SUPPORTING INFORMATION

ANTENATAL MICROBIAL COLONISATION OF MAMMALIAN GUT

Elisa Borghi[1§*], Valentina Massa[1§], Marco Severgnini[2], Grazia Fazio[3], Laura Avagliano[1], Elena Menegola[4], Gaetano Bulfamante[1], Giulia Morace[1], Francesca Borgo[1]

[1] *Dipartimento di Scienze della Salute, Università degli Studi di Milano, 20142 Milan, Italy;*
[2] *Institute of Biomedical Technologies, National Research Council, 20090 Segrate, Italy;*
[3] *Centro Ricerca Tettamanti, Clinica Pediatrica, Dipartimento di Medicina e Chirurgia, Università degli Studi di Milano-Bicocca, Fondazione MBBM, Monza, Italy*
[4] *Dipartimento di Scienze e Politiche Ambientali, Università degli Studi di Milano, 20133 Milan, Italy;*

Supplementary statistical methods

15 samples were grouped in three categories, according to: (I) the type of tissue sampled (i.e.: amniotic liquid, placenta and intestine); (II) dam (i.e.: dam A for foetus 1-3, or dam B for foetus 4-5); (III) foetus.

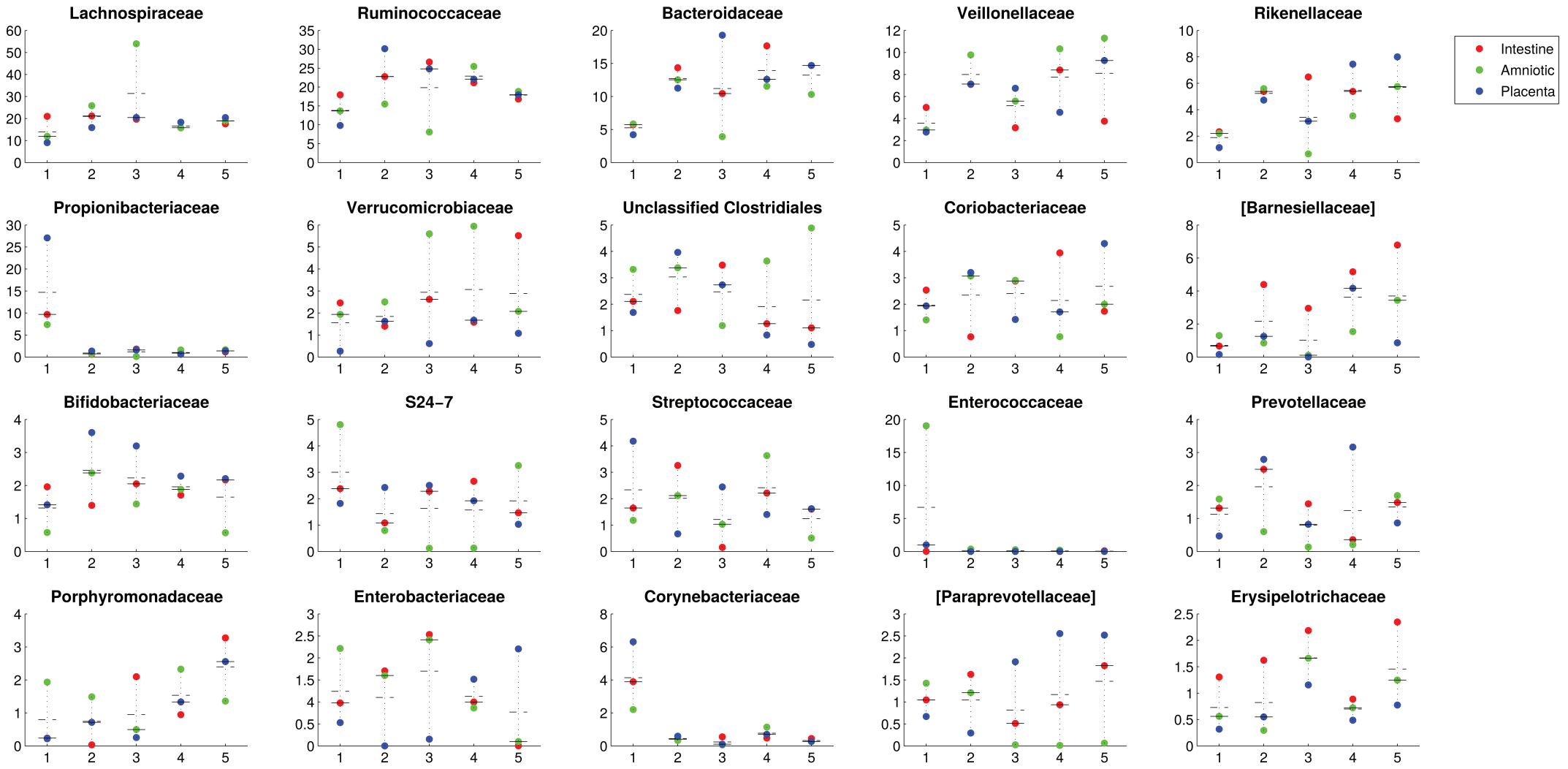
Differences in the microbial communities were evaluated on the basis of biodiversity (α-diversity) and taxonomic composition (β-diversity). Due to the limited number of samples per category, for both α- and β-diversity we applied multiple statistical procedures, flanking the commonly used statistical tests to a strategy based on distributions of distances among samples within and between experimental groups. First, the distance between a single sample against the others was calculated. Then, distances between samples belonging to the same ("intra-category" distance) or to a different ("inter-category" distance) experimental group (i.e.: tissue, dam or foetus) were calculated.

For the α-diversity, we evaluated:

- Results deriving from a MonteCarlo permutation-based, non-parametric t-test;
- Results deriving from the evaluation of the absolute difference for diversity indexes (i.e.: chao1, Shannon index, observed species and Faith's phylogenetic distance). A Mann-Whitney U-test was applied for comparing the distributions of "intra-" and "inter-category" distances.

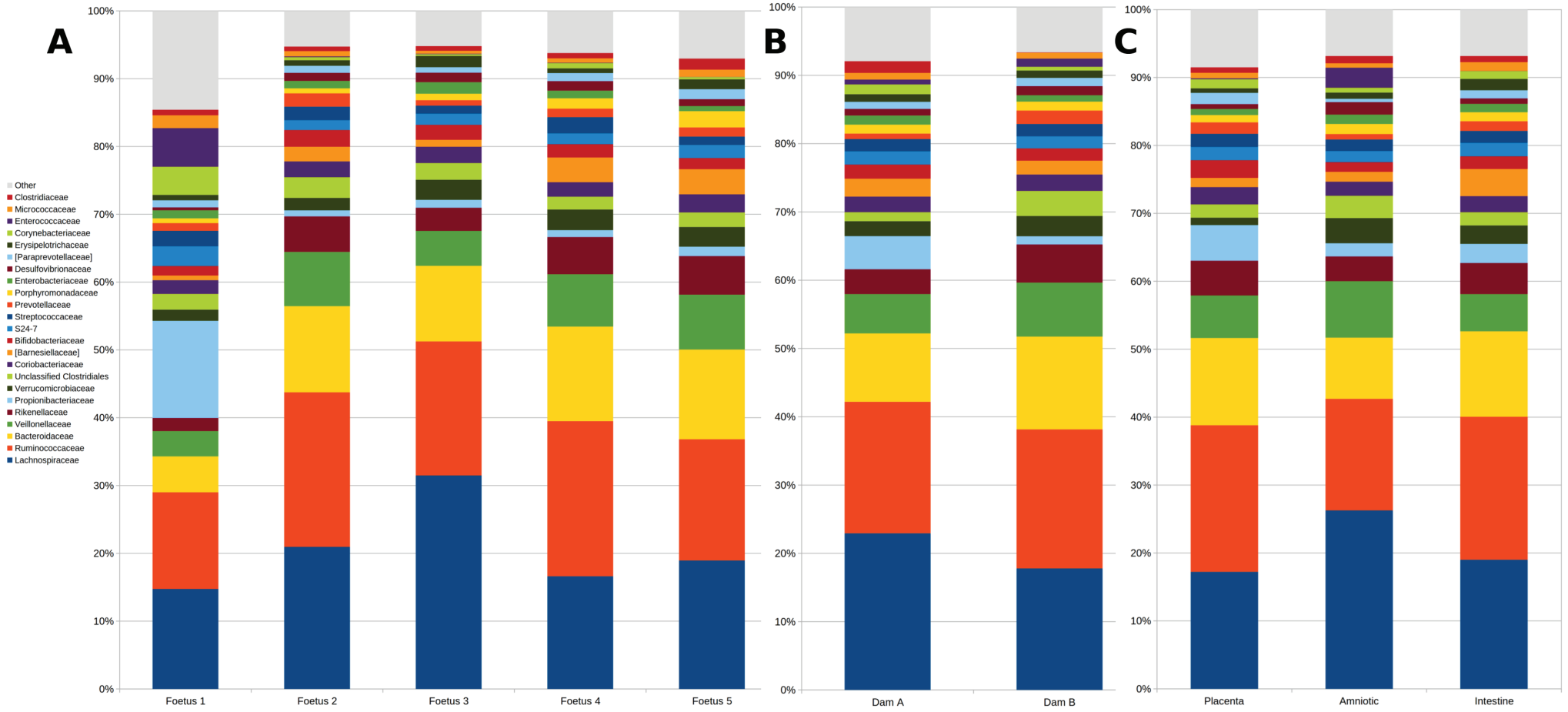
For the β-diversity, we evaluated:

- The distance matrix partitioning among sources of variations (permutation test with pseudo-F ratios) implemented in the "adonis" test contained in R package "vegan";
- The comparison between distributions of intra- and inter-category Unifrac distances (both weighted or unweighted); the function "make_distance_boxplots.py" in QIIME (http://qiime.org/scripts/make_distance_boxplots.html), implementing a permutation-based non-parametric test on distance distributions, was used to assess a significant clustering among groups.

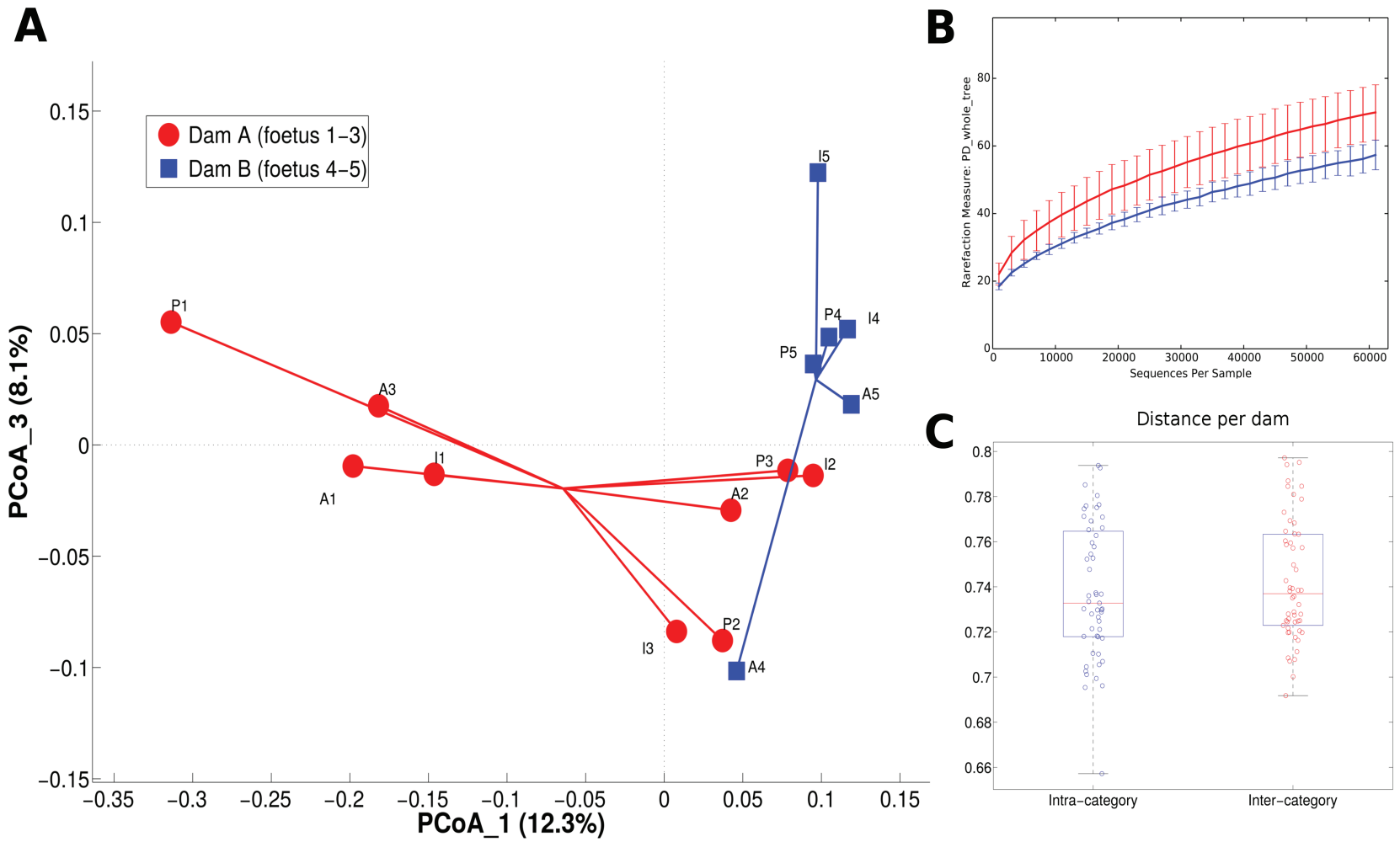


Supplementary figure 1- Relative abundance of most represented bacterial families (per foetus)

Plots show relative abundances of the main 20 bacterial families separated by foetus number. Values corresponding to the same tissue (i.e. intestine, amniotic fluid, and placenta) are represented by the same colour (red, green and blue, respectively). The solid black line represents the median of the values per each foetus, whereas the dashed black line shows the average



Supplementary figure 2 – Average relative abundance of tissue, dam and foetus microbiota
Average relative abundance of the main microbial families in (A) foetuses, (B) dams and (C) tissues. Samples considered for the analysis were 3 for each foetus, 9 for dam A, 6 for dam B and 5 for each tissue.

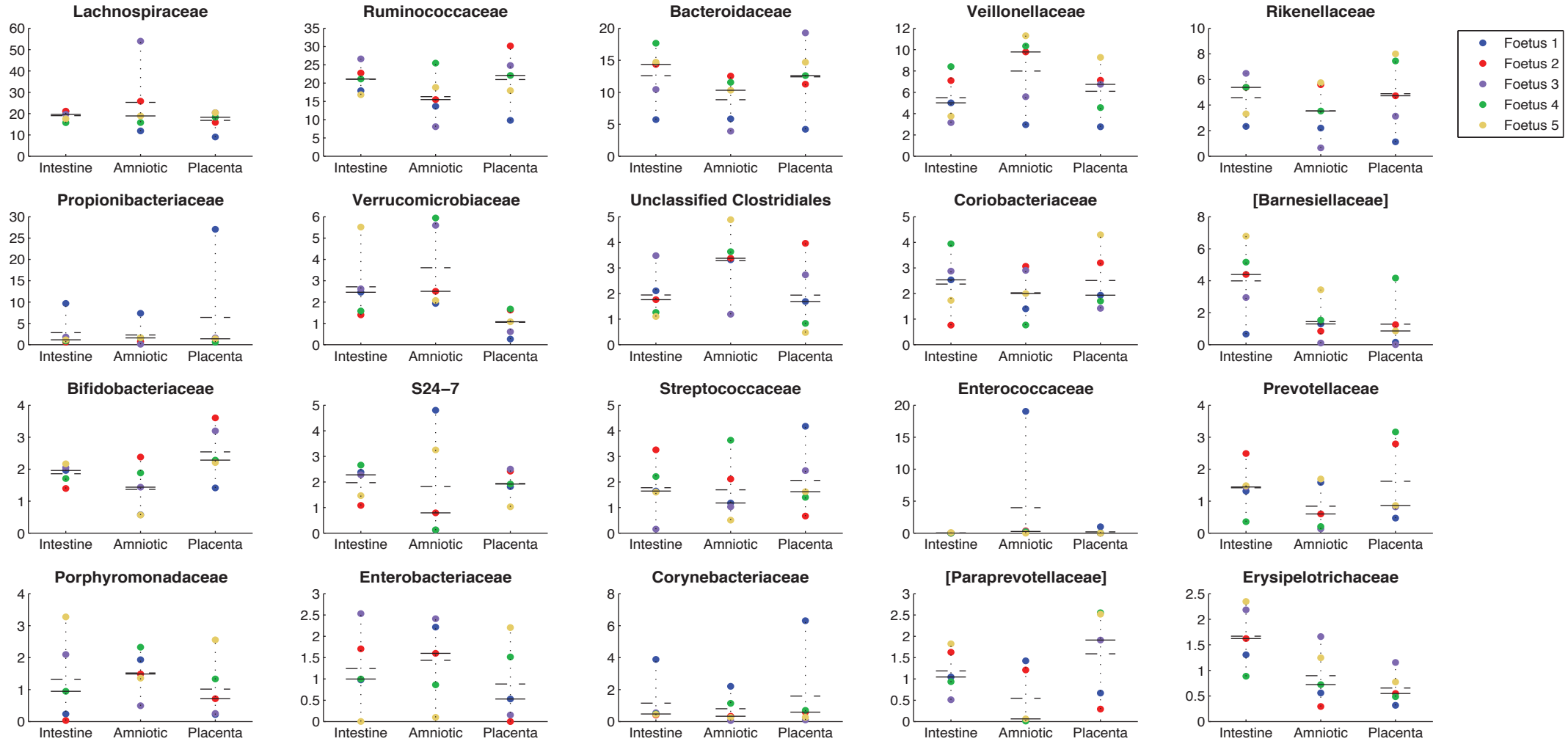


Supplementary figure 3- Dam-effect on microbiota

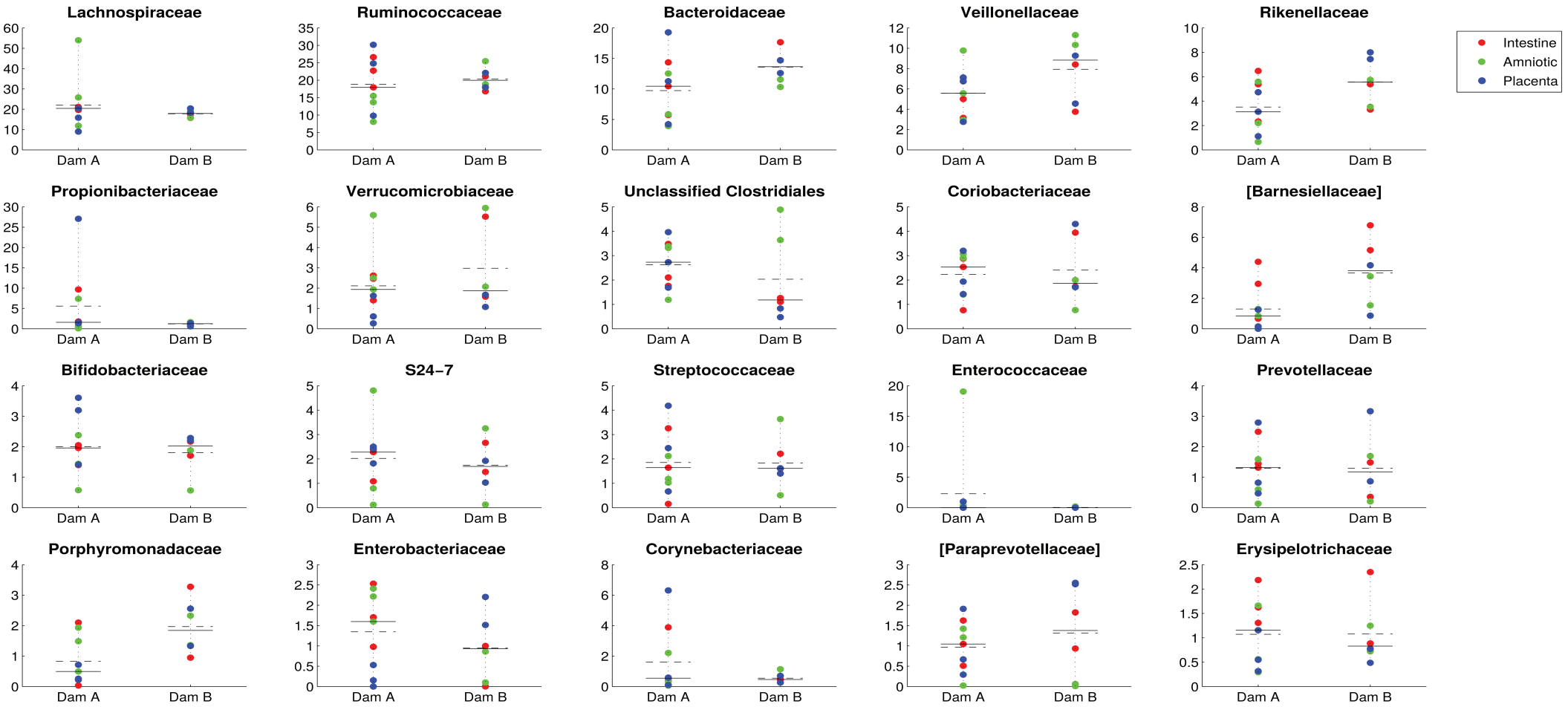
(A) Principal coordinates analysis (PCoA) of the weighted Unifrac distances; PCoA components 1 and 3 are reported. Samples are grouped based on dam.

(B) Y-axis depicts α-diversity rarefaction curves according to Faith's phylogenetic distance (PD whole tree); numbers of sequences per sample are reported on X-axis.

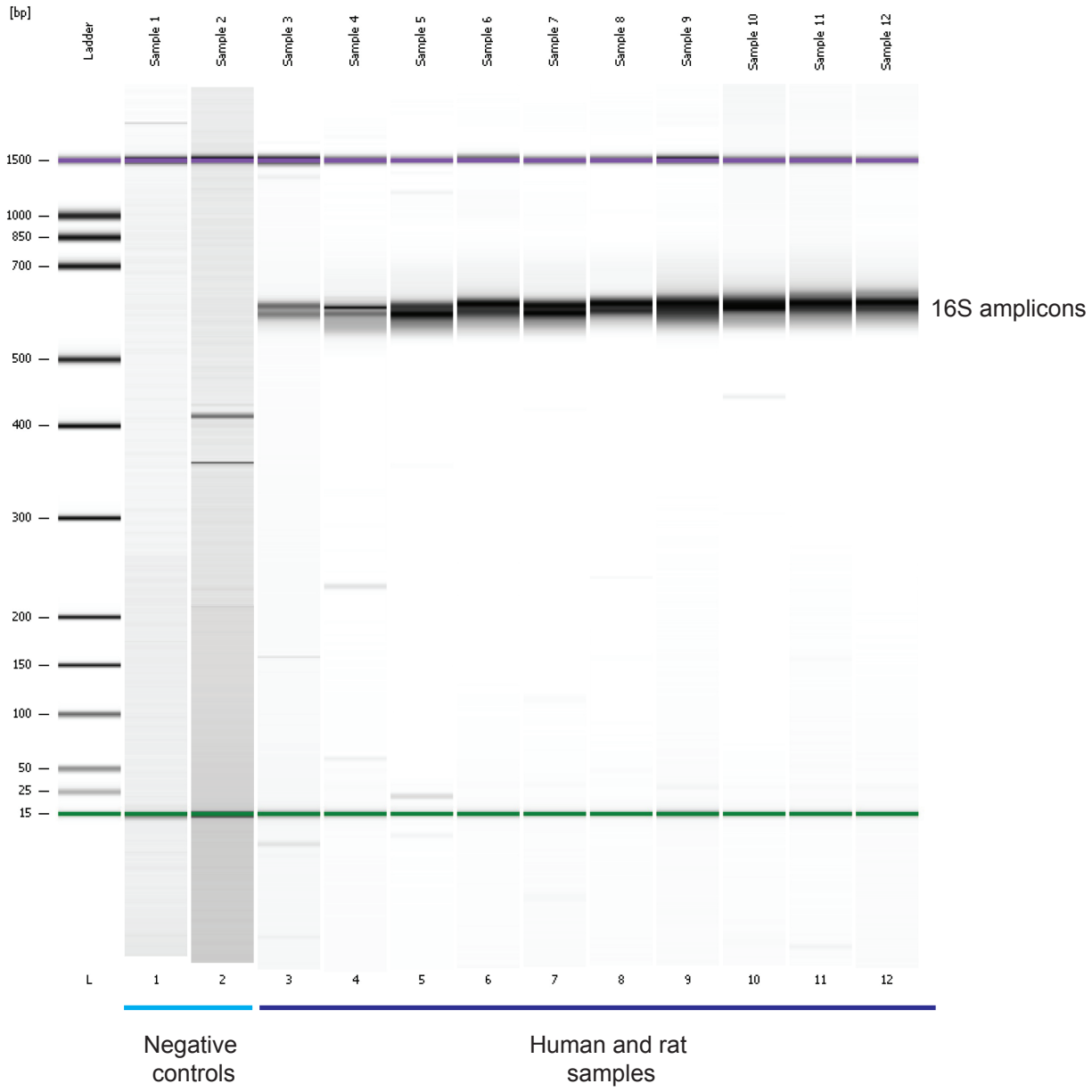
(C) Boxplots of intra- and inter-category unweighted Unifrac distances among samples; categories are based on dam.



Supplementary figure 4- Relative abundance of most represented bacterial families (per tissue)
Plots show relative abundances of the main 20 bacterial families separated by tissue type. Values corresponding to the same foetus are represented by the same colour. The solid black line represents the median of the values per each tissue, whereas the dashed black line shows the average.



Supplementary figure 5- Relative abundance of most represented bacterial families (per dam)
Plots of relative abundances of the main 20 families of the microbiota separated on dam. Relative abundance corresponding to intestine, amniotic liquid and placenta are represented in red, green and blue, respectively. The solid black line represents the median of the values per each dam, whereas the dashed black line shows the average.



Supplementary figure 6- Representative electrophoresis of 16S amplicons during library check
Agilent 2100 Bioanalyzer system was used to asses exact product size and quantity of amplicons after 16S rRNA gene library preparation, using Agilent DNA 1000 Kit, for the separation, sizing and quantification of dsDNA fragments from 25 to 1000 bp (Agilent Technologies, Santa Clara, CA, USA). First lane (L), molecular weight ladder; lines 1 and 2, negative controls (empty tubes that have been processed in parallel during tissue recovery and DNA extraction); lines 3 to 12, human and rat samples. Upon sequencing, a number of reads to the limit of detection, despite the technical concentration step, has been obtained for negative controls. Known environmental contaminants were never observed.