# 1. Introduction

Early life stress, especially if chronic, is related to negative behavioral, cognitive and health outcomes. Entering center-based childcare (CC) at three months of life, the normal age in the Netherlands, is an important change or ‘life event’ for an infant. CC includes stressors such a long maternal separations, new and frequently changing caregivers and peers, and adaptation to a new physical environment. From our own and other studies, we know that entering childcare produces significant increases in cortisol levels as compared to levels in the home environment, and that cortisol continues to increase until at least a month after entering [@Ahnert2004; @Albers2016; @Bernard2015; @Watamura2010; @Waynforth2007]. CC has also been related to more health symptoms and illnesses including diarrhea and communicable diseases such as respiratory illnesses and otitis media [@Beijers2011; @Loeb2007].

A growing body of evidence indicates that the intestinal microbiota might play an important role in how such early life stressors influence future health outcomes and behavior. Firstly, in human studies gut microbiota composition is associated with several pediatric disorders including allergic and autoimmune diseases [@Videhult2016]. For example, @Azad2015 found a lower richness and increased *Enterobacteriaceae/Bacteroidetes* ratio in infant stool to be associated with subsequent food sensitization in the population based Canadian Healthy Infant Longitudinal Development (CHILD) birth cohort study. Secondly, animal studies indicate that stress affects the gut microbiota, which in turn leads to inflammatory immune processes and even behavioral changes. For example, rodent studies showed that early life stress induced by maternal separation led to alternations of the intestinal barrier function and the microbiota [@Rea2016]. These alternations were accompanied by behavioral changes that in one study have been found to be mediated by the intestinal bacteria [@DePalma2015]. Interestingly, the administration of probiotics could prevent or ameliorate stress induced effects [@Rea2016]. In rhesus monkeys, maternal separation also altered the composition of the intestinal microbiota of the offspring by decreasing the number *Lactobacilli* [@Bailey1999]. However, we lack studies that investigated the effect of early life stress on the development of the gut microbiota composition in humans and specifically the effect of CC on the infant gut microbiota composition has not been studied at all.  
Among the better studied factors that influence the development of the infant gut microbiota are diet and specifically breastfeeding. Breastfeeding increases the abundance of Lactobacilli and Bifidobacteria species when compared to formula fed infants [@Gregory2016; @Videhult2016]. @Yamada2017 described how oligosaccharides present in human breast milk selectively stimulate the growth of Bifidobacteria. Finally, there is evidence suggesting that breast milk contains bacteria that colonize the infant gut. For instance, @Pannaraj2017 describe that on average, 27.7% of the infant gut microbiome resembles the microbiome of the mother’s breast milk and 10.3% resembles the microbiome of the mother’s areolar skin. In contrast to early life stress, breastfeeding is known to be a health promoting factor in early life [@Hoddinott2008].  
Altogether, this research suggests that early life stress might impair the development of the intestinal microbiota with possible long term effects on physical health and behavior and that breastfeeding might be a protective factor against the negative influence of stress. In this study we investigated whether the entrance to CC leads to changes in the intestinal microbiota and whether breastfeeding plays a role in this relationship. We obtained stool samples from infants who entered CC (*n*=49) and control infants (*n*=49) before the entrance at 10 weeks of age (T0) and four weeks after the entrance (T1). We hypothesized (1) that we find no differences in the composition of the intestinal microbiota between infants who entered CC and control infants at T0 but that we find differences at at T1. The effect of CC might be moderated by a buffering effect of breastfeeding on microbiota changes. Therefore, (2) we anticipate a smaller change in microbiota composition in the breastfed group compared to the non breastfed group of infants that attended CC.

# 2. Methods

## 2.1 Participants

As part of an ongoing longitudinal study (Bibo study), 220 mothers were followed since the third trimester of pregnancy, to investigate the influences of early caregiving factors on the early development of children. Uncomplicated singleton pregnancy, proficiency in the Dutch language, no drug use and the absence of physical and mental health problems were criteria for initial inclusion. Eight of the 220 women were excluded due to preterm birth or for other medical reasons. In addition, 19 mothers discontinued participation in the study during the first three postpartum months because of personal circumstances. All infants were healthy and born at full term ( 37 weeks). Infants born by cesarean delivery (n=4) and who had used antibiotics (n=4) were excluded. Out of a total of 9 fecal samples that were available per infant, 2 were selected for this study: Ten weeks post-partum before entrance to CC (T0) and 4 weeks after the entrance (T1). If there was no sample obtained at both times points, the infant was excluded. Seven samples were excluded for technical reasons. For these samples, the Pearson correlation of between duplicate bacterial profiles was not reached. The final sample included 49 infants per group. The study has been approved by the ethical Committee of the Faculty of Social Sciences, Radboud University Nijmegen (ECG / AvdK/07.563) and informed consent was obtained from each mother.

## 2.2 Procedure

### 2.2.1 Breastfeeding

Information about breastfeeding was collected through the use of diaries. The mothers received diaries towards the end of their pregnancy with instructions to take weekly notes about breastfeeding and bottle-feeding from week 1 until week 27 after birth. For each week, the average number of breast- and/or bottle- feedings per day were noted. We used the average number of breastfeedings per day until timepoint PRE and the average number of breastfeedings per day between timepoints PRE and POST as confounding variables in the analyses.

### 2.2.2 Faeces collection, DNA isolation and microbiota profiling

The parents were instructed to collect the fecal samples at home and to store them at -20°C. Samples were kept in coolers for transportation and then stored at -20°C and later at -80°C before being processed at the Microbiology Laboratory at Wageningen University. DNA isolation from faecal samples has been described in detail elsewhere [@Salonen2010]. In brief, DNA was isolated using a combination of column purification and Repeated-Bead-Beating. Purity and concentration of DNA were assessed with a Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific, Wilmington, USA). DNA was then stored at -20 C° pending microbial analysis. The analysis was then performed utilizing a previously benchmarked custom made, phylogenetic microarray, the Human Intestinal Tract Chip (HITChip) [@Rajilic-Stojanovic2009; @Claesson2009a]. The HITchip contains a duplicated set of 3,631 probes, which target the 16S rRNA gene sequences of more than 1,000 intestinal bacterial phylotypes [@Rajilic-Stojanovic2009]. After extraction of DNA, the full-length 16S rRNA gene was amplified, which was followed by in vitro transcription and labelling of the resulting RNA with Cy3/Cy5 before hybridization to the array. Each sample was hybridized at least twice to ensure reproducibility. Duplicate hybridizations with a Pearson correlation <.98 were not considered for further analysis. By pre-processing the probe-level measurements with minimum–maximum normalization and Robust Probabilistic Averaging probe summarization [@Lahti2011; @Lahti2013] the microbiota profiles were generated into three phylogenetic levels: Order-like 16S rRNA gene sequence groups (level 1); genus-like 16S rRNA gene sequence groups with a sequence similarity >90% (level 2); and phylotype-like 16S rRNA gene sequence groups with a sequence similarity >98% (level 3). For our study, we focus on the genus-level variation referred to as species and relatives (‘et rel.’) and use log10-transformed signals as a proxy for bacterial abundance.

### 2.2.3 Additional variables

Maternal age as well as infant birthweight, age in weeks at the entrance to CC and sex were obtained as potential confounding variables. Table 1 shows descriptive statistics for the variables we accounted for.

## 2.3 Statistical analysis

All analyses were performed in R version 3.5 [@R-base].

### 2.3.1 Data preparation and visualization

The continuous covariates “age” and “breastfeeding” were standardized to ease interpretability and convergence of hierarchical linear models. There were missing values for 5 infants for the variables “breastfeeding” and “sibling” each. The missing values were imputed using predictive mean matching as implemented in the package *mice* [@buurenFlexibleImputationMissing2019], which resulted in 10 imputed data sets. Probes were counted in each sample to measure richness, by using an 80% quantile threshold for detection. We calculated alpha diversity (Shannon) and applied centered-log-ratio (clr) transformation [@aitchisonStatisticalAnalysisCompositional1986] to HITchip signals using the *microbiome* package [@R-microbiome]. Euclidian distance of clr-transformed data results in Aitchiison distance [@aitchisonStatisticalAnalysisCompositional1986]. Therefore, in order to visualize the effect of CC on overall composition over time, we could apply principal component analyses to clr-transformed HITchip signals. The resulting biplots are shown in figure 1.

### 2.3.2 Statistical models

To determine whether and how much entrance to CC affects microbiota overall community composition, we performed a two way permutational multivariate analysis of variance (PERMANOVA) based on Aitchison distance. The PERMANOVA was performed using the function *adonis* from the package *vegan* [@R-vegan] with CC (CC versus HOME) interacting with time (pre versus post) as predictors controlling for breastfeeding and age. We performed differential abundance testing using Bayesian hierarchical generalized linear models with the generalized normal distribution as implemented in the *BRMS* package [@R-brms\_b]. Compared to the Gaussian distribution, the generalized Gaussian distribution has an additional parameter , which allows to model skewed distributions. Therefore, it can better deal with the distributional properties of clr-transformed bacterial abundances. Shrinkage was applied to the intercept and the slopes of the covariates and the predictor “Time”. Regularizing priors were chosen that make the model skeptical of an effect of childcare over time. For each genus, the same model specification was used. If a model did not converge, a stronger regularizing prior was required for the standard deviation of the varying intercept (see supplementary material for exact model specification). We used the Gaussian distribution to regress alpha-diversity on the same set of predictors. We could pass the multiple imputed objects created by *mice* directly to the fitting function in the package *brms*. Finally, we used the Random Forests (RF) algorithm implemented in the package *ranger* [@R-ranger] to evaluate whether we can predict childcare entrance based on 130 genus abundances for both time points. RF is a tree based ensemble learning method that is well suited for classification based on microbial abundances of samples [@knightsSupervisedClassificationHuman2011]. We randomly selected 80% of the collected samples that constituted the training data set. We first tuned the RF models based on out-of-bag error. Node splitting was based on the gini criterion. Then we evaluated whether we can correctly classify CC based on 130 genus abundances using the hold out set.

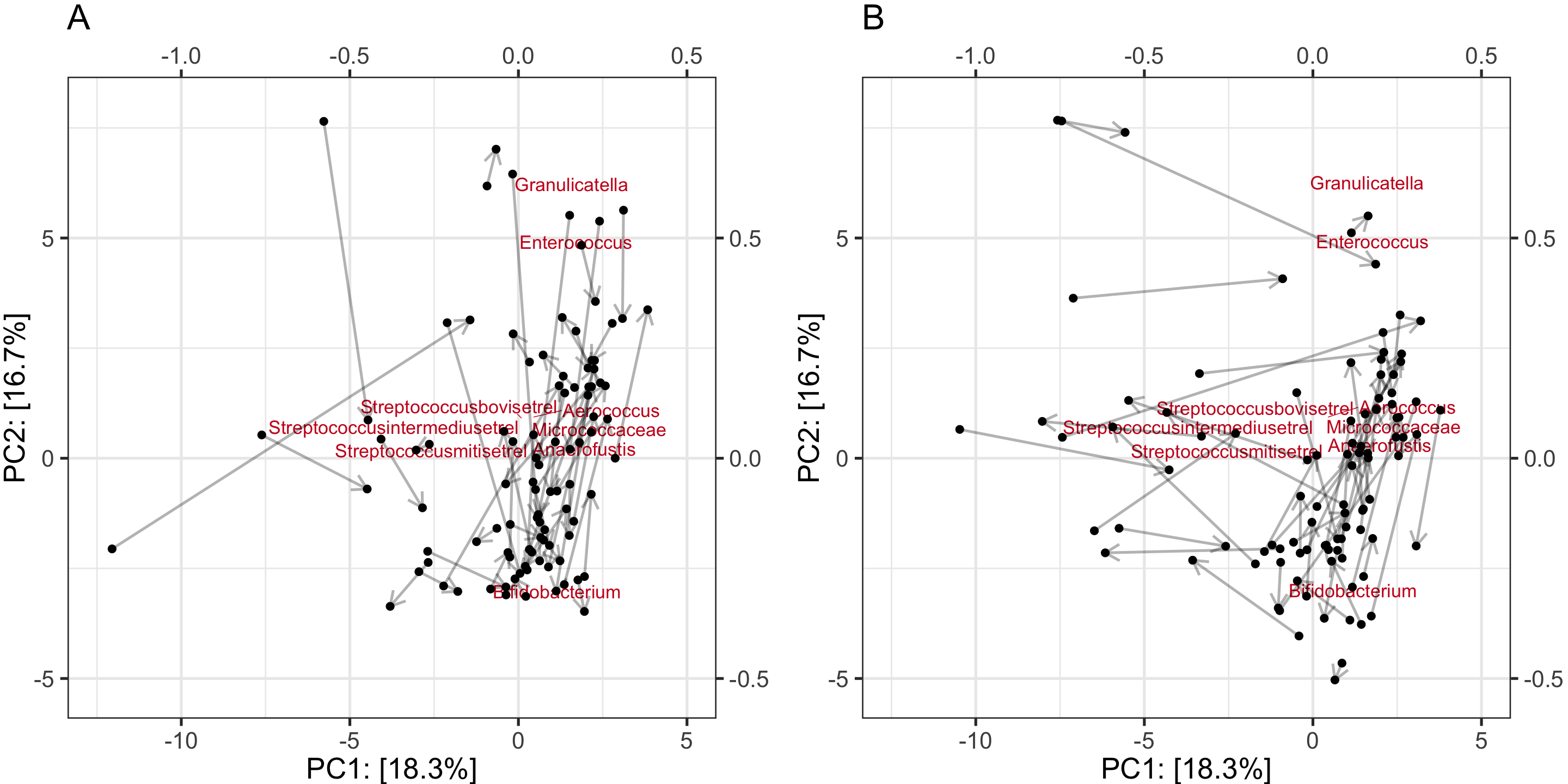
# 3. Results

# 3. Results

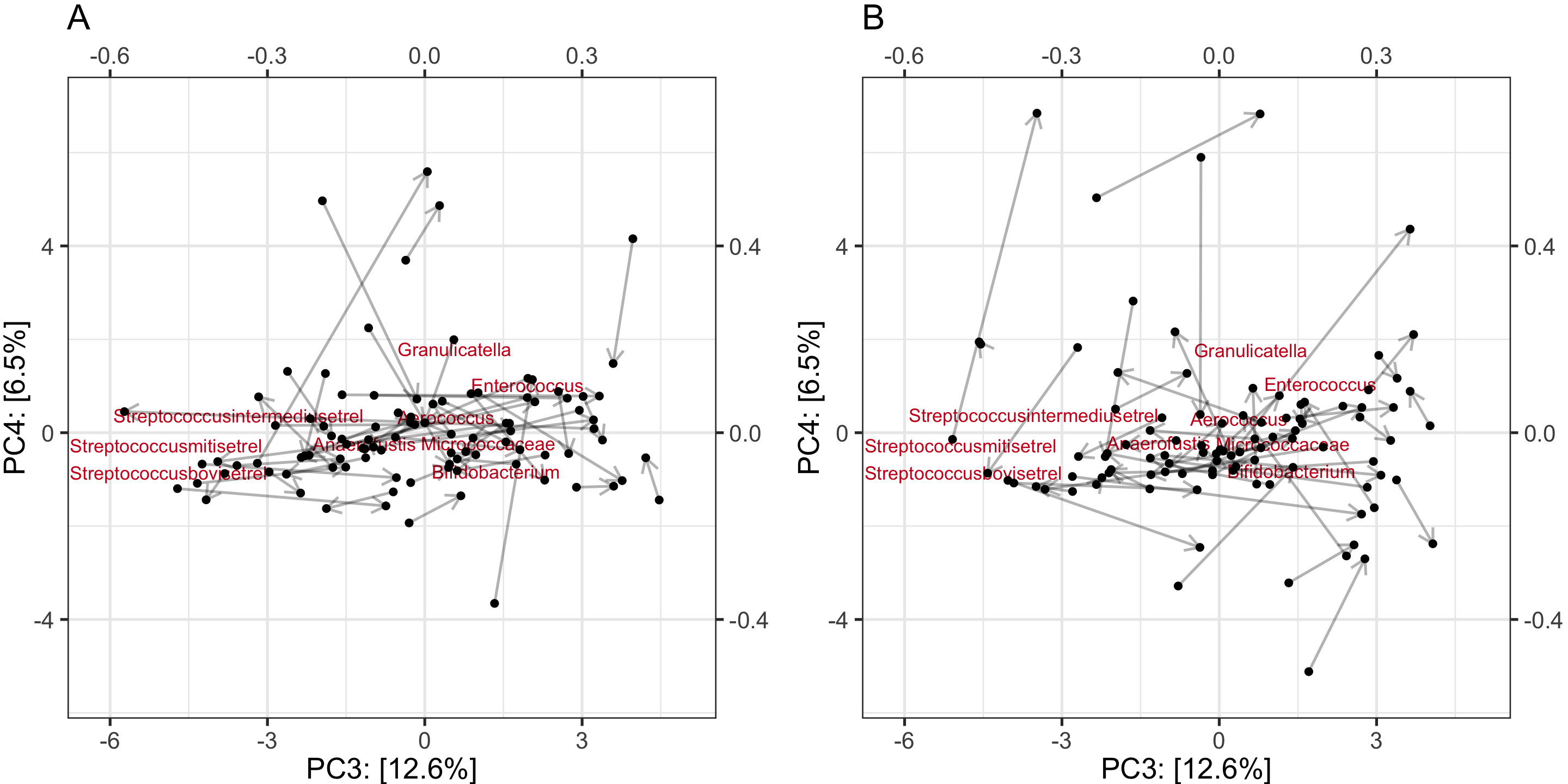
## 3.1. Microbiota composition

Thirteen genus like groups from the *Actinobacteria*, *Firmicutes*, *Proteobacteria* and phyla showed an average abundance of % in at least 20% of the samples (Figure 1). Overall the microbiota was dominated by *Bifidobacterium* with an average abundance of more than 50%. Followed by facultative anaerobic *Bacilli* (*Streptococcus spp*, *Enterococcus*, *Lactobacilllus* and *Granulicatella*). The general variation of relative abundance of all taxa was quite high, with *Bifidobacterium* for instance ranging from 0.2% to 89%.

Figures 2 and 3 show microbiota composition (Aitchison distance) for the first four principal components within the CC (A) and within the HOME (B) group. The starting points of the arrows indicate the microbiota composition in space at time-poin PRE, whereas the endpoint corresponds the composition at time-poin POST. There appear to be no differences in location between CC and HOME at PRE or POST. Also, we did not identify a clearly uniform direction of the shifts over time in either group. Figure 2 shows that within the CC group, a few infants score relatively high on PC2 and all but one move towards the center indicating shifts in the relative abundance of *Granulicatella* and *Enterococcus*, which load high on PC2. Altogether, microbiota composition development between time points appears to be highly individual and unsystematic.



First two principal components illustrating development of microbiota composition over time within CC (A) and HOME (B).



Third and fourth principal component illustrating development of microbiota composition over time within CC (A) and HOME (B).

## 3.2 Effects of CC on microbiota composition

Table 1 shows demographic variables for both groups. There was a significant difference in age between groups at PRE (*t*(62.43) = -4.54, *p* < .001) and at POST (*t*(60.82) = -4.88, *p* < .001) according to Welch’s t-test. Besides that, there were no differences between groups for any of the remaining variables. To account for the differences in age and the naturally varying amounts of breastfeeding, we used both variables as covariates in all linear models.

Cell Contents |————————-| | Count | | Expected Values | | Std Residual | |————————-|

Total Observations in Table: 96

| df\_complete$csection

|  |  |  |  |
| --- | --- | --- | --- |
| df\_complete$CC | 0 | 1 | Row Total |
| 0 | 45 | 3 | 48 |
| 43.500 | 4.500 |  |  |
| 0.227 | -0.707 |  |  |
| ————— | ———– | ———– | ———– |
| 1 | 42 | 6 | 48 |
| 43.500 | 4.500 |  |  |
| -0.227 | 0.707 |  |  |
| ————— | ———– | ———– | ———– |
| Column Total | 87 | 9 | 96 |
| ————— | ———– | ———– | ———– |

Statistics for All Table Factors

## Pearson’s Chi-squared test

Chi^2 = 1.103448 d.f. = 1 p = 0.2935106

## Pearson’s Chi-squared test with Yates’ continuity correction

Chi^2 = 0.4904215 d.f. = 1 p = 0.4837394

## Fisher’s Exact Test for Count Data

Sample estimate odds ratio: 2.126341

Alternative hypothesis: true odds ratio is not equal to 1 p = 0.4860486 95% confidence interval: 0.4217293 13.97658

Alternative hypothesis: true odds ratio is less than 1 p = 0.9206787 95% confidence interval: 0 10.39182

Alternative hypothesis: true odds ratio is greater than 1 p = 0.2430243 95% confidence interval: 0.5242304 Inf

Minimum expected frequency: 4.5

Cells with Expected Frequency < 5: 2 of 4 (50%)

(#tab:unnamed-chunk-8)

*Descriptive statistics for demographic variables of infants and mothers included in the present study.*

|  |  |  |
| --- | --- | --- |
|  | CC (n = 49) | HOME (n = 49) |
| **Gender** |  |  |
| male | 29 | 25 |
| female | 20 | 24 |
| **Age PRE** |  |  |
| mean (sd) | 87.8 16.0 | 76.7 6.3 |
| min | 58 | 68 |
| max | 123 | 90 |
| **Age POST** |  |  |
| mean (sd) | 118.4 16.1 | 106.5 5.9 |
| min | 90 | 97 |
| max | 154 | 116 |
| **Maternal Education** |  |  |
| mean (sd) | 32.9 3.0 | 32.2 3.6 |
| min | 25.1 | 24.9 |
| max | 42.0 | 40.1 |
| **Birthweight** |  |  |
| mean (sd) | 3630.4 508.9 | 3636.0 438.4 |
| min | 2708 | 2810 |
| max | 4600 | 4700 |
| **Breastfeeding (Birth - PRE)** |  |  |
| mean (sd) | 5.4 2.9 | 5.7 2.3 |
| min | 0 | 0 |
| max | 11.4 | 8.9 |
| **Breastfeeding (PRE - POST)** |  |  |
| mean (sd) | 4.0 2.9 | 3.8 2.8 |
| min | 0 | 0 |
| max | 8.5 | 8.2 |
| **Sibling** |  |  |
| yes | 25 | 32 |
| no | 23 | 17 |
| C-section |  |  |
| yes | 6 | 3 |
| no | 42 | 45 |

*Note.* CC = childcare. Breastfeeding refers to the average number of breast-feedings per day. Age refers to the age in days.

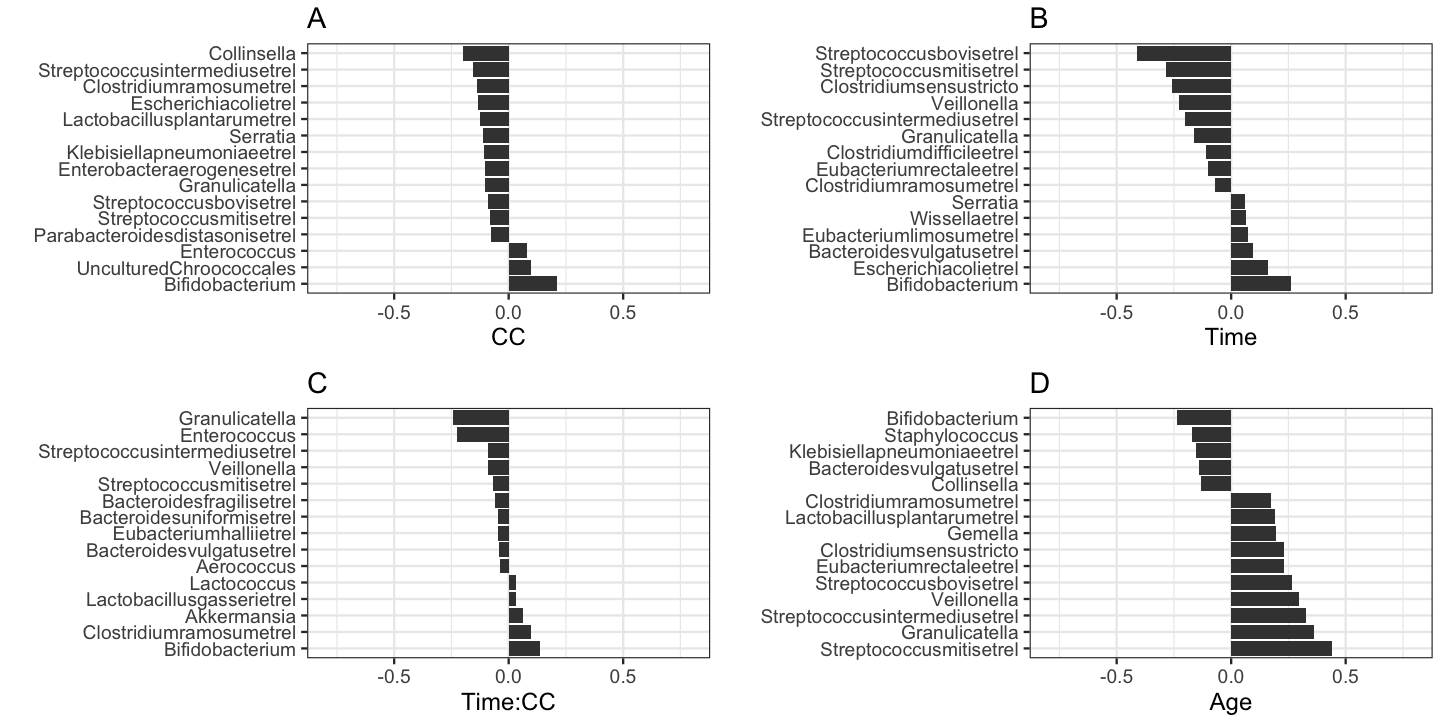
### 3.2.1 Permutational multivariate ANOVA

We compared the overall community composition using PERMANOVA based on Aitchison distance metric. An assumption for PERMANOVA is multivariate homogeneity of group dispersions (variances) [@andersonPermutationalMultivariateAnalysis2017]. We used the function *betadisper* [@R-vegan], which utilizes the *PERMDISP2* procedure as implemented by Marti Anderson and found that this assumption was met for the factors “CC” *F*(1,194) = 1.74, *p* = .188, “Time” *F*(1,194) = 1.73, *p* = .190, “Csection” *F*(1,194) = 0.47, *p* = .494 and the subgroups that result out of the interaction of “Time” and “CC” *F*(3,192) = 1.19, *p* = .313. However, the dispersions between the groups of “Sibling” were not homogeneous *F*(1,194) = 8.24, *p* = .004 and therefore the resulting test statistics should be interpreted with caution. We did not find a significant effect of CC over time on overall community composition (see table 2). Breastfeeding, sibling, c-section and age significantly predicted overall community composition. Figure x shows the genera that mostly changed as a function of each significant predictor. Despite the low number of infants born by cesarian section, PERMANOVA indicates that cesarean section significantly influences gut microbiota composition with the strongest effect on Bifidobacteria.

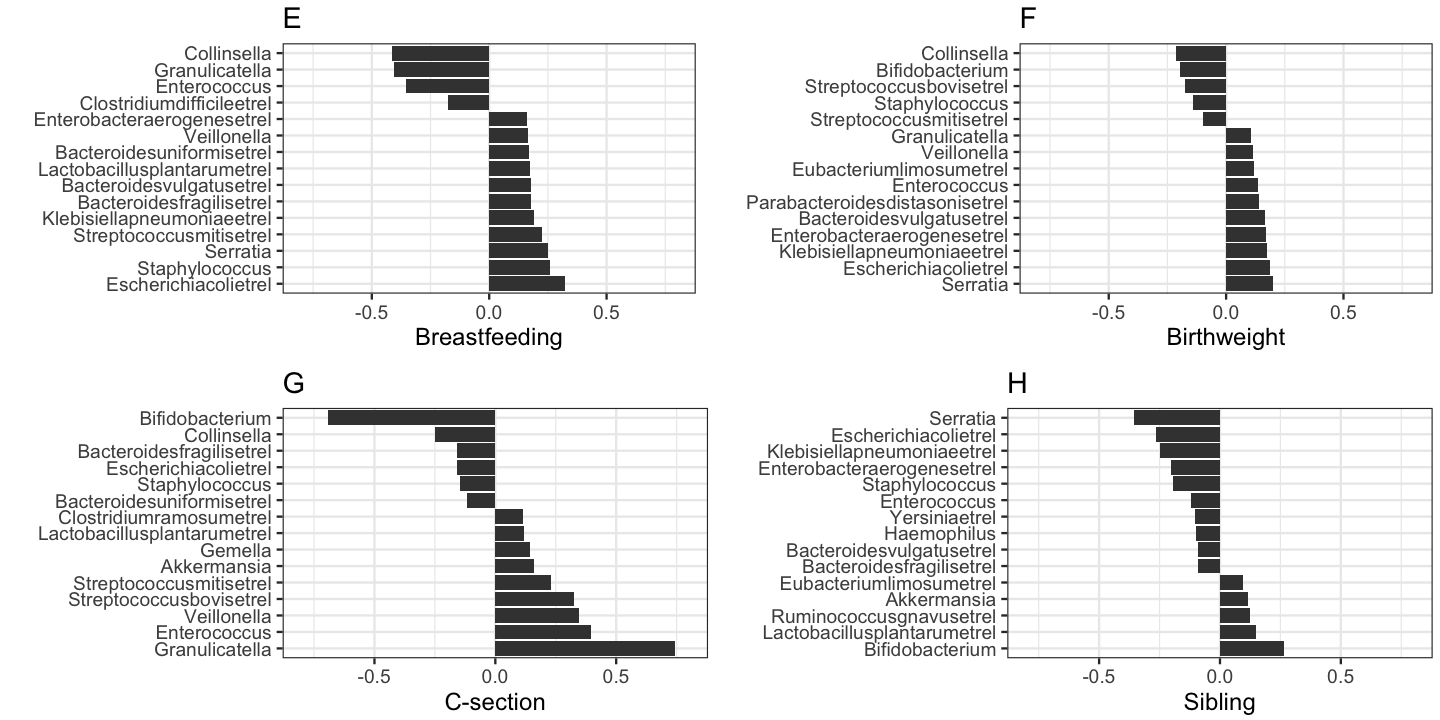
(#tab:unnamed-chunk-9)

*Model Output PERMANOVA*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Model Parameter | Sum of Squares | Mean Sum of Squares | F | Df | p | R Square |
| Time | 58.44 | 58.444 | 1.51 | 1.00 | 0.116 | 0.01 |
| CC | 52.64 | 52.636 | 1.36 | 1.00 | 0.174 | 0.01 |
| Age | 79.55 | 79.553 | 2.056 | 1.00 | 0.024 | 0.01 |
| Breastfeeding | 206.11 | 206.114 | 5.326 | 1.00 | 0.001 | 0.03 |
| C-section | 125.87 | 125.873 | 3.253 | 1.00 | 0.001 | 0.02 |
| Sibling | 96.07 | 96.074 | 2.483 | 1.00 | 0.006 | 0.01 |
| Birthweight | 89.04 | 89.037 | 2.301 | 1.00 | 0.012 | 0.01 |
| Time x CC | 38.18 | 38.178 | 0.987 | 1.00 | 0.409 | 0.00 |
| Residuals | 7,236.56 | 38.698 | - | 187.00 | - | 0.91 |
| Total | 7,982.47 | - | - | 195.00 | - | 1.00 |



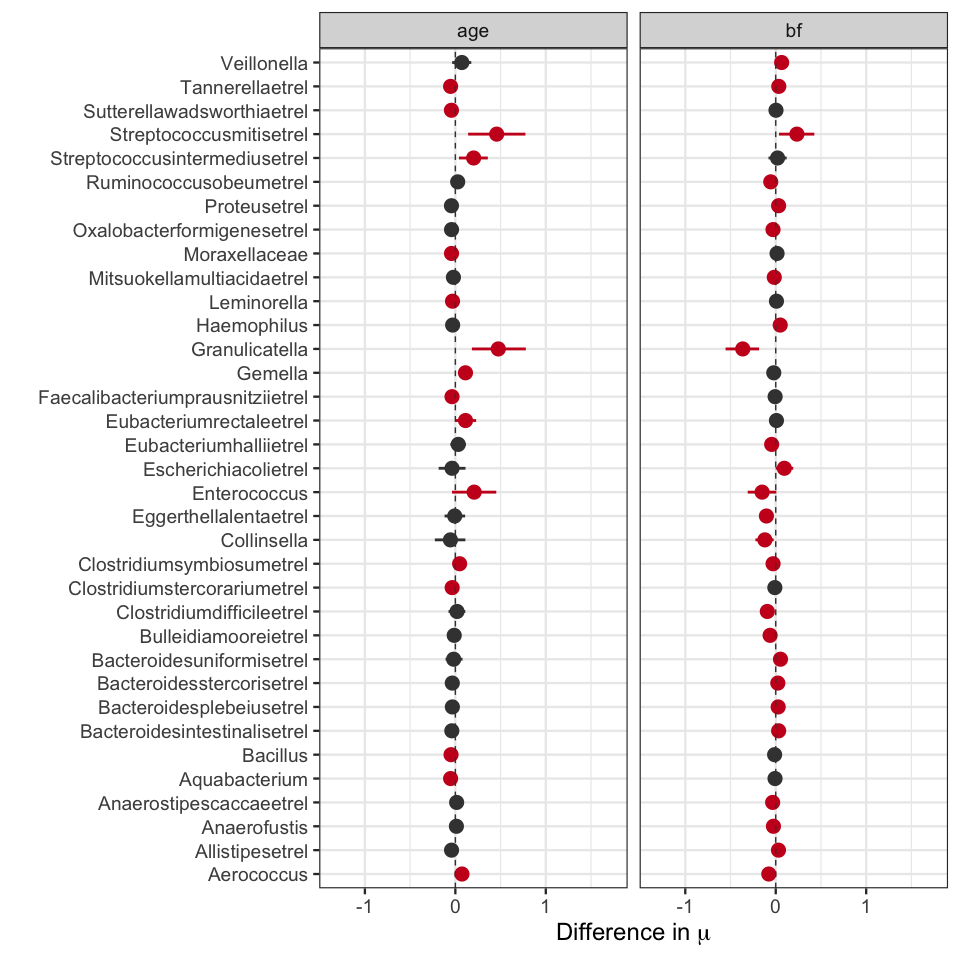
Top Taxa that differ most as a function of each predictor.



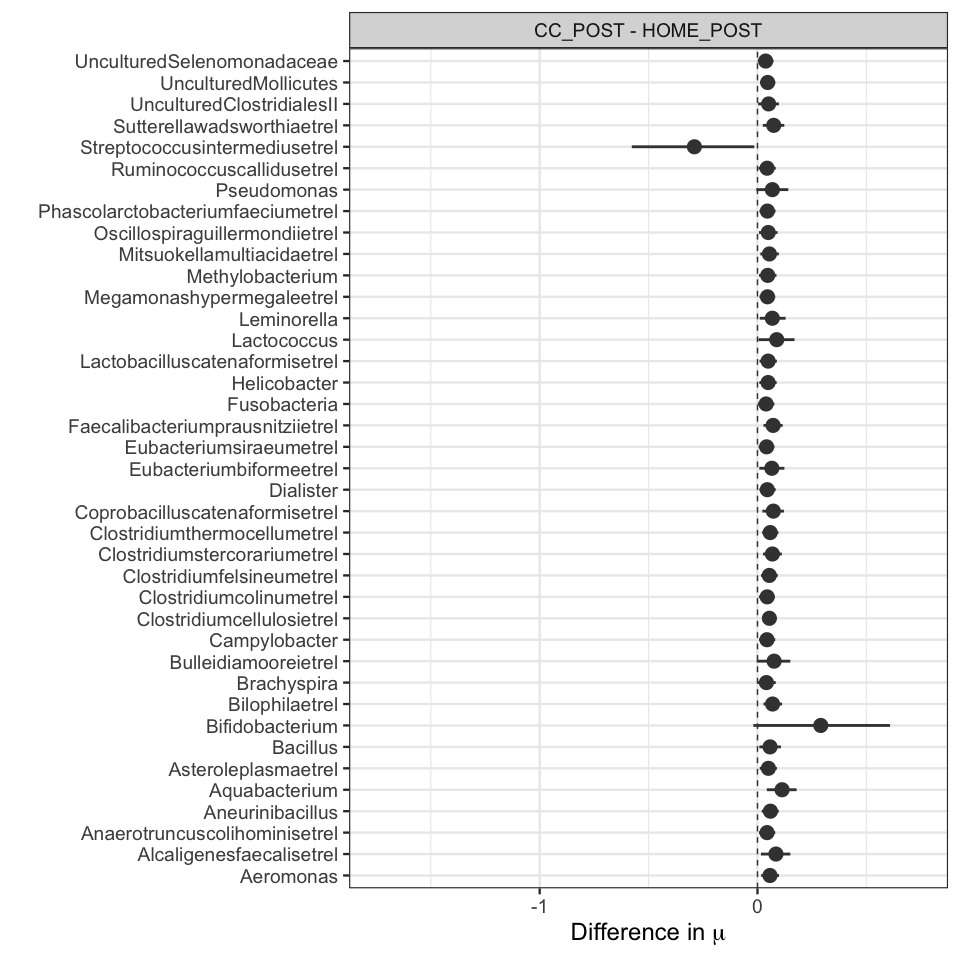
Top Taxa that differ most as a function of each predictor.

#### 3.2.2 Differential abundance testing (hierarchical GLM)

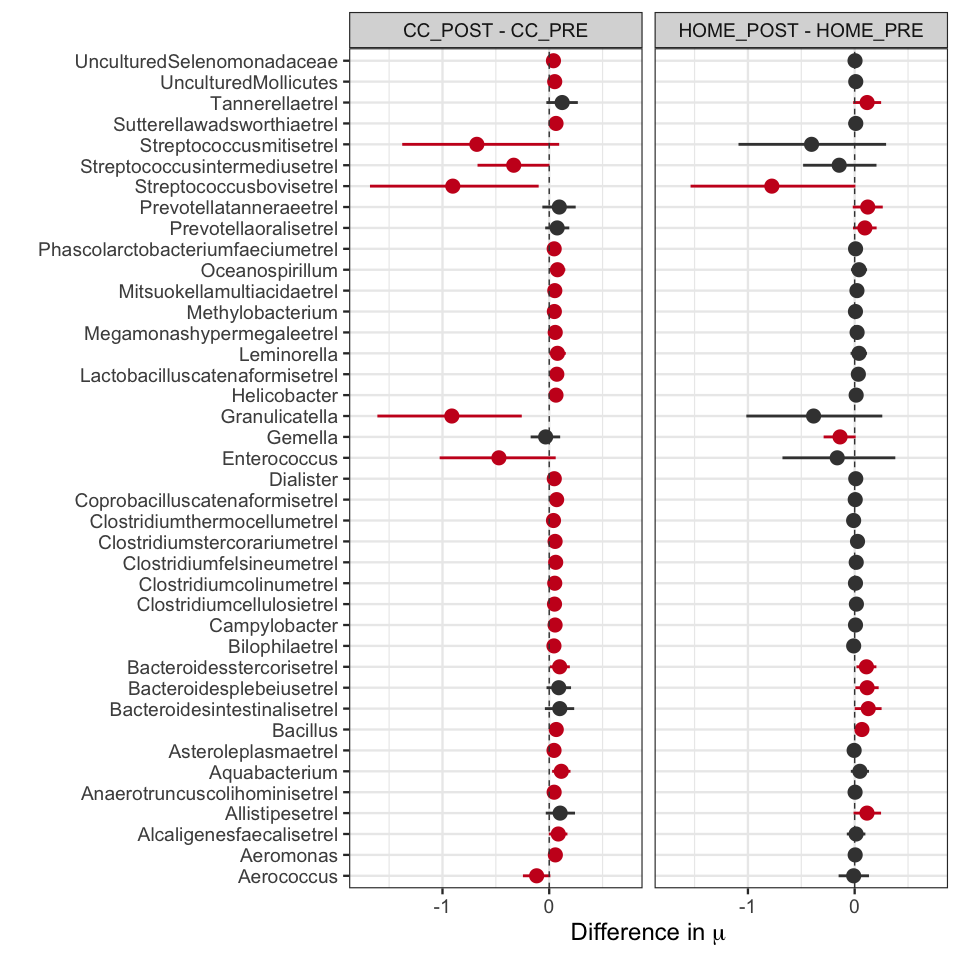
We modeled clr-transformed bacterial abundance using the generalized Gaussian distribution with constant variance and skewness parameter . The parameter is modeled as a linear function of the predictors “CC”, “Time”, “breastfeeding”, “sibling”, “csection” and “age”. Figures x-x show the arithmetric means of the posterior distribution of the differences in between (figure x) or within (figure x) the groups with 95% highest probability density interval. We show only those genera that were different with high probability ( %) in at least one comparison or predictor. Holding all other predictors constant, the GLM predicts a relative increase in the abundance of *Streptococcus mitis et rel*, *Streptococcus intermedius et rel*, *Granulicatella* and *Enterococcus* with increasing age. Breastfeeding is positively associated with *Streptococcus mitis et rel* and *Staphylococcus* and negatively associated with *Granulicatella*, *Collinsella* and *Enterococcus*. Comparing CC to HOME, we find no clear difference at time-point PRE. At time-point POST we find a relative decrease in *Streptococcus intermedius et rel*. We find a stronger decrease of *Streptococcus mitis et rel*, *Streptococcus intermedius et rel* and *Granulicatella* in the CC group compared to HOME. These findings are mostly in line with the previously presented analyses and biplots. Since there is a low number number of c-sections in our group (9), the posterior distribution of the group difference between cesction is expected to be wide reflecting high uncertainty. Furthermore, the skeptical prior assigns highest probability to the effect being equal to zero. Nevertheless, the model detects a clear decrease in *Bifidobacterium* for infants born by c-section, which is less strong for those infants with siblings.



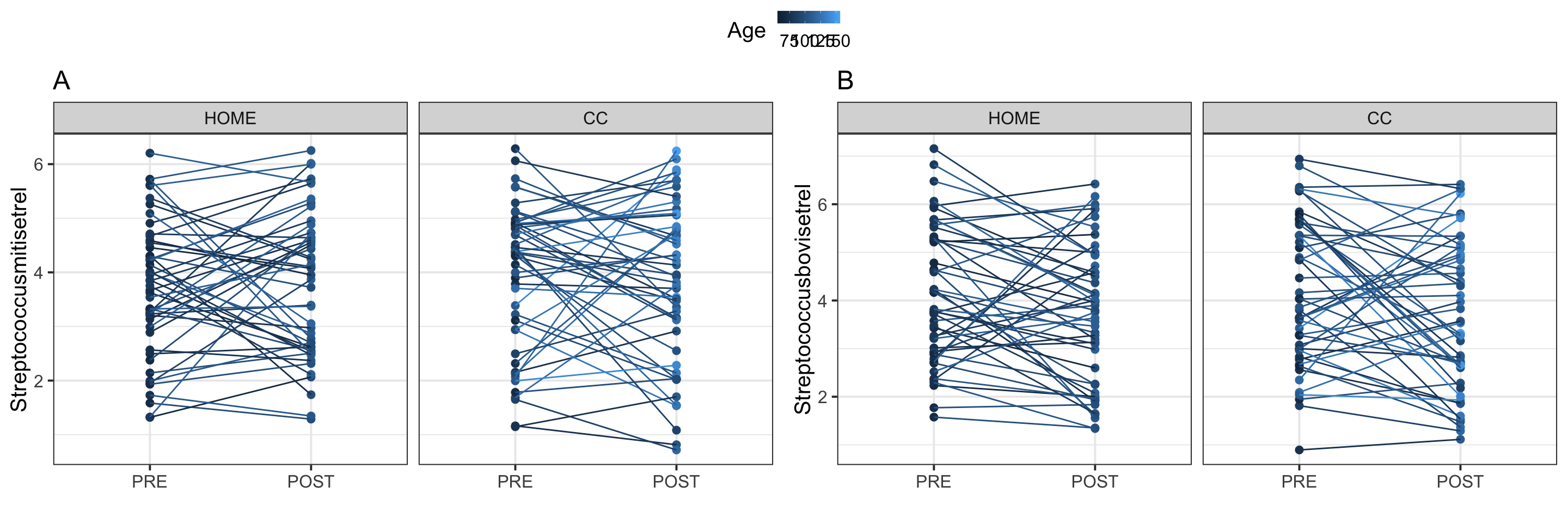
Posterior distribution of the slopes of age and breastfeeding. Red color indicated that 95% of the posterior distribution is greater or smaller than zero.



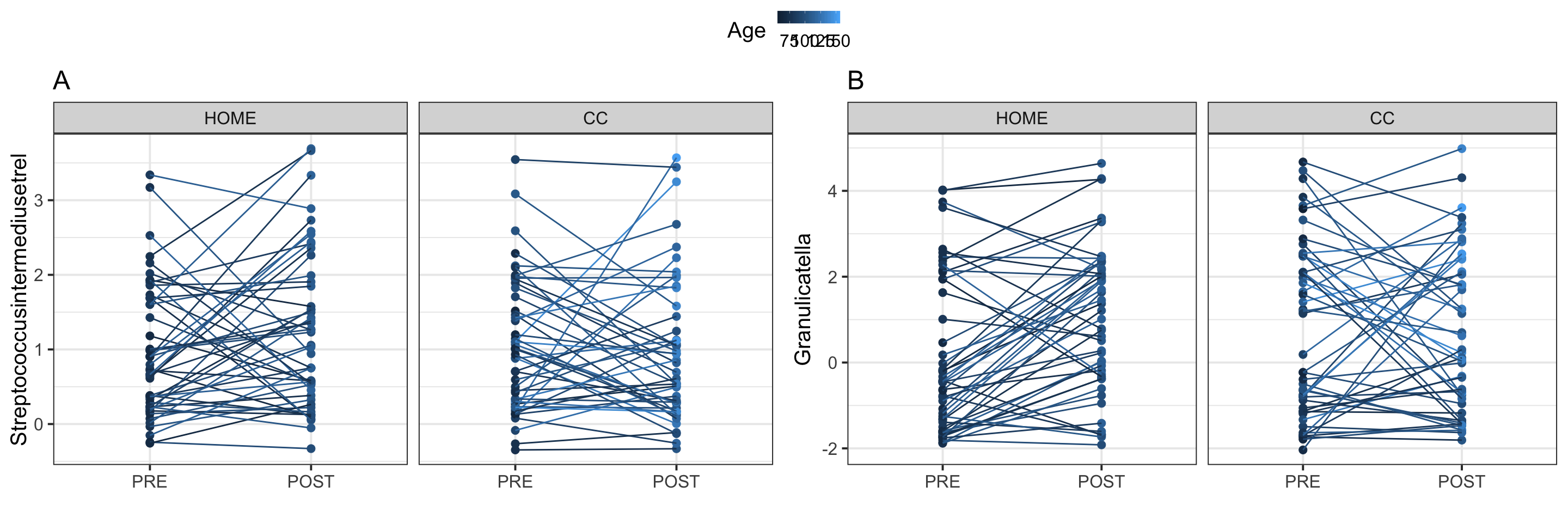
Posterior distribution of the difference in mu between CC and HOME. Red color indicated that 95% of the posterior distribution is greater or smaller than zero.



Posterior distribution of the difference in mu within CC and HOME. Red color indicated that 95% of the posterior distribution is greater or smaller than zero.



Individual paths between PRE and POST samples for Streoptococcus mitis et rel (A) and Streptococcus bovis et rel (B).



Individual paths between PRE and POST samples for Streoptococcus intermedius et rel (A) and Granulicatella (B).

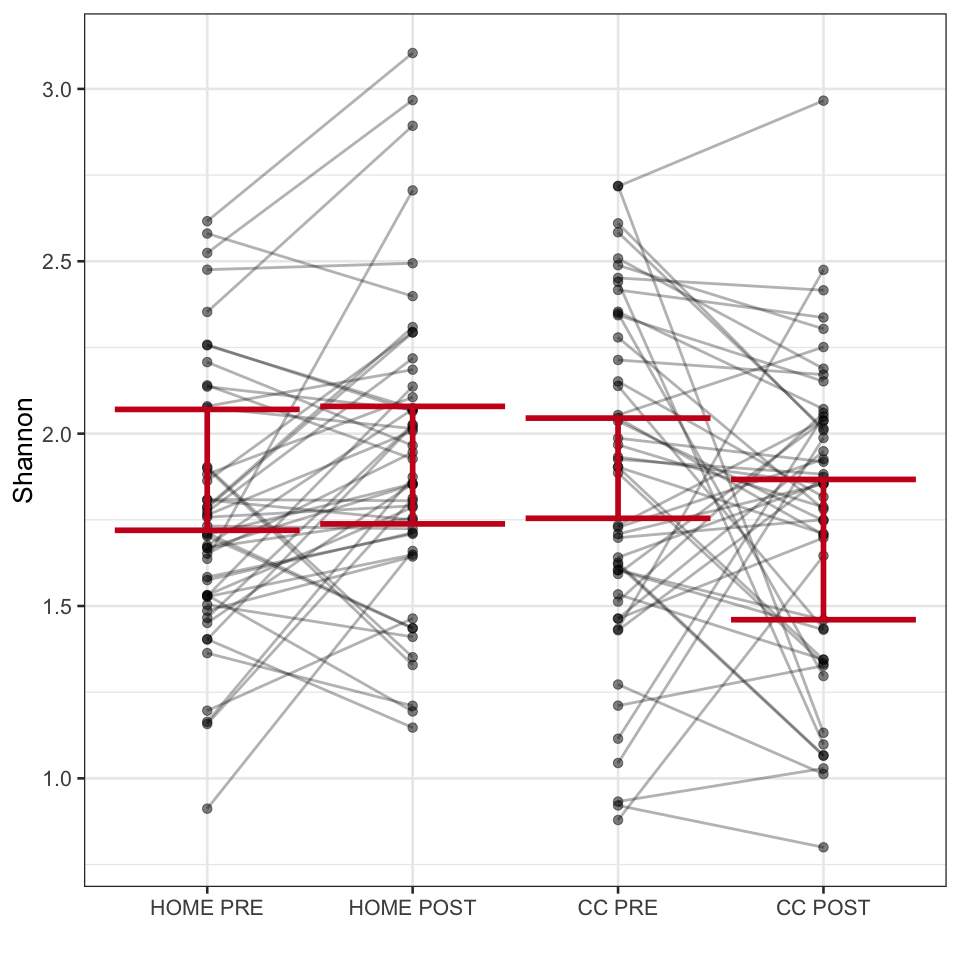
#### 3.2.3 Alpha diversity (hierarchical LM)

Alpha-diversity was assumed to be Gaussian distributed with constant variance . Table x shows the estimated difference in the parameter of the generalized Gaussian distribution between groups as well as the estimated . There was no difference in alpha diversity within the HOME group or between HOME and CC before childcare entrance. However, comparing within CC or between HOME and CC after entrance, we see that diversity is lower in the CC group with high certainty suggesting that CC lead to a decrease in alpha diversity. Figure 3 shows alpha diversity for each subject for each subgroup. It furthermore shows the highest probability density interval of for each group. We see that was highest in the CC group before entrance and lowest of all groups after CC attendance. However, we see again large individual variation within each group. Besides that, cesarian section leads to an increase in alpha-diversity, whereas having a sibling is associated with a decreased alpha-diversity.

(#tab:unnamed-chunk-17)

*Estimated model parameters alpha diversity.*

|  |  |  |  |
| --- | --- | --- | --- |
| Parameter | Mean | 95% HPDI | P(Parameter < 0) |
| Age | 0.07 | [-0.042, 0.191] | 0.11 |
| Bf | -0.06 | [-0.133, 0.004] | 0.97 |
| CCPOST - CCPRE | -0.24 | [-0.462, -0.018] | 0.98 |
| CCPOST - HOMEPOST | -0.25 | [-0.423, -0.069] | 1.00 |
| CCPRE - HOMEPRE | 0.00 | [-0.174, 0.175] | 0.50 |
| Csection | 0.45 | [0.147, 0.765] | 0.00 |
| Csection:sibling | -0.04 | [-0.473, 0.418] | 0.56 |
| HOME\_POST - HOME\_PRE | 0.01 | [-0.217, 0.225] | 0.47 |
| Sibling | -0.09 | [-0.243, 0.056] | 0.88 |
| Sigma | 0.33 | [0.275, 0.398] | 0.00 |



Observed values of Shannon alpha-diversity (points), individual paths and 95% highest probability density interval of the posterior distribution of mu when age and breastfeeding are kept at sample average.

## 3.4 Random Forest

According to our hypotheses, we would expect to be able to classify whether an infant belongs to the CC group at time-point POST. In contrast, at time-point PRE we would expect prediction accuracy to be lower since there should be no differences between CC and HOME. Neither the PRE model, nor the POST model achieved a higher prediction accuracy than 0.51 suggesting that there was no systematic effect of childcare entrance on microbiota composition.

# 4. Discussion

We utilized both multivariate and univariate statistical approaches to test our hypotheses that CC has an effect on gut microbiota composition . Our analyses did not indicate a significant effect of CC over time and therefore both null hypotheses could not be rejected for our sample. Unexpectedly, breastfeeding and CC together explained only small amounts of variation in overall microbiota composition.  
This result is in contrast to the previously cited rodent studies [@Rea2016], but in line with @Falony2016 who recently explored the drivers of microbiota variation and similarly showed very small effect sizes for all covariates under investigation. In total, 69 of 503 (13.7%) metadata variables were significantly associated with microbiota composition and these had a cumulative (non-redundant) effect size on community variation of only 7.63%. Thus, more than 92% of microbiota variation remained unexplained suggesting additional contribution from unknown factors, such as stochastic effects and biotic interactions.  
In animal experiments the variable under investigation is usually the only factor that is changed, while in human populations the hosts show much more variation with regards to genetics and lifestyle, while the microbiota is subject to much more interactions with the environment [@Falony2016; @Rea2016].  
Nevertheless, the structure, function and stability of the infant microbiota is significantly different in composition as well as diversity from that of adults [@DeMuinck2018]. Especially in the first few weeks the gut microbiota composition is highly unstable and individual. This period is followed up by a more stable phase dominated by *Bifidobacterium* species when breastfed and gradually goes to a more mature and diverse community during weaning due to the introduction of solid foods. It is yet unclear at what age the microbiota fully matures, but estimates range from 1 to 3 years of age [@DeMuinck2018]. This difference in microbiota dynamics in the first months of life could explain the small effect sizes we observed. Because the microbiota is so unstable it might be that there are no general effects of microbiota composition, in other words: Every infant responds in a unique way. This seems to be the case as can be seen by figures 5-14 where we do observe an average response, but a high percentage of infants shows an opposite effect. Inspecting the same figures for the remaining 120 bacterial groups, we also find highly variable trends between T0 and T1, which explains why we don’t observe an overall general effect. Another line of reasoning could be that, because of the lower stability of the infant microbiota compared to adults, it is therefore more easily ‘pushed’ or shifted due to an environmental influence such as the entrance to CC. This could still be the case. However, our data suggests the effects of this ‘push’ then are highly individual.  
Another observation is that unfortunately the period of the introduction of the first solid foods often coincided with the period under investigation, possibly combining multiple additional variables that affect the microbiota, thus obscuring separate effects. Further research would elucidate whether this is the case and we could then make classes for instance based on the percentage of solid food.  
Contrasting with previous studies we found no significant differences in the abundance of *Bifidobacterium* between breastfed and non-breastfed infants (see figure 15). This can be explained by the addition of prebiotic fibers in infant formula’s such as Galactooligosaccharides (GOS) and Fructooligosaccharides (FOS) [@Borewicz2018]. These fibers have been added in 2002 to most infant formulae (personal communication) to emulate the bifidogenic effect of breastmilk.  
In addition to biological explanations for our observations, we encountered several methodological limitations. First, only a limited number of infants were not breastfed per group (18 infants who entered childcare and 15 infants who did not enter childcare). Larger sample sizes might be required to find an effect of breastfeeding or interactions between breastfeeding and CC. Secondly, the distributional properties of the taxa could not be addressed by solely using a gaussian model. Despite log transforming the abundances, the assumptions of normality of residuals and heterogeneity of variance often could not be met for our LMEs. We used non-parametric tests for exploratory purposes, which are not affected by the distributional properties but these do not take into account individual variation, which is a strength of LME when testing repeated measurements [@Barr2013]. A possible solution could be to analyze the data taking into consideration the individual distribution for each bacterial genus e.g. by using different link functions for the generalized linear mixed effects models. It should be noted though that for the purpose of inference as opposed to prediction, the assumptions of normality of residuals and homogeneity of variance are less important and at least mild violations should be unproblematic [@Gelman2007].  
In conclusion, based on our analyses it seems that childcare attendance does not influence the microbiota in a large and distinct way, beyond normal temporal variation. This does not mean that childcare attendance does not have an effect on the microbiota, but might imply that its effect on the microbiota are either very individual or are only observable at a different level, such as through bacterial metabolites.