Gut dysbiosis with Bacilli dominance and accumulation of fermentation products precedes late-onset sepsis in preterm infants

Graspeuntner S*[1], Waschina S*[2], Künzel S[3], Twisselmann N[4], Rausch T K[4,5], Cloppenborg-Schmidt K[6], Zimmermann J[2], Viemann D[7], Herting E[4], Göpel W[4], Baines JF[3,5], Kaleta C[2], Rupp J[1,8], Härtel C[4], Pagel J[1,4,8]

*contributed equally

[1]Department of Infectious Diseases and Microbiology, University of Lübeck, Lübeck, Germany.

[2]Research Group Medical Systems Biology, Christian-Albrechts-University Kiel, Kiel, Germany.

[3]Max Planck Institute for Evolutionary Biology, Evolutionary Genomics, Plön, Germany.

[4]Department of Pediatrics, University of Lübeck, Lübeck, Germany.

[5]Institute for Medical Biometry and Statistics, University of Lübeck, Lübeck, Germany.

[6]Institute for Experimental Medicine, Christian Albrechts University of Kiel, Kiel, Germany.

[7]Department of Pediatric Pneumology, Allergy and Neonatology, Hannover Medical School, Hannover, Germany.

[8]German Center for Infection Research (DZIF), partner site Hamburg-Lübeck-Borstel-Riems, Lübeck, Germany.

Corresponding author: Julia Pagel, MD, Phone: +49 451 500 42801, Fax: +49 451 500 42804, e-mail: julia.pagel@uksh.de, Department of Pediatrics, University of Lübeck, Ratzeburger Allee 160, 23538 Lübeck, Germany

Running head: Gut dysbiosis in preterm neonatal sepsis

Key points: In a large cohort of preterm neonates, a Bacilli-mediated intestinal dysbiosis with a lack of anaerobic bacteria and an accumulation of fermentation products preceded the development of sepsis. This data offers potential biomarkers and therapeutic targets for prevention of sepsis.

Abstract

Background: Gut dysbiosis has been suggested as a major risk factor for the development of late-onset sepsis (LOS), a main cause of mortality and morbidity in preterm infants. We aimed to assess specific signatures of the gut microbiome including metabolic profiles in preterm infants <34 weeks of gestation preceding LOS.

Methods: In a single center cohort fecal samples of preterm infants were prospectively collected during the period of highest vulnerability for LOS (day 7, 14, 21 of life). Following 16S rRNA gene profiling, we assessed microbial community function using microbial metabolic network modeling. Data were adjusted for gestational age and use of probiotics.

Results: We studied stool samples of 71 preterm infants with LOS and 164 unaffected controls (no LOS/necrotizing enterocolitis). The bacteria isolated in diagnostic blood culture in most cases corresponded to the genera in the gut microbiome. LOS cases had a decelerated development of microbial diversity. Before onset of disease, LOS cases had specific gut microbiome signatures with higher abundance of Bacilli (specifically coagulase-negative Staphylococci, CoNS) and a lack of anaerobic bacteria. *In-silico* modeling of bacterial community metabolism suggested accumulation of the fermentation products ethanol and formic acid in LOS cases before the onset of disease.

Conclusions: Intestinal dysbiosis preceding LOS is characterized by an accumulation of Bacilli and their fermentation products and a paucity of anaerobic bacteria. Early microbiome and metabolic patterns may become a valuable biomarker to guide individualized prevention strategies of LOS in highly vulnerable populations.

Key words: gut; microbiome; preterm infant; dysbiosis; bacterial metabolism

Introduction

Preterm infants are at high risk for late-onset sepsis (LOS; occurrence after 72 hours of life), a leading cause for mortality and long-term morbidity including brain injury and chronic lung disease [1-4]. Hence, there is an urgent need to evaluate new diagnostic markers and individualized strategies for prevention of LOS.

The hypothesis that gut dysbiosis precedes the development of LOS is supported by several facts. First, previous reports noted alterations in intestinal microbiome composition and lower bacterial diversity prior to the onset of LOS in preterm infants [5-9]. These data are, however, inconsistent due to variable study designs and small cohort sizes. Second, the most vulnerable period for LOS is between days 7-21 of life [10]. The first weeks of life are associated with a high exposure rate to microbiome-disturbing influences such as antibiotic treatment or invasive measures [5, 11]. Third, host factors such as gestational age and immaturity of mucosal barriers contribute to the sepsis risk in mutual interaction with colonizing bacteria known to modulate immune responses, tight junction integrity, and metabolic function [12, 13]. Conventional culture studies have demonstrated concordance between blood-stream isolates and bacteria resident in the infants' gut [5, 11, 14-16]. Causative organisms are *Staphylococcus aureus* (*S. aureus*), coagulase-negative Staphylococci (CoNS), Group B streptococci (GBS), Enterococci, but also Gram-negative bacteria, i.e. Enterobacteriaceae [10, 17].

Recent advances in molecular microbiology and bioinformatics have enabled direct sequencing of bacterial DNA from stool in preterm infants to determine gut dysbiosis. Here we also investigate the yet unexplored function of the microbiota ecosystem and its impact on LOS risk. Functional community traits such as the release of metabolic by-products emerge from complex interactions between gut-inhabiting bacteria and with their nutritional environment [9, 18]. These data might improve the predictive value of microbiome patterns for LOS risk as they reflect important determinants of intestinal homeostasis [19, 20].

To confirm our hypothesis, we performed a prospective longitudinal study in a large cohort of preterm infants <34 weeks of gestational age at significant LOS risk during the neonatal period. Major readouts were microbiome signatures obtained from 16S rRNA sequencing and profiles of metabolic by-products released by intestinal bacteria using a systems biology approach.

Methods

Study cohort

Between January 2012 and January 2017 we collected fecal samples of preterm infants with a gestational age 23 0/7 to 33 6/7 weeks within our single center prospective study [Immunoregulation of Neonates] [22]. Fresh samples were collected on day 7±3, 14±3, and 21±3 of life and stored in a -80°C freezer. Infants without LOS or NEC were used as controls. Infants with lethal malformations or congenital anomalies of the gastrointestinal tract were excluded from our study cohort. Detailed information and definitions are outlined in the supplementary methods.

Ethics

Written informed consent was obtained from parents or legal representatives on behalf of the infants enrolled in our study. The study was approved by the local committee on research in human subjects at the University of Lübeck, Germany.

Bacterial DNA isolation

Fecal samples (approx. 200mg) were processed using the *PowerSoil* DNA *Isolation Kit* (MOBIO, Carlsbad, Canada). Details are outlined in the supplementary methods.

Polymerase Chain Reaction amplification and sequencing

We amplified V3/V4 16S gene sequences from each DNA sample. The primer design, polymerase chain reaction, quantification of amplicons, and library preparation were performed as described elsewhere [23]. Sequencing by synthesis runs were carried out on a MiSeq platform (Illumina®, San Diego, CA, USA) using the MiSeq Reagent Kit v3 (600 cycles). Negative extraction controls were included to verify reagents were uncontaminated.

Bioinformatics and Statistics

Bioinformatic processing, modeling of bacterial community metabolism, and statistical analysis are outlined in the supplementary methods.

Results

Clinical characteristics

We recruited a large cohort of preterm infants and prospectively collected fecal samples during the first weeks of life (n>600 infants/>2000 fecal samples). We only included infants with thorough longitudinal collection of fecal samples in the first weeks of life, i.e. 3 stool samples in the first 24 days of life, and additionally infants with LOS with at least 1 sample before diagnosis (n=71 infants with LOS between 72 h and 35 days after birth, n=164 control infants without LOS or NEC), in order to match the fecal sample of controls with the corresponding time-point of the pre-event sample of LOS cases. Hence a total of 607 fecal samples were analyzed, 574 had a positive sequencing result. 31 infants had a blood culture proven LOS with exclusively single pathogen detection [CoNS in 26 cases (*S. epidermidis* 12, *S. haemolyticus* 10, *S. hominis* 2, undefined species 2); *E. coli* 4, and *S. aureus* in 1 case; figure 1, supplemental table 1, 2].

The clinical characteristics of the study cohort are displayed in table 1a, specifically the lower gestational age in the LOS group as compared to the control group [median (IQR) 25.9 (24.8 - 28.5) vs. 30.7 (28.3 - 32.6) gestational weeks]. Infants with LOS and controls had no differences in exposure to antibiotics on day 1 of life (table 1b). LOS occurred at a median (IQR) day of life 12 (9-16), and total duration of antibiotic treatment was longer in LOS cases as compared to controls [median (IQR) 15 (7-26) vs. 5 (4-12) days].

Development of microbial diversity

The analysis of the bacterial DNA sequences from all 574 stools resulted in a median of 11,392 (IQR 6873–28,764) reads per specimen. We compared mean relative abundances of bacterial taxa for different time points of stool samples from all preterm infants. In detail, Bacilli, Actinobacteria,

Gammaproteobacteria, and Clostridia accounted for 97% of reads on day 7, 14 and 21 of life. The analysis revealed a decrease of Bacilli ($p<0.0001$) and an increase of Gammaproteobacteria ($p=0.0032$), Clostridia ($p=0.0003$), Alphaproteobacteria ($p=0.0001$), Betaproteobacteria ($p=0.0133$), and Fusobacteria ($p=0.0183$) within the first weeks of life (figure 2A). Thus, the microbiota structure substantially changed over time. At the species level, Bacilli were mostly identified Gram-positive cocci such as *S. epidermidis*, *S. haemolyticus*, *Enterococci* and *Streptococci*, Gammaproteobacteria were mainly *E. coli*, *Enterobacter hormaechei* (*E. hormaechei*), *Klebsiella oxytoca* (*K. oxytoca*), and other *Enterobacteriaceae,* Clostridia included *Veillonella* (it is debated whether *Veillonella* belongs to Clostridia taxa or rather to an own group), and Actinobacteria were dominated by *Bifidobacteria* (Figure 2B).

The overall branch length-weighted phylogenetic diversity in all infants increased over time (Linear mixed-effect model fit, $\chi^2(1)=11.857$, $p<0.0001$, figure 3).

Gut dysbiosis with Bacilli dominance precedes LOS

To characterize the intestinal bacterial composition preceding LOS, we compared stool samples from infants before onset of clinical or blood-culture proven LOS (pre-LOS, median interval 3 days from sample to event) with stools from unaffected infants at corresponding time points of sampling. Only the first episode of LOS was included. The pathogen identified by blood culture was concordant with the OTUs detected in the gut microbiota before diagnosis in 29 of 31 blood culture positive LOS-cases and in all cases of CoNS-sepsis (figure 1 and supplemental table 1, 2).

A maturational increase in phylogenetic diversity was only observed in infants who did not develop LOS ($p<0.0001$, figure 3). The diversity did not significantly change over time in infants with blood culture confirmed LOS ($p=0.2678$), or in infants with blood culture negative LOS ($p=0.4806$, figure 3).

The pre-event microbiome pattern of LOS patients was dominated by Bacilli (p=0.0330), which were largely Gram-positive bacteria such as *S. epidermidis*, other *Staphylococci*, and *Bacillales* (figure 4, supplemental figure 1). This was true for the group of infants with blood culture positive as well as blood culture negative LOS. A comparable high level of bacilli relative abundance could be observed on earlier sampling time points prior to onset of LOS. We further analyzed the taxa within the Bacilli class on species level and normalized the reads of each taxon to the total Bacilli reads for each sample. At the species level, the LOS cases (blood culture-positive and -negative group) were associated with higher abundances of *S. haemolyticus* (Mann-Whitney U-test, p=0.0200) in comparison with controls (supplemental figure 2).

Furthermore, we analyzed the colonization of the preterm infant gut over time in the first 24 days of life (figure 5). We incorporated only those control stool samples in our analysis, which have been collected within the same time span as the available pre-LOS samples. We also checked that there is no significant difference between the groups regarding the variance of day of life of the stool samples (Levene´s test: p=0.1074). LOS status significantly affected the colonization with Bacilli (p<0.0001). Control infants had a significantly reduced abundance of Bacilli over time (p=0.0019), whereas the infants with blood culture positive LOS showed an increase of their Bacilli colonization (p=0.0141) before onset of disease (figure 5). In contrast, LOS was associated with decreased colonization of anaerobic bacterial communities. Clostridia abundance increased over time in infants without LOS (p=0.0232), while there was no significant increase of Clostridia abundance over time in infants before onset of blood culture positive LOS (figure 5).

Profiles of metabolites released by bacteria preceding LOS

Based on 16S rRNA gene sequencing data from the gut microbiota communities we performed analyses of bacterial metabolism within the preterm infant gut bacterial communities using *in-silico*

metabolic models and assuming the nutrients from the human milk diet as resource for bacterial growth (supplementary table 3). Based on the computer simulations, formic acid, acetic acid, ethanol, DL-lactic acid, ammonium, L-alanine, propionic acid, and CO_2 were predicted to be the most prominent bacterial metabolites produced (figure 6, supplemental figure 3). The release of the fermentation products ethanol and formic acid were increased in infants with blood culture-confirmed LOS before disease onset compared to unaffected infants (figure 6). The simulations suggested further that the elevated production of ethanol and formic acid could be assigned to Bacilli ($p<0.0001$, figure 7), while the frequency of Actinobacteria and Gammaproteobacteria was inversely associated with ethanol production (figure 7).

When considering only those samples that are dominated by Bacilli (more than 75% of 16S reads) our model predicted still higher amounts of ethanol in blood culture positive LOS cases before the onset of sepsis compared to controls (Mann-Whitney U-Test, $p=0.029$, supplemental figure 4). The production of formic acid in Bacilli-dominated samples displayed a similar trend, although not significant ($p=0.067$). This could potentially be explained by the different taxonomic distribution of Bacilli between both groups described above (supplemental figure 2), because the different taxa might also express different metabolic phenotypes.

Discussion

In this prospective longitudinal study of preterm infants we demonstrated that the pattern of gut dysbiosis preceding LOS is characterized by a high abundance of Bacilli and accumulation of their fermentation products such as ethanol and formic acid. The strengths of our comprehensive study include the number of infants and specimens available for analysis and the novelty of evaluating the preterm gut for the pre-LOS habitat of the pathogens. The main limitation is lack of sequencing data on the genetic identity between blood culture isolates and pathogens in fecal samples.

The colonization of the preterm infant gut in the first month of life is a highly dynamic process. The gut microbial community generally matures along with a substantial increase in microbial diversity over many weeks at rates that are proportional to gestational age at birth (the communities in the more preterm infants mature at a slower rate). In our cohort, the gut was colonized with typical skin-associated bacteria in the first week of life, which were dominated by Bacilli, primarily CoNS. Within the second and third week of life proportions of Bacilli rapidly decreased and there was a switch from these rather aerobic and facultative anaerobic bacterial communities towards increasingly anaerobic bacterial communities including Gammaproteobacteria (i.e. largely *E. coli*, *Enterobacteriaceae* and *K. oxytoca*), Clostridia (including *Veillonella,* based on Greengenes database, although the taxonomy is still a matter of debate), Alphaproteobacteria, Betaproteobacteria, and Fusobacteria. Similar patterns of bacterial colonization have been shown by other groups in smaller cohorts of preterm infants [29-31].

For LOS patients, however, we noted different microbiome patterns. The first hallmark was a decelerated phylogenetic diversity of LOS patients over time, which is in line with previous studies [6, 7]. The second distinct phenomenon was over-representation of aerobic bacteria and a lack of anaerobic bacterial populations prior to the onset of disease. Hence, LOS-affected infants showed

significantly increased abundances of Bacilli, which were largely Gram-positive Staphylococci such as *S. epidermidis* or *S. haemolyticus*, and *Bacillales*. On species level, the LOS cases were associated with a higher ratio of *S. haemolyticus* in comparison with controls. These high proportions of Bacilli in the microbiome of LOS-affected infants did not decrease over time. A Bacilli-dominated microbiome after birth is known for infants delivered by Caesarean section. In our cohort, LOS cases and unaffected infants had a similar Caesarean section rate, but apparently unaffected infants display a gut microbiome with decreasing Bacilli abundance during the critical period of sepsis vulnerability. A critical microbiome-disrupting factor is exposure to antibiotics which was similar in LOS patients and controls immediately after birth but significantly different with regard to total days of therapy. The third hallmark of the pre-event microbiome of LOS infants is a diminished colonization with anaerobic bacterial communities such as Clostridia. These distinct aspects of gut dysbiosis are supported by previous studies, which showed that Bacilli such as *S. aureus* caused perturbations in the microbiota composition of the human adult colon, reducing the proportion of anaerobic species such as *Bifidobacterium* [32]. The lack of anaerobic bacteria, which have a role in prevention of bacterial translocation and stabilizing epithelial integrity, has been previously proposed to increase sepsis risk originating from the gut [33-36].

In our cohort, 31/71 LOS infants had a blood culture-proven sepsis. We identified a concordance of 93,5 % between the causative LOS organism and its identification in fecal samples prior to LOS, which is consistent with previous studies (from 64% to 95%) [11, 14-16]. In LOS cases caused by CoNS, concordance with the corresponding gut taxa was noted in all cases. Hence, our data support the concept that apart from catheter-associated blood-stream infections CoNS sepsis may evolve from gut translocation. Noteworthy, limitations in the interpretation of blood culture results need to be considered. In preterm infants, blood culture has a sensitivity of 10-20% due to low bacteremia, low sample volumes and previous antibiotic treatment [37].

By the means of a systems biology approach we evaluated whether the pre-event gut dysbiosis in LOS is also characterized by distinct patterns of microbiota metabolic processes. Modelling functional community traits such as the release of metabolic by-products revealed that LOS-affected infants showed an accumulation of the fermentation products ethanol and formic acid already prior to onset of sepsis. The increased release of formic acid and ethanol could be assigned to the abundance of Bacilli. The accumulation of those fermentation products is probably compounded by an immaturity of the intestinal lactase activity in preterm neonates. Animal studies suggested that excessive production and accumulation of bacterial fermentation products contribute to the pathogenesis of NEC [38]. Thus, fermentation products can directly cause damage to the intestinal mucosa and have indirect impact due to the induction of inflammation, which might lead to disruption of the mucosal barrier and translocation of luminal contents [38, 39]. Moreover, the distinct functional community traits preceding sepsis might impede further colonization with beneficial commensal bacteria, and thereby inhibit the diversification of the microbiota.

Thus, we propose that distinct patterns of gut dysbiosis, lead to intestinal barrier defects in preterm infants and thereby increase their risk for bacterial translocation. We suggest that a high Bacilli abundance may increase the potential for translocation through a damaged intestinal barrier. In addition, the high ratio of *S. haemolyticus* in the LOS cases might indicate that the colonization of virulent species increases the LOS risk [40].

Our novel findings on the structure of the gut microbiome in preterm infants encourage the concept that functional community traits can improve prediction models based on 16S rRNA sequencing data for LOS risk. The combination of intestinal dysbiosis and distinct metabolic interaction in a complex ecosystem may identify preterm infants at high sepsis risk before onset.

We made a comprehensive single-center attempt to identify microbiome signatures in stools from infants before LOS, however our approach has several limitations. The single-center study is of advantage for thorough sampling, rapid storage and standardized analysis of fecal samples, but our

data need to be confirmed in a multi-center study. Noteworthy, daily instead of weekly sampling would have been ideal to study the dynamics of microbiome establishment, however, this approach is less feasible in the NICU setting. Second, we did not sequence bloodstream isolates in order to compare the genotype with gut colonizing strains. Due to this limitation of our approach we cannot provide data on genetic identity of invasive isolates and pre-event colonizing species. In addition, we did not culture tips of removed central venous lines as this is not part of standardized clinical practice in our unit. Third, we noted an overwhelming role of Gram-positive cocci as causes of bloodstream infections. In the experience of the German Neonatal Network, Gram-positive cocci are the most dominating bacteria found in blood cultures of preterm infants with LOS (75-80%) [41]. To overcome the risk of contaminations, we only counted CoNS infection as "true", when infants additionally had clinical signs of infection and an elevated laboratory marker [42].

In conclusion, we demonstrated a distinct intestinal microbiome pattern of gut dysbiosis preceding LOS, which is dominated by Bacilli and a lack of anaerobic bacteria. The pre-event microbiome ecosystem of LOS infants is characterized by an accumulation of the fermentation products formic acid and ethanol. Future studies are needed to evaluate this distinct "sepsis signature" as a diagnostic biomarker and a prognostic tool for individualized strategies. Systems biology approaches can help to simulate specific supplementation strategies and its impact on the developing microbiome. Beside administration of specific probiotic compositions, selective decontamination with antibiotic treatment might be an option for high risk microbiota pattern. Mechanistic studies of sepsis in preterm neonates using cellular and animal models with relevant candidate bacterial drivers and attenuators of gut injury will have to complement this important translational approach.

Notes

Author Contributions. Experimental concept and design: JP, CH, JR, SG, SW. Performance/realization of experiments: JP, SG, KC, SK. Contribution of reagents/materials/analysis tools: JP, SG, SW, CK, JR, SK, JFB, NT, WG, EH, JZ. Data analysis: JP, CH, JR, SG, SW, TR, DV. Writing of the paper: JP, CH, SW.

Acknowledgements. We are indebted to the patients and their families for their participation and the physicians and nurses of the local NICU for their generous collaboration in this study. We would like to thank Constanze Reuter for the excellent technical support with the experiments.

Financial support. This work was supported by grants from the German Centre for Infection Research (DZIF; financed by the German Ministry of Education and Research, BMBF; partner-site Hamburg-Lübeck-Borstel-Riems) to JP, by the Collaborative Research Center “Origin and Function of Metaorganisms” (DFG support code: SFB1182) to CK, JZ, and SW, by grants from the Volkswagen Foundation [Az 90005] to DV, and by the BMBF (PRIMAL) to CH.

Potential conflicts of interest. These authors declare no conflicts of interest.

References

1. Shane AL, Sanchez PJ, Stoll BJ. Neonatal sepsis. Lancet **2017**; 390(10104): 1770-80.
2. Stichtenoth G, Demmert M, Bohnhorst B, et al. Major contributors to hospital mortality in very-low-birth-weight infants: data of the birth year 2010 cohort of the German Neonatal Network. Klinische Padiatrie **2012**; 224(4): 276-81.
3. van Vliet EO, de Kieviet JF, Oosterlaan J, van Elburg RM. Perinatal infections and neurodevelopmental outcome in very preterm and very low-birth-weight infants: a meta-analysis. JAMA pediatrics **2013**; 167(7): 662-8.
4. Bassler D, Stoll BJ, Schmidt B, et al. Using a count of neonatal morbidities to predict poor outcome in extremely low birth weight infants: added role of neonatal infection. Pediatrics **2009**; 123(1): 313-8.
5. Taft DH, Ambalavanan N, Schibler KR, et al. Center Variation in Intestinal Microbiota Prior to Late-Onset Sepsis in Preterm Infants. PloS one **2015**; 10(6): e0130604.
6. Madan JC, Salari RC, Saxena D, et al. Gut microbial colonisation in premature neonates predicts neonatal sepsis. Archives of disease in childhood Fetal and neonatal edition **2012**; 97(6): F456-62.
7. Shaw AG, Sim K, Randell P, et al. Late-Onset Bloodstream Infection and Perturbed Maturation of the Gastrointestinal Microbiota in Premature Infants. PLoS One **2015**; 10(7): e0132923.
8. Cernada M, Bauerl C, Serna E, Collado MC, Martinez GP, Vento M. Sepsis in preterm infants causes alterations in mucosal gene expression and microbiota profiles compared to non-septic twins. Sci Rep **2016**; 6: 25497.
9. Stewart CJ, Embleton ND, Marrs ECL, et al. Longitudinal development of the gut microbiome and metabolome in preterm neonates with late onset sepsis and healthy controls. Microbiome **2017**; 5(1): 75.

10. Boghossian NS, Page GP, Bell EF, et al. Late-onset sepsis in very low birth weight infants from singleton and multiple-gestation births. The Journal of pediatrics **2013**; 162(6): 1120-4, 4 e1.

11. Carl MA, Ndao IM, Springman AC, et al. Sepsis from the gut: the enteric habitat of bacteria that cause late-onset neonatal bloodstream infections. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America **2014**; 58(9): 1211-8.

12. Tarr PI, Warner BB. Gut bacteria and late-onset neonatal bloodstream infections in preterm infants. Seminars in fetal & neonatal medicine **2016**; 21(6): 388-93.

13. Warner BB, Deych E, Zhou Y, et al. Gut bacteria dysbiosis and necrotising enterocolitis in very low birthweight infants: a prospective case-control study. Lancet **2016**; 387(10031): 1928-36.

14. Graham PL, Della-Latta P, Wu F, Zhou J, Saiman L. The gastrointestinal tract serves as the reservoir for Gram-negative pathogens in very low birth weight infants. The Pediatric infectious disease journal **2007**; 26(12): 1153-6.

15. Smith A, Saiman L, Zhou J, Della-Latta P, Jia H, Graham PL. Concordance of Gastrointestinal Tract Colonization and Subsequent Bloodstream Infections With Gram-negative Bacilli in Very Low Birth Weight Infants in the Neonatal Intensive Care Unit. The Pediatric infectious disease journal **2010**; 29(9): 831-5.

16. Almuneef MA, Baltimore RS, Farrel PA, Reagan-Cirincione P, Dembry LM. Molecular typing demonstrating transmission of gram-negative rods in a neonatal intensive care unit in the absence of a recognized epidemic. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America **2001**; 32(2): 220-7.

17. Padula MA, Dewan ML, Shah SS, et al. Risk factors associated with laboratory-confirmed bloodstream infections in a tertiary neonatal intensive care unit. The Pediatric infectious disease journal **2014**; 33(10): 1027-32.

18. Flint HJ, Duncan SH, Scott KP, Louis P. Interactions and competition within the microbial community of the human colon: links between diet and health. Environmental microbiology **2007**; 9(5): 1101-11.

19. Cordero OX, Datta MS. Microbial interactions and community assembly at microscales. Current opinion in microbiology **2016**; 31: 227-34.

20. Waschina S, D'Souza G, Kost C, Kaleta C. Metabolic network architecture and carbon source determine metabolite production costs. The FEBS journal **2016**; 283(11): 2149-63.

21. Geffers C, Baerwolff S, Schwab F, GastmeierP. Incidence of healthcare-associated infections in high-risk neonates: results from the German surveillance system for very-low-birthweight infants. J Hosp Infect **2008**;68:214-21.

22. Pagel J, Hartz A, Figge J, et al. Regulatory T cell frequencies are increased in preterm infants with clinical early-onset sepsis. Clin Exp Immunol. **2016** Aug;185(2):219-27. doi: 10.1111/cei.12810.

23. Graspeuntner S, Bohlmann MK, Gillmann K, et al. Microbiota-based analysis reveals specific bacterial traits and a novel strategy for the diagnosis of infectious infertility. PLoS One **2018**; 13(1): e0191047.

24. Schloss PD, Westcott SL, Ryabin T, et al. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol **2009**; 75(23): 7537-41.

25. Magnusdottir S, Heinken A, Kutt L, et al. Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. Nature biotechnology **2017**; 35(1): 81-9.

26. Pruesse E, Peplies J, Glockner FO. SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. Bioinformatics **2012**; 28(14): 1823-9.

27. Munoz R, Yarza P, Ludwig W, et al. Release LTPs104 of the All-Species Living Tree. Systematic and applied microbiology **2011**; 34(3): 169-70.

28. Ludwig W, Strunk O, Westram R, et al. ARB: a software environment for sequence data. Nucleic Acids Res **2004**; 32(4): 1363-71.

29. Aujoulat F, Roudiere L, Picaud JC, et al. Temporal dynamics of the very premature infant gut dominant microbiota. BMC microbiology **2014**; 14: 325.

30. La Rosa PS, Warner BB, Zhou Y, et al. Patterned progression of bacterial populations in the premature infant gut. Proc Natl Acad Sci U S A **2014**; 111(34): 12522-7.

31. Korpela K, Blakstad EW, Moltu SJ, et al. Intestinal microbiota development and gestational age in preterm neonates. Sci Rep. **2018** Feb 6;8(1):2453. doi: 10.1038/s41598-018-20827-x.

32. Sannasiddappa TH, Costabile A, Gibson GR, Clarke SR. The influence of Staphylococcus aureus on gut microbial ecology in an in vitro continuous culture human colonic model system. PLoS One **2011**; 6(8): e23227.

33. Ewaschuk JB, Diaz H, Meddings L, et al. Secreted bioactive factors from Bifidobacterium infantis enhance epithelial cell barrier function. American journal of physiology Gastrointestinal and liver physiology **2008**; 295(5): G1025-34.

34. Duffy LC. Interactions mediating bacterial translocation in the immature intestine. The Journal of nutrition **2000**; 130(2S Suppl): 432S-6S.

35. Sonnenburg ED, Sonnenburg JL. Starving our microbial self: the deleterious consequences of a diet deficient in microbiota-accessible carbohydrates. Cell metabolism **2014**; 20(5): 779-86.

36. Smith PM, Howitt MR, Panikov N, et al. The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. Science **2013**; 341(6145): 569-73.

37. Yaacobi N, Bar-Meir M, Shchors I, Bromiker R. A prospective controlled trial of the optimal volume for neonatal blood cultures. The Pediatric infectious disease journal **2015**; 34(4): 351-4.

38. Lin J, Nafday SM, Chauvin SN, et al. Variable effects of short chain fatty acids and lactic acid in inducing intestinal mucosal injury in newborn rats. Journal of pediatric gastroenterology and nutrition **2002**; 35(4): 545-50.

39. Nafday SM, Chen W, Peng L, Babyatsky MW, Holzman IR, Lin J. Short-chain fatty acids induce colonic mucosal injury in rats with various postnatal ages. Pediatric research **2005**; 57(2): 201-4.

40. Soeorg H, Huik K, Parm U et al. Genetic relatedness of coagulase-negative Staphylococci from gastrointestinal tract and blood of preterm neonates with late-onset sepsis. Pediatr Infect Dis J. **2013** Apr;32(4):389-93. doi: 10.1097/INF.0b013e3182791abd.

41. Tröger B, Göpel W, Faust K et al. Risk for late-onset blood-culture proven sepsis in very-low-birth weight infants born small for gestational age: a large multicenter study from the German Neonatal Network. Pediatr Infect Dis J. **2014** Mar;33(3):238-43. doi: 10.1097/INF.0000000000000031.

42. Papadimitriou-Olivgeri I, Giormezis N, Papadimitriou-Olivgeris M et al. Number of positive blood cultures, biofilm formation, and adhesin genes in differentiating true coagulase-negative staphylococci bacteremia from contamination. Eur J Clin Microbiol Infect Dis. **2016** Jan;35(1):57-66.

Tables and Figures

1 a

	No LOS	LOS	p^a	Total
	n = 164	n = 71		n = 235
Gestational age [weeks]	30.7 (28.3 - 32.6)	25.9 (24.8 - 28.5)	<0.001	29.4 (26.6 – 32.0)
Birth weight [g]	1400.0 (1018.8 - 1688.8)	755.0 (630.0 - 1057.5)	<0.001	1220.0 (816.5 – 1490.0)
Male Gender	87 (53.05)	40 (56.34)	0.747	127 (54.04)
Multiple birth	60 (37.04)	22 (30.99)	0.459	82 (35.19)
Caesarean section	145 (89.51)	65 (91.55)	0.627	210 (90.13)
Probiotics	65 (39.63)	37 (52.11)	0.103	102 (43.4)
Apgar score, 5 min	8 (7 - 9)	7 (6 - 8)	<0.001	8 (7 - 8)
Apgar score, 10 min	9 (8 - 9)	8 (8 - 9)	<0.001	9 (8 - 9)
EOS	21 (12.8)	15 (21.13)	0.153	36 (15.32)
AIS	16 (9.82)	10 (14.08)	0.466	26 (11.11)
NEC	0 (0)	7 (9.86)	<0.001	7 (3.00)
FIP	1 (0.62)	11 (15.49)	<0.001	12 (5.15)
IVH	9 (5.56)	16 (22.54)	<0.001	25 (10.73)
SGA	18 (10.98)	14 (19.72)	0.112	32 (13.62)
LOS - day		12 (9 – 16)		

1 b

	No LOS	LOS	p^a	Total
	n = 164	n = 71		n = 235
Antenatal antibiotics	44 (26.83)	27 (38.03)	0.149	71 (30.21)
Exposure to antibiotics from day of life 1	144 (85.21)	60 (90.91)	0.344	204 (86.81)
- Duration (days)	5 (3 - 6)	5 (4 - 7)	<0.001	5 (3 - 7)
- Ampicillin	143 (99.31)	60 (100)		
- Gentamicin	142 (98.61)	60 (100)		
- Teicoplanin	1 (0.69)	2 (3.33)		
- Meropenem	2 (1.39)	1 (1.67)		
- Cefotaxime	6 (4.17)	3 (5.00)		
- Tobramycin	2 (1.39)	0 (0)		
Postnatal antibiotics Day of life 1-35	149 (90.85)	71 (100)	0.034	220 (93.62)
- Days	5 (4 - 12)	15 (7 - 26)	<0.001	8 (5 - 20.75)
- Ampicillin	144 (96.64)	66 (85.92)		
- Gentamicin	147 (98.66)	71 (100)		
- Teicoplanin	44 (29.53)	70 (98.59)		
- Meropenem	16 (10.74)	46 (64.79)		
- Vancomycin	1 (0.67)	1 (1.41)		
- Tazobactam	1 (0.67)	5 (7.04)		
- Cefuroxim	1 (0.67)	1 (1.41)		
- Erythromycin	3 (2.01)	5 (7.04)		
- Cefotaxime	7 (4.70)	11 (15.49)		
- Tobramycin	3 (2.01)	7 (9.86)		
- Linezolid	1 (0.67)	6 (8.45)		
- Other	2 (1.34)	5 (7.04)		

Table 1 a,b: Summary of patient demographics. Abbreviations: LOS, late-onset sepsis; EOS, early-onset sepsis; AIS, amnion infection syndrome; NEC, necrotizing enterocolitis (clinical NEC classified as Bell Stage II or Bell Stage III with the need for laparotomy with or without resection of necrotic gut, and the macroscopic diagnosis of NEC); FIP, focal intestinal perforation; IVH, intraventricular hemorrhage; SGA, small for gestational age. Data are described as median (IQR) or n (%). [a]Kruskal-Wallis rank sum test and χ^2 test or Fisher´s exact test for count data.

Figure Legends:

Figure 1: The dominance of opportunistic pathogens in fecal samples frequently explains the identity of bacterial species found in blood cultures. The figure shows the relative abundance (point size) of five bacterial species in fecal samples (rows), which were isolated in at least one case from blood cultures from the study cohort. Each column corresponds to a blood culture positive LOS case. The identity of the isolated pathogen is highlighted in red. The abundances are shown for the last fecal sample before the onset of sepsis in each respective case, if a pre-event stool sample was available. Otherwise, the sample closest to sepsis onset was chosen. The group CoNS (Coagulase-Negative Staphylococci) comprise the abundances of the species *Staphylococcus hominis*, *S. haemolyticus*, and *S. epidermidis*. In two blood cultures (first two columns), the exact CoNS species could not be further defined. *S. hominis* was not identified in our sequencing data from stool samples. The shown abundances correspond to sequences that were annotated with the genus *Staphylococcus* but unclassified species identity.

Figure 2: (A) Mean relative abundances of bacterial taxa on class level and species level for different time points in stool samples from preterm infants using non-parametric analysis of Covariances with gestational age and probiotic use as covariates. The comparison of relative abundances of bacterial taxa of different time points revealed a significant decrease of Bacilli and a significant increase of Gammaproteobacteria, Clostridia, Alphaproteobacteria, Betaproteobacteria, and Fusobacteria within the first month of life (day 7: n=204; day 14: n=202; day 21: n=164). (B) shows the most abundant taxa within each represented class.

Figure 3: Succession of bacterial phylogenetic diversity in the first 24 days of life of preterm infants without LOS (blue, n=164 infants), with clinical but blood culture negative LOS (red, n=32 infants),

and with culture-confirmed LOS (orange, n=31 infants). Intercepts and slopes of lines and standard error intervals (shaded areas) were calculated using linear regression models. Blood culture - confirmed LOS significantly affected the development of the intestinal bacterial diversity (linear mixed-effect model fit, p=0.007). While infants without LOS displayed an increase in bacterial diversity with time (linear mixed-effect model fit, p<0.0001), the diversity in infants who developed LOS did not significantly change over time in both, blood culture positive (p=0.2678) and blood culture negative LOS (p=0.4806).

Figure 4: Mean relative abundances of bacterial taxa on class level for according to LOS status before the onset of sepsis using non-parametric analysis of Covariances with gestational age and probiotic use as covariates. Infants developing LOS and non-LOS infants differed significantly in their abundance of Bacilli prior to the onset of disease (blood culture positive LOS: n=20; blood culture negative LOS: n=24; No LOS: n=164).

Figure 5: Relative abundances of Bacilli, Gammaproteobacteria, Clostridia, and Bacteroidia in the first 24 days of life in infants before LOS onset (blood culture negative=red: n=24; blood culture positive=orange: n=20) and infants without LOS (blue: n=162) using linear mixed effect models whilst controlling for probiotic administration and gestational age. Blood culture proven LOS significantly affected the development of Bacilli colonization (p-value<0.0001). While Bacilli abundance increased over time in blood culture proven LOS (linear mixed-effect model fit: p=0.0141), it decreased in control infants (linear mixed-effect model fit: p=0.0019). In addition, the relative abundance of Clostridia increased over time in control infants (linear mixed-effect model fit: p=0.0232) while this increase was absent in infants developing blood culture proven LOS.

Figure 6*: In-silico* estimates for the production of individual metabolites by bacteria communities depending on LOS. The simulations suggested an increased production of the fermentation products ethanol and formic acid in infants who later developed LOS (red: blood culture negative LOS, n=24; blue: blood culture-confirmed LOS, n=20) in comparison with non-affected infants (n=164). One-way non-parametric analysis of covariances was used to calculate p-values whilst controlling for probiotic administration and gestational age. Shown are only metabolites with a median production rate above 0.1 mmol per gDW (gram bacterial dry weight) in at least one of the three groups.

Figure 7: Correlations between predicted rates of formic acid and ethanol production by the bacterial community and the relative abundance of phylogenetic bacterial classes. Pearson Product-Moment Correlation tests were used to obtain p-values and arrows denote the sign of the correlation coefficient in case of a p-value less than 0.05.

Figure 1

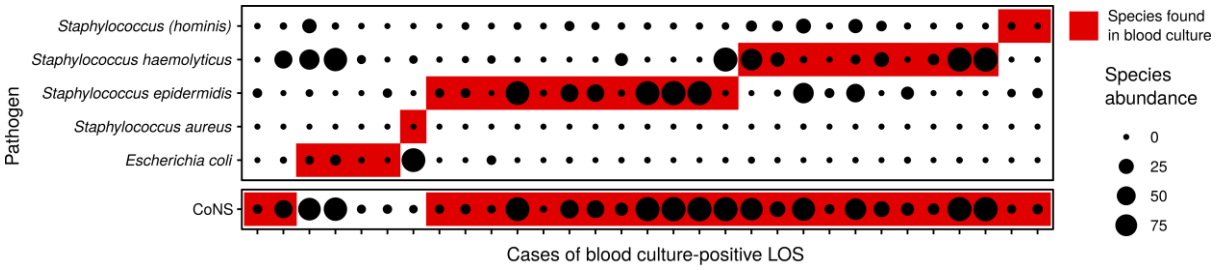

Figure 2

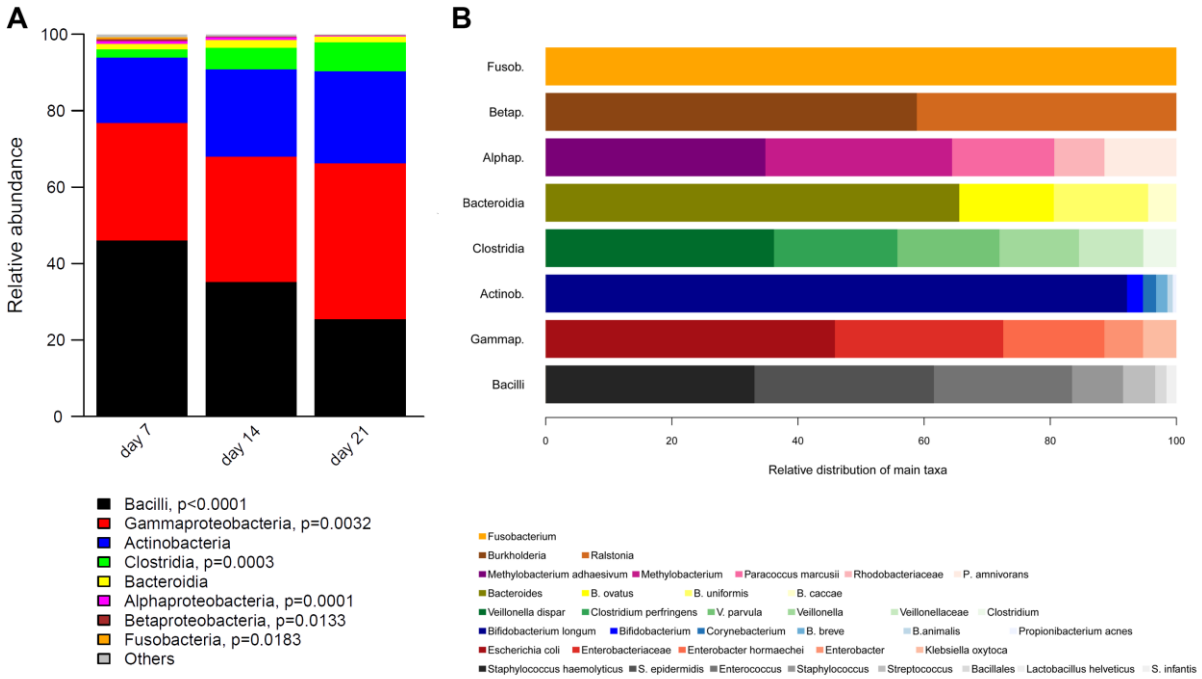

Figure 3

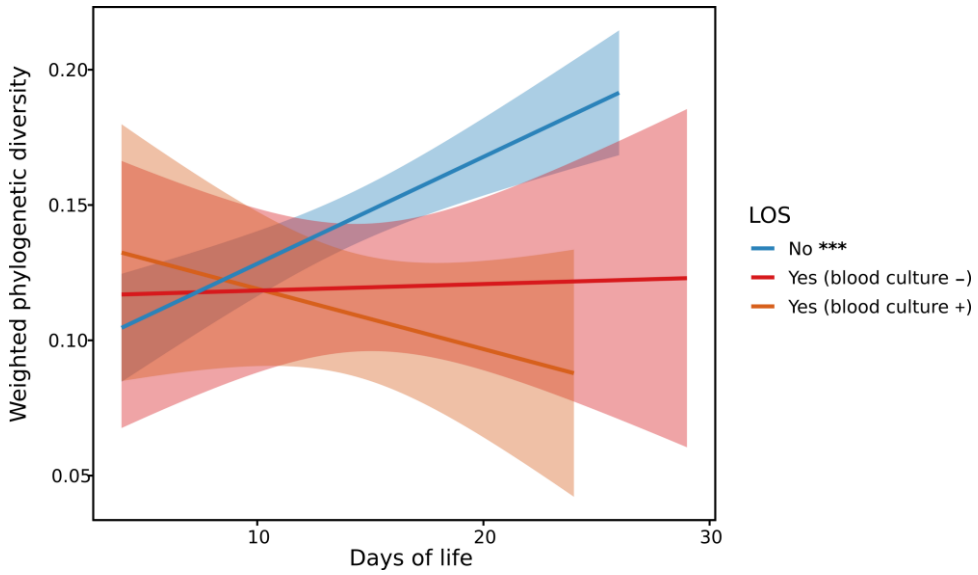

Figure 4

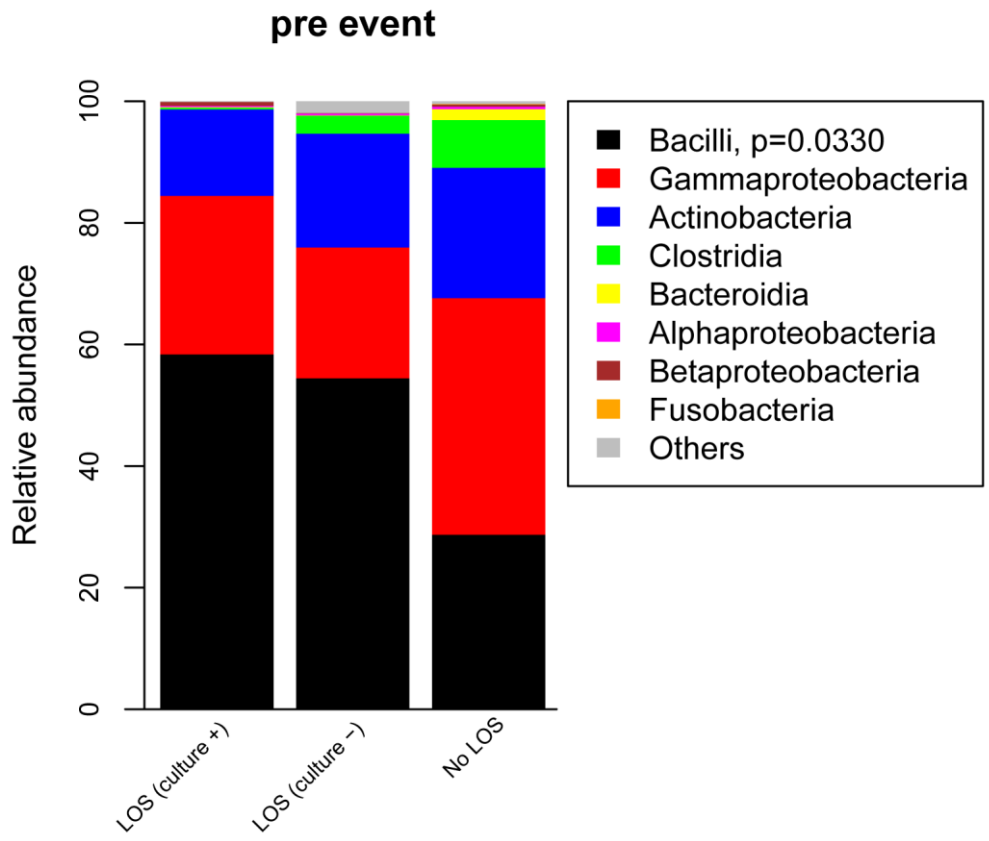

Figure 5

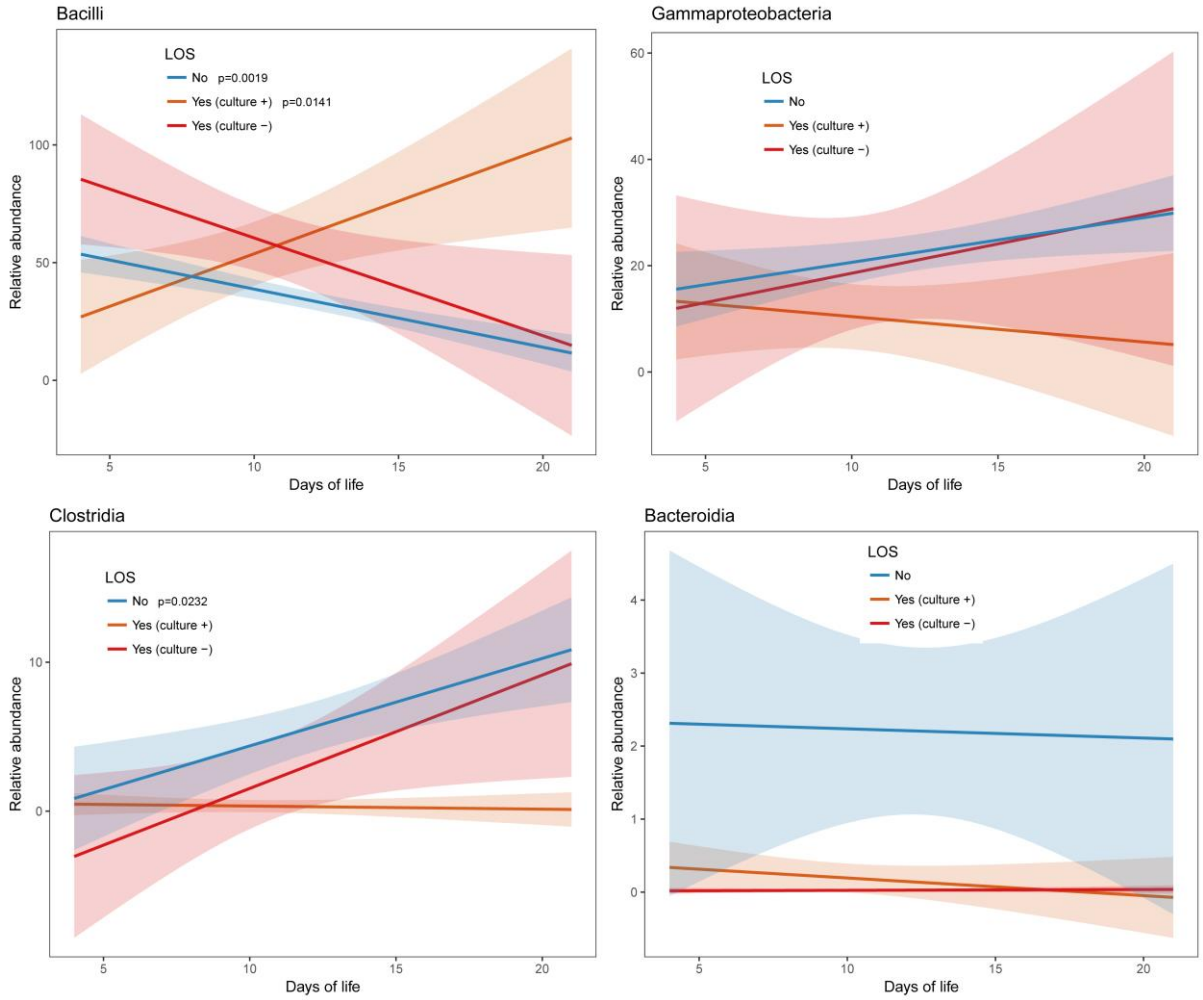

Figure 6

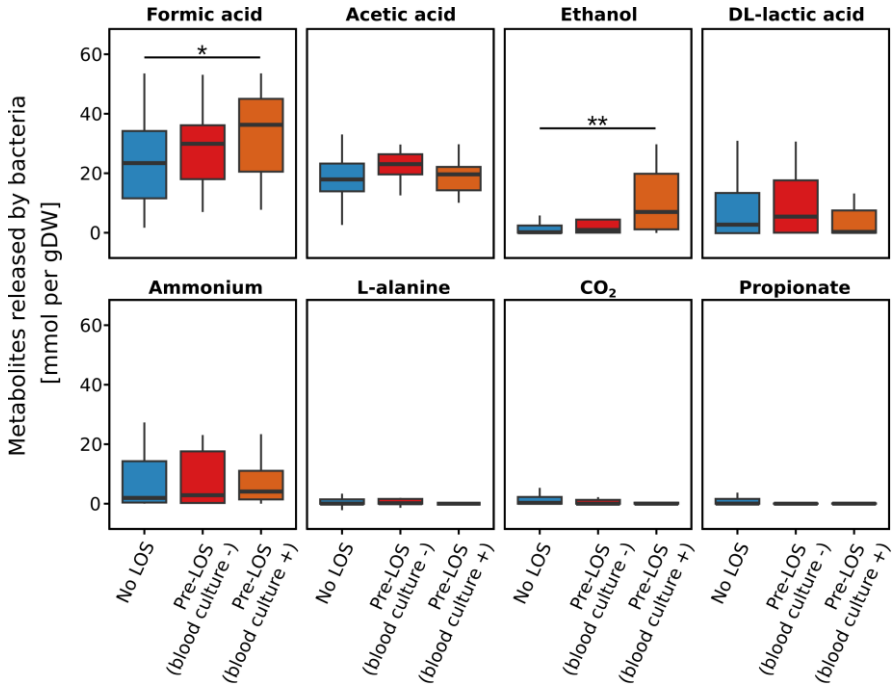

Figure 7