Relations between early entrance to childcare and infant intestinal microbiota

Henrik Eckermann[1], Gerben Hermes[2], & Carolina de Weerth[1]

[1] Radboud University Nijmegen

[2] Wageningen University

Author Note

Correspondence concerning this article should be addressed to Henrik Eckermann, Kreuzfurth 10, 47559 Kranenburg, Germany. E-mail: h.eckermann@student.ru.nl

Abstract

Attending center based childcare (CC) at three months of life can be an important life changing event that includes major stressors such as long maternal separations and frequently changing caregivers. These in turn may alter the composition of the gut microbiota with possible implications for future health outcomes. As part of an ongoing longitudinal study, we investigated whether CC compared to being cared by the mother at home alters the composition of the gut microbiota and whether breastfeeding buffers the potential effect of CC on the microbiota. Stool samples of infants who entered CC (n=49) and control infants (n=49) were obtained before and four weeks after CC entrance. We did not observe an effect of CC on overall community composition. Infants who entered CC had lower alpha diversity post CC compared to no CC or pre CC.

Keywords: microbiota, childcare, breastfeeding, early life stress

Word count: 138

Relations between early entrance to childcare and infant intestinal microbiota

3. Results

3.1. Microbiota composition

Thirteen genus like groups from the *Actinobacteria*, *Firmicutes*, *Proteobacteria* and phyla showed an average abundance of ≥ 0.05 % 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%.

Figure 2 shows 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 PRE, whereas the endpoint corresponds the composition at POST. There appear to be no differences in in location between CC and HOME at PRE or POST. Also we did not identify a uniform direction of the shifts over time in either group. Instead, microbiota composition development between time points appears to be highly individual.

3.2 Effects of CC on microbiota composition

We used age and the average number of breast-feedings per day as covariates in all linear models. Table 1 shows these and other variables for both groups. There was a significant difference in age between groups $t(62.42) = -4.54$, $p < .001$ according to Welch's t-test. Besides that, there were no differences between groups for any of the remaining variables.

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) (**???**). We used the function *betadisper* (**???**), which utilizes the *PERMDISP2* procedure as implemented by Marti Anderson and found that this assumption was met for the factors *childcare* $F(1,194) = 0.19$, $p = .188$, *time* $F(1,194) = 1.73$, $p = .190$ and the subgroups that result out of the interaction of *time* and *cc* $F(3,192) = 1.19$, $p = .313$. We did not find a significant effect of CC over time on overall community composition (see table x). Breastfeeding and age significantly predicted overall community composition. Figure x shows the genera that mostly changed as a function of each significant predictor.

3.3 Differential abundance with Bayesian GLM. We modeled clr-transformed bacterial abundance using the generalized Gaussian distribution with constant variance σ and skewness parameter α. The parameter μ is modeled as a linear function of the predictors "cc", "time", "breastfeeding" and "age". Note that μ does only represent the mean if $\alpha = 0$. Figures x-x show the arithmetric means of the posterior distribution of the differences in μ between the groups with 95% highest probability density interval. Only those genera are shown that were different with high certainty in at least one comparison. Figure x shows the comparison between CC and HOME for each time point whereas figure x shows the difference between PRE and POST within each group. There are similar trends in all comparisons for *Streptococcus intermedius et rel*, *Streptococcus mitis et rel* and *Granulicatella*. However it seems as if this change is stronger in the CC group since here these genera are decreasing over time with higher certainty ($\geq$ 95 %). Since we are looking at clr-transformed abundances it is also possible that all other genera increased relative to the above mentioned genera (**at Gerben: you have more experience with interpreting clr-transformed abundances. I am still reading about this so please**

edit this so that we make those interpretation we can make).

3.4 Alpha diversity with Bayesian. Should we maybe only report one? I reported three only be because the results are slightly different depending on the index)

Alpha diversity (Shannon) was calculated using the *microbiome* package before the clr-transformation. Alpha-diversity was assumed to be Gaussian distributed with constant variance σ and skewness parameter α. Table x shows the estimated difference in the parameter μ of the generalized Gaussian distribution between groups as well as the estimated α and σ. There was no difference in alpha diversity within the HOME group or between HOME and CC before childcare entrance. Comparing μ 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 leads 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 μ for each group. We see that μ was highest in the CC group before entrance and lowest of all groups after CC attendance. However, we see large individual variation within each group and the difference in μ is small (table x).

3.4 Random Forest

RF is a tree based ensemble learning method that is well suited for classification based on microbial abundances of samples (**???**). 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. According to our hypotheses, we would expect to be able to classify whether an infant in the test data set belongs to the CC group at T1. In contrast, at T0 we would expect prediction

accuracy to be lower since there should be no differences between CC and HOME. However, neither the T0 model, nor the model for T1 achieved a higher prediction accuracy than 0.5 suggesting that there was no systematic effect of childcare entrance on microbiota composition.

References

Figure captions

Figure 1. Development of microbiota composition over time within CC (A and C) and no CC (B and D).

Figure 2. Top Taxa that differ most as a function of each predictor.

Figure 3. Posterior distribution of the difference in mu between CC and HOME.

Figure 4. Posterior distribution of the difference in mu within CC and HOME.

Figure 5. Observed values of alpha-diversity, individual paths and posterior distribution of mu.

Table captions

Table 1. Descriptive statistics for demographic variables of infants and mothers included in the present study.

Table 2. Table x. Model Output PERMANOVA

Table 3. Estimated model parameters alpha diversity.

Table 1

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 (weeks)		
mean (sd)	12.8 ± 2.3	11.2 ± 0.9
min	8.6	10.0
max	17.9	13.1
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)	5.1 ± 2.9	6.0 ± 2.2
min	0	0
max	8.9	11.4
Breastfeeding (PRE - POST)		
mean (sd)	4.0 ± 2.8	3.8 ± 3.0
min	0	0
max	8.5	8.5

Note. CC = childcare. Breastfeeding refers to the average number of breast-feedings per day.

Table 2

Table x. Model Output PERMANOVA

Model Parameter	Sum of Squares	Mean Sum of Squares	F	Df	p	R Square
time	58.44	58.444	1.469	1.00	0.118	0.01
cc	52.64	52.636	1.323	1.00	0.192	0.01
age_d_s	79.55	79.553	2	1.00	0.023	0.01
bf_count_s	197.06	197.06	4.954	1.00	0.001	0.02
time:cc	36.23	36.229	0.911	1.00	0.506	0.00
Residuals	7,558.55	39.782	-	190.00	-	0.95
Total	7,982.47	-	-	195.00	-	1.00

Table 3

Estimated model parameters alpha diversity.

Parameter	Mean	95% HPDI	P(Parameter < 0)
Alpha	-1.60	[-6.22, 2.21]	0.75
CC_POST - CC_PRE	-0.29	[-0.52, -0.07]	0.99
CC_POST - HOME_POST	-0.22	[-0.4, -0.03]	0.99
CC_PRE - HOME_PRE	0.05	[-0.14, 0.23]	0.31
HOME_POST - HOME_PRE	-0.03	[-0.25, 0.19]	0.60
Sigma	0.33	[0.27, 0.39]	0.00

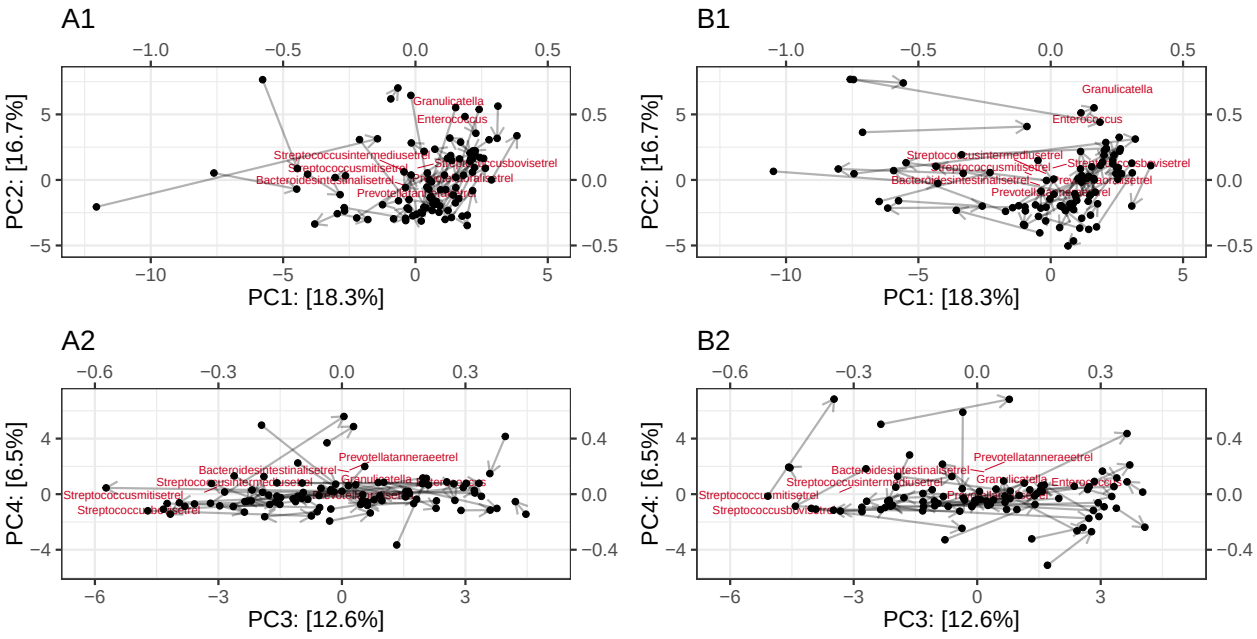

Figure 1. Development of microbiota composition over time within CC (A and C) and no CC (B and D).

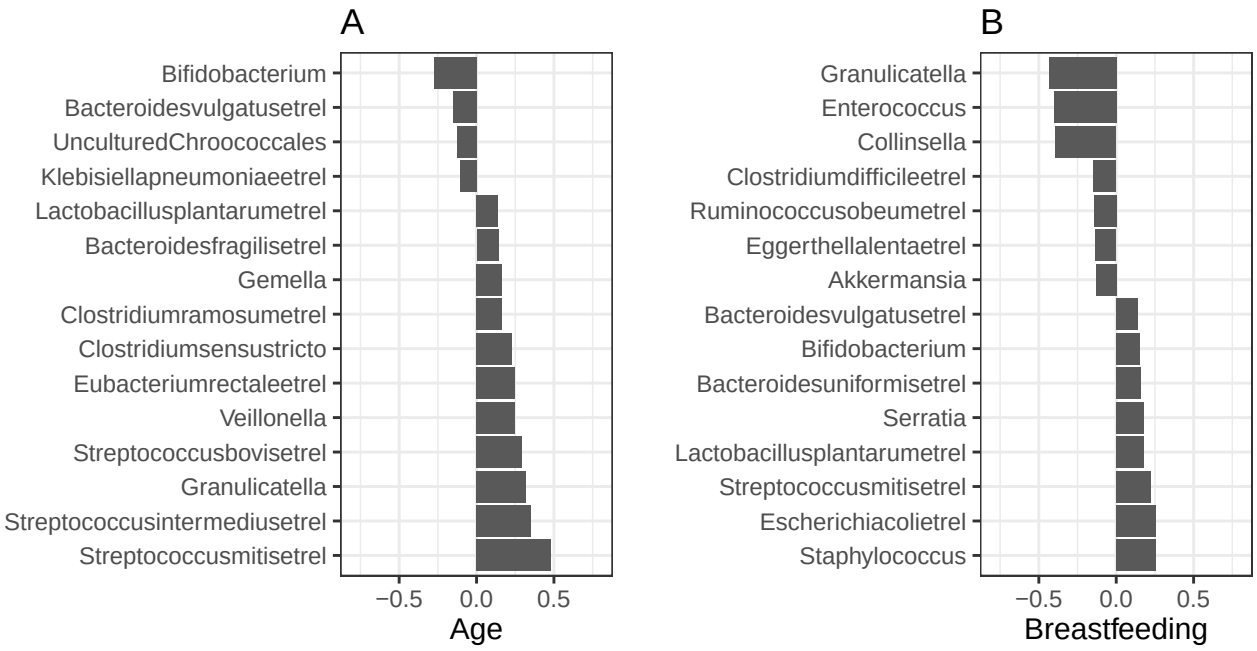

Figure 2. Top Taxa that differ most as a function of each predictor.

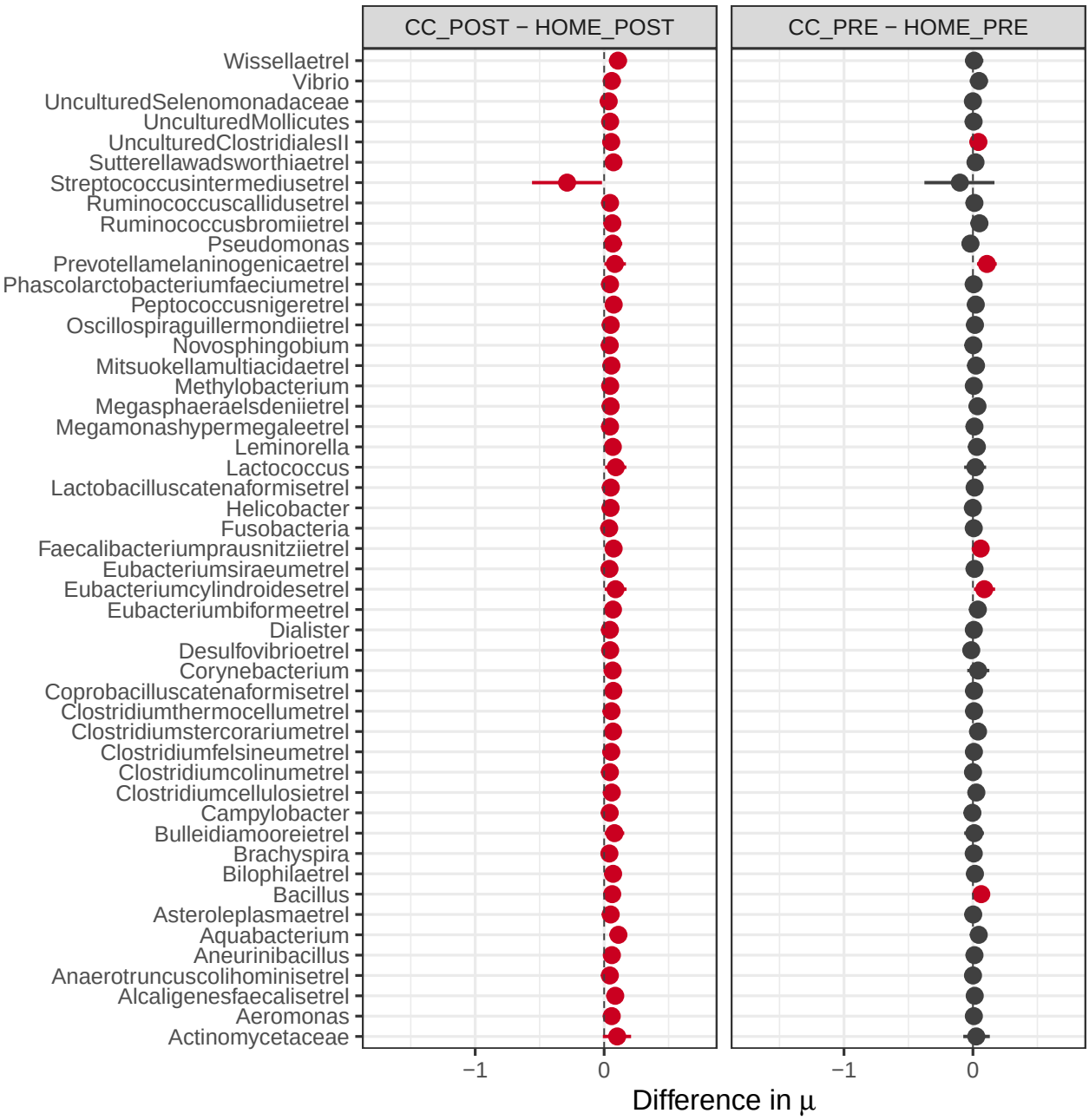

Figure 3. Posterior distribution of the difference in mu between CC and HOME.

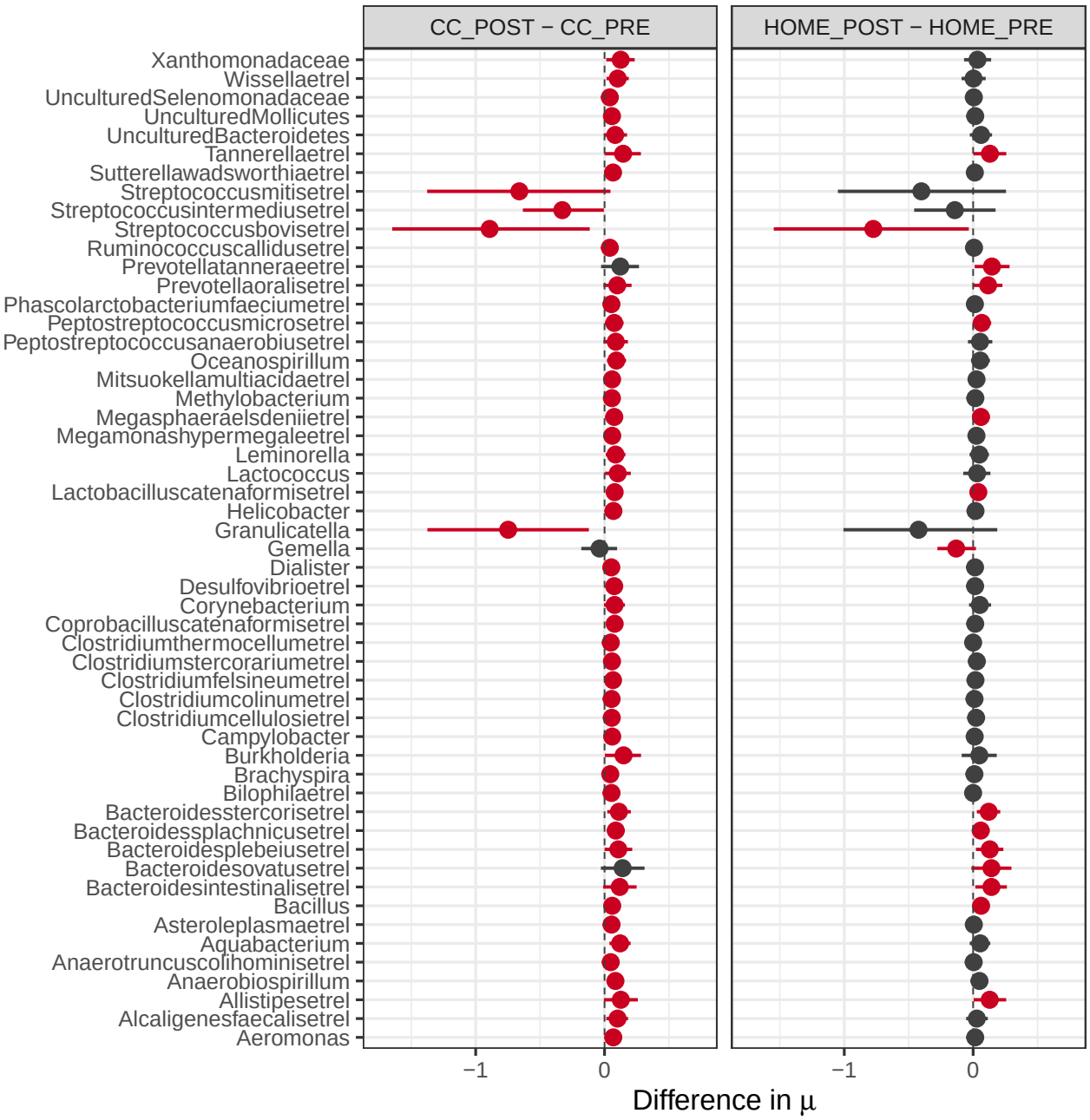

Figure 4. Posterior distribution of the difference in mu within CC and HOME.

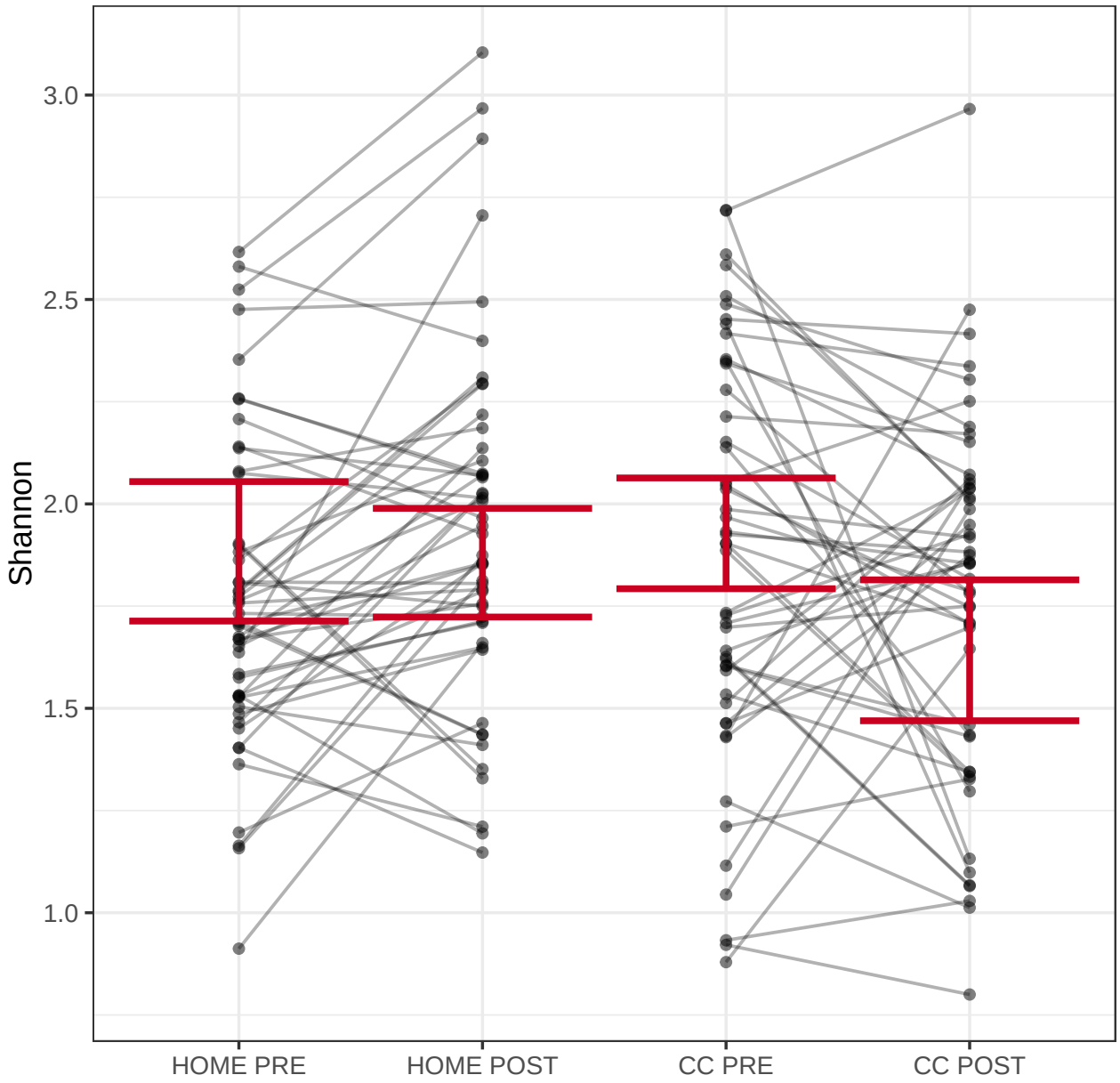

Figure 5. Observed values of alpha-diversity, individual paths and posterior distribution of mu.