Journal of Translational Medicine

RESEARCH Open Access

The gut microbiota is associated with the small intestinal paracellular permeability and the development of the immune system in healthy children during the first two years of life

Mariusz Kaczmarczyk[1], Ulrike Löber[2,3,4,5], Karolina Adamek[6], Dagmara Węgrzyn[6], Karolina Skonieczna-Żydecka[7], Damian Malinowski[8], Igor Łoniewski[7,9*], Lajos Markó[2,3,4,5,10], Thomas Ulas[11,12], Sofia K. Forslund[2,3,4,5,10,11,13] and Beata Łoniewska[6]

Abstract

Background: The intestinal barrier plays an important role in the defense against infections, and nutritional, endocrine, and immune functions. The gut microbiota playing an important role in development of the gastrointestinal tract can impact intestinal permeability and immunity during early life, but data concerning this problem are scarce.

Methods: We analyzed the microbiota in fecal samples (101 samples in total) collected longitudinally over 24 months from 21 newborns to investigate whether the markers of small intestinal paracellular permeability (zonulin) and immune system development (calprotectin) are linked to the gut microbiota. The results were validated using data from an independent cohort that included the calprotectin and gut microbiota in children during the first year of life.

Results: Zonulin levels tended to increase for up to 6 months after childbirth and stabilize thereafter remaining at a high level while calprotectin concentration was high after childbirth and began to decline from 6 months of life. The gut microbiota composition and the related metabolic potentials changed during the first 2 years of life and were correlated with zonulin and calprotectin levels. Faecal calprotectin correlated inversely with alpha diversity (Shannon index, $r = -0.30$, FDR P (Q) $= 0.039$). It also correlated with seven taxa; i.a. negatively with Ruminococcaceae ($r = -0.34$, $Q = 0.046$), and Clostridiales ($r = -0.34$, $Q = 0.048$) and positively with *Staphylococcus* ($r = 0.38$, $Q = 0.023$) and Staphylococcaceae ($r = 0.35$, $Q = 0.04$), whereas zonulin correlated with 19 taxa; i.a. with Bacillales ($r = -0.52$, $Q = 0.0004$), Clostridiales ($r = 0.48$, $Q = 0.001$) and the *Ruminococcus* (*torques* group) ($r = 0.40$, $Q = 0.026$). When time intervals were considered only changes in abundance of the *Ruminococcus* (*torques* group) were associated with changes in calprotectin ($\beta = 2.94$, $SE = 0.8$, $Q = 0.015$). The dynamics of stool calprotectin was negatively associated with changes in two MetaCyc pathways: pyruvate fermentation to butanoate ($\beta = -4.54$, $SE = 1.08$, $Q = 0.028$) and *Clostridium acetobutylicum* fermentation ($\beta = -4.48$, $SE = 1.16$, $Q = 0.026$).

*Correspondence: sanprobi@sanprobi.pl
[9] Department of Human Nutrition and Metabolomics, Broniewskiego 24, 71-460 Szczecin, Poland
Full list of author information is available at the end of the article

Conclusions: The small intestinal paracellular permeability, immune system-related markers and gut microbiota change dynamically during the first 2 years of life. The *Ruminococcus* (*torques* group) seems to be especially involved in controlling paracellular permeability. *Staphylococcus,* Staphylococcaceae, Ruminococcaceae, and Clostridiales, may be potential biomarkers of the immune system. Despite observed correlations their clear causation and health consequences were not proven. Mechanistic studies are required.

Keywords: Zonulin, Calprotectin, Gut microbiota, Gut permeability, Newborn

Background

The intestinal barrier plays important role in the defense against infections, apart from its essential nutritional, endocrine, and immune functions [1]. Multiple factors impact intestinal permeability in infants, including gestational age, mode of delivery and feeding, and various diseases [2]. Increased gut barrier permeability allows the optimal nutrient uptake and leads to increased immune tolerance [3]; on the other hand, increased permeability to foreign antigens, including intestinal bacteria, might result in inflammation and systemic hypersensitivity [3, 4]. The gut barrier of a newborn is highly permeable; importantly, the permeability decreases during a process known as "gut closure" [5]. The exact time of this process in humans, regulated by growth factors, hormones and breast milk is unknown but has been proposed to take place around 22nd week of life [6]. It could also be hypothesized that gut microbiota may be involved in this process [6, 7]. Of note, gut permeability can be assessed via the analysis of the absorption of various substances [8], as well as via the measurement of blood and stool markers, including zonulin and calprotectin [9].

Zonulin (ZON), an analog of the cholera comma toxin (ZOT, zonula occludens toxin) [10, 11] shown to regulate paracellular transport within the small bowel [12–14]. This protein is synthesized in the liver and epithelial cells of the intestine and is a component of multi-protein membrane complexes (claudin-occludin-guanylate kinase-like proteins zonula occludens (ZO)-1, ZO-2, and ZO-3) forming tight junctions (TJ) [15]. In fact, zonulin regulates the tightness of TJ, which is an important part of the proper function of the intestinal barrier [16]. Increased zonulin concentrations correlate positively with small intestinal permeability [17], a phenomenon discovered in the context of inflammatory and autoimmune disease [18–21], and could be considered as a biomarker for systemic inflammation [22].

Human calprotectin is a 24 kDa dimer formed by the protein monomers S100A8 (10,835 Da) and S100A9 (13,242 Da), and makes up to 60% of the soluble proteins in the cytosol of human neutrophils [23, 24]. The sources of calprotectin in newborns are breast milk and resident and non-resident myeloid cells [25]. In fact, fecal calprotectin is used in older children and adults as a marker for inflammatory bowel diseases (IBD); there is a body of evidence correlating calprotectin with IBD, cow-milk allergy, atopic disease and other gastrointestinal disorders [26]. Of note, calprotectin has the potential to be used as an indirect marker of gut permeability [27]. However, its role as a marker of inflammation was not confirmed in neonates [25].

In our previous study [28] we found that some maternal–fetal factors are associated with increased concentration of fecal calprotectin and zonulin in children during the first 2 years of life. Additionally, we found that after birth, zonulin levels increased up until 12 months of age, remaining high thereafter, while calprotectin levels decreased until six months of age, stabilizing thereafter [29]. Of note, in this previous study, we did not assess the fecal microbiota, which was a significant limitation; we were not able to trace the relationships between the fecal microbiome and the concentrations of the abovementioned stool markers. Therefore, in this study, we decided to analyze the microbiota in faecal samples collected longitudinally over 24 months from 21 subjects to test the hypothesis that small intestinal barrier permeability is linked to the gut microbiota. Moreover, we validated the observed results via comparison with the data obtained in the study by Willers et al. (Hannover Medical School—HMS cohort) who measured the relationship between calprotectin and gut microbiota in children during the first year of life [25].

Methods

Subjects

The present observational prospective cohort study continues research efforts in the context of intestinal barrier function in a cohort of Polish newborns, described in detail in previous publications [28, 29]. Twenty four healthy full-term newborns at the Department of Obstetrics, Gynecology and Neonatology, the Pomeranian Medical University/Independent Public Clinical Hospital No. 2 in Szczecin (PMU cohort) were initially recruited to this study. Longitudinal sampling was performed for over 24 months. Three newborns were excluded due to an inadequate number of samples. In total, 21 newborns (101 samples in total) were included in the study, with at least four longitudinal stool samples available (the sample

availability matrix is provided in Additional file 1: Figure S1). Samples were taken at the following points in time—meconium (P1, n = 8), 7 days (P2, n = 20), 1 month (P3, n = 19), 6 months (P4, n = 20), 12 months (P5, n = 21) and 24 months (P6, n = 13) (Additional file 1: Figure S2) after birth. The study was conducted using the results obtained for the time points P2-P6. The measured markers and the meconium microbiota composition (P1) reflect the perinatal period. Therefore the values obtained at this time point were used instead to validate the microbiota analysis (the meconium microbiota is expected to be clearly distinguishable from the microbiota obtained at other time points). The participants' characteristics are shown in Table 1.

All included newborns were exclusively breastfed during the first week of life. The children were healthy and did not take antibiotics at the time the samples were collected. For validation, we used independent Illumina MiSeq 16S rRNA data of the same hypervariable region (V3–V4) derived from 227 stool samples of healthy term children enrolled in the HMS cohort. They were tested at seven time points (1, 2, 10, 30, 90, 180 and 360 days of age). For the sake of compatibility with the PMU cohort, five time points were selected: 1 day (P1, n = 43), 10 days (P2, n = 42), 30 days (P3, n = 47), 180 days (P4, n = 43) and 360 days (P5, n = 52) of age.

Determination of the fecal zonulin and calprotectin content

The concentrations of zonulin and calprotectin in the fecal samples were measured as previously described [28, 29].

Sample collection, DNA extraction and sequencing

Stool samples were collected from diapers using a standardized biological material collection kit (Stool Sample Application System (SAS); Immundiagnostik, Bensheim, Germany). The samples were collected by previously trained hospital staff or by the parents according to an established procedure and stored in a refrigerator (for a maximum of eight hours) before transport. Transport to the laboratory took no longer than 60 min, at 6–8 °C. The stool samples were then frozen at −20 °C until metagenomic analyses were conducted. Microbiome DNA extraction was performed using the Genomic Mini AX Bacteria + Spin and Genomic Mini AX Soil Spin kits (A&A Biotechnology, Gdynia, Poland) following the manufacturer's protocol. DNA concentrations were determined by fluorometry (DeNovix DS-11 FX + Spectrophotometer/Fluorometer, Wilmington, DE, USA). Samples were subsequently stored at −20 °C. Metagenomic libraries of the V3–V4 hypervariable region of the 16S rRNA gene were constructed and further sequenced on the MiSeq platform (paired-end 2 × 250 bp) using V2 chemistry from Illumina (Illumina, San Diego, CA, USA). Next generation sequencing (NGS) was performed by Genomed S.A., Warsaw, Poland.

For validation, we used the independent Illumina MiSeq 16S rRNA data of the same hypervariable region (V3–V4) derived from the 227 stool samples of healthy term children enrolled in the HMS cohort.

16S sequencing sample processing

The sequences were processed using LotuS 1.62 [30]. LotuS clusters operational taxonomic units and generates taxonomic-level abundances tables. UPARSE de novo sequence clustering removed chimeric OTUs with 1470 OTUs (PMU cohort) and 923 OTUs (HMS cohort) remaining. OTU seed sequences were classified by BLAST lower ancestor comparison to SILVA (1.32). The Rarefaction Toolkit (RTK) was used to normalize the abundances on all taxonomic levels [31]. Two samples were removed due to their low number of bacterial reads from the PMU cohort (read counts of 246 and 134), resulting in a final rarefaction depth of 1117 16S reads, yielding a count of 594 genera resolved at this sequencing depth and the final sample size of 101 was obtained. Moreover, 108 samples were removed due to their low number of bacterial reads from the HMS cohort (read count less than 1000), resulting in a rarefaction depth of 1000 16S reads, yielding 484 genera and the final sample size of 227.

Table 1 Characteristics of study participants

Characteristic	P1 (n = 8)	P2 (n = 20)	P3 (n = 19)	P4 (n = 20)	P5 (n = 21)	P6 (n = 13)
Sex (F/M)	4/4	9/11	10/9	9/11	10/11	8/5
Mode of delivery (C/V)	5/3	11/9	11/8	12/8	12/9	6/7
Birth weight (g)	3472 ± 507	3320 ± 548	3311 ± 518	3322 ± 550	3341 ± 543	3352 ± 582
Body weight (kg)	–	3.36 ± 0.54	4.49 ± 0.74	8.26 ± 1.18	10.40 ± 1.29	13.00 ± 1.68
Feeding method (N/F)	8/0	20/0	15/4	9/11	4/17	0/13
Breastfeeding time (weeks)	36.2 ± 25.6	30.0 ± 26.7	27.1 ± 24.9	27.9 ± 26.9	28.9 ± 26.6	27.1 ± 26.3

C, Cesarean section, V, vaginal delivery, N, natural, F, formula; P1, meconium, P2, 7th day, P3, 1st month, P4, 6th month, P5, 12th month, P6, 24th month

The alpha diversity (Shannon, Simpson, Chao1, Evenness) was computed using RTK taking the genus median alpha diversity over ten rarefaction cycles. The beta diversity was calculated on a rarefied genus-level abundance table using the Bray–Curtis dissimilarity metrics and visualized as principal coordinate analysis (PCoA). Taxa that were present in less than 20% (PMU cohort) and 10% (HMS cohort; a lower taxa prevalence was used to allow the comparison) of samples were removed.

To predict the metabolic profile of gut microbiota, we used the PICRUSt2 tool [32]. The prediction of MetaCyc metabolic pathways was conducted on the non-rarefied OTU abundance table from which rare OTUs (present in less than 20% of samples with a minimum count of 3) were removed. In order to address the issue of compositional data, the taxa and predicted MetaCyc pathway abundances (not rarefied but filtered to match the rarefied data) were transformed by generating 128 Monte Carlo instances of the Dirichlet distribution for each gut sample, followed by center-log transformation of each instance [33]. Analyses were performed for each instance separately and the results were then averaged over 128 instances.

Power calculation

With N = 21 time series, it is true our statistical power may be limited. While this would not decrease confidence in findings where we conclude significance (type I errors) it may place us at risk for missing associations, and challenges any interpretation of negative findings. As both permeability markers and microbiome features shift over time as the gut matures and diet shifts, their association may be driven by shared dependency on this third factor. First, we consider, as the closest representative test for which power calculations are straightforward, the analysis of coupled changes in e.g. a permeability marker versus microbial taxa through (Pearson) correlation with N = 21 samples each paired across two time points, resulting in N = 21 time point differences. We consider multiple testing considerations here to be reflected as per Bonferroni correction for Nt = 50 gut genera, making for an adjusted alpha = 0.1/50 = 0.002. Under convention of considering r = 0.5 as a "large" correlation, we achieve (running as "pwr.r.test (alternative = "two.sided", n = 21, sig.level = 0.002, r = 0.5)") a statistical power of only ~ 21%, meaning it is likely that we are overly conservative due to limited sample size, and that there may be several more of these associations at least this strong which remains to be shown robust in a larger study. However, this does not detract from our confidence in the associations where we do conclude significance, as it rather is more likely we miss associations between permeability and microbiota. This has no real bearing on whether the associations are direct or indirect via time progressing both as a third factor, however. Here, support of a direct link would come from the association achieving significance not only for the correlation of differences within time courses between time points, but between donors within time points. This corresponding power calculation here is essentially the same (Nt = 21, alpha = 0.002, two-sided, taking r = 0.5 as a "large" correlation) meaning our power is also ~ 21% at concluding an effect of similar size is a direct association. However, as we are similarly underpowered to conclude significance there, we expect to underestimate rather than overestimate inference of direct versus indirect relationships. All in all thereforewe cannot be fully certain either way whether concluded associations are direct or indirect, which concern must be kept in mind in comparison of our findings to previous literature.

Statistical analysis

All statistical analysis was performed using R (version 4.0.0, R Foundation for Statistical Computing, Vienna, Austria). Changes over time in the context of permeability biomarkers (zonulin and calprotectin), the gut alpha diversity indices, gut composition principal coordinate scores, or gut taxa and pathway abundances were analyzed using mixed-effects models. Time points (P) were treated as the fixed effect and the newborn IDs as the random effect, accounting for repeated sampling of the same individual. The significance of the time point variable was tested via comparison of the likelihoods of the two nested models (likelihood ratio test, LRT): without (null model) and with (full model) the fixed factor in question. Pairwise comparisons between timepoints were performed using the emmeans package [34] and the Benjamini–Hochberg procedure for the control of the false discovery rate (FDR). In the models with biomarkers (which were rank-transformed), taxa and metabolic pathways, the mode of delivery and breastfeeding time were also included as additional covariates and thus controlled for. The LRT p-values for taxa and metabolic pathways were FDR-adjusted as above.

The gut microbial diversity, composition and metabolic pathways were analyzed with respect to zonulin and calprotectin biomarkers using a repeated measures correlation method [35]. The repeated-measures correlation (an equivalent to the linear mixed-effects model with random intercept) was used to assess the association between the levels of the biomarkers and the alpha diversity, PCoA scores, taxa and metabolic pathway abundance, accounting for repeated measures for each participant.

In addition, an association between the differences (changes) among all time points were associated using

linear mixed-effect models with random intercepts (newborn ID as a random factor). The extent of the biomarker change was defined as the dependent variable (DV). Then, the time period/interval—a categorical variable with ten levels indicating all possible differences (e.g. P2–P3, P2–P4, etc.) between the five time points (P2, P3, P4, P5 and P6, PMU cohort) or four time points (P2, P3, P4, P5, HMS cohort) and either beta-diversity (Bray–Curtis intersample composition distance) or univariate metrics of changes in microbiota composition or metabolic pathway abundance (between corresponding time points) were defined as fixed effects. To investigate the connection between the change in the overall compositional profile and the change in the levels of biomarkers, we took the within-subject Bray–Curtis distances (computed at the genus level using rarefied abundances). Fine-grained analysis was performed using changes in the taxonomic composition (from genus to phylum level) and metabolic pathways which were computed using the center-log transformed Monte Carlo instances of the Dirichlet distribution. Two sets of likelihood ratio tests were performed (with mode of delivery and breastfeeding time as covariates). In the first LRT, a significance of the Bray–Curtis distance or change in abundance was assessed under the assumption of no interaction between the fixed factors (the common slope model). The second LRT was used to determine whether the interaction was statistically significant. In case of a significant interaction effect, the emmeans package was used (emtrends function) to estimate the individual slopes over levels of the time period/interval. The p-values for the slopes (coefficients) were determined and adjusted for multiple testing to control for the FDR. No imputation was undertaken for missing data so as not to make unwarranted assumptions on the distribution of measured features. Instead these cases were consistently omitted under the settings used for the R functions employed in our analysis. Data was manipulated using Perl. A significance level of 5% was used for raw and FDR adjusted p-values.

Results

Stool levels of zonulin and calprotectin increase and decrease over time, respectively

Both stool zonulin and calprotectin concentrations changed significantly over time in the cohort (Fig. 1, LRT, df = 4, P = 2.6e−05, and df = 4, P = 10.0e−05, respectively; adjusted for the mode of delivery and breastfeeding duration).

Zonulin levels at P2 (7 days after birth) and P3 (1 month after birth) were significantly lower than those at the time points P4, P5, and P6. However, from P4 to P6 (6th and 24th months of life), no significant differences were observed (between any two time points, non-consecutive or consecutive). Thus, zonulin levels tended to increase for up to 6 months after birth (P4) and stabilize thereafter (Fig. 1a). In contrast to zonulin, calprotectin levels did not differ significantly between any consecutive time points (Fig. 1b). Moreover, up to and including P4, none of the differences in calprotectin levels between any two time points (non-consecutive or consecutive) were significant. However, the calprotectin levels at P5 were significantly lower than those at P2, P3; additionally, the calprotectin levels at P6 were also significantly lower than those at P2, P3, and P4 (Fig. 1b). Overall, these results suggest that between about 6 and 12 months of age, the zonulin levels stabilize remaining at a high level, while calprotectin concentration begins to decline. Importantly, the analysis of the HMS validation cohort (Additional file 1: Figure S3) confirmed the declining trend of calprotectin from P3 onwards; of note, a significant difference was determined between the first and sixth months of life, which altogether strongly highlights the sixth month of life as an important time point concerning changes of analyzed markers after childbirth. Interestingly, despite the observed opposite trends in the context of stool zonulin and calprotectin levels, the two markers (its changes) did not appear to correlate with each other (LRT; common slope model: df = 1, P = 0.776, β = 0.03, SE = 0.09; interaction model: df = 9, P = 0.419, Fig. 1c).

Gut microbiota diversity, taxonomy and metabolic pathways during the first two years of life

Gut microbiota diversity

The alpha diversity was measured using the Shannon index (Fig. 2a); sample diversity differed significantly between time points (LRT, df = 5, P = 6.61e−09). Specifically, P2′s alpha diversity was significantly lower than in P4, P5 and P6, whereas P3′s alpha diversity was significantly lower than in P5 and P6. We have validated those results by comparing diversity in stool samples with that in meconium in the two cohorts of children (PMU and HMS). In the PMU, the Shannon indexes in P2 (Q < 0.0001), P3 (Q = 0.0001), P4 (Q = 0.004) were significantly lower than that in P1; on the other hand, the Shannon indexes in P5 (Q = 0.109) and P6 (Q = 0.201) did not differ significantly from that. Altogether, these results indicate that the meconium samples have the highest alpha diversity; interestingly, our data suggest that the newborn's gut microbiota diversity first decreases but then is restored over time (Fig. 2a). Importantly, a similar pattern was evident in the HMS cohort data (Additional file 1: Fig. 4a), except for a clear discrepancy between the first samples' diversity (in the HMS cohort, the lowest diversity was detected at day 1, giving rise to a monotonically increasing trend in this cohort). This can occur if "high diversity" meconium samples (bacteria derived from the mother and environment following birth [36]

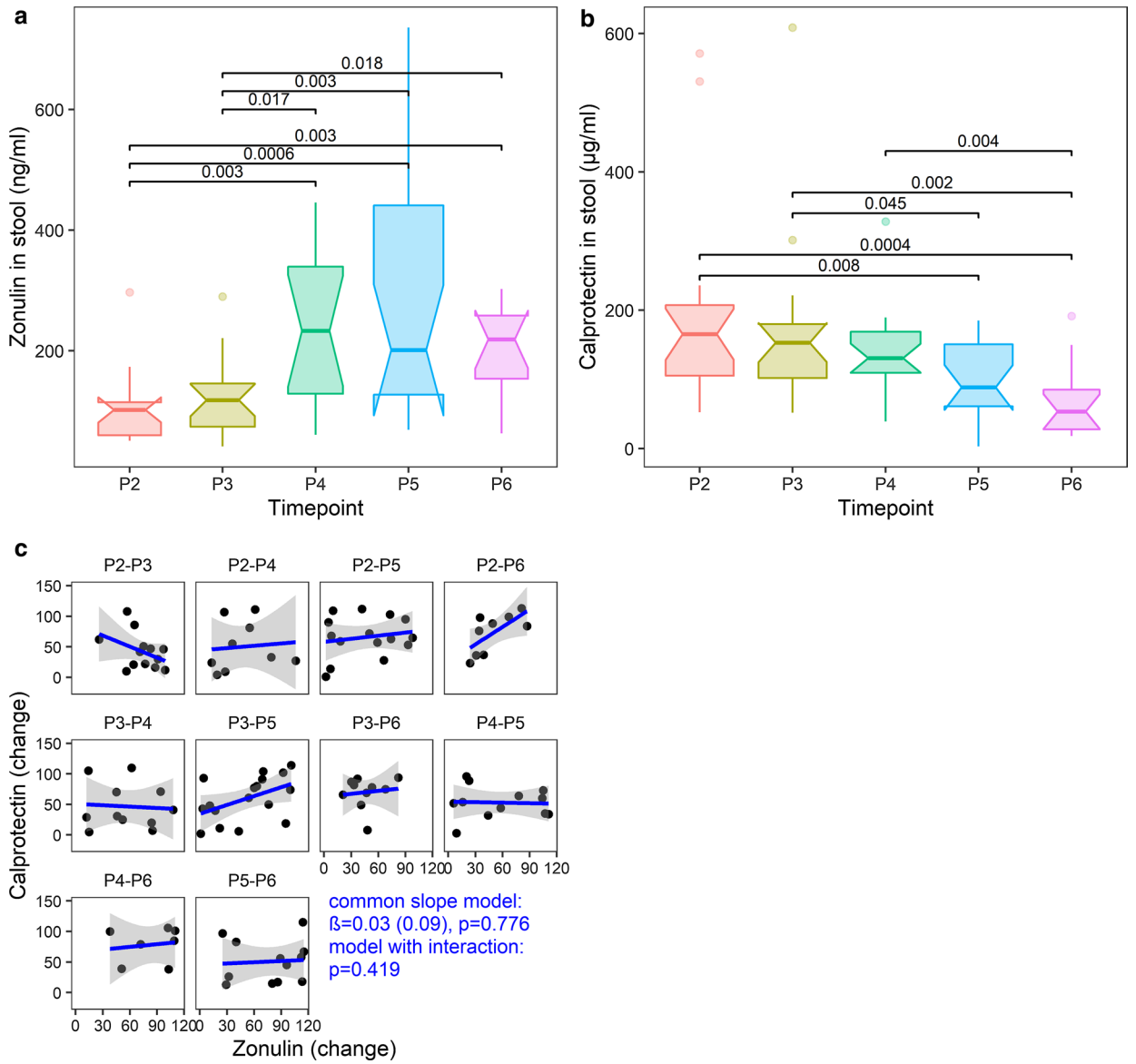

Fig. 1 Dynamics of stool zonulin and calprotectin levels. **a** zonulin, **b** calprotectin. Notched boxplots with variable widths proportional to the square-roots of the number of observations in the groups are shown; FDR adjusted p-values < 0.05 are represented. **c** correlation between the stool zonulin and calprotectin changes (rank-transformed, adjusted for the mode of delivery and breastfeeding time); P2-7th day, P3-1st month, P4-6th month, P5-12th month, P6-24th month; PX-PY-difference between the time points X and Y

are combined with early newborns' stool "low diversity" samples (in the PMU cohort, only meconium stool samples were analyzed while in the HMS cohort also the first stool samples were analyzed in some cases).

The ordination of samples in a two-dimensional space performed as per a principal coordinate (PCo) analysis on Bray–Curtis distances dissimilarity based on genus abundance showed a sequential change in the gut microbial communities (Fig. 2b), especially along the PCo2 axis, where the transition of the gut microbial composition was smooth and gradual so that the differences in the PCo2 scores were significant only between more distant time points, i.e. P2–P4/P5/P6, P3–P5/P6 and P4–P6 (Fig. 2d). The meconium community composition was markedly dissimilar from those in later samples in both cohorts**.** This was especially evident along the PCo1 axis (Fig. 2c, Additional file 1: Figure S4C): in the PMU, the PCo1 scores were significantly different among all time points except for P4. While a global multivariate test for impact of time (partial Mantel test

on Bray–Curtis dissimilarities relative to intersample time interval and donor incongruence matrices) does not achieve significance, we see significant (PCo1 scores by time, LRT, df = 5, P = 0.004; PCo2 scores by time, LRT, df = 5, P = 5.19e−08) differences between time points with regards to the first two principal coordinates of the gut taxonomic data, indicating an overall shift in sample composition over time nonetheless takes place (Fig. 2c, d). Altogether, these results align with previous findings regarding taxonomic composition of the meconium and fecal samples [25, 36], validating the microbial analysis methods used.

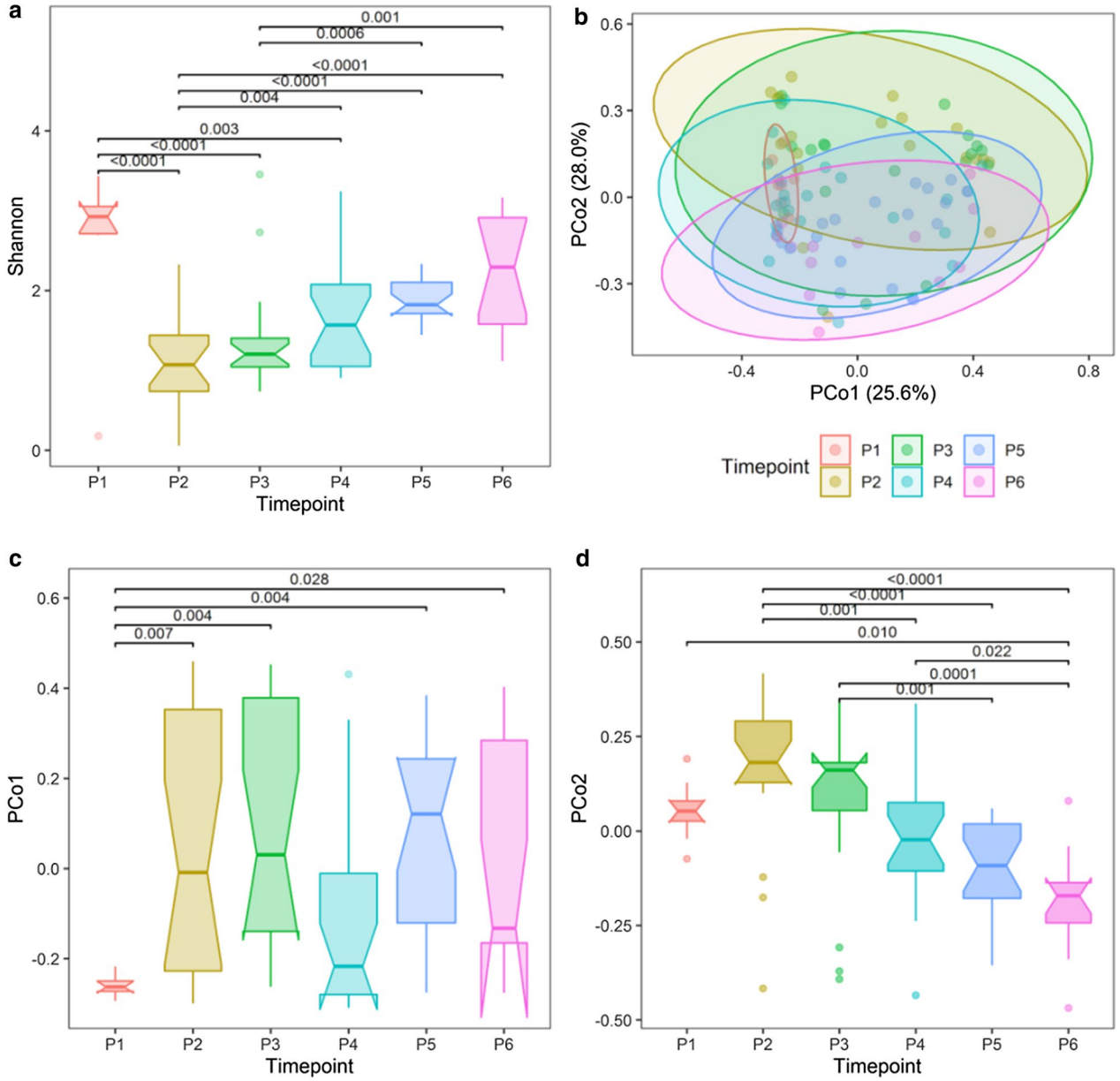

Fig. 2 Gut microbiome: alpha and beta diversity over time. **a** Shannon alpha diversity by time, LRT, df = 5, P = 5.25e−09, adjusted for breastfeeding time and mode of delivery; **b** Principal coordinate analysis plot with Bray–Curtis dissimilarity calculated from genus abundances, ellipses were drawn assuming a multivariate t-distribution; **c** PCo1 scores by time, LRT, df = 5, P = 0.004; **d** PCo2 scores by time, LRT, df = 5, P = 5.19e−08; FDR adjusted p values < 0.05 are shown, P1-meconium, P2-7th day, P3-1st month, P4-6th month, P5-12th month, P6-24th month

Gut microbiota composition and functional profiles during the first two years of life

The bacterial taxonomic and functional (PICRUSt2-predicted) development during the first 2 years of life was analyzed using linear mixed-effect modeling of the centered log-ratio transformed 128 Monte Carlo instances of the Dirichlet distribution in samples from 7th day (P2) to the 24th month (P6) of age. Twenty-seven out of 47 evaluated genera that met the prevalence criterion either increased or decreased significantly during that period (Fig. 3a). Of note, the changes in the *Staphylococcus* and *Ruminococcus* (*torques* group) abundance were approximately monotonic throughout that time (Fig. 3b). In general, the abundance tended to stabilize around 1 year of life (P5) with only mild alterations thereafter (see the column P5P6 in Fig. 3a). Interestingly, except for *Staphylococcus*, the abundance was relatively constant until the end of the first month of life (P3, see the P2P3 column in Fig. 3a). Longitudinal changes in the bacterial composition at higher taxonomic ranks are shown in Additional file 1: Figure S5.

Remarkably, fifteen genera whose abundance changed significantly over time overlapped between the PMU and HMS cohorts (Fig. 3a, Additional file 1: Figure S6). This overlap accounted for 60.0% and 51.7% of the significantly affected taxa in the PMU (P2 to P6) and HMS (P2 to P5), respectively. Moreover, for overlapping genera, the patterns of temporal changes in abundance were similar (either decreasing or increasing); the degree of changes between time points also exhibited a similar pattern (i.e. P2P3, P2P4, P3P4).

To infer the functional profile of the microbial communities, we used the PICRUSt2 software. We identified 110 MetaCyc pathways (out of 332 predicted) in the PMU cohort that were significantly associated with time. In general, almost all pathways (except PWY-7332) did not change significantly in abundance after P5, and the majority of pathways stabilized even earlier at around the 6th month of age (P4). In the HMS cohort, the analysis of predicted MetaCyc pathways revealed 223 pathways (out of 290 predicted) whose abundance changed significantly over the observation period (P2 to P5). Importantly, there were 76 pathways whose abundance changed over time in the two cohorts (Additional file 1: Figure S7), accounting for 75.2% and 34.1% of the significant pathways in the PMU and HMS cohorts, respectively. Remarkably, there was a very high degree of consistency between the cohorts regarding the direction and dynamics of the pathways' abundance.

Gut microbiota diversity, composition and metabolic pathways concerning the stool levels of zonulin and calprotectin

Next, we focused on the relationships between microbiota diversity, composition, predicted function and permeability/inflammatory markers. In a first step, we assessed whether the time intervals with significant changes in bacterial diversity, abundance and predicted metabolic functions overlapped with the time intervals associated with an increase in zonulin (i.e. P2–P4/P5/P6, P3–P4/P5/P6) and decrease in calprotectin (i.e. P2–P5/P6, P3–P5/P6, P4–P6). Then, in a second step, we used a statistical technique (repeated measures correlation) that accounts for non-independence among observations comprising the entire observation period (from P2 to P6), to investigate the associations between the bacterial diversity, abundance and predicted metabolic functions and the concentration of zonulin or calprotectin. In a third step, mixed effects linear models were used to investigate the relationship between changes in bacterial diversity, abundance and metabolic function and the corresponding changes in the concentration of zonulin and calprotectin between all time point pairs.

The compatibility of observations obtained at all stages of the analysis was considered as potential evidence of a causal relationship between the microbiota and the markers analysed. To make this analysis easier accessible, these data are included in summary tables (Additional file 1: Table S5).

Diversity

The Shannon index increased significantly in five out of six time intervals associated with a significant increase in zonulin (i.e. P2–P4/P5/P6, P3–P5/P6, Figs. 1a, 2a) and in four out of five time intervals associated with a significant decline in calprotectin (P2–P5/P6, P3–P5/P6, Figs. 1b, 2a). To investigate whether or not a relationship exists between the alpha diversity and stool zonulin/calprotectin levels, we performed correlation analysis between four alpha-diversity indices (Shannon, Simpson, Chao1, Evenness) and the levels of these biomarkers using repeated measures correlation. The zonulin levels did not correlate with any of the alpha diversity indices (after FDR adjustment). In contrast, the levels of calprotectin inversely correlated with all of the alpha diversity indices calculated (Shannon's-r = −0.30, Q = 0.039; Simpson's-r = −0.23, Q = 0.048, Chao1-r = −0.24, Q = 0.048, Evenness-r = −0.24, Q = 0.048); of note, the strongest correlation was found for the Shannon's diversity (Fig. 4a).

The PCo2 scores decreased significantly in five out of the six time intervals associated with a significant

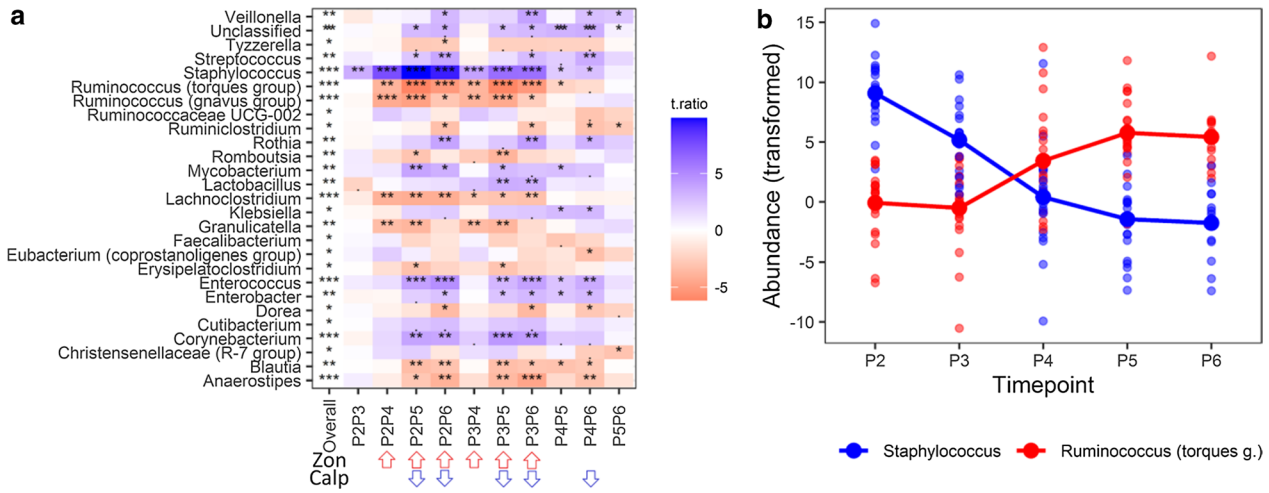

Fig. 3 Changes in the gut microbiota composition during first two years of life. **a** linear mixed effects analysis followed by pairwise comparison of time- points (adjusted for mode of delivery and breastfeeding time), **a** Genera (present in at least 20% samples) significantly associated with time, **b** Longitudinal abundance of the *Staphylococcus* and *Ruminococcus* (*torques* group). The overall p-value—a likelihood ratio test (LRT) of nested models (FDR adjusted across genera); PXPY—contrast p values between the two time points (PX and PY), FDR adjusted for all possible contrasts, t.ratio—t statistics for the contrasts estimates (a positive value, colored blue, indicates a decrease in abundance, a negative value, colored red, indicates an increase in abundance). Red arrows indicate time intervals with a significant increase of zonulin, blue arrows indicate time intervals with a significant decrease of calprotectin (see also Fig. 1). Taxa abundances (unrarefied) were transformed by generating 128 Monte Carlo instances of the Dirichlet distribution for each gut sample, followed by center-log transform of each instance. A linear mixed effects analysis was performed for each instance separately and the results were averaged over 128 instances. P2-7th day, P3-1st month, P4-6th month, P5-12th month, P6-24th month

rise in zonulin (P2–P4/P5/P6, P3–P5/P6, Figs. 1a, 2d) and in all five time intervals associated with a significant decrease in calprotectin (P2–P5/P6, P3–P5/P6, P4–P6, Figs. 1b, 2d). Besides, there was some similarity between the longitudinal zonulin and PCo1 scores pattern (Figs. 1a, 2c). Therefore, we explored the correlations between the zonulin and calprotectin levels and the per-sample positions on the first two PCoA axes using repeated measures correlation analysis. We did not find any significant correlation (zonulin and PCo1: $r = 0.06$, $P = 0.637$; zonulin and PCo2: $r = -0.23$, $P = 0.053$; calprotectin and PCo1: $r = 0.11$, $P = 0.363$; calprotectin and PCo2: $r = 0.19$, $P = 0.109$) suggesting that the overall compositional profiles (captured by the first two PCo dimensions) were not associated with the levels of zonulin or calprotectin (Fig. 4b, c). Moreover, using likelihood ratio tests of the nested models, no association the between Bray–Curtis distances and the changes in the zonulin levels was found (non-interaction model: $df = 1$, $P = 0.121$, $Q = 0.241$; interaction model: $df = 9$, $P = 0.965$, $Q = 0.965$). However, there was a significant association between the Bray–Curtis distances and the changes in calprotectin levels (time period x calprotectin change interaction model: $df = 9$, $P = 0.013$, $Q = 0.025$). The post-hoc tests revealed significant correlations between the changes in calprotectin levels and the Bray–Curtis distances for P2–P3 and P3–P4; however, they were not significant after FDR adjustment ($Q = 0.216$ and $Q = 0.110$, respectively, Fig. 4d).

Microbiota composition and functional profiles concerning stool zonulin and calprotectin

Despite the lack of associations between alpha diversity and the zonulin levels as well as between beta diversity and the levels of both biomarkers (zonulin and calprotectin) (Fig. 4), we sought to examine more particular relationships between the levels of these permeability/inflammation-associated biomarkers and the abundances of specific gut bacterial taxa in the context of five taxonomic ranks (from genus to phylum levels) as well as the PICRUSt2-predicted MetaCyc metabolic pathways.

Interestingly, the changes in abundance of several bacteria coincided with significant time dependent changes in the levels of zonulin or calprotectin (Figs. 1, 3a). For example, there was an increase in the abundance of *Lachnoclostridium*, *Ruminococcus* (*gnavus* group), Carnobacteriaceae, Lachnospiraceae, Peptostreptococcaceae, Coriobacteriales, and a decrease of Corynebacteriales among all time points associated with an increase in the zonulin concentration. Additionally, there was an increase in the abundance of *Anaerostipes*

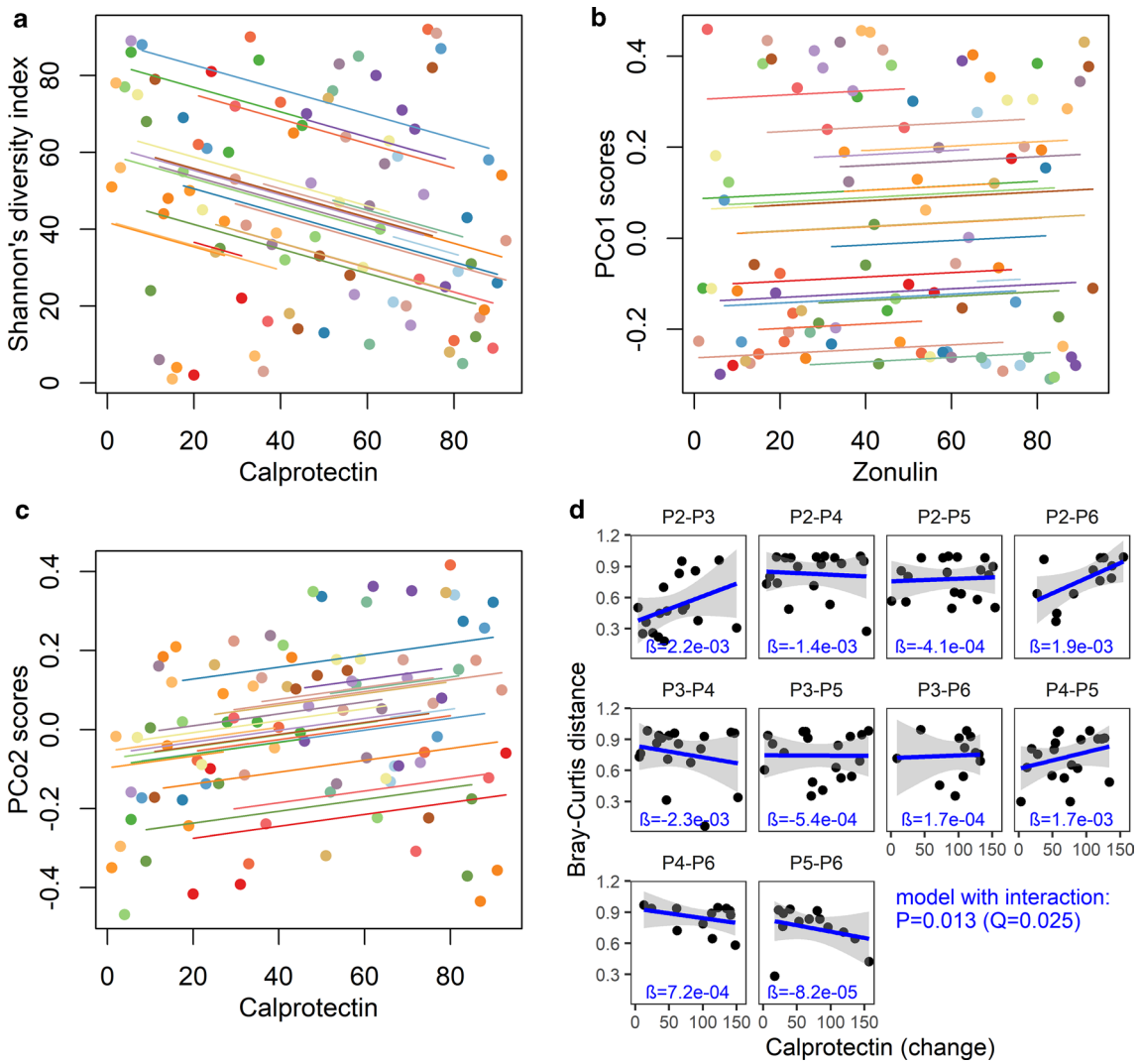

Fig. 4 Alpha and beta diversity (Bray–Curtis distance) compared to zonulin and calprotectin. **a** Repeated measures correlation between calprotectin and Shannon diversity index. **b** Repeated measures correlation between zonulin and PCo1 scores, **c** Repeated measures correlation between calprotectin and PCo2 scores, **d** Correlation between Bray–Curtis distance and calprotectin change for all possible time point pairs. Significant likelihood ratio test (LRT) indicating an interaction between time period and calprotectin change (df = 9, P = 0.013, Q = 0.025), post-hoc test revealed significant correlations between Bray–Curtis distance and calprotectin change between P2 and P3 (β = 2.18e−03, P = 0.043) and P3 and P4 (β = − 2.31e−03, P = 0.011); alpha diversity, zonulin calprotectin and calprotectin change were rank transformed, repeated measures and correlation trend lines are coloured by individual child. PCo1 and PCo2 scores were treated here as a dimensionality-reduced measure of sample composition relative to whole-dataset variability

and a decrease in the abundance of Enterococcaceae among all time points associated with a decrease in calprotectin concentration. To investigate this further, we computed the repeated measures correlations (encompassing the whole study period; Fig. 5 and Additional file 1: Table S5) between the levels of both markers and the gut microorganisms.

In the correlations between gut microorganisms and stool calprotectin, we observed significant associations at the genus level (n = 3), the family level (n = 3), and the order level (n = 1). Moreover, the levels of zonulin were correlated with the abundance of 19 microorganisms: at the genus level (n = 5), at the family level (n = 5), at the order level (n = 5) and the class level (n = 4) (Fig. 5a and Additional file 1: Table S1). Specifically, calprotectin correlated positively with *Staphylococcus*, *Enterococcus*, Staphylococcaceae, and Enterococcaceae and negatively with *Anaerostipes*, Ruminococcacea and Clostridiales, whereas zonulin correlated positively with *Lachnoclostridium*, the *Ruminococcus (gnavus* group), the

Ruminococcus (torques group), Lachnospiraceae, Peptostreptococcaceae, Ruminococcacea, Erysipelotrichales, Coriobacteriales, Clostridiales, Clostridia, Coriobacteria and negatively with *Staphylococcus, Corynebacterium*, Moraxellaceae, Staphylococcaceae, Bacillales, Pseudomonadales, Alphaproteobacteria, and Bacilli. Of note, the genus *Staphylococcus*, the families Ruminococcaceae and Staphylococcaceae, and the order Clostridiales were simultaneously and oppositely correlated with both biomarkers in the PMU cohort (Fig. 5a) and with calprotectin in the HMS cohort, exhibiting the same direction and similar magnitudes as those in the PMU cohort (Additional file 1: Table S2). Additionally, zonulin correlated significantly with 10 MetCyc pathways belonging to 4 superclasses (Superpathways: FUC-RHAMCAT-PWY, Degradation/Utilization/Assimilation: FUCCAT-PWY, PWY-6353, PWY-7237, SALVADEHYPOX-PWY, Generation of Precursor Metabolites and Energy: PWY-5676, PWY-5677, PWY-6588, PWY-7003, and Biosynthesis: PWY-6478), of which two pathways, PWY-5677 and PWY-6588, were the only pathways associated with calprotectin (Fig. 5b). The correlation between calprotectin and predicted metabolic pathways from the HMS cohort revealed 15 pathways significantly correlating with calprotectin levels; however, none of them overlapped with the correlations found in the PMU cohort (Additional file 1: Table S2).

In order to gain more insight into the causal relationship between the microbiota and zonulin/calprotectin, the changes in the gut community features (taxonomic composition and functional profile) and the changes in the levels of zonulin or calprotectin among the time points were correlated using a linear mixed-effects model. We did not find any association between changes in microbiota and changes in markers, except for the genus *Ruminococcus* (*torques* group) which was significantly associated with the changes in levels of calprotectin. he interaction model did not fit significantly better than the common (positive) slope model (Fig. 6a, Table 2). Thus, the abundance of *Ruminococcus* (*torques* group) changes in parallel with the changes in stool calprotectin levels regardless of time interval (which is especially prominent for the P2–P4, P2–P5, P3–P4, P4–P6 time intervals), thereby suggesting that both can be causatively connected. However, this observation was not confirmed in the validation analysis and must be taken with caution. The analysis in the context of higher taxonomic ranks for the PMU and HMS cohorts are shown in the Additional file 1: Figures S8–S10.

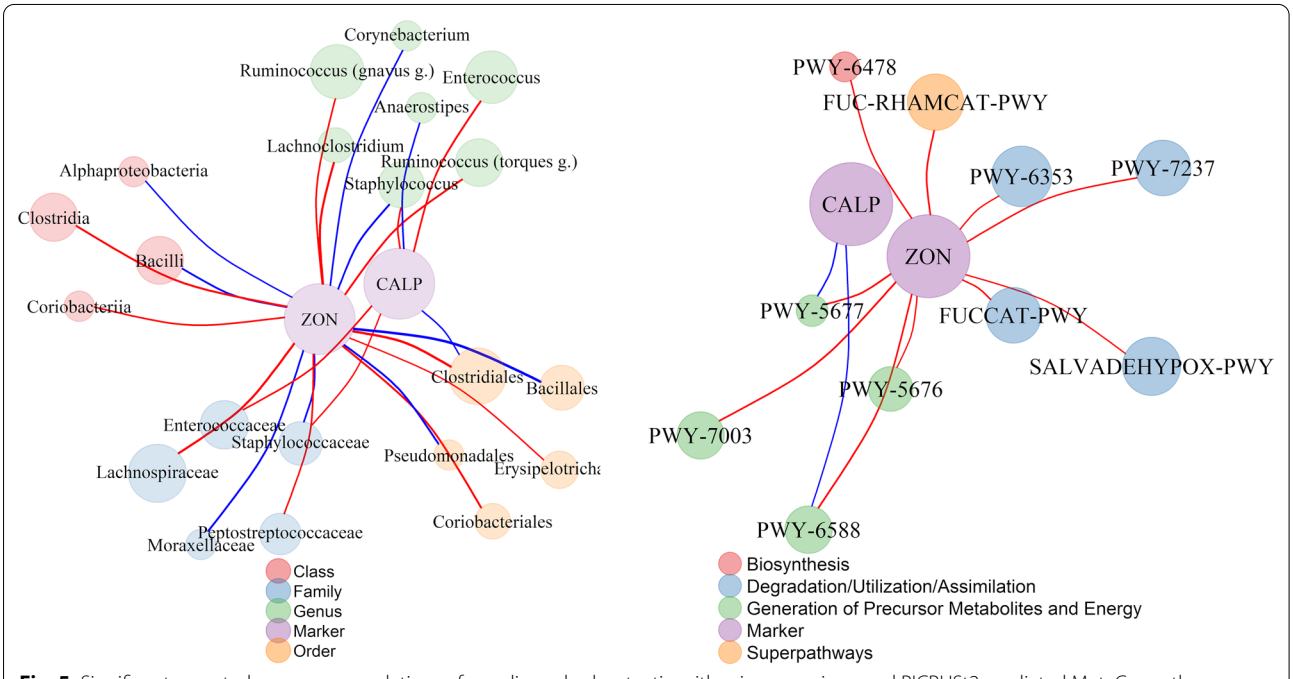

Fig. 5 Significant repeated measures correlations of zonulin and calprotectin with microorganisms and PICRUSt2 predicted MetaCyc pathways. **a** taxon correlation network, **b** MetaCyc pathway correlation network node size and edge width are proportional to microorganisms/pathways abundance and repeated measure correlation coefficients, respectively; line color indicates the negative (blue) and positive (red) correlations, node color indicates taxonomic rank or metabolic pathway categories. Taxonomic abundances (unrarefied) were transformed by generating 128 Monte Carlo instances of the Dirichlet distribution for each gut sample, followed by center-log transform of each instance

In addition to the changes in the composition of the human gut microbiota, we also investigated whether the changes in functional profiles of gut microbiota were linked to the dynamics of the zonulin or calprotectin levels. Remarkable, we found that the dynamics of stool calprotectin was negatively associated with changes in two MetaCyc metabolic pathways: pyruvate fermentation to butanoate (CENTFERM-PWY) and the superpathway of *Clostridium acetobutylicum* acidogenic fermentation (PWY-6590) (Fig. 7c, d). The common slope model was the best fit for these pathways (Table 2). For zonulin, the common slope model was the best fit for a change in the pentose phosphate pathway (PENTOSE-P-PWY) only Fig. 7a, Table 2). In addition, the dynamics of zonulin was associated with changes in 3 MetaCyc pathways: pyrimidine deoxyribonucleotides de novo biosynthesis III (PWY-6545), the superpathway of glucose and xylose degradation (PWY-6901), and the super pathway of pyridoxal 5′-phosphate biosynthesis and salvage (PWY0-845). For those 3 pathways, as illustrated in Fig. 7b for PWY-6545, the models with different slopes (i.e. including an interaction between all time point pairs and pathway changes) fitted the data significantly better than the common slope model. For example, the change in zonulin levels was associated negatively with a change in PWY-6545 (Fig. 7b, Table 2) between P3 and P4 ($P=0.001$, $Q=0.012$), yet positively between P4 and P6 ($P=0.003$, $Q=0.012$) as well as P5 and P6 ($P=0.004$, $Q=0.012$). For the other two pathways, PWY-6901 and PWY0-845 (Table 2), significant associations were found between P3 and P4 ($\beta=23.60$, $SE=7.35$, $P=0.002$, $Q=0.008$) as well as P5 and P6 ($\beta=36.08$, $SE=10.56$, $P=0.0008$, $Q=0.008$) for PWY-6901, and between P3 and P4 for the PWY0-845 ($\beta=13.07$, $SE=3.66$, $P=0.0005$, $Q=0.005$). For the full results, see Additional file 1: Tables S3 and S4.

However, it should be noted that the above results (concerning calprotectin) were not confirmed in the HMS cohort: 44 pathways showed a correlation with the dynamics of calprotectin (all under the common slope model), but the CENTFERM-PWY ($P=0.051$, $Q=0.131$) and PWY-6590 ($P=0.056$, $Q=0.140$) did not (Additional file 1: Figure S11). The relationships between the changes in abundance in the gut, the MetaCyc pathways, and the changes in calprotectin levels in the HMS cohort are shown in Additional file 1: Figure S12.

Discussion

This is the first study in which putative associations between the gut microbiota and the concentrations of zonulin and calprotectin in children's stool during the first 2 years of life have been investigated. The study of zonulin allows the non-invasive assessment of the functional state of the small intestinal paracellular permeability; additionally, the calprotectin measurements provide insight into the development of the immune system after birth. The observed changes of both markers concentration indicate that the 6th month of life can be defined as the key point for forming the small intestinal barrier and for the development of postnatal immunity. It should also be stressed that, usually, in the sixth month of life, solid food is introduced to children's diet, which can significantly impact both gut microbiota and intestinal permeability.

Although zonulin seems to be a valuable marker of small intestinal paracellular permeability, the data in the context of children aged 2 years old or younger are scarce and focused on infections [37] and prematurity [38]. The fecal zonulin concentration in children during the first 2 years of life was not reported and its role during this period is unknown. Its relationship with gut microbiota seems to be interesting [38]. Zonulin is involved in controlling the passage of molecules weighing at least 3.5 kDa [39] through the intestinal barrier via its reverse influence on TJ tightness [40, 41]. After the activation of ZO-1, zonulin controls the low-capacity "leak" type route characterized by low selectivity [42, 43]. Interactions between zonulin and gut microbiota can be bidirectional. Bacterial and gluten exposure was defined as intensive triggers of zonulin release [19, 44]. Of note, the zonulin pathway is an innate defensive mechanism of the host, able to control the gut microbiome composition via the "flushing out" of microorganisms by water secreted into the intestinal lumen following hydrostatic pressure gradients [45]. High stool zonulin levels suggest that the gut barrier allows the free flow exchange of various particles in infants. Moreover, it is likely that the commercially available ELISAs detect one or more members of the zonulin family that have not been discovered yet but play some important role during the first 2 years of life [46].

High levels of calprotectin in feces during the first months of life are associated with the regulation of the development of the immune system in neonates as well as with the adaptation to new environmental conditions [25]. The decrease in the fecal calprotectin concentration implies that inflammatory processes in the gut tend to decrease from the 6th month of age, which, with the accompanying increased paracellular intestinal permeability and bacterial alpha diversity, provides suggestion of immune tolerance. Willers et al. [25] demonstrated that mice exposed to calprotectin immediately after birth induced microbial tolerance against the first wave of microbial colonization; on the other hand, the same exposure after the neonatal period was associated with pro-inflammatory responses. Wood et al. observed an increased proportion of regulatory T cells (Tregs) during the first three weeks of life [47] which is associated

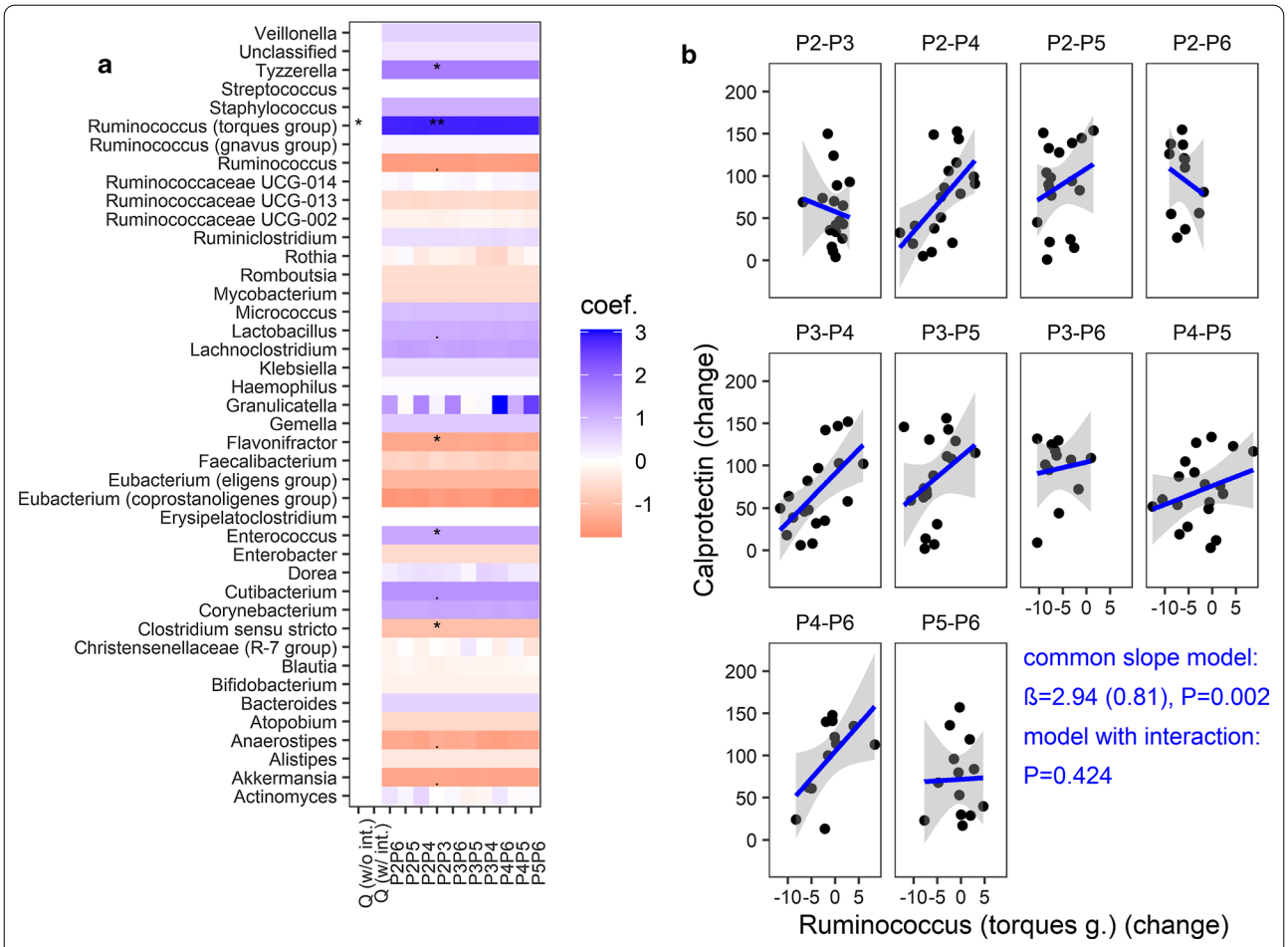

Fig. 6 Relationship between gut genus abundance and calprotectin levels (changes per time interval). **a** Linear mixed effects models were used to test for an association between the taxon abundance change and calprotectin change accounting for ten time point pairs (P2–P3, P3–P4, etc.) from the same subject**,** adjusted for mode of delivery and breastfeeding time. Two models were considered: without interaction (w/o int.) and with interaction between time point pair and taxon change (w/ int.). If the interaction term was significant—Q (w/ int.) < 0.05, individual p values (FDR adjusted) were interpreted. If the common slope model was chosen—Q (w/o int.) < 0.05 and Q (w /int) > 0.05, the same common slope coefficient β (coef.) across all time points and one p-value (P2–P3) were shown. Coefficients and Q values were averaged over 128 Monte Carlo instances of the Dirichlet distribution, followed by center-log transform of each instance, **b** the common slope model (LRT, df = 1, P = 0.002, Q = 0.015) illustrating relationships between *Ruminococcus (torques* group) change (centered log-ratio transformed) and calprotectin change (rank transformed)—based on the first Monte Carlo instance

Table 2 Association of the marker change (zonulin, calprotecin) with taxon/pathway change—a linear mixed model

Taxon/Pathway vs marker (changes)	Common slope (no interaction)			Interaction (taxon/pathway change x time interval)	
	β (SE)	P	Q	P	Q
Ruminococcus (*torques* group) vs calprotectin	2.94 (0.81)	0.002	0.015	0.424	0.910
CENTFERM-PWY vs calprotectin	− 4.54 (1.08)	0.0002	0.028	0.412	0.993
PWY-6590 vs calprotectin	− 4.48 (1.16)	0.0002	0.026	0.421	0.993
PENTOSE-P-PWY vs zonulin	9.49 (1.93)	3.98e−6	0.001	0.123	0.356
PWY-6545 vs zonulin	Na	–	–	4.15e-5	0.014
PWY-6901vs zonulin	Na	–	–	0.0005	0.046
PWY0-845 vs zonulin	Na	–	–	0.0008	0.049

β, coefficient estimate; SE, standard error; Q -FDR adjusted P value, Na, not applicable

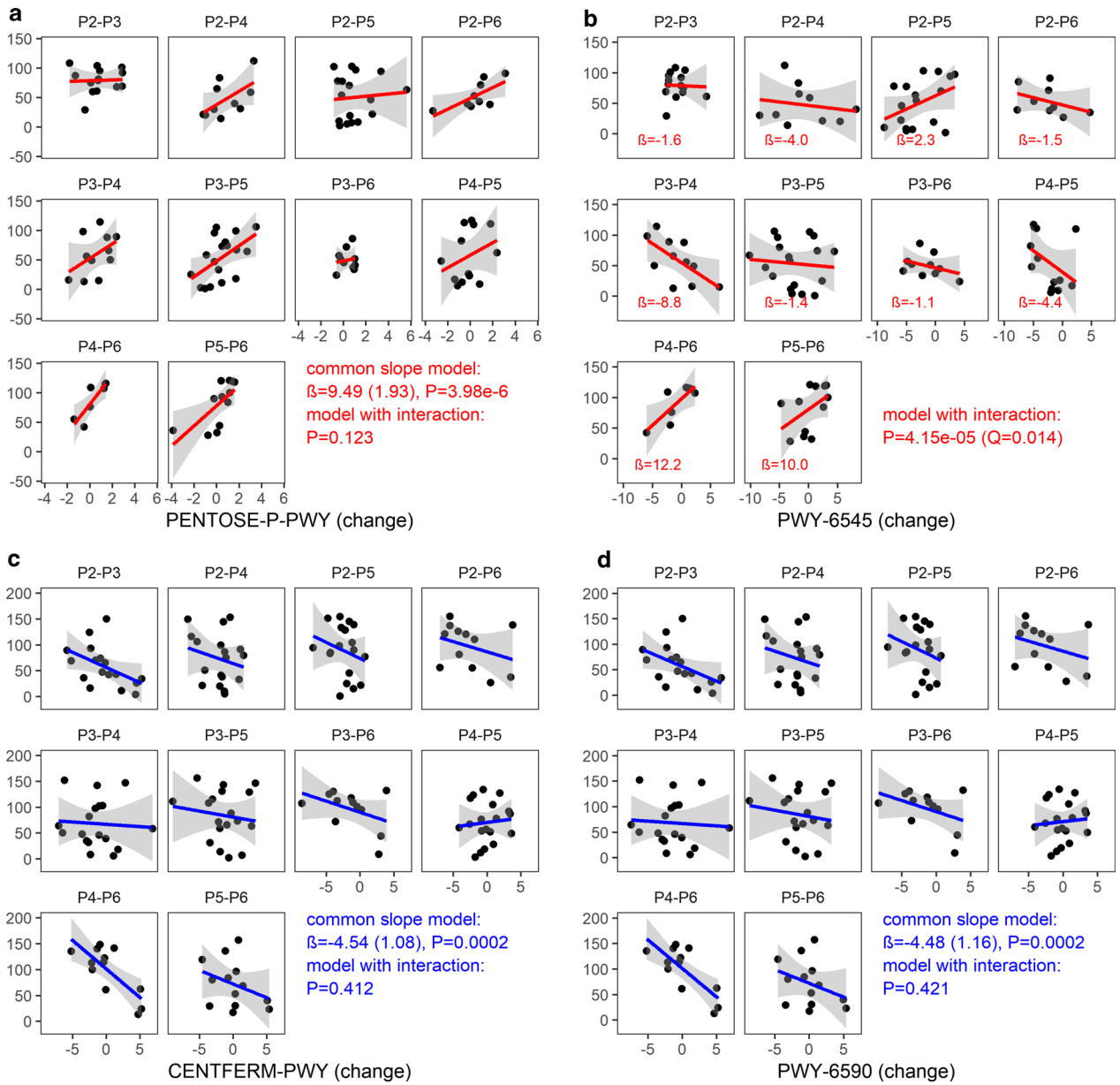

Fig. 7 Relationship between (changes in) gut abundance of inferred MetaCyc pathways and (changes in) zonulin/calprotectin levels. **a** The common slope model (LRT, df = 1, P = 3.98e−6, Q = 0.001) illustrating relationships between PENTOSE-P-PWY change (centered log-ratio transformed) and zonulin change (rank transformed), **b** the model with time point pair by PWY-6545 change interaction (LRT, df = 9, P = 4.15e−5, Q = 0.014), **c** common slope model illustrating relationships between CENTFERM-PWY change (centered log-ratio transformed) and calprotectin change (rank transformed), **d** common slope model illustrating relationships between PWY-6590 change (centered log-ratio transformed) and calprotectin change (rank transformed); Likelihood ratio test (LRT) P- and Q-values were computed based on 128 Monte Carlo instances of the Dirichlet distribution, followed by center-log transform of each instance, the plots were based on the first Monte Carlo instance

with increased immune tolerance after the birth. Our results did not fully corroborate these observations. The intestinal permeability in the first month of life, as per the zonulin concentration, is lower than that in the later period, suggesting that the possibility of antigen translocation early after birth is limited. After the sixth month of age, the intestinal paracellular permeability increases, but this does not increase inflammation, as per the calprotectin concentration, which indicates the development of immune tolerance. The differences between the two studies may be due to the fact that Willers et al. did not directly measure intestinal permeability and additionally

used an animal model that does not closely mimic the timing of human intestinal barrier development and Wood et al. measured Treg proportion only during the first three weeks of life however. As such, direct comparison of immune tolerance as reported by them and as per our present study is not possible, and we look forward to future studies bridging this gap. Moreover the role of paracellular way in translocation of bacteria and endotoxin is still the subject of debate [22], what requires that the observed results should be interpreted with caution.

In summary, the increased production of zonulin may be caused by changes in intestinal bacteria and by the introduction of gliadin into the diet. In contrast, the decrease in the calprotectin content in the stool may indicate immunological tolerance development. However, the breast milk is an essential source of calprotectin, which can significantly influence the fecal calprotectin content. This said the observed changes cannot be explained solely by environmental factors as zonulin decreases in the later years of life despite the constant consumption of gliadin. Therefore, in the context of longitudinal analyses of biomarker trajectories, we applied corrections for the time of breastfeeding and the type of delivery to minimize the influence of these confounders. To help resolve the mechanisms underlying our results, we decided to analyze whether a cause-effect relationship between the microbiota and the biomarkers would be observed. Of note, the microbial changes in the first 2 years of life are very dynamic and depend on the age, mode of delivery, feeding, antibiotic treatment, and other environmental factors [28, 29]. However, a comprehensive analysis of all factors influencing the gut microbiota is beyond the scope of this study, which is focused on the relationship between the microbiota and the fecal zonulin and calprotectin content. This said the compositional and functional comparison of the PMU and HMS cohorts' microbiota would be very interesting.

The Shannon index (more sensitive to species richness), shared 5 and 4 significant time intervals with zonulin (increasing) and calprotectin (decreasing), respectively. Similar results were observed for coordinates of the first two axes in the context of a principal coordinate analysis based on the Bray–Curtis distance, one of beta-diversity measures. However, it translated only into a negative correlation between alpha-diversity indices, the Shannon index in particular, and calprotectin. This negative correlation may be associated with the development of immunological tolerance. The microbial colonization of the intestine after birth will generate a diversity of new antigens that will play an important role in the stimulating of epithelial function and the establishing the offspring's immune system [48]. Of note, weaning (the average breastfeeding time in our study was about 30 weeks), and the resulting loss of milk-borne calprotectin, sIgA and other anti-microbial factors, along with the transition from mothers' milk to a complex diet (including solid food), can have a major impact on the dynamic microbiota development, and thus in the resulting immune responses during the neonatal period [6].

Willers et al. [25] reported that the gut microbiota's overall diversity significantly increases during the first year of life, similarly to what we observed in this study. Moreover, they observed the increased abundance of the bacterial classes Actinobacteria, Bacteroidia [49–51] and Clostridia, along with the decreased abundance of Bacilli and Gammaproteobacteria over the same time period. Of note, Willers et al. [25] demonstrated that during infancy, high abundance of Actinobacteria and low abundance of Gammaproteobacteria are linked to high fecal calprotectin levels; additionally, at the family level, fecal calprotectin promoted the higher abundance of Bifidobacteriaceae and the reduction of Enterobacteriaceae via the production of acetate [52]. Taxonomic composition observed by Willers et al. [25] might be linked to the elevation in the synthesis and reduction in the degradation of SCFAs and therefore leading to the overrepresentation of health-promoting gut microbiota metabolic functions. Our results confirm the observations obtained by Willers et al. [25] in terms of abundance; however, the correlations in the cited study were not reproduced here. Importantly, the observed differences may depend on the difference between the investigated populations (Polish and German) or/and the analytical methods used. This said the analysis of both data-sets showed similarities with respect to the correlation between the levels of calprotectin and bacteria at the family level.

Interestingly, the bacteria associated with zonulin and calprotectin (Table 3) may be involved in the production of SCFAs, which is a protective factor in the context of intestinal barrier integrity and of inflammatory processes. Particularly, the different taxonomic groups of bacteria within the class Clostridia were associated with increased zonulin and decreased calprotectin concentrations. Within this class, the family Ruminococcaceae was more abundant in women with low zonulin concentrations in a previous study; however, here, we observed the opposite pattern. Importantly, not all bacteria within this class have the same properties. For instance, the genus *Ruminococcus* (*gnavus* group)—positively correlated with fecal zonulin concentration, is increased in patients with IBS-D and Crohn's disease, which suggests its pro-inflammatory activity. An interesting result was obtained in the case of the *Ruminococcus* (*torques* group). *Ruminococcus torques*, a butyrate producer, have been found to dominate patients' gut milieu with Crohn's Disease [53, 54]. They were also studied in the context

of individuals with increased risk of upper gastrointestinal tract involvement, but with inconclusive results [55–57]. Importantly, *Ruminococcus torques* are able to utilize mucin in the human intestine, sustaining their adaptability to the human intestinal environment [54, 58, 59]. Consequently, increased mucin degradation makes luminal antigens cross the gut barrier and stimulate the immune system, leading to intestinal disorders [60]. Of note, in this study, the abundance of these above bacteria positively correlated with the concentration zonulin and the dynamics of calprotectin concentration. Therefore, bacteria from the *Ruminococcus* (*torques* group) may include species associated with small intestinal permeability and immune responses in the first 2 years of life. For deeper species resolution and assessment of the gut functional potential future shotgun metagenomics, metabolomics.

It is not possible to draw definitive conclusions regarding cause-effect relationships in a descriptive study like this, especially not given limited power due to sample size (see Methods). Moreover, we cannot exclude observed relationships could reflect hidden shared relationships with other covarying factors. However, as the association between calprotectin and *Ruminococcus* abundance remained statistically significantly also accounting for age and delivery mode under stringent FDR adjustment, we can likewise conclude it is not reducible to the role of these factors, but rather constitutes an integral part of the system, with findings and implications beyond the age/maturity aspect. Details of the associations between bacteria and markers and bacteria properties are presented in Table 3. Only some of the relationships are discussed below. In this study, the Bacilli class correlated with decreased zonulin and increased calprotectin levels. In general, this class is considered beneficial for gut health, which is in line with decreased zonulin levels. On the other hand, the family Staphylococcaceae and genus *Staphylococcus* from the Bacilli class are the most prominent bacteria in human milk, an important calprotectin source in the infant. Family Enterococcaceae belongs to opportunistic pathogens, considered as potential marker for IBD, what can explain correlation with increased fecal calprotectin concentration. However, our present results do not allow consistent conclusions on the role of bacteria in the small intestinal barrier's permeability or the immune system in the first 2 years of life. For example, the *Lachnoclostridium* genus produces SCFAs and occurs in lower amounts in patients with gastrointestinal cancers; therefore, in theory, it should decrease intestinal permeability; however, it was correlated with increased concentrations of zonulin in the neonatal stool in this study. Alphaproteobacteria occur in high numbers in patients with IBD and are negatively associated with the concentration of zonulin. The order Corynebacteriales is associated with decreased zonulin and belongs to the class Actinobacteria, which was positively correlated with calprotectin in the study by Willers et al. [25]. The order Pseudomonadales and the family Moraxellaceae are negatively correlated with zonulin and belong to Gammaproteobacteria, negatively correlated with calprotectin in the study by Willers et al. Therefore, in no case was there a full consistency of microbiota changes between particular time points and changes in the levels of zonulin and/or calprotectin. This said the validation analysis suggested that the genus *Staphylococcus* and the family Staphylococcaceae (positively correlated with the concentration of calprotectin), as well as the family Ruminococcaceae and the order Clostridiales (negatively correlated with the concentration of calprotectin), are potential candidates as markers of the development of the immune system in children under 2 years of age.

Finally, we analyzed the metabolic pathways linked with bacterial abundance (Table 4). We were able to show that the pathways involved in the production of SCFAs and the metabolism of carbohydrates were positively associated with fecal zonulin and negatively with calprotectin. The latter association suggests the role of SCFAs in the immune response. Indeed, SCFAs regulate the body's immune response through free fatty acid receptors (FFARs) that are expressed throughout the body, including along the gastrointestinal tract [61]. SCFAs inhibit the activity of histone deacetylase (HADC), which activates gene expression in a pro-inflammatory response. Butyric acid plays a special role here via the binding of two butyrate molecules in the hydrophobic part of the enzyme, which is an example of the intracellular action of SCFAs. Moreover, the secretory activity of macrophages for the production of pro-inflammatory cytokines Il-6 and IL-12 is reduced along with the increase of butyrate content. Some studies suggest that SCFAs initiate the transformation of naive T_{CD4+} cells into regulatory cells, while others indicate SCFAs stimulate Treg cells already present in the colon [62–66]. Atarashi et al. [67, 68] demonstrated that colonization of germ free rodents with appropriate strains of Clostridium influenced the multiplication and differentiation of Treg cells and stimulated anti-inflammatory cytokines synthesis, predominantly interleukin 10 (IL-10). Notably, butyric acid of all SCFAs is metabolized to the greatest extent by the colon epithelial cells. Thus, it is the main source of energy for the colonocytes as it is responsible for maintaining the integrity of the intestinal barrier. This is evidenced by the effects of butyric acid supplementation of Caco 2 cells. Butyrate increased the concentration of ZO-1 and activated AMPK (adenosine monophospate activated protein kinase), which resulted in TEER (transepithelial electrical

Table 3 Levels of associations between bacteria and markers

Bacteria (taxonomic affiliation)	Change[a]		RMC	LME	Bacteria properties
	ZON	CALP			
Class Alphaproteobacteria			Zon (−)		Exhibited numerous positive co-occurrences with other microbes [71]; enriched in vaginally delivered standard formula-fed infants three months post delivery [72]; high abundance in placentas of women with IBD, but not in stool [73]; high abundance in patients with Crohn's Disease [74]; third major constituent gut microbiome in infants with respiratory and gastrointestinal diseases [75]
Class Clostridia			Zon (+)		Dominates in infants [76]; in the formula-fed infants colonization occurred consistently throughout the 1st year of life, whereas in some breast-fed infants it was inhibited until weaning [77]; anti-inflammatory activity [78–80]
Order Clostridiales (class Clostridia)			Calp (−) Zon (+)		Butyrate production [81, 82]; suppression of proinflammatory bacteria [83]; protective role of the taxa in IBD pathogenesis [84]; induction an immune response [67]
Family Ruminococcaceae (class Clostridia)			Calp (−) Zon (+)		Ruminococcaceae is more abundant in women with low serum zonulin concentration and can enhance intestinal barrier integrity [85]; involved in modulation of gut barrier function[86]; SCFAs producer [87]; more abundant in people with diets high in carbohydrates [88, 89]
Genus *Ruminococcus (gnavus* group) (family Ruminococcaceae, order Clostridiales, class Clostridia)	Same		Zon (+)		Increased in patients with IBS-D [90] and Crohn's Disease [91]
Genus *Ruminococcus (torques* group) (family Ruminococcaceae, order Clostridiales, class Clostridia)			Zon (+)	Calp (common slope model, positive β coeff.)	Butyrate production, important role in Crohn's Disease [91]; mucin degradation [54]
Family Lachnospiraceae (order Clostridiales, class Clostridia)	Same		Zon (+)		Butyrate producer [92] induces the expression of Treg—suppression of the colonic inflammatory response [62]
Genus *Lachnoclostridium* (family Lachnospiraceae, order Clostridiales, class Clostridia)	Same		Zon (+)		Enhanced the utilization of carbohydrates [93] showed a positive association with secondary bile acids in mice [94]; SCFAs producer [95]; correlated with vitamin B6 metabolism and tryptophan metabolism in mice with colitis [96]; lower abundances was observed in neoplasms of gastrointestinal tract [97]; decreased in Autism Spectrum Disorders [98]; overdominant in population of native Alaska with high incidence of sporadic colorectal cancer [99]; positively correlated with body mass index, alanine aminotransferase, aspartate aminotransferase levels in patients with non-alcoholic fatty liver disease (NAFLD) [100]
Genus *Anaerostipes* (family Lachnospiraceae, order Clostridiales, class Clostridia)		Oppo-site	Calp (−)		Butyrate producer [101]
Family Peptostreptococcaceae (order Clostridiales, class Clostridia)	Same		Zon (+)		Overrepresented in gut microbiota of non-breastfed infants [102]; present in infants fed with standard cow's milk formula [103]; over-represented in infants living with pets and underrepresented in infants living with older siblings [104]

Table 3 (continued)

Bacteria (taxonomic affiliation)	Change[a]		RMC	LME	Bacteria properties
	ZON	CALP			
Class Bacilli			Zon (−)		Dominant in duodenum and in the jejunum; producer of enzymes, bile acid hydrolases, ACE-like inhibitors, SCFAs, hydrogen peroxide, group B vitamins, beneficial for gut health [105]; participate in metabolism of dietary components, xenobiotics and drugs helping to maintain intestinal homeostasis and host health [106, 107]
Order Bacillales (class Bacilli)			Zon (−)		Enriched in preterm infants fed with mom's own milk [108]; overrepresented in cesarean section delivered neonates [109]; second major microbiome constituent of neonatal intensive care unit rooms [110]
Family Staphylococcaceae (order Bacillales, class Bacilli)			Calp (+) Zon (−)		Present in human milk [111]
Genus *Staphylococcus* (family Staphylococcaceae, order Bacillales, class Bacilli)			Calp (+) Zon (−)		One of the most prominent bacteria of human milk [112]; correlated with higher fecal calprotectin concentration in infants [113]; pathological bacteria decreased over time after delivery [114], super antigen function stimulates the systemic secretion of IgA in neonates, protecting against allergies [115]; promotes the modulation of the infant immune system without causing an adverse inflammatory response, induction of T-reg cells via butyric acid and propionic acid [116]; participate in the saccharolytic fermentation of carbohydrates, which end products that positively affect host cells and gut bacterial community [117]
Family Enterococcaceae (class Bacilli)		Same	Calp (+)		Opportunistic pathogen [118]; potential biomarker for IBD [119]
Genus *Enterococcus* (family Enterococcaceae, class Bacilli)			Calp (+)		Higher counts in premature infants with necrotizing enterocolitis (NEC) [120]; significant positive correlations between fecal calprotectin levels and intestinal colonization levels Enterococcus in Preterm Infants during the Neonatal Period [17]; more abundant in feces of formula fed infants [121]; a leading hospital-associated pathogen [122]
Family Carnobacteriaceae (class Bacilli)	Same				potential biomarker of autoimmunological disease [123]; abundance in oral microbiome was associated with decreased risk of colorectal cancer [124]
Class Coriobacteriia			Zon (+)		Greater prevalence in formula-fed babies [125]; significantly higher in coeliac infants [126]; lower relative abundances in atopic dermatitis children [127]; increased in Crohn's Disese [128]
Order Coriobacteriales (class Coriobacteria)	Same		Zon (+)		Negatively associated with fecal protease activity associated with various gastrointestinal tract diseases [129]; can be considered as pathobionts, because their occurrence has been associated with a range of pathologies such as bacteremia, periodontitis, and vaginosis [130]; belongs to class Actinobacteria which correlates with fecal calprotectin [25]
Order Corynebacteriales	Oppo-site				Subset is pathogen [131]; belongs to class Actinobacteria which correlates with fecal calprotectin [25]

Table 3 (continued)

Bacteria (taxonomic affiliation)	Change[a]		RMC	LME	Bacteria properties
	ZON	CALP			
Genus *Corynebacterium* 1 (order Corynebacteriales)			Zon (−)		Human skin bacteria, dominant in gut microbiome of Cesarean Section (CS) born infants [132]; counts in gut microbiome of CS born children negatively associated with maternal dairy intake [133]; present in Meconium of Preterm Neonates [134]; dominant in oral microbiome of CS born infants [135]
Order Erysipelotrichales			Zon (+)		More abundant in infants at 4 weeks fed with cow-milk based formula, at 26 weeks in breast milk fed children and those receiving fructooligosaccharides (FOS) [103]; decreased in new-onset Crohn's disease [56]; lower abundance in blood of patients with cirrhosis [136]; increased in HIV infection [137]
Order Pseudomonadales			Zon (−)		Associated with normal calprotectin level in Gambian infants [138]; belongs to Gammaproteobacteria negatively correlates with calprotectin [25]; associated with neutrophilia and lower vaccine responses in infants [139]; pathogenic taxa [140]
Family Moraxellaceae			Zon (−)		Dominant phylum in human breast milk [141]; overrepresented in Crohn's disease closely associated with microaerobic energy metabolism, amino acid degradation, and energy deficiency characterized by low ATP levels [142]; lowered in colorectal cancer gut microbiome [143]; belongs to class Gammaproteobacteria which is negatively associated with calprotectin concentration [25]

[a] co-occurrence between change in the abundance and change in zonulin (P2–P4/P5/P6, P3–P4/P5/P6) or calprotectin (P2–P5/P6, P3–P5/P6, P4–P6) levels: same—a change in the bacterial abundance and change in marker level between time points occur in the same direction, opposite—a change in the bacterial abundance and change in marker level between time points occur in the opposite direction

RMC (repeated measures correlation)—correlation between the bacterial abundance and concentration of zonulin/calprotectin for paired measures assessed on six occasions (from P1 to P6, i.e., the entire observation period); LME (linear mixed-effects analysis)—an association between the taxon abundance change and marker change accounting for multiple time point pairs from the same subject, including time-point pair-specific relationships with either absence (the common slope model) or presence of the interaction (the different slopes model) between the abundance change and the time point pair); (+), positive RMC coefficient; (−), negative RMC coefficient

resistance) increase. It is also known that butyric acid regulates cell proliferation and apoptosis, protecting against colorectal cancer [69, 70]. However, the positive associations and correlations between pathways involved in SCFAs production and the zonulin elevation are difficult to explain, suggesting that other mechanisms are also involved in regulating small intestinal barrier permeability in children of up to 2 years old. It should be emphasized that the variations in the concentrations of zonulin and calprotectin overlapped with the introduction of new foods to the child's diet around the sixth month of life (addition of soup, grated apple), with the consequent alteration in the gut microbiota composition and the related metabolic functions. For example, metabolic pathways involved in the metabolism of carbohydrates, vitamin B6 and nucleotides positively correlated with the levels of zonulin, which could suggest that an "open" small intestinal paracellular permeability could play a role in the transportation of metabolic reaction products during intensive growth and development. However,

Table 4 Levels of associations between MetaCyc pathways and markers

Metabolic pathway	Change[a]		RMC	LME	Pathway
	ZON	CALP			
P-163 PWY		Same			L-lysine fermentation to acetate and butanoate
CENTFERM-PWY	Same			Calp (common slope model, negative β coeff.)	Pyruvate fermentation to butanoate
FUCCAT-PWY	Same		Zon (+)		Fucose degradation
GLCMANNANAUT-PWY	Same				Superpathway of N-acetylglucosamine, N-acetylmannosamine and N-acetylneuraminate degradation
PWY-4984	Same				Urea cycle
PWY-6588	Same		Zon (+), Calp (−)		Pyruvate fermentation to acetone
PWY-6590	Same			Calp (common slope model, negative β coeff.)	Superpathway of *Clostridium acetobutylicum* acidogenic fermentation
PWY-6608	Same				Guanosine nucleotides degradation III
PWY-7003	Same		Zon (+)		Glycerol degradation to butanol
PWY-7013	Same				(S)-propane-1,2-diol degradation
PWY-7184	Opposite				Pyrimidine deoxyribonucleotides de novo biosynthesis I
PWY-7237	Same		Zon (+)		myo-, chiro- and scyllo-inositol degradation
FUC-RHAMCAT-PWY			Zon (+)		Superpathway of fucose and rhamnose degradation
PWY-5676			Zon (+)		Acetyl-CoA fermentation to butanoate II
PWY-5677			Zon (+), Calp (−)		Succinate fermentation to butanoate
PWY-6353			Zon (+)		Purine nucleotides degradation II (aerobic)
PWY-6478			Zon (+)		GDP-D-glycero-alpha-D-manno-heptose biosynthesis
PENTOSE-P-PWY				Calp (common slope model, positive β coeff.)	Pentose phosphate pathway
PWY-6545				Zon (different slopes model) negative β coeff.: P3-P4; positive β coeff.: P4–P6, P5–P6	Pyrimidine deoxyribonucleotides de novo biosynthesis III
PWY0-845				Zon (different slopes model) positive β coeff.: P3–P4	Superpathway of pyridoxal 5′-phosphate biosynthesis and salvage
PWY-6901				Zon (different slopes model) positive β coeff.: P3–P4, P5–P6	Superpathway of glucose and xylose degradation
SALVADEHYPOX-PWY			Zon (+)		Adenosine nucleotides degradation II

[a] co-occurrence between change in the pathway abundance and change in zonulin (P2-P4/P5/P6, P3-P4/P5/P6) or calprotectin (P2-P5/P6, P3-P5/P6, P4-P6) levels: same—a change in the pathway abundance and change in marker level between time points occur in the same direction, opposite—a change in the pathway abundance and change in marker level between time points occur in the opposite direction. RMC (repeated measures correlation)—correlation between the pathway abundance and concentration of zonulin/calprotectin for paired measures assessed on six occasions (from P1 to P6, i.e., the entire observation period); LME (linear mixed-effects analysis)—an association between the pathway abundance change and marker change accounting for multiple time point pairs from the same subject, including time-point pair-specific relationships with either absence (the common slope model) or presence of the interaction (the different slopes model) between the abundance change and the time point pair); (+), positive RMC coefficient; (−), negative RMC coefficient

the paracellular transport of nutrients is known to play a rather minor role. The observed results support the concept that paracellular permeability is only the component of intestinal barrier [22] and that SCFAs affect mostly transcellular route via metabolism of enterocytes. Of note, the analysis of the validation data did not confirm our observations concerning the metabolic pathways. Therefore, shotgun sequencing and metabolomic analyses should be performed to shed more light on these processes. All associations between metabolic pathways and markers were shown in Table 4.

Our study is not without limitations. Our cohort was relatively small, though enough to observe changes in zonulin and calprotectin levels as well as in gut microbiota composition during the first 2 years of life with appropriate power. The relatively small size of our dataset (see power calculation section) does mean it is likely we may underestimate the relationship between permeability markers and the microbiota resulting in false negative findings, though we anticipate no increased risk of false positive such. However, the dataset size was at least sufficient to observe changes in zonulin and calprotectin levels as well as in gut microbiota compositions, even at this reduced power. Moreover, our results were validated using the HMS cohort as well as via the comparison between the microbiota of meconium and stool samples collected at other time points. We therefore consider our findings reliable although likely still incomplete. Moreover, our samples were somewhat heterogeneous (with respect to the delivery mode, the use of antibiotics, nutrition, and other environmental factors). Of note, all of these factors and age development can affect the gut microbiota; however, in term infants, the mode of delivery seems to impact fecal calprotectin levels only in the first week of life [25]. Moreover, in our study, we did not use direct methods to measure the gut permeability; we used biomarkers, for which data in healthy children < 2 years old are limited. In fact, there are studies reported on the fecal levels of zonulin in adults but not in children. It is known that in children, the levels of zonulin and calprotectin are incomparably higher than those in adults, changing over time, and could be influenced by e.g., breastfeeding, the dietary patterns, the mode of delivery, and the consumption of gliadin. In the current study, however, due to the small sample size, we did not analyze these biomarkers separately in the context of different delivery modes. Also, we did not collect detailed data on antibiotic exposure. Further, we did not measure these factors in blood (ethical reason—healthy children); only in stool. Moreover, since no metabolomic, lipidomic, immunological, or shotgun metagenomic analyses were performed, the results we obtained must be treated with caution. Mechanistic studies are still needed to investigate the relationship between microbiota and gut permeability and the immune system.

Conclusions

Overall, based on the results, we can conclude that the gut microbiota composition, the small intestinal paracellular permeability, as well as immune system-related markers change dynamically during the first 2 years of life. Although the gut microbiota composition and the related metabolic functions were correlated with zonulin and calprotectin levels, neither clear causation nor obvious health consequences can be proven in this study. However, our data may suggest that the *Ruminococcus* (*torques* group) might be more involved in controlling paracellular permeability during the first 2 years of life than previously considered. Additionally, our data indicate that the genus *Staphylococcus* and the family Staphylococcaceae (positively correlated with the fecal calprotectin concentration) and the family Ruminococcaceae, as well as the order Clostridiales (negatively correlated with the fecal calprotectin concentration), may be potential biomarkers of the development of the immune system development or inflammatory reactions in children younger than 2 years old. It must be emphasized that further longitudinal studies are required to investigate the mechanism behind gut permeability in children and its influence on health. The establishment of gut permeability markers optimal for children as well as germ-free animal models, together with immunological, metabolomic and lipidomic studies (including the analysis of SCFAs), are essential. This study's translational significance may follow from our observation may be that due to the of persistence of increased intercellular permeability, at least until the age of two. If these observations can be robustly replicated, it may indicate health benefits of restricting exposure of very young children the child's exposure to potential antigens, both food and environmental, should be limited, balanced against the relative benefit of such exposure under the "hygiene hypothesis" of autoimmune aetiology, at least in particular subsets of vulnerable children. Bacterial and metabolic pathways associated with permeability markers may provide a basis for the search for pro- and postbiotics, and possibly serving as markers of intestinal permeability and gut immunological status in young children that may help stratify recommendations.

Abbreviations

DV: Dependent Variable; FDR: False Discovery Rate; HMS: Hannover Medical School; IBD: Inflammatory Bowel Diseases; IBS-D: Irritable Bowel Syndrome (Diarrhea); LME: Linear Mixed-Effects Analysis; LRT: Likelihood Ratio Test; NGS: Next Generation Sequencing; OUT: Operational Taxonomic Unit; P: Time Point; PICRUSt: Phylogenetic Investigation Of Communities By Reconstruction Of Unobserved States; PMU: Pomeranian Medical University; RMC: Repeated

Measures Correlation; RTK: Rarefaction Toolkit; SAS: Stool Sample Application System; SCFAs: Short Chain Fatty Acids; TJ: Tight Junctions; ZON: Zonulin; ZOT: Zonula Ocludens Toxin.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12967-021-02839-w.

Additional file 1: Figure S1. Sample selection and availability (PMU cohort). From 100 healthy, full-term newborns during the period from March 2015 to April 2016, 18 mother + child pairs were initially selected with the highest number of samples available. Since only in two cases the delivery was natural and neither the mother nor the child was treated with antibiotics, it was decided to supplement this cohort with six pairs of mother + child who were not given antibiotics and the delivery was natural. Out of the 24 newborns that were selected in this way, three newborns (F19, F22, F24) were excluded due to inadequate number of samples. Twenty-one newborns (101 samples in total, green tiles) were included, in whom at least four longitudinal stool samples were available. **Figure S2**. Study flow chart, including zonulin and calprotectin (PMU cohort). The number of samples for downstream analyses might differ due to results out of determination limits (zonulin 800 ng/mL and calprotectin 2100 ug/mL), technical problems (small volume of collected stool specimen, inadequate amount of DNA, sequencing depth), participant attrition. **Figure S3.** Stool calprotectin level by time in the HMS cohort. Likelihood ratio test, df = 3, P = 2.56e−10, adjusted for mode of delivery Notched boxplot with variable widths proportional to the square-roots of the number of observations in the groups; FDR adjusted p-values < 0.05 are shown, P2—10th day, P3—1st month, P4—6th month, P5—12th month. **Figure S4.** Alpha and beta diversity over time in the HMS cohort. **A**—Shannon alpha diversity by time, LRT, df = 4, P < 2.2e−16, adjusted for mode of delivery, notched boxplot with variable widths proportional to the square-roots of the number of observations in the groups, **B**—Principal coordinate analysis plot with Bray–Curtis dissimilarity calculated from genus abundances, ellipses were drawn assuming a multivariate t-distribution, **C**—PCo1 scores by time, LRT, df = 4, P = 2.97e−14, adjusted for mode of delivery; **D**—PCo2 scores by time, LRT, df = 4, P = 7.31e−11; FDR adjusted p values < 0.05 are shown P1—1st day (meconium/the first stool), P2—10th day, P3—1st month, P4—6th month, P5—12th month. **Figure S5**. Gut microbiota composition change over time (PMU cohort). A linear mixed effects analysis followed by pairwise comparison of time points (adjusted for mode of delivery and breastfeeding time). The overall p-value—a likelihood ratio test (LRT) of nested models (FDR adjusted across genera); PXPY—contrast p values between the two time points (PX and PY), FDR adjusted for all possible contrasts, t.ratio—t statistics for the contrasts estimates (a positive value, colored blue, indicates a decrease abundance, a negative value, colored red, indicates increase in abundance). Taxa abundances (unrarefied) were transformed by generating 128 Monte Carlo instances of the Dirichlet distribution for each gut sample, followed by center-logtransform of each instance. A linear mixed effects analysis was performed for each instance separately and the results were averaged over 128 instances. P2—7th day, P3—1st month, P4—6th month, P5—12th month, P6—24th month. **Figure S6.** Gut microbiota composition change over time (HMS cohort). A linear mixed effects analysis followed by pairwise comparison of time points (adjusted for mode of delivery) for five taxonomic ranks (present in at least 10% samples), only taxons significantly associated with time are shown. The overall p-value—a likelihood ratio test (LRT) of nested models (FDR adjusted across taxons); PXPY—contrast p-values between the two time points (PX and PY), FDR adjusted for all possible contrasts, t.ratio—t statistics for the contrasts estimates (a positive value, colored blue, indicates a decrease in abundance, a negative value, colored red, indicates an increase in abundance). Taxa abundances (unrarefied) were transformed by generating 128 Monte Carlo instances of the Dirichlet distribution for each gut sample, followed by center-log transform of each instance. A linear mixed effects analysis was performed for each instance separately and the results were averaged over 128 instances. P2—10th day, P3—1st month, P4—6th month, P5—12th month. **Figure S7.** Predicted MetaCyc pathways that change significantly over time in both cohorts (PMU and HMS). Linear mixed effects analysis followed by pairwise comparison of time points (adjusted for mode of delivery and breastfeeding time (PMU) and mode of delivery only in the HMS cohort). The overall p-value—a likelihood ratio test (LRT) of nested models (FDR adjusted across pathways); PXPY—contrast p values between the two time points (X and Y), FDR adjusted for all possible contrasts, t.ratio—t statistics for the contrasts estimates. Pathway abundances (unrarefied) were transformed by generating 128 Monte Carlo instances of the Dirichlet distribution for each gut sample, followed by center-log transform of each instance. A linear mixed effects analysis was performed for each instance separately and the results were averaged over 128 instances. P2—7th day (PMU) or 10th day (HMS), P3—1st month, P4—6th month, P5—12th month, P6—24th month (PMU only). **Figure S8**. Taxon change versus zonulin change (PMU cohort). Linear mixed effects models were used to test for an association between the taxon abundance change and zonulin change accounting for ten time point pairs (P2-P3, P3-P4, etc.) from the same subject, adjusted for mode of delivery and breastfeeding time. Two models were considered: without interaction (w/o int.) and with interaction between time point pair and taxon change (w/ int.). If the interaction term was significant—Q (w/ int.) < 0.05, individual p values (FDR adjusted) were interpreted. If the common slope model was chosen—Q (w/o int.) < 0.05 and Q (w /int) > 0.05, the same common slope coefficient β (coef.) across all time points and one p-value (P2–P3) were shown. Coefficients and Q values were averaged over 128 Monte Carlo instances of the Dirichlet distribution, followed by center-log transform of each instance. **Figure S9.** Taxon change versus calprotectin change (PMU cohort). Linear mixed effects models were used to test for an association between the taxon abundance change and calprotectin change accounting for ten time point pairs (P2-P3, P3-P4, etc.) from the same subject, adjusted for mode of delivery and breastfeeding time. Two models were considered: without interaction (w/o int.) and with interaction between time point pair and taxon change (w/ int.). If the interaction term was significant—Q (w/ int.) < 0.05, individual p values (FDR adjusted) were interpreted. If the common slope model was chosen—Q (w/o int.) < 0.05 and Q (w /int) > 0.05, the same common slope coefficient β (coef.) across all time points and one p-value (P2-P3) were shown. Coefficients and Q values were averaged over 128 Monte Carlo instances of the Dirichlet distribution, followed by center-log transform of each instance. **Figure S10.** Taxon change versus calprotectin change (HMS cohort). Linear mixed effects models were used to test for an association between the taxon abundance change and calprotectin change accounting for six time point pairs (P2–P3, P3–P4, etc.) from the same subject, adjusted for mode of delivery. Two models were considered: without interaction (w/o int.) and with interaction between time point pair and taxon change (w/ int.). If the interaction term was significant—Q (w/ int.) < 0.05, individual p values (FDR adjusted) were interpreted. If the common slope model was chosen—Q (w/o int.) < 0.05 and Q (w /int) > 0.05, the same common slope coefficient β (coef.) across all time points and one p-value (P2-P3) were shown. Coefficients and Q values were averaged over 128 Monte Carlo instances of the Dirichlet distribution, followed by center-log transform of each instance. **Figure S11.** Pathway change vs calprotectin change (HMS cohort). Linear mixed effects models were used to test for an association between the pathway abundance change and calprotectin change accounting for six time point pairs (P2–P3, P3–P4, etc.) from the same subject, adjusted for mode of delivery. Two models were considered: without interaction (w/o int.) and with interaction between time point pair and taxon change (w/ int.). If the interaction term was significant—Q (w/int.) < 0.05, individual p values (FDR adjusted) were interpreted. If the common slope model was chosen—Q (w/o int.) < 0.05 and Q (w /int) > 0.05, the same common slope coefficient β (coef.) across all time points and one p-value (P2-P3) were shown. Coefficients and Q values were averaged over 128 Monte Carlo instances of the Dirichlet distribution, followed by center-log transform of each instance. **Figure S12.** Relationship between (changes in) gut abundance of inferred MetaCyc pathways and (changes in) calprotectin levels (HMS cohort).**A**—common slope model illustrating relationships between CENTFERM-PWY change (centered log-ratio transformed) and calprotectin change (rank transformed), **B**—common slope model

illustrating relationships between PWY-6590 change (centered log-ratio transformed) and calprotectin change (rank transformed); Likelihood ratio test (LRT) p- and Q-values were computed based on 128 Monte Carlo instances of the Dirichlet distribution, followed by center-log transform of each instance, the plots were based on the first Monte Carlo instance. **Table S1.** Repeated measures correlation of zonulin and calprotectin with microorganisms and predicted MetaCyc pathways (PMU cohort). **Table S2.** Repeated measures correlation of calprotectin with bacteria and MetaCyc pathway abundance in the PMU and HMS cohorts. **Table S3.** Pathway change vs Zonulin change (PMU cohort). **Table S4.** Pathway change vs Calprotectin change (PMU cohort).

Acknowledgements
The authors would like to thank Professor Dorothee Viemann for constructive feedback of the manuscript.

Authors' contributions
Conception and design of study; acquisition of data; analysis and/or interpretation of data: MK, UL, KA, DW, DM, KS-Ż, IŁ, LM, TU, SKF, BŁ. Drafting the manuscript; revising the manuscript critically for important intellectual content: MK, UL, KS-Ż, IŁ, LM, TU, SKF, BŁ. Approval of the version of the manuscript to be published: MK, UL, KS-Ż, IŁ, LM, TU, SKF, BŁ. All authors read and approved the final manuscript.

Funding
The research was funded by the Pomeranian Medical University in Szczecin, Poland; Project No. WLS 237–03/A/14.

Availability of data and materials
The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request. Data for validation were downloaded from www.ncbi.nlm.nih.gov/sra (BioProject accession number PRJNA514340).

Declarations

Ethics approval and consent to participate
This study was approved by the Bioethics Committee of the Pomeranian Medical University (Resolution No. KB-0012/55/14; 30.06.2014) and was conducted in accordance with the Declaration of Helsinki (2013).

Consent for publication
Written informed consent was obtained from the participants' guardians (parents). All authors gave their conent to publish the study.

Competing interests
Igor Łoniewski is a probiotic company shareholder. Karolina Skonieczna-Żydecka and Mariusz Kaczmarczyk—receive remuneration from a probiotic company. The other authors declare no conflict of interest.

Author details
[1]Department of Clinical Biochemistry, Pomeranian Medical University in Szczecin, 70-111 Szczecin, Poland. [2]Experimental and Clinical Research Center, A Cooperation of Charité - Universitätsmedizin Berlin and Max Delbrück Center for Molecular Medicine, 13125 Berlin, Germany. [3]Charité-Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität Zu Berlin, and Berlin Institute of Health, 14195 Berlin, Germany. [4]Max Delbrück Center for Molecular Medicine in the Helmholtz Association, 13125 Berlin, Germany. [5]DZHK (German Centre for Cardiovascular Research), partner site Berlin, Berlin, Germany. [6]Department of Neonatal Diseases, Pomeranian Medical University in Szczecin, 70-111 Szczecin, Poland. [7]Department of Biochemical Sciences, Pomeranian Medical University in Szczecin, 71-460 Szczecin, Poland. [8]Department of Pharmacology, Pomeranian Medical University in Szczecin, 70-111 Szczecin, Poland. [9]Department of Human Nutrition and Metabolomics, Broniewskiego 24, 71-460 Szczecin, Poland. [10]Berlin Institute of Health (BIH), 10178 Berlin, Germany. [11]Systems Medicine, German Center for Neurodegenerative Diseases (DZNE), 53127 Bonn, Germany. [12]PRECISE Platform for Single Cell Genomics and Epigenomics at the German Center for Neurodegenerative Diseases and the University of Bonn, 53127 Bonn, Germany. [13]European Molecular Biology Laboratory, Structural and Computational Biology Unit, 69117 Heidelberg, Germany.

Received: 23 January 2021 Accepted: 16 April 2021
Published online: 28 April 2021

References

1. Kerr CA, Grice DM, Tran CD, Bauer DC, Li D, Hendry P, et al. Early life events influence whole-of-life metabolic health via gut microflora and gut permeability. Crit Rev Microbiol. 2015;41:326–40.
2. Shulman RJ, Schanler RJ, Lau C, Heitkemper M, Ou C-N, Smith EO. Early feeding, antenatal glucocorticoids, and human milk decrease intestinal permeability in preterm infants. Pediatr Res. 1998;44:519–23.
3. van Elburg RM, Fetter WPF, Bunkers CM, Heymans HSA. Intestinal permeability in relation to birth weight and gestational and postnatal age. Arch Dis Child Fetal Neonatal Ed. 2003;88:F52-55.
4. Insoft RM, Sanderson IR, Walker WA. Development of immune function in the intestine and its role in neonatal diseases. Pediatr Clin North Am. 1996;43:551–71.
5. Drozdowski LA, Clandinin T, Thomson ABR. Ontogeny, growth and development of the small intestine: understanding pediatric gastroenterology. World J Gastroenterol. 2010;16:787–99.
6. Weström B, Arévalo Sureda E, Pierzynowska K, Pierzynowski SG, Pérez-Cano F-J. The immature Gut barrier and its importance in establishing immunity in newborn mammals. Front Immunol. 2020. https://doi.org/10.3389/fimmu.2020.01153/full.
7. Sharma R, Young C, Neu J. Molecular modulation of intestinal epithelial barrier: contribution of microbiota. J Biomed Biotechnol. 2010;2010:305879.
8. Khan MR, Faubion WA, Dyer R, Singh R, Larson JJ, Absah I. Role of lactulose rhamnose permeability test in assessing small bowel mucosal damage in children with celiac disease. Glob Pediatr Health. 2020 (cited 2020 Dec 3);7. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7672748/
9. Wegh CAM, de Roos NM, Hovenier R, Meijerink J, van der Vaart BI, van Hemert S, et al. Intestinal permeability measured by urinary sucrose excretion correlates with serum zonulin and faecal calprotectin concentrations in UC patients in remission. J Nutr Metab. 2019;2019:2472754.
10. Fasano A, Baudry B, Pumplin DW, Wasserman SS, Tall BD, Ketley JM, et al. Vibrio cholerae produces a second enterotoxin, which affects intestinal tight junctions. Proc Natl Acad Sci U S A. 1991;88:5242–6.
11. Baudry B, Fasano A, Ketley J, Kaper JB. Cloning of a gene (zot) encoding a new toxin produced by Vibrio cholerae. Infect Immun. 1992;60:428–34.
12. Tsukita S. Furuse M [Identification of two distinct types of four-transmembrane domain proteins, occludin and claudins: towards new physiology in paracellular pathway]. Seikagaku. 2000;72:155–62.
13. Fink MP. Intestinal epithelial hyperpermeability: update on the pathogenesis of gut mucosal barrier dysfunction in critical illness. Curr Opin Crit Care. 2003;9:143–51.
14. Sapone A, de Magistris L, Pietzak M, Clemente MG, Tripathi A, Cucca F, et al. Zonulin upregulation is associated with increased gut permeability in subjects with type 1 diabetes and their relatives. Diabetes. 2006;55:1443–9.
15. Vanuytsel T, Vermeire S, Cleynen I. The role of haptoglobin and its related protein, zonulin, in inflammatory bowel disease. Tissue Barriers. 2013;1:e27321.
16. Marlicz W, Yung DE, Skonieczna-Żydecka K, Loniewski I, van Hemert S, Loniewska B, et al. From clinical uncertainties to precision medicine: the emerging role of the gut barrier and microbiome in small bowel functional diseases. Expert Rev Gastroenterol Hepatol. 2017;11:961–78.
17. Rougé C, Butel M-J, Piloquet H, Ferraris L, Legrand A, Vodovar M, et al. Fecal calprotectin excretion in preterm infants during the neonatal period. PLOS ONE. 2010;5:e11083.
18. Watts T, Berti I, Sapone A, Gerarduzzi T, Not T, Zielke R, et al. Role of the intestinal tight junction modulator zonulin in the pathogenesis of type I diabetes in BB diabetic-prone rats. Proc Natl Acad Sci USA. 2005;102:2916–21.

19. Drago S, El Asmar R, Di Pierro M, Grazia Clemente M, Tripathi A, Sapone A, et al. Gliadin, zonulin and gut permeability: effects on celiac and non-celiac intestinal mucosa and intestinal cell lines. Scand J Gastroenterol. 2006;41:408–19.
20. Moreno-Navarrete JM, Sabater M, Ortega F, Ricart W, Fernández-Real JM. Circulating zonulin, a marker of intestinal permeability, is increased in association with obesity-associated insulin resistance. PLoS ONE. 2012;7:e37160.
21. Jayashree B, Bibin YS, Prabhu D, Shanthirani CS, Gokulakrishnan K, Lakshmi BS, et al. Increased circulatory levels of lipopolysaccharide (LPS) and zonulin signify novel biomarkers of proinflammation in patients with type 2 diabetes. Mol Cell Biochem. 2014;388:203–10.
22. Hollander D, Kaunitz JD. The "Leaky Gut": tight junctions but loose associations? Dig Dis Sci. 2020;65:1277–87.
23. Oord T, Hornung N. Fecal calprotectin in healthy children. Scand J Clin Lab Invest. 2014;74:254–8.
24. Beşer ÖF, Sancak S, Erkan T, Kutlu T, Çokuğraş H, Çokuğraş FÇ. Can fecal calprotectin level be used as a markers of inflammation in the diagnosis and follow-up of cow's milk protein allergy? Allergy Asthma Immunol Res. 2014;6:33–8.
25. Willers M, Ulas T, Völlger L, Vogl T, Heinemann AS, Pirr S, et al. S100A8 and S100A9 are Important for Postnatal Development of Gut Microbiota and Immune System in Mice and Infants. Gastroenterology. 2020;S0016508520350587.
26. Orivuori L, Mustonen K, de Goffau MC, Hakala S, Paasela M, Roduit C, et al. High level of fecal calprotectin at age 2 months as a marker of intestinal inflammation predicts atopic dermatitis and asthma by age 6. Clin Exp Allergy. 2015;45:928–39.
27. Berstad A, Arslan G, Folvik G. Relationship between intestinal permeability and calprotectin concentration in gut lavage fluid. Scand J Gastroenterol. 2000;35:64–9.
28. Łoniewska B, Węgrzyn D, Adamek K, Kaczmarczyk M, Skonieczna-Żydecka K, Adler G, et al. The influence of maternal-foetal parameters on concentrations of zonulin and calprotectin in the blood and stool of healthy newborns during the first seven days of life. An observational prospective cohort study. J Clin Med. 2019 (cited 2020 Aug 21);8. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6517987/.
29. Łoniewska B, Adamek K, Węgrzyn D, Kaczmarczyk M, Skonieczna-Żydecka K, Clark J, et al. Analysis of faecal zonulin and calprotectin concentrations in healthy children during the first two years of life. An observational prospective cohort study. J Clin Med. 2020 (cited 2020 Aug 21);9. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7141325/.
30. Hildebrand F, Tadeo R, Voigt AY, Bork P, Raes J. LotuS: an efficient and user-friendly OTU processing pipeline. Microbiome. 2014;2:30.
31. Saary P, Forslund K, Bork P, Hildebrand F. RTK: efficient rarefaction analysis of large datasets. Bioinformatics. 2017;33:2594–5.
32. Douglas GM, Maffei VJ, Zaneveld J, Yurgel SN, Brown JR, Taylor CM, et al. PICRUSt2: an improved and extensible approach for metagenome inference. bioRxiv. 2019;672295.
33. Fernandes AD, Macklaim JM, Linn TG, Reid G, Gloor GB. ANOVA-like differential expression (ALDEx) analysis for mixed population RNA-Seq. PLoS ONE. 2013;8:e67019.
34. Lenth R, Buerkner P, Herve M, Love J, Riebl H, Singmann H. emmeans: estimated marginal means, aka least-squares means. 2020 (cited 2020 Dec 3). https://CRAN.R-project.org/package=emmeans.
35. Bakdash JZ, Marusich LR. Repeated measures correlation Front Psychol. 2017. https://doi.org/10.3389/fpsyg.2017.00456/full.
36. Shen X, Wang M, Zhang X, He M, Li M, Cheng G, et al. Dynamic construction of gut microbiota may influence allergic diseases of infants in Southwest China. BMC Microbiol. 2019;19:123.
37. Tarko A, Suchojad A, Michalec M, Majcherczyk M, Brzozowska A, Maruniak-Chudek I. Zonulin. A potential marker of intestine injury in newborns. Dis Markers. 2017 (cited 2020 Aug 21);2017. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5523403/.
38. Saleem B, Okogbule-Wonodi AC, Fasano A, Magder LS, Ravel J, Kapoor S, et al. Intestinal barrier maturation in very low birthweight infants: relationship to feeding and antibiotic exposure. J Pediatr. 2017;183(31–36):e1.
39. Tripathi A, Lammers KM, Goldblum S, Shea-Donohue T, Netzel-Arnett S, Buzza MS, et al. Identification of human zonulin, a physiological modulator of tight junctions, as prehaptoglobin-2. Proc Natl Acad Sci USA. 2009;106:16799–804.
40. Fasano A. Regulation of intercellular tight junctions by zonula occludens toxin and its eukaryotic analogue zonulin. Ann N Y Acad Sci. 2000;915:214–22.
41. Wang W, Uzzau S, Goldblum SE, Fasano A. Human zonulin, a potential modulator of intestinal tight junctions. J Cell Sci. 2000;113(Pt 24):4435–40.
42. Turner JR. Intestinal mucosal barrier function in health and disease. Nat Rev Immunol. 2009;9:799–809.
43. Shen L, Weber CR, Raleigh DR, Yu D, Turner JR. Tight junction pore and leak pathways: a dynamic duo. Annu Rev Physiol. 2011;73:283–309.
44. El Asmar R, Panigrahi P, Bamford P, Berti I, Not T, Coppa GV, et al. Host-dependent zonulin secretion causes the impairment of the small intestine barrier function after bacterial exposure. Gastroenterology. 2002;123:1607–15.
45. Fasano A, Uzzau S, Fiore C, Margaretten K. The enterotoxic effect of zonula occludens toxin on rabbit small intestine involves the paracellular pathway. Gastroenterology. 1997;112:839–46.
46. Fasano A. All disease begins in the (leaky) gut: role of zonulin-mediated gut permeability in the pathogenesis of some chronic inflammatory diseases. F1000Res. 2020 (cited 2020 Aug 21);9. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6996528/.
47. Wood H, Acharjee A, Pearce H, Quraishi MN, Powell R, Rossiter A, et al. Breastfeeding promotes early neonatal regulatory T-cell expansion and immune tolerance of non-inherited maternal antigens. Allergy. 2021. https://doi.org/10.1111/all.14736.
48. Dzidic M, Boix-Amorós A, Selma-Royo M, Mira A, Collado MC. Gut microbiota and mucosal immunity in the neonate. Med Sci (Basel). 2018;6:56.
49. Grönlund MM, Lehtonen OP, Eerola E, Kero P. Fecal microflora in healthy infants born by different methods of delivery: permanent changes in intestinal flora after cesarean delivery. J Pediatr Gastroenterol Nutr. 1999;28:19–25.
50. Macpherson AJ, de Agüero MG, Ganal-Vonarburg SC. How nutrition and the maternal microbiota shape the neonatal immune system. Nat Rev Immunol. 2017;17:508–17.
51. Arboleya S, Solís G, Fernández N, de los Reyes-Gavilán CG, Gueimonde M. Facultative to strict anaerobes ratio in the preterm infant microbiota. Gut Microbes. 2012;3:583-8.
52. Fukuda S, Toh H, Hase K, Oshima K, Nakanishi Y, Yoshimura K, et al. Bifidobacteria can protect from enteropathogenic infection through production of acetate. Nature. 2011;469:543–7.
53. Martinez-Medina M, Aldeguer X, Gonzalez-Huix F, Acero D, Garcia-Gil LJ. Abnormal microbiota composition in the ileocolonic mucosa of Crohn's disease patients as revealed by polymerase chain reaction-denaturing gradient gel electrophoresis. Inflamm Bowel Dis. 2006;12:1136–45.
54. Png CW, Lindén SK, Gilshenan KS, Zoetendal EG, McSweeney CS, Sly LI, et al. Mucolytic bacteria with increased prevalence in IBD mucosa augment in vitro utilization of mucin by other bacteria. Am J Gastroenterol. 2010;105:2420–8.
55. Joossens M, Huys G, Cnockaert M, De Preter V, Verbeke K, Rutgeerts P, et al. Dysbiosis of the faecal microbiota in patients with Crohn's disease and their unaffected relatives. Gut. 2011;60:631–7.
56. Gevers D, Kugathasan S, Denson LA, Vázquez-Baeza Y, Van Treuren W, Ren B, et al. The treatment-naive microbiome in new-onset Crohn's disease. Cell Host Microbe. 2014;15:382–92.
57. Takahashi K, Nishida A, Fujimoto T, Fujii M, Shioya M, Imaeda H, et al. Reduced abundance of butyrate-producing bacteria species in the fecal microbial community in Crohn's Disease. Digestion. 2016;93:59–65.
58. Colina AR, Aumont F, Deslauriers N, Belhumeur P, de Repentigny L. Evidence for degradation of gastrointestinal mucin by Candida albicans secretory aspartyl proteinase. Infect Immun. 1996;64:4514–9.
59. Dethlefsen L, Eckburg PB, Bik EM, Relman DA. Assembly of the human intestinal microbiota. Trends Ecol Evol. 2006;21:517–23.
60. Ganesh BP, Klopfleisch R, Loh G, Blaut M. Commensal akkermansia muciniphila exacerbates gut inflammation in *Salmonella typhimurium*-infected gnotobiotic mice. PLOS ONE. 2013;8:e74963.

61. Kaemmerer E, Plum P, Klaus C, Weiskirchen R, Liedtke C, Adolf M, et al. Fatty acid binding receptors in intestinal physiology and pathophysiology. World J Gastrointest Pathophysiol. 2010;1:147–53.
62. Furusawa Y, Obata Y, Fukuda S, Endo TA, Nakato G, Takahashi D, et al. Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. Nature. 2013;504:446–50.
63. Li M, van Esch BCAM, Henricks PAJ, Folkerts G, Garssen J. The anti-inflammatory effects of short chain fatty acids on lipopolysaccharide- or tumor necrosis factor α-stimulated endothelial cells via activation of GPR41/43 and inhibition of HDACs. Front Pharmacol. 2018. https://doi.org/10.3389/fphar.2018.00533/full.
64. Chang PV, Hao L, Offermanns S, Medzhitov R. The microbial metabolite butyrate regulates intestinal macrophage function via histone deacetylase inhibition. PNAS. 2014;111:2247–52.
65. Arpaia N, Campbell C, Fan X, Dikiy S, van der Veeken J, deRoos P, et al. Metabolites produced by commensal bacteria promote peripheral regulatory T-cell generation. Nature. 2013;504:451–5.
66. Smith PM, Howitt MR, Panikov N, Michaud M, Gallini CA, Bohlooly-Y M, et al. The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. Science. 2013;341:569–73.
67. Atarashi K, Tanoue T, Shima T, Imaoka A, Kuwahara T, Momose Y, et al. Induction of colonic regulatory T cells by indigenous *Clostridium* species. Science. 2011;331:337–41.
68. Atarashi K, Tanoue T, Oshima K, Suda W, Nagano Y, Nishikawa H, et al. Treg induction by a rationally selected mixture of *Clostridia strains* from the human microbiota. Nature. 2013;500:232–6.
69. Xiao S, Jiang S, Qian D, Duan J. Modulation of microbially derived short-chain fatty acids on intestinal homeostasis, metabolism, and neuropsychiatric disorder. Appl Microbiol Biotechnol. 2020;104:589–601.
70. Canani RB, Costanzo MD, Leone L, Pedata M, Meli R, Calignano A. Potential beneficial effects of butyrate in intestinal and extraintestinal diseases. World J Gastroenterol. 2011;17:1519–28.
71. Van Goethem MW, Makhalanyane TP, Cowan DA, Valverde A. Cyanobacteria and alphaproteobacteria may facilitate cooperative interactions in niche communities. Front Microbiol. 2017. https://doi.org/10.3389/fmicb.2017.02099/full.
72. Roggero P, Liotto N, Pozzi C, Braga D, Troisi J, Menis C, et al. Analysis of immune, microbiota and metabolome maturation in infants in a clinical trial of Lactobacillus paracasei CBA L74-fermented formula. Nat Commun. 2020;11:2703.
73. Torres J, Hu J, Eisele C, Nair N, Panchal H, Bao X, et al. OP037 Infants born to mothers with inflammatory bowel disease exhibit distinct microbiome features that persist up to 3 months of life. J Crohns Colitis. 2017;11:S23–4.
74. Kaakoush NO, Day AS, Huinao KD, Leach ST, Lemberg DA, Dowd SE, et al. Microbial dysbiosis in pediatric patients with Crohn's Disease. J Clin Microbiol. 2012;50:3258–66.
75. Krishnamoorthy S, Coetzee V, Kruger J, Potgieter H, Buys EM. Dysbiosis signatures of fecal microbiota in south african infants with respiratory, gastrointestinal, and other diseases. J Pediatrics. 2020;218(106–113):e3.
76. Kongnum K, Taweerodjanakarn S, Hongpattarakere T. Longitudinal characterization of bifidobacterial abundance and diversity profile developed in Thai healthy infants. Arch Microbiol. 2020;202:1425–38.
77. Stark PL, Lee A. Clostridia isolated from the feces of infants during the first year of life. J Pediatr. 1982;100:362–5.
78. Stefka AT, Feehley T, Tripathi P, Qiu J, McCoy K, Mazmanian SK, et al. Commensal bacteria protect against food allergen sensitization. Proc Natl Acad Sci U S A. 2014;111:13145–50.
79. Nagano Y, Itoh K, Honda K. The induction of Treg cells by gut-indigenous Clostridium. Curr Opin Immunol. 2012;24:392–7.
80. Kashyap PC, Johnson S, Geno DM, Lekatz HR, Lavey C, Alexander JA, et al. A decreased abundance of clostridia characterizes the gut microbiota in eosinophilic esophagitis. Physiol Rep. 2019;7:e14261.
81. Pichler MJ, Yamada C, Shuoker B, Alvarez-Silva C, Gotoh A, Leth ML, et al. Butyrate producing colonic Clostridiales metabolise human milk oligosaccharides and cross feed on mucin via conserved pathways. Nat Commun. 2020;11:3285.
82. Van den Abbeele P, Belzer C, Goossens M, Kleerebezem M, De Vos WM, Thas O, et al. Butyrate-producing *Clostridium* cluster XIVa species specifically colonize mucins in an in vitro gut model. ISME J. 2013;7:949–61.
83. Ramanan D, Bowcutt R, Lee SC, Tang MS, Kurtz ZD, Ding Y, et al. Helminth infection promotes colonization resistance via type 2 immunity. Science. 2016;352:608–12.
84. Zhou Y, Xu ZZ, He Y, Yang Y, Liu L, Lin Q, et al. Gut microbiota offers universal biomarkers across ethnicity in inflammatory bowel disease diagnosis and infliximab response prediction. mSystems. 2018. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5790872/.
85. Mörkl S, Lackner S, Meinitzer A, Mangge H, Lehofer M, Halwachs B, et al. Gut microbiota, dietary intakes and intestinal permeability reflected by serum zonulin in women. Eur J Nutr. 2018;57:2985–97.
86. Chen K, Chen H, Faas MM, de Haan BJ, Li J, Xiao P, et al. Specific inulin-type fructan fibers protect against autoimmune diabetes by modulating gut immunity, barrier function, and microbiota homeostasis. Mol Nutr Food Res. 2017;61.
87. Biddle A, Stewart L, Blanchard J, Leschine S. Untangling the genetic basis of fibrolytic specialization by lachnospiraceae and ruminococcaceae in diverse gut communities. Diversity. 2013;5:627–40.
88. Lagier J-C, Million M, Hugon P, Armougom F, Raoult D. Human Gut microbiota: repertoire and variations. Front Cell Infect Microbiol. 2012. https://doi.org/10.3389/fcimb.2012.00136/full.
89. Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, et al. Enterotypes of the human gut microbiome. Nature. 2011;473:174–80.
90. Liu Y, Yuan X, Li L, Lin L, Zuo X, Cong Y, et al. Increased ileal immunoglobulin a production and immunoglobulin a-coated bacteria in diarrhea-predominant irritable bowel syndrome. Clin Transl Gastroenterol. 2020 (cited 2020 Nov 1). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7145038/.
91. Kwak MS, Cha JM, Shin HP, Jeon JW, Yoon JY. Development of a novel metagenomic biomarker for prediction of upper gastrointestinal tract involvement in patients with Crohn's Disease. Front Microbiol Front. 2020. https://doi.org/10.3389/fmicb.2020.01162/full.
92. Louis P, Flint HJ. Diversity, metabolism and microbial ecology of butyrate-producing bacteria from the human large intestine. FEMS Microbiol Lett. 2009;294:1–8.
93. Saha S, Jeon BH, Kurade MB, Govindwar SP, Chatterjee PK, Oh SE, et al. Interspecies microbial nexus facilitated methanation of polysaccharidic wastes. Biores Technol. 2019;289:121638.
94. Theriot CM, Bowman AA, Young VB. Antibiotic-induced alterations of the gut microbiota alter secondary bile acid production and allow for clostridium difficile spore germination and outgrowth in the large intestine. mSphere. 2016. https://doi.org/10.1128/mSphere.00045-15.
95. Hiippala K, Jouhten H, Ronkainen A, Hartikainen A, Kainulainen V, Jalanka J, et al. The potential of gut commensals in reinforcing intestinal barrier function and alleviating inflammation. Nutrients. 2018. https://doi.org/10.3390/nu10080988.
96. Haange S-B, Jehmlich N, Hoffmann M, Weber K, Lehmann J, von Bergen M, et al. Disease development is accompanied by changes in bacterial protein abundance and functions in a refined model of Dextran Sulfate Sodium (DSS)-induced colitis. J Proteome Res. 2019;18:1774–86.
97. Youssef O, Lahti L, Kokkola A, Karla T, Tikkanen M, Ehsan H, et al. Stool microbiota composition differs in patients with stomach, colon, and rectal neoplasms. Am J Digestive Dis. 2018;63:2950–8.
98. Ma B, Liang J, Dai M, Wang J, Luo J, Zhang Z, et al. Altered Gut microbiota in chinese children with autism spectrum disorders. Front Cell Infect Microbiol. 2019;9:40.
99. Ocvirk S, Wilson AS, Posma JM, Li JV, Koller KR, Day GM, et al. A prospective cohort analysis of gut microbial co-metabolism in Alaska Native and rural African people at high and low risk of colorectal cancer. Am J Clin Nutr. 2020;111:406–19.
100. Ren SM, Mei L, Huang H, Cao SF, Zhao RH, Zheng PY. Zheng PY [Correlation analysis of gut microbiota and biochemical indexes in patients with non-alcoholic fatty liver disease]. Zhonghua Gan Zang Bing Za Zhi. 2019;27:369–75.
101. Brahe LK, Astrup A, Larsen LH. Is butyrate the link between diet, intestinal microbiota and obesity-related metabolic diseases? Obes Rev. 2013;14:950–9.
102. Azad MB, Konya T, Maughan H, Guttman DS, Field CJ, Chari RS, et al. Gut microbiota of healthy Canadian infants: profiles by mode of delivery and infant diet at 4 months. CMAJ. 2013;185:385–94.
103. Wopereis H, Sim K, Shaw A, Warner JO, Knol J, Kroll JS. Intestinal microbiota in infants at high risk for allergy: effects of prebiotics and role in eczema development. J Allergy Clin Immunol. 2018;141(1334–1342):e5.

104. Azad MB, Konya T, Maughan H, Guttman DS, Field CJ, Sears MR, et al. Infant gut microbiota and the hygiene hypothesis of allergic disease: impact of household pets and siblings on microbiota composition and diversity. Allergy Asthma Clin Immunol. 2013;9:15.
105. Ilinskaya ON, Ulyanova VV, Yarullina DR, Gataullin IG. Secretome of intestinal bacilli: a natural guard against pathologies. Front Microbiol. 2017;8:1666.
106. Jandhyala SM, Talukdar R, Subramanyam C, Vuyyuru H, Sasikala M, Reddy DN. Role of the normal gut microbiota. World J Gastroenterol. 2015;21:8787–803.
107. Rowland I, Gibson G, Heinken A, Scott K, Swann J, Thiele I, et al. Gut microbiota functions: metabolism of nutrients and other food components. Eur J Nutr. 2018;57:1–24.
108. Cong X, Judge M, Xu W, Diallo A, Janton S, Brownell EA, et al. Influence of infant feeding type on gut microbiome development in hospitalized preterm infants. Nurs Res. 2017;66:123–33.
109. Akagawa S, Tsuji S, Onuma C, Akagawa Y, Yamaguchi T, Yamagishi M, et al. Effect of delivery mode and nutrition on gut microbiota in neonates. ANM. 2019;74:132–9.
110. Brooks B, Olm MR, Firek BA, Baker R, Geller-McGrath D, Reimer SR, et al. The developing premature infant gut microbiome is a major factor shaping the microbiome of neonatal intensive care unit rooms. Microbiome. 2018;6:112.
111. Corona-Cervantes K, García-González I, Villalobos-Flores LE, Hernández-Quiroz F, Piña-Escobedo A, Hoyo-Vadillo C, et al. Human milk microbiota associated with early colonization of the neonatal gut in Mexican newborns. PeerJ. 2020;8:e9205.
112. Kim SY, Yi DY. Analysis of the human breast milk microbiome and bacterial extracellular vesicles in healthy mothers. Exp Mol Med. 2020;52:1288–97.
113. Castanet M, Costalos C, Haiden N, Hascoet J-M, Berger B, Sprenger N, et al. Early effect of supplemented infant formulae on intestinal biomarkers and microbiota: a randomized clinical trial. Nutrients. 2020 (cited 2020 Nov 1);12. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7284641/.
114. Yang R, Gao R, Cui S, Zhong H, Zhang X, Chen Y, et al. Dynamic signatures of gut microbiota and influences of delivery and feeding modes during the first 6 months of life. Physiol Genomics. 2019;51:368–78.
115. Sharing of Bacterial Strains Between Breast Milk and Infant Feces—Virginia Martín, Antonio Maldonado-Barragán, Laura Moles, Mercedes Rodriguez-Baños, Rosa del Campo, Leonides Fernández, Juan M. Rodríguez, Esther Jiménez, 2012. (cited 2020 Nov 4). https://doi.org/10.1177/0890334411424729.
116. Turroni F, Serafini F, Mangifesta M, Arioli S, Mora D, van Sinderen D, et al. Expression of sortase-dependent pili of Bifidobacterium bifidum PRL2010 in response to environmental gut conditions. FEMS Microbiol Lett. 2014;357:23–33.
117. Tanaka M, Nakayama J. Development of the gut microbiota in infancy and its impact on health in later life. Allergol Int. 2017;66:515–22.
118. Sivaraj S, Chan A, Pasini E, Chen E, Lawendy B, Verna E, et al. Enteric dysbiosis in liver and kidney transplant recipients: a systematic review. Transpl Int. 2020;33:1163–76.
119. Lo Presti A, Zorzi F, Del Chierico F, Altomare A, Cocca S, Avola A, et al. Fecal and mucosal microbiota profiling in irritable bowel syndrome and inflammatory bowel disease. Front Microbiol. 2019. https://doi.org/10.3389/fmicb.2019.01655/full.
120. Butel MJ, Roland N, Hibert A, Popot F, Favre A, Tessedre AC, et al. Clostridial pathogenicity in experimental necrotising enterocolitis in gnotobiotic quails and protective role of bifidobacteria. J Med Microbiol. 1998;47:391–9.
121. Oswari H, Prayitno L, Dwipoerwantoro PG, Firmansyah A, Makrides M, Lawley B, et al. Comparison of stool microbiota compositions, stool alpha1-antitrypsin and calprotectin concentrations, and diarrhoeal morbidity of Indonesian infants fed breast milk or probiotic/prebiotic-supplemented formula. J Paediatr Child Health. 2013;49:1032–9.
122. Dubin K, Pamer EG. *Enterococci* and their interactions with the intestinal microbiome. Microbiol Spectr. 2014;5.
123. Liu C, Cheng L, Ji L, Li F, Zhan Y, Wu B, et al. Intestinal microbiota dysbiosis play a role in pathogenesis of patients with primary immune thrombocytopenia. Thromb Res. 2020;190:11–9.
124. Yang Y, Cai Q, Shu X-O, Steinwandel MD, Blot WJ, Zheng W, et al. Prospective study of oral microbiome and colorectal cancer risk in low-income and African American populations. Int J Cancer. 2019;144:2381–9.
125. Hopkins MJ, Macfarlane GT, Furrie E, Fite A, Macfarlane S. Characterisation of intestinal bacteria in infant stools using real-time PCR and northern hybridisation analyses. FEMS Microbiol Ecol. 2005;54:77–85.
126. Collado MC, Calabuig M, Sanz Y. Differences between the fecal microbiota of coeliac infants and healthy controls. Curr Issues Intest Microbiol. 2007;8:9–14.
127. Reddel S, Del Chierico F, Quagliariello A, Giancristoforo S, Vernocchi P, Russo A, et al. Gut microbiota profile in children affected by atopic dermatitis and evaluation of intestinal persistence of a probiotic mixture. Sci Rep. 2019;9:4996.
128. Kowalska-Duplaga K, Gosiewski T, Kapusta P, Sroka-Oleksiak A, Wędrychowicz A, Pieczarkowski S, et al. Differences in the intestinal microbiome of healthy children and patients with newly diagnosed Crohn's disease. Sci Rep. 2019;9:18880.
129. Carroll IM, Ringel-Kulka T, Ferrier L, Wu MC, Siddle JP, Bueno L, et al. Fecal protease activity is associated with compositional alterations in the intestinal microbiota. PLoS ONE. 2013;8:e78017.
130. Clavel T, Lepage P, Charrier C. The Family Coriobacteriaceae. In: Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F, editors. Prokaryotes Actinobacteria. 2014. https://doi.org/10.1007/978-3-642-30138-4_343.
131. Rahlwes KC, Sparks IL, Morita YS. Cell walls and membranes of actinobacteria. Subcell Biochem. 2019;92:417–69.
132. Yang I, Corwin EJ, Brennan PA, Jordan S, Murphy JR, Dunlop A. The infant microbiome: implications for infant health and neurocognitive development. Nurs Res. 2016;65:76–88.
133. Lundgren SN, Madan JC, Emond JA, Morrison HG, Christensen BC, Karagas MR, et al. Maternal diet during pregnancy is related with the infant stool microbiome in a delivery mode-dependent manner. Microbiome. 2018;6:109.
134. Moles L, Gómez M, Heilig H, Bustos G, Fuentes S, de Vos W, et al. Bacterial diversity in meconium of preterm neonates and evolution of their fecal microbiota during the first month of life. PLoS ONE. 2013;8:e66986.
135. Li H, Wang J, Wu L, Luo J, Liang X, Xiao B, et al. The impacts of delivery mode on infant's oral microflora. Sci Rep. 2018;8:11938.
136. Kajihara M, Koido S, Kanai T, Ito Z, Matsumoto Y, Takakura K, et al. Characterisation of blood microbiota in patients with liver cirrhosis. Eur J Gastroenterol Hepatol. 2019;31:1577–83.
137. Vujkovic-Cvijin I, Dunham RM, Iwai S, Maher MC, Albright RG, Broadhurst MJ, et al. Dysbiosis of the gut microbiota is associated with HIV disease progression and tryptophan catabolism. Sci Transl Med. 2013;5:19391.
138. Davis JCC, Lewis ZT, Krishnan S, Bernstein RM, Moore SE, Prentice AM, et al. Growth and morbidity of gambian infants are influenced by maternal milk oligosaccharides and infant gut microbiota. Sci Rep. 2017;7:40466.
139. Huda MN, Lewis Z, Kalanetra KM, Rashid M, Ahmad SM, Raqib R, et al. Stool microbiota and vaccine responses of infants. Pediatrics. 2014;134:e362–72.
140. Podany A, Rauchut J, Wu T, Kawasawa YI, Wright J, Lamendella R, et al. Excess dietary zinc intake in neonatal mice causes oxidative stress and alters intestinal host-microbe interactions. Mol Nutr Food Res. 2019;63:1800947.
141. Pannaraj PS, Li F, Cerini C, Bender JM, Yang S, Rollie A, et al. Association between breast milk bacterial communities and establishment and development of the infant gut microbiome. JAMA Pediatr. 2017;171:647–54.
142. Kim S, Thapa I, Zhang L, Ali H. A novel graph theoretical approach for modeling microbiomes and inferring microbial ecological relationships. BMC Genomics. 2019;20:945.
143. Yang J, McDowell A, Kim EK, Seo H, Lee WH, Moon C-M, et al. Development of a colorectal cancer diagnostic model and dietary risk assessment through gut microbiome analysis. Exp Mol Med. 2019;51:1–15.

Publisher's Note