Early acquisition and intestinal colonization by *E. coli* of infants born to mothers with Inflammatory Bowel Disease

Honors Thesis
Presented to the College of Agriculture and Life Sciences,
Cornell University in Partial Fulfillment of the Requirements for the Biological Sciences Honors Program

by
Daniela Lastras-Castrejon
May 2025

Supervisor: Dr. Kenneth Simpson, BVM&S, PhD

Abstract

The early life environment, when the microbiome is established, may dictate Inflammatory Bowel Disease (IBD) risk later in life. *E. coli* has emerged as a neonatal colonizer across species. To investigate *E. coli* colonization, phylogeny and virulence in response to different exposures, we analyzed metagenomic data and characterized *E. coli* isolates from mothers and their children. Taxa and alpha diversity for control mothers was higher than IBD mothers, and the opposite was true for *E. coli* abundance. An increase in taxa and alpha diversity is seen as infants age from 7 days to 4 years old with a concurrent decrease in *E. coli* abundance. In control mothers, *E. coli* were predominantly phylogroups B2 and D, whereas B1 isolates were more common in IBD mothers. In contrast, most infants were colonized by B2 isolates, independent of maternal IBD status. Furthermore, the phylogroup distribution of toddlers resembled that of control mothers. Cytotoxin and genotoxin genes were rare in maternal *E. coli* but common in infants of IBD mothers and absent in infants of control mothers. Conversely, AIEC-associated genes were more common in maternal *E. coli* particularly IBD mothers than infant *E. coli*. The *E. coli* isolates from toddler's were similar to the maternal cohort in rarity of genes encoding cytotoxins and genotoxins, but had more genes associated with AIEC. Our findings reveal that mothers and their babies are colonized by a diverse group of non-diarrheagenic *E. coli* that vary in phylogeny, genotype and virulence, refuting clonal transmission of maternal isolates.

Introduction

Inflammatory bowel disease (IBD) which encompasses Crohn's Disease (CD) and Ulcerative Colitis (UC) is a chronic inflammatory autoimmune disease of the gastrointestinal tract and occurs because of genetic susceptibility, environmental factors, and one's intestinal microbiome (Torres et al. 2017). CD has been increasing worldwide, especially in urban areas and is normally treated with immunosuppressants, biologicals, anti-adhesion molecules, and corticosteroids (Halfvarson et al. 2017 and Torres et al. 2017). Dysbiosis and fluctuations of the microbiome is often seen in patients with CD and is typically characterized by a decrease in *Bacteroides* and *Firmicutes* and an increase in *Gammaproteobacteria* and *Actinobacteria* (Halfvarson et al. 2017 and Torres et al. 2017). Furthermore, pathogens such as *Clostridium difficile*, *Helicobacter* and adherent-invasive *E. coli* have been identified in CD patients (Palmela et al. 2018). Around one third of patients with CD have *E. coli* with adherent and invasive qualities (adherent-invasive *E. coli* or AIEC) (Torres et al. 2017). AIEC can adhere to intestinal epithelial cells and replicate within macrophages, leading to increased secretion of TNFα which can trigger inflammation in the gut (Torres et al. 2017).

Much is known about the effects of IBD on the microbiome; however, less is known about the relationship between the microbiome of a pregnant IBD patient and their infant. Previous research has shown that the maternal microbiome plays a large role in a child's microbiome. For example, mother to infant microbial transmission is seen in the gut microbiome more than in the vaginal, skin, or oral microbiome (Ferretti et al. 2018). Furthermore, microbial strains acquired by infants from their mothers appear to be more persistent in the infants' gut than strains not acquired from their mothers (Ferretti et al. 2018). The impact of maternal IBD status on an infant's microbiome is beginning to emerge. For instance, pregnant IBD patients

have a lower alpha diversity in their first and second trimester compared to pregnant women without IBD (Torres et al. 2020). Furthermore, the infants of IBD mothers tend to have a lower bacterial diversity and altered bacterial composition compared to infants of non-IBD control mothers (Torres et al. 2020). The effect of IBD status of a mother on an infant was also explored by measuring levels of fecal calprotectin (FC), a biomarker of disease activity and inflammation (Kim et al. 2021). Pregnant IBD patients had higher levels of fecal calprotectin throughout pregnancy compared to pregnant non-IBD patients and infants born to IBD mothers had higher levels of FC as well (Kim et al. 2021). Studies such as these show that the disease status of a pregnant mother can affect the microbiome of their infant.

E. coli is present in 90% of the human population and it is an early colonizer in infants after birth, highlighting the importance of *E. coli* in the microbiome (Secher et al. 2016). Mucosa-associated *E. coli* is increased in CD patients (Rhodes 2007), and a third of CD patients are colonized by pathobiont AIEC. However, little is known about the relationship between *E. coli* and AIEC in the microbiome of pregnant patients with and without IBD and their offspring. This project aims to explore the relationship between the presence of *E. coli* and AIEC in pregnant mothers with and without IBD, and how IBD status of the mothers affects their infants' microbiome. The second aim is to explore the effects of the mode of delivery (C-section vs spontaneous vaginal delivery) on the infants' microbiome and *E. coli*/AIEC status. We hypothesize that IBD mothers will have more *E. coli* and AIEC than healthy mothers, and that infants born to IBD mothers will be colonized by maternal *E. coli*/AIEC. Furthermore, it is predicted that spontaneous vaginal delivery will result in a higher likelihood of having *E. coli*.

Materials and Methods

Metagenomic analysis

Shotgun metagenomic sequencing was performed for microbiota analysis, followed by taxonomic classification using Kraken by Mount Sinai as part of their MEChanisms Of disease traNsmission In Utero through the Microbiome (MECONIUM) cohort study. For this section of the study, 23 IBD mothers, 36 control mothers, 38 children from IBD mothers and 34 children from control mothers were studied. Of the 38 children from IBD mothers, 77 samples were provided (12 samples at the 7-day timepoint, 15 at 14-day, 13 at 3 months, 14 at 1 year, 6 at 18 months, 5 at 2 years, 5 at 3 years, 7 samples at 4-5 year timepoint). Of the 34 children from control mothers, 139 samples were provided (19 samples at the 7-day timepoint, 22 at 14-day, 25 at 3 months, 21 at 1 year, 13 at 18 months, 11 at 2 years, 13 at 3 years, 15 samples at 4-5 year timepoint). Children gave samples multiple times over the course of the MECONIUM cohort study. There were 17 samples from mothers who gave birth via C-section and 51 samples who gave birth via vaginal deliveries (some mothers in the cohort gave birth more than once). Of the children born to mothers who did not receive antibiotics, 112 samples were provided (16 samples at the 7-day timepoint, 17 at 14-day, 22 at 3 months, 20 at 1 year, 10 at 18 months, 9 at 2 years, 9 at 3 years, 9 samples at 4-5 year timepoint). Of the 128 children born to mothers who received antibiotics, 139 samples were provided (25 samples at the 7-day timepoint, 29 at 14-day, 17 at 3 months, 16 at 1 year, 10 at 18 months, 8 at 2 years, 10 at 3 years, 13 samples at 4-5 year timepoint).

The taxa table was normalized to one million bacterial reads per sample for bacterial community diversity analyses and ten thousand *E. coli* reads per sample for *E. coli* community diversity analyses. Alpha diversity of the microbiota was calculated using the number of

observed taxa and the Shannon diversity index. Alpha diversity differences between groups were evaluated using either the Mann–Whitney U test or Tukey's multiple comparison test on GraphPad Prism version 10.4.1. Beta diversity was calculated using the Bray–Curtis dissimilarity index, followed by principal coordinates analysis (PCoA). Both alpha and beta diversity were calculated using *mothur* and 90% confidence ellipses were generated using the mass 2 package in R. A PERMANOVA test was performed using the adonis2 function in the *vegan* package in R with 999 permutations and Bray-Curtis dissimilarity, to assess the effects of time and IBD status on microbial community composition (Schloss et al. 2009, Ripley et al. 2013, Dixon 2003). All metagenomic analyses and PCoA visualizations were provided by our collaborators at Mount Sinai.

Isolation and characterization of *E. coli*

Fecal samples from pregnant mothers in the third trimester, their infants (7 days, 14 days, 3 months, 1 year, 18 months, 4 years) were provided by the Icahn School of Medicine at Mount Sinai. These samples were collected as part of the MECONIUM cohort study at Mount Sinai. 70 fecal samples from mothers (40 with IBD and 30 without IBD), 29 samples from their infants (11 born to IBD mothers and 18 born to healthy mothers), and 43 samples from toddlers (27 born to IBD mothers and 16 born to healthy mothers) were received from Mount Sinai. These samples were stored at -80°C. To culture *E. coli*, fecal samples were incubated in a Gram-negative enrichment (GN) broth at a 37 °C incubator for 24 hours. Serial dilutions were prepared in PBS up to 10^{-7} and plated on MacConkey agar to isolate colonies. After overnight incubation at 37 °C, 10 pink colonies were randomly selected and purified on LB agar. Colonies were tested for *E. coli* using *E. coli*-specific primers (16S-1340) in a PCR assay. Afterwards, the phylogenetic group (A, B1, B2, D or F) for each isolate was determined using triplex PCR as previously

described by Clermont et al. (2013). Furthermore, the clonality of *E. coli* isolates was determined by random amplified polymorphic DNA (RAPD) PCR. Isolates with identical RAPD patterns (determined by non-quantitative visual analysis) and belonging to the same phylogenetic group were considered as the same strain. The primer used for RAPD PCR was 1283 as described by Akopyanz et al. (1992).

Virulence genes

The presence of *cnf1* and *pks* virulence genes associated with cytotoxic and genotoxic *E. coli* strains was determined by PCR. The presence of genes associated with AIEC pathotype, *lpfA141, lpfA154*, and *pduC*, was determined using PCR as previously described by Dogan et al. (2012, 2014). The presence of diarrheagenic virulence genes (*eae, stx1, stx2, lt, st, ipaH*, and *pCVD432*) was determined by PCR, with isolates containing these genes considered non-AIEC (Vidal et al. 2005).

Results

Metagenomic analysis

Principal Coordinate Analysis (PCoA) revealed that the bacterial communities from maternal samples form discrete clusters that are separate from infants. There is a slight difference in bacterial communities between control and IBD mothers, but broad overlap in the microbiome of babies independent of maternal IBD status.

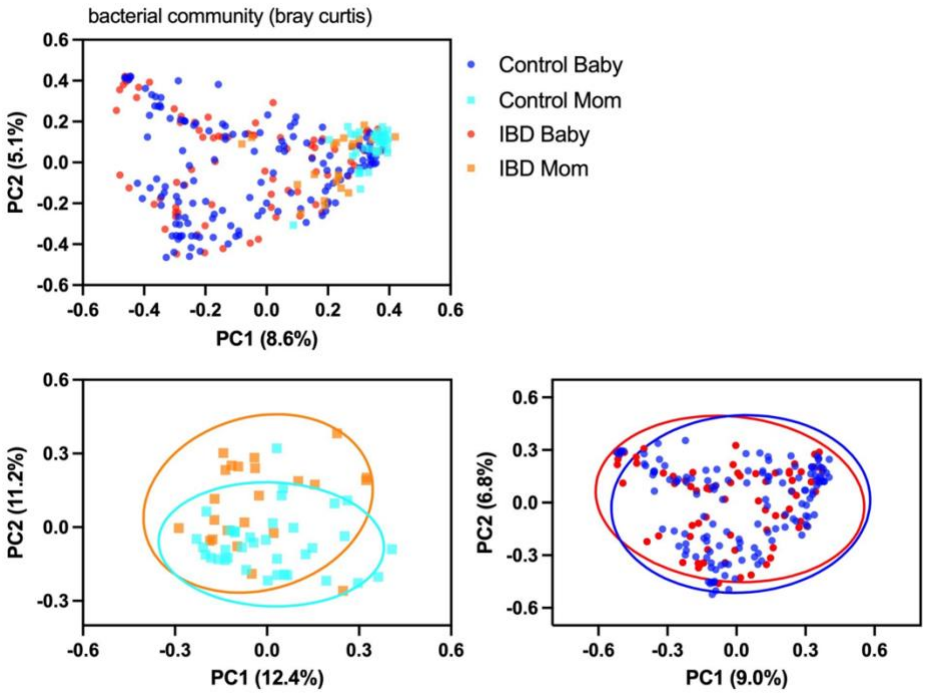


Figure 1. PCoA plot of bacterial community in IBD and non-IBD mothers and their infants, of IBD and non-IBD mothers (bottom-left), and control and infants born to IBD mothers (bottom-right). Bray-Curtis analysis used to measure dissimilarity in groups. IBD mother bacterial community had R^2 = 0.6%, F=3.32, p=0.001. IBD infants had R^2 = 0.6%, F=1.71, p=0.067 and time had R^2 = 20.7%, F=6.07, p=0.001.

PCoA analysis revealed that the *E. coli* community of maternal samples clusters separately from their infants, and a slight difference in *E. coli* communities between mothers and babies with IBD.

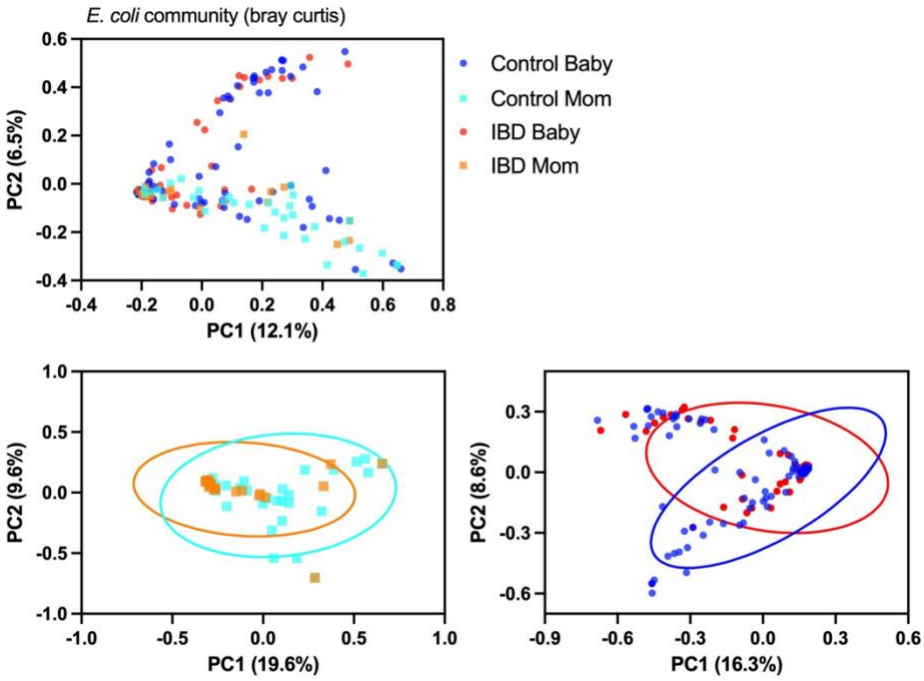


Figure 2. PCoA plot of *E. coli* sequences identified down to the strain level in IBD and non-IBD mothers and their infants, of IBD and non-IBD mothers (bottom-left), and control and infants born to IBD mothers (bottom-right). Bray-Curtis analysis used to measure dissimilarity in groups. IBD mother *E. coli* community had R^2 = 3.8%, F=2.06, p=0.059 and IBD infants had R^2 = 0.3%, F=0.69, p=0.671 and time had R^2 = 16.9%, F=4.37, p=0.00.

Observed taxa and alpha diversity was higher in control than IBD mothers. *E. coli* abundance was higher in mothers with IBD than control mothers.

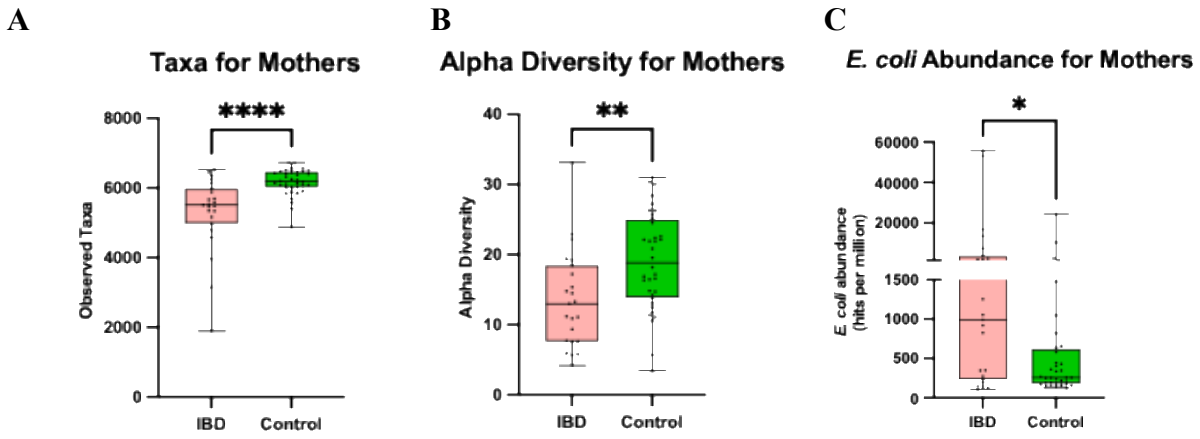
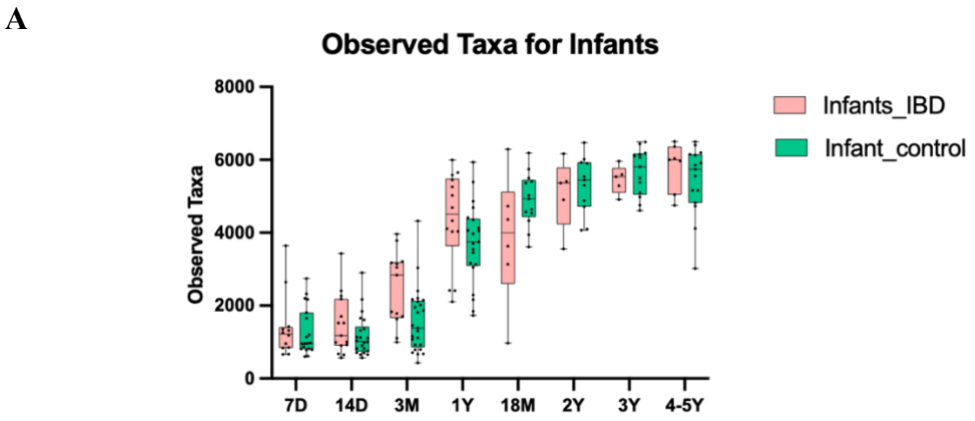
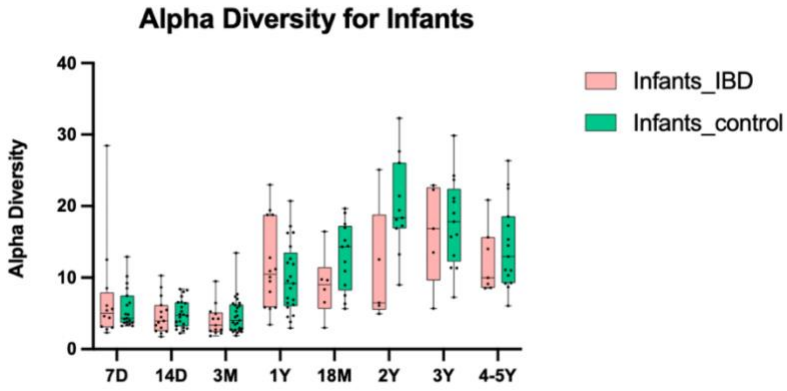


Figure 3. (**A**) Observed taxa for IBD and non-IBD mothers. Mann-Whitney test used with data shown as mean ± SD. **** $P < 0.0001$. (**B**) Alpha diversity for IBD and non-IBD mothers. Simpson Index of Diversity and Mann-Whitney test used with data shown as mean ± SD. ** $P < 0.01$. (**C**) *E. coli* abundance in hits per million for IBD and non-IBD mothers. Mann-Whitney test used with data shown as mean ± SD. * $P < 0.05$.

The observed taxa, alpha diversity, and *E. coli* abundance for infants remained similar between infants born to mothers with and without IBD. An increase in observed taxa and alpha diversity is seen in infants as they age from 7 days old to 4-5 years old. The abundance of infant *E. coli* decreased from the perinatal period to 4-5 years old.



B



C

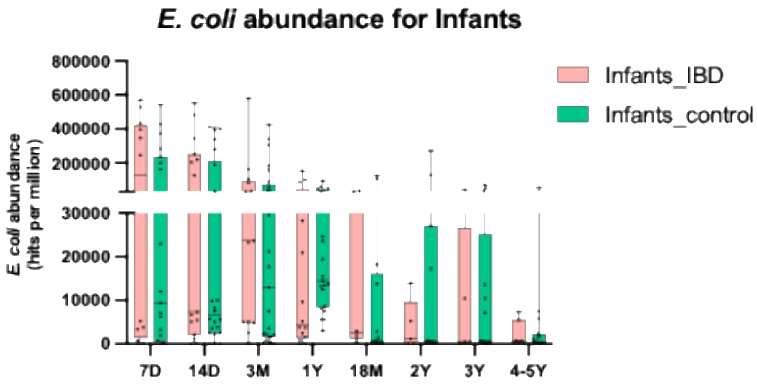


Figure 4. (**A**) Observed taxa for children of different age cohorts (7 days, 14 days, 3 months, 1 year, 18 months, 2 years, 3 years, 4-5 years) born to IBD and non-IBD mothers. Mann-Whitney test used with data shown as mean ± SD. (**B**) Alpha diversity for children born to IBD and non-IBD mothers. Simpson Index of Diversity and Mann-Whitney test used with data shown as mean ± SD. (**C**) *E. coli* abundance in hits per million for children born to IBD and non-IBD mothers. Mann-Whitney test used with data shown as mean ± SD.

Infants born via vaginal birth had more intestinal *E. coli* abundance than those born via C-section.

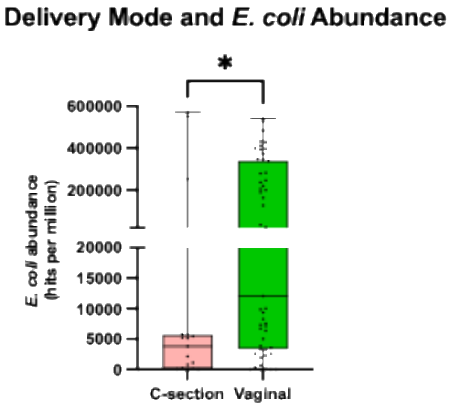


Figure 5. *E. coli* abundance for children born via C-section or vaginal birth. Mann-Whitney test used with data shown as mean ± SD. * $P < 0.05$.

E. coli abundance was similar in children who were born to mothers that did or did not receive antibiotics during birth and delivery stay at the hospital. Mothers that receive antibiotics often had C-sections as opposed to vaginal deliveries.

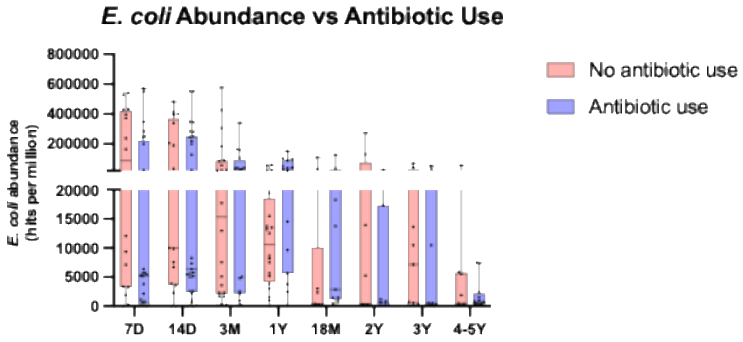


Figure 6. *E. coli* abundance for children born to mothers who did or did not receive antibiotics during delivery. Multiple Comparisons test used with data shown as mean ± SD.

Isolation and characterization of *E. coli*

Seventy fecal samples from mothers were evaluated. *E. coli* was isolated from 17/70 samples (9 from the control group and 8 from IBD mothers), with 22 different *E. coli* isolated.

RAPD PCR identified 1-3 *E. coli* strains per mother, with phylogroups A, B1, B2, D and F. In healthy mothers, *E. coli* were predominantly phylogroups B2 or D (5/12 isolates were B2 and 3/12 were D), whereas in IBD mothers, *E. coli* was mostly B1 (there were 6/10 B1 *E. coli* isolates) (Table 1).

Twenty-nine fecal samples from infants aged 7-90 days old were evaluated. *E. coli* was isolated from 7/29 infants (3 born to control mothers, 4 to IBD mothers), yielding 11 different *E. coli* isolates. RAPD PCR identified 1-2 *E. coli* strains per infant with 9 strains belonging to the phylogroup B2 (Table 1).

Forty-three fecal samples from toddlers aged 1-4 years old were evaluated. *E. coli* was isolated from 18/43 samples (5 from control mothers, 13 from IBD mothers), yielding 27 different *E. coli* isolates. RAPD PCR identified 1-3 *E. coli* strains per toddler, with phylogroups A, B1, B2, D and F. *E. coli* strains in toddlers were mostly B2 and D, and independent of maternal IBD status (Table 1).

Table 1. Isolation and characterization of *E. coli* from mothers and their babies.

Origin	Subjects	Health status	*E. coli* positive	*E. coli* strains (median, range/patient)	Phylogroup	/ individual				
						A	B1	B2	D	F
Mother	70	30 Healthy	9 (30%)	12 (1, 1-3)	1A, 2B1, 5B2, 3D, 1F	1	2	**4**	3	1
		40 IBD	8 (20%)	10 (1, 1-2)	6B1, 3B2, 1D	0	**5**	3	1	0
Baby (7 - 90 days)	29	18 infants of Healthy	3 (17%)	3 (1, 1)	2B2, 1D	0	0	**2**	1	0
		11 infants of IBD	4 (36%)	8 (2, 1-2)	7B2, 1F	0	0	**4**	0	1
Baby (1-4 years)	43	16 infants of Healthy	5 (31%)	8 (1, 1-3)	2A, 1B1, 3B2, 2D	2	1	2	2	0
		27 infants of IBD	13 (48%)	19 (1, 1-2)	2A, 4B1, 7B2, 5D, 1F	2	4	**7**	5	1

The distribution of *E. coli* phylogroups in toddlers resembles that of control mothers (Figure 7). However, the phylogroup distribution of the infants is enriched in B2 strains (Figure 7).

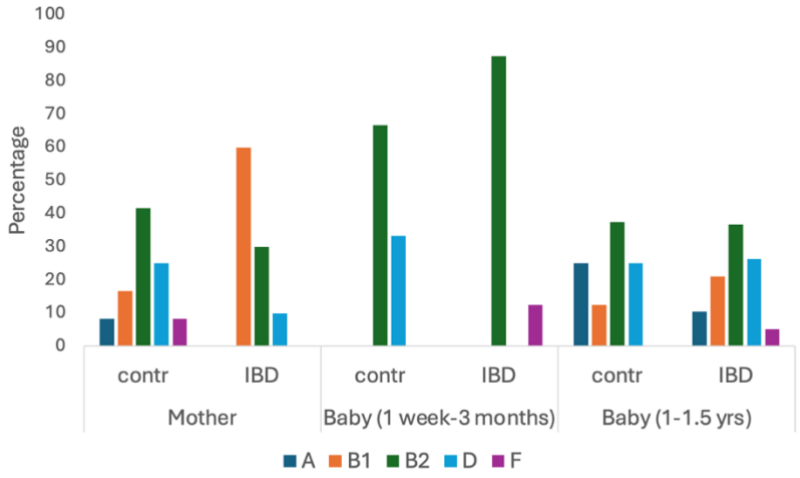


Figure 7. Distribution of *E. coli* phylogroups.

Comparison of maternal and infant *E. coli* strains: There were 5 instances when *E. coli* was isolated from a mother and her infant, with clonal *E. coli* strains identified in 1/5 mother: baby pairs (Table 2 and Figure 8).

Table 2. Characterization of *E. coli* isolated from mother-baby pairs.

Mother/Baby	ID	Disease	Phylogroup	Clonal pair
M1	907	IBD (CD)	B1	no
B1	2255 (7 days)		B1	no
	2167 (90 days)		B2	yes
	2829 (18 months)		B2	yes
M2	252	IBD (CD)	B2	yes
B2	326		B2	yes
M3	50043	IBD (UC)	B2	no
B3	414 (7days)		B2	7&90 days baby
	434 (90 days)		B2	B2
			F	
M4	871	control	A	no
B4	1886		B2	no
M5	938	control	D	no
B5	2448		B2	no

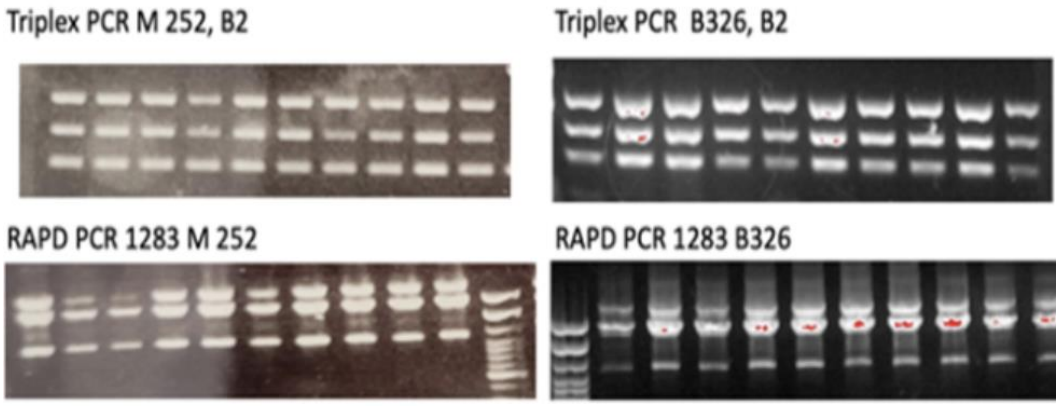


Figure 8. Gel electrophoresis of Triplex and RAPD PCR of the clonal *E. coli* strains of the mother/infant pair.

Virulence genes

Genes encoding cytotoxin and genotoxins (*cnf1, pks*) were rare in maternal *E. coli* (15% *cnf1*, 15% *pks* in control; 0% *cnf1*, 10% *pks* in IBD) but common in infants of IBD mothers (75% *cnf1*, 88% *pks*) and absent in infants of control mothers (Table 3). Conversely, AIEC associated genes (*lpfA141, lpfA154*, and *pduC*) were more common in maternal *E. coli* particularly IBD mothers (31% *lpfA141*, 31% *lpfA154*, 0% *pduC* in control; 50% *lpfA141*, 60% *lpfA154*, 30% *pduC* in IBD mothers) than infant *E. coli* (0% *lpfA141*, 33% *lpfA154*, 33% *pduC* in infants of control; 0% *lpfA141*, 13% *lpfA154*, 0% *pduC* in infants of IBD). Similar to the maternal cohort, cytotoxin and genotoxin genes (*cnf1, pks*) were rare in toddler *E. coli* (13% *cnf1*, 0% *pks* in control; 0% *cnf1,* 11% *pks* in toddlers born to IBD mothers) but toddlers had more genes associated with AIEC (0% *lpfA141*, 25% *lpfA154*, 38% *pduC* in control; 16% *lpfA141*, 53% *lpfA154*, 47% *pduC* in toddlers born to IBD mothers) (Table 3).

Table 3. Virulence genes of maternal, infant (7-90 days old), and toddler (12-18 months) *E. coli.*

			Virulence Genes					
Mother /Baby	**IBD /control**	***E. coli* phylogroup**	*cnf1*	*pks*	*lpfA141*	*lpfA154*	*pduC*	Non Lactose Fermenter
Mother	control	A						
		B1			+	+		
					+	+		
		B2			+			
			+	+				
			+	+	+			
		D						
						+		
		F				+		
	IBD	B1			+	+		
					+	+		
					+	+		
					+	+		
						+		
					+	+	+	
		B2		+				
							+	NLF
							+	
		D						
Baby (7-90 days)	control	B2						NLF
								NLF
		D				+	+	
	IBD	B2	+	+				
			+	+				
				+				
			+	+				
			+	+				
			+	+				
			+	+				
		F				+		
Baby (12-18 months)	control	A						
		B1				+		
		B2	+					
							+	
		D				+	+	
							+	
	IBD	A						
		B1				+		
						+	+	
					+	+	+	
					+	+	+	
		B2						
				+				
				+				
					+		+	
		D				+		
						+	+	
						+	+	
		F				+		

Discussion

Early life events have an important role in shaping the microbiome within the first 2-3 years of life. Events such as mode of delivery, feeding patterns, antibiotic use and maternal health status can all play an important role in the development of the gut microbiota and risk for Inflammatory Bowel Disease (IBD) later in life (Sabino et al. 2024). This study explored the effects of a mother's IBD disease status on microbiome diversity focusing on *E. coli* phylogeny and virulence. Additionally, this study explored the effect of a mother's IBD status, delivery mode, and antibiotic use on their infant's microbiome diversity. By considering maternal disease status and early life exposures such as mode of delivery and antibiotic exposure, our study provides important information of how these variables can affect the microbiome.

Metagenomic analysis

Metagenomic analysis revealed that the observed taxa and alpha diversity for control mothers was higher than IBD mothers. Previous studies have shown that pregnant IBD patients had a lower alpha diversity in their first and second trimester compared to pregnant women without IBD, which is consisted with our findings (Torres et al. 2020). Furthermore, we observed that the taxa, alpha diversity, and *E. coli* abundance for infants remained similar between infants born to IBD and non-IBD mothers. Although no significant difference in taxa and alpha diversity were observed for infants born to IBD vs non-IBD mothers, prior studies have shown that infants of IBD mothers tended to have a lower bacterial diversity and altered bacterial composition compared to infants of non-IBD mothers (Torres et al. 2020). We also observed an increase in observed taxa and alpha diversity in infants as they aged from 7 days old to 4-5 years old. This finding is consistent with previous studies which have shown an increase in bacterial alpha-

diversity in infants between their first 3 to 24 months of life (Sabino et al. 2024 and Yatsunenko et al. 2012).

Our findings indicate that *E. coli* abundance for control mothers was lower than IBD mothers, which is consistent with studies reporting an increase in mucosa-associated *E. coli*, notably AIEC, in Crohn's Disease patients (Rhodes 2007; Baumgart et al. 2008). *E. coli* has emerged as an early colonizer of neonatal infants that can promote an anaerobic environment that supports colonization by obligate anaerobic bacteria such as *Bifidobacterium, Clostridium* and *Bacteroides* (Secher et al. 2016). In our study, we observed an increase in observed taxa and alpha diversity as infants aged from 7 days old to toddlers, with a concurrent decrease in *E. coli* at the same time. This pattern provides insight into the early *E. coli* colonization that occurs in infants followed by the more complex and diverse system of bacterial colonization as a child grows and experiences changes in eating patterns from relying on breastmilk and/or formula milk to a diet higher in sugar (Yatsunenko et al. 2012).

Previous studies have shown that delivery mode (C-section vs vaginal birth) can be associated with differences in bacterial colonization, with these differences becoming less pronounced as a child ages (Stokholm et al. 2016). Consistent with our findings, previous studies have also shown higher numbers of *E. coli* with vaginal births versus C-section. Differences in bacterial colonization between C-section vs vaginal birth have also been found with other bacterial species revealing a positive correlation between *Clostridium, Enterobacter cloacae*, and *Enterococcus faecalis* species and C-sections (Stokholm et al. 2016). These findings suggest that the mode of delivery plays an important role in shaping an infant's gut microbiome.

In addition to examining the effects of delivery mode, our study also explored the relationship between *E. coli* abundance and maternal intrapartum antibiotic use. We found that *E.*

coli abundance in infants was similar regardless of whether their mothers received antibiotics during labor. It is worth noting that in our study, mothers who did receive antibiotics often had C-sections as opposed to vaginal deliveries. However, previous studies have shown that such use of antibiotics associated with C-sections have shown lower numbers of *E. coli* **(**Stokholm et al. 2016). The potential link between delivery mode, antibiotic use and changes in *E. coli* abundance raises important questions about IBD risk later in life. Future studies could explore if a lower number of *E. coli* abundance due to C-sections and intrapartum antibiotics can be used to predict an increase in IBD risk later in life. This would be an important area of research given that C-sections and exposure to antibiotics have been correlated to increase IBD risk later in life (Torres et al. 2020 and Agrawal et al. 2021).

Isolation and characterization of *E. coli*

Through metagenomic analysis we found that IBD mothers have increased abundance and diversity of *E. coli* compared to controls. However, our culture-based analysis found that IBD mothers and non-IBD mothers had roughly the same number of *E. coli* strains within the mother cohort (10/22 vs 12/22 *E. coli* strains, respectively). This discrepancy in results may be due to challenges in isolating *E. coli* from archival fecal samples. Furthermore, both infants and toddlers born to mother with IBD had a higher number of *E. coli* isolates than infants and toddlers born to non-IBD mothers. The increase of *E. coli* isolates in younger children as opposed to their mothers may be due to *E. coli*'s role as an early colonizer in the microbiome (Secher et al. 2016). Furthermore, using a chi-square test we found that there is statistical significance in phylogroup distribution between IBD mothers, IBD infants and IBD toddlers ($p < 0.001$) and there is also statistical difference in phylogroup distribution of healthy mothers, infants and toddlers ($p < 0.01$). We also hypothesized that maternal *E. coli* strains would be

directly transmitted to infants; however, we found clonal *E. coli* in only 1/5 mother baby pairs, emphasizing that the colonization of *E. coli* in the infants due to the mother's microbiome may not be as strong as predicted.

Previous literature has found that the microbiota evolves and matures by the time a child reaches three years old (Secher et al. 2016). Our finding that the distribution of *E. coli* phylogroups in toddlers is more like non-IBD mothers, than infants support the age-related transition of the microbiome.

Virulence genes

Adherent and Invasive *E. coli* (AIEC) has been linked to inflammatory bowel disease (IBD), particularly Crohn's disease (Zhang et al. 2022; Palmela et al. 2018) and colorectal cancer (CRC) (Arthur et al. 2012). The *lpfA* and *pduC* virulence genes are enriched in AIEC associated with CD (Dogan et al. 2014; Zhang et al. 2022) whereas the pks gene cluster which encodes for a genotoxin, Colibactin (Secher et al. 2016), is linked to intestinal inflammation, DNA damage and CRC (Arthur et al. 2012). Our findings suggest that phylogeny and virulence gene content of *E. coli* vary with age groups and maternal disease status. Infants (7-90 days old) showed a predominance of B2, *pks*+, *cnf*+ associated with colorectal cancer CRC, while toddlers showed a higher frequency of B2 strains containing the AIEC-associated *lpfA*+, *pduC*+ genes. Using a chi-square test, we also found that there was statistical significance between *pks*+, *cnf*+ presence in infants versus toddlers born to IBD mothers ($p < 0.0001$ and $p < 0.001$, respectively). Furthermore, IBD mothers favored B1, *lpfA*+, *pduC*+ isolates, consistent with previous studies of CD-associated *E. coli* and AIEC (Zhang et al. 2022) and non-IBD mothers favored B2, D, *lpfA*+ isolates.

Intriguingly, *pks* was only found in B2 strains, consistent with previous studies, and all the infants (7-90 days old) colonized with B2 strains born to IBD mothers, were positive for *pks*. This observation highlights a potential association between B2 phylogroups and *pks* presence, suggesting that IBD may influence whether *pks* is present.

A major limitation of this study was the low number of infants and toddlers compared to the mother cohort, which may have impacted the distribution of phylogroups observed within the cohort. This may be especially true for the case of healthy infants (7-90 days old), as we were only able to isolate three *E. coli* strains. Furthermore, there were some challenges in isolating *E. coli* from fecal samples obtained and archived by our collaborators at Mount Sinai. This may relate to direct freezing without use of cryo-preservatives.

In the future, it would be beneficial to explore the presence of virulence genes using the metagenomic data provided by Mount Sinai to supplement what has already been identified using PCR. Additionally, for samples that we were unable to isolate strains, it would be interesting to investigate metagenomic data to determine if *E. coli* strains are indeed present in those samples and to determine the phylogroups of these strains.

Acknowledgements

I would like to thank everyone in the Simpson Lab at the Cornell University College of Veterinary Medicine, especially Dr. Belgin Dogan, Dr. Shiying Zhang, and Dr. Kenneth Simpson for the opportunity to complete this project. I would also like to thank the undergraduate students in the lab for their support as well.

Citations

Agrawal M, Sabino J, Frias-Gomes C, Hillenbrand CM, Soudant C, Axelrad JE, Shah SC, Ribeiro-Mourão F, Lambin T, Peter I, Colombel JF, Narula N, Torres J. Early life exposures and the risk of inflammatory bowel disease: Systematic review and meta-analyses. EClinicalMedicine. 2021 May 15;36:100884. doi: 10.1016/j.eclinm.2021.100884. PMID: 34308303; PMCID: PMC8257976.

Akopyanz N, Bukanov NO, Westblom TU, Kresovich S, Berg DE. DNA diversity among clinical isolates of Helicobacter pylori detected by PCR-based RAPD fingerprinting. Nucleic Acids Res. 1992 Oct 11;20(19):5137-42. doi: 10.1093/nar/20.19.5137. PMID: 1408828; PMCID: PMC334296.

Arthur JC, Perez-Chanona E, Mühlbauer M, Tomkovich S, Uronis JM, Fan TJ, Campbell BJ, Abujamel T, Dogan B, Rogers AB, Rhodes JM, Stintzi A, Simpson KW, Hansen JJ,]Keku TO, Fodor AA, Jobin C. Intestinal inflammation targets cancer-inducing activity of the microbiota. Science. 2012 Oct 5;338(6103):120-3. doi: 10.1126/science.1224820. Epub 2012 Aug 16. PMID: 22903521; PMCID: PMC3645302.

Baumgart, M., Dogan, B., Rishniw, M. *et al.* Culture independent analysis of ileal mucosa

reveals a selective increase in invasive *Escherichia coli* of novel phylogeny relative to depletion of Clostridiales in Crohn's disease involving the ileum. *ISME J* **1**, 403–418 (2007). https://doi.org/10.1038/ismej.2007.52

Clermont, O.; Christenson, J.K.; Denamur, E.; Gordon, D.M. The Clermont Escherichia coli phylo-typing method revisited: Improvement of specificity and detection of new phylo-groups. Environ. Microbiol. Rep. 2013, 5, 58–65. [CrossRef]

Dixon P. VEGAN, a package of R functions for community ecology. Journal of vegetation science. 2003;14(6):927-930.

Dogan, B.; Rishniw, M.; Bruant, G.; Harel, J.; Schukken, Y.H.; Simpson, K.W. Phylogroup and lpfA influence epithelial invasion by mastitis associated Escherichia coli. Vet. Microbiol. 2012, 159, 163–170.

Dogan, B.; Suzuki, H.; Herlekar, D.; Sartor, B.R.B.; Campbell, B.J.; Roberts, C.L.; Stewart, K.; Scherl, E.J.; Araz, Y.; Bitar, P.P.; et al. Inflammation-associated adherent-invasive escherichia coli are enriched in pathways for use of propanediol and iron and M-cell translocation. Inflamm. Bowel Dis. 2014, 20, 1919–1932. [CrossRef]

Ferretti, P., Pasolli, E., Tett, A., Asnicar, F., Gorfer, V., Fedi, S., Armanini, F., Truong, D. T., Manara, S., Zolfo, M., Beghini, F., Bertorelli, R., De Sanctis, V., Bariletti, I., Canto, R., Clementi, R., Cologna, M., Crifò, T., Cusumano, G., Gottardi, S., … Segata, N. (2018). Mother-to-Infant Microbial Transmission from Different Body Sites Shapes the Developing Infant Gut Microbiome. *Cell host & microbe*, *24*(1), 133–145.e5. https://doi.org/10.1016/j.chom.2018.06.005

Halfvarson, J., Brislawn, C., Lamendella, R. *et al.* Dynamics of the human gut microbiome in

inflammatory bowel disease. *Nat Microbiol* **2**, 17004 (2017). https://doi.org/10.1038/nmicrobiol.2017.4

Kim, E. S., Tarassishin, L., Eisele, C., Barre, A., Nair, N., Rendon, A., Hawkins, K., Debebe, A., White, S., Thjømøe, A., Mørk, E., Bento-Miranda, M., Panchal, H., Agrawal, M., Patel, A., Chen, C. L., Kornbluth, A., George, J., Legnani, P., Maser, E., … Mount Sinai Road to Prevention Study Group (2021). Longitudinal Changes in Fecal Calprotectin Levels Among Pregnant Women With and Without Inflammatory Bowel Disease and Their Babies. *Gastroenterology*, *160*(4), 1118–1130.e3. https://doi.org/10.1053/j.gastro.2020.11.050

Palmela, C., Chevarin, C., Xu, Z., Torres, J., Sevrin, G., Hirten, R., Barnich, N., Ng, S. C., & Colombel, J. F. (2018). Adherent-invasive *Escherichia coli* in inflammatory bowel disease. *Gut*, *67*(3), 574–587. https://doi.org/10.1136/gutjnl-2017-314903

Rhodes J. M. (2007). The role of Escherichia coli in inflammatory bowel disease. *Gut*, *56*(5), 610–612. https://doi.org/10.1136/gut.2006.111872

Ripley B, Venables B, Bates DM, et al. Package 'mass'. Cran r. 2013;538(113-120):822.

Sabino J, Tarassishin L, Eisele C, Hawkins K, Barré A, Nair N, Rendon A, Debebe A, Picker M, Agrawal M, Stone J, George J, Legnani P, Maser E, Chen CL, Thjømøe A, Mørk E, Dubinsky M, Hu J, Colombel JF, Peter I, Torres J. Influence of Early Life Factors, including breast milk Composition, on the Microbiome of Infants Born to Mothers with and without Inflammatory Bowel Disease. J Crohns Colitis. 2023 Nov 24;17(11):1723-1732. doi: 10.1093/ecco-jcc/jjad096. Erratum in: J Crohns Colitis. 2024 Aug 6;18(7):1174. doi: 10.1093/ecco-jcc/jjae021. PMID: 37279927; PMCID: PMC10673817.

Schloss PD, Westcott SL, Ryabin T, et al. Introducing mothur: open-source, platform-

independent, community-supported software for describing and comparing microbial communities. Applied and environmental microbiology. 2009;75(23):7537-7541. doi:https://doi.org/10.1128/AEM.01541-09

Secher T, Brehin C, Oswald E. Early settlers: which *E. coli* strains do you not want at birth? Am J Physiol Gastrointest Liver Physiol. 2016 Jul 1;311(1):G123-9. doi: 10.1152/ajpgi.00091.2016. Epub 2016 Jun 10. PMID: 27288422.

Stokholm J, Thorsen J, Chawes BL, Schjørring S, Krogfelt KA, Bønnelykke K, Bisgaard H. Cesarean section changes neonatal gut colonization. J Allergy Clin Immunol. 2016 Sep;138(3):881-889.e2. doi: 10.1016/j.jaci.2016.01.028. Epub 2016 Apr 1. PMID: 27045582.

Torres, J., Hu, J., Seki, A., Eisele, C., Nair, N., Huang, R., Tarassishin, L., Jharap, B., Cote-Daigneault, J., Mao, Q., Mogno, I., Britton, G. J., Uzzan, M., Chen, C. L., Kornbluth, A., George, J., Legnani, P., Maser, E., Loudon, H., Stone, J., … Peter, I. (2020). Infants born to mothers with IBD present with altered gut microbiome that transfers abnormalities of the adaptive immune system to germ-free mice. *Gut*, *69*(1), 42–51. https://doi.org/10.1136/gutjnl-2018-317855

Torres, J., Mehandru, S., Colombel, J.-F., & Peyrin-Biroulet, L. (2017). Crohn's disease. *The Lancet*, *389*(10080), 1741–1755. https://doi.org/10.1016/S0140-6736(16)31711-1

Vidal, M., Kruger, E., Duran, C., Lagos, R., Levine, M., Prado, V., Toro, C., & Vidal, R. 2005. Single Multiplex PCR Assay To Identify Simultaneously the Six Categories of Diarrheagenic Escherichia coli Associated with Enteric Infections. J. Clin. Microbiol. 43(10): 5362-5365

Yatsunenko, T., Rey, F., Manary, M. *et al.* Human gut microbiome viewed across age and geography. *Nature* **486**, 222–227 (2012). https://doi.org/10.1038/nature11053

Zhang S, Morgan X, Dogan B, Martin FP, Strickler S, Oka A, Herzog J, Liu B, Dowd SE, Huttenhower C, Pichaud M, Dogan EI, Satsangi J, Longman R, Yantiss R, Mueller LA, Scherl EJ, Sartor RB, Simpson KW. Mucosal metabolites fuel the growth and virulence of E. coli linked to Crohn's disease. JCI Insight. 2022 May 23;7(10):e157013. doi: 10.1172/jci.insight.157013. PMID: 35413017; PMCID: PMC9220930.