

1 *Supplementary Materials*

2 **Identification of vaginal microbial communities**
 3 **associated with cervical shortening in pregnant**
 4 **women at risk for preterm birth**

5 **Monica Di Paola^{1,a}, Viola Seravalli^{2,a}, Sara Paccosi³, Carlotta Linari², Astrid Parenti³, Carlotta De**
 6 **Filippo⁴, Michele Tanturli⁵, Francesco Vitali⁴, Maria Gabriella Torcia^{5,*} and Mariarosaria Di**
 7 **Tommaso²**

8 Affiliations

9 ¹Department of Biology, University of Florence, Florence Italy.

10 ²Department of Health Sciences, Division of Obstetrics & Gynecology, University of Florence,
 11 Italy.

12 ³Department of Health Sciences, University of Florence, Italy.

13 ⁴Institute of Agricultural Biology and Biotechnology, National Research Council, Pisa, Italy.

14 ⁵Department of Experimental and Clinical Medicine, University of Florence, Florence, Italy.

15 ^a These authors have equally contributed.

16 *Corresponding author: maria.torcia@unifi.it

17

18 This file includes:

19 Supplementary Table 1 and 2

20 Supplementary Figures 1, 2, 3, 4, 5

21 **Supplementary Tables**

22

23 **Supplementary Table 1** ANCOVA analysis with microbiota as grouping variable and gestational
 24 age at sampling as covariate adjustment. Output from R software version 4.0.2 for ANCOVA
 25 analysis.

ANCOVA output from R				
P value for interaction between grouping variable (microbiota) and covariate (gestational age) is 0.185				
Anova Table (Type III test)				
	SumSq	Df	Fvalue	Pr (>F)
(Intercept)	90.05	1	2.9833	0.09131
Factor(microbiota)	154.25	4	5.1099	0.02891
Gestational age	23.35	1	0.7736	0.38400
Residuals	1297.98	40		

26

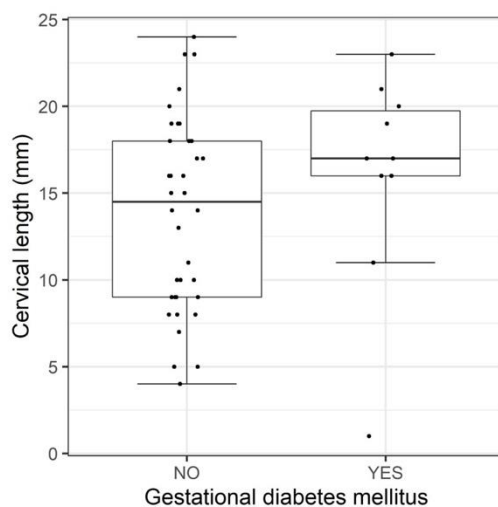
27 **Supplementary Table 2.** PERMANOVA analysis on PCoA ordination of vaginal microbiota samples.

Variable	R ²	p-value
Cervical length	0.05623	0.235
GDM	0.01743	0.544

28 **Supplementary Figures**

29

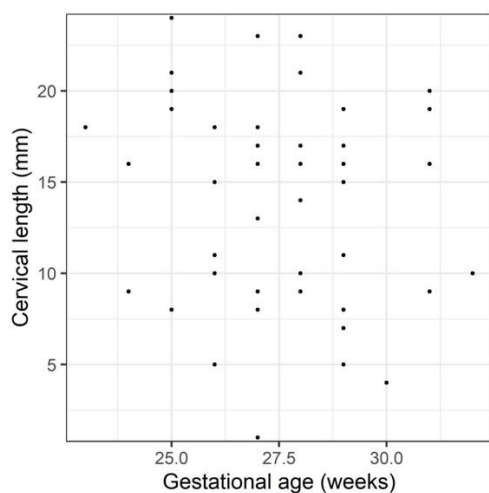
A



30

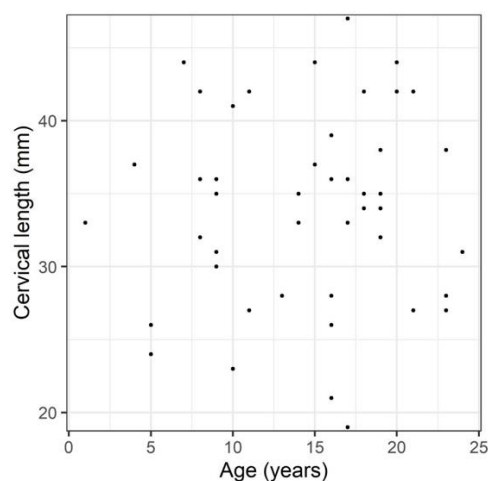
31

B



32

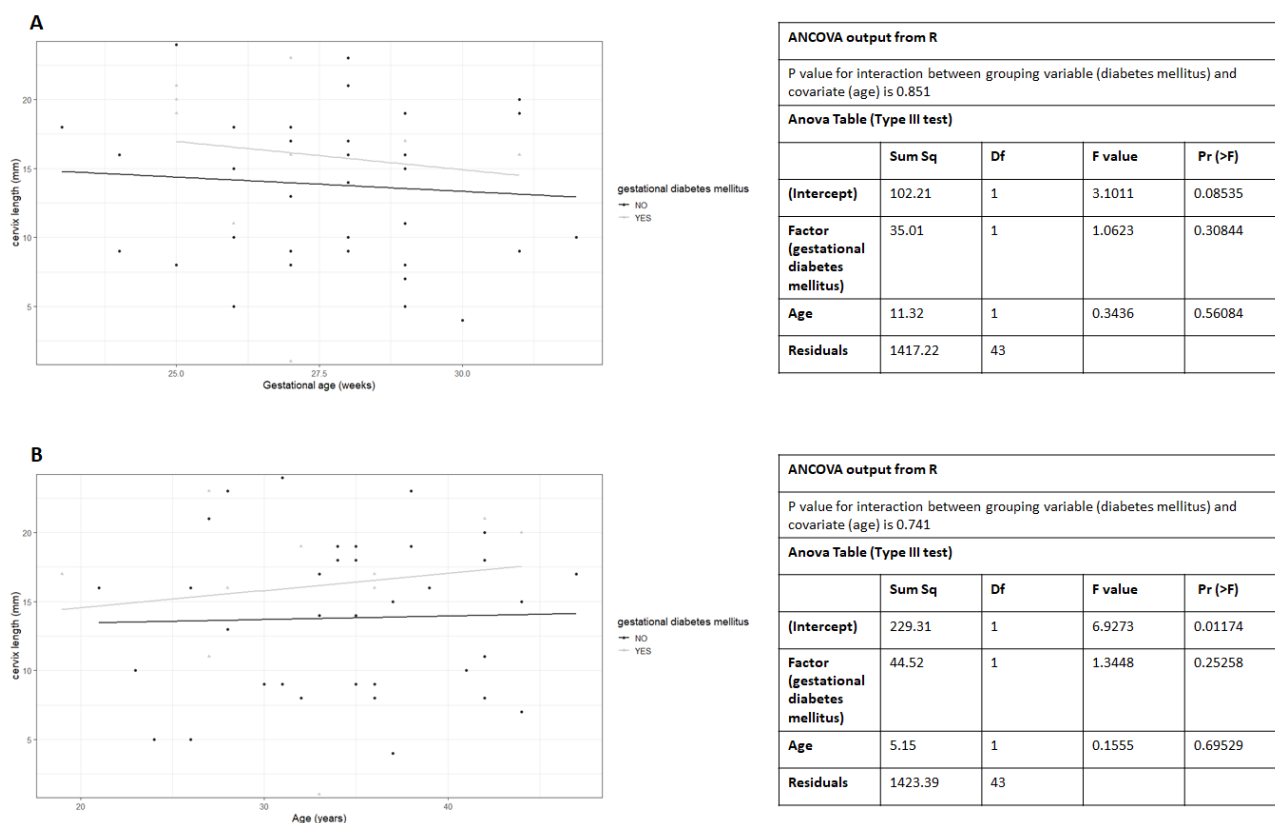
C



33 **Supplementary Figure 1. A.** Differences in cervical length (mm) between patients with or without
 34 GDM. Statistical analysis was performed by Mann Whitney test, p value = 0.161. **B.** The relationship
 35 between cervical length (mm) and gestational age (weeks) was investigated by Spearman 's rank
 36 correlation test. No statistically significant correlation was found (Spearman's rank correlation
 37 coefficient -0.108, p value = 0.475). **C.** The relationship between cervical length (mm) and age (years)
 38 was investigated by Spearman 's rank correlation test. No statistically significant correlation was
 39 found (Spearman's rank correlation coefficient 0.0451, p value = 0.766).

40

41



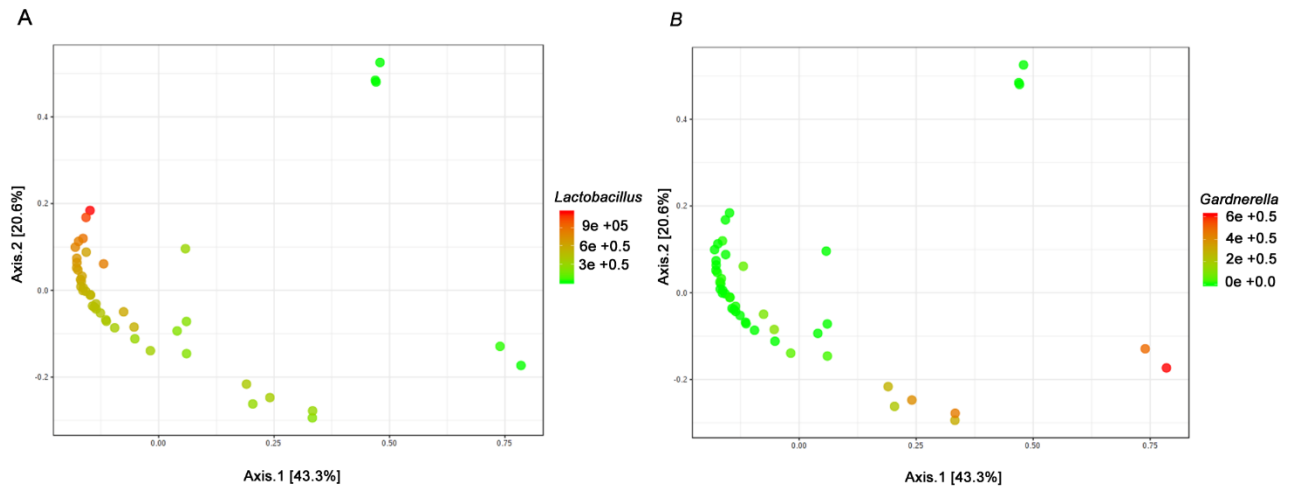
42

43 **Supplementary Figure 2.** Subgroups examination and ANCOVA analysis. **A.** The ANCOVA
 44 analysis shows that no differences in cervical length between women with or without gestational
 45 diabetes mellitus were found when the data were adjusted for gestational age at sampling (p value
 46 0.308). The same analysis shows that no correlation exists between gestational age and cervical
 47 length in each subgroup considered. In addition, the ANCOVA analysis revealed that no interaction
 48 between the covariate (gestational age at sampling) and grouping variables (gestational diabetes
 49 mellitus) exist (p value 0.851). Left panel: scatter plot with regression lines for the two groups
 50 (diabetes mellitus yes or no); right panel: output from R software version 4.0.2 for ANCOVA
 51 analysis. **B.** The ANCOVA analysis shows that no differences in cervical length between women with
 52 or without gestational diabetes mellitus were found when the data were adjusted for gestational age
 53 at sampling (p value 0.253). The same analysis shows that no correlation exists between gestational
 54 age and cervix length in each subgroup considered. In addition, the ANCOVA analysis revealed that
 55 no interaction between the covariate (gestational age at sampling) and grouping variables
 56 (gestational diabetes mellitus) exist (p value 0.741). Left panel: scatter plot with regression lines
 57 for the two groups (diabetes mellitus yes or no); right panel: output from R software version 4.0.2 for
 58 ANCOVA analysis.

59

60

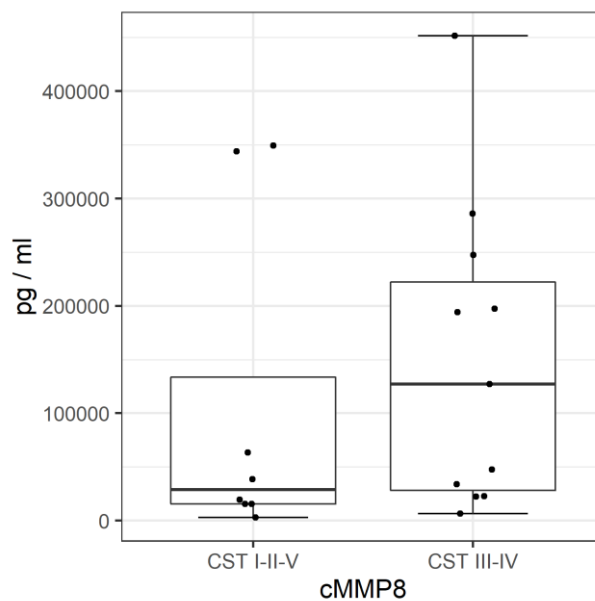
61



62

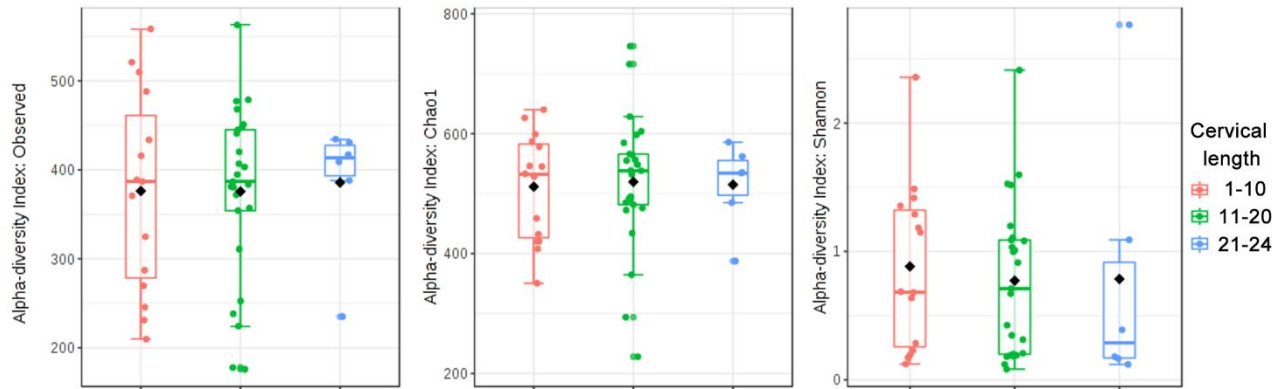
63 **Supplementary Figure 3. PCoA ordination, based on Bray Curtis dissimilarities, correlated with A.**
 64 *Lactobacillus spp.* and **B. Gardnerella spp.** abundances. Colors from red to green indicate decreasing
 65 abundance of bacterial genus. (PERMANOVA 999 permutations; R20.740 p value< 0.001).

66



67

68 **Supplementary Figure 4. Matrix metalloprotease 8 (MMP-8) concentration in vaginal samples of a**
 69 **subgroups of women.** Data from pregnant women with vaginal microbiota of CST-I, CST-II and
 70 CST-V (*Lactobacillus*- dominated community) were compared with data from women with CST-III (*L.*
 71 *iners*- dominated) and CST-IV (*Lactobacillus*- depleted). Data are presented as box and whisker plots,
 72 with boxes extending from the 25th to 75th percentile and horizontal lines representing the median.
 73 Whiskers extend 1.5 times the interquartile range from the 25th and 75 th percentile. Statistical
 74 analysis, performed by Mann-Whitney assay did not reveal significant differences.



75
76
77
78

Supplementary Figure 5. Alpha diversity measures. Box plots of observed OTUs, Chao 1, and Shannon index according cervical length classification. ANOVA test resulted not significant for all comparisons.

79



© 2020 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

80