Title
Viable bacterial colonization is highly limited in the human intestine in utero

Author
Rackaityte, Elze

Publication Date
2020

Peer reviewed|Thesis/dissertation

Viable bacterial colonization is highly limited in the human intestine in utero

by
Elze Rackaityte

DISSERTATION
Submitted in partial satisfaction of the requirements for degree of
DOCTOR OF PHILOSOPHY

in

Biomedical Sciences

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:
Adrian Erlebacher
F0E47A73FD8E419...

Adrian Erlebacher
Chair

DocuSigned by:
Susan Lynch
DocuSigned by:
Ari Molofsky
DocuSigned by:
Susan Fisher
3D7FD5EAA89E411...

Susan Lynch
Ari Molofsky
Susan Fisher

Committee Members

ACKNOWLEDGEMENTS

This thesis is the collaborative work of many people who were and continue to be the best part of the story. My excitement for science was stoked and refined by my great mentors and friends and sustains me to this day. I am so lucky to have encountered so many brilliant minds and generous hearts throughout my years at UCSF. These people have shaped me as a scientist and as a human being. Over these years, my identity has become deeply intertwined with the word scientist and I feel so lucky to follow the path of so many incredible people. I write their names in the annals of this library, though they are already filed in the halls of my heart.

I am so grateful to my thesis advisor Dr. Susan Lynch. There is no better feeling than to know you were there to support me, defend me, and work with me through the challenges that science brings forth. Throughout my PhD, your cared for my well-being and my training experience. I will miss the 4 am texts, popping into your office to chat about a crazy idea, talking on the phone, and travelling together. You have created a community of strong, bright, and hard-working scientists who are eager to help each other reach great heights, a true testament to your brilliance. You bring people up and are my absolute scientific role model. Mostly because you are simply an incredible scientist. You have keen sense of the future of science and a pulse for great ideas. There have been many times that I have approached you with a perceived scientific dead-end only to walk away with hundreds of new ideas. You have given me the intellectual space to be creative and gentle guidance toward better, more elegant studies. Debating scientific theories with you has been one of my favorite aspects of my

PhD: you pushed me to be more open-minded and rigorous. I greatly admire this quality in you and have grown so much scientifically in my six years as a graduate student in your lab. Sue, I am immeasurably grateful for your contributions to my education. I hope that ours will be a lifelong collaboration.

I thank the community of the Lynch lab past and present, especially the other graduate students and those who contributed to my thesis project. I'd like to acknowledge the post-docs, graduate students, and staff that taught me the ropes: Nik Kymes, Ricardo Valladares, Kaitlyn Lucey, Emily Cope, Din Lin, Dat Nguyen, Katie McCauley, Meera Shenoy, Jordan Mar, and Julia Durack. Kei Fujimura, my bay-mate, thank you for your no-nonsense attitude and for teaching me about the finer things. Elle Fukui, thank you for bearing with me as I threw crazy ideas at you. You are brilliant and I cannot wait to see the heights you climb! I'd like to thank the Lynch lab graduate students: Ariane Panzer, Germaine Yong, Dr. Sophie Levan. Our elaborate graduate student dinners, coffee chats, and constant distractions were such a highlight in my studies. Ariane, thank you for always bringing me up, for critiquing my work, and for being a constant sounding board. Thank you also for showing me how to scruff mice and for not giving up on me. I will miss our coffee chats and naps on the conference table. Germaine, thank you for your practical advice and for shooting down my crazy ideas, thanks also for telling me I am messy and for letting me use the plate reader a lot. Too bad I never learned how to keep Caco2s alive from you. Sophie, you have given me the best career advice and inspired me so much in your scientific ideas. I can't quite keep up with you on hikes or in the intellectual realm, but it's damn fun to try!

I am very grateful to my second home, wherever Joanna Halkias was pipetting. She taught me everything I know in immunology. From our first intestinal dissections, to constantly receiving papers to read from you, to arguing, to hugging it out, to dreaming and scheming: thank you. These years have been incredible because of you. I have learned what true grit, passion, and brilliance look like. Thank you for teaching me the language of CDs, for not giving up on me, for re-gating my flow, for constantly being positive, and for simply being a great human in my life. So many hard times were made easier because of you.

I am also very grateful to my early mentor Trevor D. Burt, who was crucial to the initiation of this project. In the Burt lab, I met some of the best scientists who significantly contributed to my training. Thank you to Norm Jones, for teaching me flow and how to compensate. Thank you to Ventura Mendoza, for digesting all those guts and for telling me to stop singing out loud and for calling EH&S when I had a blood exposure. Thank you to Dan Bunis for all your positivity and support, you are such a light and I have appreciated your compassionate lens to science so much. Thanks finally to Melissa Ng, the T cell to my dendritic cell. I simply loved doing this with you. Thank you for thinking with me, thank you for sitting next to me late into the night, thank you for teaching me so much. Discussing ideas with you has been one of the highlights of my training and I am so excited to keep doing it forever! I already miss you so much, but I cannot wait to visit and learn what you are doing with neutrophils.

I thank my thesis committee Susan Fisher, Ari Molofsky, and my thesis chair, Adrian Erlebacher. Thank you for pushing me to think rigorously and for supporting me

throughout my studies. Thanks also to my BMS program, especially Demian Sainz, who have supported me in times of administrative confusion. Thanks to my classmates, Jorge Ortiz-Carpena, Eric Wigton, Luke Smith, Adair Borges and Natanya Kerper. Staying up late with you and talking science in the outer sunset will always hold a special place in my heart.

Thanks to my friends Sophie Kaineg, Tessa Sanchez, and Casey Fitzpatrick for humoring my constant complaints. Thank you also for sharing so many special moments with me, you know that I love to be in the mix.

Thank you to Louis for bringing me dinners to the sorter, you have earned your MRS degree, with honors. I love you.

Thanks finally to my parents, for bringing me to this country so that I could attend a school like UCSF. Mama and Tėti, thank you for the freedom to pursue my ideas. And thank you for gifting me with science as a beautiful lens for viewing the world.

CONTRIBUTIONS TO THE PRESENTED WORK

The work presented in this dissertation was performed under the direct supervision and guidance of Dr. Susan V. Lynch, PhD. Additional guidance and insight was provided by Dr. Joanna Halkias, MD and thesis committee members Dr. Susan Fisher, PhD, Dr. Ari Molofsky, MD PhD, Dr. Adrian Erlebacher, MD PhD, and Dr. Trevor D. Burt, MD.

"*But seeds are invisible. They sleep deep in the heart of the Earth's darkness, until one among them is seized with the desire to awaken. Then they stretch and begin to sprout, quite timidly at first, into a charming, harmless little twig reaching toward the sun.*"

Antoine de Saint-Exupéry, *The Little Prince*

Viable bacterial colonization is highly limited in the human intestine *in utero*

Elze Rackaityte

ABSTRACT

The intestinal immune system must tolerate commensal bacteria which perform essential functions for the survival of the host. The development of intestinal immunity and the progressive colonization with bacteria, fungi, and protists (collectively known as the microbiota) occurs in concert to select for these valuable functions. Intestinal adaptive immunity develops in the human fetal intestine by 11-14 weeks gestation, yet whether viable microbes exist *in utero* and interact with the intestinal immune system is unknown. Bacterial-like morphology was identified in pockets of human fetal meconium at mid-gestation by scanning electron microscopy (n=4) and a sparse bacterial signal was detected by 16S rRNA sequencing (n=40 of 50) compared to environmental controls (n=87). Eighteen taxa were enriched in fetal meconium with *Micrococcaceae* (n=9) and *Lactobacillus* (n=6) the most abundant. Fetal intestines dominated by *Micrococcaceae* exhibited distinct patterns of T cell composition and epithelial transcription. Fetal *Micrococcus luteus*, isolated only in the presence of monocytes, grew on placental hormones, remained viable within antigen presenting cells, limited inflammation *ex vivo*, and possessed genomic features linked with survival in the fetus. Thus, viable bacteria are highly limited in the fetal intestine at mid-gestation, though strains with immunomodulatory capacity are detected in subsets of specimens. Future studies will investigate whether fetal bacteria could be utilized to rationally seed microbiomes and repurposed for drug delivery to the fetus.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

Material for this chapter was modified from:

Rackaityte E & Halkias J (2020). Mechanisms of Fetal Tolerance and Immune Regulation. ***Frontiers in Immunology*** *11*.

1.1 Mechanisms of Fetal T cell Tolerance and Immune Regulation

A healthy human pregnancy, in which the fetus shares only half of its genes with the mother, is an impressive immunological feat. Non-inherited maternal antigens and a growing repertoire of self-antigens present a unique immune challenge to survival *in utero*. Suppression of responses to these antigens is critical to the maintenance of pregnancy [1–3], and the semi-allogeneic fetus relies on a specialized program of immune tolerance for survival *in utero*. Thus, humans have evolved redundant and dominant fetal mechanisms of tolerance that override our immune system's encoded ability to mount a protective response.

Human fetal development occurs within the anatomically distinct *in utero* environment defined primarily by the placenta and fetal membranes. Maternal immune adaptation to the semi-allogeneic pregnancy includes limitations on immune cell entry, activation, and function [4] as well as the appearance of uniquely tolerogenic cellular and molecular mechanisms (reviewed in Erlebacher, 2013). Features of pregnancy-induced immune tolerance are driven in part by the endocrine functions of the placenta as well as the state of physiologic hypoxia derived from the vascular anatomy of this organ. Finally, the placenta and fetal membranes create a protected niche which filters and limits fetal exposure to external antigens and microbes. Our understanding of placental biology has evolved from a barrier organ to one of feto-maternal communication (reviewed in PrabhuDas et al., 2015) and there is a growing appreciation for the role of the fetal immune system in the maintenance of a healthy pregnancy.

Murine models have contributed significantly to our understanding of maternal immune responses in pregnancy, however fetal immunity is poorly modeled in the mouse. Although thymus organogenesis is remarkably similar between the species, the functional output differs drastically during development, likely influenced by the relatively short murine gestation in comparison to that of humans. The first wave of murine T cells to exit the thymus are TCR$\gamma\delta$ thymocytes destined for the skin around embryonic day 15 [7,8]. These cells are subsequently replaced by increasing thymopoeisis of conventional TCR$\alpha\beta$ T cells which continue to populate the periphery until the end of the first week of life [9]. In humans, TCR$\gamma\delta$ and TCR$\alpha\beta$ T cells, including regulatory T cells, exit the fetal thymus simultaneously and comparatively earlier than in mice (around 12-14 weeks of gestation; 10–12). Therefore, unlike mice, most T cell development in humans occurs *in utero.* Mice depend on a sustained thymic output of naïve T cells throughout their lifetime [13] and neonatal thymectomy results in severe impairment of immune responses to infection and autoimmunity [14–16]. In contrast, humans primarily rely on expansion of existing naïve T cells post-natally, as incidental neonatal thymectomies during cardiac surgery do not give rise to autoimmune disease or an increased susceptibility to infection [13,17,18]. Given these differences between mice and humans, this chapter will primarily focus on human adaptive immune development.

In humans, all cellular components of innate and adaptive immunity are present in the developing fetus. Adaptive immunity results from antigen-specific activation of T cells and B cells followed by the generation of long-lived memory cells capable of a robust recall response. Innate immunity, triggered by molecular pattern molecules,

provides rapid protection from pathogens and clears dying or damaged self-cells. Innate immune cells are also keystone initiators of the adaptive arm of immunity. T cells are classically activated by cognate peptide-MHC interaction with the T cell receptor (TCR), a process directed by professional antigen presenting cells (APCs) in the periphery. Fetal APCs, such as dendritic cells, can be activated in response to pathogen associated molecular pattern molecules (PAMPs), migrate between tissues and lymphatics, and robustly activate naïve T cells [19]. These data indicate that fetal innate immune cells possess the capacity to direct T cell activation and differentiation *in utero*.

Fetal immunity serves specialized evolutionary goals adapted to the unique demands of gestation and thus differs significantly from that of the adult. Immune tolerance, while critical to survival *in utero,* must transition to a program of protective immunity in preparation for birth. The neonatal window bridges these two programs in remarkable concert, allowing for the establishment of tolerance to commensals, while also protecting the newborn from infection. In mice, this critical window of immune development is brief: tolerance to skin commensals for example, is preferentially established during the first week of life, coinciding with an abrupt influx of regulatory T cells (T_{reg}; 20). In humans, this developmental transition may be longer and likely begins *in utero*, as maternal environmental exposures (e.g. to farming) result in a greater abundance of highly suppressive T_{reg} cells in neonatal umbilical cord blood [21], and childhood exposures protect from allergic disease in adult life [22–24]. However, newborn protective immunity is not fully effective because neonates and infants exhibit a higher susceptibility to infection, the leading causes of mortality in these age groups [25].

This presumed failure of the infant immune system has been bolstered by studies of umbilical cord blood, in which the majority of T cells are naïve and fail to mount a pro-inflammatory response [26], and APCs exhibit an impaired ability to activate T cells [27]. Umbilical cord blood, collected immediately after delivery, contains circulating fetal immune cells and is stereotypically distinct from the development of circulating neonatal immune cells in the first months of life [28]. In contrast to immune cells in the blood, the infant intestine possesses a sizeable proportion of memory T cells [29–32] and lymphoid-derived fetal dendritic cells are capable of T cell activation [19]. These recent reports of functional memory T cells capable of pro-inflammatory cytokine production in the fetal intestine indicate that the fetal immune system is not in essence immature, but is rather compartmentalized in its specialized function. Many of the unique features of the fetal immune system are lost with advancing age, for example the proclivity of naïve T cells to differentiate into T_{reg} cells [33], the presence of innate-like CD4 and CD8 T cells [29,34,35], and the thymus-derived capacity for T cell IL-8 production [36,37]. The aim of this chapter is to summarize the latest findings on the composition and function of fetal T cell immunity with a specific focus on mechanisms of tolerance and immune regulation (**Figure 1.1**). We discuss how the intra-uterine environment influences fetal immunity through unique adaptations and also examine experimental models to investigate these paradigms. Lastly, we consider the consequences of dysregulation of fetal immunity and how current findings could be leveraged to protect the fetus from perinatal inflammatory pathologies, such as preterm birth.

1.1.1 Fetal T cell immunity

Thymic development begins by week eight of human gestation, and the first T cells begin to populate the periphery by 12-14 weeks of gestation [10,38,39]. Unlike mice, both $\gamma\delta$ and $\alpha\beta$ T cells emigrate from the thymus simultaneously [7,8,38] and the appearance of human T_{reg} cells coincides with that of naïve T cells [11,12,16]. Fetal T cell colonization in the periphery occurs in a state of relative lymphopenia in which naïve cells composed primarily of recent thymic emigrants begin to populate lymphoid and mucosal niches. Naïve T cells undergo rapid proliferation in response to homeostatic signals [40] similar to that seen in postnatal mice [41]. While the vast majority of T cells in cord blood possess a naïve phenotype, healthy term cord blood contains memory T cells with inflammatory effector functions similar to those observed in adult, albeit in very low proportion [42]. Fetal adaptive immune memory was first reported in the fetal intestine [43–45], and memory T cells predominate in the infant and pediatric intestine [46], suggesting that early life adaptive memory is particularly abundant in mucosal tissues.

1.1.1.1 *Regulatory T cells*

Immunotolerance is essential to the maintenance of pregnancy and the T_{reg} cell is the lynchpin of adaptive tolerance due to its unique ability to suppress the activation, proliferation, and effector functions of a wide range of immune cells. In the fetus, T_{reg} cells (defined in humans by expression of FoxP3, CD25, and low or absent expression of CD127 [47,48]) are strikingly abundant in peripheral lymphoid organs during the second trimester of human gestation, in stark contrast to neonatal and adult lymph nodes and adult peripheral blood cells [2,12,49,50]. Although thymic output of T_{reg} cells is similar *in utero* and after birth [33], fetal naïve T cells display an increased propensity to differentiate

into T_{reg} cells upon antigen encounter in the periphery (induced T_{reg}; iT_{reg}; [33]. Levels of TGFβ are higher in fetal than in adult lymph nodes, which potentiates the generation of iT_{reg} cells, and distinct fetal hematopoietic stem cells give rise to fetal T cells with the unique ability to differentiate into T_{reg} cells [51]. A recent study demonstrated that fetal naïve T cells are already poised for tolerance and share a partial transcriptome and epigenome similar to that of T_{reg} cells, such as the heightened expression of the transcription factor *Helios* required for fetal iT_{reg} function [52]. These studies indicate that both cell-intrinsic and cell-extrinsic mechanisms promote the generation of a predominantly tolerant peripheral T cell response, essential for the maintenance of pregnancy.

Mold and colleagues demonstrated that fetal T cells rapidly proliferate to maternal and self-antigens in the absence of T_{reg} cells, indicating that fetal immune tolerance is an active process [33]. The contribution of T_{reg} cells to perinatal immune tolerance is highlighted by IPEX (immune dysregulation, polyendocrinopathy, entheropathy, X-linked) syndrome, resulting from the loss of the T_{reg} lineage-defining transcription factor FOXP3. Onset of IPEX-related autoimmunity occurs in the second trimester, coinciding with the emergence of peripheral T cells and underscoring the importance of T_{reg}- mediated tolerance to the survival of the fetus [52,53]. That fetal T cells are actively suppressed by T_{reg} cells points to a broader capacity of fetal responses that are kept under tight control.

1.1.1.2 *T helper cell type I (Th1) cells*

The need for active tolerance to self- and maternal antigens suggests that protective Th1 responses may be detrimental to a healthy pregnancy. Indeed, human

cord blood from infants born preterm exhibits an enrichment of Th1 cells [54–56], thus implicating Th1 cells in the pathophysiology of the premature termination of pregnancy [57]. However, a fetal immune response with Th1 polarization is also generated in response to maternal infections [58–61], indicating that protective immunity can be elicited under specific *in utero* conditions and is not always associated with a negative outcome.

Contrary to the predominantly naïve phenotype of T cells in cord blood, ~50% of CD4 T cells in the human fetal intestine exhibit a memory phenotype and robustly produce the Th1 cytokines IFNγ and TNFα [30,31,62]. I recently contributed to work reporting that the majority of fetal effector memory T cells in the intestine express the transcription factor PLZF, are transcriptionally distinct from either conventional memory T cells or innate-like T cells, and are absent from the adult intestine [62]. A greater proportion of PLZF$^+$ CD4$^+$ T cells produced Th1 cytokines as compared to conventional CD4 memory T cells, and this was most evident in the small intestine. Further, IFN-γ–producing PLZF$^+$ CD4$^+$ T cells were enriched in the cord blood of infants born preterm as well as in infants with gastroschisis, a congenital defect defined by chronic inflammation originating from the intestine. These data suggest that dysregulation of PLZF$^+$ CD4$^+$ T cells may contribute to pathologic systemic fetal immune activation.

While PLZF expression is sufficient to confer a memory phenotype and effector function in murine T cells [63,64], its contribution to human T cell lineage commitment is not as clearly defined. Unlike classic innate-like T cells, functional maturation of fetal PLZF$^+$ CD4$^+$ T cells is not evident in the thymus and PLZF$^+$ CD4$^+$ T cells display a

polyclonal TCR repertoire (28). Further, the capacity for activation in response to both TCR-dependent and cytokine-dependent signaling is also present in adult intestinal memory T cells [65]. Thus, fetal PLZF$^+$ CD4$^+$ T cells share functional and transcriptional attributes with both conventional and innate-like T cells and may serve as a link between innate and adaptive immunity in the fetus.

A number of innate-like T cells capable of Th1 cytokine production are present in the human fetus, including $\gamma\delta$ T cells, invariant natural killer T (iNKT) cells and mucosa-associated invariant T (MAIT) cells. These cells share certain key features: a semi-invariant TCR repertoire, non-classical MHC restriction, and rapid production of inflammatory cytokines such as IFNγ and TNFα [66]. While iNKTs are enriched in mucosal tissues such as the intestine and the liver, they represent a small proportion of the total T cell compartment (<1%) in both fetal and adult tissues [67]. Low proportions of Vα7.2$^+$ CD161$^+$ T cells with many attributes and characteristics consistent with MAIT cells were reported in the human fetus (<1% of CD3 cells) and were enriched in the fetal liver, small intestine, and lung [68]. However, subsequent studies reported that a minority of fetal Vα7.2$^+$ CD161$^+$ T cells reacted with the riboflavin-bound MR1 tetramer [69]. These findings indicate that broader specificity may exist in semi-invariant innate-like subsets in the fetus. Fetal semi-invariant $\gamma\delta$ T cells are functionally pre-programmed to primarily produce IFNγ and have the capacity to generate Th1 responses to CMV [70]. However, recent reports suggest that the repertoire of Vγ9Vδ2 T cells is distinct to the fetus and differs from that of the adult [71]. In sum, a substantial number and diversity of conventional and innate-like fetal cells with the capacity for Th1 cytokine production are

present in fetal tissues, suggesting a protective role at barrier sites during the critical neonatal window of development.

1.1.1.3 *T helper cell type II (Th2) cell*

Historically, the perinatal immune system has been considered skewed toward Th2 responses and away from Th1 responses. However, much of the data supporting this notion was derived from murine studies. Immunization in neonatal mice favors Th2 memory pool formation [72,73], even when immunized with a Th1 skewing adjuvant [74], and murine thymocytes exhibit an epigenetic predisposition for Th2 differentiation [75]. Mirroring findings in neonatal mice, human naïve T cells derived from fetal cord blood and infant adenoids exhibited higher expression of GATA3 than that of adult T cells, and predominantly produced an unglycosylated isoform of IL-4 that is absent from adult Th2 cells [76]. Similarly, the IL-13 locus of cord blood T cells is characterized by open chromatin and permissive epigenetic marks [77], suggesting a predisposition for Th2 differentiation. Human neonatal vaccine responses elicit predominantly Th2 type immunity[78], similar to that seen in mice. However, Th1 and Th2 cytokine responses have been generated from cord blood naïve T cells *in vitro* [79] and neonatal vaccination with Bacillus Calmette-Guérin (BCG) produces an adult-like Th1 response [58,59]. Similarly, helminth and mycobacterial antigens induce an adult-like CD4 Th1 skewed response in cord blood of infants in regions endemic for schistosomiasis, filariasis, and tuberculosis [60]. Placental malaria results in increased IFNγ- and TNFα-producing CD4 effector memory T cells that proliferate to malaria-specific antigens upon re-stimulation [61], indicating that Th1 immunity is possible in response to the appropriate stimuli. Th2 cell function and specificity have not been extensively investigated in the fetus outside

of cord blood, though IL-4 producing cells are present in the fetal intestine [30,62]. In injury contexts, Th2 cells reduce inflammation and promote tissue regeneration (reviewed in Gieseck et al., 2018), which may be co-opted *in utero* to promote organogenesis. Thus, while some neonatal vaccine responses induce a Th2 bias, a Th1 response can be elicited under specific conditions, broadening our prevailing understanding of human fetal adaptive immunity.

1.1.1.4 *T helper cell type 17 (Th17) cells*

The role of Th17 cells has not been extensively studied in the context of fetal development. Term infant cord blood is uniquely enriched in a population of $CD161^+$ CD4 T cells which preferentially gives rise to Th17 cells under *in vitro* differentiation conditions [81]. Further, cord blood naïve CD4 T cells exhibit an enhanced capacity for Th17 differentiation compared to adult cells [82]. However, term cord blood CD4 T cells are not capable of either IL-17A or IL-17F production despite the presence of $CCR6^+$ effector memory cells with high levels of *RORC* expression [42]. In contrast to circulating fetal immune cells, fetal IL-17 production is restricted to $CD4^+$ T cells expressing the transcription factor PLZF in the mesenteric lymph node and to a lesser extent in the small intestine, supporting the spatial segregation of fetal effector function [62]. Although IL-17 can be produced by human immune cell subsets other than T cells [83], IL-17 production was not detected in fetal $\gamma\delta$TCR T cells [84] and the capacity of human fetal memory CD8 T cells to produce IL-17 has not been reported.

1.1.1.5 *CD8 T cells*

Cord blood naïve CD8 T cells are transcriptionally and epigenetically distinct from that of the adult with a proclivity for proliferation and innate immune responses. These

cells express neutrophil-associated and antimicrobial peptide genes, but exhibit a decreased cytotoxic response in comparison to adult cells [85]. Fetal CD8 T cells that express KIR/NKG2A and Eomes possess innate-like properties and produce IFNγ in response to IL-12 and IL-18 [34,35]. In parallel, the fetus is also capable of conventional CD8 T cell responses comparable to those of the adult in response to specific maternal infections. Fetal CD8 T cells mount an antiviral-specific response to hepatitis B [86], hepatitis C [87], HIV [88–90], and cytomegalovirus infection, and effectively differentiate into memory cells capable of perforin-production [91]. Placental malaria also induces antigen-specific memory CD8 T cell proliferation, though this does not result in increased inflammatory cytokine production [61]. Antigen-specific CD8 T cell responses in the fetus and neonate indicate that early-life immunization is possible against specific pathogens. The fetal cytokine environment may limit the activation of CD8 T cells, as TGFβ has been shown to suppress cytotoxic CD8 T cell differentiation induced by IL-15 [92]. Fetal CD8 T cells may be highly atypical, as they possess the capacity to produce IL-4 when stimulated in the absence of IL-2 and IL-12, a feat reminiscent of iNKT cells that is absent from adult CD8 T cells [93]. In mice, neonatal CD8 T cells originate from distinct hematopoietic stem cell sources than adult CD8 T cells and exhibit rapid and short-lived responses as compared to adults [94,95], exhibiting a layering of immune function similar to that of human fetal CD4 T cells [51]. Together, this evidence points to fetal-specific features for innate-like CD8 T cells which likely play a role in perinatal immune protection.

1.1.2 Fetal mechanisms of T cell regulation

The remarkable potential of fetal T cells to generate inflammatory responses suggests that strong and redundant mechanisms must exist to preserve a tolerogenic environment required for a successful pregnancy. Indeed, the fetal immune system has adapted and repurposed unique cell-intrinsic and cell-extrinsic means to limit inflammation (**Figure 1.1**).

1.1.2.1 *Cell-intrinsic mechanisms*

Fetal naïve T cells have a greater capacity to differentiate into T_{reg} cells, which may in part be explained by a higher sensitivity to TGFβ indicated by high levels of SMAD2/SMAD3 phosphorylation in the unstimulated state [49]. Lin28b, a repressor of let-7 microRNAs that target TGFβ signaling mediators, is highly expressed in fetal naïve T cells and is required for the increased propensity toward Treg cell differentiation [49]. Unlike adult naïve T cells, fetal naïve T cells have an increased propensity to differentiate into T_{reg} cells even in the absence of high levels of exogenous TGFβ [49], suggestive of additional cell intrinsic mechanisms such as a poised epigenome. The predisposition of fetal naïve cells to differentiate into T_{reg} cells is driven by higher expression of and chromatin accessibility at the Helios (IKZF2) locus, and ablation of Helios in fetal naïve T cells results in poor T_{reg} differentiation and impaired function [96]. These studies point to additional cell-intrinsic differences in fetal naïve T cells that contribute to a predominant program of immune tolerance and regulation in response to antigen encounter *in utero.*

T cell activation is governed by productive and simultaneous TCR- and co-stimulatory molecule signaling. Molecules that inhibit T cell activation (co-inhibitory) are

expressed either in conjunction with co-stimulatory molecules or following successful activation [97]. Fetal intestinal memory PLZF$^+$ T cells express high levels of the surface C-type lectin receptor CD161 [62]. In adults, engagement of CD161 inhibits human NK cells [98–100], but exerts co-stimulatory effects on iNKT cells [101], MAIT cells [102], and displays no consistent influence on polyclonal T cell activation [98,103]. It is thus striking that ligation of CD161 by two different monoclonal antibodies inhibited IFNγ production in response to TCR activation in fetal PLZF$^+$ T cells [29]. These reports point to a unique repurposing of the CD161 axis in the fetal context, the signaling mechanisms of which remain to be explored.

Despite strong inhibition following TCR-dependent stimulation, CD161 ligation has no effect on cytokine-mediated T cell activation, which remains a critical area of study. The differential enrichment of numerous negative T cell regulators (*DUSP4, DUSP5, DUSP6, LRIG1,* and *DTX1*) in the gene signature of fetal intestinal PLZF$^+$ CD4$^+$ memory T cells [62] suggests the existence of additional cell-intrinsic mechanisms of regulation to promote immune homeostasis *in utero*. In particular, the transcriptome of fetal PLZF$^+$ CD4$^+$ T cells was differentially enriched for multiple surface molecules with known inhibitory functions such as PD1 [62]. Thus, determining the extent to which known and novel co-inhibitory molecules can limit fetal T cell activity is a promising therapeutic avenue for birth-related pathologies, as many inhibitors to these receptor-ligand interactions are already in use and continue to be developed for cancer immunotherapy.

1.1.2.2 *Cell-extrinsic mechanisms*

T cell immunity is regulated in large part by the surrounding environment, such as the antigen presenting cells which directly engage with and efficiently activate T

cells. Given their central role in T cell activation and differentiation, it is thus not surprising that they also contribute to fetal immune regulation. For example, type II fetal dendritic cells (fDC2) limit the ability of T cells to produce IFNγ, IL-2, IL-17, TNFα, and other inflammatory cytokines [19]. Inhibition of TNFα production by fetal T cells is achieved through upregulation of arginase on fDC2s [19], however the mechanisms of inhibition of other inflammatory cytokines have yet to be determined. Production of IFNγ by fetal T cells may be limited by intestinal macrophages expressing high levels of LLT1, the natural ligand of CD161 [62]. Further, suppressive functions of intestinal macrophages are well described in adults [104–106] and it is likely that similar mechanisms are also employed by the fetal immune system.

Additional human myeloid cell populations contribute to T cell tolerance *in utero*. Myeloid-derived suppressor cells (MDSCs) are a heterogeneous subset of monocytic and granulocytic/neutrophilic cells that limit T cell and NK cell immunity in tumors and during the course of infection [107]. Neutrophilic MDSCs capable of suppressing Th1, Th2, and Th17 responses are enriched in cord blood and rapidly decrease to adult levels during infancy [108,109], with levels dropping drastically after the first month of life [110]. These MDSCs are capable of phagocytosis of *E. coli*, yet produce high levels of TGF and strongly suppress T cell proliferation [111], a process which requires cell-cell contact [110]. Further, BCL-2 mediates lower rates of apoptosis among cord blood MDSCs, suggesting that their prolonged survival may promote a tolerizing environment even after the resolution of bacterial infection [111]. It is plausible that C-type lectins may also be involved in regulating fetal MDSCs, because anti-inflammatory IL-10 production in

these cells is regulated by co-triggering TLR-MyD88- and C-type lectin receptor-Syk-dependent pathways [112]. Neutrophils migrate to tissues at steady-state and support organ function [113], yet their role in fetal development has not been investigated.

The lymphocyte compartment contributes to the modulation of fetal T cells through the activity of T_{reg} cells and regulatory B (B_{reg}) cells. Fetal tolerance is maintained in large part by active T_{reg} cell-mediated suppression of T cell activation [33], yet whether fetal T_{reg} cells exhibit classic or unique suppressive mechanisms has not been formally demonstrated. Further, while T_{reg} cell specificity for non-inherited maternal antigens has been proposed, the question of antigen-specific suppression requires further study. B_{reg} cells (reviewed in Mauri and Bosma, 2012) are enriched in the fetus and an immunosuppressive role for B_{reg} cells has been proposed in maternal tolerance [115,116]. In neonatal mice, IL-10-producing B_{reg} cells limit the ability of dendritic cells to produce inflammatory cytokines, which can be induced by broad TLR stimulation of dendritic cells and is paradoxically induced by type I interferon conditioning of B cells in acute inflammatory contexts [117,118]. In humans, B_{reg} cells are highly abundant in cord blood, produce IL-10, and are capable of suppressing CD4 T cell responses in an IL-10-dependent manner through CTLA4 and CD40 mechanisms [119]. Cord blood B_{reg} cells exhibit higher IgM expression as compared to adult blood and can inhibit both Th1 and Th2 responses [120], however IgM-expressing B cells are more prevalent in infant than in fetal intestines [121]. Thus, fetal lymphocytes contribute to the maintenance of tolerance in pregnancy through the generation of T_{reg} and B_{reg} cells.

Human and murine neonatal peripheral blood cells exhibit high proportions of transferrin receptor expressing ($CD71^+$) erythrocytes with the unique ability to suppress T cells [122]. Similar to fetal dendritic cells [19], $CD71^+$ erythrocytes mediate suppression through arginase-2, suggesting that arginine depletion is a key mechanism of immune regulation in the fetal environment [122]. Although the CD71 axis has not been explored in human fetal tissues, depletion of $CD71^+$ erythrocytes results in myeloid cell activation and increased production of TNFα in the intestine of neonatal mice. Because activation of intestinal myeloid cells in response to $CD71^+$ cell depletion does not occur in germ-free or antibiotic treated neonatal mice, this suggests that $CD71^+$ cells may contribute to tolerance to commensal colonization, a critical immune adaptation in early life [122]. Further, a subset of $CD71^+$ erythrocytes which express high levels TGFβ may contribute to the increased conversion of naïve T cells into T_{reg} cells. Ablation of T_{reg} cells results in higher proportion of $CD71^+$ erythrocytes and these erythrocytes compensate by increasing expression of TGFβ after transient $CD71^+$ cell depletion [123]. This interdependence of neonatal T_{reg} cells and $CD71^+$ erythrocytes points to repurposing of conserved axes that limit T cell immunity.

1.1.3 *In utero* environment contributes to the regulation of fetal T cells

The highly controlled biological environment *in utero* permits the development of the fetal immune system in a state of active immune suppression, while also promoting fetal-specific mechanisms of immune protection. For example, physiologic intra-uterine hypoxia contributes to fetal immune regulation by limiting the ability of myeloid cells to activate T cells (reviewed in Sitkovsky and Lukashev, 2005). Tissue hypoxia also

enhances T_{reg} numbers and function [125], pointing to multiple mechanisms by which fetal hypoxia contributes to immune tolerance *in utero.*

The developing human fetus is exposed to increasing levels of placental hormones with potent immunomodulatory effects throughout gestation. The placenta produces the steroid hormones estrogen [predominantly estradiol (E2) and estriol (E3)] and progesterone (P4), which are critical to maternal tolerance of the fetus [126–134]. Despite high levels of placental hormones in the fetal circulation, little is known about their effect on fetal T cells. P4 promotes differentiation of naïve cord blood T cells into T_{reg} cells and suppresses the differentiation of Th17, yet has little effect on adult T cells, suggesting that the sensitivity to P4 is lost in adulthood [128]. A similar immunosuppressive role has been described for E2 in mice [126,135] [136], yet the direct effect of E2 on human fetal T cells remains to be explored. In addition to fetal-specific, cell-intrinsic mechanisms of regulation, the development of human T cell immunity is strongly shaped and influenced by a regulated set of exposures within the *in utero* environment.

1.1.4 Consequences of fetal immune dysregulation

Despite dominant and redundant mechanisms that control T cell immunity *in utero*, fetal T cell activation can occur in pathologic contexts and contribute to the initiation of the fetal inflammatory response. Preterm birth (PTB), defined as birth before 37 weeks of gestation, affects 1 of every 9 newborns, and is the leading cause of childhood death under the age of 5 worldwide [137,138]. Preterm labor (PTL), the main cause of PTB, is more frequent when a fetal inflammatory response is elicited [139,140],

and most of the complications associated with death and disability in preterm infants are triggered by inflammation [56,141–144]. Despite the impact of infection and inflammation on the survival and long-term outcome of preterm infants, little is known about the mechanisms that drive fetal immune responses and their contribution to perinatal immune dysregulation.

Increased proportions of activated T cells and reduced suppressive T_{reg} cell activity are evident during fetal inflammation [54–56,145] and maternal-reactive fetal memory Th1 cells are implicated in the pathophysiology of PTB [1]. Activated CD4 and CD8 T cells derived from the cord blood of preterm infants induce uterine cell line contractility *in vitro*, an effect which can be rescued by anti-TNFα or anti-IFNγ antibodies [1]. Thus, T cell activation and specifically Th1 cytokines may drive the fetal inflammatory response associated with PTB. Animal models of inflammatory PTB identify the fetal intestine as a site of immune activation [146–149], indicating that mucosal T cells in contact with amniotic fluid are ideally situated to contribute to the fetal inflammatory response.

Pathology associated with activation of fetal intestinal T cells is exemplified in necrotizing enterocolitis (NEC), a co-morbidity of PTB. NEC infants exhibit higher proportions of effector memory CD4 T cells in the lamina propria of the intestine and these cells have a heightened capacity to produce TNFα and reduced capacity for IL-10 production [30]. While moderate TNFα levels promote intestinal epithelial maturation, dysregulation of intestinal T cells leads to tissue-damaging effects of TNFα [30]. A breakdown in cell-intrinsic mechanisms of human fetal tolerance has not been investigated in this context, but lack of sufficient T_{reg} populations may contribute to

disease development in NEC [150]. Murine models of NEC demonstrate that CD4 T cell activation is critical to the initiation of inflammation [151]. Transfer of NEC intestinal lymphocytes was sufficient to induce Th17 inflammation in otherwise naïve mice. Decreased T_{reg} cell levels were observed and could be rescued by retinoic acid supplementation [151], suggesting that loss of tolerance is a key tenet in this pathology. Ample evidence supports a role for dysregulation of T cell immunity in perinatal pathologies, therefore understanding the mechanisms that keep T cells in check may uncover novel therapeutics for these diseases.

1.1.5 Conclusion

The developing human immune system is not forged in an antigenic vacuum, but rather educated in the presence of self-, maternal-, and environmental antigens. It is therefore not surprising that protective adaptive immunity critical to perinatal surveillance and defense is generated concurrently with a program of immune tolerance *in utero.* The ability of fetal T cells to mount a protective and, in pathogenic contexts, dangerous inflammatory response reinforces the existence of fetal-specific adaptations of immune regulation. Cell intrinsic mechanisms of control are further bolstered by tolerizing pregnancy factors such as the hormonal milieu, hypoxia, and relative depravation of environmental antigens. Insight into these tolerizing mechanisms and their broader applicability to limit human immunity will likely have a major impact on human health.

Early-life immune exposures and responses strongly shape lifelong health, thus increasing the urgency to understand the development of human fetal immunity. There

is a growing need to develop methods to study human samples and for innovative model systems to explore the mechanisms of disease. Discoveries in human fetal immunology for the past decades have primarily come from studies in human cord blood, and will continue to serve as an important investigative tool for biomarker and therapeutic development. The study of perinatal inflammatory pathologies relies primarily on large animal models (e.g. sheep, macaques) that possess a full adaptive immune response similar to that of the developing human to explore mechanisms of injury. While cord blood will continue to inform predictive models for disease development, a global understanding of fetal immune development and its progression into childhood and beyond is still lacking. The study of human fetal tissue has served a critical, non-redundant role in our understanding of human *in utero* development, leading to groundbreaking discoveries in fetal immunity and organogenesis [19,30,31,33,51,62,121,152]. When obtained through ethics-board approved protocols and informed consent, donated human fetal tissue offers a wealth of information which has yet to be fully interrogated. Continuing advances in single cell phenotyping as well as cell and tissue culture will accelerate investigation of human development and the translational impact of these discoveries. Both protective and tolerogenic human T cell immunity begins *in utero* and the opportunity to intervene in immune developmental trajectories before the onset of disease underscores the importance of studying fetal T cell behavior and regulation.

1.1.6 Figure

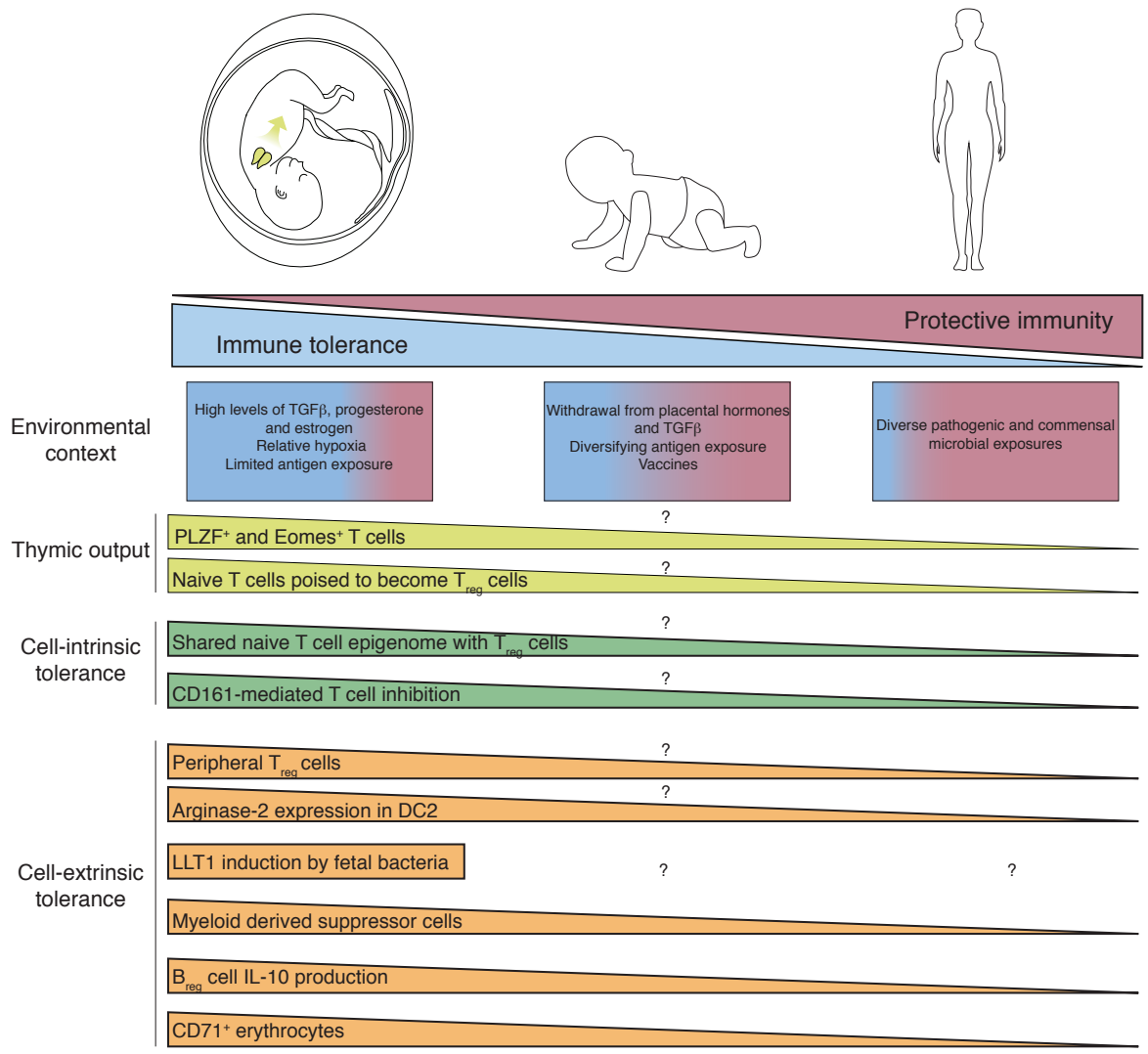

Figure 1.1 Mechanisms of Fetal T Cell Tolerance and Immune Regulation Thymic emigrants populate the periphery and encounter tissue-specific environments which differ between fetal, infant, and adult. T cell differentiation and function are governed by diverse cell-intrinsic and -extrinsic mechanisms of immune regulation, which are subsequently lost as in utero pressures for tolerance give way to the need for post-natal protective immunity.

1.2 Early Life Microbiome Impacts Immune Development

Our bodies are inhabited by vast microbial ecosystems, which are shaped by time[153], diet[154], and environmental exposures[155]. This ancillary, multi-cellular, cross-kingdom organ is capable of sensing environmental conditions and shaping human cell metabolism, immunity, and physiology[156]. The long-held assumption that the uterine environment is sterile has been challenged by many groups, because bacteria were detected in the placenta[157–159], fetal membranes[160], amniotic fluid[159], and human meconium[159,161–163] (the first neonatal bowel movement composed of swallowed amniotic fluid). These studies are fundamental, because ecologically, pioneering species colonizing a pristine habitat can play a substantial role in defining ecosystem conditions and the trajectory of primary succession, or the process of species accumulation and replacement over time[164]. For example, bacteria detected at 4 days of age in neonates correlate with community structure at 120 days of age[165]. Studies in neonates have underscored the importance of microbial pioneers on chronic inflammatory disease development in later life. Depletion of key species in very early-life predicts allergic recurrent wheeze development at age one year[166] and relative risk for asthma development at age 4 years[24]. Further, meconium of high-risk for asthma newborns is compositionally distinct from healthy control subjects[167]. Gut microbiome metabolites drive systemic effects on host immunity[168-169], such as promotion of T_{reg} cell development[170]. Consistent with these observations, fecal water of high-risk for asthma neonates suppresses Treg development *ex vivo*[24], suggesting that neonatal gut metabolites can drive tolerance breakdown associated with chronic inflammatory disease. This growing evidence suggests that complex host-microbe interactions that

define immunological health and disease may begin *in utero*[157,159,162] and therefore, chronic inflammatory diseases could be prevented before birth.

1.2.1 Neonatal gut microbiota perturbation is associated with childhood chronic inflammatory disease

Progressive assembly of the intestinal microbiome in early-life occurs in parallel with significant developmental hallmarks for the host, such as microvilli maturation, homeostatic epithelial turnover, and maturation of the gut-associated lymphoid tissue[171,172]. Commensal microbes are essential for the development of mucosal immune homeostasis and oral tolerance[171]. For example, germ-free mice exhibit polarization towards inflammatory responses[173], and regain normal immune homeostasis only when their gut microbiota are restored in early life[174]. Inappropriate neonatal microbiome assembly not only impacts early immune development, but also associates with chronic inflammatory disease development in later life[173,175,176].

I contributed to a study of a large cross-sectional birth cohort (n=130, median age=35 days) identified three distinct neonatal gut microbiota states (NGM), one of which was associated with a significant increased relative-risk of allergic sensitization and asthma development at age 2 years and 4 years, respectively. Neonates in this group lacked exposure to furred pets and were predominantly male, consistent with epidemiological factors associated with heightened risk of allergic sensitization in childhood[24]. These taxonomically distinct gut microbiota exhibited differential metabolic outputs and the sterile fecal water of high and low risk neonates induced significantly distinct $CD4^+$ T cell responses. Specifically, the high-risk NGM induced $CD4^+IL4^+$ pro-inflammatory T cells

and suppressed T_{reg} cells. T_{reg} cell suppression could be recapitulated *ex vivo* by exposure to 12,13 DiHOME, an oxylipin significantly increased in relative concentration in the high-risk neonates. This study indicated that the neonatal gut microbiota and its associated metabolic products influence T cell differentiation and function[24]. In a separate study, we have demonstrated that the meconium microbiota of high-risk for asthma neonates is compositionally distinct from healthy control subjects and that the former exhibit a significantly delayed gut microbiome developmental trajectory[167]. Therefore, peri-natal exposures program progressive gut microbiota composition and its metabolic capacity. Microbial metabolites, in turn, influence immune development and susceptibility for inflammatory disease development in childhood. It is not yet known whether microbial intestinal colonization is seeded *in utero* and whether these pioneering species drive mucosal immune development earlier than previously appreciated.

1.2.2 Evidence for and against bacterial presence *in utero*

The sterility of the *in utero* environment was initially proposed in 1900 by H. Tissier who was unable to culture bacteria from placentas on blood agar plates. Throughout the remainder of the century a growing understanding of pathologic bacterial presence in the reproductive environment was established, which led to lifesaving treatments. During the explosion of the microbiome field, partly due to the advance of 16S amplicon sequencing, groups began reporting the presence of variable bacteria in the placenta[157,159,177]. This suggested that, as least in the later phases of pregnancy, the uterine environment was not sterile. However, several latter studies contradicted this

finding[178–180]. These studies attributed much of the bacterial signal to contamination and/or bacteria commonly found in the vaginal tract. Because of the possibility of vaginal bacterial presence in the placenta and the fact that most placentas are vaginally delivered, the question of bacterial presence is very difficult to resolve by sequencing. While next-generation sequencing has spurred radical growth in the microbiome field, its limitations have been exposed in the interrogation of low-burden microbiomes[181,182]. Thus, studies of such niches require exceptional rigor in sample handling, data generation, analyses and interpretation, and must provide additional independent corroborating evidence to reject or support the proposed hypothesis. Today, sequencing is simply insufficient to address the sterility *in utero* hypothesis. Nevertheless, immunological and microbiological studies point to the presence of bacteria in the *in utero* compartment.

1.2.2.1 Immunological evidence for bacteria in utero

Fetal development occurs in a physically protected environment: the placenta and fetal membranes control the majority of interaction between the external environment experienced by the mother and the fetus. Many have therefore assumed that the antigenic encounters faced by the fetus were primarily restricted to self- and non-inherited-maternal-antigens. However, this assumption ignores long-standing observations that the fetus is not sheltered from all external antigenic stimuli, but rather that many antigens can indeed enter the fetal compartment. From studies measuring the T cell proliferative response in cord blood, responses to ovalbumin, cockroach extract, house dust mite (HDM) extract, and mouse extract were observed and in some cases was heightened in infants that went on to develop allergic disorders[183,184]. In a

landmark study, Prescott and colleagues demonstrated inflammatory Th2-skewed cord blood responsiveness to house-dust mite allergen and *Fel d* cat allergen[185], suggesting that not only proliferation but productive inflammation can be generated in response to allergens. A higher exposure to HDM during pregnancy may inhibit the development of CD4 memory cells and is associated with a lack of Th1 cells in cord blood[186]. Indeed, specific T cell clones to HDM were identified in cord blood of infants at high risk for atopy and these clones produced very high levels of IL-10 as well as exhibited a Th2 skew, though production of IFNγ was also observed[187]. The proliferative and cytokine response in cord blood T cells to allergens suggested that the fetus may encounter these antigens *in utero*, but may also be a result of an innate-like T cell responsiveness to stimulation. Direct observation of HDM allergen in amniotic fluid[188] and cord blood[186] challenges the latter hypothesis, suggesting that the fetus is indeed susceptible to transplacental and transamniotic allergen exposure. These early studies suggested that allergens and pollutants can enter the fetal compartment and generate inflammatory responses from the fetus and underscores the fact that the physical barrier in which the fetus develops is not absolute. While bolstering the molecular barriers to entry into the fetus is an exciting field of research, these studies further suggest that tolerizing the fetal response to antigens encountered *in utero* may lead to real therapeutic benefits for allergic disease.

If allergens and pollutants can penetrate the fetal compartment, it is plausible that factors that promote immune health and development are also permitted entry. Human milk oligosaccharides (HMOs), the complex carbohydrates highly abundant in human

milk, are detected in maternal circulation and urine as early as the first trimester[189] and recently were detected by high-performance liquid chromatography and mass spectroscopy in amniotic fluid of healthy term pregnancies[190]. HMOs promote intestinal immune tolerance and appropriate interactions with commensal microbiota[191]. By week 15 of gestation, the fetus begins to swallow large amounts amniotic fluid, thus the discovery of HMOs in amniotic fluid may suggest that the immune system is already being primed for future bacterial interactions. Supporting the fact that microbiomes are present in early life, bacteria and fungi are detected in meconium, the first stool of infants composed of swallowed amniotic fluid[167,192,193]. Initial bacterial encounters in the intestine may have important consequences on immunity because meconium microbiomes of high-risk for asthma infants is distinct from healthy controls[167]. Given the presence of memory T cells in the fetal intestine[29], it is therefore plausible that swallowed bacterial antigens may educate this mucosal T cell compartment before birth. These immunological data support the notion that bacterial antigens may be present *in utero*, however, they do not indicate whether viable microbiomes are present.

1.2.2.1 Microbiological evidence for bacteria in utero

Bacteria have been detected in amniotic fluid and the placenta through culture-independent studies of these tissues[159,160,194]. However, several studies have challenged the placental findings and attributed microbiome conclusions to contaminants arising from lab buffers[178–180]. In a large survey of pre-term and term placentas, de Goffau and colleagues were not able to identify a bacterial signal that could be excluded from contamination from biopsy or lab buffers in placental terminal villi from term and pre-term placentas[178]. The authors concluded that "*bacterial infection*

of the placenta is not a common cause of adverse pregnancy outcome," which is refuted by decades of clinical data that demonstrates bacterial inclusions in pre-term placentas through expert pathologist examinations[195]. Furthermore, microscopic evidence for intracellular bacterial signal in healthy term placentas delivered by C-section has been reported[196]. Microscopic evidence is imperative, as sequencing methods are noisy in low-abundance microbiomes and lead to false-positive signals in buffers[181,182]. It is also plausible that placental terminal villi, which bathe in pools of maternal blood during gestation, are indeed sterile. The presence of bacteria at this interface may be dangerous and likely highly controlled by maternal and fetal immune systems. A study sampling various anatomic regions of the placenta, with associated controls for contamination, found little signal at this site and a higher signal in fetal membranes and the maternal basal plate[197]. Thus, the presence of a microbiome at the placental interface is far from a settled issue. Simply relying on culture-independent methods is no longer sufficient to investigate the presence or absence of a microbiome and investigators should consider combining multiple lines of corroborating evidence such as *in situ* microscopy, cultivation, sequencing, and immune correlates as they answer this important question.

Beyond the placenta, fetal tissues may also harbor bacteria. Bacteria could transiently enter the placenta and establish niches in mucosal surfaces such as the intestine. Several groups, including ours, have reported the presence of microbes in human meconium[159,161–163,198,199]. Human meconium microbiota is taxonomically simple and most similar to that found in amniotic fluid[159]. Bacterial colonization in meconium is

crucial not only for appropriate mucosal immune development, but also because it can dictate microbial community structure in later life; for example, bacteria detected at 4 days of age in neonates correlate with community structure at 120 days of age[165]. While colonization through the birth canal during vaginal delivery may serve as a significant microbial exposure, a recent well-powered study posits that meconium microbiota is taxonomically distinct from maternal vaginal microbiota regardless of delivery mode, supporting the idea that these communities are founded *in utero*[199]. Thus, it is of great immunological and microbiological interest to determine whether bacteria are present in the fetal intestine during gestation.

1.3 Aims of the Study

The ability of the fetal immune system to respond to a broad suite of antigens indicates that immune memory can be shaped by fetal exposures. Thus, the aim of this study is to investigate whether bacteria are present in the fetal intestine and whether these bacteria relate to and modulate intestinal immunity. This study further aimed to set a standard for investigation of low-burden microbiomes through development of novel sequencing techniques and collection of multiple lines of corroborating evidence.

CHAPTER 2: Viable bacterial colonization is highly limited in the human intestine *in utero*

Material for this chapter was modified from:

Rackaityte E, Halkias J, Fukui EM, Mendoza VF, Hayzelden C, Crawford ED, Burt TD, Lynch SV (2020). Viable bacterial colonization is limited in the human intestine *in utero*. *Nature Medicine*. *26*(4), 599-607.

2.1 Abstract

Mucosal immunity develops in the human fetal intestine by 11-14 weeks gestation, yet whether viable microbes exist *in utero* and interact with the intestinal immune system is unknown. Bacterial-like morphology was identified in pockets of human fetal meconium at mid-gestation by scanning electron microscopy (n=4) and a sparse bacterial signal was detected by 16S rRNA sequencing (n=40 of 50) compared to environmental controls (n=87). Eighteen taxa were enriched in fetal meconium with *Micrococcaceae* (n=9) and *Lactobacillus* (n=6) the most abundant. Fetal intestines dominated by *Micrococcaceae* exhibited distinct patterns of T cell composition and epithelial transcription. Fetal *Micrococcus luteus*, isolated only in the presence of monocytes, grew on placental hormones, remained viable within antigen presenting cells, limited inflammation *ex vivo*, and possessed genomic features linked with survival in the fetus. Thus, viable bacteria are highly limited in the fetal intestine at mid-gestation, though strains with immunomodulatory capacity are detected in subsets of specimens.

2.2 Introduction

Mucosal immunity is evident in the human fetal intestine by the end of the first trimester[1,2]. The developing intestine is populated by migrating dendritic cells capable of responding to microbial stimuli and initiating robust T cell responses [200]. By week 13 of gestation, memory T cells are abundant in the human fetal intestine [29–31,44,45,201], possess pro-inflammatory potential [29], and influence epithelial maturation [30]. These cells also clonally expand in response to foreign antigens [31], suggesting that the presence of select microbes may influence prenatal T cells.

Recent evidence for bacterial presence *in utero* comes from DNA-based, culture-independent studies of the placenta [159,160,194] and amniotic fluid [159], though other studies have refuted the presence of bacteria at these sites and attributed signal to extraction kit contaminants [178,180,202]. However, whether microbes exist within the human fetal intestine and influence the earliest stages of mucosal immune development has not been examined. Neonatal meconium, the first stool of a newborn, is comprised of amniotic fluid swallowed during gestation and contains a very simple microbiota [167,199]. Heightened risk of chronic inflammatory disease in childhood, such as asthma, is associated with maternal lifestyle factors (e.g. farming; Yu et al., 2018b) and with a distinct and perturbed neonatal meconium [167], the metabolic products of which induce inflammation *ex vivo* [24]. Thus, early-life gut microbiomes have the potential to influence immunity in later life, prompting our hypothesis that specific and highly limited immunomodulatory microbes might be present in the fetal intestine and contribute to pre-natal immune priming.

2.3 Results

To determine whether bacteria are present in the intestine *in utero*, we first performed direct visualization of human fetal intestines obtained from terminated pregnancies (**CHAPTER 3**; 18-23 weeks of gestation). Terminal ileum intestinal segments were fixed and thin sectioned for light microscopy. Fluorescent *in situ* hybridization for eubacteria of 5 μm sections of fetal ileum suggested an extremely sparse bacterial signal (**Figure 2.1a-c**). Because rare signal is further diluted by thin-sectioning required for light microscopy, scanning electron microscopy (SEM) was performed on four independent fetal terminal ileum specimens; environmental exposure was minimized by ligation of the intestinal segments prior to processing (**Figure 2.2 a, CHAPTER 3**). In three of four independent fetal specimens (**Figure 2.2 b-c**; Specimens 1-3), clusters of tightly packed cellular structures morphologically and proportionally consistent with bacterial cocci were observed in discrete, isolated pockets of meconium, deeply embedded within existing mucin structures *in situ* (**Figure 2.2 b**). Specimen 4 had limited meconium in the lumen as evidenced by exposed epithelial cell structures; clusters of cocci were not observed in this specimen (**Figure 2.2 c**). Confirming bacterial localization to meconium, these coccoid structures were not observed in sub-epithelial regions, such as the lamina propria or muscularis (**Figure 2.2 c-iv**). Thus, discrete clusters of cellular structures consistent with coccoid bacterial morphology, embedded within isolated pockets of fetal intestinal meconium are evident during the second trimester of human gestation.

We established a bank of human fetal small intestine meconium samples (n=50 subjects; n=149 samples, n=87 technical and procedural controls; **Figure 2.3**; **Table**

2.1; **CHAPTER 3**) to quantify and identify these bacteria using molecular techniques. Irrespective of the small intestine segment sampled, and consistent with our SEM observations, total bacterial burden by 16S rRNA copy number was low and variable in fetal meconium, but significantly greater than that of extraction buffer, procedural swab, hospital room air swab, blank cotton swab, or fetal kidney controls (9 out of 13 meconium specimens were greater that the 75th percentile *16S rRNA* copy number in controls; **Figure 2.1 d-e**). These data suggested that bacterial presence, if any, was extremely low and nearing the limit of molecular detection. To enhance bacterial signal prior to V4 *16S rRNA* gene amplification, human mitochondrial 16S DNA (mtDNA) was depleted using Cas9 targeting (Depletion of Abundant Sequences by Hybridization[204], DASH; **CHAPTER 3**); this did not alter the profile of detected bacteria as compared to band selection and gel extraction (**Figure 2.4 a-c**). In *16S rRNA* datasets controlled for environmental and procedural contamination (**Table 2.1-2.2**, **CHAPTER 3**), a simple bacterial profile was identified in 40 of 50 subjects comprising a median of 23.5 operational taxonomic units (OTUs) with ≥ 5 sequence read counts per sample (**Table 2.1-2.2, Figure 2.5 a**). Bacterial profiles were consistent in replicate samples along the length of the intestine within subjects (n=108 samples; LME p=0.78**; Figure 2.5 b**) and inter-sample profile distances were greater than intra-sample distances, indicating that the signal detected was unlikely due to uniform contamination (**Figure 2.5 c**). Thus, subsequent analyses focused on the mid-segment (n=40) of the small intestine.

Eighteen taxa were significantly enriched in fetal meconium (DESEQ2; L2FC≥2, FDR<0.05) compared to procedural swabs and kidney controls (**Figure 2.2 d**). Distinct

bacterial profiles were evident and defined by the dominant organism detected (PERMANOVA R^2 = 0.18, p=1e-5; **Figure 2.6 a**). *Lactobacillus* OTU12 and *Micrococcaceae* OTU10 represented the two highest ranked fetal meconium taxa by relative abundance (**Figure 2.5 d**) and the dominant taxon within distinct subsets of samples (the OTU with the greatest proportion of *16S rRNA* reads in a given sample; *Lactobacillus*-meconium, LM; n=6, or *Micrococcaceae*-meconium, MM; n=9; **Figure 2.5 e**). The remaining samples were variably dominated by other bacterial taxa (Other-meconium, OM; n=25), including distinct taxa within *Lactobacillus* and *Micrococcaceae*, as well as *Bacteroides*, *Bifidobacteria*, and *Prevotella* (**Figure 2.5 e**). OM samples represented the majority of meconium studied and their 16S rRNA profiles were similar to that of biological contamination controls (**Figure 2.5 f-g**). Only LM and MM samples (n=15) exhibited significantly distinct bacterial profiles from OM samples, procedural and kidney controls (PERMANOVA, R^2=0.167, p=1e-5; **Figure 2.5 f-h**), and from a variety of technical controls (n=48, **Figure 2.5 i**). *Lactobacillus* OTU12 and *Micrococcaceae* OTU10 were not identified as contaminants using stringent thresholds (*decontam* R package; *p* threshold=0.6, **Table 2.3**). Thus, the majority of human fetal intestinal samples at this stage of gestation produced a signal that could either be attributable to noise associated with molecular detection methods in low-burden samples and/or to a lack of bacteria. Nonetheless, the identification of coccoid structures *in situ* and the observation that approximately 30% of fetal intestinal specimens produced a bacterial profile distinct from that of biological and technical controls led us to determine whether

the fetal intestinal immune context related to variance in meconium bacterial detection in paired samples.

Taking advantage of intestinal immune profiling data generated at the time of specimen collection, we examined the composition of lamina propria (LP) T cells paired with meconium *16S rRNA* data (n=22). Confirming recent findings[29], PLZF$^+$ CD161$^+$ CD4$^+$ Vα7.2$^-$ TCR$\alpha\beta^+$ T cells were highly abundant in the fetal lamina propria in contrast to mesenteric lymph node and spleen (**Figure 2.7**, **Figure 2.8 a**). We noted a significant relationship between meconium bacterial profile and LP PLZF$^+$ CD161$^+$ T cells (PERMANOVA R^2= 0.11, p=0.0004, **Figure 2.6 b**). LP samples with the highest PLZF$^+$ CD161$^+$ T cell proportion were associated with meconium dominated by *Micrococcus* OTU10 (MM; **Figure 2.6 b**) and MM samples exhibited significantly higher proportions of PLZF$^+$ CD161$^+$ T cells compared to all other samples (**Figure 2.6 c, Figure 2.8 b-c**).

Fetal intestinal memory T cells, the majority of which express PLZF and CD161[6], have recently been reported to support epithelial stem cell function[30]. Therefore, paired epithelial cell layer transcriptomes associated with MM versus other specimens were analyzed (n=7, n=6, respectively). MM-associated epithelium (MM-E) samples exhibited distinct transcriptional programs from meconium samples dominated by all other taxa (PERMANOVA p=0.02 R^2=0.16, **Figure 2.6 d-g**) and from LM-associated epithelium (LM-E, n=3) and OM-associated epithelium (OM-E, n=3) groups (**Figure 2.8 d**). Gene set enrichment analysis (GSEA) identified genes associated with intestinal epithelial stem cells, transit amplifying cells, and secretory progenitors as enriched in MM-E (e.g. *LGR5*, *SOX9*, *NOTCH1*, *NOTCH4* ; **Figure 2.6 g-h**, **Table 2.4**), consistent with the

ability of fetal memory T cells to promote epithelial stem cell function. *MUC3A* was downregulated in MM-E (**Figure 2.6 f**), yet transcripts associated with TLR-signaling (*NFKB2*, *TNFSF15*), phagolysosome function (*NOS2*), immune cell chemoattraction (*CXCL1-3* and *CCL20*), and macrophage inhibition (*CD200*) were enriched (**Figure 2.6 f-h, Table 2.4**). These transcripts indicated distinct programs of immune cell recruitment and regulation in the context of a nutritionally limited intestinal niche in samples with *Micrococcaceae* OTU10 present.

To determine whether intestinal *Micrococaceae* was viable, we attempted isolation from cryopreserved MM fetal meconium samples with the highest read counts for this taxon. *Micrococaceae* isolates could not be recovered using traditional selective media for this genus and were only obtained under culture conditions that mimicked the fetal intestinal environment (**Table 2.5**), including addition of placental steroid hormones or THP1 human monocyte cells, suggesting they may represent microbial selective pressures *in utero*. *Lactobacillus* could not be cultured from MM samples using these culture methods or traditional selective media for the genus. Fetal bacteria isolated in monocyte co-cultures were classified as *Micrococcus* using full-length *16S rRNA* sequence to interrogate the SILVA database (fetal isolate Micro36; **Table 2.5**). Additional strains were also isolated in the presence of placental hormones (**Table 2.5**). The V4 region of the *Micrococcus* isolate exhibited high homology with OTU10 (97%; **Figure 2.9 a**, **Figure 2.10, Table 2.5**).

We hypothesized that fetal *Micrococcus luteus* exhibits a fitness advantage over phylogenetic relatives under culture conditions that mimic the intestinal environment *in*

utero. Thus, we examined its growth and that of two reference strains MicroRef1 (ATCC 4698) and MicroRef2 (ATCC 12698) in the presence of peak third trimester cord blood concentrations [205] of progesterone and/or β-estradiol. Micro36 exhibited the unique ability to grow on progesterone and β-estradiol in carbon limiting media, albeit to low cell densities (**Figure 2.9 b**; **Figure 2.11 a-b**). In carbon-rich conditions, consistent with previously reported bacteriostatic effects of steroid hormones [206], progesterone and β-estradiol (but not □-estradiol alone) universally inhibited growth of all three *M. luteus* strains (**Figure 2.9 b; Figure 2.11 c-f**). These data suggest that conditions of low substrate availability coupled with pregnancy hormones permit limited growth of specific fetal bacterial isolates, offering an explanation for the control of bacterial burden in the intestine during human gestation.

The necessity of monocytes for initial *Micrococcus* isolation (**Table 2.5**) suggested the capacity for survival within phagocytic cells. Isolated primary human fetal intestinal HLA-DR$^+$ antigen presenting cells (APCs) were cleared of intracellular bacteria (**CHAPTER 3**), and incubated with fetal *Micrococcus* isolates to permit phagocytosis followed by gentamicin protection assays. At 24h, 1x10^7 CFU mL^{-1} of Micro 36 was recovered and the fetal isolate remained viable in APCs at 48h at 1x10^6 CFU mL^{-1} (**Figure 2.9 c**), indicating a capacity for prolonged intracellular survival. Control reference strains MicroRef1 and to a lesser extent MicroRef2 were non-viable under comparable conditions (**Figure 2.9 c**). Similar results were obtained using a RAW264.7 macrophage cell line with an additional *E. coli* control (**Figure 2.11 g**) and gentamicin resistance did not develop in the time course of either of these experiments (**Figure**

2.11 h-i). The ability of this fetal *Micrococcus* strain to persist inside phagocytes offers a potential mechanism of protected entry into the fetal intestine.

Whole genome sequencing of Micro36 (**Table 2.6**) permitted high resolution taxonomy of the isolate and identified shared and unique genomic features when compared to phylogenetic relatives. Micro36 exhibited 96.9% whole genome average nucleotide identity (ANI) to a reference genome of *M. luteus* and clustered by whole genome ANI with other human, but not environmental *M. luteus* isolates (**Figure 2.9 d**, **Table 2.7**). Pan-genomic analysis of our fetal *Micrococcus* and all available *Micrococcus* genomes identified shared single-copy genes (**Figure 2.12**) used to build highly resolved phylogeny (bootstrap value = 1 for relevant clade, **Figure 2.9 e**). Using a 96.5% ANI speciation cut-off [207], Micro36 was classified as a strain of *M. luteus*.

Compared to *M. luteus* (MicroRef1), Micro36 exhibited 425 unique genes, 256 of which were annotated (**Table 2.8**). Genomic features of Micro36 included two sterol carrier proteins and a putative steroid ketoisomerase, which typically facilitates degradation of steroid hormones. The genome also encoded reactive oxygen and nitrogen radical reducing enzymes, and genes in the catechol pathway. While the broader prevalence of these genes is yet to be determined, these data offer plausible mechanisms by which Micro36 may grow on placental hormones [208] (**Figure 2.9 b**), remain viable in phagocytes [209] (**Figure 2.9 c**), and under conditions of hypoxia associated with elevated *NOS2* [210] in MM-associated epithelia (**Figure 2.6 f**).

To determine whether fetal *Micrococcus luteus* isolates are found in post-natal infant samples, we utilized publicly available 16S rRNA data from three independent

early-life cohorts [24,167,211]. Sequences exhibiting ≥97% homology to our fetal isolates were detected throughout early life (up to 12 months; **Table 2.9**); however, sequences with the highest homology (≥99%, **Table 2.9**) were primarily found in infant meconium (first stool) samples (**Figure 2.13 a**). *M. luteus* was found in low in abundance in infant samples but highest in post-natal meconium in two independent metagenomic cohorts [212,213] (**Figure 2.13 b-c**). These species were detected on maternal chest and in vaginal introitus at delivery and were not highly abundant in maternal stool (**Figure 2.13 b-c**). Among our fetal meconium specimens, neither the number of detected OTUs per sample nor the relative abundance of *Micrococcaceae* OTU10 were significantly correlated with gestational age when all samples were considered (**Figure 2.13 d-e**). However, a positive correlation between OTU10 relative abundance and gestational age was evident within the MM group (Pearson's r=0.5, p=0.1; **Figure 2.13 f**). This suggests that intestinal Micro36 or highly related strains may increase during gestation, persist at least until birth and may be succeeded in the post-natal period by phylogenetically related species.

MM was associated with a distinct program of fetal epithelial gene expression (**Figure 2.6**). We thus examined the capacity of fetal *M. luteus* to induce characteristic features by profiling the transcriptome of primary human fetal intestinal epithelial cells (n=2) exposed to Micro36 for four hours *in vitro*. Transcriptional differences were observed when Micro36 exposed epithelia were compared to media controls (**Figure 2.14 a**). As expected, short-term exposure to planktonic bacterial cultures *in vitro* did not fully recapitulate the global fetal intestinal transcriptome patterns observed in MM-E

(**Figure 2.14 a**). Nonetheless, Micro36 exposure induced the expression of *TLR6* and its downstream regulator *NKFB* was enriched in MM-E (**Figure 2.13 a-b**). These data suggest that even following short-term fetal bacterial exposure, fetal intestinal epithelial cells exhibit transcriptional responses to *Micrococcus* that partially recapitulate features observed in MM-E.

The heightened expression of immune cell recruitment and regulatory mediators in MM-E (**Figure 2.6 f**), led us to assess the capacity of Micro36 to influence primary fetal HLA-DR$^+$ antigen presenting cell (APC) function obtained from spleen (**Figure 2.15**). Without decreasing cell viability (**Figure 2.16 a**), Micro36 and two reference strains induced fetal APC production of cytokines associated with maturation of intestinal macrophages (GM-CSF and G-CSF) as well as IL-10 (**Figure 2.14 c-e**), which promote a tolerogenic environment [104–106]. Micro36 induced lower levels of TNF□ compared with reference strains, indicating its ability to limit APC inflammation (**Figure 2.14 f**). LLT1, the natural ligand for the fetal-specific inhibitory C-type lectin CD161, is expressed on fetal intestinal macrophages [29] and can be induced upon TLR activation of APCs [214]. Given the capacity of fetal *M. luteus* to induce fetal epithelial *TLR6 in vitro*, we examined whether it could elicit LLT1 expression on primary human fetal splenic APCs. Compared with a phylogenetically related strain, only fetal Micro36 induced LLT1 expression and in proportion with multiplicity of infection, albeit to lower levels than observed in lamina propria APCs *ex vivo* (**Figure 2.14 g-i**).

As ligation of CD161 inhibits IFNγ production by fetal intestinal PLZF$^+$ CD161$^+$ T cells [29], we hypothesized that Micro36 may specifically regulate the inflammatory

potential of these T cells. Sorted splenic APCs (**Figure 2.16 b**) pre-conditioned with Micro36 or MicroRef1 were co-incubated with autologous, fetal intestinal effector memory T cells (>99% purity), the majority of which expressed PLZF and CD161[29] (**Figure 2.16 c-d**). Micro36 exposure resulted in a significant reduction of IFNγ production by these T cells as compared to MicroRef1 (**Figure 2.14 j**), indicating induction of immunotolerance by fetal *M. luteus*. While exerting an effect on $PLZF^+$ T cell function, Micro36 did not impact the proportional accumulation of these cells or regulatory T (Treg) cells (**Figure 2.16 e-f**), and did not influence the production of IL17A, IL17F, GMCSF, IL-4, IL-10, IL-13, or TNFα after five days of APC-T cell co-cultures (**Figure 2.16 g-m**). Together data suggest that fetal intestinal immune cells are capable of mounting an inflammatory response to bacteria and that fetal *M. luteus* may circumvent this by inducing tolerogenic APCs and inhibiting IFNγ production by fetal memory T cells.

2.4 Discussion

Whether bacteria are present *in utero* is contentious because of the inherent limitations of molecular methods that are commonly used to identify bacteria in low-burden environments. Background noise and false positives in technical negative controls are common when bacterial burden is extremely low, therefore simple removal of all taxa detected in these controls is not deemed appropriate [215]. In our study, despite improving the current molecular methods to boost bacterial signal (**CHAPTER 3**), 16S rRNA sequencing data remained noisy and, for the majority (70%) of samples, we identified a sparse bacterial signal that was indistinguishable from procedural or fetal kidney controls. Using a taxon filtering approach that focused on the signal detected in the majority of negative controls led to the identification of a *Micrococcus* taxon as one of a small number of discriminant taxa enriched in meconium samples. With this molecular data as a guide for fetal bacterial species isolation, we subsequently cultured *M. luteus* from parallel preserved samples that never encountered extraction buffers during the course of sample processing. The molecular bacterial signal was proportionately more sparse than the signal identified by microscopy, which indicates that our removal of contaminant taxa may have suppressed a broader bacterial signal and supports previous reports of false-positive classification of contaminants in buffer controls [215]. Our molecular data identified only 17 additional taxa not attributable to contamination in our cohort of 50 fetal intestinal specimens, suggesting that conditions in the fetal gut highly limit bacteria, the mechanisms of which warrant further investigation.

While there is debate regarding the best methods to address contamination in low-burden bacterial environments, our data suggest that current molecular methods alone are insufficient to support or reject the *in utero* sterility hypothesis. By combining molecular bacterial detection, immune correlates, microscopy, strain isolation, and *ex vivo* experiments, our study provides direct and indirect evidence for the presence of sparse but viable bacteria in the human fetal intestine at mid-gestation with the capacity to limit inflammatory potential by fetal immune T cell populations. While it is possible that the bacterial signal identified may arise from a contamination from a source not investigated in this study, in our judgement, the corroborating evidence suggests that restriction of bacterial entry into the human fetal intestine is not absolute. We note that the lack of clinical data associated with specimens in our cohort (as mandated by our institutional protocols) limits our ability to examine pregnancy features associated with identification of *Micrococcus*, an important consideration for subsequent studies.

Fetal *Micrococcus* most likely arises from maternal cervico-vaginal microbiomes, which commonly house this genus [216,217]. While our fetal *Micrococcus* isolate exhibited genome similarity to vaginal *Micrococcus* isolates, it also encoded strain-specific genes not found in these strains, which may provide survival advantages under the strong selective conditions of the fetal intestine. Indeed, fetal *M. luteus,* which encoded a ketoisomerase putatively involved in steroid metabolism, exhibited the unique capacity for limited growth in the presence of progesterone and β-estradiol in low-nutrient conditions. This this observation coupled with the fact that *M. luteus* can exist in a dormant, viable but non-culturable state under starvation conditions [218], may allow it to

persist under conditions of limited nutrition and pregnancy hormone exposure in the fetal intestine. Consistent with this hypothesis, intestinal epithelial transcriptome analysis suggested lower expression of microbial nutritional substrates including the mucin glycoprotein *MUC3A* in fetal samples in which *M. luteus* was found. Collectively, these observations offer a potential explanation for detection of *M. luteus* in specific subsets of fetal intestinal samples in which prevailing conditions of nutrient limitation and pregnancy hormones may contribute to its limited presence.

Bacterial presence may not be pervasive in the second trimester, yet in our study *Micrococcus* was associated with the immunological status of the intestinal epithelium and the lamina propria memory T cell compartment. *Micrococcus* in the fetal intestine modulates mucosal immunity and reciprocally, the immune system influences which microbes are tolerated by the host [219]. Thus, it is plausible that epithelial and lamina propria immunity additionally select for specific *M. luteus* strains. In turn *Micrococcus* limits the inflammatory ability of these cells, which may foster a tolerant environment that permits its survival *in utero*. However, we recognize that other developmental factors such as stem cell niche [51], the predisposition for fetal T cells to develop into regulatory T cells [33], and antigens from swallowed amniotic fluid [220] also shape prenatal immunity.

Recent studies of fetal immunity have led to the hypothesis that bacterial signals *in utero* initiate an adaptive immune response, including T cell activation. Fetal T cells respond to non-inherited maternal- and self- antigens and are capable of memory formation in the intestine[29,31,32]. The presence of bacteria in the fetal intestine suggests

that bacterial antigens may also contribute to T cell activation, as fetal intestinal T cells do not exclusively exhibit a tolerogenic phenotype. Their ability to produce inflammatory cytokines in the absence of systemic inflammation indicates intestinal compartmentalization of immune response *in utero*, which may be essential for tolerance or clearance of fetal intestinal bacteria. *Micrococcus* enrichment in the fetal gut associated with increased proportions of IFNγ-producing mucosal memory $PLZF^+$ $CD161^+$ T cells[6] and only the fetal *Micrococcus* isolate reduced IFNγ production by these T cells. While fetal *Micrococcus* likely elicits a number of responses, the specific induction of LLT1 on antigen presenting cells identifies a potential bacterial mechanism of immune regulation that is unique to fetal adaptive immunity. Thus, immunological memory to fetal *Micrococcus* may begin *in utero*.

How the fetal intestine limits bacterial presence remains underexplored, though the ability of specific bacteria to persist in nutrient limiting conditions, grow on pregnancy hormones and survive within phagocytes offer plausible mechanisms for survival *in utero*. The implications of *in utero* bacterial interactions or lack thereof on long-term health remain to be determined.

2.5 Acknowledgements

We express gratitude to the tissue donors. Thanks to M. Ng, H. Boushey and A. Molofsky for thoughtful critique of this manuscript. UCSF flow core provided instrumentation assistance, which was supported by National Institutes of Health (NIH) grant no. P30 DK063720 and NIH Shared Instrument Grant no. 1S10OD021822-01. The FE-SEM and supporting facilities were obtained under National Science Foundation (NSF)-MRI award no. 0821619 and NSF-EAR award no. 0949176, respectively. This study was supported in part by the UCSF Clinical and Translational Science Institute Pilot Award for Basic and Translational Investigators grant no. 2014908 to JH. ER was supported by NSF Graduate Research Fellowship grant no. 1650113 and by the National Institute of Allergy and Infectious Diseases (NIAID) of the NIH F31AI136336. JH was supported by NIH NIAID grant no. K08 AI128007. SVL was supported by NIH NIAID grant no. AI114271 and UG30D023282. We thank C. Domingo for support of the electron microscopy core facility at SFSU and D. Mars for assistance with FE-SEM sample preparation. We thank J. L. DeRisi at UCSF for generously sharing custom-made T7 RNA polymerase with our group. We express gratitude to E. Stapleton for illustration assistance and to J. M. Rackaitis for support during these studies. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NSF or the NIH.

2.6 Figures

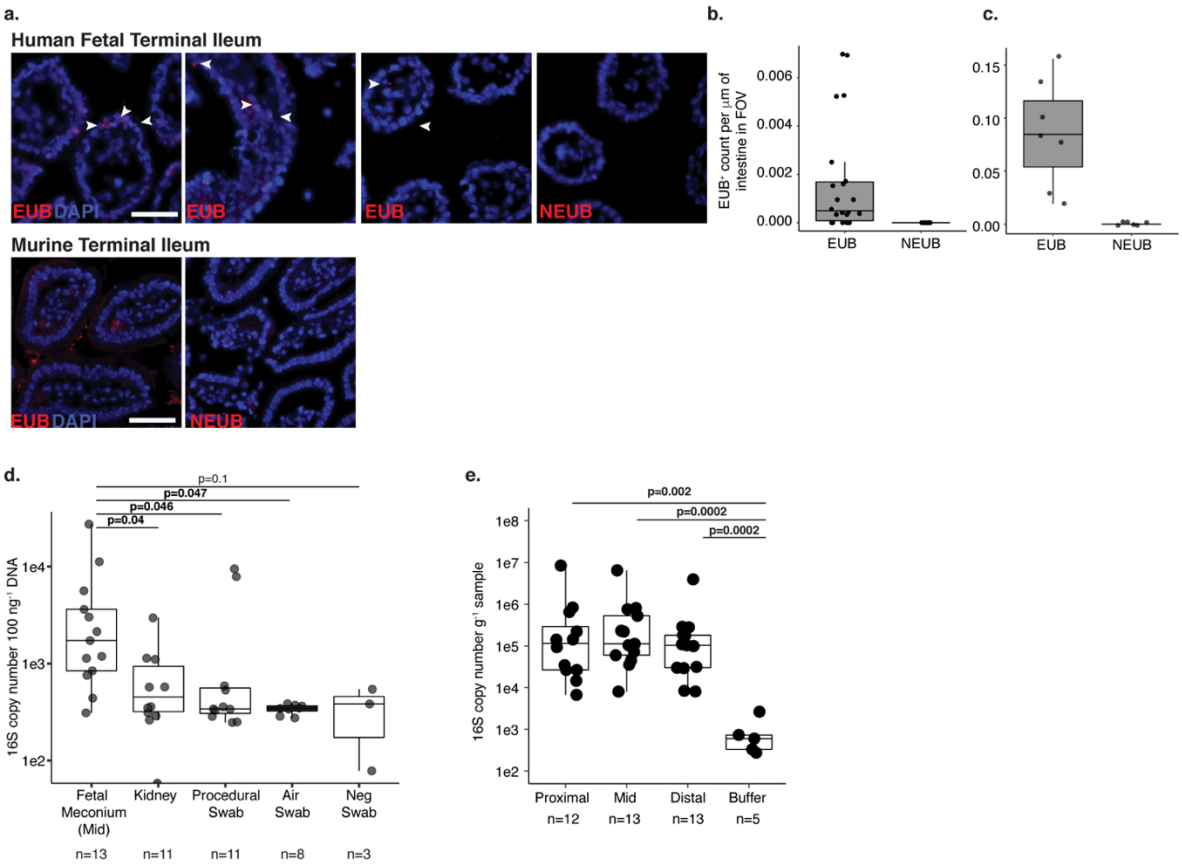

Figure 2.1 Low-burden bacterial signal in fetal meconium. a. Fluorescent *in situ* hybridization probes targeting eubacteria (EUB) or non-targeting probe (NEUB) in 5 μm cryosections of human fetal (top panel) or murine (bottom panel) terminal ileum at 400x magnification. Arrowheads indicate EUB-positive findings in fetal sections. Scale bar corresponds to 50 mm. Quantification of independent fields of view (FOV) per mm within a representative **b.** human fetal (n=28) or **c.** murine (n=13) intestinal segments, where n is an independent segment of the intestine. For a-b, representative images of three fetal specimens and three mice. **d.** Total 16S copy number per 100 ng gDNA in meconium from mid-section of the fetal small intestine (n=13), fetal kidney (n=11), and procedural (n=11), air (n=8), or blank (n=3) swab was quantified by qPCR of DNA extracts using a standard curve where n represents biologically independent samples; two-sided Satterthwaite's method on linear mixed effects model to test for significance. **e.** Total 16S copy number per gram frozen sample in meconium from proximal (n=12), mid (n=13), and distal (n=13) sections of the fetal small intestine where n represents independent biological samples across the length of the intestine or extraction buffer (n=5, n represents biologically independent samples) was quantified by qPCR of DNA extracts using a standard curve; two-sided Wilcoxon rank sum test for significance compared to buffer control. Boxplots indicate the median (center), the 25th and 75th percentiles, and the smallest and largest values within 1.5× the interquartile range (whiskers).

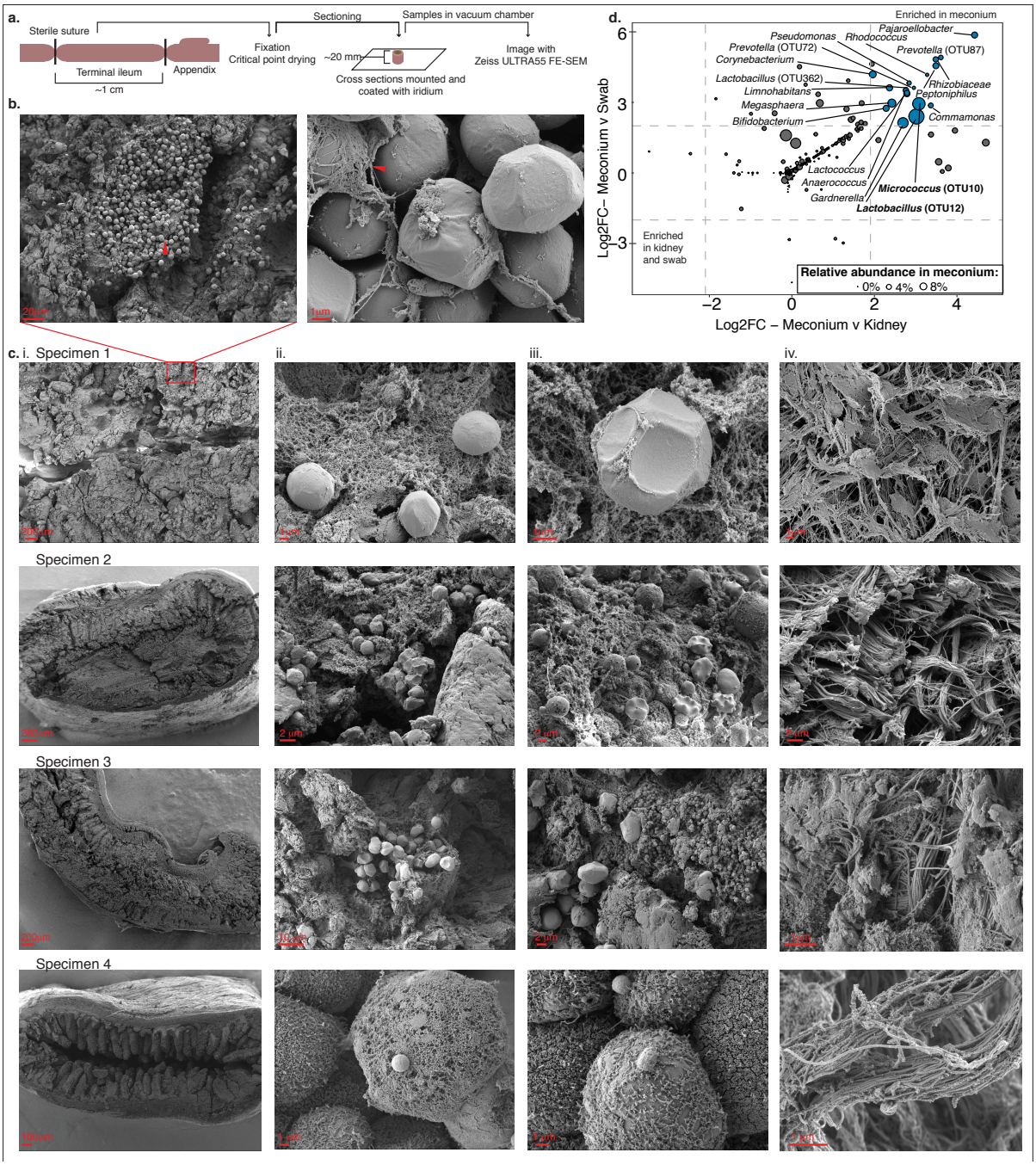

Figure 2.2 Rare bacterial structures in fetal meconium a. Schematic of sample preparation method of fetal intestines for scanning electron microscopy. **b.** Representative scanning electron micrographs of fetal intestinal lumen, arrowheads indicate pockets of bacterial-like morphology in meconium at 3 000 (left) and mucin-embedded structures at 50 000 (right) times magnification. **c.** Scanning electron micrographs of four biologically independent fetal intestinal specimens (i.) at low magnification, (ii. -iii.) two independent regions within intestinal lumen, and (iv.) sub-epithelial region outside of the lumen. Scale bars indicate size. Experiments in b-c were repeated 4 times. **d.** Significantly enriched taxa (Log2-fold change> 2, false discovery

rate <0.05) in meconium (n=40) as compared to both kidney (n=7) and procedural swab (n=14) controls after removal of technical negative OTUs where n indicates biologically independent specimens. Dots represent differential taxa and are scaled by percent relative abundance in meconium; top abundant taxa are labeled. DESEQ2 of unnormalized reads was used to determine Log2-fold change and a two-sided false discovery rate.

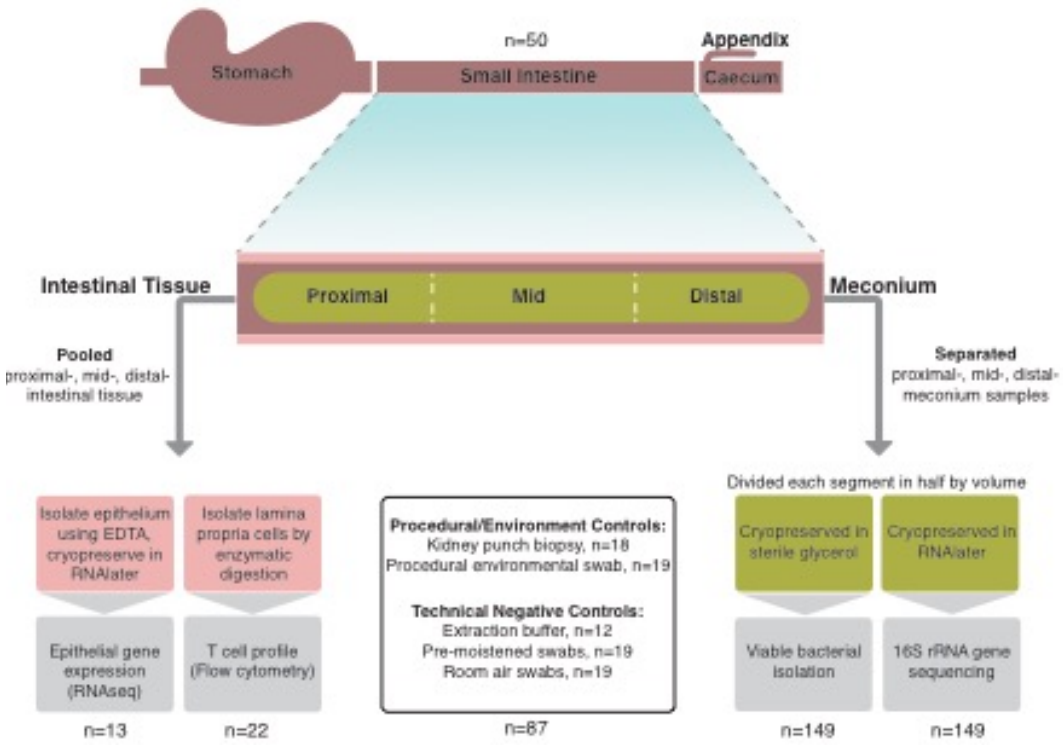

Figure 2.3 Collection method for fetal intestinal sample bank Uninterrupted small intestine sections were divided into equal thirds and internal contents (meconium) cryopreserved for either genomic DNA extraction (in RNAlater) or bacterial isolation (in 50% v/v glycerol). Remaining intestinal tissue from all three sections was pooled and washed with EDTA to recover epithelium (preserved in RNAlater for subsequent RNAseq analysis) and enzymatically digested to isolate lamina propria cells (for immediate analysis by flow cytometry). Internal kidney punch biopsies and surgical environmental swabs served as procedural or environmental controls. Extraction buffer, pre-moistened swabs, and pre-moistened swabs held in the surgical room air for 30 seconds served as technical negative controls.

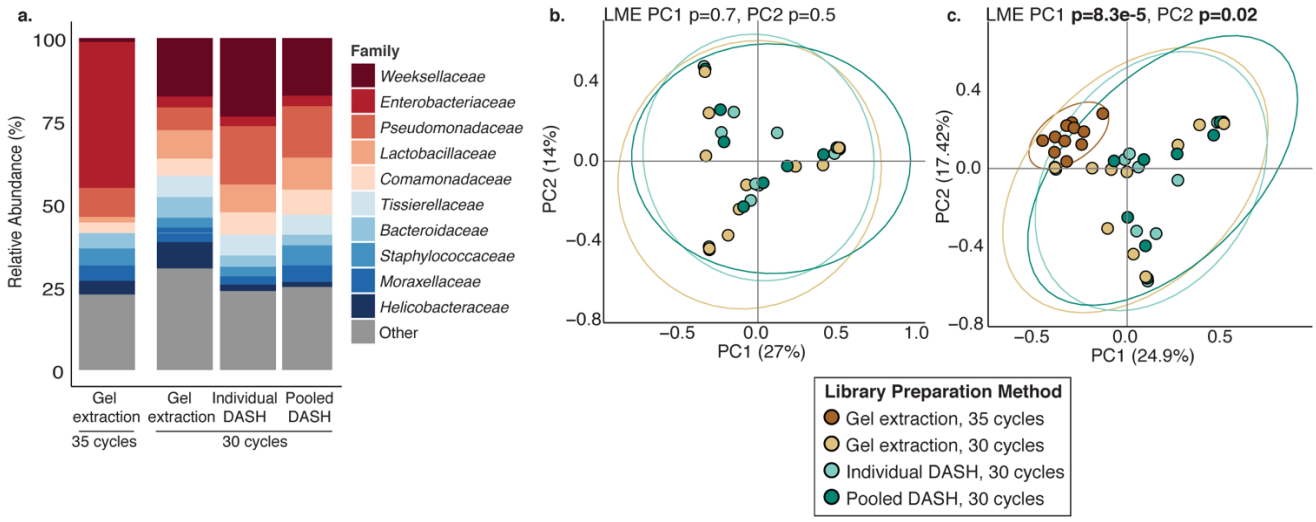

Figure 2.4 Depletion of mtDNA by Cas9 does not alter bacterial composition after 30 cycles of amplification 16S rRNA V4 profiling of a subset (n=10) of banked fetal meconium samples using different library preparation methods: gel extraction and 30 or 35 cycles of amplification, or 30 cycles combined with DASH performed on individual samples (Individual DASH) or on the library pool (Pooled DASH). **a.** Expansion in *Enterobacteriaceae* family is detected in 35-cycle amplification method, while small expansion of *Pseudomonadaceae* is detected post-DASH. Principal coordinates analysis of Bray Curtis distances of libraries using **b.** 30 cycles of amplification or **c.** 30 and 35 cycles of amplification, latter to provide an outgroup known to skew bacterial composition. Ellipses indicate 95% confidence intervals. All p-values were calculated using two-sided Satterthwaite's method on Linear Mixed Effects (LME) modeling to correct for n=10 paired samples that underwent multiple library preparation methods.

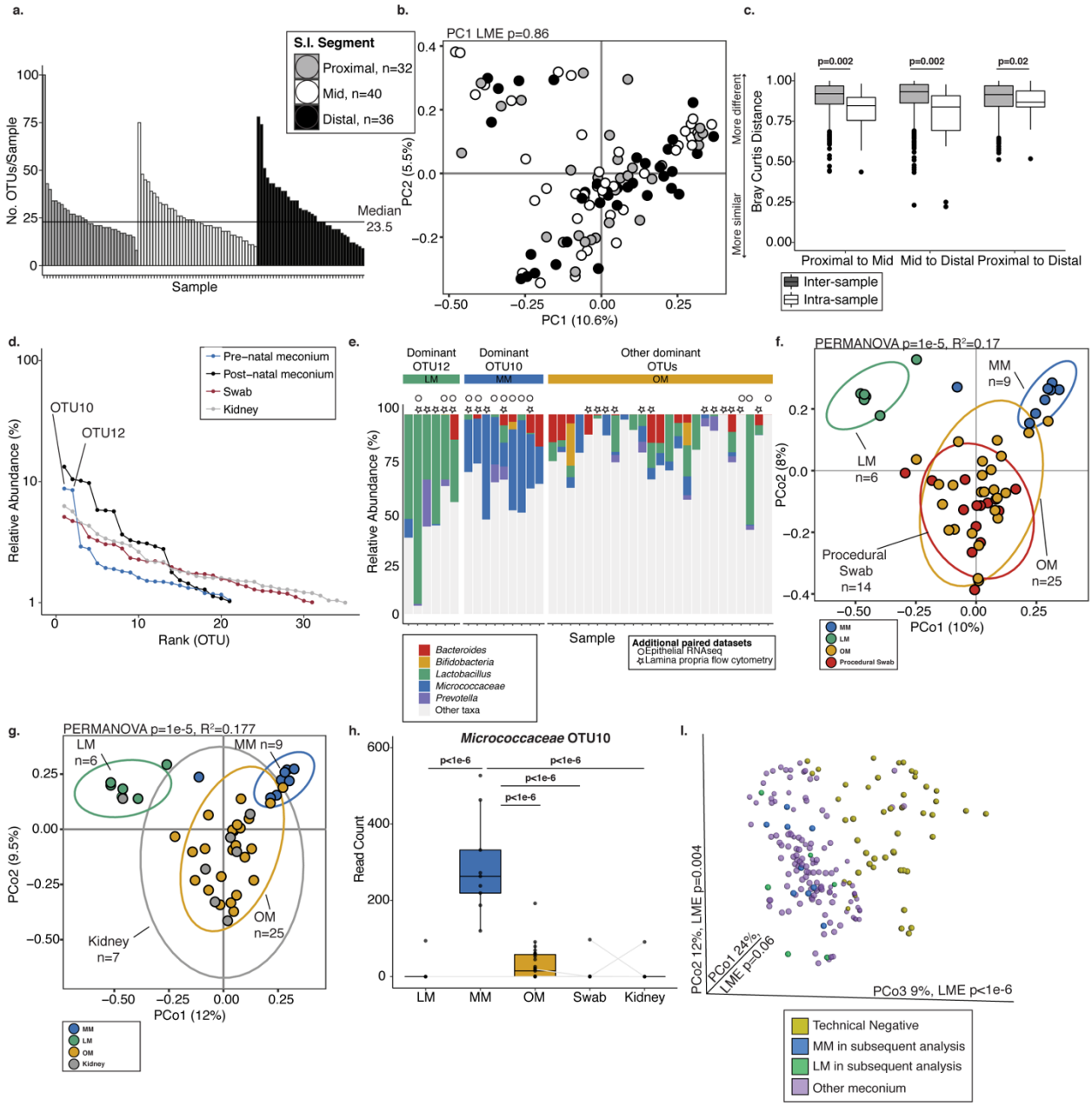

Figure 2.5 Sparse bacterial signal distinct from background detected in fetal meconium a. Histogram of number of operational taxonomic units (OTUs) per sample detected in fetal meconium from proximal-, mid-, or distal-segments of the small intestine after technical control filtering. **b.** Principal coordinates analysis (PCoA) of Bray Curtis distances of rareified bacterial profiles of proximal- (n=32), mid- (n=40), and distal- (n=36) sections of the intestine. **c.** Inter- and intra-sample Bray Curtis distances between indicated comparisons of proximal- (n=32), mid- (n=40), and distal- (n=36) intestinal sections. For a-b, n represents biologically independent fetal samples across the length of the intestine. **d.** Bacterial abundance ranks in fetal meconium, post-natal meconium, procedural swab, and kidney control. **e.** Relative abundance of select genera among samples dominated by OTU12, OTU10, or other OTUs. Symbols

indicate samples with paired immunological datasets. PCoA of Bray Curtis distances of Lactobacillus-meconium (LM, n=6), *Micrococcaceae*-meconium (MM, n=9), or Other-meconium (OM, n=25) compared to **f.** procedural swab (n=14) or **g.** fetal kidney (n=7) control, where n represents biologically independent fetal samples. **h.** Normalized read counts for *Micrococcocaceae* OTU10 in LM (n=6), MM (n=9), OM (n=25), swab (n=14), and fetal kidney (n=7) control samples, where n represents biologically independent fetal samples; samples collected from the same fetus indicated by grey line, when possible. **i.** PCoA of Bray Curtis distances of unrareified and unfiltered bacterial profiles of meconium (n=138 biologically independent fetal samples across three segments of the intestine) with technical negative controls (n=48 biologically independent samples including extraction buffer, room air swab, pre-moistened swabs). LM (n=6) and MM (n=9) samples identified in later analyses are highlighted; significance was measured using two-sided Satterthwaite's method on linear mixed effects model to test for significance and correct for repeated measures in b,h; ANOVA of linear mixed effects model was used to test for significance and correct for repeated measures in i, two-sided t-test was used for c, PERMANOVA was used in f-g. Boxplots indicate the median (center), the 25th and 75th percentiles, and the smallest and largest values within 1.5× the interquartile range (whiskers).

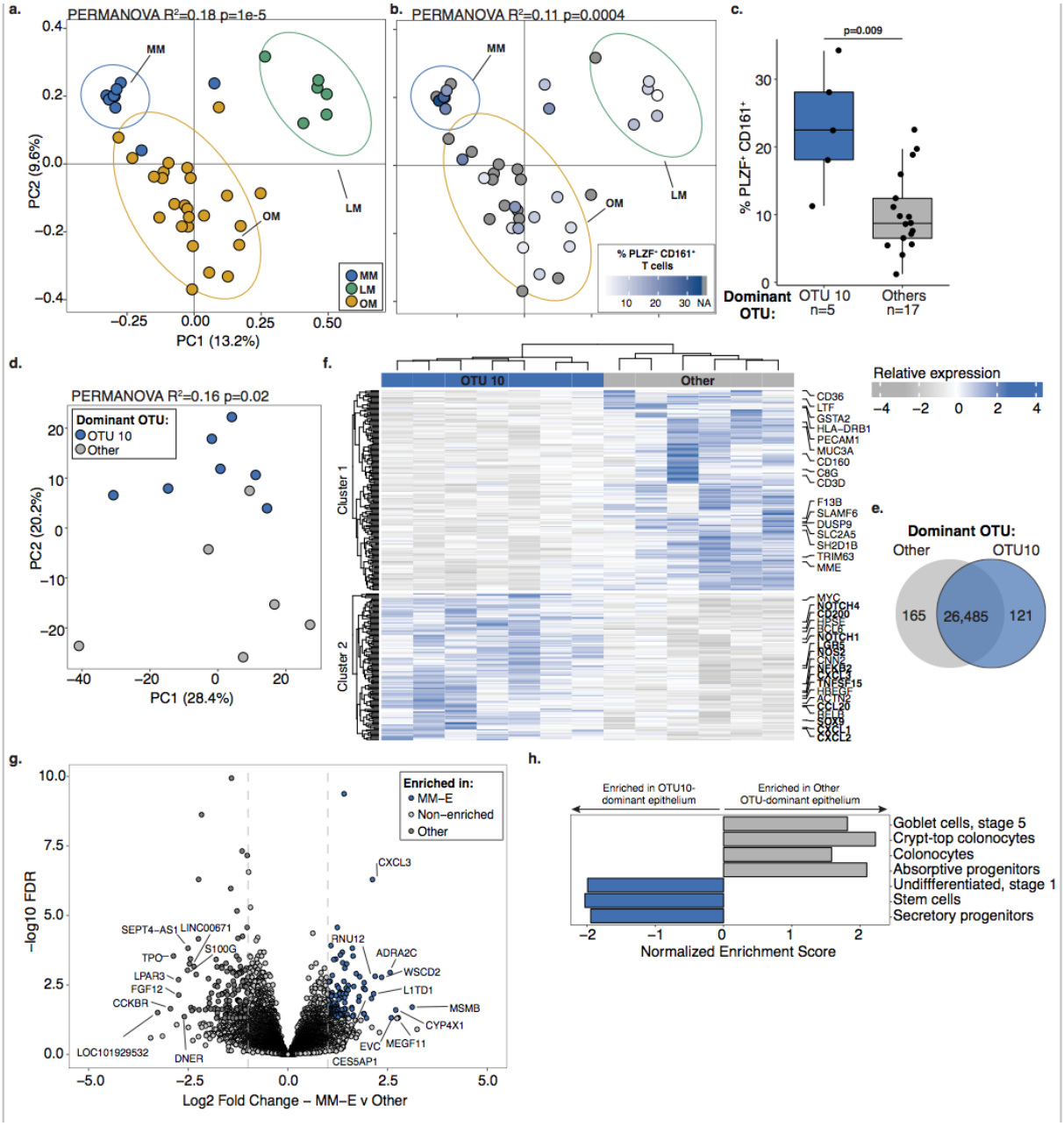

Figure 2.6 Divergent immune cell phenotypes are associated with *Micrococcaceae* relative enrichment in fetal meconium PCoA of Bray Curtis distances of 16S rRNA profiles colored by **a.** meconium dominated by *Lactobacillus* OTU12 (LM, n=6), *Micrococcaceae* OTU10 (MM, n=9), or other taxa (OM, n=25) or **b.** the proportion of PLZF$^+$ CD161$^+$ T cells among live, TCRβ^+, Vα7.2$^-$, CD4$^+$ cells in intestinal lamina propria (LP) paired with LM (n=5), MM (n=5) or OM (n=12) samples indicated by ellipses at 95% confidence. **c.** Proportion of PLZF$^+$ CD161$^+$ T cells of live, CD4$^+$ TCRβ^+ Vα7.2$^-$ cells in LP among samples associated with meconium dominated by OTU10 (MM, n=5) or other OTUs (n=17). **d.** Principal components (PC) analysis of Euclidean distances of top 10000 variable genes (by coefficient of variation) in OTU10-dominated meconium associated epithelium (OTU10, MM-E, n=7) or other OTU-

dominated meconium associated epithelium (Other, n=6) as determined by RNA sequencing. **e.** Venn diagram, **f.** heatmap with labeled immune pathway transcripts, and **g.** volcano plot of top differentially expressed genes between MM-E (n=7, log2 fold change >1, FDR < 0.05) and other OTU-dominated meconium associated epithelium (n=6, log2 fold change<1, FDR < 0.05). **h.** Normalized enrichment scores of gene set enrichment analysis of transcripts associated with epithelial cell states. For a-g, n indicates biologically independent fetal samples. PERMANOVA test for significance for a-b, d. Two-sided Wilcoxon rank sum test was used for c. DESEQ2 was used to calculate significant genes using a two-sided false discovery rate and log2 fold change. Each dot represents one independent biological replicate in a-d and one transcript in g. Boxplot indicates the median (center), the 25th and 75th percentiles, and the smallest and largest values within 1.5× the interquartile range (whiskers).

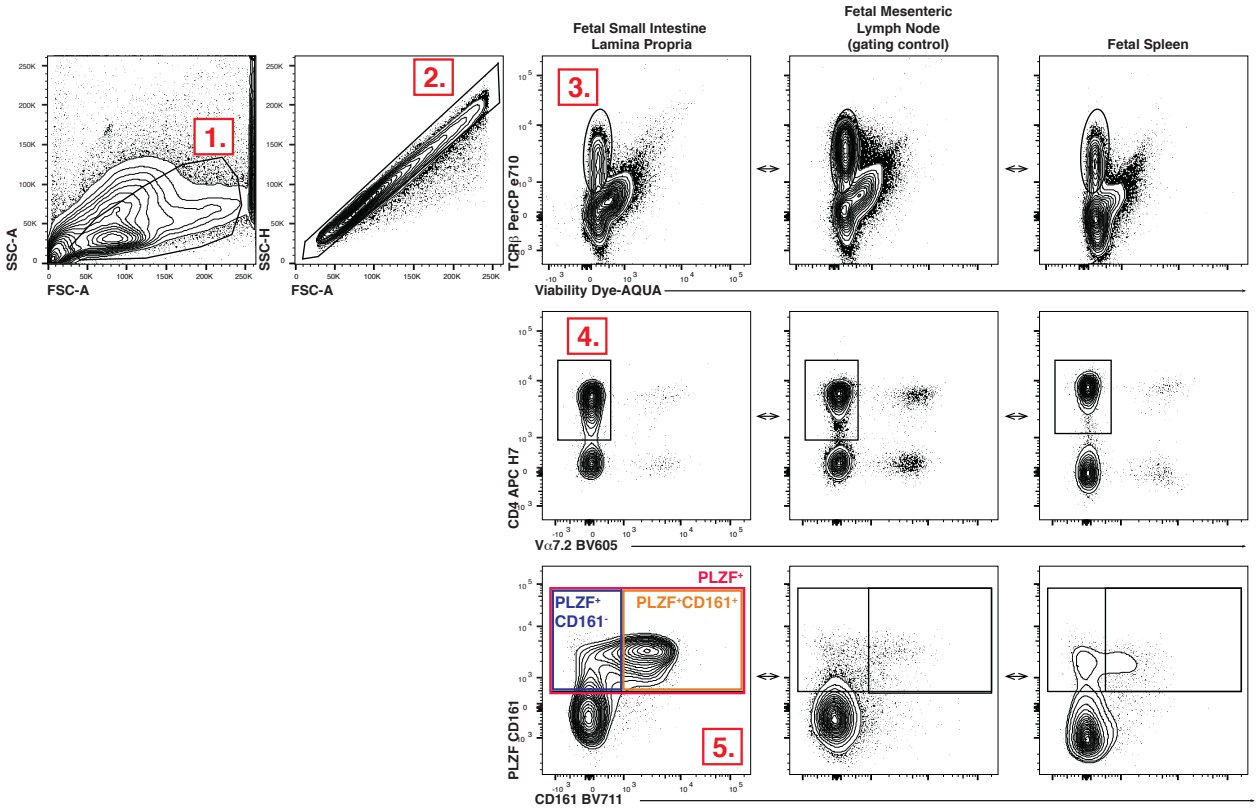

Figure 2.7 Gating strategy for T cell profile assessment Gating strategy for identification of PLZF^{+} CD161^{+} CD4^{+} αβT cells. Cells were gated on 1- lymphocytes, 2- singlets, 3- live cells expressing TCRβ, 4- CD4 expressing cells that were excluded of the dominant invariant chain expressed on mucosa-associated invariant T cells, Vα7.2. 5^{-} PLZF^{+}, PLZF^{+} CD161^{+} or PLZF^{+} CD161^{-} cells. All gating was set on mesenteric lymph node (MLN) internal controls and when available, splenic internal controls (SPL).

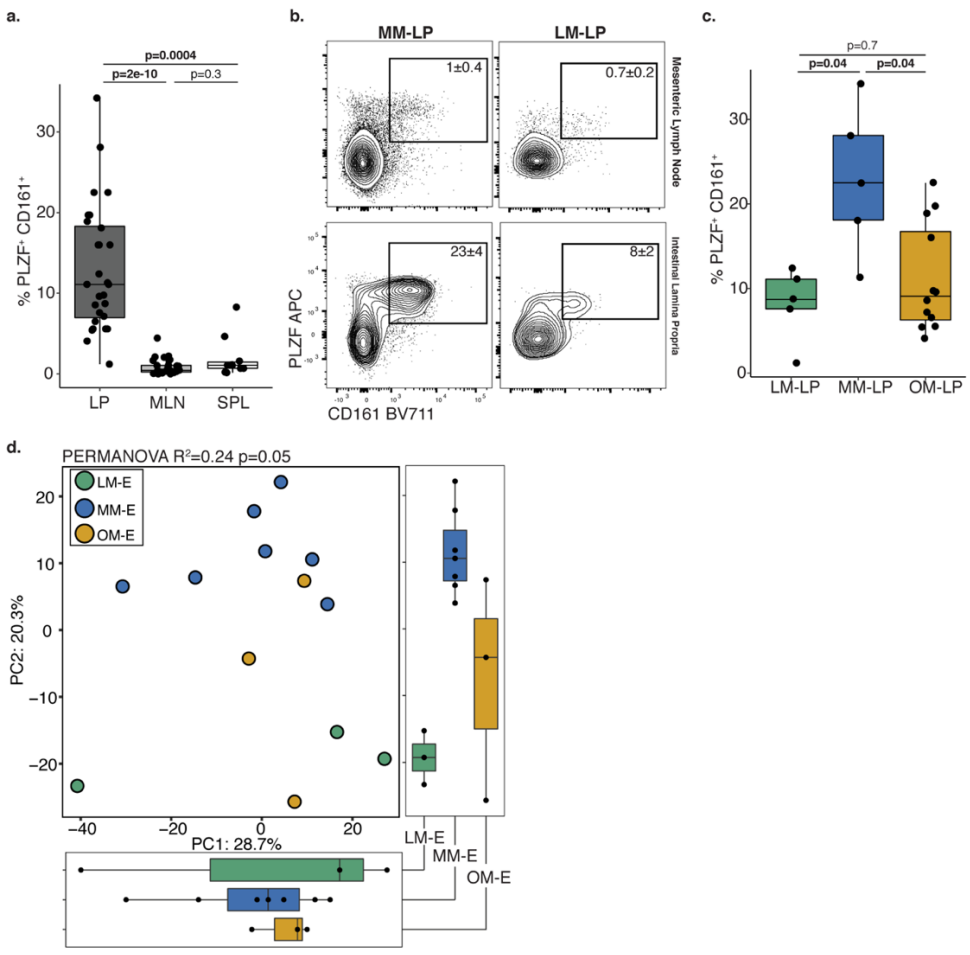

Figure 2.8 Divergent epithelial transcriptome and lamina propria T cells in samples associated with LM, MM, or OM a. Proportion of PLZF$^+$ CD161$^+$ T cells among live, TCRβ^+, Vα7.2$^-$, CD4$^+$ cells in intestinal lamina propria (LP, n=28), mesenteric lymph node (MLN, n=27), and spleen (SPL, n=10) where n is a biologically independent fetal sample. **b.** Representative flow plots of mesenteric lymph node (top panel, gating control) or intestinal lamina propria (bottom panel) associated with MM and LM. Experiments were repeated 5 independent times for each LM and MM associated samples with similar results. **c.** Proportion of PLZF$^+$ CD161$^+$ T cells in intestinal lamina propria paired with LM, MM, or OM (LM-LP, n=5; MM-LP, n=5; OM-LP, n=12) among live, TCRβ^+, Vα7.2$^-$, CD4$^+$ cells. **d.** Principal components (PC) analysis of Euclidean distances of top 10000 variable genes (by coefficient of variation) in LM associated epithelium (LM-E, n=3), MM associated epithelium (MM-E, n=7), or OM associated epithelium (OM-E, n=3) as determined by RNA sequencing where n represents a biologically independent fetal sample. Kruskal-Wallis ANOVA, with Dunnet's correction for multiple comparisons was used for a, c. PERMANOVA test for significance in d. Each dot represents a biological replicate. Boxplots indicate the median (center), the 25th and 75th percentiles, and the smallest and largest values within 1.5× the interquartile range (whiskers).

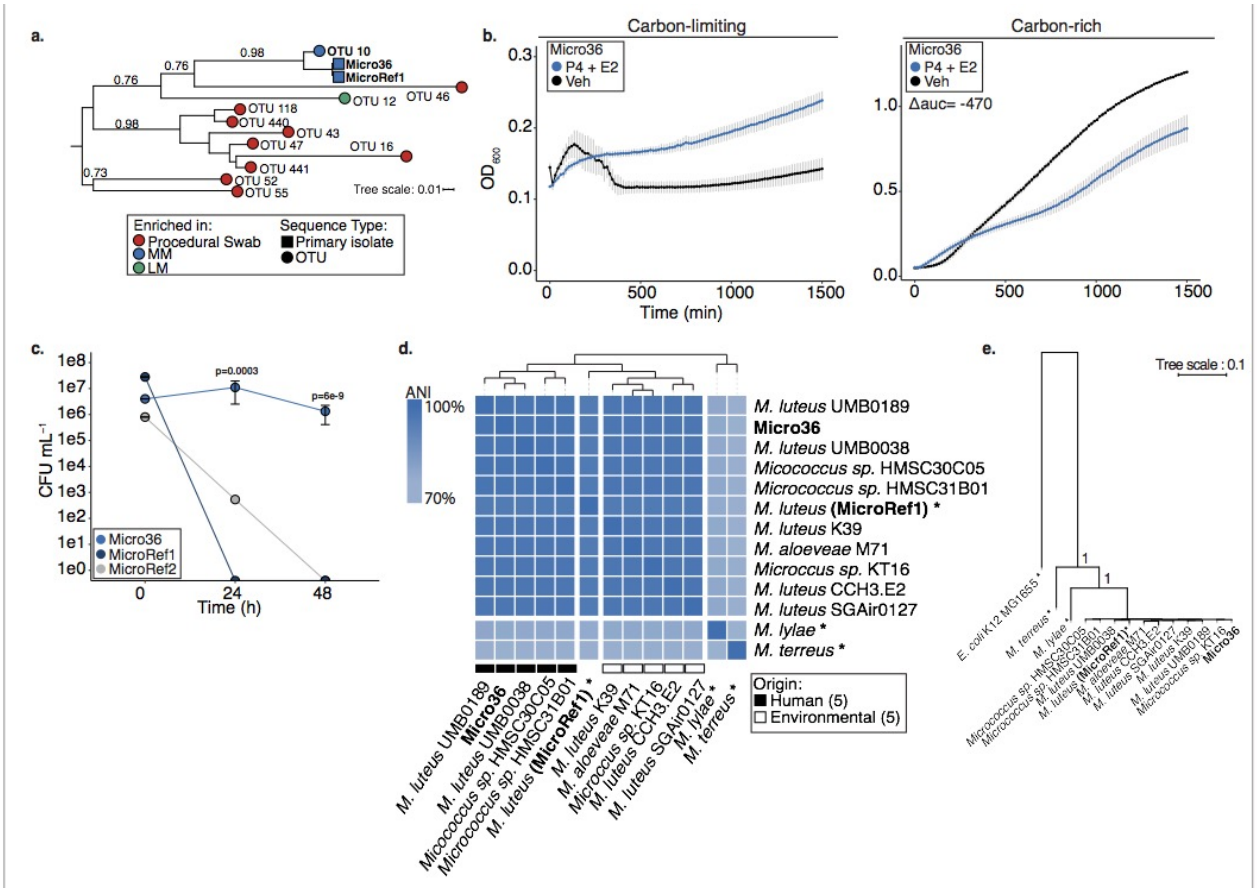

Figure 2.9 *Micrococcus* isolate from fetal meconium exhibits adaptation to the fetal environment a. Phylogenetic tree of 16S V4 rRNA gene sequences from *Lactobacillus*-enriched meconium (LM; green), *Micrococcaceae*-enriched meconium (MM; blue), or procedural swab (red) enriched OTUs (circles) and primary isolate (square) from fetal meconium (Micro36 and reference strain for *Micrococcus luteus* (MicroRef1). **b.** Effects of 10^{-5} M progesterone (P4) and 10^{-6} M β-Estradiol (E2) on the growth of Micro36 compared to ethanol vehicle control in carbon-limiting media (mineral salt media, left) or carbon-rich media (brain heart infusion; right) at 37 °C. Representative growth curves of three independent experiments measured by optical density at 600nm (OD600), error bars denote standard error of the mean (SEM) from center mean between three technical experiments. For carbon-rich media conditions, integral of logistic regression model fitting was used to calculate area under the curve (auc) and difference relative to vehicle control is reported as Δauc. Intracellular survival of **c.** Micro36, MicroRef1, MicroRef2 in primary human antigen presenting cells isolated from the fetal intestine. Representative data of three biologically independent fetal specimens, error bars indicate SEM from center mean of three cell culture replicates. ANOVA of a generalized linear model of log(CFU+1) against MicroRef1 for each timepoint was used to calculate significance. **d.** Whole genome average nucleotide identity (ANI) of all available genomes in *Micrococcus* and Micro36 isolate. When available strain origin is represented, hierarchical clustering was performed on average nucleotide identity, asterisk (*) indicates a reference or a representative genome for the taxon. **e**. Phylogenetic tree of conserved single-copy genes across all publicly available

genomes within *Micrococcus* and fetal meconium isolate Micro36 with *E. coli* K12 MG1655 outgroup. For a and e, branch lengths are scaled to the mean number of nucleotide substitutions per site and bootstrap values are represented for relevant nodes.

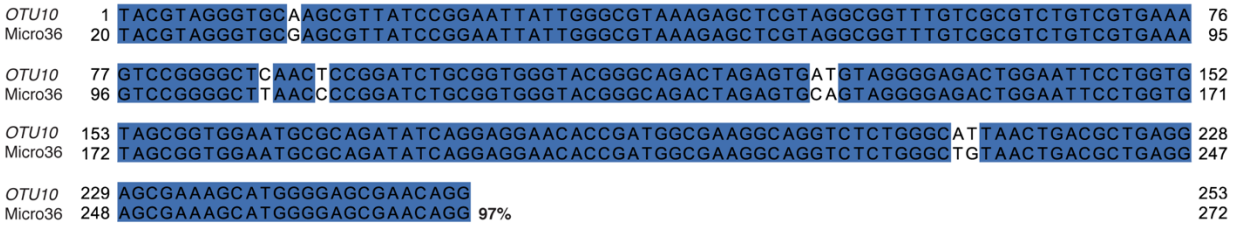

Figure 2.10 *Micrococcus* fetal isolate exhibits high 16S rRNA V4 sequence identity to fetal meconium OTUs Alignment of 16S V4 rRNA gene sequences of Micro36 to OTU10. Percentage indicates identity to representative OTU sequence.

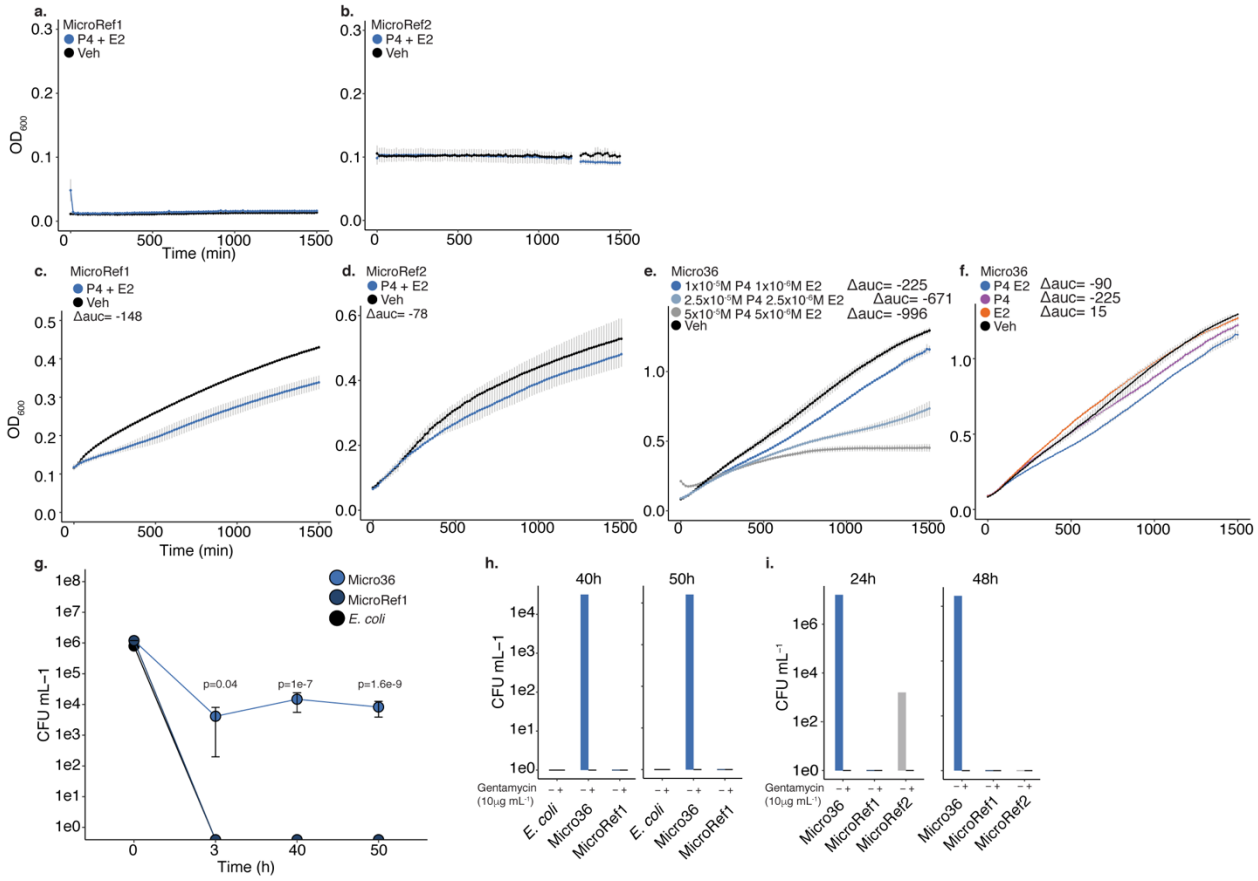

Figure 2.11 Fetal meconium *Micrococcus* isolate exhibits adaptation to the fetal environment Effects of 10^{-5} M progesterone (P4) and 10^{-6} M β-Estradiol (E2) on the growth **a.** MicroRef1 or **b.** MicroRef2 in carbon limiting media or of **c.** MicroRef1, **d.** MicroRef2, **e**. Micro36 with indicated concentrations of P4 and E2 or **f.** combinations of hormones compared to ethanol vehicle control, in carbon-rich media at 37 °C. Representative growth curves of three independent experiments measured by optical density at 600nm (OD600), error bars denote standard error of the mean (SEM) between three technical experiments. For carbon-rich media conditions, integral of logistic regression model fitting was used to calculate area under the curve (auc) and change with respect to vehicle control is reported as Δauc. **g.** Intracellular survival of Micro36 or MicroRef1 or *E. coli* in RAW264.3 cells. ANOVA of generalized linear model of log(CFU+1) against *E. coli* for each timepoint was used to calculate significance. Error bars indicate SEM around center mean of n=3 independent cell culture experiments. Growth of indicated strains on media with (+) or without (-) gentamycin (10μg mL^{-1}) following 24-50 hours of intracellular growth in **h.** RAW264.7 cells or **i.** primary human fetal intestinal antigen presenting cells.

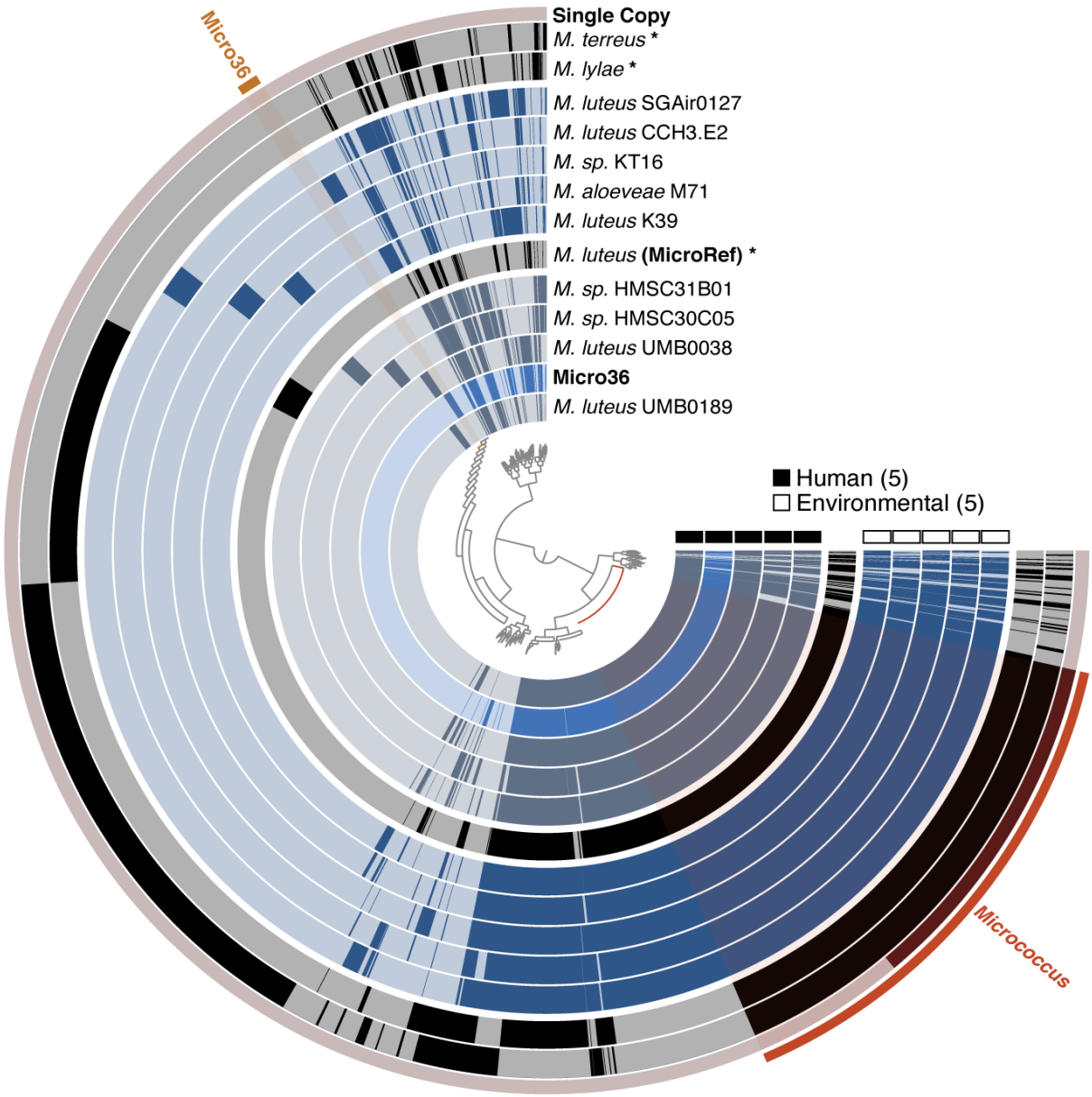

Figure 2.12 Genomic features of fetal *Micrococcus* isolate Alignment of all publicly available *Micrococcus* genomes; single copy *Micrococcus* genes used for phylogeny (inset) and genes unique to Micro36 isolate are highlighted. Figure was generated using the *Anvi'o* package; each radial layer represents a genome; representative or reference genomes are colored in black indicated with asterisk; inner dendrogram represents hierarchical clustering of amino acid sequences based on their sequence composition and distribution across genomes; genomes are organized based on gene clusters they share using Euclidian distance and Ward ordination; outer ring represents single copy genes predicted using hidden markov model in *Anvi'o* package.

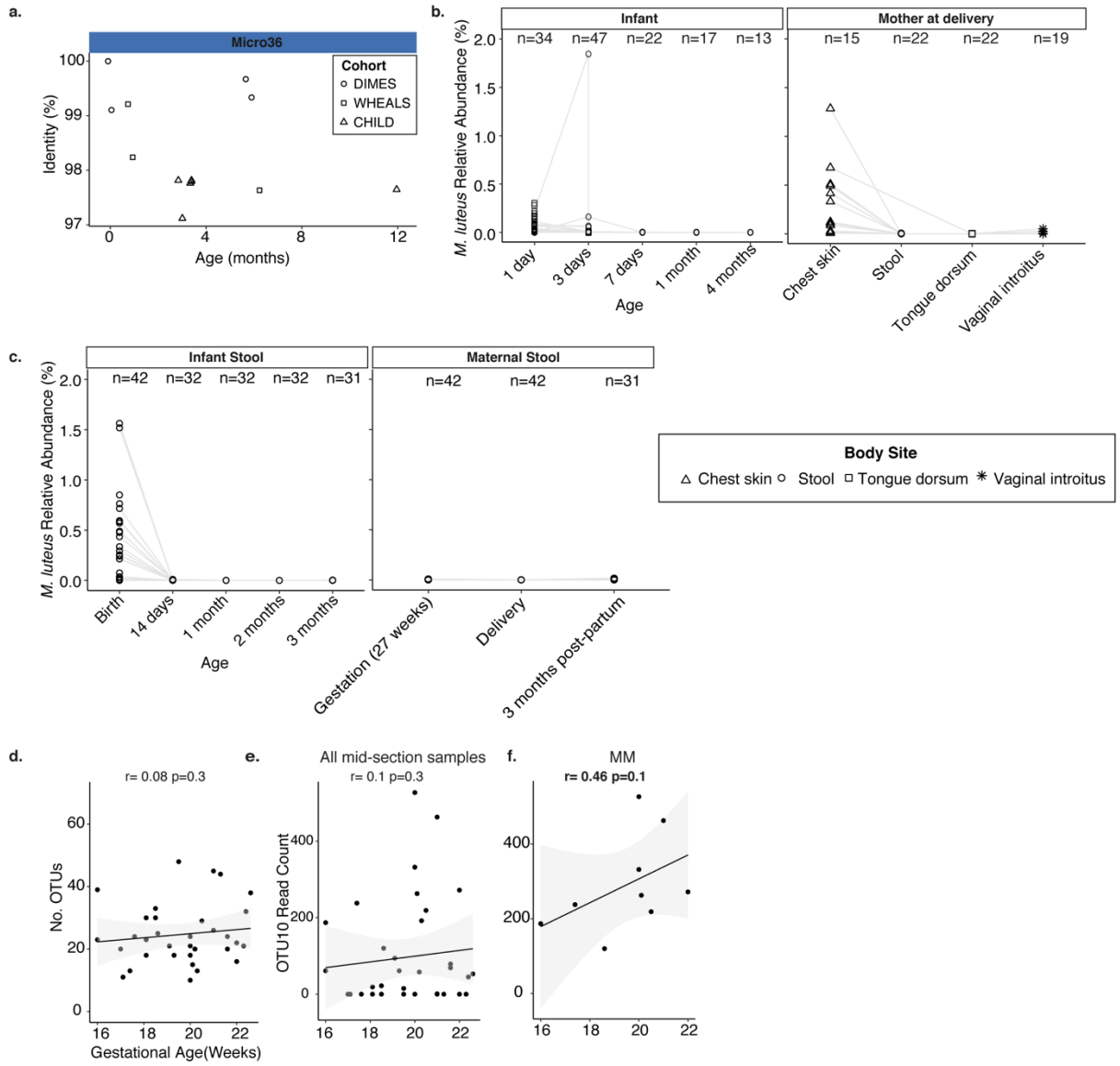

Figure 2.13 Prevalence of M. luteus in infants and mothers a. Percent identity of samples to 16S rRNA gene of Micro36 in three independent infant stool cohorts. Each symbol represents a sample with a positive hit (>97% sequence identity); symbol shape indicates cohort. Relative abundance of *Micrococcus luteus* in metagenomic sequencing cohorts across **b.** body sites at delivery in mother and infant within four months after birth, or **c.** in maternal stool around delivery and infant stool within the first three months of life. Metagenomic sequences obtained from two independent studies were classified using a custom kraken2 database including the fetal *M. luteus* Micro 36 genome. Correlation of gestational age with **d.** total number of OTUs or **e.** *Micrococcaceae* OTU10 count in mid-section meconium samples (n= 35 biologically independent fetal specimens) or **f.** among *Micrococaceae* meconium (MM, n=9 biologically independent fetal specimens). Pearson's product-moment correlation coefficient and a one-sided t-distribution p-value is reported for d-f.

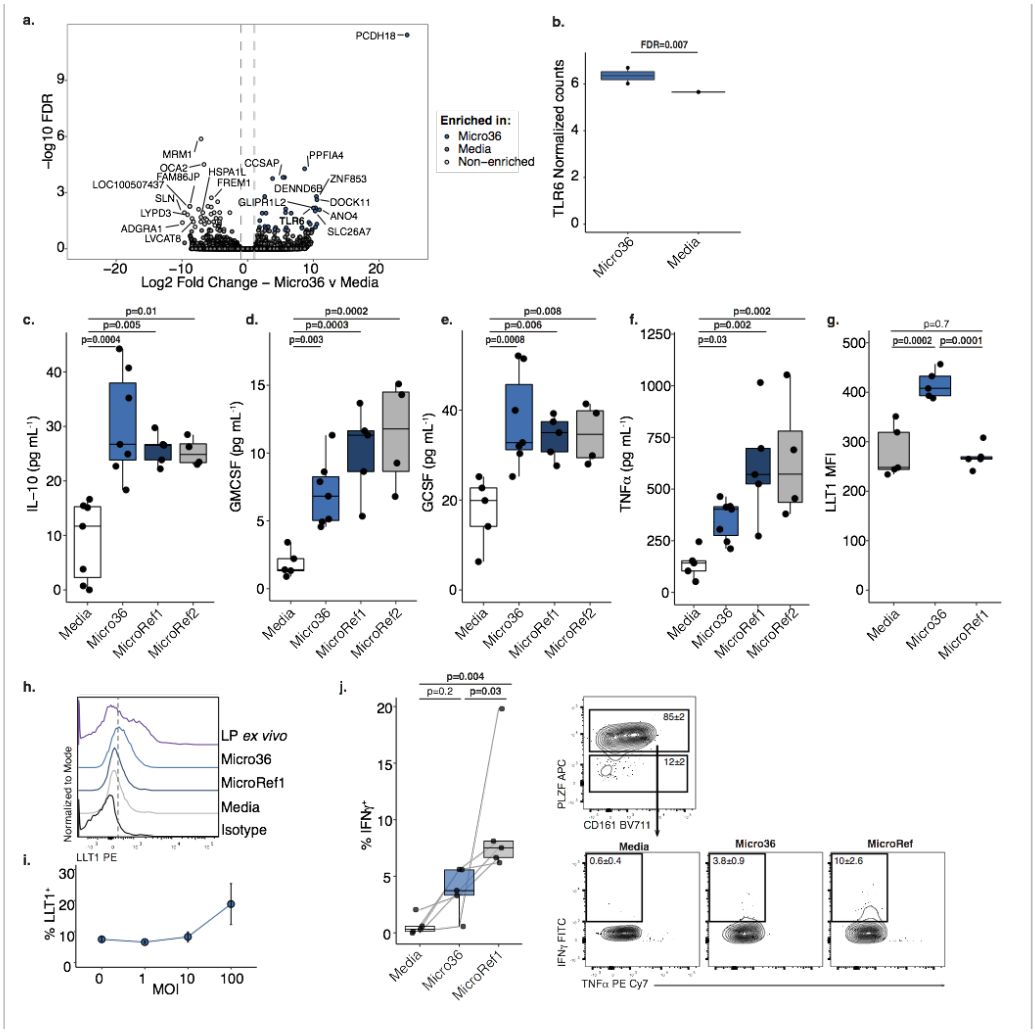

Figure 2.14 Fetal *Micrococcus* isolate promotes immunotolerance phenotypes *in vitro* a. Volcano plot of significantly (false discovery rate, FDR, <0.05) and differentially (Log2FoldChange |1|) expressed genes and **b.** normalized read counts of TLR6 in primary human fetal intestinal epithelial cells after Micro36 treatment versus media control by RNAseq. For a-b, FDR-adjusted p-values were calculated using two-sided DESEQ2 algorithm, n=2 biologically independent fetal samples. Concentrations of **c.** IL-10, **d.** GM-CSF, **e.** G-CSF, or **f.** TNFα in supernatants of fetal splenic antigen presenting cells following four hours of exposure to media (n=7) or *Micrococcus* (Micro36 n=7, MicroRef1 n=5, MicroRef2 n=4) strains, where n represents biologically independent fetal specimens for the indicated treatment. **g.** Mean fluorescence intensity (MFI) and **h.** representative histograms of five experiments of LLT1 expression of live, lin$^-$, CD45$^+$, HLA-DR$^+$ splenocytes of n=5 biologically independent fetal specimens exposed to media, Micro36, MicroRef1 or unstimulated lamina propria (LP) antigen presenting cells. **i.** Multiplicity of infection (MOI) of Micro36 relative to proportion of LLT1$^+$ live, lin$^-$, CD45$^+$, HLA-DR$^+$ splenocytes, each dot (center) represents mean of n=3 biologically independent fetal specimens and error bars indicate standard error of the mean. **j.** Intracellular INFγ production among pure intestinal effector memory T cells in mixed lymphocyte reactions with sorted, autologous lin$^-$, CD45$^+$, HLA-DR$^+$ antigen

presenting cells that were pre-exposed to media or *Micrococcus* (Micro36, MicroRef1) strains. Left, Percent IFNγ^+ T cells among live, TCRβ^+, CD4$^+$, Vα7.2$^-$, PLZF$^+$, n=5 biologically independent fetal specimens. Right, Representative flow plots of five experiments of sorted effector memory T cells T cells (top) and intracellular cytokine, IFNγ and TNFα, expression (bottom); numbers indicate mean proportion and standard error of the mean (SEM). Lines connecting dots indicate same fetal specimen across treatments. Two-sided Satterthwaite's method on linear mixed effects model was used to test for significance, controlling for repeated measures of cell donor, for c-g, j; Positive LME residuals are plotted for c-f. Each dot represents an independent fetal sample. Boxplots indicate the median (center), the 25th and 75th percentiles, and the smallest and largest values within 1.5× the interquartile range (whiskers).

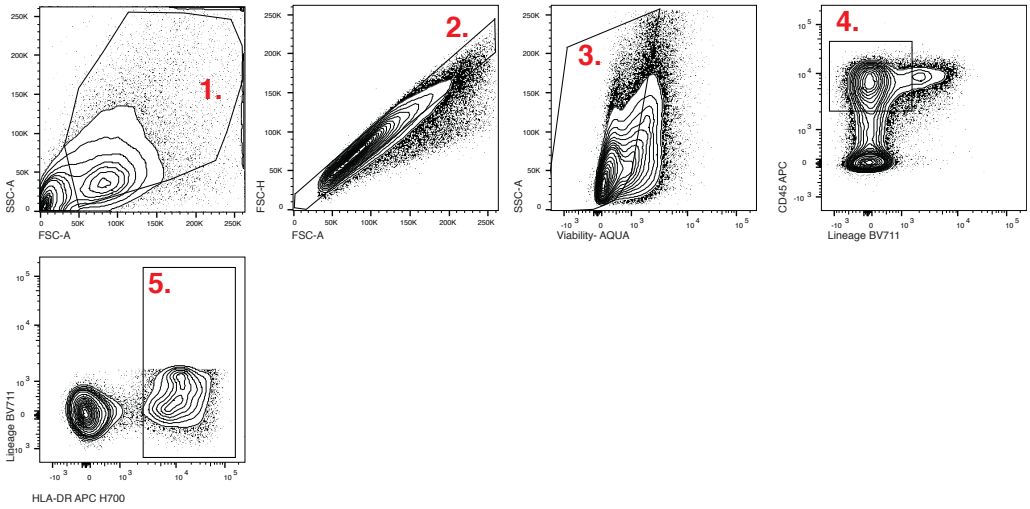

Figure 2.15 Gating strategy for antigen presenting cell phenotypes Gating strategy for identification of fetal splenic antigen presenting cells. Cells were gated on 1-lymphocytes, 2-singlets, 3-live, 4-lineage (CD3, CD56, CD20, CD19)- and CD45+, 5-HLA-DR+ cells.

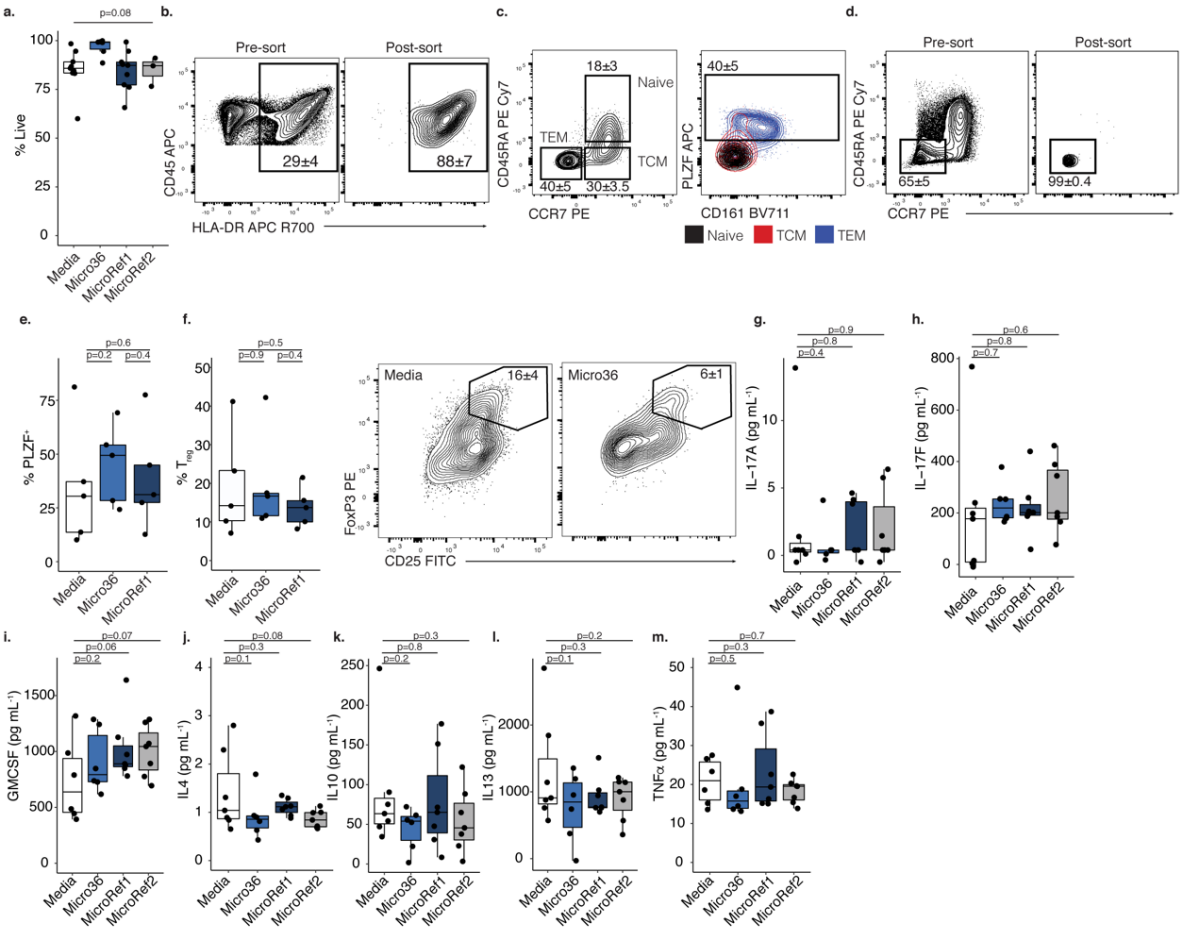

Figure 2.16 Fetal *Micrococcus* isolate promotes distinct APC and T cell phenotypes a. Proportion of live cells after treatment with media (n=9) or *Micrococcus* (Micro36 n=6, MicroRef1 n=9, MicroRef2 n=3) strains, where n represents biologically independent fetal specimens for the indicated treatment. ANOVA test for significance. **b.** HLA-DR^{+} CD45^{+} lin^{-} cells pre- (left) and post- (right) fluorescence activated cell sorting (FACS). **c.** Proportion of naïve (CD45RA^{+} CCR7^{+}), central memory (TCM, CD45RA^{-} CCR7^{+}), and effector memory T cells (TEM, CD45RA^{-} CCR7^{-}) among live, TCRβ^{+}, CD4^{+} cells (left panel) and PLZF and CD161 expression among memory subsets, numbers indicate proportion in TEM (right panel). **d.** Pre- (left) and post- (right) FACS of effector memory T cells. **e.** Proportion of PLZF^{+} T cells or **f.** left, proportion of CD25^{hi} FoxP3^{+} regulatory T cells (T_{regs}) and right, representative flow plots of FoxP3 and CD25 expression among intestinal live, TCRβ^{+}, CD4^{+}, Vα7.2^{-}, cells after exposure to splenic APCs pretreated with media or *Micrococcus* (Micro36, MicroRef1) strains for n=5 biologically independent fetal specimens. Concentration of **g.** IL-17A, **h.** IL-17F, **i.** GM-CSF, **j.** IL-4, **k.** IL-10, **l.** IL-13, **m.** TNFα in culture supernatants of lamina propria T cell co-cultures with splenic antigen presenting cells pre-exposed to media (n=7) or *Micrococcus* (Micro36 n=6, MicroRef1 n=7, MicroRef2 n=7) strains, where n represents biologically independent fetal specimens for the indicated treatment. For b-d, f numbers indicate mean proportion and standard error of the mean (SEM) representative of five independent experiments. For e-f, g-m two-sided Satterthwaite's method on linear

mixed effects model was used to test for significance between strains, controlling for repeated measures of cell donor. Positive LME residuals are plotted for g-m. Each dot represents an independent fetal sample, unless otherwise indicated. Boxplots indicate the median (center), the 25th and 75th percentiles, and the smallest and largest values within 1.5× the interquartile range (whiskers).

2.7 Tables

Table 2.1. Fetal Meconium Bank

Fetal meconium	*No. samples banked*	*No. samples successfully sequenced **	*Median 16S rRNA V4 Sequencing Read Count Per Sample prior to technical negative filtering*	*Median 16S rRNA V4 Sequencing Read Count Per Sample after technical negative filtering*	*Reads remaining after filtering (%)*	*No. samples with sequences remaining after technical negative filtering and rareifying*
Technical negative controls	**50**	**48**	**35,906**			
Extraction Buffer	12	12	12,957			
Room Air swab	19	19	39,651			
Pre-moistened swab	19	17	40,805			
Technical positive controls	**3**	**3**	**281,284**	**13,766**	**4.89**	**3**
Mock community	3	3	281,284	13,766	4.89	3
Procedural Environment controls	**37**	**35**	**23,300**	**1,822**	**7.82**	**21**
Procedural swab	19	17	27,534	2,847	10.34	14
Kidney	18	18	12,510	645	5.16	7
Fetal meconium	**150**	**107**	**30,508**	**4,424**	**14.50**	**108**
Proximal	50	47	33,391	3,301	9.89	32
Mid	50	46	30,247.5	4,786	15.82	40
Distal	50	45	29,826	4,488	15.05	36

Table 2.2 Filtered and unfiltered OTU tables with respect to technical negative controls: OTUs filtered

OTU ID	Kingdom	Phylum	Class	Order	Family	Genus
OTU_1	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Enterobacteriales*	*Enterobacteriaceae*	*Escherichia/Shigella*
OTU_1098	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	*Burkholderia-Caballeronia-Paraburkholderia*
OTU_1121	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Enterobacteriales*	*Enterobacteriaceae*	NA
OTU_1136	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Pseudomonadales*	*Pseudomonadaceae*	*Pseudomonas*
OTU_1167	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	NA
OTU_1174	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	NA
OTU_1182	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Pseudomonadales*	*Pseudomonadaceae*	NA
OTU_1199	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	NA
OTU_15	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Xanthomonadales*	*Xanthomonadaceae*	*Stenotrophomonas*
OTU_17	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Oceanospirillales*	*Halomonadaceae*	*Halomonas*
OTU_2	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	*Burkholderia-Caballeronia-Paraburkholderia*
OTU_24	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	*Ralstonia*
OTU_263	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Pseudomonadales*	*Pseudomonadaceae*	*Pseudomonas*
OTU_3	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Pseudomonadales*	*Pseudomonadaceae*	*Pseudomonas*
OTU_4	*Bacteria*	*Proteobacteria*	*Deltaproteobacteria*	*Myxococcales*	*Myxococcaceae*	*Myxococcus*
OTU_498	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	*Burkholderia-Caballeronia-Paraburkholderia*
OTU_5	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	*Delftia*
OTU_6	*Bacteria*	*Verrucomicrobia*	*Verrucomicrobiae*	*Verrucomicrobiales*	*Akkermansiaceae*	*Akkermansia*
OTU_63	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Propionibacteriales*	*Propionibacteriaceae*	*Cutibacterium*
OTU_7	*Bacteria*	*Firmicutes*	*Bacilli*	*Bacillales*	*Staphylococcaceae*	*Staphylococcus*
OTU_9	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	*Sphaerotilus*
OTU_933	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Pseudomonadales*	*Pseudomonadaceae*	*Pseudomonas*

Table 2.3. OTUs identified as contaminants using *decontam* package

OTU	Prevalence in Meconium	Prevalence in Technical Negative Controls	Kingdom	Phylum	Class	Order	Family	Genus
OTU_5	45	46	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	*Delftia*
OTU_1	45	46	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Enterobacteriales*	*Enterobacteriaceae*	*Escherichia/Shigella*
OTU_1136	45	46	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Pseudomonadales*	*Pseudomonadaceae*	*Pseudomonas*
OTU_3	45	46	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Pseudomonadales*	*Pseudomonadaceae*	*Pseudomonas*
OTU_263	44	42	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Pseudomonadales*	*Pseudomonadaceae*	*Pseudomonas*
OTU_1174	38	42	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	*NA*
OTU_2	36	42	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	*Burkholderia-Caballeronia-Paraburkholderia*
OTU_4	35	32	*Bacteria*	*Proteobacteria*	*Deltaproteobacteria*	*Myxococcales*	*Myxococcaceae*	*Myxococcus*
OTU_1167	30	34	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	*NA*

OTU	Prevalence in Meconium	Prevalence in Technical Negative Controls	Kingdom	Phylum	Class	Order	Family	Genus
OTU_6	30	33	*Bacteria*	*Verrucomicrobia*	*Verrucomicrobiae*	*Verrucomicrobiales*	*Akkermansiaceae*	*Akkermansia*
OTU_1182	29	32	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Pseudomonadales*	*Pseudomonadaceae*	*NA*
OTU_15	27	30	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Xanthomonadales*	*Xanthomonadaceae*	*Stenotrophomonas*
OTU_24	26	30	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	*Ralstonia*
OTU_933	26	34	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*NA*	*NA*	*NA*
OTU_8	24	23	*Bacteria*	*Firmicutes*	*Bacilli*	*Lactobacillales*	*Streptococcaceae*	*Streptococcus*
OTU_13	20	19	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Corynebacteriales*	*Corynebacteriaceae*	*Lawsonella*
OTU_495	19	20	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	*Limnohabitans*
OTU_1199	17	29	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	*NA*
OTU_63	17	26	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Propionibacteriales*	*Propionibacteriaceae*	*Cutibacterium*
OTU_17	15	24	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Oceanospirillales*	*Halomonadaceae*	*Halomonas*
OTU_19	15	14	*Bacteria*	*NA*	*NA*	*NA*	*NA*	*NA*

OTU	Prevalence in Meconium	Prevalence in Technical Negative Controls	Kingdom	Phylum	Class	Order	Family	Genus
OTU_30	15	9	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Bacteroidaceae*	*Bacteroides*
OTU_37	15	14	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	*Comamonas*
OTU_16	14	17	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Pasteurellales*	*Pasteurellaceae*	*Haemophilus*
OTU_498	14	38	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	*Burkholderia-Caballeronia-Paraburkholderia*
OTU_1184	13	20	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Pseudomonadales*	*Pseudomonadaceae*	*Pseudomonas*
OTU_994	12	19	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Xanthomonadales*	*Xanthomonadaceae*	*Stenotrophomonas*
OTU_1120	11	7	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Enterobacteriales*	*Enterobacteriaceae*	*NA*
OTU_26	11	9	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Family_XI*	*Finegoldia*
OTU_27	11	7	*Bacteria*	*Firmicutes*	*Bacilli*	*Bacillales*	*Family_XI*	*Gemella*
OTU_29	11	10	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Bacteroidaceae*	*Bacteroides*
OTU_49	10	13	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	*Undibacterium*

OTU	Prevalence in Meconium	Prevalence in Technical Negative Controls	Kingdom	Phylum	Class	Order	Family	Genus
OTU_544	10	8	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Pasteurellales*	*Pasteurellaceae*	*Haemophilus*
OTU_842	10	9	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Xanthomonadales*	*Xanthomonadaceae*	*Stenotrophomonas*
OTU_39	9	8	*Bacteria*	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Rhizobiaceae*	*NA*
OTU_45	9	9	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*Blautia*
OTU_54	9	10	*Bacteria*	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Xanthobacteraceae*	*Bradyrhizobium*
OTU_1083	8	6	*Bacteria*	*Proteobacteria*	*Alphaproteobacteria*	*Sphingomonadales*	*Sphingomonadaceae*	*NA*
OTU_38	8	6	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Neisseriaceae*	*Neisseria*
OTU_429	8	4	*Bacteria*	*Firmicutes*	*Bacilli*	*Lactobacillales*	*NA*	*NA*
OTU_71	8	6	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Family_XI*	*Anaerococcus*
OTU_31	7	17	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Pseudomonadales*	*Moraxellaceae*	*Acinetobacter*
OTU_50	7	5	*Bacteria*	*Epsilonbacteraeota*	*Campylobacteria*	*Campylobacterales*	*Helicobacteraceae*	*Helicobacter*
OTU_57	7	9	*Bacteria*	*Firmicutes*	*Bacilli*	*Lactobacillales*	*Lactobacillaceae*	*Lactobacillus*
OTU_1138	6	3	*Bacteria*	*Firmicutes*	*Bacilli*	*Lactobacillales*	*NA*	*NA*

OTU	Prevalence in Meconium	Prevalence in Technical Negative Controls	Kingdom	Phylum	Class	Order	Family	Genus
OTU_974	6	4	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Bacteroidaceae*	*Bacteroides*
OTU_100	5	4	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Marinifilaceae*	*Odoribacter*
OTU_115	5	4	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Pseudomonadales*	*Moraxellaceae*	*Enhydrobacter*
OTU_178	5	4	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Bacteroidaceae*	*Bacteroides*
OTU_47	5	9	*Bacteria*	*Proteobacteria*	*Alphaproteobacteria*	*Sphingomonadales*	*Sphingomonadaceae*	*Novosphingobium*
OTU_61	5	4	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Family_XI*	*Ezakiella*
OTU_64	5	5	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Actinomycetales*	*Actinomycetaceae*	*Actinomyces*
OTU_78	5	7	*Bacteria*	*Firmicutes*	*Negativicutes*	*Selenomonadales*	*Veillonellaceae*	*Veillonella*
OTU_80	5	6	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Bacteroidaceae*	*Bacteroides*
OTU_89	5	4	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Muribaculaceae*	*NA*
OTU_93	5	7	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*Agathobacter*
OTU_1233	4	3	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*NA*
OTU_169	4	10	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Actinomycetales*	*Actinomycetaceae*	*Actinomyces*
OTU_34	4	12	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Micrococcales*	*Micrococcaceae*	*Rothia*

OTU	Prevalence in Meconium	Prevalence in Technical Negative Controls	Kingdom	Phylum	Class	Order	Family	Genus
OTU_459	4	5	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Xanthomonadales*	*Xanthomonadaceae*	*Stenotrophomonas*
OTU_785	4	2	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*ASF356*
OTU_82	4	4	*Bacteria*	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Beijerinckiaceae*	*Methylobacterium*
OTU_1036	3	5	*Bacteria*	*Proteobacteria*	*Alphaproteobacteria*	*Sphingomonadales*	*Sphingomonadaceae*	*Sphingomonas*
OTU_1052	3	10	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Pseudomonadales*	*Moraxellaceae*	*Acinetobacter*
OTU_110	3	13	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	*NA*
OTU_1123	3	5	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Bifidobacteriales*	*Bifidobacteriaceae*	*NA*
OTU_121	3	4	*Bacteria*	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Hyphomicrobiaceae*	*NA*
OTU_1222	3	4	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Pasteurellales*	*Pasteurellaceae*	*NA*
OTU_1223	3	3	*Bacteria*	*Proteobacteria*	*Deltaproteobacteria*	*Myxococcales*	*Myxococcaceae*	*NA*
OTU_131	3	3	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Streptomycetales*	*Streptomycetaceae*	*Streptomyces*
OTU_132	3	5	*Bacteria*	*Deinococcus-Thermus*	*Deinococci*	*Thermales*	*Thermaceae*	*Thermus*

OTU	Prevalence in Meconium	Prevalence in Technical Negative Controls	Kingdom	Phylum	Class	Order	Family	Genus
OTU_134	3	3	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Prevotellaceae*	*Prevotella_9*
OTU_152	3	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*Shuttleworthia*
OTU_154	3	2	*Bacteria*	*Proteobacteria*	*Alphaproteobacteria*	*Rhodobacterales*	*Rhodobacteraceae*	*NA*
OTU_197	3	4	*Bacteria*	*Firmicutes*	*Bacilli*	*Lactobacillales*	*Aerococcaceae*	*Abiotrophia*
OTU_390	3	2	*Archaea*	*Nanoarchaeaeota*	*Woesearchaeia*	*NA*	*NA*	*NA*
OTU_56	3	5	*Bacteria*	*Firmicutes*	*Bacilli*	*Lactobacillales*	*Carnobacteriaceae*	*Granulicatella*
OTU_659	3	2	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*Faecalibacterium*
OTU_73	3	5	*Bacteria*	*Proteobacteria*	*Alphaproteobacteria*	*Caulobacterales*	*Caulobacteraceae*	*Brevundimonas*
OTU_74	3	6	*Bacteria*	*Firmicutes*	*Negativicutes*	*Selenomonadales*	*Veillonellaceae*	*Dialister*
OTU_75	3	2	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Flavobacteriales*	*Weeksellaceae*	*Chryseobacterium*
OTU_77	3	7	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Flavobacteriales*	*Weeksellaceae*	*Chryseobacterium*
OTU_94	3	3	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*Ruminococcaceae_UCG-014*

OTU	Prevalence in Meconium	Prevalence in Technical Negative Controls	Kingdom	Phylum	Class	Order	Family	Genus
OTU_96	3	5	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Marinifilaceae*	*Odoribacter*
OTU_97	3	3	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*NA*
OTU_98	3	3	*Bacteria*	*Epsilonbacteraeota*	*Campylobacteria*	*Campylobacterales*	*Helicobacteraceae*	*Helicobacter*
OTU_1033	2	1	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Corynebacteriales*	*Corynebacteriaceae*	*Corynebacterium*
OTU_105	2	3	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*Ruminococcus_2*
OTU_1098	2	31	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	*Burkholderia-Caballeronia-Paraburkholderia*
OTU_1110	2	3	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*Blautia*
OTU_1151	2	3	*Bacteria*	*Firmicutes*	*Bacilli*	*Lactobacillales*	*Streptococcaceae*	*Streptococcus*
OTU_1163	2	5	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*NA*
OTU_117	2	11	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	*NA*

OTU	Prevalence in Meconium	Prevalence in Technical Negative Controls	Kingdom	Phylum	Class	Order	Family	Genus
OTU_1179	2	3	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Pseudomonadales*	*Moraxellaceae*	*Acinetobacter*
OTU_1224	2	1	*Bacteria*	*Firmicutes*	*Bacilli*	*Lactobacillales*	*NA*	*NA*
OTU_138	2	4	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Xanthomonadales*	*Xanthomonadaceae*	*NA*
OTU_139	2	4	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Hydrogenophilaceae*	*Hydrogenophilus*
OTU_144	2	1	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Rikenellaceae*	*Rikenella*
OTU_147	2	1	*Bacteria*	*Firmicutes*	*Bacilli*	*Lactobacillales*	*Streptococcaceae*	*Streptococcus*
OTU_148	2	3	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Peptostreptococcaceae*	*Romboutsia*
OTU_150	2	2	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Prevotellaceae*	*Alloprevotella*
OTU_153	2	2	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Micrococcales*	*Micrococcaceae*	*Rothia*
OTU_158	2	2	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Clostridiaceae_1*	*Clostridium_sensu_stricto_1*
OTU_161	2	5	*Bacteria*	*Proteobacteria*	*Alphaproteobacteria*	*Sphingomonadales*	*Sphingomonadaceae*	*Qipengyuania*
OTU_171	2	3	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Corynebacteriales*	*Corynebacteriaceae*	*Turicella*
OTU_173	2	2	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*Fastidiosipila*

OTU	Prevalence in Meconium	Prevalence in Technical Negative Controls	Kingdom	Phylum	Class	Order	Family	Genus
OTU_174	2	1	*Bacteria*	*Cyanobacteria*	*Oxyphotobacteria*	*Synechococcales*	*Cyanobiaceae*	*Cyanobium_PCC-6307*
OTU_175	2	5	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*Dorea*
OTU_182	2	2	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Family_XI*	*Anaerococcus*
OTU_202	2	2	*Bacteria*	*Firmicutes*	*Bacilli*	*Bacillales*	*Bacillaceae*	*NA*
OTU_211	2	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Family_XI*	*Anaerococcus*
OTU_228	2	2	*Bacteria*	*Proteobacteria*	*Alphaproteobacteria*	*Sphingomonadales*	*Sphingomonadaceae*	*NA*
OTU_238	2	6	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Bifidobacteriales*	*Bifidobacteriaceae*	*Bifidobacterium*
OTU_301	2	3	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Neisseriaceae*	*NA*
OTU_320	2	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*ASF356*
OTU_378	2	2	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Flavobacteriales*	*Flavobacteriaceae*	*Flavobacterium*
OTU_388	2	3	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Pseudomonadales*	*Moraxellaceae*	*Acinetobacter*
OTU_43	2	13	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	*Achromobacter*
OTU_520	2	2	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Micrococcales*	*Intrasporangiaceae*	*Janibacter*

OTU	Prevalence in Meconium	Prevalence in Technical Negative Controls	Kingdom	Phylum	Class	Order	Family	Genus
OTU_615	2	2	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*Ruminococcaceae_UCG-014*
OTU_69	2	3	*Bacteria*	*Actinobacteria*	*Coriobacteriia*	*Coriobacteriales*	*Coriobacteriaceae*	*Collinsella*
OTU_715	2	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*Lachnospiraceae_NK4A136_group*
OTU_747	2	4	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	*NA*
OTU_756	2	3	*Bacteria*	*Firmicutes*	*Bacilli*	*Lactobacillales*	*Lactobacillaceae*	*Lactobacillus*
OTU_91	2	3	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	*Massilia*
OTU_980	2	1	*Bacteria*	*Fusobacteria*	*Fusobacteriia*	*Fusobacteriales*	*Leptotrichiaceae*	*Sneathia*
OTU_165	1	3	*Bacteria*	*Epsilonbacteraeota*	*Campylobacteria*	*Campylobacterales*	*Helicobacteraceae*	*Helicobacter*
OTU_102	1	2	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Chitinophagales*	*Chitinophagaceae*	*NA*
OTU_104	1	4	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Bacteroidaceae*	*Bacteroides*
OTU_1118	1	2	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Frankiales*	*Geodermatophilaceae*	*Modestobacter*
OTU_1135	1	1	*Bacteria*	*Proteobacteria*	*Alphaproteobacteria*	*Rhodobacterales*	*Rhodobacteraceae*	*Paracoccus*

OTU	Prevalence in Meconium	Prevalence in Technical Negative Controls	Kingdom	Phylum	Class	Order	Family	Genus
OTU_1165	1	1	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Frankiales*	*Geodermatophilaceae*	*Geodermatophilus*
OTU_118	1	2	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*NA*
OTU_1189	1	2	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Micrococcales*	*Microbacteriaceae*	*NA*
OTU_1206	1	1	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Micrococcales*	*Microbacteriaceae*	*NA*
OTU_1210	1	1	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Enterobacteriales*	*Enterobacteriaceae*	*NA*
OTU_1221	1	1	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Bacteroidaceae*	*Bacteroides*
OTU_123	1	3	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Rikenellaceae*	*Alistipes*
OTU_1243	1	1	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Tannerellaceae*	*Parabacteroides*
OTU_129	1	2	*Bacteria*	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Rhizobiales_Incertae_Sedis*	*Alsobacter*
OTU_136	1	6	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Corynebacteriales*	*Mycobacteriaceae*	*Mycobacterium*
OTU_143	1	2	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Tannerellaceae*	*Parabacteroides*
OTU_156	1	2	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*NA*
OTU_162	1	2	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Micrococcales*	*Micrococcaceae*	*Kocuria*
OTU_168	1	1	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	*NA*

OTU	Prevalence in Meconium	Prevalence in Technical Negative Controls	Kingdom	Phylum	Class	Order	Family	Genus
OTU_184	1	3	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Methylophilaceae*	*NA*
OTU_188	1	1	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Prevotellaceae*	*Prevotella_6*
OTU_194	1	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*NA*
OTU_196	1	2	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Methylococcales*	*Methylomonaceae*	*Methylomonas*
OTU_198	1	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*Subdoligranulum*
OTU_199	1	2	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*NA*
OTU_203	1	1	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Flavobacteriales*	*Crocinitomicaceae*	*Fluviicola*
OTU_204	1	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*Lachnoclostridium*
OTU_206	1	1	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Pseudomonadales*	*Pseudomonadaceae*	*Pseudomonas*
OTU_209	1	3	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*NA*
OTU_213	1	1	*Bacteria*	*Firmicutes*	*Bacilli*	*Bacillales*	*Paenibacillaceae*	*Brevibacillus*
OTU_214	1	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*NA*
OTU_215	1	1	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Xanthomonadales*	*Xanthomonadaceae*	*Thermomonas*
OTU_217	1	2	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*Lachnospiraceae_N*

OTU	Prevalence in Meconium	Prevalence in Technical Negative Controls	Kingdom	Phylum	Class	Order	Family	Genus
								K4A136_group
OTU_218	1	2	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	*Cupriavidus*
OTU_220	1	2	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Christensenellaceae*	*Christensenellaceae_R-7_group*
OTU_233	1	2	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*Ruminiclostridium_6*
OTU_234	1	3	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Rikenellaceae*	*Alistipes*
OTU_237	1	2	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Xanthomonadales*	*Xanthomonadaceae*	*Arenimonas*
OTU_247	1	1	*Bacteria*	*Proteobacteria*	*Alphaproteobacteria*	*Caulobacterales*	*Caulobacteraceae*	*Asticcacaulis*
OTU_253	1	1	*Bacteria*	*Firmicutes*	*Erysipelotrichia*	*Erysipelotrichales*	*Erysipelotrichaceae*	*Erysipelotrichaceae_UCG-003*
OTU_264	1	1	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Prevotellaceae*	*Prevotella*
OTU_265	1	1	*Bacteria*	*Proteobacteria*	*Alphaproteobacteria*	*Reyranellales*	*Reyranellaceae*	*Reyranella*
OTU_266	1	2	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Corynebacteriales*	*Corynebacteriaceae*	*Corynebacterium*

OTU	Prevalence in Meconium	Prevalence in Technical Negative Controls	Kingdom	Phylum	Class	Order	Family	Genus
OTU_267	1	2	*Bacteria*	*Firmicutes*	*Bacilli*	*Lactobacillales*	*Streptococcaceae*	*Streptococcus*
OTU_277	1	1	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Methylococcales*	*Methylomonaceae*	*NA*
OTU_278	1	2	*Bacteria*	*Firmicutes*	*Bacilli*	*Bacillales*	*Listeriaceae*	*Listeria*
OTU_293	1	1	*Bacteria*	*Firmicutes*	*Erysipelotrichia*	*Erysipelotrichales*	*Erysipelotrichaceae*	*Catenibacterium*
OTU_303	1	2	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*NA*
OTU_318	1	1	*Bacteria*	*Acidobacteria*	*Blastocatellia_(Subgroup_4)*	*Blastocatellales*	*Blastocatellaceae*	*Blastocatella*
OTU_326	1	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*Ruminococcaceae_UCG-005*
OTU_331	1	3	*Bacteria*	*Firmicutes*	*Bacilli*	*Lactobacillales*	*Carnobacteriaceae*	*Alloiococcus*
OTU_342	1	2	*Bacteria*	*Firmicutes*	*Bacilli*	*Lactobacillales*	*Lactobacillaceae*	*Lactobacillus*
OTU_350	1	1	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Rhodocyclaceae*	*NA*
OTU_353	1	1	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Corynebacteriales*	*Dietziaceae*	*Dietzia*
OTU_393	1	1	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Prevotellaceae*	*Alloprevotella*
OTU_397	1	1	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*

OTU	Prevalence in Meconium	Prevalence in Technical Negative Controls	Kingdom	Phylum	Class	Order	Family	Genus
OTU_401	1	1	*Bacteria*	*Planctomycetes*	*Planctomycetacia*	*Isosphaerales*	*Isosphaeraceae*	*NA*
OTU_404	1	1	*Bacteria*	*Firmicutes*	*Negativicutes*	*Selenomonadales*	*Acidaminococcaceae*	*Phascolarctobacterium*
OTU_409	1	1	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Prevotellaceae*	*Alloprevotella*
OTU_413	1	1	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Actinomycetales*	*Actinomycetaceae*	*Actinomyces*
OTU_434	1	1	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Methylophilaceae*	*NA*
OTU_447	1	1	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Frankiales*	*Sporichthyaceae*	*NA*
OTU_456	1	2	*Bacteria*	*Patescibacteria*	*Saccharimonadia*	*Saccharimonadales*	*NA*	*NA*
OTU_465	1	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*NA*
OTU_467	1	3	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	*Cupriavidus*
OTU_471	1	1	*Bacteria*	*Deferribacteres*	*Deferribacteres*	*Deferribacterales*	*Deferribacteraceae*	*Mucispirillum*
OTU_489	1	1	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Micrococcales*	*Microbacteriaceae*	*NA*
OTU_491	1	1	*Bacteria*	*Patescibacteria*	*Saccharimonadia*	*Saccharimonadales*	*Saccharimonadaceae*	*Candidatus_Saccharimonas*

OTU	Prevalence in Meconium	Prevalence in Technical Negative Controls	Kingdom	Phylum	Class	Order	Family	Genus
OTU_51	1	1	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Xanthomonadales*	*Rhodanobacteraceae*	*Luteibacter*
OTU_529	1	1	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	*Sutterella*
OTU_542	1	2	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Enterobacteriales*	*Enterobacteriaceae*	*Providencia*
OTU_561	1	2	*Bacteria*	*Firmicutes*	*Erysipelotrichia*	*Erysipelotrichales*	*Erysipelotrichaceae*	*Turicibacter*
OTU_573	1	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Family_XI*	*Peptoniphilus*
OTU_580	1	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*Agathobacter*
OTU_605	1	1	*Bacteria*	*NA*	*NA*	*NA*	*NA*	*NA*
OTU_633	1	1	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Xanthomonadales*	*Xanthomonadaceae*	*Xanthomonas*
OTU_643	1	1	*Bacteria*	*Proteobacteria*	*NA*	*NA*	*NA*	*NA*
OTU_648	1	1	*Bacteria*	*Tenericutes*	*Mollicutes*	*Mollicutes_RF39*	*NA*	*NA*
OTU_658	1	1	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Pseudomonadales*	*Moraxellaceae*	*Moraxella*
OTU_688	1	2	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*NA*
OTU_717	1	1	*Bacteria*	*Patescibacteria*	*Parcubacteria*	*Candidatus_Ryanbacteria*	*NA*	*NA*
OTU_726	1	1	*Bacteria*	*NA*	*NA*	*NA*	*NA*	*NA*

OTU	Prevalence in Meconium	Prevalence in Technical Negative Controls	Kingdom	Phylum	Class	Order	Family	Genus
OTU_746	1	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*Ruminococcaceae_UCG-013*
OTU_76	1	7	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*Anaerostipes*
OTU_770	1	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*Lachnospiraceae_NK4A136_group*
OTU_81	1	2	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*Blautia*
OTU_854	1	3	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Neisseriaceae*	*Neisseria*
OTU_912	1	1	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Rikenellaceae*	*Alistipes*
OTU_920	1	2	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*NA*
OTU_988	1	1	*Bacteria*	*Acidobacteria*	*Subgroup_6*	*NA*	*NA*	*NA*
OTU_1008	0	1	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Propionibacteriales*	*Propionibacteriaceae*	*Friedmanniella*
OTU_1018	0	6	*Bacteria*	*Proteobacteria*	*Alphaproteobacteria*	*Sphingomonadales*	*Sphingomonadaceae*	*Novosphingobium*
OTU_1023	0	2	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*Lachnoanaerobaculum*
OTU_1028	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*NA*

OTU	Prevalence in Meconium	Prevalence in Technical Negative Controls	Kingdom	Phylum	Class	Order	Family	Genus
OTU_1034	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*Lachnoclostridium*
OTU_1053	0	2	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Bacteroidaceae*	*Bacteroides*
OTU_107	0	2	*Bacteria*	*Cyanobacteria*	*Oxyphotobacteria*	*Nostocales*	*Nostocaceae*	*NA*
OTU_1071	0	3	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	*NA*
OTU_108	0	1	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Frankiales*	*NA*	*NA*
OTU_1105	0	2	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Bacteroidaceae*	*Bacteroides*
OTU_1106	0	1	*Bacteria*	*Deinococcus-Thermus*	*Deinococci*	*Deinococcales*	*Deinococcaceae*	*Deinococcus*
OTU_1109	0	1	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Actinomycetales*	*Actinomycetaceae*	*Actinomyces*
OTU_114	0	2	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Clostridiaceae_1*	*Clostridium_sensu_stricto_10*
OTU_1152	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*NA*
OTU_1153	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*Subdoligranulum*
OTU_1177	0	1	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Micrococcales*	*Microbacteriaceae*	*Microbacterium*

OTU	Prevalence in Meconium	Prevalence in Technical Negative Controls	Kingdom	Phylum	Class	Order	Family	Genus
OTU_1202	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*NA*
OTU_1228	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*NA*
OTU_124	0	1	*Bacteria*	*NA*	*NA*	*NA*	*NA*	*NA*
OTU_125	0	2	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*NA*
OTU_140	0	2	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Corynebacteriales*	*Nocardiaceae*	*Rhodococcus*
OTU_146	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Clostridiaceae_1*	*Clostridium_sensu_stricto_1*
OTU_155	0	3	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*Subdoligranulum*
OTU_164	0	2	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*NA*	*NA*	*NA*
OTU_179	0	1	*Bacteria*	*NA*	*NA*	*NA*	*NA*	*NA*
OTU_183	0	1	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Prevotellaceae*	*Prevotella*
OTU_210	0	1	*Bacteria*	*Firmicutes*	*Bacilli*	*Bacillales*	*Staphylococcaceae*	*Jeotgalicoccus*
OTU_226	0	2	*Bacteria*	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Rhizobiaceae*	*Allorhizobium-Neorhizobium-Pararhizo*

OTU	Prevalence in Meconium	Prevalence in Technical Negative Controls	Kingdom	Phylum	Class	Order	Family	Genus
								bium-Rhizobium
OTU_232	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*Ruminococcus_1*
OTU_236	0	3	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*Ruminococcus_1*
OTU_240	0	1	*Bacteria*	*Firmicutes*	*Negativicutes*	*Selenomonadales*	*Veillonellaceae*	*Megasphaera*
OTU_249	0	2	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Bacteroidaceae*	*Bacteroides*
OTU_252	0	2	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Rikenellaceae*	*Alistipes*
OTU_255	0	1	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	*Polynucleobacter*
OTU_262	0	1	*Bacteria*	*Patescibacteria*	*Parcubacteria*	*NA*	*NA*	*NA*
OTU_272	0	1	*Bacteria*	*Deinococcus-Thermus*	*Deinococci*	*Deinococcales*	*Deinococcaceae*	*Deinococcus*
OTU_279	0	1	*Bacteria*	*Epsilonbacteraeota*	*Campylobacteria*	*Campylobacterales*	*Helicobacteraceae*	*Helicobacter*
OTU_280	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*NA*
OTU_283	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*Lachnospiraceae_ND3007_group*

OTU	Prevalence in Meconium	Prevalence in Technical Negative Controls	Kingdom	Phylum	Class	Order	Family	Genus
OTU_288	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Peptostreptococcaceae*	*Clostridioides*
OTU_292	0	1	*Archaea*	*Euryarchaeota*	*Methanobacteria*	*Methanobacteriales*	*Methanobacteriaceae*	*Methanobacterium*
OTU_308	0	1	*Bacteria*	*NA*	*NA*	*NA*	*NA*	*NA*
OTU_317	0	2	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Bacteroidaceae*	*Bacteroides*
OTU_323	0	1	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Xanthomonadales*	*Xanthomonadaceae*	*Pseudoxanthomonas*
OTU_336	0	1	*Bacteria*	*NA*	*NA*	*NA*	*NA*	*NA*
OTU_337	0	3	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Prevotellaceae*	*NA*
OTU_339	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*Ruminococcaceae_UCG-014*
OTU_352	0	1	*Bacteria*	*Patescibacteria*	*ABY1*	*Candidatus_Magasanikbacteria*	*NA*	*NA*
OTU_359	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*NA*
OTU_360	0	2	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Corynebacteriales*	*Mycobacteriaceae*	*Mycobacterium*
OTU_365	0	1	*Bacteria*	*Planctomycetes*	*Planctomycetacia*	*Gemmatales*	*Gemmataceae*	*Fimbriiglobus*
OTU_368	0	1	*Bacteria*	*Nitrospirae*	*Nitrospira*	*Nitrospirales*	*Nitrospiraceae*	*Nitrospira*

OTU	Prevalence in Meconium	Prevalence in Technical Negative Controls	Kingdom	Phylum	Class	Order	Family	Genus
OTU_374	0	1	*Bacteria*	*Gemmatimonadetes*	*Gemmatimonadetes*	*Gemmatimonadales*	*Gemmatimonadaceae*	*Gemmatirosa*
OTU_380	0	2	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*Ruminococcaceae_UCG-002*
OTU_385	0	2	*Bacteria*	*Dependentiae*	*Babeliae*	*Babeliales*	*NA*	*NA*
OTU_396	0	1	*Bacteria*	*Proteobacteria*	*Alphaproteobacteria*	*Azospirillales*	*Azospirillaceae*	*Skermanella*
OTU_399	0	1	*Bacteria*	*Omnitrophicaeota*	*NA*	*NA*	*NA*	*NA*
OTU_427	0	3	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*Roseburia*
OTU_433	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*Lachnospiraceae_NK4B4_group*
OTU_437	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*Ruminococcaceae_UCG-002*
OTU_449	0	1	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Frankiales*	*Sporichthyaceae*	*hgcI_clade*
OTU_454	0	1	*Bacteria*	*Proteobacteria*	*Alphaproteobacteria*	*Sphingomonadales*	*Sphingomonadaceae*	*Sphingobium*
OTU_487	0	1	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Marinifilaceae*	*Odoribacter*

OTU	Prevalence in Meconium	Prevalence in Technical Negative Controls	Kingdom	Phylum	Class	Order	Family	Genus
OTU_494	0	2	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Nitrosomonadaceae*	*DSSD61*
OTU_499	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*NA*
OTU_525	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*Blautia*
OTU_540	0	1	*Bacteria*	*Firmicutes*	*Erysipelotrichia*	*Erysipelotrichales*	*Erysipelotrichaceae*	*Holdemanella*
OTU_546	0	1	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Micrococcales*	*Microbacteriaceae*	*NA*
OTU_549	0	1	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Pasteurellales*	*Pasteurellaceae*	*NA*
OTU_567	0	1	*Bacteria*	*Cyanobacteria*	*Oxyphotobacteria*	*Chloroplast*	*NA*	*NA*
OTU_575	0	2	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*NA*
OTU_587	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*Faecalibacterium*
OTU_588	0	1	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*NA*	*NA*	*NA*
OTU_590	0	1	*Bacteria*	*Verrucomicrobia*	*Verrucomicrobiae*	*Opitutales*	*Opitutaceae*	*NA*
OTU_593	0	1	*Bacteria*	*Firmicutes*	*Negativicutes*	*Selenomonadales*	*Veillonellaceae*	*NA*
OTU_596	0	1	*Bacteria*	*Verrucomicrobia*	*Verrucomicrobiae*	*Pedosphaerales*	*Pedosphaeraceae*	*NA*
OTU_610	0	1	*Archaea*	*Euryarchaeota*	*Methanobacteria*	*Methanobacteriales*	*Methanobacteriaceae*	*Methanobrevibacter*
OTU_612	0	2	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*NA*

OTU	Prevalence in Meconium	Prevalence in Technical Negative Controls	Kingdom	Phylum	Class	Order	Family	Genus
OTU_614	0	2	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Micrococcales*	*Micrococcaceae*	*Kocuria*
OTU_620	0	1	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Barnesiellaceae*	*Barnesiella*
OTU_624	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*CAG-352*
OTU_632	0	3	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*Roseburia*
OTU_645	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*Ruminococcaceae_NK4A214_group*
OTU_65	0	4	*Bacteria*	*Proteobacteria*	*Alphaproteobacteria*	*Rhodobacterales*	*Rhodobacteraceae*	*NA*
OTU_654	0	1	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Bacteroidaceae*	*Bacteroides*
OTU_657	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*Lachnospira*
OTU_697	0	1	*Bacteria*	*Firmicutes*	*Erysipelotrichia*	*Erysipelotrichales*	*Erysipelotrichaceae*	*Erysipelatoclostridium*
OTU_700	0	2	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*Lachnoclostridium*
OTU_719	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*Coprococcus_2*
OTU_725	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*Ruminiclostridium_5*

OTU	Prevalence in Meconium	Prevalence in Technical Negative Controls	Kingdom	Phylum	Class	Order	Family	Genus
OTU_728	0	1	*Bacteria*	*Bacteroidetes*	*Bacteroidia*	*Bacteroidales*	*Tannerellaceae*	*Parabacteroides*
OTU_733	0	2	*Bacteria*	*Actinobacteria*	*Actinobacteria*	*Frankiales*	*Geodermatophilaceae*	*NA*
OTU_734	0	1	*Bacteria*	*Proteobacteria*	*Deltaproteobacteria*	*Myxococcales*	*Polyangiaceae*	*Pajaroellobacter*
OTU_735	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*Ruminococcaceae_UCG-003*
OTU_757	0	1	*Bacteria*	*Proteobacteria*	*Alphaproteobacteria*	*Rhodobacterales*	*Rhodobacteraceae*	*NA*
OTU_759	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*Lachnoclostridium*
OTU_765	0	1	*Bacteria*	*NA*	*NA*	*NA*	*NA*	*NA*
OTU_766	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*Coprococcus_1*
OTU_802	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*NA*
OTU_817	0	1	*Bacteria*	*Firmicutes*	*Erysipelotrichia*	*Erysipelotrichales*	*Erysipelotrichaceae*	*Faecalitalea*
OTU_825	0	1	*Bacteria*	*Proteobacteria*	*Alphaproteobacteria*	*Sphingomonadales*	*Sphingomonadaceae*	*Sandaracinobacter*
OTU_826	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*NA*
OTU_845	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*Ruminococcaceae_UCG-002*

OTU	Prevalence in Meconium	Prevalence in Technical Negative Controls	Kingdom	Phylum	Class	Order	Family	Genus
OTU_847	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*NA*
OTU_851	0	2	*Bacteria*	*Actinobacteria*	*Coriobacteriia*	*Coriobacteriales*	*Eggerthellaceae*	*Eggerthella*
OTU_880	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*Ruminiclostridium_5*
OTU_92	0	6	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*Fusicatenibacter*
OTU_938	0	2	*Bacteria*	*Firmicutes*	*Bacilli*	*Lactobacillales*	*NA*	*NA*
OTU_950	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Ruminococcaceae*	*Ruminococcaceae_UCG-013*
OTU_955	0	2	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*NA*
OTU_972	0	2	*Bacteria*	*Proteobacteria*	*Alphaproteobacteria*	*NA*	*NA*	*NA*
OTU_981	0	2	*Bacteria*	*Proteobacteria*	*Gammaproteobacteria*	*Betaproteobacteriales*	*Burkholderiaceae*	*Janthinobacterium*
OTU_984	0	1	*Bacteria*	*Firmicutes*	*Clostridia*	*Clostridiales*	*Lachnospiraceae*	*Lachnospiraceae_UCG-004*

Table 2.4. Significantly and differentially expressed genes in MM-E

Gene	Base Mean	Log2 Fold Change	False discovery rate (FDR)	Enriched in
MSMB	20.77445742	3.117731787	0.019852934	MM-E
MEGF11	117.7691835	2.766463847	0.048267547	MM-E
CYP4X1	135.7096101	2.705286143	0.025299859	MM-E
EVC	14.76866706	2.593066198	0.048301412	MM-E
ADRA2C	27.10304841	2.563361071	0.001139399	MM-E
WSCD2	79.20988249	2.350003797	0.001685297	MM-E
RNU12	26.92945062	2.187166188	0.001546647	MM-E
L1TD1	54.63412587	2.149166755	0.006582495	MM-E
CXCL3	2006.38609	2.120774947	5.13E-07	MM-E
CES5AP1	29.86863134	2.095816568	0.009728374	MM-E
LOC388882	30.39688524	1.969917451	0.049356541	MM-E
C1orf95	47.03475841	1.935602934	0.004617	MM-E
NMUR2	84.02048635	1.910981979	0.002629131	MM-E
LINC00612	20.89157035	1.909723185	0.034777996	MM-E
ACTN2	30.85197838	1.901259569	0.034187892	MM-E
FOXQ1	1002.484627	1.896798711	0.004116525	MM-E
MIR378D2	31.95163204	1.800125461	0.001602875	MM-E
SLC34A2	135.8512812	1.785834304	0.0083238	MM-E

Gene	Base Mean	Log2 Fold Change	False discovery rate (FDR)	Enriched in
NPPC	28.28336026	1.775652564	0.012013733	MM-E
NOS2	188.2398668	1.711376214	0.001243684	MM-E
TDRD5	48.39430894	1.649255382	0.003605139	MM-E
EVA1C	14.96941265	1.645148621	0.038556838	MM-E
ARC	1047.557695	1.633586575	0.000289185	MM-E
CDH12	235.201314	1.623396577	0.042129767	MM-E
RNVU1-14	335.5012872	1.61531332	0.000151358	MM-E
FAM3B	372.7071886	1.604792565	0.023295708	MM-E
SLC14A2	179.0450388	1.601108977	0.020496534	MM-E
LAMA2	211.6001534	1.509456015	0.003288415	MM-E
CCL20	1034.912748	1.507070713	0.004868315	MM-E
MROH9	15.24900784	1.486778965	0.036059217	MM-E
PVT1	178.2650316	1.472803576	0.000552766	MM-E
LOC100126784	21.53005017	1.465059953	0.049531789	MM-E
KCNH3	259.0312309	1.458290295	0.006723442	MM-E
RGMB-AS1	170.1776329	1.457264537	0.004193801	MM-E
KY	182.2281869	1.451951908	0.006566206	MM-E
VASN	150.74635	1.446054809	0.023347001	MM-E
CXCL5	52.0069157	1.427127286	0.013751959	MM-E
MSI1	235.9239227	1.426118815	0.000790334	MM-E

Gene	Base Mean	Log2 Fold Change	False discovery rate (FDR)	Enriched in
AXIN2	673.8127907	1.423066757	0.009000692	MM-E
LGR5	6770.669082	1.417096206	0.023528454	MM-E
FOSL1	783.7443709	1.410629829	0.023295708	MM-E
CD200	258.6104938	1.408496617	4.22E-10	MM-E
NR4A1	16002.10353	1.402830481	0.002350027	MM-E
RUFY4	17.20109869	1.388649265	0.047318916	MM-E
MIR614	340.8345125	1.38513969	0.006612009	MM-E
MYC	5121.974183	1.373964374	0.000504209	MM-E
DFNB31	239.8696049	1.363669077	0.000210311	MM-E
GPC4	288.2755406	1.359956153	0.034347274	MM-E
SOX9	4824.532981	1.349253157	0.02495019	MM-E
POU3F1	53.26705622	1.337270095	0.037082197	MM-E
NKILA	70.94079094	1.320206873	0.005715301	MM-E
CXCL2	8413.567882	1.315686392	0.009413871	MM-E
TNFSF15	2348.942639	1.314867019	0.000264883	MM-E
C1orf194	24.12520063	1.302414928	0.037191309	MM-E
MEX3A	100.9009908	1.300450966	0.005983972	MM-E
PPRC1	3621.314014	1.297109754	0.006582495	MM-E
THBS1	2564.974704	1.287327665	0.036217138	MM-E
EGOT	83.04091814	1.274049095	0.010553993	MM-E

Gene	Base Mean	Log2 Fold Change	False discovery rate (FDR)	Enriched in
CDR2L	550.7081739	1.265481249	0.003999657	MM-E
HES1	16156.85891	1.251967799	0.046432346	MM-E
C2orf48	40.66839017	1.24949804	0.034508671	MM-E
SULT1C4	89.68137102	1.248110671	0.02495019	MM-E
FBXO16	62.70234755	1.244819588	0.034616073	MM-E
RHOV	329.4369846	1.240455429	0.020283939	MM-E
NFKB2	1730.570338	1.238873117	2.71E-05	MM-E
MYO3A	100.9725082	1.23404845	0.003387712	MM-E
KCNAB3	19.95257349	1.233785463	0.039942035	MM-E
CHAC1	839.3504513	1.233112786	0.038663057	MM-E
SLC9C1	44.83016393	1.227803824	0.022771769	MM-E
LOC101928100	203.1875517	1.226462341	0.01651249	MM-E
PABPC4L	42.95792415	1.221144442	0.02985373	MM-E
LOC642846	269.3201154	1.217898224	0.028634861	MM-E
ZSCAN12P1	310.7459748	1.212163196	0.022887021	MM-E
NOTCH1	903.164258	1.206468812	0.007660305	MM-E
HPSE	68.80417668	1.198906097	0.019188032	MM-E
ZNRF3	1158.894619	1.196460911	0.000374289	MM-E
NOTCH4	68.71244682	1.180718079	0.04193255	MM-E
CCAT1	150.4531906	1.173186106	0.033842873	MM-E

Gene	Base Mean	Log2 Fold Change	False discovery rate (FDR)	Enriched in
C7orf61	44.69240109	1.172641189	0.024427925	MM-E
HBEGF	6004.766903	1.169953074	0.024427925	MM-E
TRPV4	107.1513861	1.166981025	0.025439261	MM-E
ZDHHC8P1	499.4746283	1.166778112	0.032539458	MM-E
NKD2	422.4839946	1.166206021	0.038312136	MM-E
FAM24B-CUZD1	45.94355138	1.164488006	0.014173988	MM-E
RCOR2	77.07351505	1.149996315	0.028289839	MM-E
ADRA2A	528.7215772	1.146877577	0.000386643	MM-E
SQLE	4620.107617	1.122557529	0.036059217	MM-E
SHROOM4	241.1739707	1.121991456	0.021381045	MM-E
SIX5	209.0972063	1.111906383	0.001440849	MM-E
MIRLET7BHG	80.02496285	1.109237687	0.049498848	MM-E
NHS	380.3744953	1.108446824	0.005983972	MM-E
RGMB	2554.416581	1.104442946	0.007530364	MM-E
MFAP4	1118.519486	1.102353846	0.04642672	MM-E
ELN	107.7633429	1.097606939	0.024047856	MM-E
RELB	1612.819814	1.094605784	0.039942035	MM-E
PER1	2129.224593	1.093817111	0.042724765	MM-E
LINC01465	82.08417136	1.091836603	0.019738723	MM-E
BCL6	747.745828	1.091592839	0.008075922	MM-E

Gene	Base Mean	Log2 Fold Change	False discovery rate (FDR)	Enriched in
ROBO2	478.6162243	1.080660657	0.021666486	MM-E
RAPGEFL1	877.8112844	1.078081462	0.000121659	MM-E
MTSS1L	64.81303817	1.078054145	0.024427925	MM-E
CXCL1	316.5221903	1.072020903	0.01143484	MM-E
FOXO6	147.4705253	1.071026887	0.046017196	MM-E
NR4A3	924.0833874	1.069900871	0.02495019	MM-E
LOC221946	87.66861362	1.069606203	0.007193488	MM-E
C18orf54	69.14691846	1.069199001	0.0265975	MM-E
LDLR	19827.14419	1.066898114	0.02311695	MM-E
COL9A2	448.1749336	1.06395361	0.035800078	MM-E
OVGP1	146.1512368	1.063437227	0.044637723	MM-E
MNS1	145.3310743	1.062685426	0.038035122	MM-E
C5orf34	75.26522578	1.05839987	0.029602423	MM-E
PNPLA3	105.7239956	1.058159763	0.002350027	MM-E
CEP112	47.85624303	1.053534605	0.032224833	MM-E
CNN2	796.3330662	1.051758346	0.002343668	MM-E
OR51E1	266.5929011	1.03307216	0.012787193	MM-E
RASL10B	142.9873434	1.031805176	0.037326936	MM-E
EP400NL	82.45755555	1.025600213	0.002430343	MM-E
EDN1	36686.0727	1.020016951	0.016641151	MM-E

Gene	Base Mean	Log2 Fold Change	False discovery rate (FDR)	Enriched in
PROX1	566.6513161	1.012743699	0.007403603	MM-E
HSD17B7P2	384.638959	1.010193815	0.034364881	MM-E
EXPH5	303.3190883	1.004550872	0.023295708	MM-E
LOC100506159	83.38815143	-7.731491434	0.001664978	Other
LOC101929532	12.25884812	-3.271501596	0.031268763	Other
CCKBR	19.71276889	-2.949425335	0.022753338	Other
TPO	41.21778229	-2.874074974	0.000289185	Other
LPAR3	31.29012013	-2.747105023	0.00205187	Other
FGF12	22.8626664	-2.74137009	0.00731161	Other
DNER	100.4381543	-2.601030215	0.043903117	Other
S100G	214.2947825	-2.520432284	0.000941493	Other
SEPT4-AS1	202.4789386	-2.508530549	0.000151358	Other
LINC00671	58.73732319	-2.470582441	0.000551717	Other
G6PC	116.3431418	-2.448846773	0.000366253	Other
MSLN	10404.70716	-2.435469684	0.020705922	Other
SLC30A8	12.97525828	-2.409511432	0.026568863	Other
CDH13	20.55368637	-2.368457545	0.02311695	Other
SEPT4	1078.319906	-2.361534848	0.000689032	Other
REN	1106.237982	-2.303696313	0.001335369	Other
APOA4	313587.2217	-2.246615271	6.97E-05	Other

Gene	Base Mean	Log2 Fold Change	False discovery rate (FDR)	Enriched in
ASPDH	103.4840262	-2.242175242	5.05E-07	Other
HPR	48.67123183	-2.232378435	0.02495019	Other
CYP3A4	1239.396816	-2.168179032	2.38E-09	Other
HBM	14.30172398	-2.071179863	0.024704911	Other
CYP1A1	69.64980339	-2.002660297	0.020952676	Other
ACSM2B	61.51831137	-1.995327226	0.001880838	Other
SLN	113.2386159	-1.978396981	0.009583085	Other
TMEM132B	36.33778204	-1.96927762	0.01444188	Other
SPOCK3	46.66677238	-1.936475981	0.023295708	Other
FRRS1	96.37041043	-1.918180864	0.006207235	Other
PM20D1	110.9959034	-1.861184318	0.02314163	Other
PCDH7	130.4058889	-1.828967745	0.003351294	Other
GIMAP6	108.0298791	-1.819767613	0.021381045	Other
CYP2B7P	2298.309465	-1.803681768	0.000383257	Other
ADGRF1	36.21198055	-1.79249042	0.049356541	Other
FZD4	143.9315225	-1.781027586	0.000707942	Other
BAAT	75.45925735	-1.772677193	0.032224833	Other
LINC00113	34.82936275	-1.770971817	0.045965619	Other
SLC22A7	51.25961483	-1.769799416	0.003617832	Other
ADGRL3	23.18351944	-1.760434284	0.023931085	Other

Gene	Base Mean	Log2 Fold Change	False discovery rate (FDR)	Enriched in
BHMT	176.5482602	-1.74909455	0.012393421	Other
SERPINE2	2506.302309	-1.740899985	0.002350027	Other
APOC3	61341.46476	-1.740812624	0.007403603	Other
C8G	3297.63904	-1.662980381	0.006905677	Other
APOA1	563298.695	-1.662045653	0.000689032	Other
APOA1-AS	262632.168	-1.658310606	0.000790334	Other
TREML3P	36.04101379	-1.657380343	0.023894411	Other
LINGO1	504.6214903	-1.653123842	0.006582495	Other
LOC400043	68.75223492	-1.584653818	0.049008515	Other
PNCK	2558.444023	-1.581419425	0.005983972	Other
SLAMF6	35.80522583	-1.571243193	0.036859377	Other
CYSLTR1	16.36262195	-1.563946957	0.039141175	Other
HLA-DRB1	6939.346936	-1.556607372	0.017328679	Other
SPANXN2	57.48240259	-1.554974401	0.000524701	Other
CD160	117.1330868	-1.548254226	0.036859377	Other
SERPINI1	217.0178297	-1.540844601	0.005370809	Other
CTNNA2	114.4950503	-1.53983477	0.011431278	Other
SUMO4	18.38031778	-1.538024194	0.044637723	Other
GK3P	12.87230763	-1.535705025	0.034777996	Other
FBLN2	99.39215515	-1.530116918	0.026643777	Other

Gene	Base Mean	Log2 Fold Change	False discovery rate (FDR)	Enriched in
OTOP3	2425.97822	-1.529996127	0.001036958	Other
P2RX2	700.3330788	-1.528418593	0.02222647	Other
TRIB2	51.82179511	-1.524811553	0.021381045	Other
PHLDA3	358.1944413	-1.52069258	0.007172422	Other
STAR	40.92258061	-1.491894385	0.003288415	Other
PPP2R2B	86.07701172	-1.491246132	0.035237109	Other
GGT3P	15.12709481	-1.482813729	0.021163852	Other
GLYAT	88.97291203	-1.479232014	0.028533883	Other
F13B	351.5148538	-1.477869913	0.01741047	Other
P2RY12	15.98399178	-1.47679084	0.02197507	Other
GUCA2B	2476.986919	-1.473251291	0.046966106	Other
FSTL1	107.918045	-1.466419836	0.011820222	Other
SLC17A1	32.90639648	-1.452075212	0.036938585	Other
NPL	1805.225372	-1.448716059	0.000790334	Other
ZG16	552.6271651	-1.444933655	0.031452691	Other
DGAT2	3215.740069	-1.43578093	1.08E-06	Other
CD3D	154.0437705	-1.425815251	0.02516504	Other
ENPP7	26477.35996	-1.417537789	1.15E-10	Other
GIMAP7	81.72257311	-1.40189198	0.036217138	Other
C11orf86	2882.237789	-1.390762807	0.007191281	Other

Gene	Base Mean	Log2 Fold Change	False discovery rate (FDR)	Enriched in
B4GALNT3	159.0488561	-1.387044352	0.027834692	Other
SLC2A7	147.6506746	-1.385526136	0.001139399	Other
TSPAN32	721.9800235	-1.38088981	0.049338392	Other
CLRN1-AS1	135.6507833	-1.371491453	0.002308061	Other
A2M	113.8245729	-1.3699158	0.001036958	Other
APOA2	3930.269802	-1.360038398	0.041246749	Other
LOC102467214	591.6898506	-1.349249074	0.003181393	Other
BEGAIN	195.4977337	-1.344619174	0.004664911	Other
CD244	120.7936264	-1.340007202	0.006300033	Other
PECAM1	59.93169942	-1.335010167	0.032181214	Other
SLC38A3	80.79145313	-1.324450063	0.03803893	Other
ADA	1090.962379	-1.320408571	0.036524945	Other
SCAMP5	632.6485602	-1.294812818	0.049338392	Other
TMIGD1	2413.829351	-1.290833365	0.000289185	Other
RAB42	1910.144097	-1.288877287	0.001456839	Other
DNAH5	36.61597782	-1.285839262	0.021381045	Other
UNC93A	2410.486932	-1.279685325	6.92E-06	Other
A4GALT	202.2729765	-1.278425408	0.02495019	Other
SPIRE1	108.9921683	-1.265155604	0.010989371	Other
MPP1	3207.997032	-1.254618678	6.66E-05	Other

Gene	Base Mean	Log2 Fold Change	False discovery rate (FDR)	Enriched in
CIDEC	2895.15844	-1.24741492	0.004791614	Other
C1orf105	71.00812306	-1.244113987	0.011417985	Other
GPX3	656.031556	-1.243917526	0.03789463	Other
STK32C	687.208622	-1.242899889	0.001139399	Other
SERPINC1	30.43534461	-1.242400818	0.036338603	Other
PTH1R	505.8365223	-1.239908421	0.011474006	Other
A2M-AS1	33.37117747	-1.230850791	0.03998404	Other
SLC13A5	393.7810276	-1.230644387	0.011820222	Other
SH2D1B	77.6739858	-1.229054038	0.012169427	Other
SLC15A1	13988.87574	-1.228065691	0.000390238	Other
DZIP1	155.2557832	-1.224471511	0.011806775	Other
MYL3	339.9312772	-1.220981451	0.001243684	Other
GDPD2	1370.518852	-1.208030719	0.040227078	Other
TMPRSS13	518.3440117	-1.205941868	0.012253047	Other
SLC34A3	3487.169128	-1.205388353	0.001139399	Other
PDZRN3-AS1	41.18949121	-1.201420875	0.023295708	Other
XKRX	277.7556639	-1.195472062	0.040227078	Other
REEP2	570.4384779	-1.193696281	0.000696165	Other
CYP19A1	223.9732122	-1.192912714	0.04642672	Other
GS1-594A7.3	37.92982163	-1.191618786	0.046966106	Other

Gene	Base Mean	Log2 Fold Change	False discovery rate (FDR)	Enriched in
CAPN3	11155.38665	-1.187414424	0.001036958	Other
CLRN1	397.6014107	-1.18667447	0.016749901	Other
TM4SF5	2057.943428	-1.179735374	0.02176872	Other
IKZF3	127.211745	-1.17916675	0.027661786	Other
SLC28A2	1312.078578	-1.178781071	0.019852934	Other
LINC01503	111.5072691	-1.170088728	0.003434017	Other
SLC2A5	3141.699838	-1.168014928	0.030626258	Other
VSIG2	177.0573714	-1.167674965	0.032224833	Other
GZMA	266.8907295	-1.164146777	0.043154303	Other
APOBEC1	1737.124109	-1.162666641	0.001139399	Other
GPAT3	1425.442552	-1.151533457	4.83E-08	Other
KCNG1	2140.092515	-1.141219894	5.65E-05	Other
CPVL	20111.03679	-1.133500113	0.002430343	Other
MAMDC4	58789.85242	-1.121766888	0.002142034	Other
ZCCHC24	772.4770488	-1.121499256	0.039942035	Other
GDA	5534.145395	-1.118155384	0.045524289	Other
AKR1D1	435.6516732	-1.108013954	0.016724497	Other
FRMD1	516.3820233	-1.0974623	0.019251563	Other
LTF	396.3659527	-1.096864131	0.018616273	Other
ST8SIA4	123.8597446	-1.089463042	0.003617832	Other

Gene	Base Mean	Log2 Fold Change	False discovery rate (FDR)	Enriched in
DUSP9	96.88920722	-1.088429072	0.047064771	Other
CD36	1609.654398	-1.083746647	0.022771769	Other
TRPM4	336.2205185	-1.080559064	0.000940356	Other
G0S2	12000.1537	-1.075712882	0.02495019	Other
ABCG5	1823.894132	-1.075387727	0.003419271	Other
SECTM1	184.8290217	-1.073486203	0.038730378	Other
TRIM63	284.1529105	-1.058301643	0.012893891	Other
PAPL	176.4468344	-1.056627137	0.042129767	Other
LOC100133286	3725.864492	-1.055845861	0.037243134	Other
MME	16681.54852	-1.054953946	0.010880562	Other
AFP	16097.07746	-1.054627186	0.024051223	Other
PDZK1	2671.534262	-1.051569416	0.002308061	Other
SLC28A1	2495.342817	-1.05100622	0.003639841	Other
MUC3A	4563.341101	-1.043389905	0.000790334	Other
TGM1	47.63354355	-1.040883844	0.022390358	Other
GSTA2	4518.484777	-1.037559264	0.046565377	Other
SHBG	1696.898935	-1.025727184	0.022873592	Other
P4HA2-AS1	248.8460005	-1.025200732	2.71E-05	Other
REEP1	2556.959636	-1.022618313	0.006028051	Other
PNMA1	1299.461174	-1.021912171	6.95E-08	Other

Gene	Base Mean	Log2 Fold Change	False discovery rate (FDR)	Enriched in
AMN	69661.57922	-1.021425064	0.005972991	Other
CHST2	67.21812567	-1.017567206	0.034445612	Other
PFKFB4	1222.511201	-1.009454331	0.002974447	Other
GGT1	11180.37668	-1.007645822	0.038398224	Other
LAMTOR5-AS1	124.9938266	-1.007083954	0.001664978	Other
MFI2-AS1	2670.219203	-1.002768247	0.002974447	Other
SCN9A	60.131233	-1.002651626	0.033386005	Other
SLC43A2	6836.316022	-1.000754388	0.005136422	Other

Table 2.5. Fetal Meconium Isolates

Isolate Identifier	Meconium SampleID	Meconium subset by 16S V4 rRNA Sequencing	Isolation Media	Isolation Conditions	Subculture Coditions	SILVA Taxonomy	% Taxonomy identity	% Identity to OTU10	% Identity to OTU12
M45	1539J	MM	cRPMI + THP1 monocytes	48h 37C stationary conditions	BHI, 48h 37C stationary conditions	Bacteria;Proteobacteria;Gammaproteobacteria;Pseudomonadales;Pseudomonadaceae;Pseudomonas;	99.075	75	79.835
M46	1539J	MM	cRPMI + THP1 monocytes	48h 37C stationary conditions	BHI, 48h 37C stationary conditions	Bacteria;Proteobacteria;Gammaproteobacteria;Pseudomonadales;Pseudomonadaceae;Pseudomonas;	98.4536	74.699	79.508
Micro36	1543J	MM	cRPMI + THP1 monocytes	48h 37C stationary conditions	BHI, 48h 37C stationary conditions	Bacteria;Actinobacteria;Actinobacteria;Micrococcales;Micrococcaceae;Micrococcus;	99.8304	**97.233**	76.285
M44	1539J	MM	cRPMI + THP1 monocytes	48h 37C stationary conditions	BHI, 48h 37C stationary conditions	Bacteria;Proteobacteria;Gammaproteobacteria;Xanthomonadales;Xanthomonadaceae;Stenotrophomonas;	99.7927	79.051	76.078
M47	1539J	MM	cRPMI + THP1 monocytes	48h 37C stationary conditions	BHI, 48h 37C stationary conditions	Bacteria;Proteobacteria;Gammaproteobacteria;Xanthomonadales;Xanthomonadaceae;Stenotrophomonas;	99.793	79.051	76.078

Isolate Identifier	Meconium SampleID	Meconium subset by 16S V4 rRNA Sequencing	Isolation Media	Isolation Conditions	Subculture Coditions	SILVA Taxonomy	% Taxonomy identity	% Identity to OTU10	% Identity to OTU12
M52	1539J	MM	cRPMI + THP1 monocytes	48h 37C stationary conditions	BHI, 48h 37C stationary conditions	Bacteria;Proteobacteria;Gammaproteobacteria;Xanthomonadales;Xanthomonadaceae;Stenotrophomonas;	99.5889	79.051	76.078
M155	1539J	MM	cRPMI + THP1 monocytes	48h 37C stationary conditions	BHI, 48h 37C stationary conditions	Bacteria;Proteobacteria;Gammaproteobacteria;Xanthomonadales;Xanthomonadaceae;Stenotrophomonas;	99.7923	79.051	76.078
M49	1539J	MM	cRPMI + THP1 monocytes	48h 37C stationary conditions	BHI, 48h 37C stationary conditions	Bacteria;Proteobacteria;Gammaproteobacteria;Xanthomonadales;Xanthomonadaceae;Stenotrophomonas;	98.9765	78.656	75.686
M53	1544J	MM	cRPMI + THP1 monocytes	48h 37C stationary conditions	BHI, 48h 37C stationary conditions	Bacteria;Proteobacteria;Gammaproteobacteria;Xanthomonadales;Xanthomonadaceae;Stenotrophomonas;	99.0702	78.656	75.686
M37	1543J	MM	cRPMI + THP1 monocytes	48h 37C stationary conditions	BHI, 48h 37C stationary conditions	Bacteria;Actinobacteria;Actinobacteria;Micrococcales;Microbacteriaceae;Microbacterium;	99.5767	92.095	75.494
M38	1543J	MM	cRPMI + THP1 monocytes	48h 37C stationary conditions	BHI, 48h 37C stationary conditions	Bacteria;Actinobacteria;Actinobacteria;Micrococcales;Microbacteriaceae;Microbacterium;	99.7901	92.095	75.494

Isolate Identifier	Meconium SampleID	Meconium subset by 16S V4 rRNA Sequencing	Isolation Media	Isolation Conditions	Subculture Coditions	SILVA Taxonomy	% Taxonomy identity	% Identity to OTU10	% Identity to OTU12
M39	1543J	MM	cRPMI + THP1 monocytes	48h 37C stationary conditions	BHI, 48h 37C stationary conditions	Bacteria;Actinobacteria;Actinobacteria;Micrococcales;Microbacteriaceae;Microbacterium;	99.894	92.095	75.494
MicroRef	NA	NA	NA	Obtained from ATCC, Cat. No. 4698	BHI, 48h 37C stationary conditions	Bacteria;Actinobacteria;Actinobacteria;Micrococcales;Micrococcaceae;Micrococcus;	100	**97.233**	76.285

Table 2.6. Fetal Meconium Isolate Whole Genome Sequencing Statistics

Isolate Identifier	Meconium SampleID	Meconium subset by 16S V4 rRNA Sequencing	No. contigs	Total genome length	N50	N75	L50	L75	Largest Contig	GC (%)	Reads mapped to contig assembly (%)	x Mean Coverage (Standard Deviation)	NCBI Accession
Micro36	1543J	MM	101	2584920	59431	27520	14	31	205052	73	99.17	511.3 (332.3)	PRJNA498337

Table 2.7. Average nucleotide identity and coverage of Micro36 against all available genomes in *Micrococcus*

Average nucleotide identity (%)	Micro36	*M. aloeveae* M.71	*M. luteus* CCH3 E2	*M. luteus* K39	*M. luteus* NCTC 2665	*M. luteus* SGAir0127	*M. luteus* UMB0038	*M. luteus* UMB0189	*M. lylae* NBRC 15355	*M. terreus* CG M.CC 1.7054	*Micococcus sp* HMSC30C05	*Micrococcus sp* KT16	*Micrococcus sp* HMSC31B01
Micro36	100	97.9974632	98.0426837	98.0584686	96.8318581	98.0360053	98.4259949	98.7903333	79.5441421	76.3321721	98.844429	98.0907957	98.7530279
M. aloeveae M.71	98.058359 7	100	98.2177074	98.2522981	96.951068	98.1041768	98.0323145	98.0554157	79.5200384	76.4159475	98.044476	98.1644046	98.1657169
M. luteus CCH3 E2	98.133942 8	98.1773169	100	98.0411098	97.0526662	98.2727932	98.0219056	98.1267268	79.8534474	76.8618655	98.1046368	98.0996323	98.1242562
M. luteus K39	98.073570 9	98.2583714	98.0613886	100	96.9305351	98.0665795	98.3225642	98.1132818	79.9813266	76.5418922	98.1832695	98.0729653	98.3366949
M. luteus NCTC 2665	96.890628 1	96.9782001	96.9734035	96.9312467	100	97.1005789	96.8776828	96.8099104	79.4501095	76.5501732	96.8119252	96.9880309	96.9651932
M. luteus SGAir0127	98.135118 3	98.0976999	98.2995674	98.0946873	97.1321742	100	98.1772406	98.1140785	79.7867218	76.6759268	98.3413836	98.0944176	98.3678172

Average nucleotide identity (%)	Micro36	*M. aloeveae* M.71	*M. luteus* CCH3 E2	*M. luteus* K39	*M. luteus* NCTC 2665	*M. luteus* SGAir0127	*M. luteus* UMB0038	*M. luteus* UMB0189	*M. lylae* NBRC 15355	*M. terreus* CG M.CC 1.7054	*Micococcus sp* HMSC30C05	*Micrococcus sp* KT16	*Micrococcus sp* HMSC31B01
M. luteus UMB0038	98.340404	97.9101996	97.9579452	98.2649475	96.720887	98.1521142	100	98.4231542	79.5841112	76.4308441	98.5050303	97.962887	98.3978776
M. luteus UMB0189	98.701493	97.9904329	98.0826111	98.0003536	96.7558408	98.0725415	98.4038791	100	79.8542671	76.66552	98.8244134	98.0234393	98.6148494
M. lylae NBRC 15355	79.590697	79.596312	80.0220329	79.8794871	79.2570298	79.6977507	79.5073341	79.6598372	100	76.7967064	81.1839407	79.5266632	80.9103028
M. terreus CG M.CC 1.7054	76.459476	76.3096553	76.8020678	76.4750488	76.538867	76.5018682	76.5411418	76.4877541	76.8581604	100	77.6194206	76.3994446	77.4471057
Micococcus sp HMSC30C05	98.796282	97.9767778	98.1086493	98.0463374	96.765823	98.145107	98.4663876	98.8571791	80.4421481	77.1884204	100	97.9964088	98.7269168
Micrococcus sp KT16	98.158084	98.2533806	98.1960056	98.1892312	97.0347042	98.1321247	98.0689642	98.1618233	79.4923301	76.5689704	98.0163331	100	98.2473815

Average nucleotide identity (%)			*M. aloeverae* M.71		*M. luteus* CCH3 E2		*M. luteus* K39		*M. luteus* NCTC 2665		*M. luteus* SGAir0127		*M. luteus* UMB0038		*M. luteus* UMB0189		*M. lylae* NBRC 15355		*M. terreus* CG M.CC 1.7054		*Micrococcus sp* HMSC30C05		*Micrococcus sp* KT16		*Micrococcus sp* HMSC31B01	
Micrococcus sp HMSC31B01	Micro36	98.67																								
		9903	98.063		98.08975		98.19644		96.76977		98.16425		98.38640		98.66344		80.23635		77.28796		98.785		98.0992			
		1	257		19		79		74		25		3		89		15		98		6975		954		100	

Table 2.8. Unique annotated predicted protein sequences in Micro36 genome as compared to MicroRef1

Unique ID	Gene Cluster ID	COG Function Accession	COG Function
4500	GC_00002362	COG3255	Putative sterol carrier protein
4530	GC_00002392	COG3255	Putative sterol carrier protein
4114	GC_00001976	COG1468|COG1518	CRISPR/Cas system-associated exonuclease Cas4, RecB family|CRISPR/Cas system-associated endonuclease Cas1
4363	GC_00002225	COG1468|COG1518	CRISPR/Cas system-associated exonuclease Cas4, RecB family|CRISPR/Cas system-associated endonuclease Cas1
4481	GC_00002343	COG1343	CRISPR/Cas system-associated endoribonuclease Cas2
4339	GC_00002201	COG1203	CRISPR/Cas system-associated endonuclease/helicase Cas3
4313	GC_00002175	COG4823	Abortive infection bacteriophage resistance protein
4645	GC_00002507	COG1225	Peroxiredoxin
4565	GC_00002427	COG0695	Glutaredoxin
4393	GC_00002255	COG2128	Alkylhydroperoxidase family enzyme, contains CxxC motif
4680	GC_00002542	COG2128	Alkylhydroperoxidase family enzyme, contains CxxC motif
4186	GC_00002048	COG3485	Protocatechuate 3,4-dioxygenase beta subunit
4310	GC_00002172	COG3485	Protocatechuate 3,4-dioxygenase beta subunit

Unique ID	Gene Cluster ID	COG Function Accession	COG Function
4455	GC_00002317	COG0346	Catechol 2,3-dioxygenase or other lactoylglutathione lyase family enzyme
4529	GC_00002391	COG0346	Catechol 2,3-dioxygenase or other lactoylglutathione lyase family enzyme
4595	GC_00002457	COG0179	2-keto-4-pentenoate hydratase/2-oxohepta-3-ene-1,7-dioic acid hydratase (catechol pathway)
4237	GC_00002099	COG1878	Kynurenine formamidase
4118	GC_00001980	COG1359	Quinol monooxygenase YgiN
4230	GC_00002092	COG1249	Pyruvate/2-oxoglutarate dehydrogenase complex, dihydrolipoamide dehydrogenase (E3) component or related enzyme
4410	GC_00002272	COG2041	Periplasmic DMSO/TMAO reductase YedYZ, molybdopterin-dependent catalytic subunit
4654	GC_00002516	COG1853	NADH-FMN oxidoreductase RutF, flavin reductase (DIM6/NTAB) family
4632	GC_00002494	COG1018	Ferredoxin-NADP reductase
4224	GC_00002086	COG1012	Acyl-CoA reductase or other NAD-dependent aldehyde dehydrogenase

Unique ID	Gene Cluster ID	COG Function Accession	COG Function
4154	GC_00002016	COG0654	2-polyprenyl-6-methoxyphenol hydroxylase and related FAD-dependent oxidoreductases
4255	GC_00002117	COG0654	2-polyprenyl-6-methoxyphenol hydroxylase and related FAD-dependent oxidoreductases
4584	GC_00002446	COG0654	2-polyprenyl-6-methoxyphenol hydroxylase and related FAD-dependent oxidoreductases
4638	GC_00002500	COG3336	Cytochrome c oxidase assembly factor CtaG
4733	GC_00002595	COG0604	NADPH:quinone reductase or related Zn-dependent oxidoreductase
4415	GC_00002277	COG4641	Spore maturation protein CgeB
4401	GC_00002263	COG1674	DNA segregation ATPase FtsK/SpoIIIE and related proteins
4567	GC_00002429	COG1475	Chromosome segregation protein Spo0J, contains ParB-like nuclease domain
4374	GC_00002236	COG2856	Zn-dependent peptidase ImmA, M78 family
1791	GC_00000803	COG0133	Tryptophan synthase beta chain
1792	GC_00000803	COG0133	Tryptophan synthase beta chain
4092	GC_00001954	COG2066	Glutaminase
4482	GC_00002344	COG0076	Glutamate or tyrosine decarboxylase or a related PLP-dependent protein

Unique ID	Gene Cluster ID	COG Function Accession	COG Function
4150	GC_00002012	COG3842	ABC-type Fe3+/spermidine/putrescine transport systems, ATPase components
4168	GC_00002030	COG0697	Permease of the drug/metabolite transporter (DMT) superfamily
4358	GC_00002220	COG0697	Permease of the drug/metabolite transporter (DMT) superfamily
4703	GC_00002565	COG0028	Acetolactate synthase large subunit or other thiamine pyrophosphate-requiring enzyme
4521	GC_00002383	COG1396|COG2856	Transcriptional regulator, contains XRE-family HTH domain|Zn-dependent peptidase ImmA, M78 family
4648	GC_00002510	COG1396|COG2856	Transcriptional regulator, contains XRE-family HTH domain|Zn-dependent peptidase ImmA, M78 family
4526	GC_00002388	COG1457	Purine-cytosine permease or related protein
4301	GC_00002163	COG1780	Protein involved in ribonucleotide reduction
4317	GC_00002179	COG0015	Adenylosuccinate lyase
4616	GC_00002478	COG1082	Sugar phosphate isomerase/epimerase
4676	GC_00002538	COG1082	Sugar phosphate isomerase/epimerase
4392	GC_00002254	COG1172	Ribose/xylose/arabinose/galactoside ABC-type transport system, permease component
4204	GC_00002066	COG2814	Predicted arabinose efflux permease, MFS family
4615	GC_00002477	COG4154	L-fucose mutarotase/ribose pyranase, RbsD/FucU family

Unique ID	Gene Cluster ID	COG Function Accession	COG Function
4151	GC_00002013	COG1216	Glycosyltransferase, GT2 family
4369	GC_00002231	COG1216	Glycosyltransferase, GT2 family
4429	GC_00002291	COG1216	Glycosyltransferase, GT2 family
4568	GC_00002430	COG1929	Glycerate kinase
4575	GC_00002437	COG0738	Fucose permease
4160	GC_00002022	COG1879	ABC-type sugar transport system, periplasmic component, contains N-terminal xre family HTH domain
4640	GC_00002502	COG1129	ABC-type sugar transport system, ATPase component
4451	GC_00002313	COG0363	6-phosphogluconolactonase/Glucosamine-6-phosphate isomerase/deaminase
4297	GC_00002159	COG1940	Sugar kinase of the NBD/HSP70 family, may contain an N-terminal HTH domain
4173	GC_00002035	COG2610	H+/gluconate symporter or related permease
4213	GC_00002075	COG2226	Ubiquinone/menaquinone biosynthesis C-methylase UbiE
4281	GC_00002143	COG2978	p-Aminobenzoyl-glutamate transporter AbgT
4306	GC_00002168	COG0314	Molybdopterin synthase catalytic subunit
4471	GC_00002333	COG1977	Molybdopterin converting factor, small subunit
4292	GC_00002154	COG0303	Molybdopterin biosynthesis enzyme

Unique ID	Gene Cluster ID	COG Function Accession	COG Function
		COG0315|C	Molybdenum cofactor biosynthesis enzyme|Molybdopterin
4747	GC_00002609	OG0521	biosynthesis enzyme MoaB
4735	GC_00002597	COG2896	Molybdenum cofactor biosynthesis enzyme MoaA
			NADPH-dependent 2,4-dienoyl-CoA reductase, sulfur reductase,
4448	GC_00002310	COG0446	or a related oxidoreductase
4504	GC_00002366	COG2267	Lysophospholipase, alpha-beta hydrolase superfamily
4307	GC_00002169	COG1024	Enoyl-CoA hydratase/carnithine racemase
4591	GC_00002453	COG1024	Enoyl-CoA hydratase/carnithine racemase
			Crotonobetainyl-CoA:carnitine CoA-transferase CaiB and related
4701	GC_00002563	COG1804	acyl-CoA transferases
			Acyl-CoA dehydrogenase related to the alkylation response
1483	GC_00000649	COG1960	protein AidB
			Acyl-CoA dehydrogenase related to the alkylation response
1484	GC_00000649	COG1960	protein AidB
			Acyl-CoA dehydrogenase related to the alkylation response
4104	GC_00001966	COG1960	protein AidB
			Acyl-CoA dehydrogenase related to the alkylation response
4336	GC_00002198	COG1960	protein AidB

Unique ID	Gene Cluster ID	COG Function Accession	COG Function
4400	GC_00002262	COG1960	Acyl-CoA dehydrogenase related to the alkylation response protein AidB
4433	GC_00002295	COG2030	Acyl dehydratase
4570	GC_00002432	COG2030	Acyl dehydratase
4644	GC_00002506	COG2057	Acyl CoA:acetate/3-ketoacid CoA transferase, beta subunit
4468	GC_00002330	COG1788	Acyl CoA:acetate/3-ketoacid CoA transferase, alpha subunit
4176	GC_00002038	COG0183	Acetyl-CoA acetyltransferase
4249	GC_00002111	COG0183	Acetyl-CoA acetyltransferase
4727	GC_00002589	COG0183	Acetyl-CoA acetyltransferase
4290	GC_00002152	COG0318	Acyl-CoA synthetase (AMP-forming)/AMP-acid ligase II
4298	GC_00002160	COG0318	Acyl-CoA synthetase (AMP-forming)/AMP-acid ligase II
4431	GC_00002293	COG0318	Acyl-CoA synthetase (AMP-forming)/AMP-acid ligase II
4551	GC_00002413	COG0599|COG2267	Uncharacterized conserved protein YurZ, alkylhydroperoxidase/carboxymuconolactone decarboxylase family|Lysophospholipase, alpha-beta hydrolase superfamily
4200	GC_00002062	COG1028	NAD(P)-dependent dehydrogenase, short-chain alcohol dehydrogenase family
4515	GC_00002377	COG1028	NAD(P)-dependent dehydrogenase, short-chain alcohol dehydrogenase family

Unique ID	Gene Cluster ID	COG Function Accession	COG Function
4726	GC_00002588	COG1028	NAD(P)-dependent dehydrogenase, short-chain alcohol dehydrogenase family
4439	GC_00002301	COG0456	Ribosomal protein S18 acetylase RimI and related acetyltransferases
4630	GC_00002492	COG0456	Ribosomal protein S18 acetylase RimI and related acetyltransferases
4402	GC_00002264	COG0684	Regulator of RNase E activity RraA
4601	GC_00002463	COG0215	Cysteinyl-tRNA synthetase
4195	GC_00002057	COG1514	2'-5' RNA ligase
4273	GC_00002135	COG1514	2'-5' RNA ligase
4403	GC_00002265	COG0030	16S rRNA A1518 and A1519 N6-dimethyltransferase RsmA/KsgA/DIM1 (may also have DNA glycosylase/AP lyase activity)
4139	GC_00002001	COG0456|COG0510	Ribosomal protein S18 acetylase RimI and related acetyltransferases|Thiamine kinase and related kinases
4216	GC_00002078	COG1670	Protein N-acetyltransferase, RimJ/RimL family
4491	GC_00002353	COG1670	Protein N-acetyltransferase, RimJ/RimL family
4334	GC_00002196	COG4977	Transcriptional regulator GlxA family, contains an amidase domain and an AraC-type DNA-binding HTH domain

Unique ID	Gene Cluster ID	COG Function Accession	COG Function
4198	GC_00002060	COG2865	Predicted transcriptional regulator, contains HTH domain
4327	GC_00002189	COG1321	Mn-dependent transcriptional regulator, DtxR family
4548	GC_00002410	COG1321	Mn-dependent transcriptional regulator, DtxR family
4477	GC_00002339	COG2508	DNA-binding transcriptional regulator, PucR family
4203	GC_00002065	COG0789	DNA-binding transcriptional regulator, MerR family
4389	GC_00002251	COG0789	DNA-binding transcriptional regulator, MerR family
4202	GC_00002064	COG1846	DNA-binding transcriptional regulator, MarR family
4253	GC_00002115	COG1846	DNA-binding transcriptional regulator, MarR family
4424	GC_00002286	COG1846	DNA-binding transcriptional regulator, MarR family
4354	GC_00002216	COG0583	DNA-binding transcriptional regulator, LysR family
4385	GC_00002247	COG0583	DNA-binding transcriptional regulator, LysR family
4599	GC_00002461	COG0583	DNA-binding transcriptional regulator, LysR family
4330	GC_00002192	COG1414	DNA-binding transcriptional regulator, IclR family
4376	GC_00002238	COG1414	DNA-binding transcriptional regulator, IclR family
4561	GC_00002423	COG1414	DNA-binding transcriptional regulator, IclR family
4472	GC_00002334	COG1733	DNA-binding transcriptional regulator, HxlR family
4614	GC_00002476	COG1802	DNA-binding transcriptional regulator, GntR family
4148	GC_00002010	COG1937	DNA-binding transcriptional regulator, FrmR family
4579	GC_00002441	COG2186	DNA-binding transcriptional regulator, FadR family

Unique ID	Gene Cluster ID	COG Function Accession	COG Function
4681	GC_00002543	COG0640	DNA-binding transcriptional regulator, ArsR family
4441	GC_00002303	COG1309	DNA-binding transcriptional regulator, AcrR family
4554	GC_00002416	COG1309	DNA-binding transcriptional regulator, AcrR family
4686	GC_00002548	COG1309	DNA-binding transcriptional regulator, AcrR family
4325	GC_00002187	COG3226	DNA-binding transcriptional regulator YbjK
2163	GC_00000989	COG2207	AraC-type DNA-binding domain and AraC-containing proteins
2164	GC_00000989	COG2207	AraC-type DNA-binding domain and AraC-containing proteins
4522	GC_00002384	COG2852	Very-short-patch-repair endonuclease
4119	GC_00001981	COG0338	Site-specific DNA-adenine methylase
4414	GC_00002276	COG0338	Site-specific DNA-adenine methylase
4101	GC_00001963	COG1961	Site-specific DNA recombinase related to the DNA invertase Pin
4717	GC_00002579	COG1961	Site-specific DNA recombinase related to the DNA invertase Pin
4126	GC_00001988	COG1204	Replicative superfamily II helicase
4299	GC_00002161	COG1204	Replicative superfamily II helicase
4528	GC_00002390	COG1195|COG3593	Recombinational DNA repair ATPase RecF|Predicted ATP-dependent endonuclease of the OLD family, contains P-loop ATPase and TOPRIM domains
4384	GC_00002246	COG0758	Predicted Rossmann fold nucleotide-binding protein DprA/Smf involved in DNA uptake

Unique ID	Gene Cluster ID	COG Function Accession	COG Function
		COG0178|C	Excinuclease UvrABC ATPase subunit|Excinuclease UvrABC
4738	GC_00002600	OG0178	ATPase subunit
4323	GC_00002185	COG2356	Endonuclease I
4379	GC_00002241	COG1484	DNA replication protein DnaC
4239	GC_00002101	COG2189	Adenine specific DNA methylase Mod
4322	GC_00002184	COG2189	Adenine specific DNA methylase Mod
4602	GC_00002464	COG0122	3-methyladenine DNA glycosylase/8-oxoguanine DNA glycosylase
4240	GC_00002102	COG1061|C OG3886	Superfamily II DNA or RNA helicase|HKD family nuclease
4263	GC_00002125	COG1061|C OG3886	Superfamily II DNA or RNA helicase|HKD family nuclease
4672	GC_00002534	COG1619	Muramoyltetrapeptide carboxypeptidase LdcA (peptidoglycan recycling)
4133	GC_00001995	COG0438	Glycosyltransferase involved in cell wall bisynthesis
4178	GC_00002040	COG0438	Glycosyltransferase involved in cell wall bisynthesis
4501	GC_00002363	COG0438	Glycosyltransferase involved in cell wall bisynthesis
4603	GC_00002465	COG0438	Glycosyltransferase involved in cell wall bisynthesis
4641	GC_00002503	COG0438	Glycosyltransferase involved in cell wall bisynthesis

Unique ID	Gene Cluster ID	COG Function Accession	COG Function
4748	GC_00002610	COG1209	dTDP-glucose pyrophosphorylase
4495	GC_00002357	COG1088	dTDP-D-glucose 4,6-dehydratase
		COG1091|C	dTDP-4-dehydrorhamnose reductase|dTDP-4-dehydrorhamnose
4149	GC_00002011	OG1898	3,5-epimerase or related enzyme
4319	GC_00002181	COG1083	CMP-N-acetylneuraminic acid synthetase
4512	GC_00002374	COG1292	Choline-glycine betaine transporter
4426	GC_00002288	COG0791	Cell wall-associated hydrolase, NlpC family
		COG0438|C	
		OG1196|CO	Glycosyltransferase involved in cell wall bisynthesis|Chromosome
4105	GC_00001967	G4641	segregation ATPase|Spore maturation protein CgeB
		COG0489|C	Chromosome partitioning ATPase, Mrp family, contains Fe-S
4421	GC_00002283	OG3944	cluster|Capsular polysaccharide biosynthesis protein
			Peptidoglycan/LPS O-acetylase OafA/YrhL, contains
		COG1835|C	acyltransferase and SGNH-hydrolase
4467	GC_00002329	OG2755	domains|Lysophospholipase L1 or related esterase
		COG0517|C	
		OG1082|CO	CBS domain|Sugar phosphate isomerase/epimerase|Sialic acid
4111	GC_00001973	G2089	synthase SpsE, contains C-terminal SAF domain

Unique ID	Gene Cluster ID	COG Function Accession	COG Function
4700	GC_00002562	COG4948	L-alanine-DL-glutamate epimerase or related enzyme of enolase superfamily
4311	GC_00002173	COG1192	Cellulose biosynthesis protein BcsQ
4404	GC_00002266	COG1192	Cellulose biosynthesis protein BcsQ
4147	GC_00002009	COG1651	Protein-disulfide isomerase
4371	GC_00002233	COG0229	Peptide methionine sulfoxide reductase MsrB
2533	GC_00001174	COG0225	Peptide methionine sulfoxide reductase MsrA
2534	GC_00001174	COG0225	Peptide methionine sulfoxide reductase MsrA
4488	GC_00002350	COG0542	ATP-dependent Clp protease ATP-binding subunit ClpA
4344	GC_00002206	COG0428	Zinc transporter ZupT
4122	GC_00001984	COG4264	Siderophore synthetase component
4605	GC_00002467	COG2072	Predicted flavoprotein CzcO associated with the cation diffusion facilitator CzcD
4312	GC_00002174	COG2375	NADPH-dependent ferric siderophore reductase, contains FAD-binding and SIP domains
4348	GC_00002210	COG1914	Mn2+ and Fe2+ transporters of the NRAMP family
1473	GC_00000644	COG0053	Divalent metal cation (Fe/Co/Zn/Cd) transporter
1474	GC_00000644	COG0053	Divalent metal cation (Fe/Co/Zn/Cd) transporter
4524	GC_00002386	COG1230	Co/Zn/Cd efflux system component

Unique ID	Gene Cluster ID	COG Function Accession	COG Function
4679	GC_00002541	COG1230	Co/Zn/Cd efflux system component
4691	GC_00002553	COG1230	Co/Zn/Cd efflux system component
4652	GC_00002514	COG0798	Arsenite efflux pump ArsB, ACR3 family
4690	GC_00002552	COG0803	ABC-type Zn uptake system ZnuABC, Zn-binding component ZnuA
4091	GC_00001953	COG4149	ABC-type molybdate transport system, permease component
4541	GC_00002403	COG0725	ABC-type molybdate transport system, periplasmic component
4272	GC_00002134	COG1108	ABC-type Mn2+/Zn2+ transport system, permease component
4508	GC_00002370	COG1121	ABC-type Mn2+/Zn2+ transport system, ATPase component
4107	GC_00001969	COG0477	MFS family permease
4131	GC_00001993	COG0477	MFS family permease
4740	GC_00002602	COG0477	MFS family permease
4751	GC_00002613	COG0477	MFS family permease
4752	GC_00002614	COG0477	MFS family permease
4123	GC_00001985	COG0476|C OG0607	Molybdopterin or thiamine biosynthesis adenylyltransferase|Rhodanese-related sulfurtransferase
4437	GC_00002299	COG1321|C OG1918	Mn-dependent transcriptional regulator, DtxR family|Fe2+ transport system protein FeoA
4117	GC_00001979	COG0783	DNA-binding ferritin-like protein (oxidative damage protectant)

Unique ID	Gene Cluster ID	COG Function Accession	COG Function
4316	GC_00002178	COG3486	Lysine/ornithine N-monooxygenase
4146	GC_00002008	COG2366	Acyl-homoserine lactone (AHL) acylase PvdQ
4742	GC_00002604	COG0702	Uncharacterized conserved protein YbjT, contains NAD(P)-binding and DUF2867 domains
4419	GC_00002281	COG4127	Predicted restriction endonuclease, Mrr-cat superfamily
4698	GC_00002560	COG0667	Predicted oxidoreductase (related to aryl-alcohol dehydrogenase)
4754	GC_00002616	COG0667	Predicted oxidoreductase (related to aryl-alcohol dehydrogenase)
4355	GC_00002217	COG4122	Predicted O-methyltransferase YrrM
4581	GC_00002443	COG1569	Predicted nucleic acid-binding protein, contains PIN domain
4324	GC_00002186	COG3618	Predicted metal-dependent hydrolase, TIM-barrel fold
4559	GC_00002421	COG2159	Predicted metal-dependent hydrolase, TIM-barrel fold
4664	GC_00002526	COG0645	Predicted kinase
4380	GC_00002242	COG0673	Predicted dehydrogenase
4473	GC_00002335	COG0673	Predicted dehydrogenase
4377	GC_00002239	COG2388	Predicted acetyltransferase, GNAT superfamily
4398	GC_00002260	COG1473	Metal-dependent amidase/aminoacylase/carboxypeptidase
4372	GC_00002234	COG1253	Hemolysin or related protein, contains CBS domains
4502	GC_00002364	COG1253	Hemolysin or related protein, contains CBS domains
4351	GC_00002213	COG3391	DNA-binding beta-propeller fold protein YncE

Unique ID	Gene Cluster ID	COG Function Accession	COG Function
4208	GC_00002070	COG1106	ATPase/GTPase, AAA15 family
4180	GC_00002042	COG2141	Flavin-dependent oxidoreductase, luciferase family (includes alkanesulfonate monooxygenase SsuD and methylene tetrahydromethanopterin reductase)
4409	GC_00002271	COG2141	Flavin-dependent oxidoreductase, luciferase family (includes alkanesulfonate monooxygenase SsuD and methylene tetrahydromethanopterin reductase)
4250	GC_00002112	COG0596|COG0627	Pimeloyl-ACP methyl ester carboxylesterase|S-formylglutathione hydrolase FrmB
4333	GC_00002195	COG5519	Uncharcterized protein, DUF927 family
4713	GC_00002575	COG4683	Uncharacterized protein
4744	GC_00002606	COG4748	Uncharacterized protein
4532	GC_00002394	COG3238|OG3238	Uncharacterized membrane protein YdcZ, DUF606 family|Uncharacterized membrane protein YdcZ, DUF606 family
4406	GC_00002268	COG4763	Uncharacterized membrane protein YcfT
4184	GC_00002046	COG4243	Uncharacterized membrane protein
4613	GC_00002475	COG3759	Uncharacterized membrane protein
4209	GC_00002071	COG2318	Uncharacterized damage-inducible protein DinB (forms a four-helix bundle)

Unique ID	Gene Cluster ID	COG Function Accession	COG Function
4578	GC_00002440	COG1479	Uncharacterized conserved protein, contains ParB-like and HNH nuclease domains
4342	GC_00002204	COG1652|COG1652|COG3629	Nucleoid-associated protein YgaU, contains BON and LysM domains|Nucleoid-associated protein YgaU, contains BON and LysM domains|DNA-binding transcriptional activator of the SARP family
4639	GC_00002501	COG4585	Signal transduction histidine kinase
4411	GC_00002273	COG0394	Protein-tyrosine-phosphatase
4704	GC_00002566	COG3451	Type IV secretory pathway, VirB4 component
4368	GC_00002230	COG1989	Prepilin signal peptidase PulO (type II secretory pathway) or related peptidase
4099	GC_00001961	COG4962	Pilus assembly protein, ATPase of CpaF family
4145	GC_00002007	COG4961	Flp pilus assembly protein TadG
4183	GC_00002045	COG4965	Flp pilus assembly protein TadB
4197	GC_00002059	COG1002	Type II restriction/modification system, DNA methylase subunit YeeA
4284	GC_00002146	COG3587|COG3587	Restriction endonuclease|Restriction endonuclease
4233	GC_00002095	COG2076	Multidrug transporter EmrE and related cation transporters

Unique ID	Gene Cluster ID	COG Function Accession	COG Function
4721	GC_00002583	COG2076	Multidrug transporter EmrE and related cation transporters
4585	GC_00002447	COG0251	Enamine deaminase RidA, house cleaning of reactive enamine intermediates, YjgF/YER057c/UK114 family
4408	GC_00002270	COG1132	ABC-type multidrug transport system, ATPase and permease component
118	GC_00000009	COG2801	Transposase InsO and inactivated derivatives
119	GC_00000009	COG2801	Transposase InsO and inactivated derivatives
120	GC_00000009	COG2801	Transposase InsO and inactivated derivatives
121	GC_00000009	COG2801	Transposase InsO and inactivated derivatives
122	GC_00000009	COG2801	Transposase InsO and inactivated derivatives
123	GC_00000009	COG2801	Transposase InsO and inactivated derivatives
4268	GC_00002130	COG2801	Transposase InsO and inactivated derivatives
4381	GC_00002243	COG2801	Transposase InsO and inactivated derivatives
4387	GC_00002249	COG2801	Transposase InsO and inactivated derivatives
4620	GC_00002482	COG2801	Transposase InsO and inactivated derivatives
4707	GC_00002569	COG2801	Transposase InsO and inactivated derivatives
4756	GC_00002618	COG2963|COG3547	Transposase and inactivated derivatives|Transposase
4582	GC_00002444	COG2963	Transposase and inactivated derivatives

Unique ID	Gene Cluster ID	COG Function Accession	COG Function
146	GC_00000014	COG3328	Transposase (or an inactivated derivative)
3074	GC_00001444	COG3547	Transposase
4496	GC_00002358	COG4584	Transposase
4503	GC_00002365	COG0675	Transposase
4516	GC_00002378	COG4584	Transposase
4712	GC_00002574	COG3378	Phage- or plasmid-associated DNA primase
4646	GC_00002508	COG3311	Predicted DNA-binding transcriptional regulator AlpA
4110	GC_00001972	COG0582	Integrase
4552	GC_00002414	COG0582	Integrase

Table 2.9. Micro36 sequences in post-natal infant cohorts

SampleID	Query Sequence	Count	Similiarity (%)	Length (bp)	Mismached	Gaps	Participant	Age of Stool (months)	Sex	Cohort
A219	Micro36	1	99.21	253	2	0	A219	1	Male	WHEALS
DB23M00	Micro36	37	99.09865	253	2.3	0	DB23	0	Male	DIMES
DB29M00	Micro36	1	100	253	0	0	DB29	0	Male	DIMES
B10M06	Micro36	3	99.34	253	1.67	0	PB10	6	Female	DIMES
DB11M06	Micro36	23	99.67261	253	0.83	0	DB11	6	Female	DIMES
A233	Micro36	86	98.23488	253	4.44	0.02	A233	1	Male	WHEALS
A295	Micro36	4	97.6275	253	6	0	A295	6	Male	WHEALS
ERR756407	Micro36	3	97.81	137	3	0	7-048261-1-4_2	3	Male	CHILD
ERR756501	Micro36	1	97.76	134	3	0	7-096066-1-4_2	3	Female	CHILD
ERR756508	Micro36	1	97.79	136	3	0	7-108700-1-4_2	3	Female	CHILD
ERR756880	Micro36	5	97.644	135.8	3.2	0	7-193789_2	12	Male	CHILD
ERR756922	Micro36	1	97.12	139	2	1	7-208786-1-4_2	3	Female	CHILD
ERR756996	Micro36	1	97.81	137	3	0	7-236436-1-3_2	3	Male	CHILD

CHAPTER 3: MATERIALS AND METHODS

Material for this chapter was modified from:

Rackaityte E, Halkias J, Fukui EM, Mendoza VF, Hayzelden C, Crawford ED, Burt TD, Lynch SV (2020). Viable bacterial colonization is limited in the human intestine *in utero*. *Nature Medicine*. *26*(4), 599-607.

3.1 Human Samples and Consent

Donated human fetal tissue (small intestine, mesenteric lymph node, spleen) was obtained under the auspices of UCSF Committee on Human Research (CHR) approved protocols after written informed consent at 20 ± 2.2 gestational weeks from the Department of Obstetrics, Gynecology and Reproductive Science at San Francisco General Hospital from terminated pregnancies. Exclusion criteria were: (1) known maternal or intrauterine infection, (2) intrauterine fetal demise, and/or (3) known or suspected chromosomal abnormality[6]. No Human Patient Information (HPI) is associated with the data presented. All sample collection methods comply with the Helsinki Declaration principles. Samples were transported in media on ice and processed within 2 hours after collection.

3.2 Sample Collection for Fetal Meconium Cohort

Uninterrupted stomach to caecum sections (fetal intestine), kidneys, spleens, and mesenteric lymph nodes were collected by a single operator using sterile tools within 10 minutes of termination procedure and placed into sterile containers with pre-aliquoted complete RPMI (cRPMI) media composed of: RPMI media (GIBCO) without antibiotics, 10% fetal bovine serum (GIBCO), 1 mM sodium pyruvate (Life Technologies), 2 mM L-glutamine (Life Technologies), 1 x non-essential amino acids (Life Technologies), and 10 mM HEPES (Life Technologies). Sterile cotton swabs were pre-moistened with sterile 1 x phosphate-buffered saline (PBS) and stored in containers until used to vigorously sample the surgical tray for 30 seconds, thus sampling both the hospital environment and any contaminants arising from the procedure; swabs were immediately

snapped off into sterile tubes containing 500 μL of pre-aliquoted, sterile RNAlater. Blank swabs were prepared as described above, but immediately snapped off into RNAlater, without sampling the surgical tray. Air swabs were prepared as described above, but held in surgical room air for 30 seconds, before immediately being snapped off into RNAlater. All specimens were immediately placed on ice and transported to the laboratory. Intestinal sections were dissected to remove the mesentery and the muscularis in a sterile petri dish in a biosafety laminar flow cabinet. Separate sterile tools were used to divide the small intestine into three equal sections and new sterile tools were used to scrape internal contents, termed fetal meconium, of each section into sterile 1 x PBS (**Figure 2.3**). Fetal meconium was homogenized by vigorous pipetting in sterile 1 x PBS, pelleted by centrifugation at 3000 x g for 10 minutes, and re-suspended in 1 mL of sterile 1 x PBS. Half of fetal meconium suspension (by volume) was added to RNAlater (Ambion), while the remainder was re-suspended in sterile 50% (v/v) glycerol. Sterile tools were used to remove kidney capsule of the fetal kidney in a sterile petri and separate sterile tools were used to biopsy the internal kidney tissue, which was immediately placed in RNAlater. Fetal meconium samples, kidney specimens, procedural swabs, and blank swabs were cryopreserved at -80 °C, within 2 hours of the termination procedure. Additional splenic and intestinal samples were collected in the manner described above for *ex vitro* APC and T cell experiments. In total 77 fetal specimens were used in this study.

3.3 16S rRNA Gene Burden and Sequencing

3.3.1 DNA extraction

Genomic DNA (gDNA) from fetal meconium samples, kidney specimens, procedural swabs, and blank swabs was extracted using a modified cetyltrimethylammonium bromide (CTAB)-buffer-based protocol exactly as previously described[24] along with buffer controls. Buffers were prepared using HPLC-grade chemicals in a BSL2 biosafety cabinet and autoclaved before use.

3.3.2 16S rRNA gene burden qPCR analysis

16S rRNA gene copy number was assessed by quantitative PCR (Q-PCR) using the 16S rRNA universal primers and TaqMan probes, as previously described[221]. Briefly, total 16S rRNA gene copy number was calculated against a standard curve of known 16S rRNA copy numbers (1×10^2–1×10^9). Q-PCR was performed in triplicate 20 μl reactions containing final concentrations of 1 ×TaqMan Universal Master Mix (Life Technologies), 100 ng of extracted genomic DNA, 900 nM of each primer, P891F (5'-seq-3'F) and P1033R (5'-seq-3'R) and 125 nM of UniProbe under the following conditions: 50 °C for 2 min, 95 °C for 10 min, followed by 40 cycles of denaturation at 95 °C for 15 s, and annealing and extension at 60 °C for 1 min, along with no-template control and 8 standards. Copy number was normalized either by 100ng of input DNA, when possible. When too little DNA was obtained, such as in the case of the buffers, 10μL of DNA extract was added to the PCR reaction and copy number was normalized by weight of frozen sample.

3.3.3 Depletion of Abundant Sequences by Hybridization (DASH)

Depletion of human 16S mitochondrial DNA (mtDNA) using single guide RNA (sgRNA) targeting of Cas9 was performed as previously described[222]. Briefly, 54 sgRNAs targeting the human mtDNA were transcribed from pooled sgRNA templates using custom T7 RNA polymerase generously provided by the DeRisi laboratory at UCSF. sgRNAs were purified and concentrated using a column-based RNA purification kit with DNAse treatment (Zymo) and incubated with purified Cas9 (Berkeley Macrolab) for 10 minutes at 37°C. sgRNA-loaded Cas9 was incubated with either meconium genomic DNA (gDNA) or pooled library of 16S rRNA V4 amplicon (see below) for 2 hours at 37°C. Cas9 was deactivated by boiling the *in vitro* reaction at 98°C for 10 minutes and Ampure XP beads (Agencourt) were used to purify the amplicon DNA. To test the effects of DASH on bacterial community composition, a subset of meconium samples from our bank (n=10) was depleted of mtDNA either from individual meconium gDNA (individual DASH) prior to 30-cycle amplification or from the pooled library of 30-cycle amplicons (pooled DASH). DASH bacterial profiles were compared to 30-cycle or 35-cycle amplicons that were depleted of mtDNA by gel extraction, using a gel extraction kit (Quiagen). For sequencing of the entire bank of fetal meconium gDNA, individual DASH was implemented on all samples including buffer blanks and contamination swabs.

3.3.4 Sequencing preparation

The V4 region of the depleted genomic DNA was amplified using primers designed by Caporaso *et al*[223] using PCR conditions and protocol as described in Fujimura *et al*[24].

Briefly, samples were amplified in heptuplicate from a single mastermix per template, aliquoted into 384-well plates, and included a negative control reaction for each template mastermix and each reverse barcoded primer. PCR reactions were performed in 25□L volumes using 0.025 U Takara Hot Start ExTaq (Takara Mirus Bio Inc.), 1X Takara buffer with $MgCl_2$, 0.4 pmol μl^{-1} of F515 and barcoded R806 primers, 0.56 mg/ml of bovine serum albumin (BSA; Roche Applied Science), 200 μM of dNTPs and 10 ng of DASH gDNA. PCR conditions were: initial denaturation (98 °C, 2 min), 30 cycles of 98 °C (20 s), annealing at 50 °C (30 s), extension at 72 °C (45 s) and final extension at 72 °C (10 min), except in validation of DASH protocol (see above), where 35 cycles of amplification were also used. Amplicons were pooled and verified using a 2% TBE agarose e-gel (Life Technologies), purified using AMPure SPRI beads (Beckman Coulter), quality checked using Bioanalyzer DNA 1000 Kit (Agilent) and quantified using the Qubit 2.0 Fluorometer and the dsDNA HS Assay Kit (Life Technologies). Amplicons were pooled at equimolar amounts to create the sequencing library, with the exception of buffer controls, which did not yield enough amplicon and were pooled at the average volume. A mock community (BEI Resources HM-277D) composed of equal genomic concentration of bacterial genomic DNA was sequenced for each amplification plate to monitor and standardize data between amplification plates. Denatured libraries were diluted to 2 nM and were loaded onto the Illumina MiSeq cartridge at 5 pM with 15% (v/v) denatured 12.5 pM PhiX spike-in for sequencing. Complete fetal meconium bank of samples was sequenced on one 250 x 250 base pair Illumina MiSeq run.

3.3.5 Sequence data processing and quality control

Paired-end reads were assembled using FLASH v1.2.11[224] requiring a minimum base pair overlap of 200 and de-multiplexed by barcode using QIIME (Quantitative Insights Into Microbial Ecology, v1.9.1)[225]. Quality filtering was accomplished using USEARCH v8.0.1623 to remove reads with >2 expected errors[226]. Quality reads were de-replicated at 100% sequence identity, clustered at 97% sequence identity into operational taxonomic units (OTUs), filtered of chimeric sequences, and mapped back to resulting OTUs using USEARCH. Taxonomy was assigned to the OTUs using SILVA database.

3.3.6 Fetal meconium data analysis

OTUs detected in greater than 50% of extraction buffer, blank swab, and air swab controls were removed from all samples prior to further filtering. OTUs comprising fewer than 5 reads and fewer than 0.0001% of the total read counts across all samples were removed. Additional buffer contaminants were identified using *decontam* package[182] in R. Resulting sequence reads were normalized by multiply rarefying to 1,000 reads per sample as previously described, to assure reduced data were representative of the fuller data for each sample[24]. Dominant taxa were identified for each rarefied sample by determining the OTU with the greatest number of reads per sample.

Post-natal meconium data analysis. 16S rRNA gene V4 amplicon sequencing profiles of meconium collected at birth was obtained from the European Nucleotide Archive (ENA) under accession number PRJEB20766 and post-processed as described above for fetal meconium. OTUs were re-picked with combined fetal and post-natal meconium datasets combined. Infant stool samples with high identify to fetal isolates were

identified by first trimming the appropriate variable region (depending on study) from full-length 16S rRNA gene Micro36 sequences. These sequences were then aligned using BLASTn to publicly available infant stool cohorts[24,167,211] with accession numbers PRJEB13896, PRJEB20766, PRJEB8463; sequences with >97% identity and >99% coverage were identified.

3. 4 Immune Cell Isolation

Uninterrupted stomach to caecum sections of the fetal small intestine were dissected in cold 1x PBS (see above). The intestine was cut into 1cm sections and washed three times with 1mM DTT in 1x PBS for 10 minutes at 37°C to remove mucus. The epithelial layer was dissociated with three washes of 1mM EDTA in 1x PBS for 20 minutes at 37 °C and the latter wash was preserved in RNAlater (Ambion) at -80°C for RNAseq. The remaining lamina propria cells were dissociated with freshly prepared 1mg/mL Collagenase IV (Gibco) and 10mg mL^{-1} DNAse (Roche) in cRPMI for 30 minutes at 37°C, in a shaking water bath at 200 rpm. Mesenteric lymph node and spleen cells were isolated by a 30-minute digestion in Collagenase IV media as described above and then gently pressed through a 70□m strainer. Cells were separated in a 20%-40%-80% Percoll density gradient at 400 x g for 40 minutes: T cells were recovered at the 40-80% interface, while antigen presenting cells were recovered at the 20-40% interface. All cells were washed twice with cRPMI media. Viability was measured with propidium iodide (Sigma Aldrich) and AQUA dye (Invitrogen) using flow cytometry.

3.5 Epithelial Cell RNA Sequencing

Cryopreserved epithelial cell layers (in RNAlater, Ambion) were lysed using QIAshredder (QIAGEN) columns and RNA was extracted using RNAqueous kit (ThermoFisher). RNA was quantified using Qubit RNA HS Assay (ThermoFisher), normalized, and converted to cDNA using SMARTer cDNA Synthesis Kit (Takara Bio) using 7 cycles of amplification. RNA and cDNA quality was determined by Bioanalyzer (Agilent). cDNA was fragmented, ligated with Illumina adapters using Nextera XT kit (Illumina), following manufacturer's instructions, and sequenced on NovaSeq6000 sequencer using two lanes. Paired-end 100 by 100 bp reads were obtained, demultiplexed, quality filtered, removed of Illumina adapters using TrimGalore (github.com/FelixKrueger/TrimGalore), and aligned to the human genome (Hg38 release) using STAR[227] with ENCODE recommended parameters. Features were assigned to transcripts using featureCounts[228], normalized using DESEQ2[229]. Differential expression was evaluated using DESEQ2 genes with at least 20 reads per gene in respective sample grouping. Log-normalized read counts were obtained from DESEQ2 package, genes were filtered for presence in 75% of samples per comparison group, top variable genes were identified by the coefficient of variance and used to calculate principal components of Euclidean distances.

3.6 Fluorescence *In Situ* Hybridization

Murine and human fetal terminal ileum was fixed in Carnoy fixative to preserve the mucous layer[230], embedded in Tissue-Tek OCT (VWR) medium, and cryosectioned to 5 □m sections using a cryostat. Sections were thawed, were post-fixed with acetone for

15 minutes, and rinsed with 1x PBS. Slides were incubated with sterile-filtered 100□L of probe solution containing 35% formamide, as previously described[230]. Hybridizations were performed for 10 hours at 48°C, followed by a washing step for one hour at the same temperature, as previously described[230]. Hybridization probes were utilized at 0.5 □M final concentration and included fluorescently-labeled oligos eubacterial (EUB) /5Cy3/GC TGC CTC CCG TAG GAG T/3Cy3Sp/[231] or non-targeting (NEUB) /5Cy3/AC TCC TAC GGG AGG CAG C/3Cy3Sp/[231]. Slides were mounted in Vectashield with DAPI (Vector Laboratories) and imaged at 400x and 1000x magnification using epifluorescence Keyence Microscope BZ-X700. Quantification of images was performed in ImageJ software using the set scale function to calibrate pixels to □m units, freehand selection tool was used to trace the perimeter of each villi, and tracing lengths were measured and summed for each section. The point tool was used to manually count EUB or NEUB signal.

3.7 Electron Microscopy

Terminal ileum of fetal intestines was dissected and ligated with sterile suture to prevent contamination of the internal lumen. Ligated samples were immediately immersed in 2.5% (v/v) electron microscopy (EM) grade glutaraldehyde fixative (Sigma Aldrich) in 1x PBS solution and incubated overnight at room temperature with agitation. Samples were washed twice with 1x PBS for 15 minutes and dehydrated with a series of ethanol baths. Samples were then critical point dried (Tousimisautosamdri-815), sliced open with a clean razorblade, mounted in conductive silver epoxy (Ted Pella, Inc.), and coated with 15-30 nm of iridium (Cressington 208-HR sputter coater). Electron

micrographs were recorded using a Carl Zeiss ULTRA55 FE-SEM at accelerating voltages in the range 1.24-3.9 keV, working distances of 4.8-9.2 mm, and 20-60 μm diameter apertures with high-current mode. Post-processing of images was not performed. Specimens were stored in a vacuum chamber to avoid contamination between imaging sessions.

3.8 Bacterial Isolation

Punch biopsies were taken from three samples of cryopreserved meconium with highest read counts for *Micrococcus* using a sterile surgical punch biopsy tool (Integra Miltex, Plainsboro, NJ) in clean biosafety cabinet. Three independent fetal meconium samples were used for isolation. Punch biopsies of *Micrococcus* enriched meconium were incubated in antibiotic-free cRPMI with or without 2 x10^6 THP1 human monocyte feeder cells for 48 hours at 37 °C in ambient atmospheric stationary conditions. Single colonies were isolated after transfer to brain heart infusion (BHI; TekNova) agar plates and single colonies were picked. Colony sequencing (Quintara Biosciences) was performed using the full length 16S rRNA gene using primer pairs 27F (5'-seq-3') and 1492R (5'-seq-3')[232]. Full-length gene was assembled using Clustal Omega and taxonomy was determined by SINA[233] against the curated SILVA database. Reference strains were obtained from American Type Culture Collection for *Micrococcus luteus* (MicroRef1, ATCC 4698; MicroRef2 ATCC 12698) and grown by ATCC's protocol.

3.9 Bacterial Whole Genome Sequencing and Comparative Genomics

3.9.1 Whole genome sequencing and assembly

Twenty-four-hour cultures of Micro36 were obtained in media and culture conditions as described above, and DNA was extracted using CTAB-based protocol as described above. Genomic DNA (gDNA) was fragmented and Illumina adapters were ligated using Nextera XT (Illumina) kit following manufacturer's instructions. gDNA library quality was verified by gel-electrophoresis Bioanalyzer (Agilent) and was sequenced on Illumina MiSeq using a MiSeq Reagent Kit v3 (Illumina) with 300 x 300bp paired-end reads. Reads were removed of adapters and quality filtered using TrimGalore. When possible, paired-end reads were assembled using FLASh[224] for use as a single-ended library for assembly using SPAdes[234] genome assembler. Genome assembly quality was determined by QUAST[235] and genomes were submitted NCBI Prokaryotic Genome Annotation Pipeline (PGAP). Annotation was performed locally using NCBI COG database in anvi'o package[236].

3.9.2 Comparative genomics

Micrococcus genomes were downloaded from NCBI using NCBI genome download tool (github.com/kblin/ncbi-genome-download) and imported into anvi'o pangenome analysis environment[236]. Average nucleotide identity and coverage was calculated using ANIb within *pyani* package (widdowquinn.github.io/pyani/)[237]. Single copy genes[238] were identified for all relevant genomes within anvi'o environment, aligned using MUSCLE[239], phylogenetic trees were constructed using FastTree2[240], and visualized in iTOL[241].

Post-natal data analysis. A custom *kraken2*[242] database was created by adding Micro36 genome contigs to the standard database. Maternal and infant stool and various body

site bacterial metagenomic reads[212,213] and public metadata were obtained from NCBI SRA in FASTQ format using accession numbers PRJNA475246 and PRJNA352475. Percent relative abundance of *M. luteus* per sample was obtained using *kraken2* software was used to classify metagenomic reads against the custom database using a minimum base quality threshold of 20 and a confidence threshold of 95%.

3.10 Bacterial Growth Curves

Liquid cultures of *Micrococcus* strains were grown for 24-48 hours at 37°C in BHI. Cultures were normalized to 0.05 optical density at A_{600nm} (OD_{600}) and incubated with indicated molar concentrations of progesterone (Tocris Bioscience) and 17β-estradiol (Tocris Bioscience) or equal volume of absolute ethanol vehicle (Sigma Aldrich), in respective culture media (see above). To test whether bacterial isolates were capable of growth with progesterone and 17β-estradiol as the sole carbon source, bacterial growth curves were performed in freshly prepared mineral salt media[243] supplemented with $1x10^{-5}$M progesterone and $1x10^{-6}$M 17β-estradiol or equal volume of absolute ethanol vehicle at a normalized starting OD_{600} of 0.1. Bacterial cultures were then incubated in a Cytation3 spectrophometer (BioTek) at 37°C for 35 hours, and OD_{600} was recorded every 15 minutes.

3.11 Gentamicin Protection Assay

Intracellular lifestyle of bacterial isolates was determined by gentamicin protection assays as described previously[244]. Primary human antigen presenting cells from fetal spleen were enriched by negative selection using Easy Step Human Biotin Isolation kit (STEMCELL Technologies) and biotin-conjugated mouse anti-human mAbs for CD3,

CD56, CD19, and CD20. Isolated cells were incubated for 24h in cRPMI with penicillin and streptomycin at 4°C. Fetal antigen presenting cells or RAW 264.7 macrophage cells (ATCC) were seeded in each well of a 96-well plate and incubated for two hours at 37°C 5% CO_2 with bacterial isolate overnight cultures at a multiplicity of infection (MOI) of 10. Non-adherent bacteria were removed by washing three times with 1x PBS and incubating for 30 minutes with cRPMI supplemented with 50μg mL^{-1} gentamicin. Cells were then incubated with 10μg mL^{-1} gentamicin supplemented cRPMI for 3, 24, 40, 48 or 50 hours at 37°C 5% CO_2. Intracellular bacteria were recovered by lysing eukaryotic cells with sterile 1% (v/v) Triton X (Sigma Aldrich) solution for 10 minutes, with lysis was visually confirmed by light microscope. CFUs were counted from serial dilutions of lysate, grown on either BHI (see above) agar plates for *Micrococcus* exposed cells. *Escherichia coli* strain DH10B was used as a negative control. Lysate was plated on respective media agar plates with 10μg mL^{-1} gentamicin to determine acquisition of antibiotic resistance.

3.12 Antibodies and Flow Cytometry

Extracellular staining of isolated cells was performed in 2% FBS in PBS with 1mM EDTA (staining buffer) with human Fc blocking antibody (STEMCELL Technologies) and with fluorochrome-conjugated antibodies against surface markers, as previously described[6]. Intracellular protein were detected in fixed, permeabilized cells using the Foxp3/Transcription Factor Staining Buffer set (Tonbo Biosciences), as previously described[6]. Mouse anti-human monoclonal antibodies used in this study: TCRβ PerCP Cy5.5 (Clone IP26, eBioscience Cat. No. 46-9986-42, dilution 1:100), Vα7.2 BV605

(Clone 3C10, BioLegend Cat. No. 351720, dilution 1:100), CD4 APC H7 (Clone L200, BD Pharmingen Cat. No. 560837, dilution 1:100), CD8α FITC and PE Cy7(Clone B7-1, BD Pharmingen Cat. No. 347313, dilution 1:100), CD45RA PE Cy7 (Clone HI100, BD Pharmingen Cat. No. 555489, dilution 1:100), CCR7 PE (Clone G043H7, BioLegend Cat. No. 353208, dilution 1:100), PLZF-APC (Clone 6318100, R&D Cat. No. IC2944A, dilution 1:50), CD161-BV711 (Clone DX12, BD Biosciences Cat. No. 563865, dilution 1:50), CD25 FITC (Clone 2A3, BD Biosciences Cat. No. 347643, 1:100), FoxP3 PE (Clone PCH101, eBioscience Cat. No. 12-4776-42, 1:50), IFNγ-FITC (Clone 25723.11, BD Biosciences Cat. No. 340449, 1:50), TNFα-PE Cy7 (Clone MAB11, BD Pharmingen Cat. No. 557647, 1:50), CD45 APC (Clone HI30, Tonbo Cat. No. 20-0459, 1:100), HLA-DR APC-R700 (Clone G46-6, BD Cat. No. 565127, dilution 1:100), CD3 biotin (Clone OKT3, eBioscience Cat. No. 13-0037-82, dilution 1:100), CD19 biotin (Clone HIB19, BioLegend Cat. No. 203304, dilution 1:100), CD20 biotin (Clone 2H7, eBioscience Cat. No. 13-0209-82, dilution 1:100), CD56 biotin (Clone NCAM16.2, BD Cat. No. 555515, dilution 1:100), LLT1 PE (Clone 402659 R&D Cat. No. FAB3480P, dilution 1:50)[6]. Biotin antibodies were detected with streptavidin conjugated to BV711 (BD Biosciences Cat. No. 563262, 1:200), as previously described[6]. Dead cells were excluded from analysis using Aqua LIVE/DEAD Fixable Dead Cell Stain Kit (Invitrogen) stain. All data were acquired with BD LSR/Fortessa Dual SORP using FACS Diva software (BD Biosciences) and analyzed with FlowJo (TreeStar) software.

3.13 *Ex vivo* Intestinal Epithelial Cell Transcriptomics after Bacterial Isolates Exposure

EDTA washes containing fetal intestinal epithelial cells (see above) were washed with 1x PBS, passed through 40 □m strainer, and plated on Collagen I coated 96-well plates (Corning) in cRPMI containing 5 ng mL^{-1} epidermal growth factor (Gibco). Cells were incubated overnight at 37°C 5%CO_2 4% O_2, to mimic hypoxic conditions in the fetal intestine[245] and non-adherent cells were removed. Cells were allowed to differentiate for five days or until 80% confluence, with media replacement every two days. Cells were incubated with a multiplicity of infection of 10 of bacterial isolates in cRPMI for 4 at 37°C 5%CO_2 4% O_2. After 4h, cells were preserved in RNAlater and RNA was prepared for sequencing as described above.

3.14 *Ex vivo* Antigen Presenting Cell Activation with Bacterial Isolates

Antigen presenting cells from fetal spleen were enriched by negative selection using Easy Step Human Biotin Isolation kit (STEMCELL Technologies) as described above. Cells were seeded into 96-well plates and incubated with multiplicity of infection of 10 of bacterial isolates in cRPMI for 4 hours at 37°C 5%CO_2 4% O_2, to mimic hypoxic conditions in the fetal intestine[245] and normalize for bacterial growth.

3.15 *Ex vivo* Autologous Mixed Lymphocyte Reactions

Lamina propria T cells were enriched using Easy Sep Human T cell isolation kit (STEMCELL Technologies), effector memory cells were sorted to >99% purity (**Extended Data 8i**) using BD Aria Fusion SORP, and cells were labeled with cell trace violet (Invitrogen). Splenic antigen presenting cells autologous to isolated T cells were

enriched as described above, sorted to >96% purity (**Extended Data 8j**), and exposed to bacterial isolates as described above. Bacteria were removed with three washes of cRPMI supplemented with penicillin and streptomycin. Sorted, labeled effector memory T cells were incubated with pre-exposed antigen presenting cells in a 2:1 ratio in cRPMI with supplemented with 10ng mL^{-1} IL-2 (PeproTech), 10ng mL^{-1} IL-7 (PeproTech), 2 μg mL^{-1} purified anti-CD28 (Clone CD28.2, BD Pharmigen Cat. No. 555725), 2 μg mL^{-1} purified anti-CD49d (Clone, BD Pharmingen Cat No. 555501), and 10 μg mL^{-1} gentamicin for three days at 37°C 5% CO_2 4% O_2. Cells were incubated with 10μg mL^{-1} Brefeldin A (Sigma Aldrich) in the same media for 4 hours at 37°C 5% CO_2 4% O_2 and were subsequently stained for intracellular cytokine production as described above. Mixed lymphocyte reactions as described above were extended to 5 days with enriched T cells and antigen presenting cells, and T cell proportions were measured using flow cytometry as described above.

3.16 Statistical Analysis

Shannon's diversity index was calculated in Qiime and student's, Welch's, or Wicoxon t-tests were calculated in R, depending on the distribution. Bray Curtis distance matrices were calculated in QIIME to assess compositional dissimilarity between samples and visualized using principal coordinates analysis (PCoA) plots in R. Permutational multivariate analysis of variance (PERMANOVA) was performed using *Adonis* function of vegan package[246] in R to determine factors that significantly ($p<0.05$) explained variation in microbiota β-diversity. In cases where replicates were included, linear mixed effects modeling was used to determine significance using the R package lmerTest[247].

Ranked abundance curve fit to geometric or log-series functions was determined by Bayesian Information Criterion (BIC) to evaluate models generated from fitsad function in vegan R package. To determine which OTUs differed in relative abundance between contamination swab and meconium, unnormalized read counts were transformed using DESEQ2 in QIIME to identify log-fold change enrichment and corrected for multiple hypothesis testing using the false-discovery rate ($q<0.05$). Growth curves were modeled using a logistic regression in R package growthcurver[248], integral of the best fit regression was used to calculate the area under the curve (auc), and auc of vehicle was subtracted from hormone treatment controls according to the following formula:

$$\sum_{t_i=1}^{n} Hormone\ treatment(t_i) - Vehicle\ (t_i)\ \Delta i$$

Significance in gentamicin protection assays was evaluated by transforming colony forming unit (CFU) counts using $\log_{10}(CFU + 1)$ and applying a generalized linear model to transformed data. Significance in *ex vivo* immune cell assays was evaluated using linear mixed effect modeling to account for cell donor correlations and where indicated, residuals are plotted. Except where indicated, all analyses were performed using R statistical programming language in the Jupyter Notebook environment.

3.17 Data Availability Statement

All sequencing data associated with this study has been made publicly available. 16S rRNA bacterial profiling data generated in this study is available in the EMBLI-EBI ENA repository accession #PRJEB25779 (https://www.ebi.ac.uk/ena). *De novo* assembled genomes were deposited at DDBJ/ENA/GenBank under the accession number

VFQL00000000 for Micro36. The genome version described in this paper is version VFQL01000000 for Micro36. Raw sequence reads used for genome assembly were deposited in NCBI SRA under BioProject accession # PRJNA498337 for Micro36, respectively. RNA sequencing dataset is available in NCBI under PRJNA506292 accession. All additional datasets and materials are available from the corresponding author upon request; requests are promptly reviewed by the University of California San Francisco Innovation Office to verify if the request is subject to any intellectual property or confidentiality obligations. Any data and materials that can be shared will be released via a Material Transfer Agreement.

CHAPTER 4: FUTURE DIRECTIONS

This work represents the first report of a sparse bacterial presence in the fetal intestine. These bacteria are correlated with and modulate a specific subset of T cells abundant in the intestine and exhibit the capacity to persist intracellularly. It is plausible that fetal bacteria are transiently present and are cleared, thus future investigation of their prolonged effect should be attempted. For example, this could be accomplished through formal demonstration of LLT1 ligand induction by Micro36 and the subsequent inhibition of $PLZF^+$ $CD161^+$ T cell IFNγ production using blocking antibodies. Further, determining the antigen specificity of fetal intestinal and circulating T cells is imperative. Because many PLZF T cells exhibit innate-like responses, demonstrating the requirement of bacterial antigen presentation would bolster the argument that bacterial-specific adaptive responses are generated *in utero*. From a broader perspective the question of bacterial-reactive adaptive cells in fetal blood remains an important area of investigation. This could be addressed for example, by screening for cord blood antibodies that are reactive to the human microbiome peptidome and hits would indicate humoral immunity to bacteria. Currently, identifying peptides that elicit T cell responses is difficult because of the requirement of peptide processing and presentation. However, minimal antigen processing machinery systems could be developed *in vitro* and express peptidomes of interest for baiting T cells. The identity of bacterial antigens that elicit immune memory responses is central to the understanding of the long-term consequences of host-microbe interactions.

This work is also among the first to show that non-pathogenic bacteria capable of immune suppression persist intracellularly, indicating that intracellular commensals may

be a broader biological phenomenon. Several reports have indicated that identical strains are found between the mother's intestine and breast milk, suggesting that bacteria are capable of trafficking between niches. It is plausible that this is facilitated by intracellularization and commandeering of existing cell trafficking mechanisms. Through microscopy, single-cell RNA sequencing and genetic knock-outs, mechanisms of Micro36 intracellular survival could be determined. This would lead to further elucidation of mechanisms of tolerance promotion. Through genetic labeling, Micro36 could be introduced *in vivo* in pregnant mice and traced throughout the fetus and mother. Genetic manipulation of the intracellular pathway could be co-opted for drug delivery to the fetus. For example, Micro36 could be utilized to express peanut proteins and elicit tolerant immune responses in fetuses at high-risk of peanut allergy.

While bacteria in the fetus may be a rare phenomenon, a clear conclusion of this work is that antibacterial forces are a critical feature of reproduction. Bacteria will rapidly double in growth when encountering an environment favorable to their growth. We did not observe bacterial overgrowth in the fetus indicating that survival in the fetal environment is limited to highly adapted bacteria. Exploration of reproductive antibacterial immunity remains a key question of this work. For example, determining the mechanisms of immune and microbiome changes triggered by pregnancy could lead to discovery of novel antimicrobials. This antimicrobial immunity is likely to be innate, due its importance in reproduction. Selection of the microbiome for the offspring is a key evolutionary goal and determining the seeding organisms could lead to strategies for rationally re-seeding microbiomes in other contexts.

REFERENCES

1. Frascoli, M. *et al.* Alloreactive fetal T cells promote uterine contractility in preterm labor via IFN-γ and TNF-α. *Sci. Transl. Med.* **10**, eaan2263 (2018).

2. Michaëlsson, J., Mold, J. E., McCune, J. M. & Nixon, D. F. Regulation of T Cell Responses in the Developing Human Fetus. *J. Immunol.* **176**, 5741–5748 (2006).

3. Kemp, M. W. Preterm birth, intrauterine infection, and fetal inflammation. *Front. Immunol.* **5**, 574 (2014).

4. Nancy, P. *et al.* Chemokine Gene Silencing in Decidual Stromal Cells Limits T Cell Access to the Maternal-Fetal Interface. *Science (80-.).* **336**, 1317–1321 (2012).

5. Erlebacher, A. Immunology of the Maternal-Fetal Interface. *Annu. Rev. Immunol.* **31**, 387–411 (2013).

6. PrabhuDas, M. *et al.* Immune mechanisms at the maternal-fetal interface: Perspectives and challenges. *Nature Immunology* **16**, 328–334 (2015).

7. Ikuta, K. *et al.* A developmental switch in thymic lymphocyte maturation potential occurs at the level of hematopoietic stem cells. *Cell* **62**, 863–874 (1990).

8. Havran, W. L. & Allison, J. P. Developmentally ordered appearance of thymocytes expressing different T-cell antigen receptors. *Nature* **335**, 443–5 (1988).

9. Friedberg, S. H. & Weissman, I. L. Lymphoid tissue architecture. II. Ontogeny of peripheral T and B cells in mice: evidence against Peyer's patches as the site of generation of B cells. *J. Immunol.* **113**, 1477–1492 (1974).

10. Lobach, D. F. & Haynes, B. F. Ontogeny of the human thymus during fetal development. *J. Clin. Immunol.* **7**, 81–97 (1987).

11. Cupedo, T., Nagasawa, M., Weijer, K., Blom, B. & Spits, H. Development and activation of regulatory T cells in the human fetus. *Eur. J. Immunol.* **35**, 383–390 (2005).

12. Darrasse-Jeze, G. Ontogeny of CD4+CD25+ regulatory/suppressor T cells in human fetuses. *Blood* **105**, 4715–4721 (2005).

13. den Braber, I. *et al.* Maintenance of Peripheral Naive T Cells Is Sustained by Thymus Output in Mice but Not Humans. *Immunity* **36**, 288–297 (2012).

14. Miller, J. F. A. P. Effect of neonatal thymectomy on the immunological responsiveness of the mouse. *Proc. R. Soc. London. Ser. B. Biol. Sci.* **156**, 415–428 (1962).

15. Adkins, B. T-cell function in newborn mice and humans. *Immunol. Today* **20**, 330–335 (1999).

16. Asano, M. Autoimmune disease as a consequence of developmental abnormality of a T cell subpopulation. *J. Exp. Med.* **184**, 387–396 (1996).

17. Mancebo, E. *et al.* Longitudinal analysis of immune function in the first 3 years of life in thymectomized neonates during cardiac surgery. *Clin. Exp. Immunol.* **154**, 375–383 (2008).

18. Prelog, M. *et al.* Thymectomy in early childhood: Significant alterations of the CD4+CD45RA+CD62L+ T cell compartment in later life. *Clin. Immunol.* **130**, 123–

132 (2009).

19. McGovern, N. *et al.* Human fetal dendritic cells promote prenatal T-cell immune suppression through arginase-2. *Nature* **546**, 662–666 (2017).

20. Scharschmidt, T. C. *et al.* A Wave of Regulatory T Cells into Neonatal Skin Mediates Tolerance to Commensal Microbes. *Immunity* **43**, 1011–1021 (2015).

21. Yu, J. *et al.* Maternal exposure to farming environment protects offspring against allergic diseases by modulating the neonatal TLR-Tregs-Th axis. *Clin. Transl. Allergy* **8**, 1–13 (2018).

22. Riedler, J. *et al.* Exposure to farming in early life and development of asthma and allergy: A cross-sectional survey. *Lancet* **358**, 1129–1133 (2001).

23. Eriksson, J. *et al.* Growing up on a farm leads to lifelong protection against allergic rhinitis. *Allergy Eur. J. Allergy Clin. Immunol.* **65**, 1397–1403 (2010).

24. Fujimura, K. E. *et al.* Neonatal gut microbiota associates with childhood multisensitized atopy and T cell differentiation. *Nat. Med.* **22**, 1187–1191 (2016).

25. Liu, L. *et al.* Global, regional, and national causes of under-5 mortality in 2000–15: an updated systematic analysis with implications for the Sustainable Development Goals. *Lancet* **388**, 3027–3035 (2016).

26. Chen, L., Cohen, A. C. & Lewis, D. B. Impaired allogeneic activation and T-helper 1 differentiation of human cord blood naive CD4 T cells. *Biol. Blood Marrow Transplant.* **12**, 160–171 (2006).

27. Velilla, P. A., Rugeles, M. T. & Chougnet, C. A. Defective antigen-presenting cell function in human neonates. *Clin. Immunol.* **121**, 251–259 (2006).

28. Olin, A. *et al.* Stereotypic Immune System Development in Newborn Children. *Cell* **174**, 1277-1292.e14 (2018).

29. Halkias, J. *et al.* CD161 contributes to prenatal immune suppression of IFNγ-producing PLZF+ T cells. *J. Clin. Invest.* (2019). doi:10.1172/JCI125957

30. Schreurs, R. R. C. E. *et al.* Human Fetal TNF-α-Cytokine-Producing CD4+ Effector Memory T Cells Promote Intestinal Development and Mediate Inflammation Early in Life. *Immunity* **0**, 1–15 (2019).

31. Li, N. *et al.* Memory CD4+ T cells are generated in the human fetal intestine. *Nat. Immunol.* (2019). doi:10.1038/s41590-018-0294-9

32. Stras, S. F. *et al.* Maturation of the Human Intestinal Immune System Occurs Early in Fetal Development. *Dev. Cell* **51**, 357-373.e5 (2019).

33. Mold, J. E. *et al.* Maternal Alloantigens Promote the Development of Tolerogenic Fetal Regulatory T Cells in Utero. *Science (80-.).* **322**, 1562–1565 (2008).

34. Min, H. S. *et al.* MHC Class II-Restricted Interaction between Thymocytes Plays an Essential Role in the Production of Innate CD8 + T Cells . *J. Immunol.* **186**, 5749–5757 (2011).

35. Jacomet, F. *et al.* Evidence for eomesodermin-expressing innate-like CD8+ KIR/NKG2A+ T cells in human adults and cord blood samples. *Eur. J. Immunol.* **45**, 1926–1933 (2015).

36. Gibbons, D. *et al.* Interleukin-8 (CXCL8) production is a signatory T cell effector function of human newborn infants. *Nat. Med.* **20**, 1206–1210 (2014).

37. Das, A. *et al.* Adaptive from Innate: Human IFN-γ + CD4 + T Cells Can Arise Directly from CXCL8-Producing Recent Thymic Emigrants in Babies and Adults . *J. Immunol.* **199**, 1696–1705 (2017).

38. Haynes, B. F. & Heinly, C. S. Early human T cell development: Analysis of the human thymus at the time of initial entry of hematopoietic stem cells into the fetal thymic microenvironment. *J. Exp. Med.* **181**, 1445–1458 (1995).

39. Farley, A. M. *et al.* Dynamics of thymus organogenesis and colonization in early human development. *Dev.* **140**, 2015–2026 (2013).

40. Schönland, S. O. *et al.* Homeostatic control of T-cell generation in neonates. *Blood* **102**, 1428–1434 (2003).

41. Min, B. *et al.* Neonates support lymphopenia-induced proliferation. *Immunity* **18**, 131–140 (2003).

42. Zhang, X. *et al.* CD4 T Cells with Effector Memory Phenotype and Function Develop in the Sterile Environment of the Fetus. *Sci. Transl. Med.* **6**, 238ra72-238ra72 (2014).

43. Spencer, J., Dillon, S. B., Isaacson, P. G. & Macdonald, T. T. T cell subclasses in fetal ileum. *Clin. Exp. Immunol.* **65**, 553–558 (1986).

44. Spencer, J., MacDonald, T. T., Finn, T. & Isaacson, P. G. The development of gut associated lymphoid tissue in the terminal ileum of fetal human intestine. *Clin.*

Exp. Immunol. **64**, 536–43 (1986).

45. Howie, D. *et al.* Extrathymic T Cell Differentiation in the Human Intestine Early in Life. *J. Immunol.* **161**, 5862–5872 (1998).

46. Thome, J. J. C. *et al.* Early-life compartmentalization of human T cell differentiation and regulatory function in mucosal and lymphoid tissues. *Nat. Med.* **22**, 72–77 (2016).

47. Seddiki, N. *et al.* Expression of interleukin (IL)-2 and IL-7 receptors discriminates between human regulatory and activated T cells. *J. Exp. Med.* **203**, 1693–1700 (2006).

48. Liu, W. *et al.* CD127 expression inversely correlates with FoxP3 and suppressive function of human CD4+ T reg cells. *J. Exp. Med.* **203**, 1701–1711 (2006).

49. Bronevetsky, Y., Burt, T. D. & McCune, J. M. Lin28b Regulates Fetal Regulatory T Cell Differentiation through Modulation of TGF-β Signaling. *J. Immunol.* **197**, 4344–4350 (2016).

50. Takahata, Y. *et al.* CD25+CD4+ T cells in human cord blood: An immunoregulatory subset with naive phenotype and specific expression of forkhead box p3 (Foxp3) gene. *Exp. Hematol.* **32**, 622–629 (2004).

51. Mold, J. E. *et al.* Fetal and Adult Hematopoietic Stem Cells Give Rise to Distinct T Cell Lineages in Humans. *Science (80-.).* **330**, 1695–1699 (2010).

52. Ng, M. S. F., Roth, T. L., Mendoza, V. F., Marson, A. & Burt, T. D. Helios enhances the preferential differentiation of human fetal CD4+ naïve T cells into

regulatory T cells. *Sci. Immunol.* **4**, 1–20 (2019).

53. Bacchetta, R., Barzaghi, F. & Roncarolo, M. G. From IPEX syndrome to FOXP3 mutation: A lesson on immune dysregulation. *Ann. N. Y. Acad. Sci.* **1417**, 5–22 (2016).

54. Luciano, A. A., Yu, H., Jackson, L. W., Wolfe, L. A. & Bernstein, H. B. Preterm labor and chorioamnionitis are associated with neonatal T cell activation. *PLoS One* **6**, 2–6 (2011).

55. Crespo, M. *et al.* Neonatal T-cell maturation and homing receptor responses to Toll-like receptor ligands differ from those of adult naive T cells: Relationship to prematurity. *Pediatr. Res.* **71**, 136–143 (2012).

56. Duggan, P. J. *et al.* Intrauterine T-cell activation and increased proinflammatory cytokine concentrations in preterm infants with cerebral lesions. *Lancet* **358**, 1699–1700 (2001).

57. Frascoli, M. *et al.* Alloreactive fetal T cells promote uterine contractility in preterm labor via IFN-γ and TNF-α. *Sci. Transl. Med.* **10**, eaan2263 (2018).

58. Vekemans, J. *et al.* Neonatal bacillus Calmette-Guérin vaccination induces adult-like IFN-gamma production by CD4+ T lymphocytes. *Eur. J. Immunol.* **31**, 1531–5 (2001).

59. Marchant, A. *et al.* Newborns develop a Th1-type immune response to Mycobacterium bovis bacillus Calmette-Guerin vaccination. *J. Immunol.* **163**, 2249–2255 (1999).

60. Malhotra, I. *et al.* In utero exposure to helminth and mycobacterial antigens generates cytokine responses similar to that observed in adults. *J. Clin. Invest.* **99**, 1759–1766 (1997).

61. Odorizzi, P. M. *et al.* In utero priming of highly functional effector T cell responses to human malaria. *Sci. Transl. Med.* **10**, eaat6176 (2018).

62. Halkias, J. *et al.* CD161 contributes to prenatal immune suppression of IFNγ-producing PLZF+ T cells. *J. Clin. Invest.* (2019). doi:10.1172/JCI125957

63. Kovalovsky, D. *et al.* The BTB–zinc finger transcriptional regulator PLZF controls the development of invariant natural killer T cell effector functions. *Nat. Immunol.* **9**, 1055–1064 (2008).

64. Savage, A. K. *et al.* The Transcription Factor PLZF Directs the Effector Program of the NKT Cell Lineage. *Immunity* **29**, 391–403 (2008).

65. Okazawa, A. *et al.* Th1-mediated intestinal inflammation in Crohn's disease may be induced by activation of lamina propria lymphocytes through synergistic stimulation of interleukin-12 and interleukin-18 without T cell receptor engagement. *Am. J. Gastroenterol.* **97**, 3108–3117 (2002).

66. Seyda, M., Elkhal, A., Quante, M., Falk, C. S. & Tullius, S. G. T Cells Going Innate. *Trends Immunol.* **37**, 546–556 (2016).

67. Loh, L., Ivarsson, M., Michaëlsson, J., Sandberg, J. & Nixon, D. Invariant natural killer T cells developing in the human fetus accumulate and mature in the small intestine. *Mucosal Immunol.* **7**, (2014).

68. Leeansyah, E., Loh, L., Nixon, D. F. & Sandberg, J. K. Acquisition of innate-like microbial reactivity in mucosal tissues during human fetal MAIT-cell development. *Nat. Commun.* **5**, 3143 (2014).

69. Youssef, G. Ben *et al.* Ontogeny of human mucosal-associated invariant T cells and related T cell subsets. *J. Exp. Med.* **215**, 459–479 (2018).

70. Vermijlen, D. *et al.* Human cytomegalovirus elicits fetal γδ T cell responses in utero. *J. Exp. Med.* **207**, 807–821 (2010).

71. Papadopoulou, M. *et al.* TCR Sequencing Reveals the Distinct Development of Fetal and Adult Human Vγ9Vδ2 T Cells. *J. Immunol.* **203**, 1468–1479 (2019).

72. Barrios, C. *et al.* Neonatal and early life immune responses to various forms of vaccine antigens qualitatively differ from adult responses: Predominance of a Th2-biased pattern which persists after adult boosting. *Eur. J. Immunol.* **26**, 1489–1496 (1996).

73. Forsthuber, T., Yip, H. C. & Lehmann, P. V. Induction of TH1 and TH2 immunity in neonatal mice. *Science (80-.).* **271**, 1728–1730 (1996).

74. Kovarik, J. *et al.* CpG oligodeoxynucleotides can circumvent the Th2 polarization of neonatal responses to vaccines but may fail to fully redirect Th2 responses established by neonatal priming. *J. Immunol.* **162**, 1611–1617 (1999).

75. Rose, S., Lichtenheld, M., Foote, M. R. & Adkins, B. Murine Neonatal CD4 + Cells Are Poised for Rapid Th2 Effector-Like Function . *J. Immunol.* **178**, 2667–2678 (2007).

76. Hebel, K. *et al.* CD4 + T Cells from Human Neonates and Infants Are Poised Spontaneously To Run a Nonclassical IL-4 Program . *J. Immunol.* **192**, 5160–5170 (2014).

77. Webster, R. B., Rodriguez, Y., Klimecki, W. T. & Vercelli, D. The human IL-13 locus in neonatal CD4+ T cells is refractory to the acquisition of a repressive chromatin architecture. *J. Biol. Chem.* **282**, 700–709 (2007).

78. Siegrist, C. A. Vaccination in the neonatal period and early infancy. *Int. Rev. Immunol.* **19**, 195–219 (2000).

79. Lewis, D. B. *et al.* Cellular and molecular mechanisms for reduced interleukin 4 and interferon-γ production by neonatal T cells. *J. Clin. Invest.* **87**, 194–202 (1991).

80. Gieseck, R. L., Wilson, M. S. & Wynn, T. A. Type 2 immunity in tissue repair and fibrosis. *Nat. Rev. Immunol.* **18**, 62–76 (2018).

81. Cosmi, L. *et al.* Human interleukin 17-producing cells originate from a CD161 +CD4+ T cell precursor. *J. Exp. Med.* **205**, 1903–1916 (2008).

82. Black, A., Bhaumik, S., Kirkman, R. L., Weaver, C. T. & Randolph, D. A. Developmental regulation of Th17-cell capacity in human neonates. *Eur. J. Immunol.* **42**, 311–319 (2012).

83. Cupedo, T. *et al.* Human fetal lymphoid tissue-inducer cells are interleukin 17-producing precursors to RORC+ CD127+ natural killer-like cells. *Nat. Immunol.* **10**, 66–74 (2009).

84. Dimova, T. *et al.* Effector Vγ9Vδ2 T cells dominate the human fetal γδ T-cell repertoire. *Proc. Natl. Acad. Sci.* **112**, E556–E565 (2015).

85. Galindo-Albarrán, A. O. *et al.* CD8+ T Cells from Human Neonates Are Biased toward an Innate Immune Response. *Cell Rep.* **17**, 2151–2160 (2016).

86. Hong, M. *et al.* Trained immunity in newborn infants of HBV-infected mothers. *Nat. Commun.* **6**, (2015).

87. Babik, J. M., Cohan, D., Monto, A., Hartigan-O'Connor, D. J. & McCune, J. M. The human fetal immune response to hepatitis C virus exposure in utero. *J. Infect. Dis.* **203**, 196–206 (2011).

88. Luzuriaga, K. *et al.* HIV-1-specific cytotoxic T lymphocyte responses in the first year of life. *J. Immunol.* **154**, 433–443 (1995).

89. Sanchez-Merino, V., Nie, S. & Luzuriaga, K. HIV-1-Specific CD8 + T Cell Responses and Viral Evolution in Women and Infants . *J. Immunol.* **175**, 6976–6986 (2005).

90. Legrand, F. A. *et al.* Strong HIV-1-specific T cell responses in HIV-1-exposed uninfected infants and neonates revealed after regulatory T cell removal. *PLoS One* **1**, (2006).

91. Marchant, A. *et al.* Mature CD8(+) T lymphocyte response to viral infection during fetal life. *J.Clin.Invest* **111**, 1747–1755 (2003).

92. Kim, Y. J., Stringfield, T. M., Chen, Y. & Broxmeyer, H. E. Modulation of cord blood CD8+ T-cell effector differentiation by TGF-β1 and 4-1BB costimulation.

Blood **105**, 274–281 (2005).

93. Zhang, Y. *et al.* Cord blood CD8 + T cells have a natural propensity to express IL-4 in a fatty acid metabolism and caspase activation-dependent manner. *Front. Immunol.* **9**, 1–13 (2018).

94. Wang, J. *et al.* Fetal and adult progenitors give rise to unique populations of CD8+ T cells. *Blood* **128**, 3073–3082 (2016).

95. Smith, N. L. *et al.* Developmental Origin Governs CD8 + T Cell Fate Decisions during Infection. *Cell* **174**, 117-130.e14 (2018).

96. Ng, M. S. F., Roth, T. L., Mendoza, V. F., Marson, A. & Burt, T. D. Helios enhances the preferential differentiation of human fetal CD4+ naïve T cells into regulatory T cells. *Sci. Immunol.* **4**, 1–20 (2019).

97. Chen, L. & Flies, D. B. Molecular mechanisms of T cell co-stimulation and co-inhibition. *Nat. Rev. Immunol.* **13**, 227–242 (2013).

98. Aldemir, H. *et al.* Cutting Edge: Lectin-Like Transcript 1 Is a Ligand for the CD161 Receptor. *J. Immunol.* **175**, 7791–7795 (2005).

99. Germain, C. *et al.* Induction of lectin-like transcript 1 (LLT1) protein cell surface expression by pathogens and interferon-γ contributes to modulate immune responses. *J. Biol. Chem.* **286**, 37964–37975 (2011).

100. Rosen, D. B. *et al.* Cutting Edge: Lectin-Like Transcript-1 Is a Ligand for the Inhibitory Human NKR-P1A Receptor. *J. Immunol.* **175**, 7796–7799 (2005).

101. Exley, M., Porcelli, S., Furman, M., Garcia, J. & Balk, S. CD161 (NKR-P1A) Costimulation of CD1d-dependent Activation of Human T Cells Expressing Invariant Vα24JαQ T Cell Receptor α Chains. *J. Exp. Med.* **188**, 867 LP – 876 (1998).

102. Fergusson, J. R. *et al.* CD161 defines a transcriptional and functional phenotype across distinct human T cell lineages. *Cell Rep.* **9**, 1075–1088 (2014).

103. Rosen, D. B. *et al.* Functional Consequences of Interactions between Human NKR-P1A and Its Ligand LLT1 Expressed on Activated Dendritic Cells and B Cells. *J. Immunol.* **180**, 6508–6517 (2008).

104. Denning, T. L., Wang, Y. C., Patel, S. R., Williams, I. R. & Pulendran, B. Lamina propria macrophages and dendritic cells differentially induce regulatory and interleukin 17-producing T cell responses. *Nat. Immunol.* **8**, 1086–1094 (2007).

105. Egea, L., Hirata, Y. & Kagnoff, M. F. GM-CSF: A role in immune and inflammatory reactions in the intestine. *Expert Rev. Gastroenterol. Hepatol.* **4**, 723–731 (2010).

106. Krause, P. *et al.* IL-10-producing intestinal macrophages prevent excessive antibacterial innate immunity by limiting IL-23 synthesis. *Nat. Commun.* **6**, (2015).

107. Veglia, F., Perego, M. & Gabrilovich, D. Myelod derived suppressor cells coming of age. *Nat. Immunol.* **19**, 108–119 (2019).

108. Rieber, N. *et al.* Neutrophilic myeloid-derived suppressor cells in cord blood modulate innate and adaptive immune responses. *Clin. Exp. Immunol.* **174**, 45–52 (2013).

109. He, Y. M. *et al.* Transitory presence of myeloid-derived suppressor cells in neonates is critical for control of inflammation. *Nat. Med.* **24**, 224–231 (2018).

110. Gervassi, A. *et al.* Myeloid derived suppressor cells are present at high frequency in neonates and suppress in vitro t cell responses. *PLoS One* **9**, 1–7 (2014).

111. Leiber, A. *et al.* Neonatal myeloid derived suppressor cells show reduced apoptosis and immunosuppressive activity upon infection with Escherichia coli. *Eur. J. Immunol.* **47**, 1009–1021 (2017).

112. Zhang, X., Majlessi, L., Deriaud, E., Leclerc, C. & Lo-Man, R. Coactivation of Syk Kinase and MyD88 Adaptor Protein Pathways by Bacteria Promotes Regulatory Properties of Neutrophils. *Immunity* **31**, 761–771 (2009).

113. Casanova-Acebes, M. *et al.* Neutrophils instruct homeostatic and pathological states in naive tissues. *J. Exp. Med.* **215**, 2778–2795 (2018).

114. Mauri, C. & Bosma, A. Immune Regulatory Function of B Cells. *Annu. Rev. Immunol.* **30**, 221–241 (2012).

115. Rolle, L. *et al.* Cutting Edge: IL-10-Producing Regulatory B Cells in Early Human Pregnancy. *Am. J. Reprod. Immunol.* **70**, 448–453 (2013).

116. Fettke, F., Schumacher, A., Costa, S. D. & Zenclussen, A. C. B cells: The old new players in reproductive immunology. *Front. Immunol.* **5**, 1–10 (2014).

117. Sun, C. M., Deriaud, E., Leclerc, C. & Lo-Man, R. Upon TLR9 signaling, CD5+ B cells control the IL-12-dependent Th1-priming capacity of neonatal DCs. *Immunity* **22**, 467–477 (2005).

118. Zhang, X. *et al.* Type I interferons protect neonates from acute inflammation through interleukin 10-producing B cells. *J. Exp. Med.* **204**, 1107–1118 (2007).

119. Sarvaria, A. *et al.* IL-10+ regulatory B cells are enriched in cord blood and may protect against cGVHD after cord blood transplantation. *Blood* **128**, 1346–1361 (2016).

120. Esteve-Solé, A. *et al.* Characterization of the highly prevalent regulatory CD24hiCD38hi B-Cell population in human cord blood. *Front. Immunol.* **8**, 1–12 (2017).

121. Stras, S. F. *et al.* Maturation of the Human Intestinal Immune System Occurs Early in Fetal Development. *Dev. Cell* **51**, 357-373.e5 (2019).

122. Elahi, S. *et al.* Immunosuppressive CD71 + erythroid cells compromise neonatal host defence against infection. *Nature* **504**, 158–162 (2013).

123. Shahbaz, S. *et al.* CD71 + VISTA + erythroid cells promote the development and function of regulatory T cells through TGF-β. *PLoS Biol.* **16**, 1–26 (2018).

124. Sitkovsky, M. & Lukashev, D. Regulation of immune cells by local-tissue oxygen tension: HIF1α and adenosine receptors. *Nat. Rev. Immunol.* **5**, 712–721 (2005).

125. Facciabene, A. *et al.* Tumour hypoxia promotes tolerance and angiogenesis via CCL28 and T reg cells. *Nature* **475**, 226–230 (2011).

126. Polanczyk, M. J. *et al.* Cutting Edge: Estrogen Drives Expansion of the CD4 + CD25 + Regulatory T Cell Compartment . *J. Immunol.* **173**, 2227–2230 (2004).

127. Salem, M. L. Estrogen, a double-edged sword: Modulation of TH1- and TH2-mediated inflammations by differential regulation of TH1/TH2 cytokine production. *Curr. Drug Targets Inflamm. Allergy* **3**, 97–104 (2004).

128. Lee, J. H., Ulrich, B., Cho, J., Park, J. & Kim, C. H. Progesterone promotes differentiation of human cord blood fetal T cells into T regulatory cells but suppresses their differentiation into Th17 cells. *J. Immunol.* **187**, 1778–1787 (2011).

129. Miyaura, H. & Iwata, M. Direct and Indirect Inhibition of Th1 Development by Progesterone and Glucocorticoids. *J. Immunol.* **168**, 1087–1094 (2002).

130. Hughes, G. C., Clark, E. A. & Wong, A. H. The intracellular progesterone receptor regulates CD4+ T cells and T cell-dependent antibody responses. *J. Leukoc. Biol.* **93**, 369–375 (2013).

131. Gilmore, W., Weiner, L. P. & Correale, J. Effect of estradiol on cytokine secretion by proteolipid protein-specific T cell clones isolated from multiple sclerosis patients and normal control subjects. *J. Immunol.* **158**, 446–51 (1997).

132. Maret, A. *et al.* Estradiol enhances primary antigen-specific CD4 T cell responses and Th1 development in vivo. Essential role of estrogen receptor α expression in hematopoietic cells. *Eur. J. Immunol.* **33**, 512–521 (2003).

133. Tai, P. *et al.* Induction of regulatory T cells by physiological level estrogen. *J. Cell. Physiol.* **214**, 456–464 (2008).

134. Fox, H. S., Bond, B. L. & Parslow, T. G. Estrogen regulates the IFN-gamma

promoter. *J. Immunol.* **146**, 4362 LP – 4367 (1991).

135. Tai, P. *et al.* Induction of regulatory T cells by physiological level estrogen. *J. Cell. Physiol.* **214**, 456–464 (2008).

136. Chen, R.-Y. *et al.* Estradiol Inhibits Th17 Cell Differentiation through Inhibition of RORγT Transcription by Recruiting the ERα/REA Complex to Estrogen Response Elements of the RORγT Promoter . *J. Immunol.* **194**, 4019–4028 (2015).

137. Beck, S. *et al.* The worldwide incidence of preterm birth: A systematic review of maternal mortality and morbidity. *Bull. World Health Organ.* **88**, 31–38 (2010).

138. Liu, L. *et al.* Global, regional, and national causes of under-5 mortality in 2000–15: an updated systematic analysis with implications for the Sustainable Development Goals. *Lancet* **388**, 3027–3035 (2016).

139. Romero, R. *et al.* A fetal systemic inflammatory response is followed by the spontaneous onset of preterm parturition. *Am. J. Obstet. Gynecol.* **179**, 186–193 (1998).

140. Romero, R. *et al.* Prevalence and Clinical Significance of Sterile Intra-amniotic Inflammation in Patients with Preterm Labor and Intact Membranes. *Am. J. Reprod. Immunol.* **72**, 458–474 (2014).

141. Dammann, O. & Leviton, A. Brain damage in preterm newborns: Might enhancement of developmentally regulated endogenous protection open a door for prevention? *Pediatrics* **104**, 541–550 (1999).

142. Dammann, O. & Leviton, A. Maternal Intrauterine Infection, Cytokines, and Brain

Damage in the Preterm Newborn. *Pediatr. Res.* **42**, 1–8 (1997).

143. Viscardi, R. M. Perinatal inflammation and lung injury. *Semin. Fetal Neonatal Med.* **17**, 30–35 (2012).

144. Takahashi, N. *et al.* Cytokine profiles of seventeen cytokines, growth factors and chemokines in cord blood and its relation to perinatal clinical findings. *Cytokine* **49**, 331–337 (2010).

145. Rueda, C. M. *et al.* Effect of chorioamnionitis on regulatory T cells in moderate/late preterm neonates. *Hum. Immunol.* **76**, 65–73 (2015).

146. Wolfs, T. G. A. M. *et al.* Chorioamnionitis-induced fetal gut injury is mediated by direct gut exposure of inflammatory mediators or by lung inflammation. *Am. J. Physiol. Gastrointest. Liver Physiol.* **306**, G382–G393 (2014).

147. Wolfs, T. G. A. M. *et al.* IL-1α mediated chorioamnionitis induces depletion of FoxP3+ cells and ileal inflammation in the ovine fetal gut. *PLoS One* **6**, 1–9 (2011).

148. Maneenil, G. *et al.* Oral, nasal and pharyngeal exposure to lipopolysaccharide causes a fetal inflammatory response in sheep. *PLoS One* **10**, 1–11 (2015).

149. Nikiforou, M. *et al.* Intra-amniotic Candida albicans infection induces mucosal injury and inflammation in the ovine fetal intestine. *Sci. Rep.* **6**, 1–9 (2016).

150. Weitkamp, J. H. *et al.* Necrotising enterocolitis is characterised by disrupted immune regulation and diminished mucosal regulatory (FOXP3)/effector (CD4, CD8) T cell ratios. *Gut* **62**, 73–82 (2013).

151. Egan, C. E. *et al.* Toll-like receptor 4-mediated lymphocyte influx induces neonatal necrotizing enterocolitis. *J. Clin. Invest.* **126**, 495–508 (2016).

152. Rackaityte, E. *et al.* Viable bacterial colonization is highly limited in the human intestine in utero. *Nat. Med.* 1–44 (2020). doi:10.1038/s41591-020-0761-3

153. Koenig, J. E. *et al.* Succession of microbial consortia in the developing infant gut microbiome. *Proc. Natl. Acad. Sci. U. S. A.* **108 Suppl**, 4578–4585 (2011).

154. Turnbaugh, P. J. *et al.* The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. *Sci. Transl. Med.* **1**, 6ra14 (2009).

155. Fujimura, K. E. *et al.* House dust exposure mediates gut microbiome Lactobacillus enrichment and airway immune defense against allergens and virus infection. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 805–10 (2014).

156. Brugman, S., Perdijk, O., van Neerven, R. J. J. & Savelkoul, H. F. J. Mucosal Immune Development in Early Life: Setting the Stage. *Arch. Immunol. Ther. Exp. (Warsz).* (2015). doi:10.1007/s00005-015-0329-y

157. Aagaard, K. *et al.* The placenta harbors a unique microbiome. *Sci. Transl. Med.* **6**, 237ra65 (2014).

158. Fardini, Y., Chung, P., Dumm, R., Joshi, N. & Han, Y. W. Transmission of diverse oral bacteria to murine placenta: Evidence for the oral microbiome as a potential source of intrauterine infection. *Infect. Immun.* **78**, 1789–1796 (2010).

159. Collado, M. C., Rautava, S., Aakko, J., Isolauri, E. & Salminen, S. Human gut

colonisation may be initiated in utero by distinct microbial communities in the placenta and amniotic fluid. *Sci. Rep.* **6**, 23129 (2016).

160. Steel, J. H. *et al.* Bacteria and Inflammatory Cells in Fetal Membranes Do Not Always Cause Preterm Labor. *Pediatr. Res.* **57**, 404–411 (2005).

161. Jiménez, E. *et al.* Is meconium from healthy newborns actually sterile? *Res. Microbiol.* **159**, 187–193 (2008).

162. Ardissone, A. N. *et al.* Meconium microbiome analysis identifies bacteria correlated with premature birth. *PLoS One* **9**, 1–8 (2014).

163. Gosalbes, M. J. *et al.* Meconium microbiota types dominated by lactic acid or enteric bacteria are differentially associated with maternal eczema and respiratory problems in infants. *Clin. Exp. Allergy* **43**, 198–211 (2013).

164. Duram, L. A. *Encyclopedia of organic, sustainable, and local food*. (ABC-CLIO, 2010).

165. EGGESBø, M. *et al.* Development of gut microbiota in infants not exposed to medical interventions. *APMIS* **119**, 17–35 (2011).

166. Arrieta, M. *et al.* Early infancy microbial and metabolic alterations affect risk of childhood asthma. **7**, (2015).

167. Durack, J. *et al.* Delayed gut microbiota development in high-risk for asthma infants is temporarily modifiable by Lactobacillus supplementation. *Nat. Commun.* **9**, (2018).

168. Wang, Z. *et al.* Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* **472**, 57–63 (2011).

169. Cahenzli, J., Köller, Y., Wyss, M., Geuking, M. B. & McCoy, K. D. Intestinal microbial diversity during early-life colonization shapes long-term IgE levels. *Cell Host Microbe* **14**, 559–570 (2013).

170. Smith, P. M. *et al.* The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. *Science* **341**, 569–73 (2013).

171. Pollard, M. & Sharon, N. Responses of the Peyer's Patches in Germ-Free Mice to Antigenic Stimulation. *Infect. Immun.* **2**, 96–100 (1970).

172. Hooper, L. V & Macpherson, A. J. Immune adaptations that maintain homeostasis with the intestinal microbiota. *Nat. Rev. Immunol.* **10**, 159–169 (2010).

173. Herbst, T. *et al.* Dysregulation of allergic airway inflammation in the absence of microbial colonization. *Am. J. Respir. Crit. Care Med.* **184**, 198–205 (2011).

174. Olszak, T. *et al.* Microbial exposure during early life has persistent effects on natural killer T cell function. *Science* **336**, 489–93 (2012).

175. Russell, S. L. *et al.* Early life antibiotic-driven changes in microbiota enhance susceptibility to allergic asthma. *EMBO Rep.* **13**, 440–7 (2012).

176. Lynch, S. V *et al.* Effects of early-life exposure to allergens and bacteria on recurrent wheeze and atopy in urban children. *J. Allergy Clin. Immunol.* **134**, 593-601.e12 (2014).

177. Rautava, S., Collado, M. C., Salminen, S. & Isolauri, E. Probiotics Modulate Host-Microbe Interaction in the Placenta and Fetal Gut: A Randomized, Double-Blind, Placebo-Controlled Trial. *Neonatology* **102**, 178–184 (2012).

178. de Goffau, M. C. *et al.* Human placenta has no microbiome but can contain potential pathogens. *Nature* (2019). doi:10.1038/s41586-019-1451-5

179. Salter, S. J. *et al.* Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* **12**, 87 (2014).

180. Lauder, A. P. *et al.* Comparison of placenta samples with contamination controls does not provide evidence for a distinct placenta microbiota. *Microbiome* 1–11 (2016). doi:10.1186/s40168-016-0172-3

181. Minich, J. J. *et al.* Quantifying and Understanding Well-to-Well Contamination in Microbiome Research. *mSystems* **4**, 1–13 (2019).

182. Davis, N. M., Proctor, D. M., Holmes, S. P., Relman, D. A. & Callahan, B. J. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* **6**, 226 (2018).

183. Kondo, N. *et al.* Cord blood lymphocyte responses to antigens for the prediction of allergy. *Pediatr. Asthma, Allergy Immunol.* **12**, 61–66 (1998).

184. Miller, R. L. *et al.* Prenatal exposure, maternal sensitization, and sensitization in utero to indoor allergens in an inner-city cohort. *Am. J. Respir. Crit. Care Med.* **164**, 995–1001 (2001).

185. Prescott, S. L. *et al.* Development of allergen-specific T-cell memory in atopic and

normal children. *Lancet* **353**, 196–200 (1999).

186. Hagendorens, M. M. *et al.* Prenatal exposure to house dust mite allergen (Der p 1), cord blood T cell phenotype and cytokine production and atopic dermatitis during the first year of life. *Pediatr. Allergy Immunol.* **15**, 308–315 (2004).

187. Prescott, S. L. *et al.* Transplacental priming of the human immune system to environmental allergens: universal skewing of initial T cell responses toward the Th2 cytokine profile. *J. Immunol.* **160**, 4730–7 (1998).

188. Holloway, J. A. *et al.* Detection of house-dust-mite allergen in amniotic fluid and umbilical-cord blood. *Lancet* **356**, 1900–1902 (2000).

189. Hallgren, P., Lindberg, B. S. & Lundblad, A. Quantitation of some urinary oligosaccharides during pregnancy and lactation. *J. Biol. Chem.* **252**, 1034–1040 (1977).

190. Wise, A. *et al.* Infants are exposed to human milk oligosaccharides already in Utero. *Front. Pediatr.* **6**, 1–4 (2018).

191. Plaza-Díaz, J., Fontana, L. & Gil, A. Human milk oligosaccharides and immune system development. *Nutrients* **10**, (2018).

192. Ardissone, A. N. *et al.* Meconium microbiome analysis identifies bacteria correlated with premature birth. *PLoS One* **9**, 1–8 (2014).

193. Willis, K. A. *et al.* Fungi form interkingdom microbial communities in the primordial human gut that develop with gestational age. *bioRxiv* 621235 (2019). doi:10.1101/621235

194. Aagaard, K. *et al.* The placenta harbors a unique microbiome. *Sci. Transl. Med.* **6**, 237ra65 (2014).

195. Stout, M. J. *et al.* Identification of intracellular bacteria in the basal plate of the human placenta in term and preterm gestations. *YMOB* **208**, 226.e1-226.e7 (2013).

196. Seferovic, M. D. *et al.* Visualization of microbes by 16S in situ hybridization in term and preterm placentas without intraamniotic infection. *Am. J. Obstet. Gynecol.* **221**, 146.e1-146.e23 (2019).

197. Parnell, L. A. *et al.* Microbial communities in placentas from term normal pregnancy exhibit spatially variable profiles. *Sci. Rep.* 1–11 (2017). doi:10.1038/s41598-017-11514-4

198. Rromano-Keeler, J. *et al.* Early life establishment of site-specific microbial communities in the gut. *Gut Microbes* **5**, 37–41 (2014).

199. Chu, D. M. *et al.* Maturation of the infant microbiome community structure and function across multiple body sites and in relation to mode of delivery. *Nat. Med.* **23**, 314-326 13 (2017).

200. McGovern, N. *et al.* Human fetal dendritic cells promote prenatal T-cell immune suppression through arginase-2. *Nature* (2017). doi:10.1038/nature22795

201. Podd, B. S. *et al.* T Cells in Cryptopatch Aggregates Share TCR Variable Region Junctional Sequences with T Cells in the Small Intestinal Epithelium of Mice. *J. Immunol.* **176**, 6532–6542 (2006).

202. Lim, E. S., Rodriguez, C. & Holtz, L. R. Amniotic fluid from healthy term pregnancies does not harbor a detectable microbial community. 4–11 (2018).

203. Yu, J. *et al.* Maternal exposure to farming environment protects offspring against allergic diseases by modulating the neonatal TLR-Tregs-Th axis. *Clin. Transl. Allergy* **8**, 1–13 (2018).

204. Gu, W. *et al.* Depletion of Abundant Sequences by Hybridization (DASH): using Cas9 to remove unwanted high-abundance species in sequencing libraries and molecular counting applications. *Genome Biol.* **17**, 41 (2016).

205. Trotter, A., Maier, L., Grill, H.-J., Wudy, S. A. & Pohlandt, F. 17β-Estradiol and Progesterone Supplementation in Extremely Low-Birth-Weight Infants. *Pediatr. Res.* **45**, 489–493 (1999).

206. Yotis, W. & Stanke, R. Bacteriostatic action of progesterone on staphylococci and other microorganisms. *J. Bacteriol.* **92**, 1285–1289 (1966).

207. Varghese, N. J. *et al.* Microbial species delineation using whole genome sequences. *Nucleic Acids Res.* **43**, 6761–6771 (2015).

208. Benach, J. *et al.* Structure of bacterial 3β/17β-hydroxysteroid dehydrogenase at 1.2 Å resolution: A model for multiple steroid recognition. *Biochemistry* **41**, 14659–14668 (2002).

209. Hillas, P. J., Soto Del Alba, F., Oyarzabal, J., Wilks, A. & Ortiz De Montellano, P. R. The AhpC and AhpD antioxidant defense system of Mycobacterium tuberculosis. *J. Biol. Chem.* **275**, 18801–18809 (2000).

210. Dons, L. E. *et al.* Role of the Listeria monocytogenes 2-Cys peroxiredoxin homologue in protection against oxidative and nitrosative stress and in virulence. *Pathog. Dis.* **70**, 70–74 (2014).

211. Arrieta, M.-C. *et al.* Early infancy microbial and metabolic alterations affect risk of childhood asthma. *Sci. Transl. Med.* **7**, 307ra152-307ra152 (2015).

212. Ferretti, P. *et al.* Mother-to-Infant Microbial Transmission from Different Body Sites Shapes the Developing Infant Gut Microbiome. *Cell Host Microbe* **24**, 133-145.e5 (2018).

213. Yassour, M. *et al.* Strain-Level Analysis of Mother-to-Child Bacterial Transmission during the First Few Months of Life. *Cell Host Microbe* **24**, 146-154.e4 (2018).

214. Rosen, D. B. *et al.* Functional Consequences of Interactions between Human NKR-P1A and Its Ligand LLT1 Expressed on Activated Dendritic Cells and B Cells. *J. Immunol.* **180**, 6508–6517 (2008).

215. Minich, J. J. *et al.* Quantifying and Understanding Well-to-Well Contamination in Microbiome Research. *mSystems* **4**, e00186-19 (2019).

216. Martín, R. *et al.* Characterization of indigenous vaginal lactobacilli from healthy women as probiotic candidates. *Int. Microbiol.* **11**, 261–266 (2008).

217. Chen, C. *et al.* The microbiota continuum along the female reproductive tract and its relation to uterine-related diseases. *Nat. Commun.* **8**, 875 (2017).

218. Mukamolova, G. V. *et al.* The rpf gene of Micrococcus luteus encodes an essential secreted growth factor. *Mol. Microbiol.* **46**, 611–621 (2002).

219. Duerkop, B. A., Vaishnava, S. & Hooper, L. V. Immune Responses to the Microbiota at the Intestinal Mucosal Surface. *Immunity* **31**, 368–376 (2009).

220. Pastor-vargas, C. *et al.* Detection of major food allergens in amniotic fluid : initial allergenic encounter during pregnancy. **27**, 716–720 (2016).

221. Iwai, S. *et al.* The lung microbiome of Ugandan HIV-infected pneumonia patients is compositionally and functionally distinct from that of San Franciscan patients. *PLoS One* **9**, (2014).

222. Gu W, Crawford ED, O'Donovan BD, Wilson MR, Chow ED, Retallack H, D. J. Depletion of Abundant Sequences by Hybridization (DASH): Using Cas9 to remove unwanted high- abundance species in sequencing libraries and molecular counting applications. *Genome Biol.* **17**, (2016).

223. Caporaso, J. G. *et al.* Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci.* **108**, 4516–4522 (2011).

224. Magoč, T. & Salzberg, S. L. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**, 2957–63 (2011).

225. Caporaso, J. G. *et al.* correspondence QIIME allows analysis of high- throughput community sequencing data Intensity normalization improves color calling in SOLiD sequencing. *Nat. Publ. Gr.* **7**, 335–336 (2010).

226. Edgar, R. C. & Flyvbjerg, H. Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics* **31**, 3476–3482 (2015).

227. Dobin, A. *et al.* STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–

21 (2013).

228. Liao, Y., Smyth, G. K. & Shi, W. FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).

229. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 1–21 (2014).

230. Johansson, M. E. V & Hansson, G. C. Preservation of Mucus in Histological Sections, Immunostaining of Mucins in Fixed Tissue, and Localization of Bacteria with FISH. in *Mucins: Methods and Protocols* (eds. McGuckin, M. A. & Thornton, D. J.) 229–235 (Humana Press, 2012). doi:10.1007/978-1-61779-513-8_13

231. Vaishnava, S. *et al.* The Antibacterial Lectin RegIII. *Science (80-.).* **334**, 255–258 (2011).

232. Weisburg, W. G., Barns, S. M., Pelletier, D. A. & Lane, D. J. 16S ribosomal DNA amplification for phylogenetic study. *Weisburg, WG* **173**, 697–703 (1991).

233. Pruesse, E., Peplies, J. & Glöckner, F. O. SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* **28**, 1823–1829 (2012).

234. Bankevich, A. *et al.* SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *J. Comput. Biol.* **19**, 455–477 (2012).

235. Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. QUAST: Quality assessment tool for genome assemblies. *Bioinformatics* **29**, 1072–1075 (2013).

236. Eren, A. M. *et al.* Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* **3**, e1319 (2015).

237. Pritchard, L., Glover, R. H., Humphris, S., Elphinstone, J. G. & Toth, I. K. Genomics and taxonomy in diagnostics for food security: Soft-rotting enterobacterial plant pathogens. *Anal. Methods* **8**, 12–24 (2016).

238. Rinke, C. *et al.* Insights into the phylogeny and coding potential of microbial dark matter. *Nature* **499**, 431–437 (2013).

239. Edgar, R. C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).

240. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2 - Approximately maximum-likelihood trees for large alignments. *PLoS One* **5**, (2010).

241. Letunic, I. & Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* **44**, W242–W245 (2016).

242. Wood, D. E. & Salzberg, S. L. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol.* **15**, R46 (2014).

243. Sekar, S., Mahadevan, S., Kumar, S. S. D. & Mandal, A. B. Thermokinetic responses of the metabolic activity of Staphylococcus lentus cultivated in a glucose limited mineral salt medium. *J. Therm. Anal. Calorim.* **104**, 149–155 (2011).

244. Subashchandrabose, S., Smith, S. N., Spurbeck, R. R., Kole, M. M. & Mobley, H.

L. T. Genome-Wide Detection of Fitness Genes in Uropathogenic Escherichia coli during Systemic Infection. *PLoS Pathog.* **9**, 1–15 (2013).

245. Ducsay, C. A. *et al.* Gestational Hypoxia and Developmental Plasticity. *Physiol. Rev.* **98**, 1241–1334 (2018).

246. Oksanen, J. *et al.* Vegan: Community Ecology Package. R Package Version. 2.0-10. *CRAN* (2013).

247. Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. **lmerTest** Package: Tests in Linear Mixed Effects Models. *J. Stat. Softw.* **82**, (2017).

248. Sprouffske, K. & Wagner, A. Growthcurver: An R package for obtaining interpretable metrics from microbial growth curves. *BMC Bioinformatics* **17**, 17–20 (2016).

Publishing Agreement

It is the policy of the University to encourage open access and broad distribution of all theses, dissertations, and manuscripts. The Graduate Division will facilitate the distribution of UCSF theses, dissertations, and manuscripts to the UCSF Library for open access and distribution. UCSF will make such theses, dissertations, and manuscripts accessible to the public and will take reasonable steps to preserve these works in perpetuity.

I hereby grant the non-exclusive, perpetual right to The Regents of the University of California to reproduce, publicly display, distribute, preserve, and publish copies of my thesis, dissertation, or manuscript in any form or media, now existing or later derived, including access online for teaching, research, and public service purposes.

DocuSigned by:

051FFDA1E5444EF...
Author Signature

9/3/2020
Date