DISTANCE-BASED INTRACLASS CORRELATION COEFFICIENT (ICC)

Consider the simple random effects model

$$y_{ij} = \mu + \alpha_j + \epsilon_{ij},\ i = 1, \cdots, N$$

where y_{ij} is the observed measurement for ith subject and jth replicate measurement, μ is the population mean, $\alpha_j \sim N(0, \sigma_b^2)$ is the subject-level random effect shared by all measurements for ith subject and $\epsilon_{ij} \sim N(0, \sigma_\epsilon^2)$ is the measurement error. σ_b^2 and σ_ϵ^2 reflect biological and technical variability respectively. The ICC is then defined as

$$ICC = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_\epsilon^2}.$$

To estimate ${\sigma_\epsilon}^2$, ${\sigma_b}^2$, we replace $E(WSS)$, $E(T\,SS)$ with sample WSS, $T\,SS$. For pairwise distances, the within-subject sum of squares (WSS) is defined in terms of the sum of the squares of distances within subjects

$$WSS_d = \sum_{i=1}^{N} \sum_{j,k \in \mathcal{A}(i), j>k} d_{jk}^2 / n_i,$$

where $\mathcal{A}(i)$ contains the indices for samples from subject i, and the total sum of squares (TSS) is defined as

$$TSS_d = \sum_{j>k} d_{jk}^2 / n.$$

It is easy to verify, when the distance is Euclidean, WSS_d and TSS_d is reduced to WSS and TSS. Thus we have the distance-based ICC estimate as

$$dICC = \frac{\hat{\sigma_b^2}}{\hat{\sigma_b^2} + \hat{\sigma_\epsilon^2}},$$

where

$$\begin{aligned} \hat{\sigma_\epsilon^2} &= \frac{WSS_d}{\sum_i (n_i - 1)} \\ \hat{\sigma_b^2} &= \frac{TSS_d - (n-1)\sigma_\epsilon^2}{n - m}. \end{aligned} \tag{1}$$

With this definition, we can see the $dICC \in (0, 1)$ and is comparable to the univariate ICC.

Reference: Distance-based Intraclass Correlation Coefficient for Complex Multivariate Measurements. Chen J. et al., to be submitted.

SUPPLEMENTAL FIGURE LEGENDS

Supplemental Figure 1. Principal coordinates analysis plot of the unweighted UniFrac distance by subject after removal of outliers.

Supplemental Figure 2. Percent of microbial variability explained by subject, sample collection type (treatment), and day of freezing was calculated using a distance-based coefficient of determination R^2 for beta-diversity estimates from unweighted UniFrac, generalized UniFrac, weighted UniFrac, and Bray-Curtis (BC) distance.

Supplemental Figure 3. Average alpha diversity for the fecal collection methods estimated with observed OTUs (3A) and the Shannon index (3B).

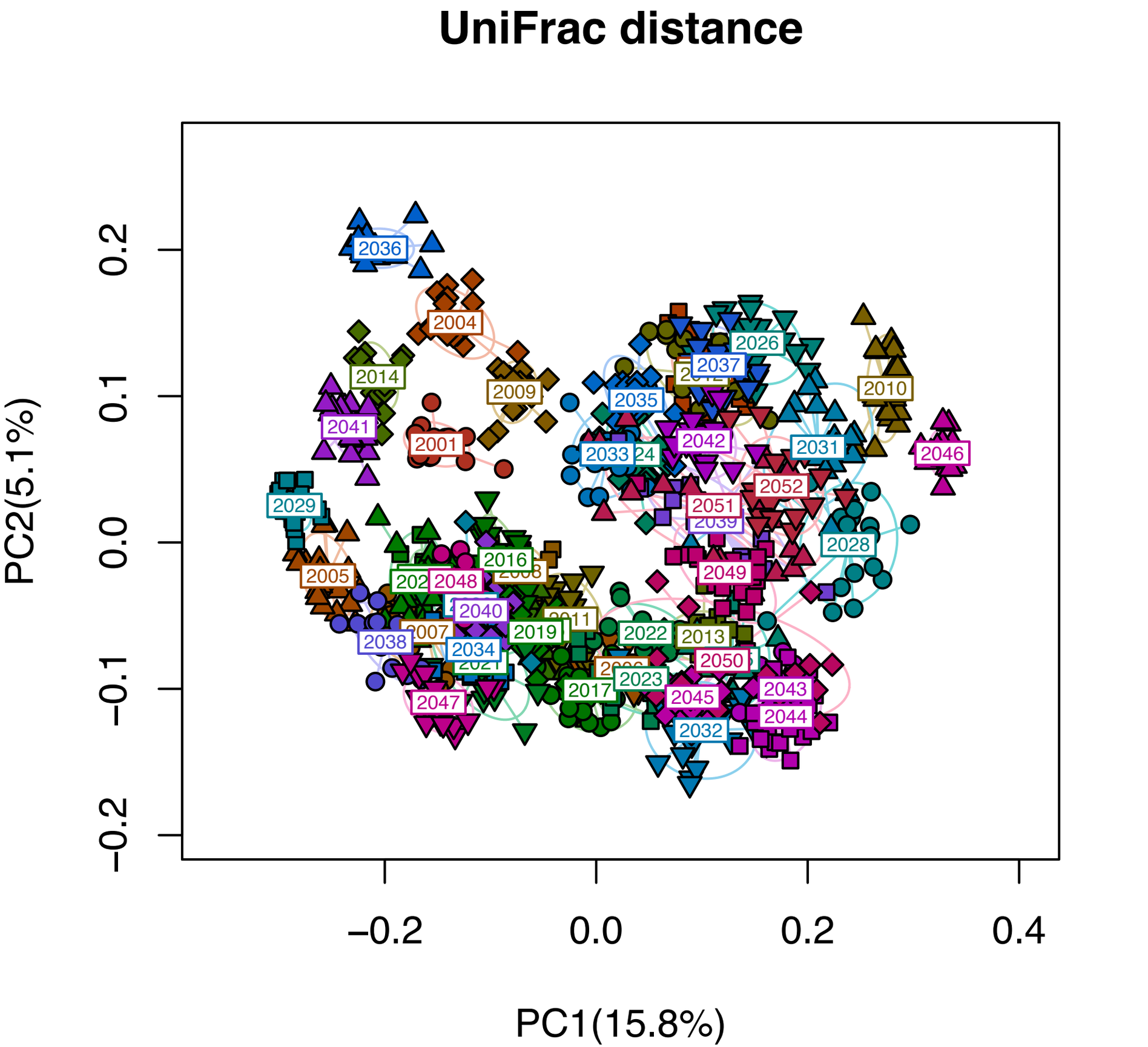
Supplemental Figure 4. Relative abundance at the phylum level by individual for Day 0 replicates (4A) and averaged for each collection method and freezing timepoint (4B).

Supplemental Figure 5. Technical reproducibility of Day 0 replicates (5A) and Day 4 replicates (5B) for the relative abundance at the genus level by fecal sample collection method using intraclass correlation coefficients.

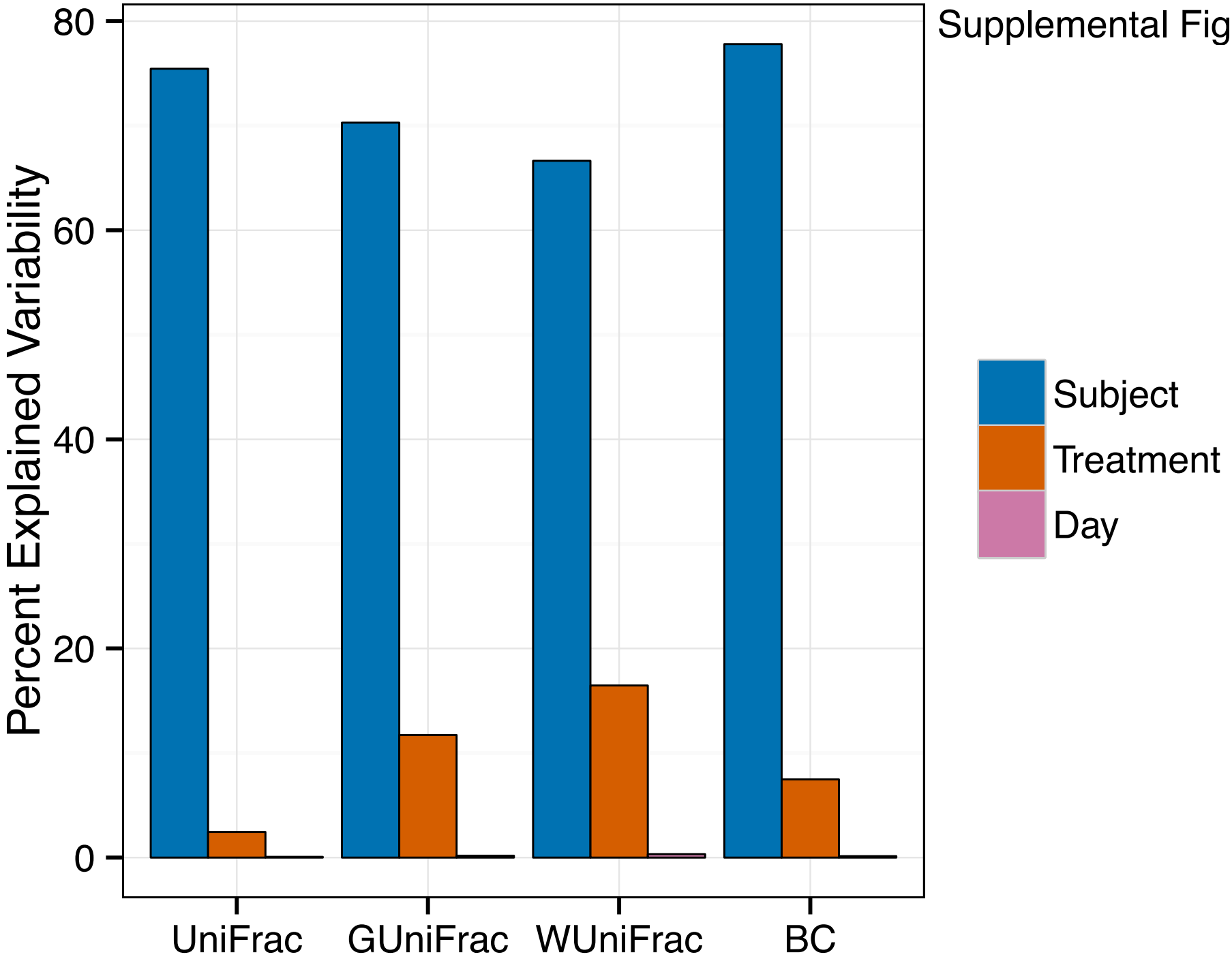
Supplemental Figure 6. Stability of fecal samples collection methods (i.e., Day 4 fecal samples compared to Day 0 fecal samples) for the relative abundance at the genus level by sample collection method using intraclass correlation coefficients.

Supplemental Figure 7. Accuracy of Day 0 fecal samples compared to the “gold standard” no solution sample frozen immediately. Intraclass correlation coefficients (7A) and Spearman correlations (7B) were calculated for the relative abundance at the genus level by sample collection method.

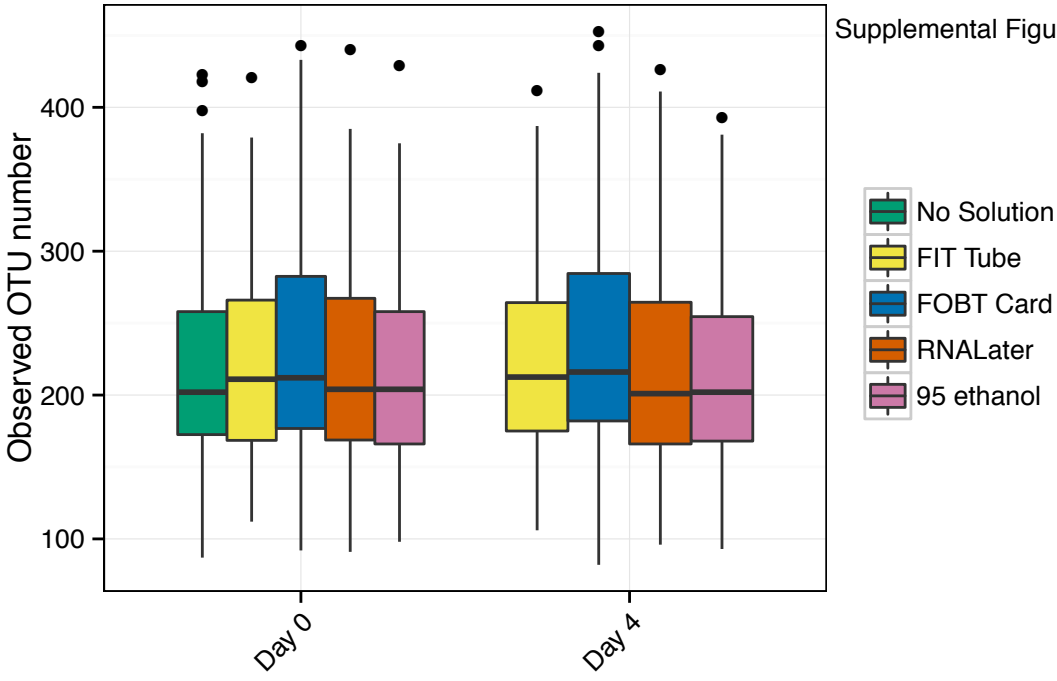
Supplemental Figure 1



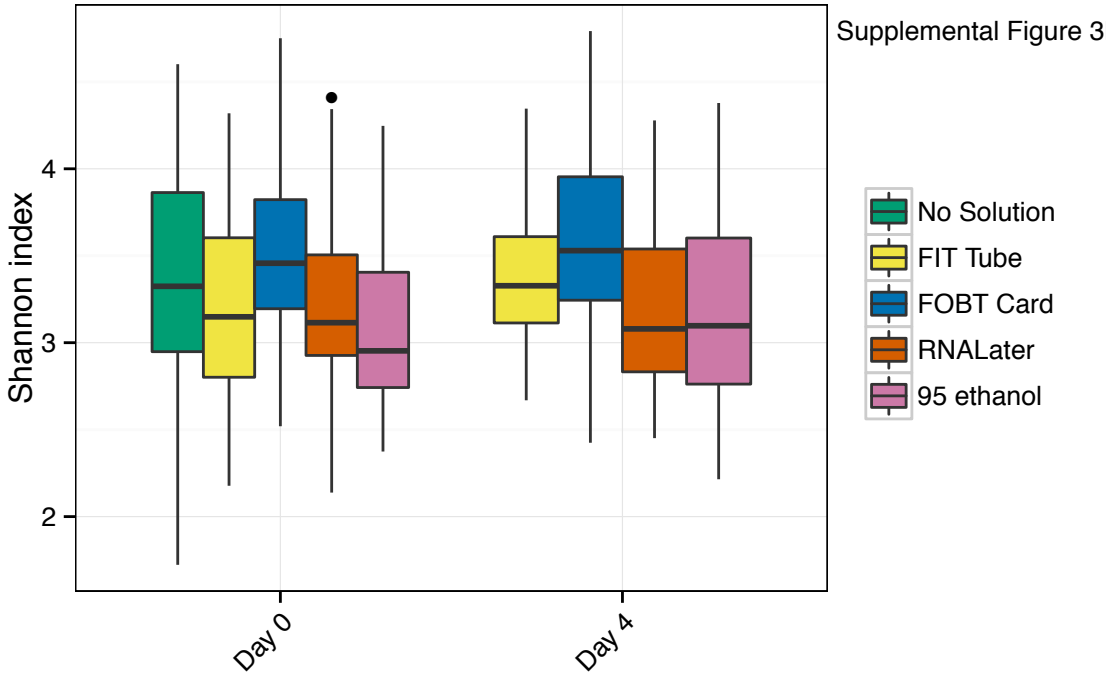
Supplemental Figure 2



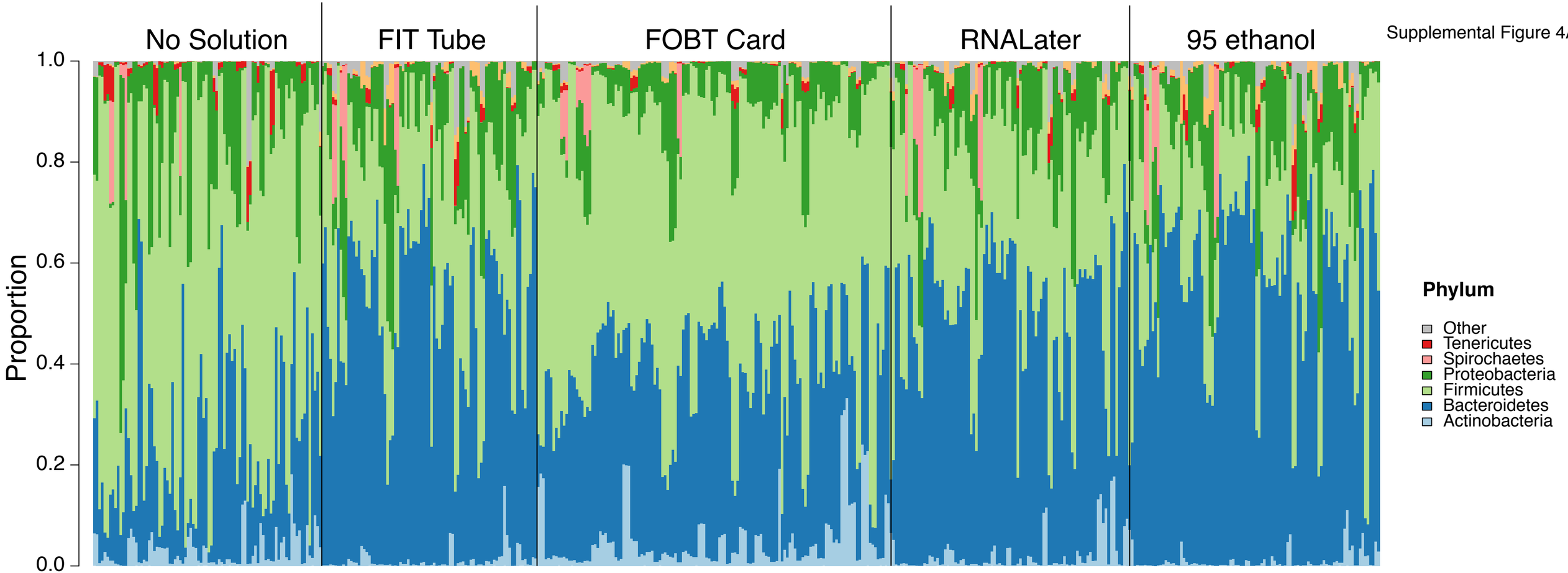
Supplemental Figure 3A



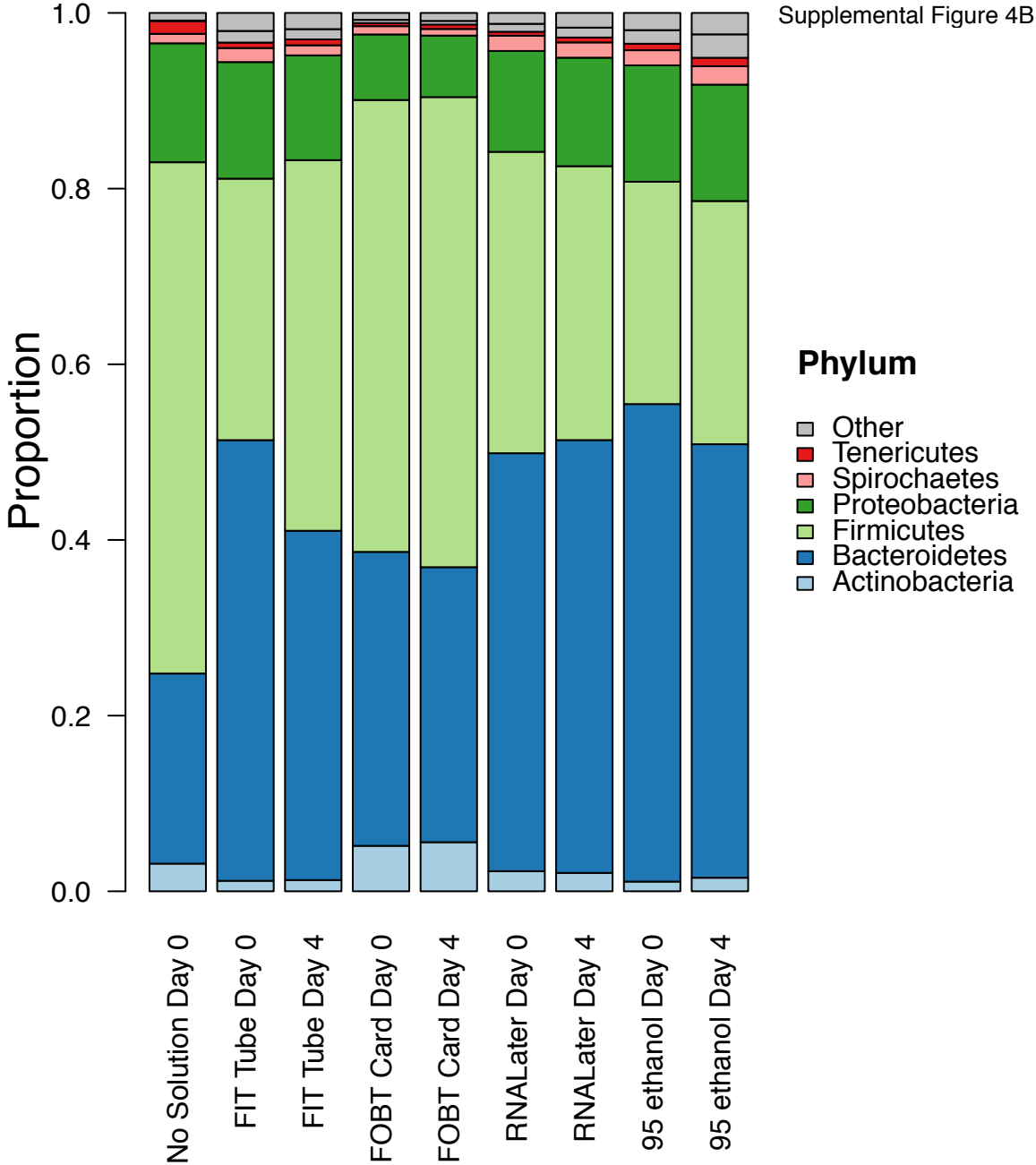
Supplemental Figure 3B



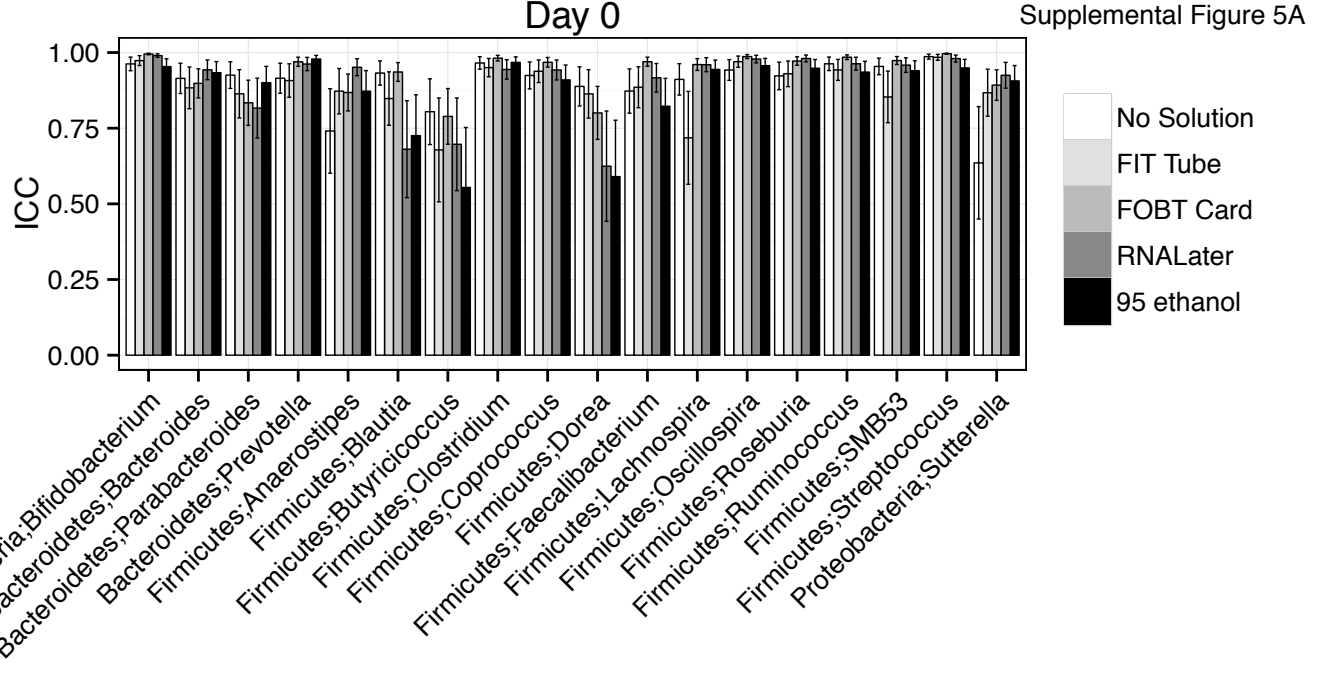
Supplemental Figure 4A



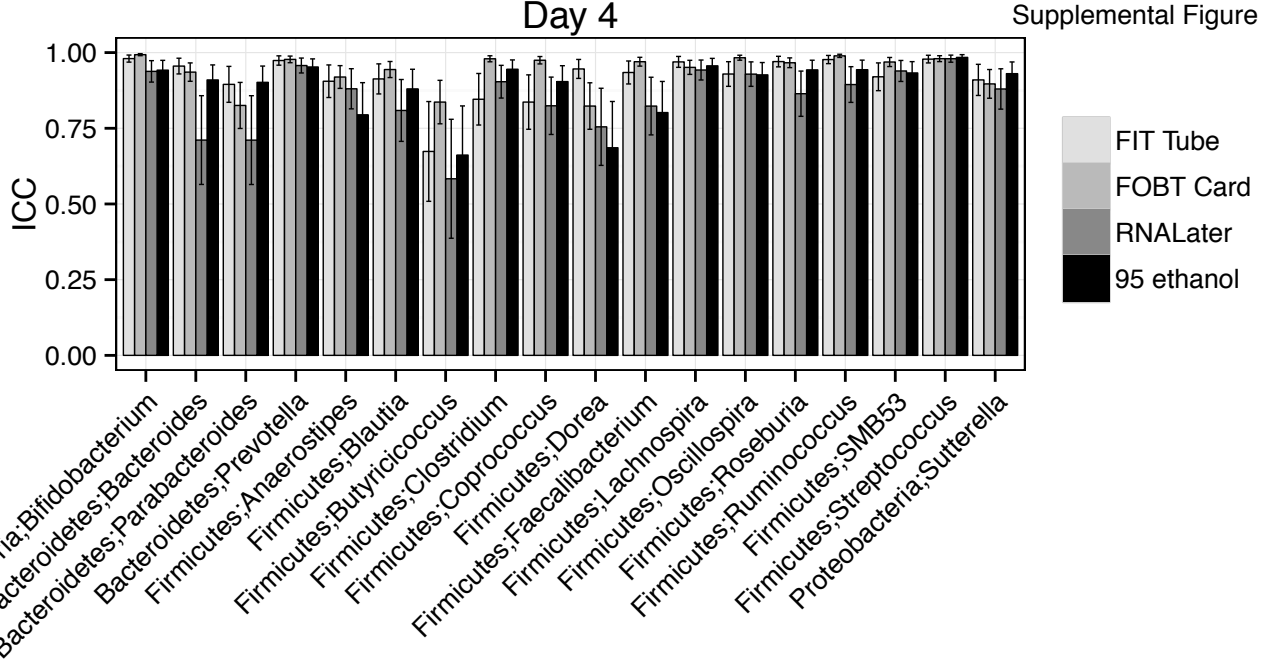
Supplemental Figure 4B



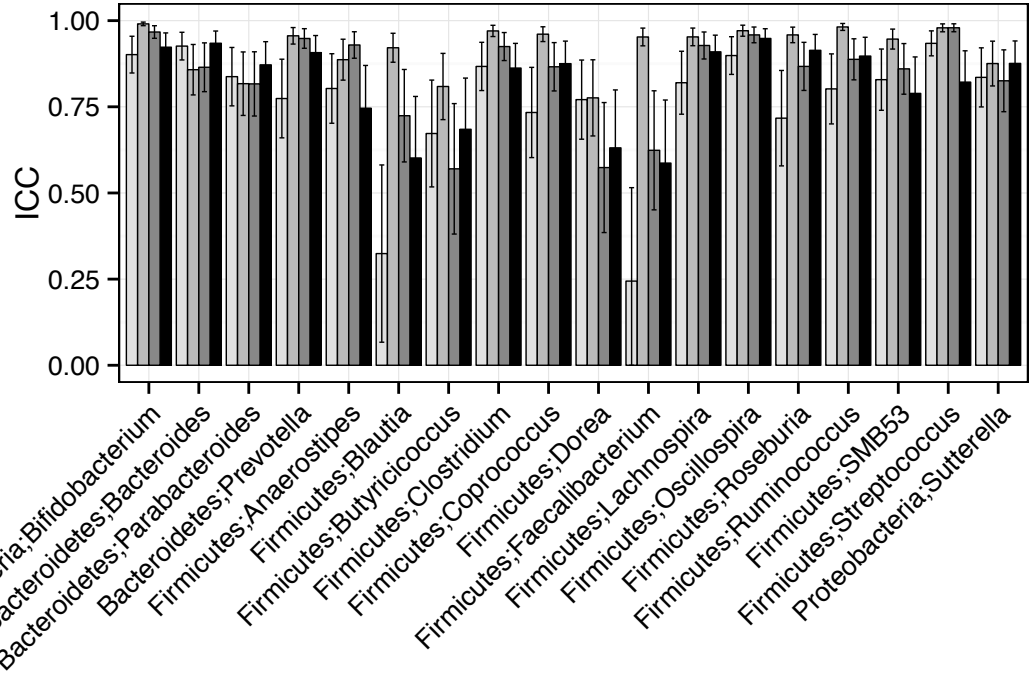
Supplemental Figure 5A



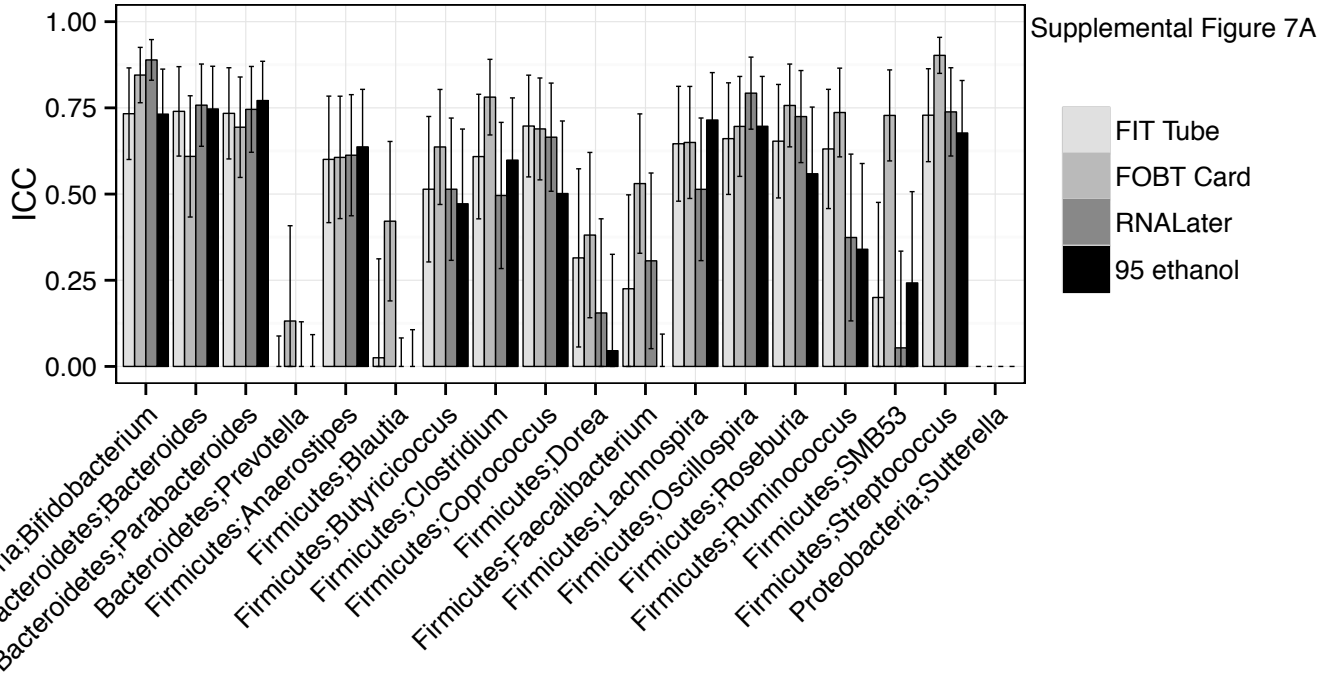
Supplemental Figure 5B

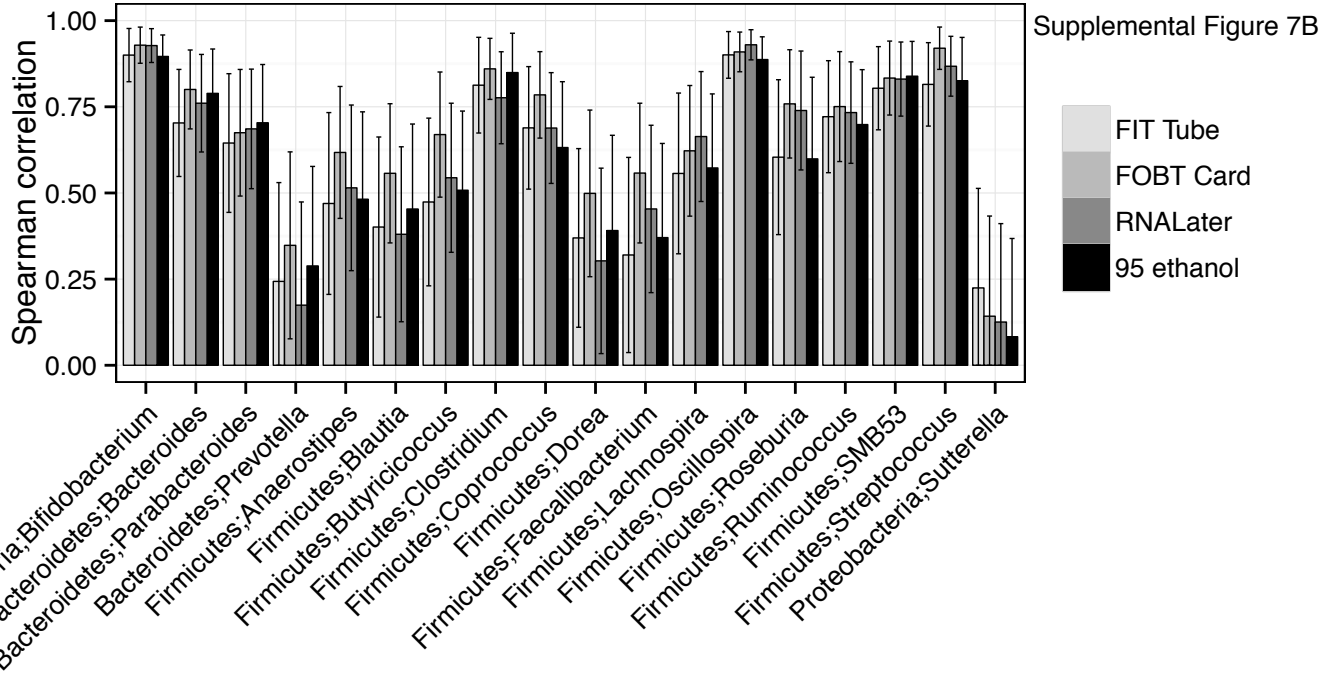


Supplemental Figure 6



Supplemental Figure 7A




Supplemental Figure 7B
Spearman correlation
1.00
0.75
0.50
0.25
0.00
FIT Tube
FOBT Card
RNALater
95 ethanol
Actinobacteria;Bifidobacterium
Bacteroidetes;Bacteroides
Bacteroidetes;Parabacteroides
Bacteroidetes;Prevotella
Firmicutes;Anaerostipes
Firmicutes;Blautia
Firmicutes;Butyricicoccus
Firmicutes;Clostridium
Firmicutes;Coprococcus
Firmicutes;Dorea
Firmicutes;Faecalibacterium
Firmicutes;Lachnospira
Firmicutes;Oscillospira
Firmicutes;Roseburia
Firmicutes;Ruminococcus
Firmicutes;SMB53
Firmicutes;Streptococcus
Proteobacteria;Sutterella