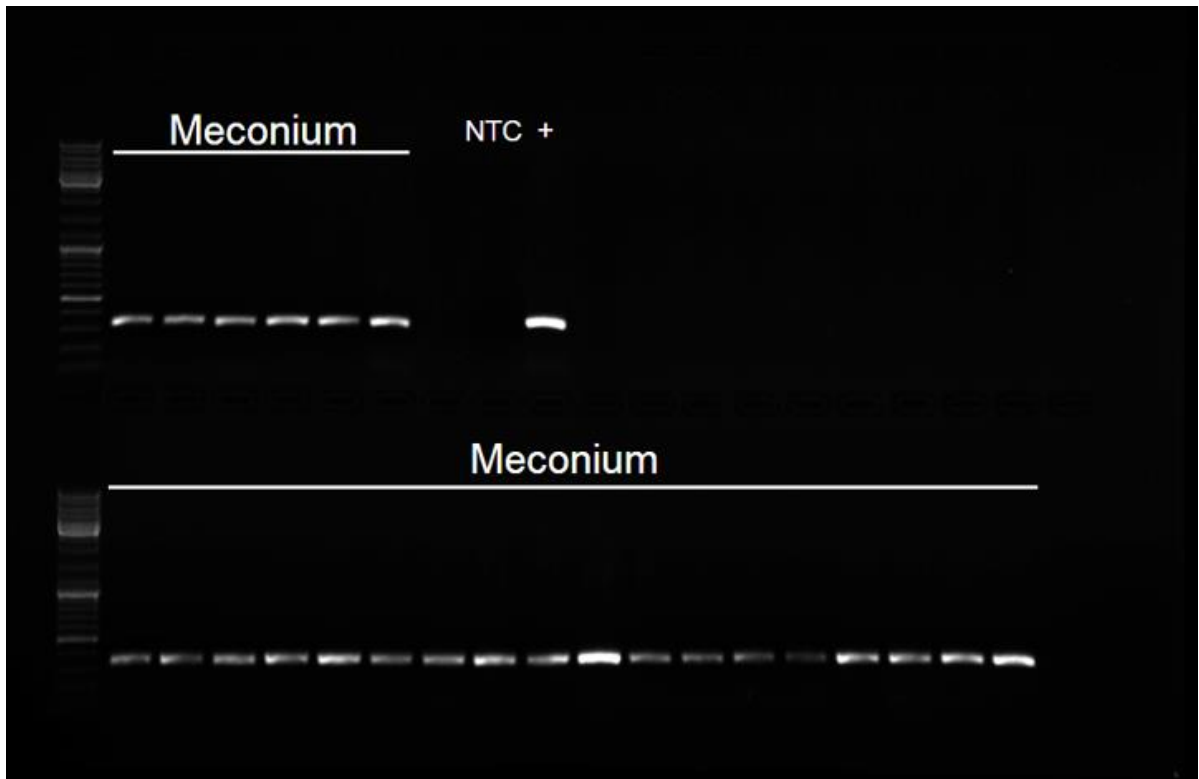
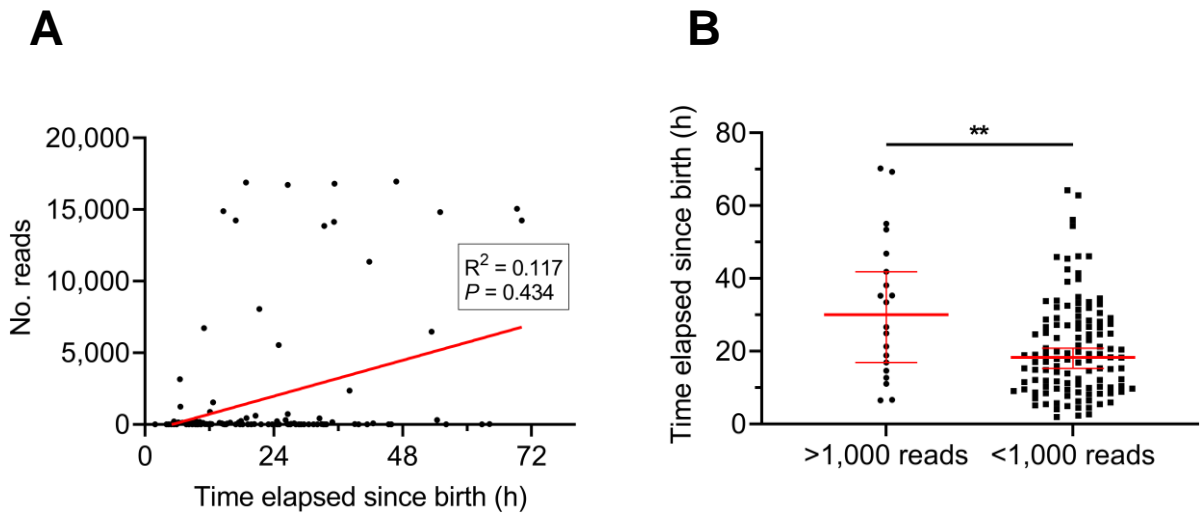


Supplementary Table S1: Primers used in this study (I = inosine; Y = C/T; K = G/T; R = A/G; S = G/C).

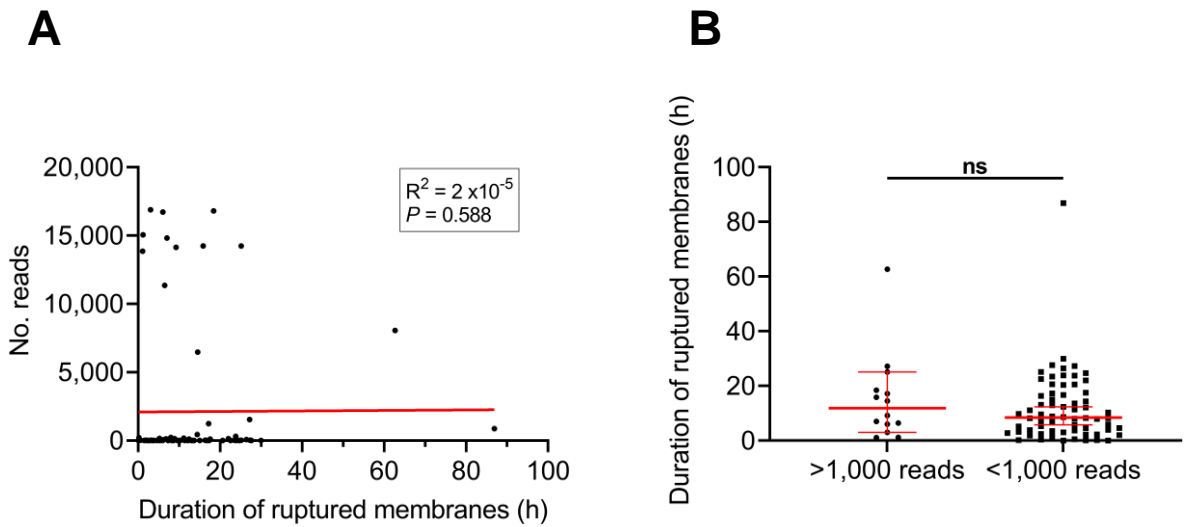
Target	Forward primer	Reverse primer
<i>cpn60</i> (containing Illumina index primer landing sites; underlined)	M279: 5'- <u>TCG TCG GCA GCG TCA GAT GTG</u> <u>TAT AAG AGA CAG GAI</u> III GCI GGI GAY GGI ACI ACI AC-3'	M280: 5'- <u>GCT TCG TGG GCT CGG AGA TGT GTA</u> <u>TAA GAG ACA</u> GYK IYK ITC ICC RAA ICC IGG IGC YTT-3'
	M1612: 5'- <u>TCG TCG GCA GCG TCA GAT GTG</u> <u>TAT AAG AGA CAG GAI</u> III GCI GGY GAC GGY ACS ACS AC-3'	M1613: 5'- <u>GCT TCG TGG GCT CGG AGA TGT</u> <u>GTA TAA GAG ACA</u> GCG RCG RTC RCC GAA GCC SGG IGC CTT-3'
Human mitochondrial cytochrome c oxidase I	JH0241: 5'-CAC CTT CTT CGA CCC CGC CG-3'	JH0242: 5'-TGC TTC CGT GGA GTG TGG CG-3'
16S rRNA gene (qPCR)	SRV3-1: 5'-CGG YCC AGA CTC CTA C-3'	SRV3-2: 5'-TTA CCG CGG CTG CTG GCA C-3'
<i>cpn60</i> (containing M13 sequencing primer landing sites; underlined)	H729: 5'-CGC CAG GGT TTT CCC AGT CAC <u>GAC</u> GAI III GCI GGI GAY GGI ACI ACI AC-3'	H730: 5'-AGC GGA TAA CAA TTT CAC ACA GGA YKI YKI TCI CCR AAI CCI GGI GCY TT-3'



Supplementary Figure S1 All meconium samples contain amplifiable DNA: (A) Representative ethidium bromide gel confirming amplification of the human mitochondrial COX-1 gene from meconium samples. All samples were PCR-positive for COX-1, ruling out PCR inhibition as a cause of low read numbers observed in meconium microbiome profiles (NTC = no-template control; + = positive control, total genomic DNA extracted from human vaginal swab fluid).



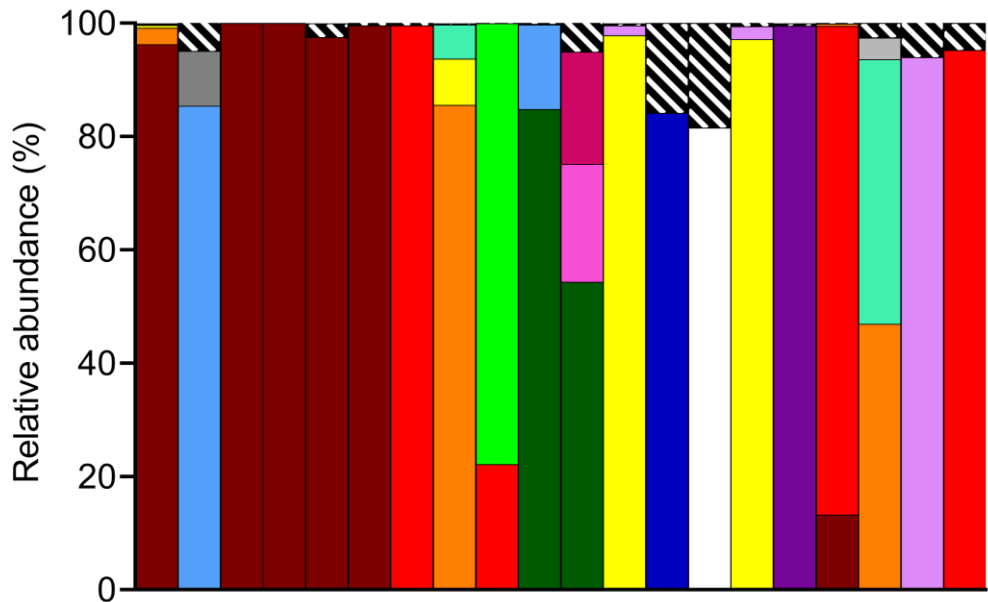
Supplementary Figure S2 Longer intervals from birth to sample collection in meconium samples with higher read numbers: (A) Number of hours elapsed from birth to sample collection is plotted against read counts for 136 meconium samples. Time since birth showed a weakly positive but non-significant correlation with read numbers ($R^2 = 0.117$; Spearman's rank, $P = 0.434$, $\rho = 0.096$). (B) Number of hours elapsed from birth to collection time for all meconium samples. Mean number of hours elapsed since birth was significantly higher for samples with >1,000 reads (Mann-Whitney U test, $P < 0.01$, $U = 766.5$). Red line indicates median ($\pm 95\%$ CI).



Supplementary Figure S3 No difference in duration of ruptured membranes in meconium samples with higher read numbers: (A) Number of hours elapsed from membrane rupture to birth is plotted against read counts for meconium samples obtained from infants delivered vaginally or by emergent caesarean ($n = 80$). Duration of ruptured membranes showed no significant correlation with read numbers ($R^2 = 2 \times 10^{-5}$; Spearman's rank, $P = 0.588$, $\rho = 0.061$). (B) Number of hours elapsed from membrane rupture to birth for all meconium samples grouped by read numbers. Mean number of hours between membrane rupture and birth was not significantly different (Mann-Whitney U test, $P = 0.454$, $U = 402$). Red line indicates median ($\pm 95\%$ CI).

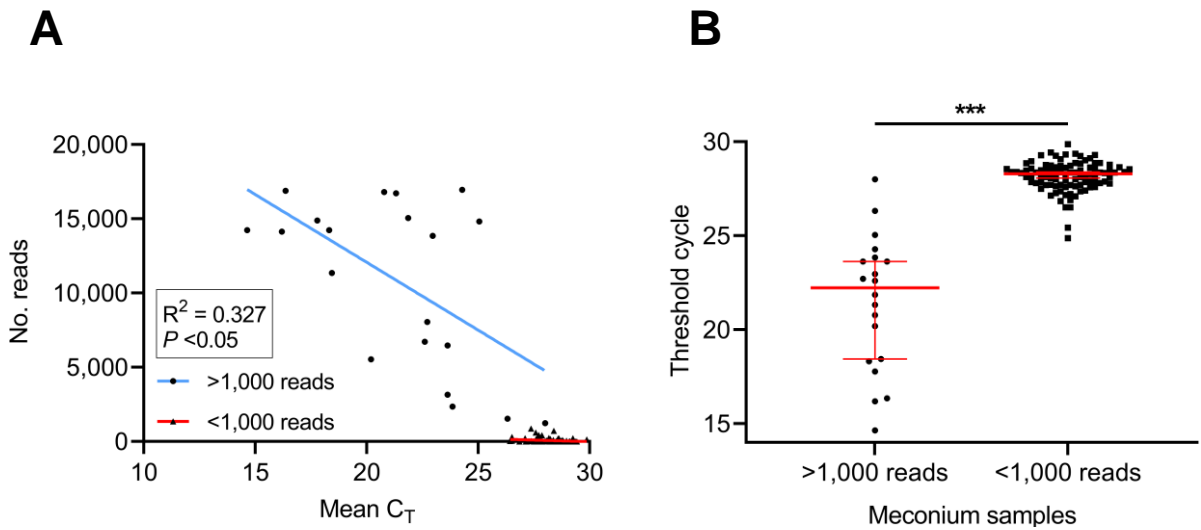
Supplementary Table S2: Read count summaries for the most common nearest neighbours found in at least 4 samples or controls across the entire dataset. Note: % similarity denotes the range of nucleotide similarities of all ASVs collapsed into a given nearest neighbour to the best match in cpnDB (* = single ASV).

Nearest neighbour (NN)	# Samples (control)	# Reads (control)			% Similarity
		Total	Median	Range	
<i>Hyphomicrobium zavarzinii</i>	24 (3)	397 (156)	12 (57)	107 - 2 (77 - 22)	80.6 - 80.2
<i>Shigella sonnei</i>	18 (4)	81,037 (449)	12 (89)	16961 - 1 (271 - 1)	100 - 98.4
<i>Bradyrhizobium arachidis</i>	14 (2)	149 (6)	7 (3)	44 - 2 (3)	90.0 - 89.6
<i>Chelatococcus</i> spp. GW1	11 (1)	137 (2)	10 (2)	26 - 1 (2)	77.7 - 76.9
<i>Lactobacillus crispatus</i>	8 (2)	172 (17)	3 (9)	143 - 1 (15 - 2)	100 - 99.6
<i>Enterococcus faecalis</i>	8 (0)	2,239 (0)	130 (n/a)	1,453 - 3 (n/a)	99.6 - 99.2
<i>Ralstonia insidiosa</i>	8 (0)	120 (0)	10 (n/a)	39 - 2 (n/a)	99.2 - 98.4
<i>Parabacteroides</i> spp. D25	7 (0)	19,373 (0)	60 (n/a)	11,854 - 2 (n/a)	100 - 99.2
<i>Escherichia coli</i>	6 (0)	38,537 (0)	5,710 (n/a)	14832 - 12 (n/a)	99.2 - 98.4
<i>Lactobacillus gasseri</i>	5 (0)	652 (0)	7 (n/a)	420 - 1 (n/a)	100 - 97.2
<i>Bacteroides vulgatus</i>	5 (0)	18,826 (0)	1,130 (n/a)	11,031 - 7 (n/a)	100 - 99.6
<i>Propionibacterium</i> spp. 434-HC2	4 (1)	40 (7)	11 (7)	15 - 3 (7)	100 - 98.8
<i>Shigella flexneri</i>	5 (0)	13,241 (0)	2 (0)	13,138 - 1 (n/a)	99.6 - 99.2
<i>Moraxella osloensis</i>	5 (0)	361 (0)	29 (n/a)	194 - 2 (n/a)	97.6 - 96.4
<i>Pseudomonas lactis</i>	5 (0)	259 (0)	24 (n/a)	147 - 6 (n/a)	96.4 - 96.0
<i>Streptococcus mitis</i>	4 (0)	359 (0)	25 (n/a)	306 - 3 (n/a)	100 - 92.8
<i>Rothia mucilaginosa</i>	4 (0)	7,173 (0)	1,207 (n/a)	4,736 - 24 (n/a)	98.8 - 85.9
<i>Lysobacter spongijicola</i>	4 (0)	124 (0)	28 (n/a)	62 - 6 (n/a)	* 89.9
<i>Arthrobacter nitrophenolicus</i>	3 (1)	79 (5)	22 (5)	54 - 3 (5)	* 78.7

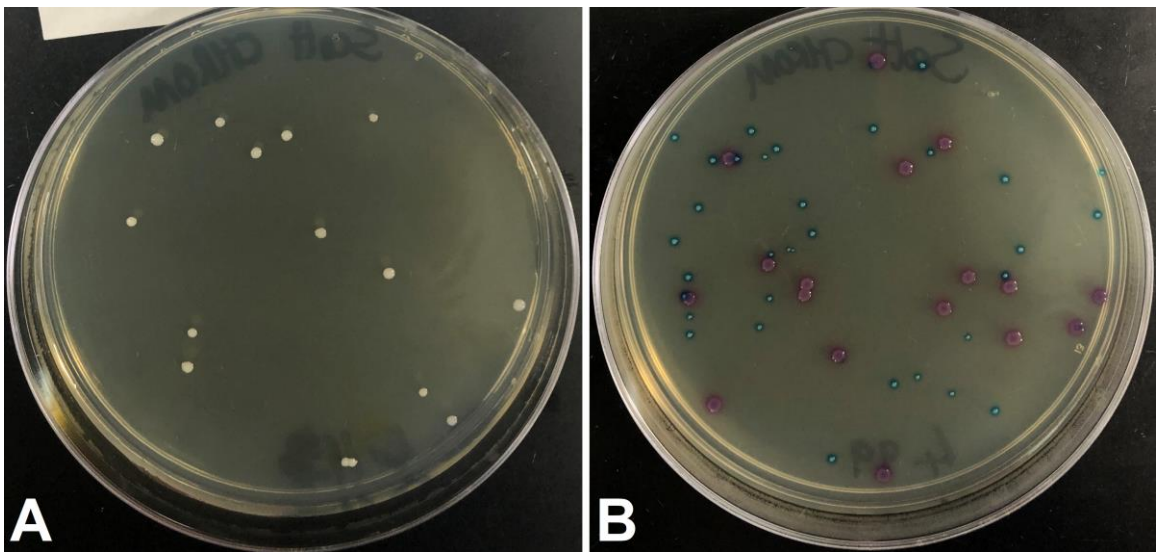


- Shigella sonnei*
- Escherichia coli*
- Parabacteroides* spp. D25
- Bacteroides vulgatus*
- Streptococcus parasanguinis*
- Shigella flexneri*
- Bifidobacterium longum* subsp. *longum*
- Rothia mucilaginosa*
- Porphyromonas uenonis*
- Clostridiales* spp. KA00274
- Enterococcus faecalis*
- Bacteroides uniformis*
- Haemophilus parainfluenzae*
- Flavobacterium psychrolimnae*
- Bifidobacterium adolescentis*
- Streptococcus pseudopneumoniae*
- Other

Supplementary Figure S4 High read number meconium samples are comprised of common enteric species often found in stool: Microbiome profiles of the 20 meconium samples yielding >1,000 *cpn60* reads, expressed as relative abundance. All but one of these samples are dominated by a single NN, demonstrating the simplicity of the neonatal microbiome compared to that of older infants and adults. Common bacterial species known to inhabit the gastrointestinal tract of humans, such as *E. coli* / *Shigella* spp., *Bacteroides vulgatus* and *Streptococcus* spp. are abundant across this subset of samples.



Supplementary Figure S5 Read numbers and 16S rRNA qPCR cycle thresholds are well-correlated: (A) Mean 16S rRNA qPCR threshold cycles (CT; from duplicate reactions) plotted against number of reads for 114 samples with both data available. CT values of samples with >1,000 reads show a significant inverse correlation with read number ($R^2 = 0.327$; Spearman's rank, $P < 0.05$, $\rho = -0.447$) while those with <1,000 reads do not ($R^2 = 0.035$; Spearman's rank, $P = 0.252$, $\rho = -0.116$). (B) Meconium samples with >1,000 reads had a significantly lower threshold cycle than those with <1,000 reads (Mann-Whitney U test, $P < 0.01$, $U = 38.5$). Red line indicates median ($\pm 95\%$ CI).



Supplementary Figure S6 Undiluted meconium cultures yield very few colonies: Two representative plates of CHROM Orientation agar inoculated with meconium (resuspended in PBS without serial dilution) incubated at 37°C overnight under aerobic conditions. Plates are representative of the entire dataset and exhibit a small number of colonies of (A) *Staphylococcus epidermidis*, (B) *Escherichia coli* (purple) and *Enterococcus faecalis* (blue).

Supplementary Table S6: Birth months and delivery modes for infants from whose meconium clonal isolates of *Enterococcus faecalis* were cultured. Infants in each clonal group were born at similar times, providing a potential explanation for the isolation of identical *Enterococcus* strains.

Clonal Group	Sample ID	Birth month	Delivery mode
1	0037	Apr '18	Vaginal
	0141	May '18	Vaginal
2	0499	Aug '18	Vaginal
	0510	Sep '18	Vaginal
3	0187	Jun '18	Vaginal
	0223	Jun '18	C/S elective
	0366	Aug '18	C/S elective
	0349	Aug '18	Vaginal
4	0014	Apr '18	C/S emergency
	0030	Apr '18	C/S elective
	0081	Apr '18	C/S elective
	0265	Jul '18	C/S emergency