Figure 5D; $p = 0.041$, 0.017, 0.027, and 0.001, respectively). We conclude that beneficial microbes expressing all HMO-utilization genes in breastfed infants silence intestinal Th2 and Th17 responses. We also describe a previously unrecognized induction of IFNβ by beneficial microbes, although the direct effect on T cells remains elusive. Thus, it is not the simple presence of bifidobacteria that is responsible for the immune effects but the metabolic partnership between the bacteria and HMOs.

B. infantis EVC001 fecal water skews T cell polarization

To better understand the direct effects of bifidobacterial metabolites and enteric cytokines on T cells, we flow sorted naive $CD4^+$ T cells from a healthy adult donor and polarized these cells using standard cytokine combinations (Cano-Gamez et al., 2020). We also added fecal waters (1:100 dilution) collected from either *B. infantis* EVC001-supplemented or control children lacking *B. infantis* (Figure 6A). To evaluate the cellular states of polarized T cells, we applied a targeted multiomics approach quantifying 259 mRNA molecules with known functions in T cells and 10 surface proteins detected by oligo-coupled antibodies (AbSeq, Rhapsody; BD Biosciences) (Mair et al., 2020). The UMAP embeddings of cells were largely similar across Th0, Th1, and Th2 conditions but slightly different for Th17 and induced regulatory T cells (iTreg) states (Figure 6B). Using a graphical abstraction and clustering method, PAGA (Wolf et al., 2019), we compared the T cell states by polarizing condition and, importantly, uncovered differences between T cells polarized in the presence of

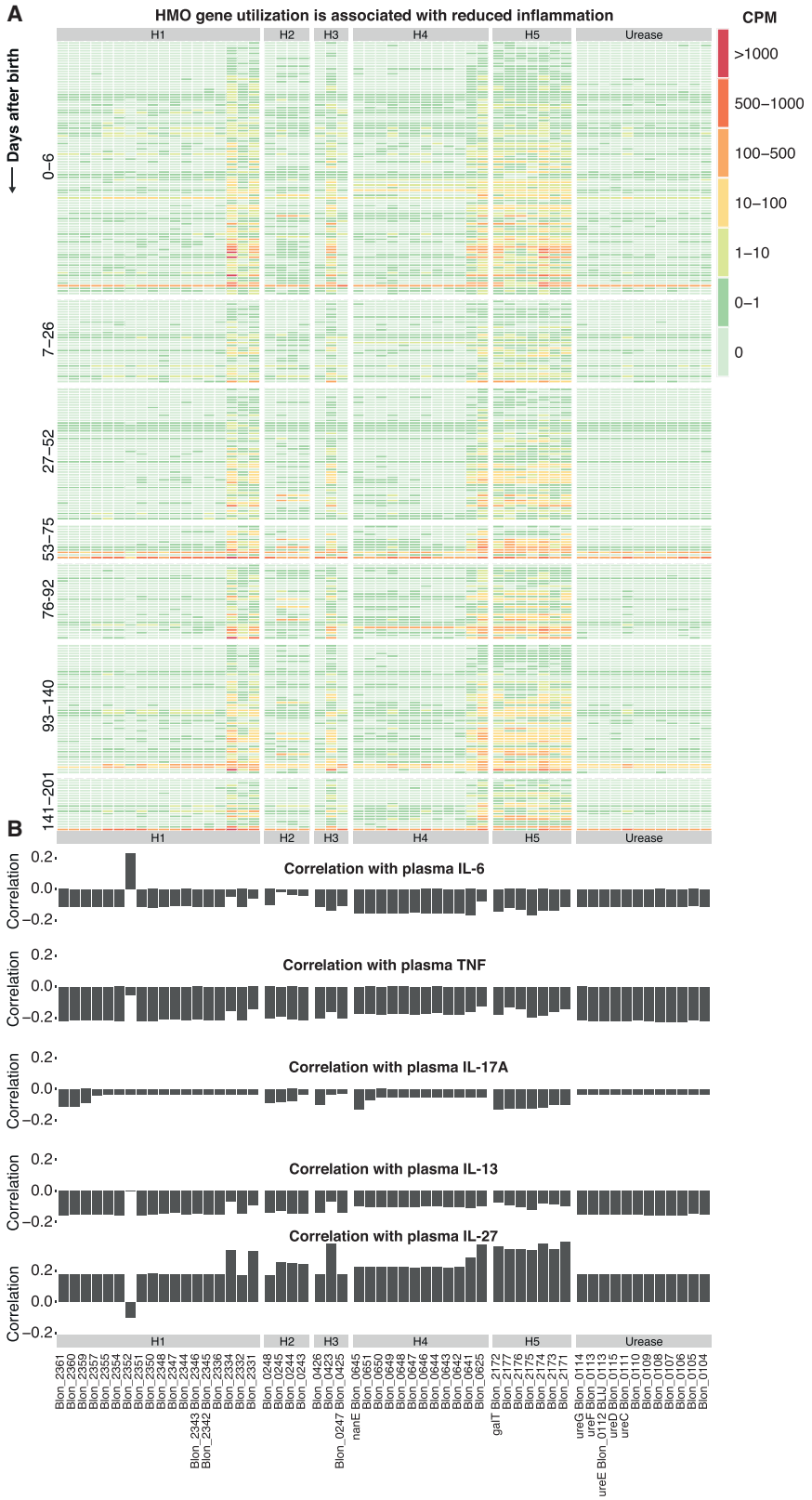


Figure 4. HMO-utilization genes are associated with reduced markers of inflammation

(A) Heatmap showing abundance of the indicated HMO-utilization genes (columns) in the gut metagenome of fecal samples at the collected time interval after birth (horizontal groups).

(B) Spearman correlation coefficients between the indicated plasma cytokine levels (NPX) and abundance of HMO-utilization genes (CPM). IL-13, $p = 7.79e^{-8}$; IL-6, $p = 1.56e^{-6}$; TNF, $p = 8.85e^{-17}$.

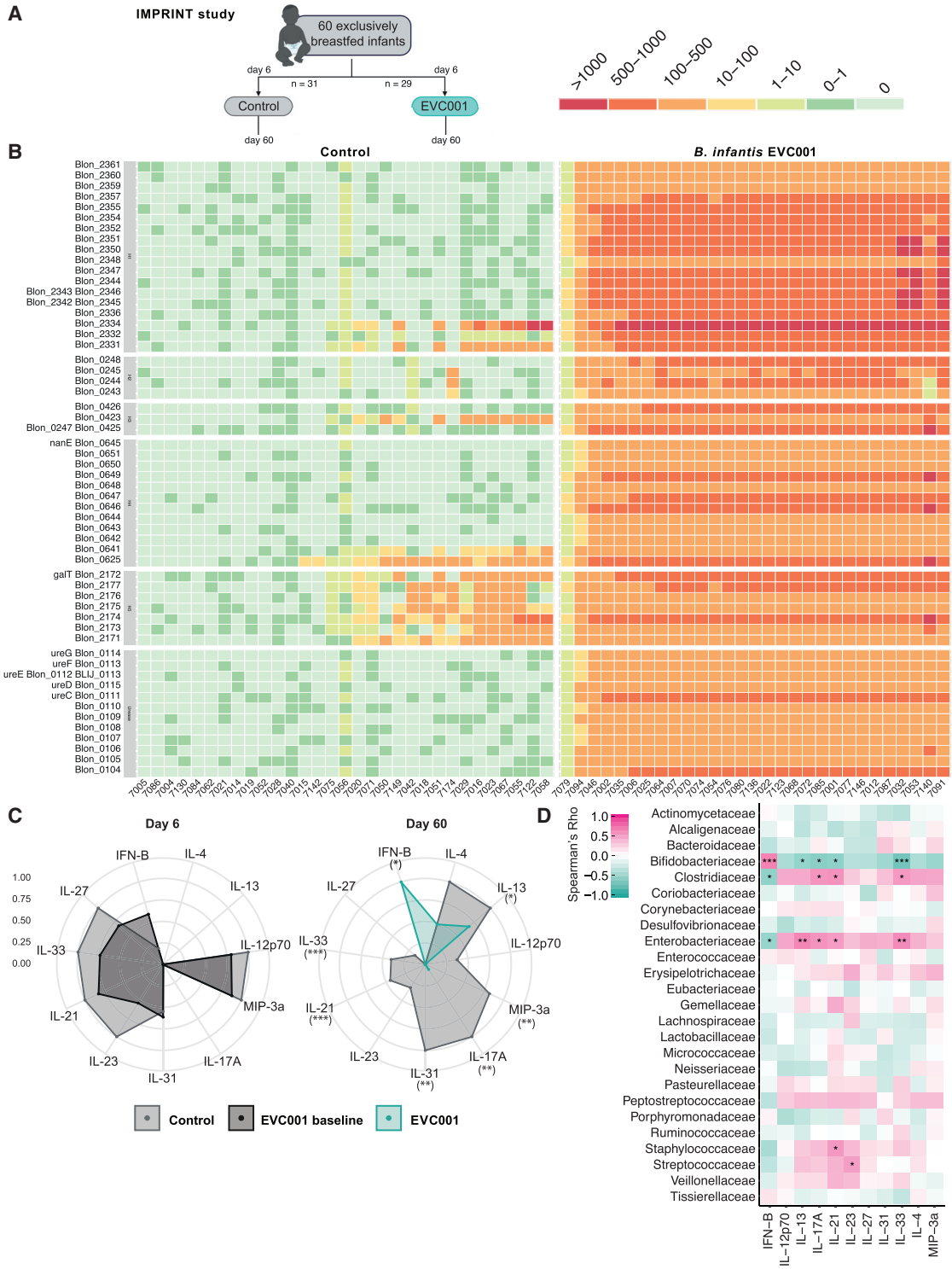


Figure 5. *B. infantis* EVC001 expresses all HMO-utilization genes and reduces intestinal inflammation in breastfed infants

(A) IMPRINT study design and randomization.

(B) Abundance (CPM) of HMO-utilization genes in the metagenome of EVC001-supplemented and control infants. Microbiome alpha diversity of control and *B.infantis* EVC001-fed infants at day 60 postnatal found in Figure S2A.

(legend continued on next page)

B. infantis EVC001 or control fecal water (Figure 6C). Induced Th1-, Th2-, and iTreg-polarized states were comparable between *B. infantis* EVC001 and control fecal water cultures (Figure 6C). Th0 cells, on the other hand, cultured without any polarizing cytokines in the presence of fecal waters from control infants lacking *B. infantis* assumed a Th2-like state, whereas fecal water from infants fed *B. infantis* EVC001 induced a Th1-like state in these Th0 cells (Figure 6C). Differentially expressed genes involved Th1-associated GZMA, GZMB, TNF, and STAT1 in cells exposed to *B. infantis* EVC001 fecal water, whereas IL23R was highly overexpressed in cells exposed to control fecal water (Figure 6D). We supplemented Th0 cultures with IFNβ only, but this did not replicate the effect on Th1/Th2 skewing (Figure S3A).

Apart from this skewing toward Th1, we also noted a difference in Th17-polarized states in *B. infantis* EVC001 fecal water cultures (Figure 6C). Specifically, naive T cells polarized toward Th17 in the presence of fecal waters from control infants had elevated markers of activation and proliferation. Markers such as Ki67 when compared to cells polarized toward Th17 in the presence of *B. infantis* EVC001 fecal water were reduced in contrast to controls (Figure S3B). Collectively, these findings suggest that *B. infantis* EVC001 metabolites or enteric cytokines induced by the presence of *B. infantis* EVC001 exert a polarizing effect on naive $CD4^+$ T cells that favors Th1 polarization, corroborating a mechanism of a silencing effect on fecal IL-13 and IL-17 *in vivo*.

B. infantis EVC001 metabolite ILA induces galectin-1 on Th2 and Th17 cells

Next, we assessed fecal metabolites in samples collected from infants fed *B. infantis* EVC001 and control infants, respectively. A total of 564 biochemicals were significantly different between these fecal samples. Metabolites within the tryptophan metabolism pathway were particularly enriched. The most overrepresented tryptophan metabolite in EVC001-fed infant feces compared to controls was ILA (Figure 6E; $p = 5.89 \times 10^{-8}$; FDR) (Ehrlich et al., 2020; Meng et al., 2020). Importantly, bifidobacteria-derived ILA has recently been shown to bind both the AhR and hydrocarboxylic acid receptor 3 (HCAR3) and modulates monocyte responses to lipopolysaccharide (LPS) (Laursen et al., 2020). $CD4^+$ T cells do not express HCAR3 but do express the AhR (Uhlen et al., 2019), and we tested the impact of ILA on T cell polarization *in vitro* using the same polarizing cytokine conditions as above but replacing fecal water with ILA alone (1 mM). We found a number of mRNA transcripts induced, and these differed between Th1- and Th0-, Th2-, and Th17-polarizing conditions. In the presence of ILA, Th0, Th2, and Th17 cells upregulated the chemokine receptor CXCR3 often associated with Th1 cells and granzyme B (Figure 6F). Moreover, these cells all strongly upregulated the negative regulator of T cell activation, LGALS1 (galectin-1) (Th0, $p = 2.19e^{-42}$; Th2, $p = 2.47e^{-269}$; Th17, $p = 7.62e^{-41}$), suggesting an additional pathway for regulating pathogenic Th2 and Th17 immune responses in newborns (Figure 6F). Also, in culture experiments using fecal water from *B. infantis* EVC001-supplemented children, upregulation of galectin-1 was seen (Figure 6D), further suggesting that ILA-mediated signaling explains some of the effects of *B. infantis* EVC001 supplementation in breastfed infants. In an animal model of zymosan-induced peritonitis, galectin-1 has been reported to induce IL-27 and IL-10 and act through IFNβ-dependent reprogramming of tissue macrophages and be essential in order to resolve inflammation (Yaseen et al., 2020). Our findings of HMO-metabolizing microbes and the induction of tolerogenic responses are concurrent and associated with elevated IL-27 and IFNβ and ILA-mediated upregulation of galectin-1 on $CD4^+$ T cells.

Taken together, our results indicate that during the first weeks of life there are transient immune responses to colonizing microbes, centered on mucosal surfaces. The colonization of the gut microbiome plays an integral role in immune responses, which is likely itself influenced by this process as well as other important determinants of health such as antibiotic use and breastfeeding. Specific bacteria, particularly those expressing HMO-utilization genes, have nutritional advantages in breastfed infants and influence immune-microbe interactions by dampening inflammatory responses, in particular Th2- and Th17-type responses in favor of Th1 and regulatory T cells. Key metabolites such as ILA exert direct regulatory effects on Th2 and Th17 cells such as induction of regulatory galectin-1, known to limit T cell activation, and collectively these layered effects of beneficial microbes and their metabolites on the developing immune system early in life have potential long-term consequences for the risk of developing immune-mediated diseases.

DISCUSSION

It is increasingly clear that early-life immune-microbe interactions influence the risk of immune-mediated diseases later in life; however, the exact mechanisms remain elusive. Results herein extend our previous understanding of an immunological sequence of events, triggered by microbial colonization, that results either in a balanced immune-microbe relationship or varying degrees of intestinal and systemic inflammation and perturbed immune cell regulation. Most notably, we show that within the T cell compartment, low abundance of bifidobacterial species and/or lack of HMO-utilization capacity is associated with intestinal inflammation, driven by aberrant Th2 and Th17 responses. Fecal waters from infants colonized with *B. infantis* EVC001, a strain of *Bifidobacterium* that harbors the full genetic capacity to break down all glycosidic bonds of HMOs (Duar et al., 2020a, 2020b), polarized naive T cells toward Th1 *in vitro*, whereas fecal waters from infants not colonized with *B. infantis* EVC001 polarized Th2- and Th17-phenotype cells. Early in life, newly generated naive T cells populate mucosal tissues and differentiate into memory T cells (Thome

(C) Fecal cytokines at baseline (day 6) and post treatment with *B. infantis* EVC001 or no supplementation. Cytokines are measured as pg/mg of feces; median values were log transformed and scaled from 0 to 1. Median and standard deviation of individual cytokines in each feeding group at day 6 (baseline) and day 60 postnatal found in Table S2.

(D) Spearman correlation coefficients between fecal cytokine levels and bacterial abundance. *p = 0.05, **p = 0.01, ***p = 0.001. Correlation of Bifidobacteriaceae relative abundance and fecal IFNβ concentration postnatal found in Figure S2B.

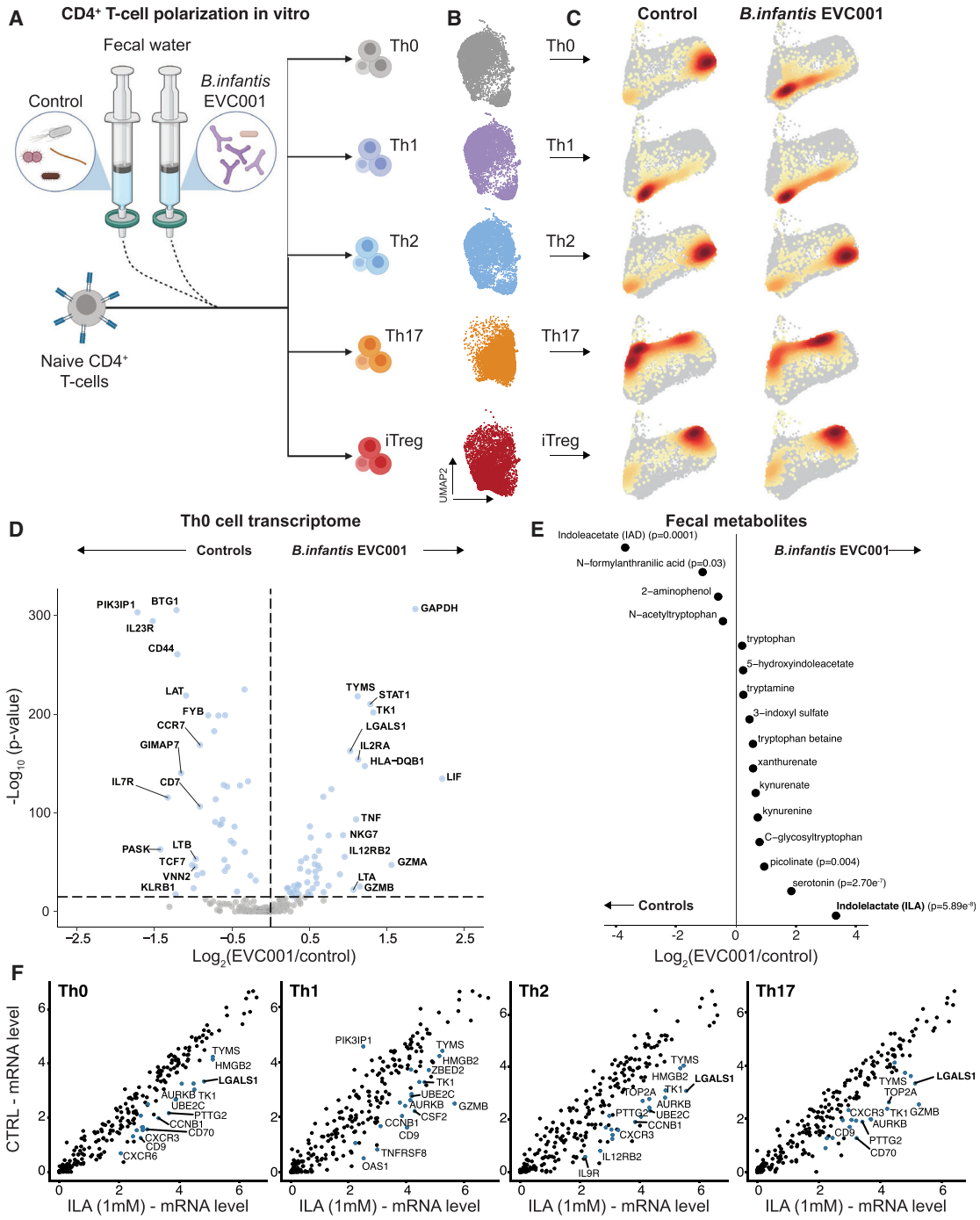


Figure 6. CD4⁺ T cell polarization under the influence of microbial metabolites

(A) CD4⁺ T cell polarization *in vitro* in the presence of fecal water from infants supplemented with *B. infantis* EVC001 or control (no supplement).

(B) UMAP plots of polarized T cells analyzed by targeted single cellmRNA-sequencing (sc-mRNA-seq).

(C) PAGA plots of T cells polarized in the presence of fecal water from infants given *B. infantis* EVC001 supplementation or control. Coloring by cell density is from gray (low) to red (high). IFNβ culture condition and top DEGs among Th17 cell states found in Figure S3.

(D) Volcano plot showing differentially expressed mRNA in Th0 cells cultured with fecal water from infants given *B. infantis* EVC001 supplementation or control.

(E) Fecal tryptophan metabolites measured on day 21 from EVC001-treated and control children; p values indicate mean comparison EVC001 versus control samples.

(F) T cells polarized as in (B) but, instead of fecal water, supplemented by ILA (1 mM) or no supplement and mRNA expression in individual cells quantified by targeted sc-mRNA-seq. Mean expression from 878/362 cells (Th0), 395/697 cells (Th1), 1,073/1,922 cells (Th2), and 861/403 cells (Th17).

et al., 2016), suggesting a window of opportunity for the microbiota to influence mucosal immunity.

A prevailing theory derived from mouse data postulates that timed reactivity to the colonizing microbes establishes a healthy immune-microbe interface in the gut through the induction of tolerance (Knoop et al., 2017). The transient immune cell activation events described here are reminiscent of this transient "weaning reaction" to colonizing microbes in mice (Al Nabhani et al., 2019), but the timing is clearly different in human newborns and the reaction is uncoupled from weaning. Most infants in our cohort breastfed beyond the 3-month window during which the most dramatic immune cell responses seem to occur. Furthermore, the types of reactions seen in human newborns are qualitatively distinct and involve different cell populations as compared to the changes seen in mice, not only due to differences between blood analyses in humans and intestinal analyses in mice (P.B. et al., unpublished results).

Tolerance induction to the microbiota is key to prevent tissue damage and inflammation and here we make observations that indicate how such tolerance could be induced in human infants. We find that Bifidobacteriaceae able to metabolize HMOs are associated with reduced pro-inflammatory markers and conversely elevated proteins such as IL-10 and IL-27 are associated with regulatory T cells. Also, memory Treg frequency was inversely correlated with pro-inflammatory monocyte abundance and activated T cell population abundances in children with abundant bifidobacteria, a regulatory relationship that is lost in children lacking such beneficial microbes. Infants colonized with Bifidobacteriaceae are known to produce high levels of short-chain fatty acids (SCFAs) (Frese et al., 2017), and these are important inducers of regulatory T cells in the gut (Smith et al., 2013). Moreover, we uncover an additional possible inducer of tolerance, namely intestinal IFNβ, which was produced at a significantly higher concentration in infants fed *B. infantis* EVC001 compared to infants that were not colonized with *B. infantis*. This finding is consistent with a recent report showing that specific Bifidobacteriaceae maintain host health during viral infection by inducing IFNβ production in colonic DCs in mice (Stefan et al., 2020). Also, other studies have described a role for microbiota-induced IFNβ in determining immune responses to viral challenges (Abt et al., 2012; Schaupp et al., 2020). IFNβ therapy in patients with multiple sclerosis induces IL-10 production by regulatory T cells (Byrnes et al., 2002) and in mice IFNβ induces regulatory T cells (Dikopoulos et al., 2005), all suggesting that intestinal IFNβ production induced, directly or indirectly by colonizing beneficial microbes is another mechanism of ensuring intestinal tolerance.

Although the *B. infantis* EVC001 strain is not detectable in both cohorts studied, the focus on HMO-utilization genes allows for direct comparisons across these children. Importantly, the similarities in immune system states in relation to HMO-utilization gene abundance serve as independent confirmation of our findings. It is also intriguing, because this suggests that the imprinting effect on developing infant immune systems is not reliant on specific strains of microbes but core functional properties of the metagenome, such as the ability to metabolize HMOs and produce key downstream metabolites such as ILA.

Organic acids, acetate, and propionate have previously been implicated in mitigating lung inflammation in animal models (Trompette et al., 2014) and food allergy in human infants (Sandin et al., 2009). ILA, produced in breastfed infants colonized with *B. infantis*, has been shown to decrease enteric inflammation through activation of AhR and Nrf-2 although the immune system changes were not resolved (Ehrlich et al., 2018; Meng et al., 2020). Further supporting the role of bifidobacteria-derived ILA is a recent report showing induced IL-22 production in $CD4^+$ T cells and modulation of monocyte TNFα responses upon LPS stimulation through the AhR and HCAR3 (Laursen et al., 2020). Our data are in line with this study but add more immunological details, such as the ILA-mediated direct effects on Th2 and Th17 cells and the upregulation of a negative regulator, galectin-1. This finding is also interesting in relation to data in patients with celiac disease, where substantial upregulation of galectin-1 has been shown to induce tolerogenic intestinal responses (Sundblad et al., 2018). We believe that this mechanism of ensuring intestinal tolerance early in life, here shown to be induced by a specific bacterial metabolite, should be investigated further and its therapeutic potential be tested in subsequent clinical trials.

Importantly, bifidobacterial species differ in their ability to utilize HMOs (Sela, 2011). Specifically, all proteins needed to transport HMOs and enzymes required to break down glycosidic bonds in HMOs are conserved only among some members of the *B. longum* subsp. *infantis* lineage (LoCascio et al., 2007, 2010; Sela et al., 2008), in which *B. infantis* HMO-utilization loci have been identified (Albert et al., 2019), although disparities in HMO catabolism between different *B. infantis* strains have only recently been identified, which impact their ecological fitness in the infant gut (Duar et al., 2020a). Our data elucidate a link between the overall ability of the microbiome to access and metabolize HMOs and decreased levels of mucosal and systemic inflammation. Indeed, the correlation of HMO-utilization genes, specifically H5 gene abundance and the decrease in Th2-related cytokines with increased IL-27, is important given recent findings by Duar et al. that H5 is a key ecological determinant of fitness for *Bifidobacterium* species in the infant gut (Duar et al., 2020a) and this fitness advantage is likely both metabolic and dependent on the induction of immunological tolerance.

Collectively, these data provide additional details to the role of beneficial microbes during early-life immune system imprinting and propose possible explanations to previously reported correlative data showing that infants colonized early in life with *Bifidobacterium* species are less likely to develop immune-mediated diseases (Arrieta et al., 2015, 2018; Vatanen et al., 2016). Furthermore, these results highlight the importance of early microbiome colonization during a key window of immunological development where opportunities exist for supplementing the infant gut microbiome with potential benefits to the young child both early in life and long term.

Limitations of the study

Given the dramatic differences between infants supplemented with *B. infantis* EVC001 and control children, it will be important to perform longitudinal systems-level analyses of immune system development of which enrollment in such a study is ongoing. Also, one major limitation in studying newborn children is the inability to

access the intestinal tissue, where immune-microbe interactions occur. The fact that systemic perturbations are still visible is a testament to the wide-ranging effects of immune-microbe interactions at mucosal surfaces on host physiology. The fact that mucosal-specific immune cells are circulating in the blood is known in celiac disease (Han et al., 2013) and inflammatory bowel disease (IBD) (Gorfu et al., 2009) patients. It is also important to note that this current study does not directly assess health outcomes and benefits of expanded *B. infantis* or other HMO-metabolizing bacteria but this is inferred from a wealth of previous associations between microbiome development and immune-mediated diseases (Arrieta et al., 2015; Vatanen et al., 2016), and a more recent study pointing specifically at *B.* subsp. *infantis* associated with a reduced risk of atopy (Seppo et al., 2021).

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Born-immune newborn cohort study
  - IMPRINT study
- METHOD DETAILS
  - Blood immune cell profiling by mass cytometry
  - Antibodies and reagents for mass cytometry
  - Born-immune plasma protein profiling
  - Born-immune fecal metagenomics
  - T cell polarization experiments
  - Targeted transcriptome and protein by BD Rhapsody single cell RNA sequencing
  - IMPRINT—absolute quantification of B. infantis by quantitative real-time PCR
  - IMPRINT—fecal cytokine measurements
  - IMPRINT—fecal water preparation
  - IMPRINT—fecal metabolomics
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Olink preprocessing
  - Mass cytometry preprocessing
  - Born-Immune metagenome data—quality filtering and host removal
  - Born-Immune metagenome data—taxonomic and functional profiling
  - Bifidobacteria abundance correlation
  - HMO correlation
  - Targeted transcriptomics processing
  - Analysis of Seurat object with targeted data
  - Partition-based graph abstraction of single-cell data
  - mRNA-seq data analysis
  - Gene set enrichment analysis (GSEA)
  - IMPRINT—fecal microbiome analyses
  - IMPRINT—quality filtering and removal of human sequences from shotgun metagenomes
  - IMPRINT—fecal cytokine analysis
  - IMPRINT—fecal metabolomics analyses
  - IMPRINT—statistical analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cell.2021.05.030.

ACKNOWLEDGMENTS

The authors thank all families involved in the Born Immune study as well as the clinical staff involved in consenting and collecting samples. The authors thank the parents and their infants enrolled in the IMPRINT clinical trial for collecting information and samples with methodological detail. The IMPRINT study work used the Vincent J. Coates Genomics Sequencing Laboratory at the University of California, Berkeley, supported by NIH instrumentation grant S10 OD018174. Funding was awarded to P.B. from the European Research Council (StG 2015-677559), Swedish Research Council (2015-03028 and 2019-01495), Marianne & Marcus Wallenberg Foundation (2017.0127), Torsten Söderberg Foundation (M41/18), and Karolinska Institutet. We thank the SciLifeLab Plasma Profiling Facility for generating Olink data and National Genomics Infrastructure for sequencing.

AUTHOR CONTRIBUTIONS

B.M.H. and P.B. conceived the study. B.M.H., J.T.S., M.A.U., J.B.G., and S.A.F. designed the IMPRINT study, whereas P.B., E.H., and K.B. devised the Born Immune study. R.D.M., S.C., J.P., and H.K.B. generated IMPRINT data. T T.L., J.M., J.W., C.H.M., C.P., and A.O. generated all other experimental data. Y.A., A.K.B., A.G., E.H., and K.B. gathered clinical data in the Born Immune study. B.M.H., J.T.S., J.B.G., S.A.F., A.A., L.R., Z.T., Y.C., Y.A., and P.B. analyzed the data. P.B. and B.M.H. wrote the manuscript with support from L.R. and A.M.E. All authors approved the manuscript for publication.

DECLARATION OF INTERESTS

P.B., A.O., J.M., and T.L. are co-founders of Cytodelics AB (Stockholm, Sweden). P.B. is an advisor to Scailyte AG (Zurich, Switzerland) and Kancera AB (Stockholm, Sweden). R.D.M., S.C., J.P., H.K.B., S.A.F., and B.M.H. are employees of Evolve BioSystems. J.T.S. received funding to conduct the IMPRINT trial and A.M.E. received funding to assist in writing the manuscript. J.B.G. is a co-founder of Evolve BioSystems. S.A.F. and B.M.H. serve as adjunct assistant professors in the Food Science and Technology Department, University of Nebraska, Lincoln. P.B. and B.M.H. are co-inventors on a patent application related to this work.

INCLUSION AND DIVERSITY

We worked to ensure gender balance in the recruitment of human subjects. One or more of the authors of this paper self-identifies as a member of the LGBTQ+ community. One or more of the authors of this paper self-identifies as an underrepresented ethnic minority in science.

Received: October 27, 2020
Revised: March 19, 2021
Accepted: May 19, 2021
Published: June 17, 2021

REFERENCES

Abrahamsson, T.R., Jakobsson, H.E., Andersson, A.F., Björkstén, B., Engstrand, L., and Jenmalm, M.C. (2014). Low gut microbiota diversity in early infancy precedes asthma at school age. Clin. Exp. Allergy *44*, 842–850.

Abt, M.C., Osborne, L.C., Monticelli, L.A., Doering, T.A., Alenghat, T., Sonnenberg, G.F., Paley, M.A., Antenus, M., Williams, K.L., Erikson, J., et al. (2012). Commensal bacteria calibrate the activation threshold of innate antiviral immunity. Immunity *37*, 158–170.

Albert, K., Rani, A., and Sela, D.A. (2019). Comparative pangenomics of the mammalian gut commensal *Bifidobacterium longum*. Microorganisms *8*, 7.

Al Nabhani, Z., Dulauroy, S., Marques, R., Cousu, C., Al Bounny, S., Déjardin, F., Sparwasser, T., Bérard, M., Cerf-Bensussan, N., and Eberl, G. (2019). A weaning reaction to microbiota is required for resistance to immunopathologies in the adult. Immunity *50*, 1276–1288.e5.

Amnon, Amir (2017). Deblur Rapidly Resolves Single-Nucleotide Community Sequence Patterns. mSystems, In press. https://doi.org/10.1128/mSystems.00191-16.

Apte, R.N., and Voronov, E. (2008). Is interleukin-1 a good or bad 'guy' in tumor immunobiology and immunotherapy? Immunol. Rev. *222*, 222–241.

Arrieta, M.-C., Stiemsma, L.T., Dimitriu, P.A., Thorson, L., Russell, S., Yurist-Doutsch, S., Kuzeljevic, B., Gold, M.J., Britton, H.M., Lefebvre, D.L., et al.; CHILD Study Investigators (2015). Early infancy microbial and metabolic alterations affect risk of childhood asthma. Sci. Transl. Med. *7*, 307ra152.

Arrieta, M.-C., Arévalo, A., Stiemsma, L., Dimitriu, P., Chico, M.E., Loor, S., Vaca, M., Boutin, R.C.T., Morien, E., Jin, M., et al. (2018). Associations between infant fungal and bacterial dysbiosis and childhood atopic wheeze in a nonindustrialized setting. J. Allergy Clin. Immunol. *142*, 424–434.e10.

Avershina, E., Storrø, O., Øien, T., Johnsen, R., Pope, P., and Rudi, K. (2014). Major faecal microbiota shifts in composition and diversity with age in a geographically restricted cohort of mothers and their children. FEMS Microbiol. Ecol. *87*, 280–290.

Azad, M.B., Konya, T., Maughan, H., Guttman, D.S., Field, C.J., Chari, R.S., Sears, M.R., Becker, A.B., Scott, J.A., and Kozyrskyj, A.L.; CHILD Study Investigators (2013). Gut microbiota of healthy Canadian infants: profiles by mode of delivery and infant diet at 4 months. CMAJ *185*, 385–394.

Beghini, F., McIver, L.J., Blanco-Míguez, A., Dubois, L., Asnicar, F., Maharjan, S., Mailyan, A., Manghi, P., Scholz, M., Thomas, A.M., et al. (2021). Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. Elife *10*, e65088.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.

Boylen, Evan (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nature *37* (8), In press. https://doi.org/10.1038/s41587-019-0209-9.

Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. Nat. Biotechnol. *34*, 525–527.

Brodin, P., Duffy, D., and Quintana-Murci, L. (2019). A call for blood—in human immunology. Immunity *50*, 1335–1336.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. *36*, 411–420.

Byrnes, A.A., McArthur, J.C., and Karp, C.L. (2002). Interferon-β therapy for multiple sclerosis induces reciprocal changes in interleukin-12 and interleukin-10 production. Ann. Neurol. *51*, 165–174.

Call, L., Stoll, B., Oosterloo, B., Ajami, N., Sheikh, F., Wittke, A., Waworuntu, R., Berg, B., Petrosino, J., Olutoye, O., and Burrin, D. (2018). Metabolomic signatures distinguish the impact of formula carbohydrates on disease outcome in a preterm piglet model of NEC. Microbiome *6*, 111.

Cano-Gamez, E., Soskic, B., Roumeliotis, T.I., So, E., Smyth, D.J., Baldrighi, M., Willé, D., Nakic, N., Esparza-Gordillo, J., Larminie, C.G.C., et al. (2020). Single-cell transcriptomics identifies an effectorness gradient shaping the response of CD4$^+$ T cells to cytokines. Nat. Commun. *11*, 1801.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Lozupone, C.A., Turnbaugh, P.J., Fierer, N., and Knight, R. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proc. Natl. Acad. Sci. USA *108* (*Suppl 1*), 4516–4522.

Casaburi, G., and Frese, S.A. (2018). Colonization of breastfed infants by *Bifidobacterium longum* subsp. *infantis* EVC001 reduces virulence gene abundance. Hum. Microbiome J. *9*, 7–10.

Casaburi, G., Duar, R.M., Vance, D.P., Mitchell, R., Contreras, L., Frese, S.A., Smilowitz, J.T., and Underwood, M.A. (2019). Early-life gut microbiome modulation reduces the abundance of antibiotic-resistant bacteria. Antimicrob. Resist. Infect. Control *8*, 131.

Casaburi, G., Duar, R.M., Brown, H., Mitchell, R.D., Kazi, S., Chew, S., Cagney, O., Flannery, R.L., Sylvester, K.G., Frese, S.A., et al. (2021). Metagenomic insights of the infant microbiome community structure and function across multiple sites in the United States. Sci. Rep. *11*, 1472.

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics *34*, i884–i890.

Chen, Y., Lakshmikanth, T., Mikes, J., and Brodin, P. (2020). Single-cell classification using learned cell phenotypes. bioRxiv. https://doi.org/10.1101/2020.07.22.216002.

Costea, P.I., Zeller, G., Sunagawa, S., Pelletier, E., Alberti, A., Levenez, F., Tramontano, M., Driessen, M., Hercog, R., Jung, F.-E., et al. (2017). Towards standards for human fecal sample processing in metagenomic studies. Nat. Biotechnol. *35*, 1069–1076.

Czajkowski, M.D., Vance, D.P., Frese, S.A., and Casaburi, G. (2019). GenCoF: a graphical user interface to rapidly remove human genome contaminants from metagenomic datasets. Bioinformatics *35*, 2318–2319.

Davis, M.M., and Brodin, P. (2018). Rebooting human immunology. Annu. Rev. Immunol. *36*, 843–864.

Dikopoulos, N., Bertoletti, A., Kröger, A., Hauser, H., Schirmbeck, R., and Reimann, J. (2005). Type I IFN negatively regulates CD8$^+$ T cell responses through IL-10-producing CD4$^+$ T regulatory 1 cells. J. Immunol. *174*, 99–109.

Dominguez-Bello, M.G., Godoy-Vitorino, F., Knight, R., and Blaser, M.J. (2019). Role of the microbiome in human development. Gut *68*, 1108–1114.

Duar, R.M., Henrick, B.M., Casaburi, G., and Frese, S.A. (2020a). Integrating the ecosystem services framework to define dysbiosis of the breastfed infant gut: the role of *B. infantis* and human milk oligosaccharides. Front. Nutr. *7*, 33.

Duar, R.M., Casaburi, G., Mitchell, R.D., Scofield, L.N.C., Ortega Ramirez, C.A., Barile, D., Henrick, B.M., and Frese, S.A. (2020b). Comparative genome analysis of *Bifidobacterium longum* subsp. *infantis* strains reveals variation in human milk oligosaccharide utilization genes among commercial probiotics. Nutrients *12*, 3247.

du Pré, M.F., van Berkel, L.A., Ráki, M., van Leeuwen, M.A., de Ruiter, L.F., Broere, F., Ter Borg, M.N.D., Lund, F.E., Escher, J.C., Lundin, K.E.A., et al. (2011). CD62L(neg)CD38$^+$ expression on circulating CD4$^+$ T cells identifies mucosally differentiated cells in protein fed mice and in human celiac disease patients and controls. Am. J. Gastroenterol. *106*, 1147–1159.

Ehrlich, A.M., Henrick, B., Pacheco, A., Taft, D., Xu, G., Huda, N., Lozada-Contreras, M., Goodson, M., Slupsky, C., Mills, D., et al. (2018). *Bifidobacterium* grown on human milk oligosaccharides produce tryptophan metabolite indole-3-lactic acid that significantly decreases inflammation in intestinal cells *in vitro*. FASEB J. *32*, lb359.

Ehrlich, A.M., Pacheco, A.R., Henrick, B.M., Taft, D., Xu, G., Huda, M.N., Mishchuk, D., Goodson, M.L., Slupsky, C., Barile, D., et al. (2020). Indole-3-lactic acid associated with *Bifidobacterium*-dominated microbiota significantly decreases inflammation in intestinal epithelial cells. BMC Microbiol. *20*, 357.

Franzosa, E.A., McIver, L.J., Rahnavard, G., Thompson, L.R., Schirmer, M., Weingart, G., Lipson, K.S., Knight, R., Caporaso, J.G., Segata, N., and Huttenhower, C. (2018). Species-level functional profiling of metagenomes and metatranscriptomes. Nat. Methods *15*, 962–968.

Frese, S.A., Hutton, A.A., Contreras, L.N., Shaw, C.A., Palumbo, M.C., Casaburi, G., Xu, G., Davis, J.C.C., Lebrilla, C.B., Henrick, B.M., et al. (2017). Persistence of supplemented *Bifidobacterium longum* subsp. *infantis* EVC001 in breastfed infants. mSphere *2*, e00501-17.

Fukuda, S., Toh, H., Hase, K., Oshima, K., Nakanishi, Y., Yoshimura, K., Tobe, T., Clarke, J.M., Topping, D.L., Suzuki, T., et al. (2011). Bifidobacteria can

protect from enteropathogenic infection through production of acetate. Nature *469*, 543–547.

Gorfu, G., Rivera-Nieves, J., and Ley, K. (2009). Role of beta7 integrins in intestinal lymphocyte homing and retention. Curr. Mol. Med. *9*, 836–850.

Grześkowiak, Ł., Collado, M.C., Mangani, C., Maleta, K., Laitinen, K., Ashorn, P., Isolauri, E., and Salminen, S. (2012). Distinct gut microbiota in southeastern African and northern European infants. J. Pediatr. Gastr. Nutr. *54*, 812–816.

Gu, J., Ni, X., Pan, X., Lu, H., Lu, Y., Zhao, J., Guo Zheng, S., Hippen, K.L., Wang, X., and Lu, L. (2017). Human CD39^{hi} regulatory T cells present stronger stability and function under inflammatory conditions. Cell. Mol. Immunol. *14*, 521–528.

Han, A., Newell, E.W., Glanville, J., Fernandez-Becker, N., Khosla, C., Chien, Y.H., and Davis, M.M. (2013). Dietary gluten triggers concomitant activation of CD4$^+$ and CD8$^+$ $\alpha\beta$ T cells and $\gamma\delta$ T cells in celiac disease. Proc. Natl. Acad. Sci. USA *110*, 13073–13078.

Henrick, B.M., Chew, S., Casaburi, G., Brown, H.K., Frese, S.A., Zhou, Y., Underwood, M.A., and Smilowitz, J.T. (2019). Colonization by *B. infantis* EVC001 modulates enteric inflammation in exclusively breastfed infants. Pediatr. Res. *86*, 749–757.

Huda, M.N., Lewis, Z., Kalanetra, K.M., Rashid, M., Ahmad, S.M., Raqib, R., Qadri, F., Underwood, M.A., Mills, D.A., and Stephensen, C.B. (2014). Stool microbiota and vaccine responses of infants. Pediatrics *134*, e362–e372.

Hviid, A., Svanström, H., and Frisch, M. (2011). Antibiotic use and inflammatory bowel diseases in childhood. Gut *60*, 49–54.

Ioannidis, M., Cerundolo, V., and Salio, M. (2020). The immune modulating properties of mucosal-associated invariant T cells. Front. Immunol. *11*, 1556.

Jacomy, M., Venturini, T., Heymann, S., and Bastian, M. (2014). ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. PLoS One *9*, e98679.

Jost, T., Lacroix, C., Braegger, C.P., and Chassard, C. (2012). New insights in gut microbiota establishment in healthy breast fed neonates. PLoS One *7*, e44595.

Katoh, K. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Briefings in Bioinformatics. https://doi.org/10.1093/bib/bbx108.

Knoop, K.A., Gustafsson, J.K., McDonald, K.G., Kulkarni, D.H., Coughlin, P.E., McCrate, S., Kim, D., Hsieh, C.-S., Hogan, S.P., Elson, C.O., et al. (2017). Microbial antigen encounter during a preweaning interval is critical for tolerance to gut bacteria. Sci. Immunol. *2*, eaao1314.

Laforest-Lapointe, I., and Arrieta, M.-C. (2017). Patterns of early-life gut microbial colonization during human immune development: an ecological perspective. Front. Immunol. *8*, 788.

Laursen, M.F., Sakanaka, M., von Burg, N., Mörbe, U., Andersen, D., Moll, J.M., Pekmez, C.T., Rivollier, A., Michaelsen, K.F., Mølgaard, C., et al. (2020). Breastmilk-promoted bifidobacteria produce aromatic amino acids in the infant gut. bioRxiv. https://doi.org/10.1101/2020.01.22.914994.

Lewis, Z.T., Totten, S.M., Smilowitz, J.T., Popovic, M., Parker, E., Lemay, D.G., Van Tassell, M.L., Miller, M.J., Jin, Y.-S., German, J.B., et al. (2015). Maternal fucosyltransferase 2 status affects the gut bifidobacterial communities of breastfed infants. Microbiome *3*, 13.

Li, S., Rouphael, N., Duraisingham, S., Romero-Steiner, S., Presnell, S., Davis, C., Schmidt, D.S., Johnson, S.E., Milton, A., Rajam, G., et al. (2014). Molecular signatures of antibody responses derived from a systems biology study of five human vaccines. Nat. Immunol. *15*, 195–204.

Lloyd-Price, J., Mahurkar, A., Rahnavard, G., Crabtree, J., Orvis, J., Hall, A.B., Brady, A., Creasy, H.H., McCracken, C., Giglio, M.G., et al. (2017). Strains, functions and dynamics in the expanded Human Microbiome Project. Nature *550*, 61–66.

LoCascio, R.G., Ninonuevo, M.R., Freeman, S.L., Sela, D.A., Grimm, R., Lebrilla, C.B., Mills, D.A., and German, J.B. (2007). Glycoprofiling of bifidobacterial consumption of human milk oligosaccharides demonstrates strain specific, preferential consumption of small chain glycans secreted in early human lactation. J. Agric. Food Chem. *55*, 8914–8919.

LoCascio, R.G., Desai, P., Sela, D.A., Weimer, B., and Mills, D.A. (2010). Broad conservation of milk utilization genes in *Bifidobacterium longum* subsp. *infantis* as revealed by comparative genomic hybridization. Appl. Environ. Microbiol. *76*, 7373–7381.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Lozupone, Catherine (2010). UniFrac: an effective distance metric for microbial community comparison. The ISME Journal. https://doi.org/10.1038/ismej.2010.133.

Lundberg, M., Eriksson, A., Tran, B., Assarsson, E., and Fredriksson, S. (2011). Homogeneous antibody-based proximity extension assays provide sensitive and specific detection of low-abundant proteins in human blood. Nucleic Acids Res. *39*, e102.

Maggi, L., Santarlasci, V., Capone, M., Peired, A., Frosali, F., Crome, S.Q., Querci, V., Fambrini, M., Liotta, F., Levings, M.K., et al. (2010). CD161 is a marker of all human IL-17-producing T-cell subsets and is induced by RORC. Eur. J. Immunol. *40*, 2174–2181.

Mair, F., Erickson, J.R., Voillet, V., Simoni, Y., Bi, T., Tyznik, A.J., Martin, J., Gottardo, R., Newell, E.W., and Prlic, M. (2020). A targeted multi-omic analysis approach measures protein expression and low-abundance transcripts on the single-cell level. Cell Rep. *31*, 107499.

Meng, D., Sommella, E., Salviati, E., Campiglia, P., Ganguli, K., Djebali, K., Zhu, W., and Walker, W.A. (2020). Indole-3-lactic acid, a metabolite of tryptophan, secreted by *Bifidobacterium longum* subspecies *infantis* is anti-inflammatory in the immature intestine. Pediatr. Res. *88*, 209–217.

Mikes, J., Olin, A., Lakshmikanth, T., Chen, Y., and Brodin, P. (2019). Automated cell processing for mass cytometry experiments. Methods Mol. Biol. *1989*, 111–123.

Mohammadkhah, A.I., Simpson, E.B., Patterson, S.G., and Ferguson, J.F. (2018). Development of the gut microbiome in children, and lifetime implications for obesity and cardiometabolic disease. Children (Basel) *5*, 160.

Narasimhan, P.B., Marcovecchio, P., Hamers, A.A.J., and Hedrick, C.C. (2019). Nonclassical monocytes in health and disease. Annu. Rev. Immunol. *37*, 439–456.

Nguyen, M., Holdbrooks, H., Mishra, P., Abrantes, M.A., Eskew, S., Garma, M., Oca, C.-G., McGuckin, C., Hein, C.B., Mitchell, R.D., et al. (2021). Impact of probiotic *B. infantis* EVC001 feeding in premature infants on the gut microbiome, nosocomially acquired antibiotic resistance, and enteric inflammation. Front. Pediatr. *9*, 618009.

Olin, A., Henckel, E., Chen, Y., Lakshmikanth, T., Pou, C., Mikes, J., Gustafsson, A., Bernhardsson, A.K., Zhang, C., Bohlin, K., and Brodin, P. (2018). Stereotypic immune system development in newborn children. Cell *174*, 1277–1292.e14.

Pou, C., Nkulikiyimfura, D., Henckel, E., Olin, A., Lakshmikanth, T., Mikes, J., Wang, J., Chen, Y., Bernhardsson, A.K., Gustafsson, A., et al. (2019). The repertoire of maternal anti-viral antibodies in human newborns. Nat. Med. *25*, 591–596.

Price, Morgan (2009). FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol Biol Evol. https://doi.org/10.1093/molbev/msp077.

Renz, H., and Skevaki, C. (2021). Early life microbial exposures and allergy risks: opportunities for prevention. Nat. Rev. Immunol. *21*, 177–191.

Rhoads, J.M., Collins, J., Fatheree, N.Y., Hashmi, S.S., Taylor, C.M., Luo, M., Hoang, T.K., Gleason, W.A., Van Arsdall, M.R., Navarro, F., and Liu, Y. (2018). Infant colic represents gut inflammation and dysbiosis. J. Pediatr. *203*, 55–61.e3.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. *43*, e47.

Rodriguez, L., Pekkarinen, P.T., Lakshmikanth, T., Tan, Z., Consiglio, C.R., Pou, C., Chen, Y., Mugabo, C.H., Nguyen, N.A., Nowlan, K., et al. (2020). Systems-level immunomonitoring from acute to recovery phase of severe COVID-19. Cell Rep. Med. *1*, 100078.

Roos, S., Dicksved, J., Tarasco, V., Locatelli, E., Ricceri, F., Grandin, U., and Savino, F. (2013). 454 pyrosequencing analysis on faecal samples from a randomized DBPC trial of colicky infants treated with *Lactobacillus reuteri* DSM 17938. PLoS One *8*, e56710.

Sandin, A., Bråbäck, L., Norin, E., and Björkstén, B. (2009). Faecal short chain fatty acid pattern and allergy in early childhood. Acta Paediatr. *98*, 823–827.

Schaupp, L., Muth, S., Rogell, L., Kofoed-Branzk, M., Melchior, F., Lienenklaus, S., Ganal-Vonarburg, S.C., Klein, M., Guendel, F., Hain, T., et al. (2020). Microbiota-induced type I interferons instruct a poised basal state of dendritic cells. Cell *181*, 1080–1096.e19.

Scholz, M., Ward, D.V., Pasolli, E., Tolio, T., Zolfo, M., Asnicar, F., Truong, D.T., Tett, A., Morrow, A.L., and Segata, N. (2016). Strain-level microbial epidemiology and population genomics from shotgun metagenomics. Nat. Methods *13*, 435–438.

Sela, D.A. (2011). Bifidobacterial utilization of human milk oligosaccharides. Int. J. Food Microbiol. *149*, 58–64.

Sela, D.A., and Mills, D.A. (2010). Nursing our microbiota: molecular linkages between bifidobacteria and milk oligosaccharides. Trends Microbiol. *18*, 298–307.

Sela, D.A., Chapman, J., Adeuya, A., Kim, J.H., Chen, F., Whitehead, T.R., Lapidus, A., Rokhsar, D.S., Lebrilla, C.B., German, J.B., et al. (2008). The genome sequence of *Bifidobacterium longum* subsp. *infantis* reveals adaptations for milk utilization within the infant microbiome. Proc. Natl. Acad. Sci. USA *105*, 18964–18969.

Seppo, A.E., Bu, K., Jumabaeva, M., Thakar, J., Choudhury, R.A., Yonemitsu, C., Bode, L., Martina, C.A., Allen, M., Tamburini, S., et al. (2021). Infant gut microbiome is enriched with *Bifidobacterium longum* ssp. *infantis* in Old Order Mennonites with traditional farming lifestyle. Allergy. Published online April 27, 2020. https://doi.org/10.1111/all.14877.

Shin, N.-R., Whon, T.W., and Bae, J.-W. (2015). Proteobacteria: microbial signature of dysbiosis in gut microbiota. Trends Biotechnol. *33*, 496–503.

Smilowitz, J.T., Moya, J., Breck, M.A., Cook, C., Fineberg, A., Angkustsiri, K., and Underwood, M.A. (2017). Safety and tolerability of *Bifidobacterium longum* subspecies *infantis* EVC001 supplementation in healthy term breastfed infants: a phase I clinical trial. BMC Pediatr. *17*, 133.

Smith, P.M., Howitt, M.R., Panikov, N., Michaud, M., Gallini, C.A., Bohlooly-Y, M., Glickman, J.N., and Garrett, W.S. (2013). The microbial metabolites, short-chain fatty acids, regulate colonic T_{reg} cell homeostasis. Science *341*, 569–573.

Sonnenburg, E.D., and Sonnenburg, J.L. (2019). The ancestral and industrialized gut microbiota and implications for human health. Nat. Rev. Microbiol. *17*, 383–390.

Stefan, K.L., Kim, M.V., Iwasaki, A., and Kasper, D.L. (2020). Commensal microbiota modulation of natural resistance to virus infection. Cell *183*, 1312–1324.e10.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA *102*, 15545–15550.

Sundblad, V., Quintar, A.A., Morosi, L.G., Niveloni, S.I., Cabanne, A., Smecuol, E., Mauriño, E., Mariño, K.V., Bai, J.C., Maldonado, C.A., and Rabinovich, G.A. (2018). Galectins in intestinal inflammation: galectin-1 expression delineates response to treatment in celiac disease patients. Front. Immunol. *9*, 379.

Thome, J.J.C., Bickham, K.L., Ohmura, Y., Kubota, M., Matsuoka, N., Gordon, C., Granot, T., Griesemer, A., Lerner, H., Kato, T., and Farber, D.L. (2016). Early-life compartmentalization of human T cell differentiation and regulatory function in mucosal and lymphoid tissues. Nat. Med. *22*, 72–77.

Traag, V.A., Waltman, L., and van Eck, N.J. (2019). From Louvain to Leiden: guaranteeing well-connected communities. Sci. Rep. *9*, 5233.

Trompette, A., Gollwitzer, E.S., Yadava, K., Sichelstiel, A.K., Sprenger, N., Ngom-Bru, C., Blanchard, C., Junt, T., Nicod, L.P., Harris, N.L., and Marsland, B.J. (2014). Gut microbiota metabolism of dietary fiber influences allergic airway disease and hematopoiesis. Nat. Med. *20*, 159–166.

Truong, D.T., Franzosa, E.A., Tickle, T.L., Scholz, M., Weingart, G., Pasolli, E., Tett, A., Huttenhower, C., and Segata, N. (2015). MetaPhlAn2 for enhanced metagenomic taxonomic profiling. Nat. Methods *12*, 902–903.

Uhlen, M., Karlsson, M.J., Zhong, W., Tebani, A., Pou, C., Mikes, J., Lakshmikanth, T., Forsström, B., Edfors, F., Odeberg, J., et al. (2019). A genome-wide transcriptomic analysis of protein-coding genes in human blood cells. Science *366*, eaax9198.

Underwood, M.A., German, J.B., Lebrilla, C.B., and Mills, D.A. (2015). *Bifidobacterium longum* subspecies *infantis*: champion colonizer of the infant gut. Pediatr. Res. *77*, 229–235.

Vatanen, T., Kostic, A.D., d'Hennezel, E., Siljander, H., Franzosa, E.A., Yassour, M., Kolde, R., Vlamakis, H., Arthur, T.D., Hämäläinen, A.-M., et al.; DIABIMMUNE Study Group (2016). Variation in microbiome LPS immunogenicity contributes to autoimmunity in humans. Cell *165*, 842–853.

Walters, W., Hyde, E.R., Berg-Lyons, D., Ackermann, G., Humphrey, G., Parada, A., Gilbert, J.A., Jansson, J.K., Caporaso, J.G., Fuhrman, J.A., et al. (2015). Improved bacterial 16S rRNA gene (V4 and V4-5) and fungal internal transcribed spacer marker gene primers for microbial community surveys. mSystems *1*, e00009-15.

Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: large-scale singlecell gene expression data analysis. Genome Biology *19*, 15.

Wolf, F.A., Hamey, F.K., Plass, M., Solana, J., Dahlin, J.S., Göttgens, B., Rajewsky, N., Simon, L., and Theis, F.J. (2019). PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. Genome Biol. *20*, 59.

Wood, D.E., Lu, J., and Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. Genome Biol. *20*, 257.

Yaseen, H., Butenko, S., Polishuk-Zotkin, I., Schif-Zuck, S., Pérez-Sáez, J.M., Rabinovich, G.A., and Ariel, A. (2020). Galectin-1 facilitates macrophage reprogramming and resolution of inflammation through IFN-β. Front. Pharmacol. *11*, 901.

Yoshida, H., and Hunter, C.A. (2015). The immunobiology of interleukin-27. Annu. Rev. Immunol. *33*, 417–443.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Mass cytometry—broad extended panel		
Anti-human CD3e (UCHT1), Sm-154	Fluidigm	Cat# 317302; RRID: AB_571927
Anti-human CD4 (RPA-T4), Purified	Biolegend	Cat# 300502; RRID: AB_314070
Anti-human CD5 (UCHT2), Purified	Biolegend	Cat# 300602; RRID: AB_314088
Anti-human CD7 (CD7-6B7), Purified	Biolegend	Cat# 343102; RRID: AB_1659214
Anti-human CD8 (SK1), Purified	Biolegend	Cat# 344702; RRID: AB_1877104
Anti-human CD9 (SN4 C3-3A2), Purified	eBiosciences	Cat# 14-0098-82; RRID: AB_657777
Anti-human CD11B (Mac-1) – 209Bi	Fluidigm	Cat# 3209003B; RRID:AB_2687654
Anti-human CD11c (Bu15), Purified	Biolegend	Cat# 337202; RRID: AB_1236381
Anti-human CD14 (M5E2), Purified	Biolegend	Cat# 301802; RRID: AB_314184
Anti-human CD15 (W6D3), Purified	Biolegend	Cat# 323002; RRID: AB_756008
Anti-human CD16 (3G8), Bi-209	Fluidigm	Cat# 3209002B, RRID: AB_2756431
Anti-human CD19 (HIB19), Nd-142	Fluidigm	Cat# 3142001B; RRID:AB_2651155
Anti-human CD20 (2H7), Purified	Biolegend	Cat# 302302; RRID: AB_314250
Anti-human CD22 (HIB22), Purified	Biolegend	Cat# 302502; RRID: AB_314264
Anti-human CD24 (ML5), Purified	Biolegend	Cat# 311102; RRID: AB_314851
Anti-human CD25 (2A3), Sm-149	Fluidigm	Cat# 3149010B, RRID: AB_2756416
Anti-human CD26 (BA5b), Purified	Biolegend	Cat# 302702; RRID: AB_314286
Anti-human CD27 (L128), Er-167	Fluidigm	Cat# 3167006B; RRID: AB_2811093
Anti-human CD29 (TS2/16), Purified	Biolegend	Cat# 303002; RRID: AB_314318
Anti-human CD31 (WM59) – Purified	BioLegend	Cat# 303102; RRID:AB_314328
Anti-human CD33 (WM53), Purified	Biolegend	Cat# 303402; RRID: AB_314346
Anti-human CD34 (581), Purified	Biolegend	Cat# 343502; RRID: AB_1731898
Anti-human CD38 (HIT2), Purified	Biolegend	Cat# 303502; RRID: AB_314354
Anti-human CD39 (A1), Purified	Biolegend	Cat# 328202; RRID: AB_940438
Anti-human CD44 (BJ18) – Purified	BioLegend	Cat# 338802; RRID:AB_1501199
Anti-human CD45 (HI30), Y-89	Fluidigm	Cat# 3089003B; RRID: AB_2661851
Anti-human CD45RA (HI100), Tm-169	Fluidigm	Cat# 3169008B; RRID: N/A
Anti-human CD49d (9F10), Pr-141	Fluidigm	Cat# 3141004B; RRID: N/A
Anti-human CD56 (NCAM16.2), Purified	BD	Cat# 559043; RRID: AB_397180
Anti-human CD57 (HCD57), Purified	Biolegend	Cat# 322302; RRID: AB_535988
Anti-human CD64 (10.1), Purified	Biolegend	Cat# 305002, RRID: AB_314486
Anti-human CD99 (HCD99), Purified	Biolegend	Cat# 318002; RRID: AB_604112
Anti-human CD123 (6H6), Purified	Biolegend	Cat# 306002; RRID: AB_314576
Anti-human CD127 (A019D5), Ho-165	Fluidigm	Cat# 3165008B; RRID: AB_2868401
Anti-human CD161 (HP-3G10), Purified	Biolegend	Cat# 339902; RRID: AB_2661837
Anti-human HLA-ABC (W6/32), Purified	Biolegend	Cat# 311402; RRID: AB_1076699
Anti-human HLA-DR (L243), Purified	Biolegend	Cat# 307602; RRID: AB_314680
Anti-human IgD (IA6-2), Purified	Biolegend	Cat# 348202; RRID: AB_10550095
Anti-human Siglec-8 (837535), Purified	R&D Systems	Cat# MAB7975; RRID: N/A
Anti-human TCRgd (5A6.E9), Purified	Fischer Scientific	Cat# TCR1061; RRID: AB_223500
Biological samples		
Whole blood samples from newborns and parents	Karolinska University Hospital	N/A
Fecal samples from newborn children	Karolinska University Hospital	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
Bovine Serum Albumin	Sigma-Aldrich	Cat# A3059; RRID: N/A
Cell-ID Intercalator-Ir	Fluidigm	Cat# 201192B; RRID: N/A
Cell-ID 20-Plex Pd Barcoding Kit	Fluidigm	Cat# 201060; RRID: N/A
DMSO	Sigma-Aldrich	Cat# D8418; RRID: N/A
EDTA	Rockland	Cat# MB-014; RRID: N/A
EQ Four Element Calibration Beads	Fluidigm	Cat# 201078; RRID: N/A
FBS	Sigma-Aldrich	Cat# 12103C; RRID: N/A
Fc Receptor (FcR) blocking buffer	Cytodelics	Customized
Maxpar Water	Fluidigm	Cat# 201069; RRID; N/A
Maxpar X8 Multimetal Labeling Kit (40 rxn)	Fluidigm	Cat# 201300; RRID; N/A
Metal isotopes as chloride salts (In-115, Gd-155, Gd-157, Dy-161, Dy-163, Yb-173)	Trace Sciences International	Customized
Paraformaldehyde	VWR	Cat# 16005; RRID: N/A
Penicillin-streptomycin	Sigma-Aldrich	Cat# P4333; RRID: N/A
Protein Stabilizer PBS	Candor Bioscience	Cat# 131125, RRID: N/A
PBS 1X	Rockland	Cat# MB-008; RRID: N/A
RPMI 1640 medium	Sigma-Aldrich	Cat# R848; RRID: N/A
Sodium Azide	Sigma-Aldrich	Cat# 71289; RRID: N/A
Whole blood (human) processing kit	Cytodelics	Cat# hC001-500; RRID: N/A
Critical commercial assays		
Inflammation I panel	Olink AB	N/A
Cardiovascular (CVD) II panel	Olink AB	N/A
Cardiovascular (CVD) III panel	Olink AB	N/A
Immune Response panel	Olink AB	N/A
Other		
BenchBot robot	Agilent technologies	Customized
Bravo liquid handling platform	Agilent technologies	Customized
CyTOF 2 mass cytometer	Fluidigm	N/A
EL406 Washer Dispenser	BioTek	Customized
pluriStrainer Mini, 40 μm	pluriSelect	Cat# 43-10040-70; RRID: N/A
Polypropylene tubes	Sarstedt	Cat# 55526; RRID: N/A
TC20 automated cell counter	BioRad	N/A
Vspin microplate centrifuge	Agilent technologies	Customized
Software and algorithms		
Seven Bridges Platform	-	https://www.sevenbridges.com/platform/
R 4.0.2		https://www.r-project.org/
Python 3.7		https://www.python.org/
Seurat		https://cran.r-project.org/web/packages/Seurat/index.html
DESeq2	Love et al., 2014	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
tmod 0.46.2		https://cran.r-project.org/package=tmod
kallisto	Bray et al., 2016	https://github.com/pachterlab/kallisto
fgsea	Subramanian et al., 2005	http://bioconductor.org/packages/release/bioc/html/fgsea.html
Scanpy	Wolf et al., 2018	https://github.com/theislab/Scanpy
PAGA	Wolf et al., 2019	https://github.com/theislab/paga

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
BD Rhapsody antibodies and reagents		
BD AbSeq Hu CCR7(CD197)	BD	Cat # 940394
BD AbSeq Hu CD137	BD	Cat # 940055
BD AbSeq Hu CD28	BD	Cat # 940017
BD AbSeq Hu CD11a	BD	Cat # 940077
BD AbSeq Hu TCR galma/delta	BD	Cat # 940057
BD AbSeq Hu CD103	BD	Cat # 940067
BD AbSeq Hu CD27	BD	Cat # 940018
BD AbSeq Hu CD39	BD	Cat # 940073
BD AbSeq Hu CD62L	BD	Cat # 940041
BD AbSeq Hu CD4	BD	Cat # 940001
BD AbSeq Hu CD8	BD	Cat # 940003
BD AbSeq Hu CD3	BD	Cat # 940000
BD AbSeq Hu CD16	BD	Cat # 940006
BD AbSeq Hu CD56	BD	Cat # 940007
BD AbSeq Hu CD161	BD	Cat # 940070
BD AbSeq Hu CD45RA	BD	Cat # 940011
BD AbSeq Hu CD38	BD	Cat # 940013
BD AbSeq Hu HLA-DR	BD	Cat # 940010
BD single cell multiplexing kit	BD	Cat # 633781
BD Rhapsody cartridge reagent kit	BD	Cat # 633731
BD Rhapsody cDNA kit	BD	Cat # 633773
BD Rhapsody targeted mRNA and AbSeq amplification kit	BD	Cat # 633774
BD PharMingen Stain Buffer (FBS)	BD	Cat # 554656
DRAQ7	BD	Cat # 564904
BD Rhapsody Human T Cell Expression Panel	BD	Cat # 633751
Critical commercial assays		
Zirconium beads 0,1 mm	QIAGEN	13118-400
Lysozyme	ThermoFisher	89833
Proteinase K	QIAGEN	19133
ThruPLEX® DNA-seq 96D	Takara	R400676
Agencourt AMPure XP	Beckmann Coulter	A63881
Qubit Assay Tubes,	ThermoFisher	Q32856
Qubit dsDNA HS Assay Kit	ThermoFisher	Q32854
QIAamp Fast DNA Stool Mini Kit	QIAGEN	51604
Agilent High Sensitivity DNA Kit	Agilent	5067-4626
96 microTUBE Plate	Covaris	520078

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by Lead Contact, Petter Brodin (petter.brodin@ki.se).

Materials availability

This study did not generate new unique reagents.

Data and code availability

Raw and processed data is available for download in our Mendeley Data repository: https://doi.org/10.17632/gc4d9h4x67.2. The bulk mRNA sequencing data has not been deposited in a public repository due to the sensitive nature of this dataset but are available from the corresponding author on request. Scripts for recreating figures in the paper: https://github.com/rodriluc/Bifido_newborns/

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Born-immune newborn cohort study

The study was performed in accordance with the declaration of Helsinki and the study protocol was approved by the regional ethical board in Stockholm, Sweden (DNR: 2009/2052–31/3, 2014/921-32 and 2016/512-31/1). After obtaining informed consent from parents, blood samples from newborns (n = 208) were collected at the Karolinska University Hospital. We also collected fecal samples from infants, either at the time of clinical visits and frozen directly at −80 C or collected at home frozen at −20 C and brought to the clinic by parents. Age of the newborns were measured as days after birth and ranging from 0 to 1452 days. Clinical metadata such as mode of delivery, sex, nutrition, growth and medications were gathered in a clinical database: https://brodinlab.com/newborns/

IMPRINT study

Details of the study design and procedures used to collect these samples has been reported elsewhere (ClinicalTrials.gov: NCT02457338) (Frese et al., 2017; Smilowitz et al., 2017). Briefly, exclusively breastfed term infants were randomly selected to receive 1.8×10^{10} colony-forming units (CFU) activated *B. infantis* EVC001 daily for 21 days (EVC001) starting at day 7 postnatal or to receive breast milk alone (control) and followed up to postnatal day 60 (Frese et al., 2017). All mothers received lactation support throughout the study. The demographic information (e.g., age, sex, and gestational age) was collected from each participant. All aspects of the study were approved by the University of California Davis Institutional Review Board (IRB Number: ID 631099) and all participants provided written informed consent. Here fecal samples from individual subjects were chosen at random and made up a subset of the original study participants. Fecal samples from randomly selected infants who were fed EVC001 (n = 20) and control infants (n = 20) on days 6 (Baseline) 40 and 60 postnatal were collected and analyzed for fecal enteric cytokine concentrations (described below). Day 21 fecal samples were used for non-targeted fecal metabolomics analysis (described below). Day 21 metagenomics, which have been reported elsewhere (Casaburi and Frese, 2018) were used for correlative analyses. 16S amplicon date previously described (Frese et al., 2017) was used for diversity and correlative analyses. All sequencing libraries generated in this study have been deposited with the NCBI SRA (PRJNA390646) and are publicly available.

METHOD DETAILS

Blood immune cell profiling by mass cytometry

Blood samples drawn from newborns and parents were mixed with a stabilizer (Brodin et al., 2019) (one of the components of Whole blood processing kit; Cytodelics AB, Sweden) either immediately or within 1-3 hours post blood draw and cryopreserved as per the manufacturer's recommendations. Samples were thawed, and cells were fixed/RBCs lysed using WASH # 1 and WASH # 2 buffers (Whole blood processing kit; Cytodelics AB, Sweden) as per the manufacturer's recommendations. This was performed a few days prior to barcoding and staining of cells. Post fix/lysis of cells, $\sim$1-2x10^6 cells/sample were plated onto a 96 well round bottom plate using standard cryoprotective solution (10% DMSO and 90% FBS) and cryopreserved at −80°C. At the time of experimentation, cells were thawed at 37°C using RPMI medium supplemented with 10% fetal bovine serum (FBS), 1% penicillin-streptomycin and benzonase (Sigma-Aldrich, Sweden). Briefly, cells were barcoded using automated liquid handling robotic system (Agilent technologies) (Mikes et al., 2019) using the Cell-ID 20-plex Barcoding kit (Fluidigm Inc.) as per the manufacturer's recommendations. Samples were pooled batch wise by keeping together the longitudinal samples from each newborn baby or parent in the same batch. Cells were then washed, FcR blocked for 12 min at room temperature, following which cells were incubated for another 30 min at 4°C after addition of a cocktail of metal conjugated antibodies targeting the surface antigens. Cells were washed twice with CyFACS buffer (PBS with 0.1% BSA, 0.05% sodium azide and 2mM EDTA) and fixed overnight using 2% formaldehyde made in PBS (VWR, Sweden). The broad extended panel of antibodies used are listed in Table S1. For acquisition by CyTOF, cells were stained with DNA intercalator (0.125 μM Iridium-191/193 or MaxPar® Intercalator-Ir, Fluidigm) in 2% formaldehyde made in PBS for 20 min at room temperature. Cells were washed twice with CyFACS buffer, once with PBS and twice with milliQ water. Cells were mixed with 0.1X Norm Beads (EQ™ Four Element Calibration Beads, Fluidigm) filtered through a 35μm nylon mesh and diluted to 1000,000 cells/ml. Samples were acquired using super samplers connected to our CyTOF2 mass cytometers (Fluidigm Inc.) using CyTOF software version 6.0.626 with noise reduction, a lower convolution threshold of 200, event length limits of 10-150 pushes, a sigma value of 3, and flow rate of 0.045 ml/min.

Antibodies and reagents for mass cytometry

The panel of monoclonal antibodies used for this study are indicated in the Key resources table. Monoclonal antibodies were either purchased pre-conjugated from Fluidigm or obtained in carrier/protein-free buffer as purified antibodies that were then coupled to lanthanide metals using the MaxPar X8 polymer conjugation kit (Fluidigm Inc.) as per the manufacturer's recommendations.

Following the protein concentration determination by measurement of absorbance at 280nm on a nanodrop, the metal-labeled antibodies were diluted in Candor PBS Antibody Stabilization solution (Candor Bioscience, Germany) for long-term storage at 4°C.

Born-immune plasma protein profiling

Plasma protein data was generated using Olink assays, a proximity extension assay (Olink AB, Uppsala) (Lundberg et al., 2011) For analysis, 20μL of plasma from each sample was thawed and sent for analysis, either at the plasma protein profiling platform, Science for Life Laboratory, Stockholm or Olink AB in Uppsala. In these assays, plasma proteins are dually recognized by pairs of antibodies coupled to a cDNA-strand that ligates when brought into proximity by its target, extended by a polymerase and detected using a Biomark HD 96.96 dynamic PCR array (Fluidigm Inc.). Four Olink panels (CVD 2, CVD 3, Inflammation and Immune response) have been used as indicated in Key resources table, capturing a total of 355 unique proteins in each plasma sample.

Born-immune fecal metagenomics

DNA was extracted as in IHMS DNA extraction protocol #8 (Costea et al., 2017). According to protocol recommendations 0.2g of faeces were used. Briefly, samples were treated with lysozyme solution and subjected to bead beating using zirconium beads. After centrifugation, DNA is extracted from supernatants with QIAamp Fast DNA Stool Mini Kit (QIAGEN, Cat No. 51604). After DNA extraction, collected DNA was quantified with Qubit (ThermoFisher, Cat No. Q32851) and 10ng were subjected to mechanical fragmentation with the Covaris Focused-ultrasonicator to ensure that fragment sizes were compatible with Illumina sequencing (~300bp average). Sequencing adapters and sample barcodes were incorporated to the DNA fragments using ThruPLEX DNA-seq kit (Rubicon Genomics, Cat No. R400406). ThruPLEX DNA-seq products were purified and size selected by AMPure beads (Beckman Coulter, Cat No. B23318), and DNA concentration and size distribution were inspected with the Qubit dsDNA HS Assay Kit (ThermoFisher, Cat No. Q32851) and the Agilent 2100 Bioanalyzer High Sensitivity kit (Agilent Technologies, Cat No. 5067-4626), respectively. Purified ThruPLEX DNA-seq products were then equimolarly pooled in 4 lanes and subjected to NovaSeq 6000 S4 Illumina Sequencing at the Science for Life Laboratory, Stockholm, Sweden.

T cell polarization experiments

PBMC derived $CD3^+CD4^+CD45RA^+$ naive T cells (CD8/14/19/56 negative) into an enriched culture medium including RPMI 1640 + 10 % FBS + NEAA + 1% Pen-strep + 55 μM β-mercaptoethanol. Cells are added at a concentration of 2 $x10^5$ cells/ml and pre-washed and resuspended 2.5ml T-Activator Dynabeads are added to obtain a bead:cell ratio of 1:2. One milliliter of bead/cell suspension was added to 24-well plate and the following polarizing supplements Th0: No cytokines, Th1: IL-12 (50 ng/ml), Anti-human IL-4 antibody (1 μg/ml) Th2: IL-4 (10 ng/ml), Anti-human IFN-γ antibody (1 μg/ml), Th17: IL-6 (50 ng/ml), IL-23 (20 ng/ml), IL-1β (10 ng/ml), TGF-β1 (5 ng/ml), Anti-hum IL-4 (1 μg/ml), Anti-human IFN-γ antibody (1 μg/ml), iTreg: TGF-β1 (5 ng/ml), IL-2 (10 ng/ml) IFN-β: IFN-β (10 ng/ml). Cells are incubated at 37°C, 5% CO_2 humidified incubator for 5 days and harvested.

Targeted transcriptome and protein by BD Rhapsody single cell RNA sequencing

Polarized T cells in each condition were labeled using BD Single-Cell Multiplexing Kit and BD AbSeq Ab-Oligos reagents strictly following the manufacturers protocol (BD Biosciences). Briefly, cells from each experiment condition were labeled with each sample tag and pooled 18 AbSeq Ab-Oligos. Each sample was then washed twice, counted and resuspended in cold BD Sample Buffer, then calculated the cell number in each sample, then pooled the required number of cells from each sample to get approximately 20,000 cells in 620 μL for each cartridge (around 9 samples for each cartridge). After priming the nanowell cartridges, the pooled sample was loaded onto BD Rhapsody cartridges and incubated at room temperature. Cell Capture Beads were prepared and then loaded onto the cartridge. According to the manufacturers protocol, cartridges were washed, cells were lysed, and Cell Capture Beads were retrieved and washed prior to performing reverse transcription and treatment with Exonuclease I. cDNA Libraries were prepared using mRNA Targeted, Sample Tag, and BD AbSeq Library Preparation with the BD Rhapsody Targeted mRNA and AbSeq Amplification Kits and protocol. In brief, cDNA targeted amplification using the Human T cell Expression Panel primers via PCR. mRNA PCR products were separated from sample tag and AbSeq products with double-sided size selection using AMPure XP magnetic beads (Beckman Coulter). mRNA and Sample Tag products were further amplified using PCR. PCR products were then purified using AMPure XP magnetic beads. Quality and quantity of PCR products were determined by using an Agilent 2100 Bioanalyzer and Qubit Fluorometer using the Qubit dsDNA HS Kit (ThermoFisher). Targeted mRNA product was diluted to 2.5 ng/μL and sample tag and AbSeq PCR products were diluted to 1ng/μL to prepare final libraries. Final libraries were indexed using PCR. Index PCR products were purified using AMPure XP magnetic beads. Quality of final libraries was assessed by using Agilent Bioanalyzer and quantified using a Qubit Fluorometer. Final libraries were diluted to 2nM for paired-end (150bp) sequencing on a NovaSeq sequencer (Illumina).

IMPRINT—absolute quantification of B. infantis by quantitative real-time PCR

As previously described in Frese et al. (2017), quantification of the total *B. infantis* was performed by quantitative real-time PCR using Blon_0915 primers Blon0915F (5′- CGTATTGGCTTTGTACGCATTT −3′), Blon0915R (5′- ATCGTGCCGGTGAGATTTAC −3′) and BI915 PRB (5′- 6-FAM–CCAGTATGG–ZEN–CTGGTAAAGTTCACTGCA–3IABkFQ). Each reaction contained 10 μL of 2 × TaqMan Universal Master Mix II with UNG master mix (Applied Biosystems), 0.9 μm of each primer, 0.25 μM probe and 5 μL of template DNA. Thermal cycling was performed on a QuantStudio 3 Real-Time PCR System and consisted of an initial UNG activation step

of 2 minute at 50°C followed by a 10-minute denaturation at 95°C succeeded by 40 cycles of 15 s at 95°C and 1 min at 60°C. Quantitative PCR was carried out using standard curves of known B. infantis EVC001 cultures prepared by serial dilution. All samples including the standard curve were ran in duplicate.

IMPRINT—fecal cytokine measurements

Interleukin (IL)-4, IL-12p70, IL-13, IL-17A, IL-21, IL-23, IL-27, IL-31, IL-33, IFNβ, and MIP3α were quantified from 80 mg of stool diluted 1:10 in Meso Scale Discovery (MSD; Rockville, MD) diluent using the U-PLEX Inflammation Panel 1 (human) Kit according to the manufacturer's instructions. Standards and samples were measured in duplicate and blank values were subtracted from all readings. Assays were performed at least twice.

IMPRINT—fecal water preparation

Historical fecal samples from term infants either colonized with *B. infantis* EVC001 or not were collected and stored in −80°C until processing. Pooled fecal samples (minimum of 3) from infants colonized with *B. infantis* EVC001 or not were weighed and diluted 25% w/v in sterile PBS and vortexed for 1 min allowing stool to thaw and form a homogeneous slurry. Fecal slurries were then centrifuged for 30min at 4,000 RPM at 4°C. The supernatant was collected and spun again for 3 hours at 12,000 RPM at 4°C. The supernatant was further collected and then serially filtered (40um cell strainer, 1um, 0.45um, and 0.22um). Filtered fecal waters were stored in −80°C until use.

IMPRINT—fecal metabolomics

Fecal samples were sent to Metabolon, Inc. (Durham, NC) for non-targeted metabolite profiling. Forty fecal samples, control (n = 20) and *B. infantis* EVC001-fed (n = 20) from day 21 postnatal were collected and processed for non-targeted metabolomics profiling, as shown previously (Call et al., 2018). Briefly, samples were exposed to a combination of aqueous and organic solvents to extract small molecules. Residual organic solvent was removed using a TurboVap (Zymark), and the fecal extracts were lyophilized and divided equally for then equal GC/MS and UPLC-MS/MS analysis in parallel. Extracts were derivatized with bistrimethyl-silyl-triflouroacetamide and analyzed using a Trace DSQ (Thermo-Finnigan) mass spectrometer Fecal extracts were analyzed under both acidic and basic conditions using an ACQUITY (Waters) UPLC and an LTQ (Thermo-Finnigan) mass spectrometer.

QUANTIFICATION AND STATISTICAL ANALYSIS

Olink preprocessing

Plasma protein data was batch corrected and normalized on the basis of NPX values across batches with available bridge samples.

Mass cytometry preprocessing

All FCS-files unrandomized using CyTOF software (version 6.0.626) were transferred with no further preprocessing. An automated cell classification supervised algorithm, Grid was used to first manually define reference subpopulations and then train a learning algorithm (XGBoost) to recognize the same subsets of cells in novel data, providing a rapid and robust cell classification method for high-dimensional Mass cytometry datasets (Chen et al., 2020). FCS files and relevant phenotypic markers were used in order to manually gate cell populations to be used as a reference, then the reference was used to train a classifier algorithm to categorize similar cells. The output results in a dataframe with samples as columns and cell sub-populations as rows. Outputs from all Mass cytometry experiments were merged, and batch differences removed using limma (Ritchie et al., 2015).

Born-Immune metagenome data—quality filtering and host removal

347 and 60 demultiplexed fastq files from the Born-immune and IMPRINT cohort respectively were downstream processed using the same pipeline and parameters. Demultiplexed samples were quality filtered using fastp v0.20.0 (Chen et al., 2018), and host contamination removed using Kraken v2.0.8_beta (Wood et al., 2019) by mapping against the NCBI's GRCh38.p13 database. Both steps were run using default settings in StaG-mwc v.0.4.1 (https://doi.org/10.5281/zenodo.1483891).

Born-Immune metagenome data—taxonomic and functional profiling

Taxonomic profiles were established using MetaPhlAn v3.0.5 (Beghini et al., 2021) and functional HMO profiles generated with HUMAnN2 v.2.8.1 (Franzosa et al., 2018) bypassing all steps except "nucleotide-search" and "evalue 0.00001" with a customized nucleotide database of HMO genes instead of the chochophlan database. RPKs from HUMAnN2 were normalized to cpms using 'humann2_renorm_table'. Both taxonomic and functional profiling were incorporated into StaG-mwc.

Bifidobacteria abundance correlation

Immune cell and plasma protein data was reduced to include infant samples collected 56 to 152 days after birth with correlating metagenomics data, resulting in n = 18 and n = 19 samples respectively. Fold change was calculated, and variables re-ordered in descending value order.

The correlation matrices were built using library *corrplot* with spearman method, for comparison purposes low bifidobacterial matrix was reordered to the level order of high bifidobacterial matrix.

HMO correlation

Samples were binned according to days after birth while HMO utilization genes were clustered in accordance with pathway and function. The CPM counts were binned as well with increasing ranges and heatmap was built using library *superheat*. Correlations with individual cytokines were based on NPX values and CPM counts. ANOVA test was performed for Spearman correlation performed between individual cytokines and HMO utilization genes.

Targeted transcriptomics processing

FASTQ files of targeted transcriptomics data were processed on the Seven Bridges platform using the Targeted Analysis Pipeline v1.9 (BD Biosciences)(https://www.sevenbridges.com). R1 and R2 are filtered removing low quality sequencing reads, checking read lengths as well as lengths of strings of identical bases. Read pair is removed if read length of R1 is less than 66 bases or R2 is less than 64 bases. R1 reads are annotated to cell label sequences and unique molecular identifiers (UMI), perfect matches are kept while others will be held for further filtering. R2 reads are annotated to oligo sequence to genes on targeted panel by Bowtie2. Then, all valid R1 and R2 read pairs are collapsed into unique raw molecules. For all analysis, we used output of distribution-based error correction (DBEC) as a means to correct for both artifacts in PCR cycles and sequencing errors. Expression matrices containing DBEC-adjusted molecule counts after sample tag assignment were used for downstream analysis.

Analysis of Seurat object with targeted data

The expression matrices were read into the R package Seurat v3 (Butler et al., 2018) where they were merged and split between RNA and antibody (Ab) assays thereafter using scripts from https://github.com/MairFlo/Targeted_transcriptomics (Mair et al., 2020). After creating Seurat object that included features that were detected in at least 3 cells and cells that were detected in at least 50 features within the RNA assay. In the first experiment relating to bifido, out of 53,184 cells, 2,911 were called as multiples and 877 events as undetermined. Multiples and undetermined cells were removed from the analysis. In the second experiment relating to ILA, out of 42,151 cells, 5,539 were called as multiples and 17 events as undetermined. On RNA assay, a natural log normalization was performed with a scale factor of 10,000 while a centered log-ratio normalization was performed on the Ab assay. Both assays were linearly scaled to remove uninteresting sources of variation like batch effect. Additionally, RNA assay was scaled to regress out the total number of molecules identified within a cell as well as the effect of GAPDH gene. The effect of the GAPDH gene was regressed out by computing the fraction of counts from that gene. All genes or proteins were used for dimensionality reduction using UMAP and clustering.

Partition-based graph abstraction of single-cell data

Partition-based graph abstraction (PAGA) (Wolf et al., 2019) was utilized to demonstrate the topology abstraction of single-cell RNA data. In brief, PCA was first applied to reduce the dimension of RNA data to 20, and then a kNN-like graph was built with the approximate nearest neighbor search. Afterward, the highly connected communities in the kNN-like graph were discovered with Leiden method (Traag et al., 2019), which were further utilized by PAGA to infer a trajectory map, which demonstrates the topology relationship of those highly connected communities. Finally, the trajectory map was used as the initial position and the scatters of single cells were embedded with ForceAtlas2 (Jacomy et al., 2014) for visualization.

mRNA-seq data analysis

Quality control for the bulk RNA-sequencing of FACS-sorted immune cell populations was provided by the National Genomics Infrastructure (NGI) at Science for Life Laboratory, Stockholm, Sweden. First, we quantified abundances of transcript sequences in FASTA format by generating abundance estimates for all samples using the Kallisto software (Bray et al., 2016). Also, gene abundance estimates were performed by summing the transcript expression (TPM) values for the transcripts of the same gene. Since DESeq2 expects count data from the Kallisto output the *tximport* package was used to convert these estimates into read counts. DESeq2 was performed as a basis for differential gene expression analysis based on the negative binomial distribution (Love et al., 2014). We employed a design to demonstrate differential gene expression between circulating CD38-CD4-CD62Lneg and memory CD4T cells over time. Low gene counts (< 100) were filtered out and variance stabilizing transformation (VST) was performed on the count data.

Gene set enrichment analysis (GSEA)

Gene set enrichment analysis (GSEA) was performed to identify transcriptomic differences occurring over time in circulating $CD38^{+}CD62L^{-}$ memory $CD4^{+}$ T cells versus Total memory $CD4^{+}$ T cells. The R package fgsea was used to find the most highly enriched hallmark pathways through *gmtPathways* function (Subramanian et al., 2005). Pathways of interest were subsequently isolated and all genes within each pathway observed using volcano plots.

IMPRINT—fecal microbiome analyses

As previously described in Casaburi et al. (2019), shotgun metagenomic libraries were prepared and sequenced from DNA previously extracted from approximately 100 mg of frozen stools collected from infants on day 21 postnatal(Frese et al., 2017). Briefly, the DNA was subjected to bead beating prior to column purification using a Zymo Fecal DNA Miniprep kit, according to the manufacturer's instructions. Metagenomic shotgun library preparation and sequencing was performed at the California Institute for Quantitative Biosciences (QB3) (University of California, Berkeley) on an Illumina HiSeq 4000 platform using a paired-end sequencing approach with a targeted read length of 150 bp and an insert size of 150bp (Casaburi et al., 2019).

Additionally, 16S rRNA libraries were generated and sequenced (Frese et al., 2017). Briefly, the V4 region of the 16S rRNA gene was amplified and sequenced using primers 515f and 806r as previously described with recent modifications (Caporaso et al., 2011; Walters et al., 2015). Paired-end DNA (300 bp) sequencing was performed at the UC Davis Genome Center on an Illumina MiSeq system.

IMPRINT—quality filtering and removal of human sequences from shotgun metagenomes

Demultiplexed fastq sequences were quality filtered, including adaptor trimming using Trimmomatic v0.36 (Bolger et al., 2014) with default parameters. Quality-filtered sequences were screened to remove human sequences using GenCoF v1.0 (Czajkowski et al., 2019) against a non-redundant version of the Genome Reference Consortium Human Build 38, patch release 7 (https://www.ncbi.nlm.nih.gov/grc/human/data?asm=GRCh38.p7)www.ncbi.nlm.nih.gov. Human sequence-filtered raw reads were deposited in the Sequence Read Archive (SRA; https://www.ncbi.nlm.nih.gov/sra) under the reference number, PRJNA390646. Taxonomic profiling of the metagenomic samples was performed using MetaPhlAn2 (Truong et al., 2015), which uses a library of clade-specific markers to provide pan-microbial (bacterial, archaeal, viral, and eukaryotic) profiling (http://huttenhower.sph.harvard.edu/metaphlan2). Strain characterization was performed using PanPhlan (Scholz et al., 2016) which is used in combination with MetaPhlAn2 to characterize strain-level variants in marker genes for a selected organism. For PanPhlan analysis, the pangenomes from *Bifidobacterium longum* https://bitbucket.org/CibioCM/panphlan https://github.com/segatalab/panphlan were used as a reference. Both MetaPhlAn2 and PanPhlan were used with their default settings as described in the updated global profiling of the Human Microbiome Project (2017)(Lloyd-Price et al., 2017).

IMPRINT—fecal cytokine analysis

All detectable biomarker values were included as continuous data in the analyses; however, values below level of detection (< 20% of all cytokines measurements) were generated below the level of quantification to justify parametric statistics. Fecal cytokine concentrations were determined using calibration curves to which electrochemiluminescence signals were backfitted. Final concentrations were calculated using the Sector Imager 2400 MSD Discovery Workbench analysis software, as was previously published (Henrick et al., 2019; Nguyen et al., 2021).

IMPRINT—fecal metabolomics analyses

Fecal compounds were identified by comparison of the raw data with Metabolon's curated library of standards. The values for compounds in the fecal samples were normalized by the dry mass of the sample and missing values were imputed with half the compound minimum. Absolute compound intensity values were used to calculate fold differences between controls and EVC001-fed samples, while for all other analyses, the values were transformed using the generalized log transformation then mean-centered and scaled by the standard deviation.

IMPRINT—statistical analysis

All statistical analyses were performed in R v3.6.2. To assess infant gut microbiome differences across ages, principle coordinate analysis (PCoA) was performed on a Bray-Curtis dissimilarity matrix calculated from infant bacterial taxa relative abundance. The first two principle coordinate axes were plotted and colored by day of life. Cytokine vs bacterial relative abundance values were Spearman correlated with false discovery rate (FDR) correction. FDR adjusted Wilcoxon rank-sum tests were used to compare baseline (Day 6) cytokine levels to Day 60 cytokine levels between treatments. Median cytokine values were calculated for Days 6 and 60, then 0-1 normalized within each cytokine for graphical visualization purposes. Demultiplexed sequencing data was downloaded from the NCBI SRA (PRJNA390646). Sequencing data from samples utilized in the present study were subjected to analysis in qiime2 (Boylen, 2019). Samples were quality filtered and denoised using deblur (Amnon, 2017). Phylogenic alignments were conducted with mafft (Katoh, 2019) and fasttree (Price, 2009) and Shannon Diversity and a Weighted Unifrac distance matrix (Lozupone, 2010) were created using the q2-diversity plugin in qiime2, rarified to the minimum number of dereplicated sequences passing quality filter for the samples (1208 reads per sample). Shannon diversity values were compared using a Wilcoxon test and graphed using ggpubr (v. 0.4.0) and ggplot2 in R (v4.0.4). Analysis of variance using distance matrices (adonis) was used to compare beta diversity by treatment group. The P values throughout the manuscript are represented by asterisks (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$).

CellPress
OPEN ACCESS

Supplemental figures

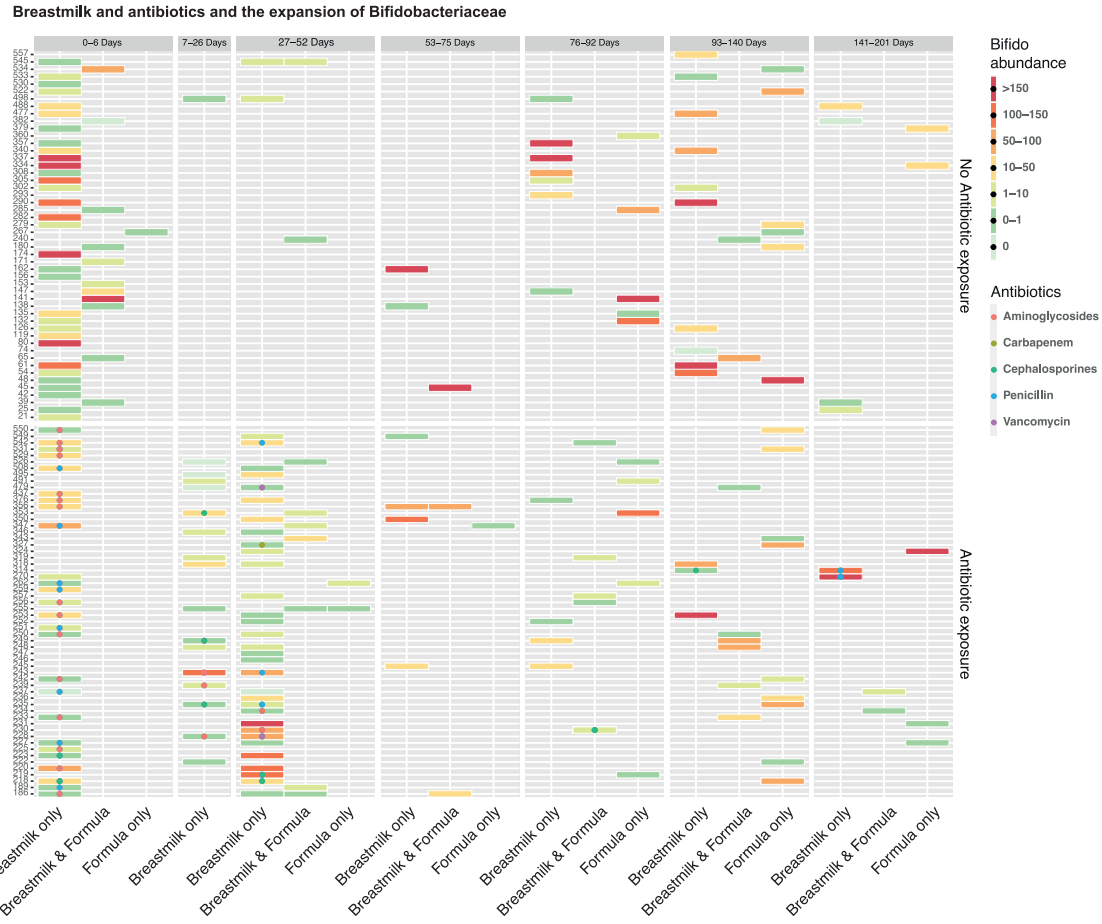


Figure S1. Bifidobacteriaceae abundance in relation to breastfeeding and antibiotic exposure, related to Figure 2

For each child (rows) with available information on feeding type (columns) and antibiotic exposure (colored dots) relative abundance of *Bifidobacteriaceae* is shown. Data are separated by time windows in early life.

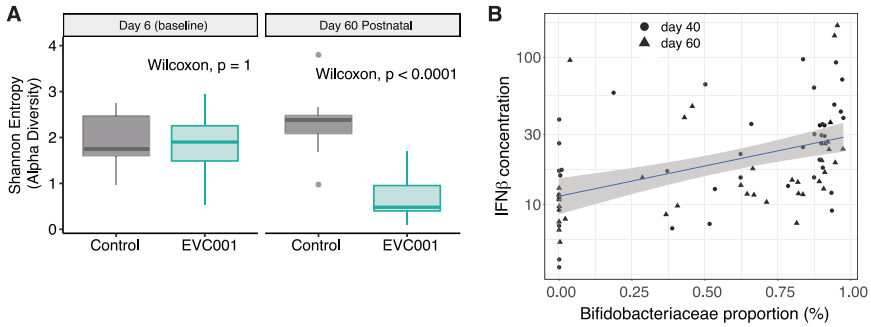


Figure S2. Microbiome associations with fecal cytokines, related to Figure 5

(A) Boxplots of Shannon Diversity index calculated on samples (n = 80) at Day 6 and Day 60, compared between groups (control, gray; EVC001-fed, teal) by Wilcox test. (B) Correlation of fecal IFNβ concentration and *Bifidobacterium* relative abundance at day 40 and 60 postnatal (R = 0.66, p = 1.2e-05). Each cytokine was tested in duplicate. Statistical analysis was completed using Wilxocon rank-sum test. *P* values were adjusted using Bonferroni-Holm method and considered statistically significant if *p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001.

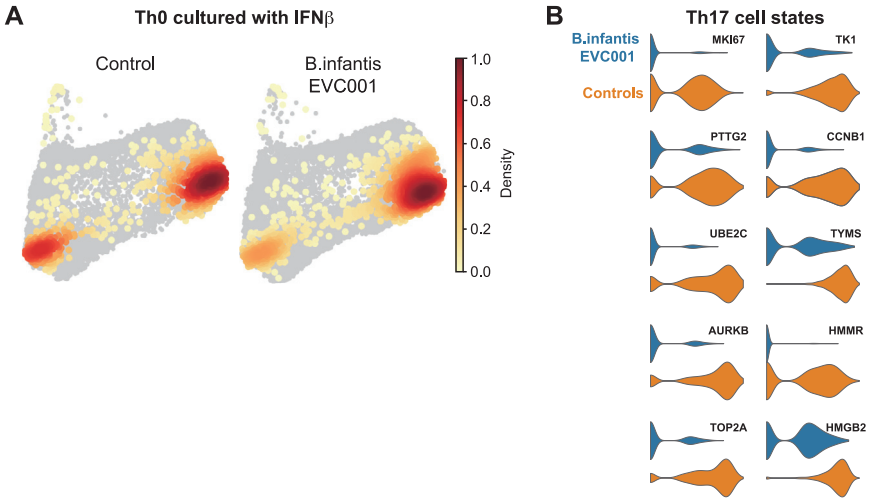


Figure S3. CD4⁺ T cell polarization under the influence of microbial metabolites and IFNβ, related to Figure 6

(A) CD4⁺ T cell polarization *in vitro* in the presence of fecal water from children given B.infantis EVC001 or control and additional IFNβ. T cells at the end of the culture shown as PAGA density plots and separated by culture condition. (B) Top genes differentially expressed among Th17-induced states in B.infantis EVC001 treated or control infants fecal water cultures.