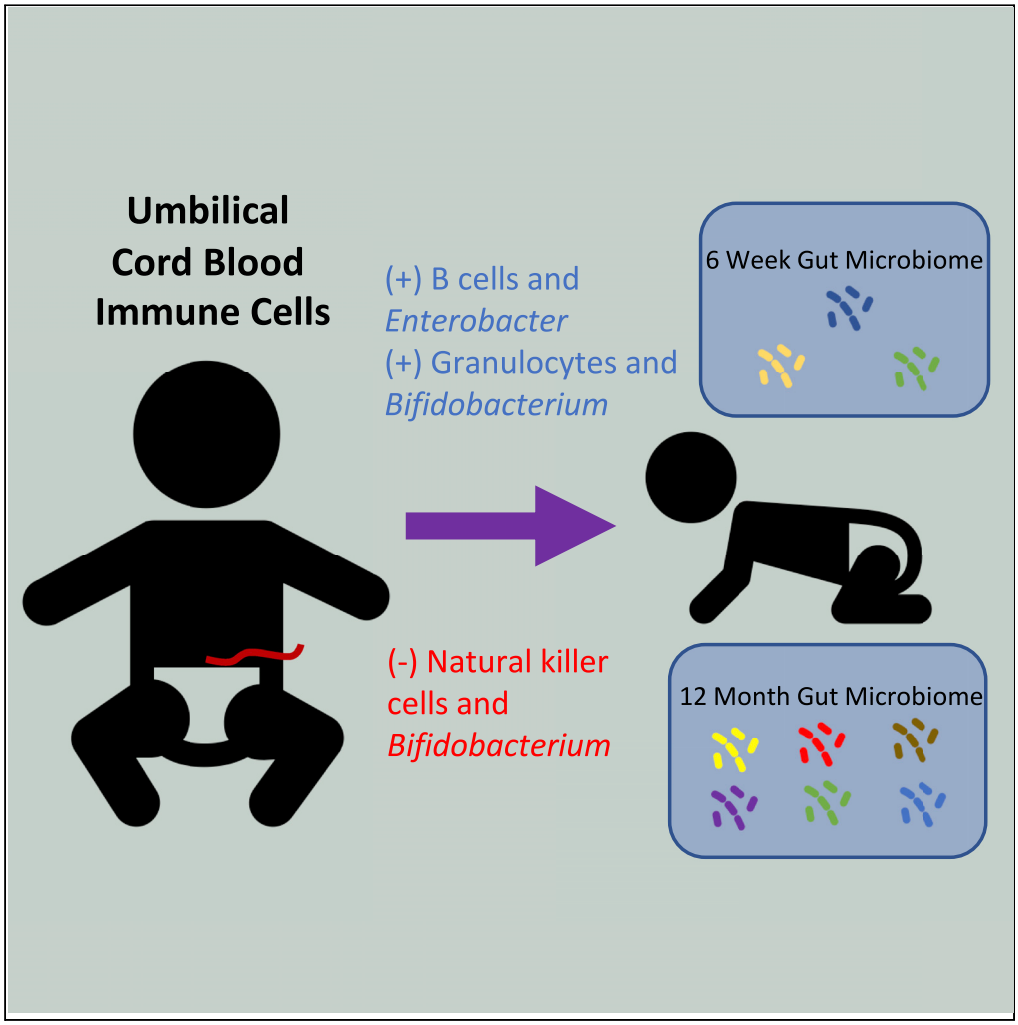
iScience

CellPress
OPEN ACCESS

Article

Umbilical cord blood immune cell profiles in relation to the infant gut microbiome



Yuka Moroishi, Lucas A. Salas, Jie Zhou, ..., Juliette Madan, Jiang Gui, Margaret R. Karagas

margaret.r.karagas@dartmouth.edu

Highlights

B cells in the umbilical cord blood are positively associated with *Enterobacter*

Monocytes and natural killer cells are negatively associated with *Bifidobacterium*

Granulocytes are positively associated with *Bifidobacterium*

Moroishi et al., iScience *26*, 105833
January 20, 2023
https://doi.org/10.1016/j.isci.2022.105833

iScience

CellPress
OPEN ACCESS

Article

Umbilical cord blood immune cell profiles in relation to the infant gut microbiome

Yuka Moroishi,[1,2] Lucas A. Salas,[1] Jie Zhou,[1,2] Emily R. Baker,[3] Anne G. Hoen,[1,2] Todd M. Everson,[4] Carmen J. Marsit,[4] Juliette Madan,[1,5] Jiang Gui,[2] and Margaret R. Karagas[1,*]

SUMMARY

During infancy, the interplay between the developing immune system and the microbiome is critical. We examined whether blood immune cell composition at birth in the umbilical cord (inferred by DNA methylation profiling) related to the early infant gut microbiome (assessed by 16S rRNA gene sequencing) among 73 infants in the New Hampshire Birth Cohort Study. We used generalized estimating equations and controlled for false discovery rate to select microbial taxa associated with immune cells. We found associations between the infant gut microbiome and immune cells, including a positive association between B cells and *Enterobacter*, a negative association between natural killer cells and *Bifidobacterium*, and a positive association between granulocytes and *Bifidobacterium*. Our findings give clues that immune profiles at the time of birth as measured in umbilical cord blood are associated with the development of the gut microbiome in early life.

INTRODUCTION

The bidirectional relationship between the infant immune system and the microbiome plays a critical role in developing both innate and adaptive immunity.[1] Neonates rely on innate immunity, including natural killer (NK) cells and monocytes, to fight against infections.[2] The umbilical cord blood provides a rich source of immune cells and comprises more NK cells and immature T-lymphocytes compared to adult peripheral blood[3,4] and are representative of a newborn's blood. Studies have demonstrated differences in immune profiles between infants born preterm and those born term, suggesting higher susceptibility of preterm infants to infections.[5–7] Furthermore, the early gut microbiome has been associated with adverse health outcomes from early life to adulthood.[8] However, it is unclear whether an infant's early immune profile influences the development of their gut microbiome.

RESULTS

Of infants in our study, 73 had provided both cord blood and stool samples at six weeks. Forty-four had provided stool samples at both 6 weeks and 12 months of age, totaling 117 stool samples. Based on 16S rRNA gene sequencing, we identified 336 microbes at the genus level or family level if the genus could not be determined. We identified seven immune cell types in cord blood from the DNA methylation data, including B cells, $CD4^+$ T-lymphocytes, $CD8^+$ T-lymphocytes, monocytes, NK cells, nucleated red blood cells (NRBC), and granulocytes.

Baseline characteristics

Our study population included more male infants than female infants (Table 1). The mean maternal enrollment age was 32.2 years, and most mothers did not smoke during infancy (Table 1). The average infant's gestational age at birth was 39.1 weeks (Table 1). The mean alpha diversities of six-week and 12-month stool microbiome samples were 1.37 and 2.08, respectively (Table 1). Estimated using DNA methylation array data, granulocytes were the highest proportion of immune cells in the cord blood, followed by $CD4^+$ T-lymphocytes and monocytes (Table 1).

Associations between cord blood immune cells and the infant gut microbiome

Based on generalized estimating equations (GEE) adjusted for delivery mode, infant sex, infant birth weight, gestational age, maternal enrollment age, maternal smoking during pregnancy, and alpha diversity of infant's

[1]Department of Epidemiology, The Geisel School of Medicine at Dartmouth, Hanover, NH 03756, USA

[2]Department of Biomedical Data Science, The Geisel School of Medicine at Dartmouth, Hanover, NH 03756, USA

[3]Department of Obstetrics and Gynecology, Dartmouth-Hitchcock Medical Center, Lebanon, NH 03766, USA

[4]Gangarosa Department of Environmental Health, Emory University Rollins School of Public Health, Atlanta, GA 30322, USA

[5]Department of Pediatrics, Children's Hospital at Dartmouth, Lebanon, NH 03766, USA

*Correspondence: margaret.r.karagas@dartmouth.edu

https://doi.org/10.1016/j.isci.2022.105833

Table 1. Baseline characteristics of participants included in the study, N = 73

Covariate	No. (%) or Mean (SD)
Infant characteristics	
Infant sex	
Male	50 (68.5%)
Female	23 (31.5%)
Birth weight (g)	3454.84 (520.64)
Gestational age (weeks)	39.11 (1.42)
Delivery mode	
Vaginal	55 (75.3%)
Cesarian	18 (24.7%)
Maternal characteristics	
Maternal enrollment age (years)	32.2 (4.2)
Maternal smoking during pregnancy[a]	
Yes	3 (4.2%)
No	68 (95.8%)
Infant gut microbiome alpha diversity	
Shannon diversity index at 6 weeks	1.37 (0.45)
Shannon diversity index at 12 months[b]	2.08 (0.39)
Cord blood immune cell proportions	
B cell	6.76 (2.57)
$CD4^+$ T-lymphocytes	13.86 (5.05)
$CD8^+$ T-lymphocytes	4.06 (2.36)
Monocyte	9.74 (2.25)
NK	6.40 (3.20)
NRBC	9.79 (8.72)
Granulocyte	51.12 (9.79)

Abbreviations: No. number; SD standard deviation; g grams; NK natural killer; NRBC nucleated red blood cells.
[a]N = 71 for maternal smoking during pregnancy due to missing.
[b]We collected 12-month microbiome samples for 44 infants.

gut, several taxa were nominally associated with immune cells ($p < 0.05$). Increased abundance of B cells in the cord blood was positively associated with an abundance of *Enterobacter* ($p = 0.002$), *Anaerostipes* ($p = 0.015$), and *Streptococcus* ($p = 0.019$) (Figure 1, Table S1). Monocytes were negatively associated with *Bifidobacterium* ($p = 0.015$) (Figure 1, Table S1). Higher proportions of NK cells were positively associated with Enterobacteriaceae ($p = 0.007$) and negatively associated with *Bifidobacterium* ($p < 0.001$) (Figure 1, Table S1). NRBC were positively associated with *Flavonifractor* ($p = 0.017$), *Blautia* ($p = 0.012$), and Lachnospiraceae ($p = 0.020$) (Figure 1,Table S1). Granulocytes were negatively associated with *Anaerostipes* ($p = 0.020$) and *Flavonifractor* ($p = 0.037$) and positively associated with *Bifidobacterium* ($p = 0.002$) (Figure 1, Table S1). No associations were found for $CD4^+$ T-lymphocytes and $CD8^+$ T-lymphocytes. After false discovery rate (FDR) correction for multiple hypotheses, three associations were found to be statistically significant: B cells and *Enterobacter*, NK cells and *Bifidobacterium*, and granulocytes and *Bifidobacterium*.

Sensitivity analyses

We conducted a sensitivity analysis on 54 infants born vaginally, because both cord blood immune cells and infant gut microbiome compositions can vary by delivery mode.[9,10] The three associations found in the original analysis were also found in the subset of infants delivered vaginally. Furthermore, some associations that were nominally significant were statistically significant at the FDR level (Table S2). These include the positive association between B cell and *Anaerostipes* ($p = 0.004$, FDR-adjusted $p = 0.036$) and the positive association between NK cells and Enterobacteriaceae ($p < 0.001$, FDR-adjusted $p = 0.002$) (Table S2). Some additional nominally associated relations were observed for NRBC and Enterobacteriaceae,

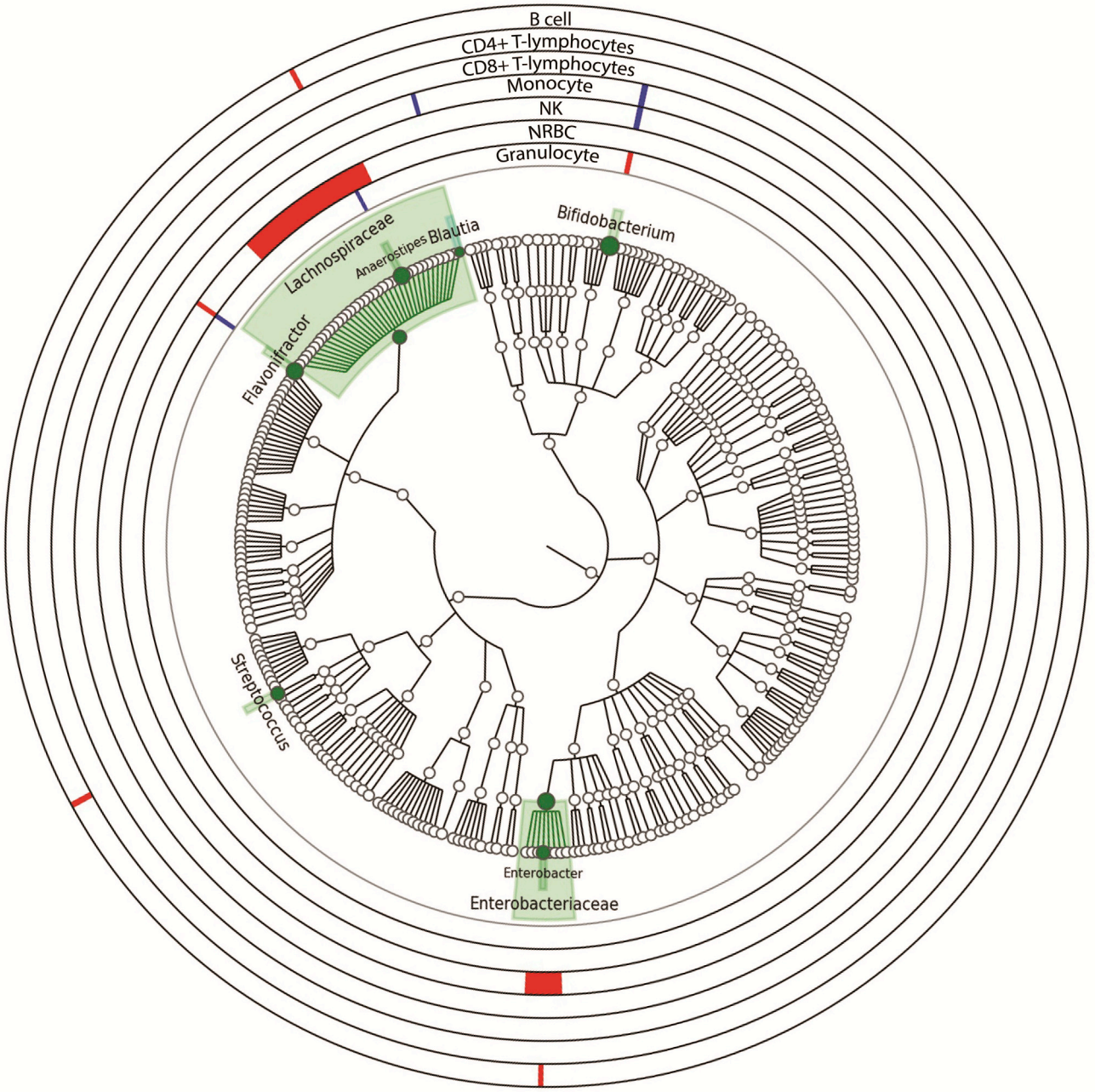


Figure 1. Associations between cord blood immune cells and bacterial family and genera in the infant gut microbiome

Cladogram plot for the relationship between cord immune cells and the gut microbiome. Red represents the positive coefficient of the immune cell on the abundance of taxa, whereas blue represents the negative coefficient of the immune cell on the abundance of taxa. Abbreviations: *NK* natural killer; *NRBC* nucleated red blood cells.

Staphylococcus, and *Clostridium sensu stricto*, but they were not statistically significant after FDR correction (Table S2). Results for single time period analyses for 6-week stool and 12-week stool are provided in Tables S3 and S4 respectively.

DISCUSSION

In our prospective study of infants in the New Hampshire Birth Cohort Study (NHBCS), we found differences in immune cell proportions in the cord blood related to the developing gut microbiome. Abundances of

B cells, monocytes, NK cells, nucleated red blood cells, and granulocytes were associated with several bacterial families and genera in the infant's gut, including *Bifidobacterium*.

The immune system in the newborn is immature and may be affected by several external factors. A previous study of infants in the NHBCS found that maternal exposure to arsenic and cadmium were associated with T cell proportions in cord blood.[11] Infants delivered following uncomplicated pregnancies and delivered vaginally demonstrated higher counts of NK cells and granulocytes because of the physical stress and normal cortisol release during labor compared to cesarean section born babies.[9,12]

Studies suggest a possible link between the early immune system and the gut microbiome; the innate and adaptive immune cells of neonates evolve with maturing intestinal epithelial cells.[13] In this bidirectional relationship, the gut microbiome aids in the maturation of the early immune system whereas the immune system regulates host-microbe symbiosis.[1,14] The immature immune system's blunted cytokine response and T cell and B cell production allows the establishment and tolerability of microbiota in the gut.[15,16] Mice and human studies have shown differential colonization of bacteria depending on presence of IgA.[17,18] Furthermore, many Th cells produce cytokines that aid in protecting the intestinal epithelium.[19] Previous studies have shown that reduction in Bifidobacteriaceae in the gut microbiome was associated with expanded populations of neutrophils, basophils (both of them granulocytes), plasmablasts, and memory $CD8^{+}$T cells, indicating an activation of the newborns' immune system.[20] Less is known about NK and *Bifidobacterium*. Cesarean section has been shown to be associated with both lower levels of leukocytes, hemoglobin, cytokines in cord blood and lower *Bifidobacterium* in transitional stool compared to vaginal deliveries.[21] On the other hand, the early colonization by *Escherichia coli* and other Enterobacteriaceae has been associated with changes in B cell modulation later in life, mainly related to allergy responses that are more common in cesarean section than in vaginal deliveries.[22] Our analysis indicates that newborns' immune proportions in cord blood have an association with later gut microbiome for infants delivered vaginally.

A strength of our study is the use of a novel procedure using DNA methylation arrays to infer immune cell composition in cord blood.[23,24] This allowed us to estimate the immune system cell proportions of newborn infants using cell DNA instead of intact cells for cell proportions quantification, using a validated method versus traditional flow cytometry.[25] Unlike traditional flow cytometry, this does not limit the collection to fresh samples with intact cell surface markers, neither to the number of color channels and antibodies selected for a specific experiment, or the potential biases during gating procedures.[26] In the future, additional cell specific analyses could be incorporated, as newer more comprehensive libraries become available.[27]

Limitations of the study

One limitation of our analyses is the sample size. Although we found associations in our sensitivity analyses on infants delivered vaginally, a more comprehensive stratified analysis with a larger sample size is warranted to find specific associations within vaginal and cesarean deliveries. Furthermore, we cannot establish a direct causal link between the immune composition and the later gut microbiome. Another limitation includes the possibility of unmeasured confounding that may impact cord blood immune cell proportions and the infant gut microbiome. Finally, we could not identify specific species and strain of bacteria associated with immune cell proportions; further analyses with metagenomics data are warranted.

In conclusion, our data indicate that early adaptations in the newborn's immune system may play a role in the infant's later establishment of some of the gut microbiome population. Additional longitudinal measurements of changes in the gut microbiota and immune cell composition may help to further understand the bidirectional relationship between immunity and the gut microbiome.

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

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2022.105833.

ACKNOWLEDGMENTS

This work was supported in part by the NIH grants P20GM104416, UH3OD023275, P01ES022832, P42ES007373, R01LM012723, and US EPA grant RD-83544201. L.A.S. was supported by CDMRP/Department of Defense (W81XWH-20-1-0778) and NIGMS (P20 GM104416-09/8299). We thank all the participants and staff of the New Hampshire Birth Cohort Study for their collaboration.

AUTHOR CONTRIBUTIONS

Conceptualization, Y.M. and M.R.K.; Investigation, E.R.B., A.G.H., C.J.M, T.M.E, J.C.M., and M.R.K.; Methodology, Y.M., L.A.S., J.Z., and J.G.; Formal Analysis, Y.M. and J.Z.; Visualization, Y.M. and J.Z., Writing – Original Draft, Y.M., L.A.S., and J.Z.; Writing – Review and Editing, Y.M., L.A.S., J.Z., E.R.B., A.G.H., C.J.M., T.M.E., J.C.M., J.G., and M.R.K.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: April 29, 2022
Revised: November 14, 2022
Accepted: December 15, 2022
Published: January 20, 2023

REFERENCES

1. Zheng, D., Liwinski, T., and Elinav, E. (2020). Interaction between microbiota and immunity in health and disease. Cell Res. *30*, 492–506. https://doi.org/10.1038/s41422-020-0332-7.

2. Yu, J.C., Khodadadi, H., Malik, A., Davidson, B., Salles, É.d.S.L., Bhatia, J., Hale, V.L., Baban, B., and Baban, B. (2018). Innate immunity of neonates and infants. Front. Immunol. *9*, 1759. https://doi.org/10.3389/fimmu.2018.01759.

3. Harris, D.T., Schumacher, M.J., Locascio, J., Besencon, F.J., Olson, G.B., DeLuca, D., Shenker, L., Bard, J., and Boyse, E.A. (1992). Phenotypic and functional immaturity of human umbilical cord blood T lymphocytes. Proc. Natl. Acad. Sci. USA *89*, 10006–10010. https://doi.org/10.1073/pnas.89.21.10006.

4. D'Arena, G., Musto, P., Cascavilla, N., Di Giorgio, G., Fusilli, S., Zendoli, F., and Carotenuto, M. (1998). Flow cytometric characterization of human umbilical cord blood lymphocytes: immunophenotypic features. Haematologica *83*, 197–203.

5. Correa-Rocha, R., Pérez, A., Lorente, R., Ferrando-Martínez, S., Leal, M., Gurbindo, D., and Muñoz-Fernández, M.Á. (2012). Preterm neonates show marked leukopenia and lymphopenia that are associated with increased regulatory T-cell values and diminished IL-7. Pediatr. Res. *71*, 590–597. https://doi.org/10.1038/pr.2012.6.

6. Marchant, E.A., Kan, B., Sharma, A.A., van Zanten, A., Kollmann, T.R., Brant, R., and Lavoie, P.M. (2015). Attenuated innate immune defenses in very premature neonates during the neonatal period. Pediatr. Res. *78*, 492–497. https://doi.org/10.1038/pr.2015.132.

7. Anderson, J., Thang, C.M., Thanh, L.Q., Dai, V.T.T., Phan, V.T., Nhu, B.T.H., Trang, D.N.X., Trinh, P.T.P., Nguyen, T.V., Toan, N.T., et al. (2021). Immune profiling of cord blood from preterm and term infants reveals distinct differences in pro-inflammatory responses. Front. Immunol. *12*, 777927. https://doi.org/10.3389/fimmu.2021.777927.

8. Milani, C., Duranti, S., Bottacini, F., Casey, E., Turroni, F., Mahony, J., Belzer, C., Delgado Palacio, S., Arboleya Montes, S., Mancabelli, L., et al. (2017). The first microbial colonizers of the human gut: composition, activities, and health implications of the infant gut microbiota. Microbiol. Mol. Biol. Rev. *81*. 000366-e117. https://doi.org/10.1128/MMBR.00036-17.

9. De Amici, D., Gasparoni, A., Chirico, G., Ceriana, P., Bartoli, A., Ramajoli, I., and Moretta, A. (1999). Natural killer cell activity and delivery: possible influence of cortisol and anesthetic agents. Neonatology *76*, 348–354. https://doi.org/10.1159/000014178.

10. Madan, J.C., Hoen, A.G., Lundgren, S.N., Farzan, S.F., Cottingham, K.L., Morrison, H.G., Sogin, M.L., Li, H., Moore, J.H., and Karagas, M.R. (2016). Association of cesarean delivery and formula supplementation with the intestinal microbiome of 6-week-old infants. JAMA Pediatr. *170*, 212–219. https://doi.org/10.1001/jamapediatrics.2015.3732.

11. Nygaard, U.C., Li, Z., Palys, T., Jackson, B., Subbiah, M., Malipatlolla, M., Sampath, V., Maecker, H., Karagas, M.R., and Nadeau, K.C. (2017). Cord blood T cell subpopulations and associations with maternal cadmium and arsenic exposures. PLoS One *12*, e0179606. https://doi.org/10.1371/journal.pone.0179606.

12. Molloy, E.J., O'Neill, A.J., Grantham, J.J., Sheridan-Pereira, M., Fitzpatrick, J.M., Webb, D.W., and Watson, R.W.G. (2004). Labor promotes neonatal neutrophil survival and lipopolysaccharide responsiveness. Pediatr. Res. *56*, 99–103. https://doi.org/10.1203/01.PDR.0000130473.30874.B6.

13. Coker, M., and Madan, J.C. (2020). The microbiome and immune system development. In The Developing Microbiome (Academic Press), pp. 43–66.

14. Gensollen, T., Iyer, S.S., Kasper, D.L., and Blumberg, R.S. (2016). How colonization by microbiota in early life shapes the immune system. Science *352*, 539–544. https://doi.org/10.1126/science.aad9378.

15. Kamada, N., Chen, G.Y., Inohara, N., and Núñez, G. (2013). Control of pathogens and pathobionts by the gut microbiota. Nat. Immunol. *14*, 685–690. https://doi.org/10.1038/ni.2608.

16. Belkaid, Y., and Hand, T.W. (2014). Role of the microbiota in immunity and inflammation. Cell *157*, 121–141. https://doi.org/10.1016/j.cell.2014.03.011.

17. Bunker, J.J., Flynn, T.M., Koval, J.C., Shaw, D.G., Meisel, M., McDonald, B.D., Ishizuka, I.E., Dent, A.L., Wilson, P.C., Jabri, B., et al. (2015). Innate and adaptive humoral responses coat distinct commensal bacteria with immunoglobulin A. Immunity *43*, 541–553. https://doi.org/10.1016/j.immuni.2015.08.007.

18. Suzuki, K., Meek, B., Doi, Y., Muramatsu, M., Chiba, T., Honjo, T., and Fagarasan, S. (2004). Aberrant expansion of segmented filamentous bacteria in IgA-deficient gut. Proc. Natl. Acad. Sci. USA *101*, 1981–1986. https://doi.org/10.1073/pnas.0307317101.

19. Tanaka, M., and Nakayama, J. (2017). Development of the gut microbiota in infancy and its impact on health in later life. Allergol. Int. *66*, 515–522. https://doi.org/10.1016/j.alit.2017.07.010.

20. Henrick, B.M., Rodriguez, L., Lakshmikanth, T., Pou, C., Henckel, E., Arzoomand, A., Olin, A., Wang, J., Mikes, J., Tan, Z., et al. (2021). Bifidobacteria-mediated immune system imprinting early in life. Cell *184*, 3884–3898.e11. https://doi.org/10.1016/j.cell.2021.05.030.

21. Werlang, I.C.R., Mueller, N.T., Pizoni, A., Wisintainer, H., Matte, U., Costa, S.H.d.A.M., Ramos, J.G.L., Goldani, M.Z., Dominguez-Bello, M.G., and Goldani, H.A.S. (2018). Associations of birth mode with cord blood cytokines, white blood cells, and newborn intestinal bifidobacteria. PLoS One *13*, e0205962. https://doi.org/10.1371/journal.pone.0205962.

22. Rudin, A., and Lundell, A.-C. (2012). Infant B cell memory and gut bacterial colonization. Gut Microb. *3*, 474–475. https://doi.org/10.4161/gmic.21419.

23. Salas, L.A., Koestler, D.C., Butler, R.A., Hansen, H.M., Wiencke, J.K., Kelsey, K.T., and Christensen, B.C. (2018). An optimized library for reference-based deconvolution of whole-blood biospecimens assayed using the Illumina HumanMethylationEPIC BeadArray. Genome Biol. *19*, 64. https://doi.org/10.1186/s13059-018-1448-7.

24. Salas, L.A., Baker, E.R., Nieuwenhuijsen, M.J., Marsit, C.J., Christensen, B.C., and Karagas, M.R. (2019). Maternal swimming pool exposure during pregnancy in relation to birth outcomes and cord blood DNA methylation among private well users. Environ. Int. *123*, 459–466. https://doi.org/10.1016/j.envint.2018.12.017.

25. Gervin, K., Salas, L.A., Bakulski, K.M., van Zelm, M.C., Koestler, D.C., Wiencke, J.K., Duijts, L., Moll, H.A., Kelsey, K.T., Kobor, M.S., et al. (2019). Systematic evaluation and validation of reference and library selection methods for deconvolution of cord blood DNA methylation data. Clin. Epigenetics *11*, 125. https://doi.org/10.1186/s13148-019-0717-y.

26. Staats, J., Divekar, A., McCoy, J.P., and Maecker, H.T. (2019). Guidelines for gating flow cytometry data for immunological assays. Methods Mol. Biol. *2032*, 81–104. https://doi.org/10.1007/978-1-4939-9650-6_5.

27. Salas, L.A., Zhang, Z., Koestler, D.C., Butler, R.A., Hansen, H.M., Molinaro, A.M., Wiencke, J.K., Kelsey, K.T., and Christensen, B.C. (2022). Enhanced cell deconvolution of peripheral blood using DNA methylation for high-resolution immune profiling. Nat. Commun. *13*, 761. https://doi.org/10.1038/s41467-021-27864-7.

28. Aryee, M.J., Jaffe, A.E., Corrada-Bravo, H., Ladd-Acosta, C., Feinberg, A.P., Hansen, K.D., and Irizarry, R.A. (2014). Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics *30*, 1363–1369. https://doi.org/10.1093/bioinformatics/btu049.

29. Pruesse, E., Quast, C., Knittel, K., Fuchs, B.M., Ludwig, W., Peplies, J., and Glöckner, F.O. (2007). SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res. *35*, 7188–7196. https://doi.org/10.1093/nar/gkm864.

30. Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., and Holmes, S.P. (2016). DADA2: high-resolution sample inference from Illumina amplicon data. Nat. Methods *13*, 581–583. https://doi.org/10.1038/nmeth.3869.

31. Gilbert-Diamond, D., Cottingham, K.L., Gruber, J.F., Punshon, T., Sayarath, V., Gandolfi, A.J., Baker, E.R., Jackson, B.P., Folt, C.L., and Karagas, M.R. (2011). Rice consumption contributes to arsenic exposure in US women. Proc. Natl. Acad. Sci. USA *108*, 20656–20660. https://doi.org/10.1073/pnas.1109127108.

32. Koestler, D.C., Jones, M.J., Usset, J., Christensen, B.C., Butler, R.A., Kobor, M.S., Wiencke, J.K., and Kelsey, K.T. (2016). Improving cell mixture deconvolution by identifying optimal DNA methylation libraries (IDOL). BMC Bioinf. *17*, 120. https://doi.org/10.1186/s12859-016-0943-7.

33. Houseman, E.A., Accomando, W.P., Koestler, D.C., Christensen, B.C., Marsit, C.J., Nelson, H.H., Wiencke, J.K., and Kelsey, K.T. (2012). DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinf. *13*, 86. https://doi.org/10.1186/1471-2105-13-86.

34. Lundgren, S.N., Madan, J.C., Emond, J.A., Morrison, H.G., Christensen, B.C., Karagas, M.R., and Hoen, A.G. (2018). Maternal diet during pregnancy is related with the infant stool microbiome in a delivery mode-dependent manner. Microbiome *6*, 109. https://doi.org/10.1186/s40168-018-0490-8.

35. Halekoh, U., Højsgaard, S., and Yan, J. (2006). The R package geepack for generalized estimating equations. J. Stat. Soft. *15*, 1–11. https://doi.org/10.18637/jss.v015.i02.

36. Liang, K.-Y., and Zeger, S.L. (1986). Longitudinal Data Analysis Using Generalized Linear Models. Biometrika *73* (1), 13–22. https://doi.org/10.1093/biomet/73.1.13.

37. Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: apractical and powerful approach to multiple testing. J. Roy. Stat. Soc. B *57*, 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

38. Buuren, S.v., and Groothuis-Oudshoorn, K. (2011). Mice: multivariate imputation by chained equations in R. J. Stat. Soft. *45*, 1–67. https://doi.org/10.18637/jss.v045.i03.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Infant stool samples	This manuscript	PRJNA296814 https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA296814
Cord blood samples	This manuscript	N/A
Critical commercial assays		
DNeasy Blood and Tissue Kits	Qiagen	Cat #: 69504
EZ DNA Methylation Kit	Zymo Research	Cat #: D5001
Infinium HumanMethylation450k	Illumina	Cat #: WG-314-1003
Infinium HumanMethylationEPIC	Illumina	Cat #: WG-317-1001
Quick-DNA Fecal/Soil Microbe Miniprep Kit	Zymo Research	Cat #: D6010
Software and algorithms		
minfi	Aryee et al.[28]	https://doi.org/10.1093/bioinformatics/btu049
FlowSorted.CordBloodCombined.450k library	Gervin et al.[25]	https://doi.org/10.1186/s13148-019-0717-y
FlowSorted. Blood.EPIC library	Salas et al.[23]	https://doi.org/10.1186/s13059-018-1448-7
SILVA database	Pruesse et al.[29]	https://doi.org/10.1093/nar/gkm864
DADA2	Callahan et al.[30]	https://doi.org/10.1038/nmeth.3869
R Version 4.1.2	R	https://www.r-project.org
RStudio Version 2021.09.2	RStudio	http://www.rstudio.com

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Margaret R. Karagas Margaret.R.Karagas@dartmouth.edu.

Materials availability

This study did not generate new unique reagents.

Data and code availability

- The microbiome data used in this study can be found at Sequence Read Archive: http://www.ncbi.nlm.nih.gov/sra under accession number PRJNA296814.
- Code required to reanalyze the data reported in this paper is available on Github: https://github.com/yukamoro/GutMicrobiome_CordBloodImmuneCell/tree/jiezhou
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Study design

This study includes mother-infant dyads from the prospective NHBCS. The NHBCS recruited pregnant women aged 18 to 45 who received prenatal care at study clinics in New Hampshire, USA starting in January 2009.[10,31] Women living in the same household since their last menstrual period, not planning to move, living in a residence served by a private water system, and with a singleton pregnancy were eligible to participate. Participants completed surveys, including questions on sociodemographic factors and lifestyles such as smoking history. Pre-pregnancy body mass index (BMI) and infant birth characteristics

were ascertained from a review of the medical delivery records. All protocols were approved by the Committee for the Protection of Human Subjects at Dartmouth College. Participants provided written informed consent upon enrollment.

METHOD DETAILS

Cord blood collection

We collected cord blood samples of participants at the time of delivery. DNA was extracted from samples using DNeasy® blood and tissue kits (Qiagen, Valencia, CA) and was bisulfite converted using the EZ DNA Methylation kits (Zymo, Irvine, CA).[24] Samples were randomized across several plates and subsequently subjected to epigenome-wide DNA methylation assessment. Infinium HumanMethylation450k (105 samples) and Infinium HumanMethylationEPIC (237 samples) (Illumina, San Diego, CA). Following standard protocols, microarrays were processed at the Biomedical Genomics Center at the University of Minnesota (Minneapolis, MN). Beta values were extracted using minfi[28] based on the ratios of the fluorescent signals (red and green) from the idat files. Umbilical cord blood cell estimates were obtained using the FlowSorted.CordBloodCombined.450k[25] library using the library derived from the IDOL procedure.[32] This method uses a highly curated reference with high purity for seven cell types (granulocytes, CD8(+) T cells, CD4(+) T cells, B cells, NK, monocytes, and nucleated red blood cells). The reference includes children from different genetic ancestries. Using the IDOL procedure, we can estimate the cell composition for these samples accurately, minimizing potential biases during the calculation. Using the Houseman constrained projection/quadratic programming, the estimation was performed through the FlowSorted. Blood.EPIC library in Bioconductor.[23,33] After the estimation, the proportions were converted to percentages with only one decimal point to reduce floating-point approximation errors.

Infant stool collection and profiling

We collected infant stool samples at approximately 6 weeks and 12 months of age.[10,34] These samples were aliquoted and frozen at −80°C, and DNA was extracted from thawed samples using Zymo DNA extraction kit (Zymo Research, Irvine, CA). OD260/280 nanodrop was used to measure sample quality and purity. Samples were sent to Marine Biological Laboratory in Woods Hole, Massachusetts, for bacterial 16S rRNA gene sequencing of the V4-V5 hypervariable region using Illumina MiSeq (Illumina, San Diego, CA). We amplified samples in triplicate with one negative control for quality control.[10] We inferred amplicon sequence variants (ASVs) using DADA2[30] and assigned taxonomies using the SILVA database.[29]

QUANTIFICATIONS AND STATISTICAL ANALYSIS

Statistical analysis

We applied GEE[35,36] to estimate the parameters in the proposed model, which can yield consistent estimates of the parameters and their variance estimates. Using this approach, let w_{it} be the abundance of a given ASV for subject i at time point t, m_{it} the total abundance of all ASVs for subject i at time point t. Binomial distribution is then employed to model the relationship between w_{it} and m_{it},

$$w_{it} \sim Binomial(m_{it}, z) \quad \text{(Equation 1)}$$

$$\log\left(\frac{z}{1-z}\right) = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_7 x_7 \quad \text{(Equation 2)}$$

where z represents the probability of a random-sampled microbe falling into the given ASV. For probability z, we assume it relates to a given immune cell through the Equation 2, where x_1 represents the concentration of a given immune cell, x_2 to x_7 are the covariates, which include respectively delivery mode, infant sex, infant birth weight, gestational age, maternal enrollment age, maternal smoking during pregnancy, and alpha diversity of the infant's gut. For each ASV, the GEE procedure was carried out using the R package *geepack*.[35] Then we control the FDR[37] to select significant ASVs under the criterion FDR <0.35. Missing data for confounding variables were assumed to be missing completely at random at imputed using multiple imputations by chained equations and the predictive mean matching method with the R package *mice*.[38]