



**HAL**  
open science

## **Female reproductive tract microbiota varies with MHC profile**

Sarah Leclaire, Mandar Bandekar, Melissah Rowe, Jarmo Ritari, Annalaura Jokiniemi, Jukka Partanen, Pia Allinen, Liisa Kuusipalo, Jukka Kekäläinen

► **To cite this version:**

Sarah Leclaire, Mandar Bandekar, Melissah Rowe, Jarmo Ritari, Annalaura Jokiniemi, et al.. Female reproductive tract microbiota varies with MHC profile. *Proceedings of the Royal Society B: Biological Sciences*, 2024, 291 (2033), 10.1098/rspb.2024.1334 . hal-04764768

**HAL Id: hal-04764768**

**<https://hal.science/hal-04764768v1>**

Submitted on 4 Nov 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Female reproductive tract microbiota varies with MHC profile

Sarah LECLAIRE<sup>1</sup>, Mandar BANDEKAR<sup>2</sup>, Melissa ROWE<sup>3</sup>, Jarmo RITARI<sup>4</sup>, Annalaura JOKINIEMI<sup>2</sup>, Jukka PARTANEN<sup>4</sup>, Pia ALLINEN<sup>5</sup>, Liisa KUUSIPALO<sup>6</sup>, Jukka KEKÄLÄINEN<sup>2</sup>

<sup>1</sup>Centre de Recherche sur la Biodiversité et l'Environnement (CRBE), UMR5300, Université Toulouse, CNRS, IRD, Toulouse INP, 118 rte de Narbonne, 31062 Toulouse, France

<sup>2</sup>Department of Environmental and Biological Sciences, University of Eastern Finland, PO Box 111, 80101 Joensuu, Finland

<sup>3</sup>Department of Animal Ecology, Netherlands Institute of Ecology (NIOO-KNAW), 6700 AB, Wageningen, The Netherlands

<sup>4</sup>Finnish Red Cross Blood Service, Research and Development, Haartmaninkatu 8, 00290 Helsinki, Finland

<sup>5</sup>Ovumia Kuopio, Ajurinkatu 16, 70110 Kuopio, Finland

<sup>6</sup>North Karelia Central Hospital, Tikkamäentie 16, 80210 Joensuu, Finland

## Abstract

Numerous studies have shown that a healthy reproductive tract microbiota is crucial for successful reproduction and that its composition is influenced by various environmental and host factors. However, it is not known whether the reproductive microbiota is also shaped by the major histocompatibility complex (MHC), a family of genes essential to differentiate “self” from “non-self” peptides to initiate an adaptive immune response. We tested the association between the follicular fluid microbiome and MHC genes in 27 women. Women with higher MHC diversity had a higher microbiome diversity, characterized by bacteria commonly associated with vaginal dysbiosis. Women with similar MHC genes were also similar in their microbiome composition, indicating that MHC composition may be a key factor in determining the bacterial assemblage in reproductive tract. Finally, the composition of the follicular fluid microbiome was similar to the vaginal microbiome, suggesting that numerous bacteria of the vagina are true inhabitants of the follicular fluid or that vaginal microbiota contaminated the follicular fluid microbiota during transvaginal collection.

33 Collectively, our results demonstrate the importance of host genetic factors in shaping the  
34 reproductive microbiota of women, and open the door for further research on the role of microbiota  
35 in mediating MHC-related variation in reproductive success.

36

37

## 38 **Introduction**

39 The reproductive microbiome is emerging as an important determinant of male and female  
40 reproductive health and performance [1–4]. For example, certain bacteria of the semen microbiota  
41 are associated with decreased sperm quality in humans and livestock (e.g., bulls, boars) [4–7], while  
42 the composition of the vaginal microbiota is associated with egg number variation in hens [8] and  
43 reproductive outcome in ewes [9]. In addition, in women, microbial imbalance in the reproductive  
44 tract is associated with a range of pregnancy disorders, including polycystic ovary syndrome,  
45 recurrent implantation failure and spontaneous preterm birth [10–12]. Given the importance of  
46 reproductive microbiota for successful host reproduction, it is critical to unravel the factors that  
47 shape its variation in microbial composition and function.

48         The reproductive microbiota is the focus of intense research, especially in women [13–16].  
49 While the vagina was the focus of early microbiology studies, the upper reproductive tract  
50 (endometrium, fallopian tubes, ovaries) has recently been shown to be also colonized by bacteria,  
51 even in the absence of symptomatic infection, thereby challenging the traditional assumption that it  
52 is sterile [17–19]. The abundance, diversity and composition of women’s reproductive microbiota is  
53 in fact a continuum along the reproductive tract, changing from the vagina to the ovaries [20]. For  
54 instance, the vaginal microbiota, which mostly comprises *Lactobacilli* in healthy women, is typically  
55 less diverse, but more abundant, than upper sites [20]. The increasing literature on the determinants  
56 of the women’s reproductive microbiota remains, however, mostly limited to the vaginal microbiota.  
57 Women’s vaginal microbiota varies with age, menstruation cycle, sexual practices, hormone levels,  
58 disease state and host genetics [21–25]. Among the host genetic factors, immune-related genes have  
59 been of special interest, because of their vital role in maintaining the intimate relationship between  
60 the host and its commensal microorganisms [26]. For instance, genetic polymorphism in interferon,  
61 interleukins and toll-like receptors are associated with the abundance of specific vaginal bacteria  
62 [22,27,28].

63         Among the host immunogenetic factors, the major histocompatibility complex (MHC; called  
64 HLA in humans) may play an important role in shaping the host-associated microbiomes [29]. The  
65 MHC is one of the most polymorphic group of genes in vertebrates and encompass the classical MHC  
66 class I and class II genes that encode for cell-surface proteins that are crucial in the recognition of  
67 “self” and “non-self” antigens by T-cells. While MHC class I molecules mainly process peptide  
68 fragments derived from intracellular proteins, MHC class II molecules predominantly process  
69 extracellular peptide fragments. The MHC class II plays therefore a traditional role in the defence  
70 against extracellular bacterial pathogens [30], but is also increasingly recognized to shape tolerance

71 towards symbiotic bacteria [26,29,31,32]. This tolerogenic effect may be, in part, mediated by the  
72 immunoglobulin A (IgA), which is a key component of adaptive immunity at mucosal tissues, such as  
73 the vagina or the gut epithelium, and which is known to promote the proliferation of beneficial  
74 bacteria [33–35]. Importantly, in several vertebrates including humans, different characteristics of  
75 the MHC class II genotype are associated with inter-individual differences in the microbiota [36–40].  
76 For instance, certain MHC class II motifs are specific to certain bacteria of the gut microbiota of  
77 sticklebacks (*Gasterosteus aculeatus*), reddish-gray mouse lemurs (*Microcebus griseorufus*) and  
78 neotropical bats (*Artibeus jamaicensis*) [41–43]. In addition, individuals with higher MHC class II  
79 diversity have lower gut microbiota diversity in sticklebacks and lemurs [41,42]. However, the  
80 association between the MHC and the reproductive tract microbiota is unknown in both humans and  
81 other species.

82           Here, we investigated the potential relationship between reproductive tract microbiota and  
83 MHC class I and class II variation in women. We focused on the microbiota of the follicular fluid,  
84 which forms the microenvironment of the oocyte during its development and maturation, and which  
85 bacterial composition, albeit poorly described, has been shown to be a key element determining  
86 pregnancy success [44–47] [but see 48,49]. Specifically, we tested whether the  $\alpha$ -diversity (within-  
87 individual microbial diversity) of the human reproductive microbiota co-vary with MHC diversity.  
88 Furthermore, as two individuals with the same MHC genotypes are hypothesized to have similar  
89 immune control of their microbiota, we predicted that similarity in microbiome composition  
90 between individuals ( $\beta$ -diversity) would covary with MHC similarity. We studied associations  
91 between microbiome and both MHC class I and class II genes; though we primarily expected to find  
92 associations between the microbiota and MHC class II genes because of their important role in the  
93 recognition of antigenic peptides derived from extracellular bacteria.

94

## 95 **Material and methods**

### 96 *Sample collection*

97 We investigated the association between follicular fluid microbiome and HLA profiles in 27 women  
98 who underwent transvaginal follicular aspiration for *in vitro* fertilization. Three women were sampled  
99 twice, with 6-16 months intervals between sampling events, leading to a total of 30 follicular fluid  
100 samples. Women were recruited via three fertility clinics, located in the Eastern Finland: North  
101 Karelia Central Hospital (Joensuu, Finland, n = 20 samples), Kuopio University Hospital (Kuopio,  
102 Finland, n = 4 samples) and InOva fertility clinic (Kuopio, Finland, n = 6 samples), between April 2017  
103 and December 2020. The women's mean age was  $30.7 \pm 0.96$  (s.e.) years. All women were of

104 apparent Finnish origin, except for one woman of Asian origin. The inclusion of this individual did not  
105 affect our results (table S1 and S2). We therefore included this participant in our final dataset. All  
106 samples were collected during the ovulation phase. Prior to follicular fluid collection, follicle  
107 maturation was hyperstimulated with follicle-stimulating hormone, and premature ovulation was  
108 prevented using a gonadotrophin-releasing hormone antagonist. When the diameter of the largest  
109 follicle reached 18–20 mm, human chorionic gonadotrophin was administered. Transvaginal follicular  
110 punctures were performed under local anesthesia, using ultrasound guidance. The collected follicular  
111 fluid was centrifuged at  $500 \times g$  for 10 min, and supernatant aliquoted and stored in liquid nitrogen  
112 for later use. Nine of the women had been diagnosed with diseases associated with female factor  
113 infertility (i.e., anovulation, polycystic ovarian syndrome, or endometriosis), whereas 18 women did  
114 not have any female factor infertility diagnosis. All the women donated a blood sample for genetic  
115 analyses.

116 For this study, follicular fluid samples were obtained transvaginally, the only reasonable  
117 method for collecting follicular fluid in living patients. However, transvaginal collection increases the  
118 risk of sample contamination by vaginal bacteria [19]. Consequently, we collected both follicular fluid  
119 samples and vaginal swabs from four additional women recruited from the InOva fertility clinic  
120 (Kuopio, Finland) in March-April 2022. Before follicular fluid collection, a sterile cotton swab  
121 (FLOQSwabs, COPAN, California) was inserted in the vagina to collect mucus from the area of the  
122 transvaginal puncture. After the collection, swabs were transferred in sterile 5 ml cryotubes and  
123 stored in liquid nitrogen. We did not carry out HLA imputation for these four women, who were  
124 sampled later than the other women. Instead, these paired samples were only used to investigate  
125 the possibility that the follicular fluid samples reflected the vaginal microbiome. At the same time,  
126 we hoped to understand the potential for contamination of follicular fluid samples from the blood, as  
127 the transvaginal collection of samples may involve a small amount of bleeding and thus may also be  
128 prone to bacterial contamination from the blood. Therefore, we performed microbiome analysis of  
129 the blood collected during the ovulation phase for six of the study participants.

130

### 131 *Microbiome analyses*

132 All DNA extractions were carried out in a sterile environment. We added 50  $\mu\text{L}$  of lysozyme (10  
133 mg/ml) to each sample of 500  $\mu\text{L}$  follicular fluid and to each vaginal swab, and 20  $\mu\text{l}$  of lysozyme (10  
134 mg/ml) to each blood sample. Samples were incubated at 37°C for 60 minutes. Next, we added 50  $\mu\text{L}$   
135 of proteinase K (20 mg/ml) and 50  $\mu\text{L}$  of RNase A to each follicular fluid and vaginal sample, and 20 $\mu\text{l}$   
136 of proteinase K and 20 $\mu\text{l}$  RNase A to each blood sample. We mixed samples by vortexing, and  
137 incubated samples for 2-minutes at room temperature (22°C). Following this chemical lysis step, we

138 further processed all samples using PureLink™ Genomic DNA Mini Kit (Invitrogen) following the kit  
139 protocol. Extracted DNA was stored at -20°C until later analysis. Follicular fluid DNA was extracted in  
140 five different batches. To control for the potential batch effect, one sample was repeatedly extracted  
141 in each batch (three times in the first batch and once in each other batch). In addition, DNA  
142 extraction was repeated twice for five other follicular fluid samples to ensure that DNA extraction  
143 was repeatable. Because commercial extraction kits can contain microbial DNA [50], we also included  
144 four additional 'blank' DNA extraction controls.

145 To characterize the bacterial community present in each sample, we performed 16S rRNA  
146 amplicon high-throughput sequencing targeting the V4 region (primers: 515F  
147 [GTGCCAGCMGCCGCGGTAA; 51] and 806R [GGACTACHVGGGTWTCTAAT; 51]. Extracted DNA was  
148 amplified and sequenced at Genome Quebec (Montréal, Canada). We included three PCR blank  
149 controls to control for the possibility of microbial contamination during PCR and sequencing. Samples  
150 were sequenced as 250bp paired-end reads using an Illumina MiSeq (n = 45 follicular fluid samples, 4  
151 vaginal swab samples, 6 blood samples, 4 extraction blanks and 3 PCR blanks).

152 We analyzed sequence data using the FROGS 4.1 pipeline (Escudié et al., 2018). Briefly,  
153 demultiplexed reads were merged using VSearch [52], with a minimum of 10bp in the overlap region  
154 and a maximum mismatch rate of 10% in the overlap region. Primers were trimmed using cutadapt  
155 and a maximum error rate of 10% [53]. Unmerged reads were excluded. Amplicon sequence variants  
156 (ASVs) were formed using the swarm denoising method with the fastidious method and with  $d = 1$   
157 [54]. Chimeras were detected using VSearch (Rognes et al., 2016). In total of 15.4% of the ASVs  
158 (representing 0.9 % of the reads) were considered as chimeras and were thus removed. We also  
159 removed ASVs that had only one read. Taxonomic assignment was then performed using the RDP  
160 classifier and the 16S SILVA v138.1 database (Quast et al., 2012). ASVs sharing same taxonomy with  
161 at least 99% sequence similarity and 99% alignment coverage were aggregated together. We  
162 removed ASVs that were taxonomically classified as mitochondria, chloroplasta or Archaea, as well as  
163 each ASV that totaled fewer than 0.0005 % of all reads, as suggested by Bokulich et al. (2013) and  
164 those that were detected in two or fewer samples. Finally, we used the isContaminant function with  
165 the prevalence option and threshold value of 0.5 (package decontam in R) to remove potential  
166 contaminants from the data set [56]. In this approach, ASVs are classified as contaminants if they are  
167 more frequently present in negative controls than in true samples [56]. The isContaminant function  
168 was run first with the DNA extraction blank controls, and second with the PCR blank controls as  
169 negative controls. Six ASVs were found to be contaminants and were removed from the dataset.  
170 They belonged to families *Burkholderiaceae*, *Pseudomonadaceae*, *Rhodobacteraceae* and *Bacillaceae*.  
171 The microbiome of the blood samples differed from that of the follicular fluid samples (permanova:

172  $F_{1,11} = 2.65$ ,  $r^2 = 0.21$ ,  $P = 0.001$ ), and had very low biomass and richness. In blood samples, number of  
173 reads ranged between 37 and 1972 reads and richness was low (number of ASVs per blood sample  
174 ranged between 4 and 14 ASVs; in total, 26 ASVs were detected in blood samples). Therefore, to  
175 remove bacterial contamination from the blood, we run also the isContaminant function with blood  
176 samples as negative controls. Six additional ASVs were found to be contaminants and were removed.  
177 They belonged to families *Maroxellaceae*, *Staphylococcaceae*, *Veillonellaceae*, *Xanthomonadaceae*  
178 and *Rhodobacteraceae*. We did not use the vaginal samples as negative controls, because the  
179 bacterial composition of the reproductive tract is a continuum along the reproductive tract [20] and  
180 therefore several abundant residents of the vagina might be true inhabitants in follicular fluid. For  
181 instance, *Lactobacillus*, which is the predominant bacteria genera in the vagina, is also an abundant  
182 genera in follicular fluid [45–47]. In the four women with both follicular fluid and vaginal swab  
183 samples, we found no differences in the composition or diversity of the microbiome between sample  
184 types (permanova:  $F_{1,6} = 330$ ,  $r^2 = 0.14$ ,  $P = 0.54$  and paired t-test:  $t_3 = 0.30$ ,  $P = 0.80$ , respectively; Fig.  
185 S1).

186 We found that the number of reads of the ASV belonging to *Escherichia-Shigella* was similar  
187 between follicular fluid samples and blank controls (mean  $\pm$  SE:  $2474 \pm 229$  vs  $1892 \pm 337$  reads,  
188 respectively; Wilcoxon rank sum test:  $W = 188$ ,  $P = 0.40$ ). Given that it has been suggested that reads  
189 attributed to *Escherichia-Shigella* may represent remnant DNA from the *Escherichia coli* used to  
190 produce *Taq* DNA polymerase [57], the ASV belonging to *Escherichia-Shigella* was excluded from our  
191 final dataset. In addition, when focusing on the sample that was extracted repeatedly, we observed  
192 high variation in the relative abundance of *Acinetobacter* between extraction replicates (Fig. S2).  
193 When considering all follicular fluid samples, the relative abundance of *Acinetobacter* was found to  
194 vary with the DNA extraction batch (Kruskal-Wallis test:  $\chi^2_4 = 21$ ,  $P = 0.0004$ ; Fig. S2). *Acinetobacter*  
195 are frequent cleanroom contaminants because they resist sterilization procedures and disinfectants  
196 [58,59]. We therefore excluded all ASVs belonging to *Acinetobacter* from the analyses ( $n = 8$  ASVs).  
197 After filtering, we obtained 118 ASVs in follicular fluid samples and mean  $\pm$  SE:  $61\,012 \pm 11\,441$   
198 sequences per follicular fluid sample (median: 21 377 sequences; range: 1 689 – 285 147 sequences)  
199 The most abundant ASVs belonged to the genus *Lactobacillus* and included *Lactobacillus iners*,  
200 *Lactobacillus jensenii*, *Lactobacillus acidophilus* / *L. crispatus*, *Lactobacillus gasseri* / *L. johnsonii*.  
201 Other abundant taxa included *Gardnerella vaginalis*, *Sneathia amnii*, *Prevotella sp*, *Atopobium*  
202 *vaginae*, *Kocuria sp.* and *Alishewanella sp.* (Fig. 1). DNA extraction replicates were repeatable  
203 (permanova testing the effect of replicate identity on microbiome composition:  $F_{5,11} = 5.79$ ,  $r^2 = 0.73$ ,  
204  $P < 0.001$ ; Fig. S3) and were therefore averaged.



205 Because phylogenetically-related bacteria are likely to share similar MHC-binding peptides  
206 [60], we integrated phylogeny into the  $\alpha$ - and  $\beta$ -diversity measures. To analyse  $\alpha$ -diversity of  
207 follicular fluid microbiota, we used weighted phylogenetic diversity (Faith's PD) based on the  
208 unrooted phylogenetic tree (function "pd" in package picante in R). Faith's PD is the sum of total  
209 phylogenetic length of ASVs in each sample. Unrooted Faith's PD has been shown to be one of the  
210 diversity measures that is highly effective in distinguishing community types in women vaginal  
211 microbiota [61]. In our dataset, ANCOM-BC analyses at the species level [62] showed that follicular  
212 fluid microbiomes with higher Faith's diversity were characterized by higher abundance of  
213 *Gardnerella vaginalis*, *Atopobium vaginae*, *Sneathia amnii*, *Prevotella sp* and *Megasphaera sp*, and  
214 lower abundance of *Lactobacillus jensenii*, *Lactobacillus acidophilus* / *L. crispatus* and *Lactobacillus*  
215 *gasseri* / *L. johnsonii*.

216 To analyse  $\beta$ -diversity, we used the phylogenetic isometric log-ratio transform (PhiLR), which  
217 allows analysis of compositional data where the parts can be related through a phylogenetic tree  
218 (Silverman et al., 2017). PhiLR has been suggested to be a replacement for the weighted UniFrac  
219 distances when accounting for the compositional nature of the data is needed (Gloor et al., 2017).  
220 ASV sequences were aligned using Decipher (Wright, 2015), and a maximum-likelihood tree  
221 constructed using the package phangorn (Schliep, 2011). This tree was used as a reference from  
222 which the Euclidean distances were calculated on phylogenetic isometric log-ratio transformed ASV  
223 abundance (i.e., the PhiLR transformation) between samples (Egozcue et al., 2003; Gloor et al., 2017)  
224 using the package PhiLR (Silverman et al., 2017). In the PhiLR transform, we added a pseudocount of  
225 0.1 reads.

226 While there have been considerable debate concerning rarefaction in microbiome datasets  
227 [63–65], we rarefied the dataset (without replacement) because of the wide variability in our library  
228 sizes, ranging from 1 808 to 259 887 sequences in follicular fluid samples (Fig. S4). However, because  
229 rarefaction can result in a substantial loss of reads [63] and can introduce a stochastic component to  
230 the analysis (because the rarefied taxa count table depends on the specific rarefaction replicate), we  
231 used bootstrapping (i.e., multiple rarefactions;  $n = 500$  rarefaction rounds), which has been proposed  
232 to recover some of the information lost through a single rarefaction [65]. Our results are robust to  
233 the library size used when rarefying (e.g. rarefaction to the smallest library size, 5000 reads or 10 000  
234 reads), except for one result which is discussed below (Table S1 and Table S2). We present results  
235 using a rarefied dataset to the smallest library size (i.e., 1689 sequences).

236

237 *HLA genotyping*

238 Details about HLA genotyping can be found in the supplementary materials. Briefly, genomic DNA  
239 was extracted and genotyped on an Illumina GlobalScreeningArray-24v2-0 + Multi-Disease beadchip  
240 at the Institute for Molecular Medicine Finland (FIMM). All genotypes were named with Illumina's  
241 GenomeStudio 2.0.3 software. PLINK v1.90b6.6 [66] was used to manage and filter the genotyping  
242 data. We carried out imputation of three classical HLA class I genes (*HLA-A*, *-B*, and *-C*) and four  
243 classical HLA class II genes (*HLA-DRB1*, *-DQA1*, *-DQB1*, and *-DPB1*) at four-digit (i.e., protein level)  
244 resolution.

245 HLA diversity of the women was calculated based on (1) the number of heterozygous sites  
246 and (2) the mean Sandberg divergence over the seven imputed HLA genes per individual. The  
247 Sandberg divergence is the pairwise Euclidean distance between the two protein sequences  
248 represented by the five physico-chemical z-descriptors [67]. For HLA genes, the Sandberg distance  
249 has been found to correlate with the total number of peptides bound by the alleles [68]. The  
250 Sandberg divergence at each HLA gene was calculated using the DistCalc function in the MHCtools  
251 package in R [69]. For each woman, the Sandberg divergence was then averaged across the seven  
252 genes. The number of heterozygous sites was correlated to the Sandberg divergence (Spearman  
253 correlation test:  $S = 837$ ,  $r = 0.74$ ,  $P < 0.0001$ ). Dissimilarity between each women pair was estimated  
254 based on (1) a pure shared/non-shared allele basis and (2) mean Sandberg dissimilarity between HLA  
255 alleles. HLA dissimilarity based on a pure shared/non-shared allele basis (hereafter called allelic  
256 dissimilarity) was calculated as  $100 * (1 - (\text{the number of shared alleles} / \text{the total number of alleles}))$ .  
257 For each HLA gene, Sandberg dissimilarity was calculated by constructing a UPGMA dendrogram of  
258 alleles (function `hclust` in R), which represents clusters of functionally similar HLA sequences. Then  
259 HLA dissimilarity between two women was estimated using the tree distance (function `treedist` in  
260 package `vegan`) [70]. In this method, the alleles of the two women are combined in a common  
261 dendrogram and the distance between two women corresponds to how much of the dendrogram  
262 height is shared between the women and how much of it is unique. Allelic dissimilarity was  
263 correlated to Sandberg dissimilarity (Mantel test:  $r = 0.91$ ,  $P = 0.001$ ). Given the limited sample size  
264 and high allelic variations across women, we did not test the associations between specific HLA  
265 alleles and specific ASVs. Because correlations between HLA characteristics and microbiome diversity  
266 might be the result of background genetic diversity, we also characterized the whole genome  
267 heterozygosity of the participants and the whole-genome dissimilarity between each female-female  
268 combination (methods are described in the supplementary materials)

269

270 *Statistical analyses*

271 We performed all statistical analyses using the R 4.3.1 statistical software (R Core Team, 2020). To  
272 determine whether HLA diversity (heterozygosity or Sandberg divergence) was associated with  
273 microbiome  $\alpha$ -diversity (Faith's PD), we used correlation tests (Spearman's correlation tests for the  
274 ordinal heterozygosity variable and Pearson's correlation test for the continuous Sandberg  
275 divergence variable). For each analysis (heterozygosity or Sandberg divergence), we performed 500  
276 iterations of the model, with each iteration corresponding to a single rarefaction run round [65]. In  
277 addition, for each of the three women who were sampled twice, we included just one sample  
278 (chosen at random) in each rarefaction round to avoid pseudoreplication. We report the mean of the  
279 statistic parameters over the 500 tests. We did not control for the DNA extraction batch, because it  
280 had no effect on the microbiome  $\alpha$ -diversity (anova test:  $F_{3,23} = 0.51$ ,  $P = 0.68$ ). Next, to characterize  
281 the specific bacterial species that were best associated with variations in HLA diversity, we used  
282 ANCOM-BC [ancombc2 function in the ANCOM package; ,62]. Since ANCOM includes an internal  
283 normalization step, we used the unrarefied data. ANCOM-BC analysis was carried out on the most  
284 abundant ASVs (i.e., ASVs representing  $> 1\%$  of the total reads and present in more than 5 samples).  
285 In ANCOM-BC, we compared women who were heterozygous at the three HLA class I genes to  
286 women who had at least one homozygous HLA class I gene.

287 To determine whether women who were more similar at HLA sites (allelic dissimilarity or  
288 Sandberg dissimilarity) had more similar microbiome composition, we used partial Mantel test.  
289 Because women whose DNA was extracted in the same extraction batch had more similar  
290 microbiome composition (mantel test:  $r = 0.11$ ,  $P = 0.039$ ), we included a binomial matrix of DNA  
291 extraction batch similarity (0: the two samples were extracted in the same batch and 1: the two  
292 samples were not extracted in the same batch) as a covariate. Similarly to the analysis on the  
293 association between HLA and microbiome diversity, we performed 500 iterations of the Mantel test,  
294 with each iteration corresponding to a single rarefaction run round [65]. In addition, for each of the  
295 three women who were sampled twice, we included just one sample (chosen at random) in each  
296 rarefaction round to avoid pseudoreplication. We report the mean of the statistic parameters over  
297 the 500 tests.

298 Because infections, autoimmune diseases or endocrine disorders may destabilize the  
299 microbiota due to alterations in the immunogenetic control [71], we expected HLA to be a weaker  
300 determinant of the microbiome in women with female factor infertility than in healthy women.  
301 Despite low sample sizes, we explored this hypothesis by testing the associations between HLA and  
302 microbiome diversity ( $\alpha$ - and  $\beta$ -diversity) in the two groups of women separately.

303

## 304 Results

### 305 *Microbiome $\alpha$ -diversity and HLA diversity*

306 Women who had a higher HLA heterozygosity or Sandberg divergence had higher Faith's PD in  
307 follicular fluid microbiome (heterozygosity:  $r = 0.39$ ,  $P = 0.050$  and Sandberg divergence:  $r = 0.44$ ,  $P =$   
308  $0.023$ , respectively). This relationship was also significant when considering HLA class I genes only  
309 (heterozygosity:  $r = 0.55$ ,  $P = 0.004$  and Sandberg divergence:  $r = 0.42$ ,  $P = 0.032$ , respectively; Fig. 2).  
310 In contrast, when considering HLA class II genes, the association between HLA diversity and  
311 microbiome diversity was not significant (heterozygosity:  $r = 0.16$ ,  $P = 0.44$  and Sandberg divergence:  
312  $r = 0.36$ ,  $P = 0.069$ ; Fig. S5). ANCOM-BC analyses carried out at the bacterial species level showed that  
313 women who were heterozygous at the three class I genes had lower relative abundance of  
314 *Lactobacillus jensenii*, but higher relative abundance of *Atopobium vaginae* and *Prevotella bivia* than  
315 women who had at least one homozygous class I genes (Fig. 3). There was no association between  
316 follicular fluid microbiome  $\alpha$ -diversity and genome-wide heterozygosity ( $r = 0.02$ ,  $P = 0.91$ ; Fig. S6).  
317 Interestingly, although  $\alpha$ -diversity of follicular fluid microbiome was not affected by the female factor  
318 infertility diagnoses (t- test:  $t_{1,24} = -0.96$ ,  $P = 0.35$ ; Cohen's  $d: -0.32 [-1.13 - 0.48]$ ; Fig. S7), the  
319 association between HLA diversity and microbiome diversity was detected in women without the  
320 diagnosis of female factor infertility only (i.e., 'healthy women') (heterozygosity at HLA class I genes:  
321  $r = 0.74$ ,  $P = 0.001$  and Sandberg divergence at HLA class I genes:  $r = 0.57$ ,  $P = 0.014$ ; heterozygosity at  
322 HLA class II genes:  $r = 0.30$ ,  $P = 0.24$  and Sandberg divergence at HLA class II genes:  $r = 0.44$ ,  $P =$   
323  $0.071$ ). In contrast, we did not detect an association between microbiome diversity and HLA class I or  
324 class II diversity in women with the diagnosis of female factor infertility ( $n = 9$  women; heterozygosity  
325 at HLA class I genes:  $r = -0.44$ ,  $P = 0.25$ ; Sandberg divergence at HLA class I genes:  $r = -0.65$ ,  $P = 0.07$ ;  
326 heterozygosity at HLA class II genes:  $r = -0.12$ ,  $P = 0.74$ ; Sandberg divergence at HLA class II genes:  $r = -$   
327  $0.10$ ,  $P = 0.79$ ).

### 328 *Microbiome similarity and HLA similarity*

329 Women who were more similar at HLA genes had more similar follicular fluid microbiomes when  
330 considering allelic dissimilarity ( $r = 0.18$ ,  $P = 0.039$ ; Fig. S8a), but not when considering Sandberg  
331 dissimilarity ( $r = 0.12$ ,  $P = 0.11$ ; Fig. S8b). However, this former result was only a non-significant trend  
332 when using the rarefied dataset (Table S2) or when excluding the Asian woman from the data (Table  
333 S2). Interestingly, when we focused on healthy women only, we found that women with more similar  
334 HLA genes had more similar follicular fluid microbiome (allelic dissimilarity:  $r = 0.31$ ,  $P = 0.010$  and  
335 Sandberg dissimilarity:  $r = 0.23$ ,  $P = 0.036$ ; Fig. 4). This effect was mainly due to HLA class II genes  
336 (HLA class II genes:  $r = 0.29$ ,  $P = 0.008$  and  $r = 0.23$ ,  $P = 0.027$ , while HLA class I genes:  $r = 0.23$ ,  $P =$

337 0.056 and  $r = 0.15$ ,  $P = 0.14$ , for allelic dissimilarity and Sandberg dissimilarity respectively; Fig. S9). In  
338 contrast, there was no association between similarity at HLA class I or class II genes and similarity in  
339 microbiome in women diagnosed with female factor infertility ( $P > 0.45$ , in all cases). Women with a  
340 higher genome-wide similarity did not have more similar follicular fluid microbiome ( $r = 0.15$ ,  $P = 0.20$   
341 when considering all women, and  $r = -0.20$ ,  $P = 0.82$  when focusing on the women with no female  
342 factor infertility), indicating that our results cannot be explained by the observed covariation  
343 between HLA and genome-wide dissimilarity and is thus HLA-specific. The  $\beta$ -diversity of follicular  
344 fluid microbiome did not vary, however, with the female factor infertility diagnoses (permanova:  $F_{1,22}$   
345 = 0.79,  $r^2 = 0.03$ ,  $P = 0.76$ ; Fig. S10).

346

## 347 Discussion

348 Here, we report the first evidence that the female reproductive tract microbiome varies with HLA.  
349 First, we found that women with higher HLA class I diversity (heterozygosity and divergence) had a  
350 more diverse follicular fluid microbiome (Faith's PD). In the context of the fight against bacterial  
351 pathogens, a higher diversity of MHC molecules is expected to lead to the recognition of a greater  
352 number of foreign peptides, and thus to the elimination of a larger number of bacterial species [68].  
353 However, while this negative association is expected for pathogenic bacteria, the opposite may be  
354 expected for commensal bacteria, to which we are immunologically tolerant. Accordingly, MHC class  
355 II expression has been shown to promote gut bacterial richness in mice [29,37]. Therefore, a positive  
356 association between MHC and microbiome diversity may arise if a higher diversity of MHC molecules  
357 facilitate the proliferation of a more diverse range of symbiotic microbes [72]. However, studies on  
358 the association between MHC diversity and microbiome diversity have yielded inconsistent results.  
359 While MHC class II diversity is positively associated with microbiome diversity in the skin of two  
360 amphibians [73,74], it is negatively associated with gut microbiome taxonomic diversity in  
361 sticklebacks (*Gasterosteus aculeatus*) and gut microbiome functional diversity in reddish-gray mouse  
362 lemurs (*Microcebus griseorufus*) [41,42]. In contrast, in neotropical bats (*Artibeus jamaicensis*),  
363 Seychelles warblers (*Acrocephalus sechellensis*) and congenic mice, MHC allelic diversity seems to  
364 play little role in shaping gut bacterial diversity [38,43,75]. Considerably more research is thus  
365 required to determine the mechanisms by which MHC diversity may influence microbiome diversity  
366 at mucosal surfaces, including in the female reproductive tract.

367 In contrast to our initial expectations, the association between HLA diversity and  
368 reproductive microbiome diversity was mainly detected at HLA class I genes. Similarly, the gut  
369 microbiome of humans show stronger association with HLA class I diversity than with HLA class II

370 diversity [40]. While MHC class II is typically involved in the interaction between the host and the  
371 microbiome [37], evidence also suggests the role of MHC class I pathway in the cross-talks between  
372 the host and its bacterial commensals [37,76,77]. In addition, in our dataset, the stronger association  
373 with HLA class I diversity may potentially be explained by the higher inter-individual variability of HLA  
374 class I diversity than HLA class II diversity.

375         We found that women with higher HLA class I diversity had higher abundances of *Prevotella*  
376 *bivia* and *Atopobium vaginae* and a lower abundance of *Lactobacillus jensenii*. In the vagina, this  
377 highly diverse microbial composition, described as “CST-IV”, is commonly associated with bacterial  
378 vaginosis [reviewed in 78], a disorder that can cause vaginal discharge and foul-smelling odor.  
379 Bacterial vaginosis is associated with sexually transmitted infections and adverse reproductive  
380 outcome, including pre-term birth, low birthweight and miscarriage [reviewed in 79]. Bacterial  
381 vaginosis may also play a role in the acquisition of human papillomavirus and the development of  
382 associated cervical neoplasia [80,81]. Interestingly, the positive association between HLA diversity  
383 and a CST-IV microbiome, might therefore explain the surprising positive association between HLA  
384 diversity and increased risk of cervical cancer [82]. Although high diversity of the vaginal microbiome  
385 is considered deleterious for fitness in women [78], high HLA diversity conversely is expected to be  
386 beneficial, because it confers enhanced resistance to several pathogens [30,83] and reduces risks of  
387 several cancers (albeit not cervical cancer) [82]. The positive association between HLA diversity and  
388 reproductive tract microbiome diversity observed here therefore raises the possibility of a potential  
389 conflict between the fitness benefits of a highly diverse HLA and of a poorly diverse reproductive  
390 microbiome. However, the effect of the CST-IV microbial state on reproduction is mixed and, in some  
391 women, this microbial state is asymptomatic [84]. Some studies suggest that CST-IV state increases  
392 the risk of adverse outcome only in interaction with a specific genotype (i.e., gene-environment  
393 interaction). For instance, Gómez et al. [85] showed that bacterial vaginosis *per se* is not associated  
394 with an increased risk of spontaneous preterm delivery, but it is its association with susceptible  
395 genotypes at several immune genes that increases the risk. It is therefore possible that women with  
396 higher HLA-diversity are better able to tolerate CST-IV state (i.e., reduced or absent pathogenesis  
397 despite high pathogen load). Similarly, HLA-heterozygosity has been shown to confer better  
398 tolerance to HIV [86].

399         We also found that HLA similar women host more similar reproductive tract microbiome than  
400 HLA-dissimilar women. This is in line with the recent findings by Andeweg et al. (2021), who showed  
401 that individuals with similar functional HLA profile have similar gut microbiota community. This result  
402 suggests that women with similar HLA profile have similar immune control of their microbiota, which  
403 is probably due to differences between HLA variants in their binding specificity to distinct peptides

404 [60]. Specific associations between certain HLA motifs and certain commensal bacteria of the gut  
405 have been detected in sticklebacks, mouse lemurs and neotropical bats [41–43].

406 In the present study, the association between HLA and microbiome ( $\alpha$ - and  $\beta$ -diversity) was  
407 detected in healthy women, but not in women who had female factor infertility. Although we found  
408 no significant differences in the microbiome composition or diversity of women with female factor  
409 infertility and that of healthy women (but see Fig. S6 and S9 which show that this lack of difference is  
410 mostly driven by the microbiome of few healthy women that looks like that of women with female  
411 factor infertility), previous studies have detected that polycystic ovary syndrome, endometriosis or  
412 tubal disease alters the reproductive microbiota [11,87,88]. These reproductive disorders are  
413 associated with changes in immunity [89–91]. The associated dysbiosis might thus be due to an  
414 alteration of the immune control of the microbiota, as suggested for other diseases [71]. Although  
415 our results should be taken with caution because of low sample sizes ( $n = 18$  healthy women and 9  
416 women with female factor infertility), they suggest that, compared to HLA genes, other deterministic  
417 factors or stochastic effects might be stronger determinants of the reproductive microbiome of  
418 women with female factor infertility.

419 We found that, in our follicular fluid samples, the most abundant ASVs belonged to the genus  
420 *Lactobacillus*. This result is consistent with other studies which showed that *Lactobacillus* is a  
421 common genus in follicular fluid [44–49]. However, we found that other abundant ASVs belonged to  
422 the genera *Gardnerella*, *Prevotella*, *Atopobium* and *Sneathia*, which contrasts with the abundant  
423 bacteria detected in other studies (e.g., *Achromobacter*, *Actinomyces*, *Propionobacterium*,  
424 *Bifidobacterium*, *Staphylococcus*, *Streptococcus* and *Enterococcus*) [44,46–49]. Recently, low-biomass  
425 microbial niches, such as the human fetus, the placenta or the endometrium, have been suggested to  
426 be sensitive to contamination and thus data misinterpretation [92–94]. The follicular fluid, like other  
427 sites of the upper reproductive tract, is likely to have low microbial biomass, and clearly, the original  
428 source of the bacteria detected in our study and others can be called into question. In the present  
429 study, the most abundant bacteria are common members of the vaginal microbiome [15,95,96], and  
430 we did not detect significant differences between the composition of follicular fluid microbiome and  
431 vaginal microbiome. In addition, our follicular fluid microbiome can be split into five configurations  
432 that correspond to the five community state types (CSTs) widely used to describe vaginal microbiota  
433 (Fig. 1) [15,97]. Samples collected transcervically or transvaginally are known to differ to samples  
434 collected using hysterectomy, laparoscopy or cesarean section [19,98]. Therefore, we cannot exclude  
435 the possibility that our follicular fluid samples were contaminated with vaginal bacterial strains at the  
436 time of puncture for transvaginal oocyte aspiration, although it is quite possible also that the bacteria  
437 observed both in follicular fluid and vagina are true inhabitants of both sites.

438 In conclusion, regardless of the true (vaginal or follicular fluid) origin of the microbiome  
439 studied, our study indicates that HLA is an important host genetic factor driving differences in the  
440 reproductive microbiota of women. The MHC is known to affect the reproductive success via  
441 numerous mechanisms, including effects on the quality of the gametes, fertilization success and  
442 embryonic development [99–101]. Because the reproductive tract microbiota is well known to  
443 modulate all these factors, our results open the door for further research on the role of the  
444 reproductive tract microbiota in mediating MHC-related variation in reproductive success.

445

#### 446 **Data availability**

447 Due to the privacy commitments of the patients and the terms of the consent, the original data  
448 cannot be made widely available to the academic community.

449

#### 450 **Acknowledgements**

451 We thank Ulrica Candolin for hosting SL during her stay at the University of Helsinki. This study was  
452 funded by grants from the Fondation Fyssen (to SL) and Academy of Finland (grant 308485 to JK). We  
453 are grateful to the genotoul bioinformatics platform Toulouse Occitanie (Bioinfo Genotoul, doi:  
454 10.15454/1.5572369328961167E12) and Sigenae group for providing computing and storage  
455 resources thanks to Galaxy instance <http://sigenae-workbench.toulouse.inra.fr>.

456

#### 457 **Ethical statement**

458 The sample collection and the experimental procedures were based on the favorable opinion of the  
459 Ethics Committee, Hospital District of Northern Savo (decision number 77/2017). An informed  
460 consent was obtained from all the subjects prior to sample collection.

461

462

#### 463 **References**

- 464 1. Kroon SJ, Ravel J, Huston WM. 2018 Cervicovaginal microbiota, women's health, and  
465 reproductive outcomes. *Fertil. Steril.* **110**, 327–336. (doi:10.1016/j.fertnstert.2018.06.036)



- 466 2. Moreno I, Simon C. 2019 Deciphering the effect of reproductive tract microbiota on human  
467 reproduction. *Reprod. Med. Biol.* **18**, 40–50. (doi:10.1002/rmb2.12249)
- 468 3. Rowe M, Veerus L, Trosvik P, Buckling A, Pizzari T. 2020 The Reproductive Microbiome: an  
469 emerging driver of sexual selection, sexual conflict, mating systems, and reproductive isolation.  
470 *Trends Ecol. Evol.*
- 471 4. Farahani L, Tharakan T, Yap T, Ramsay JW, Jayasena CN, Minhas S. 2021 The semen microbiome  
472 and its impact on sperm function and male fertility: A systematic review and meta-analysis.  
473 *Andrology* **9**, 115–144. (doi:10.1111/andr.12886)
- 474 5. Baud D, Pattaroni C, Vulliamoz N, Castella V, Marsland BJ, Stojanov M. 2019 Sperm Microbiota  
475 and Its Impact on Semen Parameters. *Front. Microbiol.* **10**. (doi:10.3389/fmicb.2019.00234)
- 476 6. Koziol JH, Sheets T, Wickware CL, Johnson TA. 2022 Composition and diversity of the seminal  
477 microbiota in bulls and its association with semen parameters. *Theriogenology* **182**, 17–25.  
478 (doi:10.1016/j.theriogenology.2022.01.029)
- 479 7. McAnally BE, Smith MS, Wiegert JG, Palanisamy V, Chitlapilly Dass S, Poole RK. 2023  
480 Characterization of boar semen microbiome and association with sperm quality parameters. *J.*  
481 *Anim. Sci.* **101**, skad243. (doi:10.1093/jas/skad243)
- 482 8. Su Y *et al.* 2021 Association of female reproductive tract microbiota with egg production in  
483 layer chickens. *GigaScience* **10**, giab067.
- 484 9. Koester LR, Petry AL, Youngs CR, Schmitz-Esser S. 2021 Ewe Vaginal Microbiota: Associations  
485 With Pregnancy Outcome and Changes During Gestation. *Front. Microbiol.* **12**.
- 486 10. Fu M, Zhang X, Liang Y, Lin S, Qian W, Fan S. 2020 Alterations in Vaginal Microbiota and  
487 Associated Metabolome in Women with Recurrent Implantation Failure. *mBio* **11**,  
488 10.1128/mbio.03242-19. (doi:10.1128/mbio.03242-19)
- 489 11. Hong X, Qin P, Huang K, Ding X, Ma J, Xuan Y, Zhu X, Peng D, Wang B. 2020 Association between  
490 polycystic ovary syndrome and the vaginal microbiome: A case-control study. *Clin. Endocrinol.*  
491 *(Oxf.)* **93**, 52–60. (doi:10.1111/cen.14198)
- 492 12. Zhao F, Hu X, Ying C. 2023 Advances in Research on the Relationship between Vaginal  
493 Microbiota and Adverse Pregnancy Outcomes and Gynecological Diseases. *Microorganisms* **11**,  
494 991. (doi:10.3390/microorganisms11040991)
- 495 13. Stumpf RM, Wilson BA, Rivera A, Yildirim S, Yeoman CJ, Polk JD, White BA, Leigh SR. 2013 The  
496 primate vaginal microbiome: Comparative context and implications for human health and  
497 disease. *Am. J. Phys. Anthropol.* **152**, 119–134. (doi:10.1002/ajpa.22395)
- 498 14. Comizzoli P, Power M. 2019 Reproductive microbiomes in wild animal species: A new  
499 dimension in conservation biology. In *Reproductive Sciences in Animal Conservation*, pp. 225–  
500 240. Springer.
- 501 15. France M, Alizadeh M, Brown S, Ma B, Ravel J. 2022 Towards a deeper understanding of the  
502 vaginal microbiota. *Nat. Microbiol.* **7**, 367–378. (doi:10.1038/s41564-022-01083-2)

- 503 16. Garcia-Garcia RM, Arias-Álvarez M, Jordán-Rodríguez D, Rebollar PG, Lorenzo PL, Herranz C,  
504 Rodríguez JM. 2022 Female reproduction and the microbiota in mammals: Where are we?  
505 *Theriogenology* **194**, 144–153. (doi:10.1016/j.theriogenology.2022.10.007)
- 506 17. Moreno I *et al.* 2016 Evidence that the endometrial microbiota has an effect on implantation  
507 success or failure. *Am. J. Obstet. Gynecol.* **215**, 684–703.  
508 (doi:https://doi.org/10.1016/j.ajog.2016.09.075)
- 509 18. Wessels JM, Domínguez MA, Leyland NA, Agarwal SK, Foster WG. 2021 Endometrial microbiota  
510 is more diverse in people with endometriosis than symptomatic controls. *Sci. Rep.* **11**, 18877.
- 511 19. Canha-Gouveia A *et al.* 2023 The female upper reproductive tract harbors endogenous  
512 microbial profiles. *Front. Endocrinol.* **14**.
- 513 20. Chen C *et al.* 2017 The microbiota continuum along the female reproductive tract and its  
514 relation to uterine-related diseases. *Nat. Commun.* **8**, 875. (doi:10.1038/s41467-017-00901-0)
- 515 21. Hickey RJ *et al.* 2015 Vaginal Microbiota of Adolescent Girls Prior to the Onset of Menarche  
516 Resemble Those of Reproductive-Age Women. *mBio* **6**, 10.1128/mbio.00097-15.  
517 (doi:10.1128/mbio.00097-15)
- 518 22. Si J, You HJ, Yu J, Sung J, Ko G. 2017 Prevotella as a Hub for Vaginal Microbiota under the  
519 Influence of Host Genetics and Their Association with Obesity. *Cell Host Microbe* **21**, 97–105.  
520 (doi:10.1016/j.chom.2016.11.010)
- 521 23. Moosa Y, Kwon D, de Oliveira T, Wong EB. 2020 Determinants of Vaginal Microbiota  
522 Composition. *Front. Cell. Infect. Microbiol.* **10**. (doi:10.3389/fcimb.2020.00467)
- 523 24. Cheng L, Gao Y, Xia Q, Wang H, Xie X, Liu Y, Shang H, Diao Y. 2021 Reproductive tract microbiota  
524 of women in childbearing age shifts upon gynecological infections and menstrual cycle. *BMC*  
525 *Microbiol.* **21**, 252. (doi:10.1186/s12866-021-02300-4)
- 526 25. Morsli M, Gimenez E, Magnan C, Salipante F, Huberlant S, Letouzey V, Lavigne J-P. 2024 The  
527 association between lifestyle factors and the composition of the vaginal microbiota: a review.  
528 *Eur. J. Clin. Microbiol. Infect. Dis.* (doi:10.1007/s10096-024-04915-7)
- 529 26. Zheng D, Liwinski T, Elinav E. 2020 Interaction between microbiota and immunity in health and  
530 disease. *Cell Res.* **30**, 492–506. (doi:10.1038/s41422-020-0332-7)
- 531 27. Mehta SD, Nannini DR, Otieno F, Green SJ, Agingu W, Landay A, Zheng Y, Hou L. 2020 Host  
532 Genetic Factors Associated with Vaginal Microbiome Composition in Kenyan Women.  
533 *mSystems* **5**, 10.1128/msystems.00502-20. (doi:10.1128/msystems.00502-20)
- 534 28. Fan W *et al.* 2021 Association between Human Genetic Variants and the Vaginal Bacteriome of  
535 Pregnant Women. *mSystems* **6**, 10.1128/msystems.00158-21. (doi:10.1128/msystems.00158-  
536 21)
- 537 29. Roland MM *et al.* 2023 B-cell-specific MhcII regulates microbiota composition in a primarily IgA-  
538 independent manner. *Front. Immunol.* **14**. (doi:10.3389/fimmu.2023.1253674)
- 539 30. Khan T, Rahman M, Ahmed I, Al Ali F, Jithesh PV, Marr N. 2022 Human leukocyte antigen class II  
540 gene diversity tunes antibody repertoires to common pathogens. *Front. Immunol.* **13**.

- 541 31. Winkler P, Ghadimi D, Schrezenmeir J, Kraehenbuhl J-P. 2007 Molecular and Cellular Basis of  
542 Microflora-Host Interactions<sup>1,2</sup>. *J. Nutr.* **137**, 756S-772S. (doi:10.1093/jn/137.3.756S)
- 543 32. Roland MM, Mohammed AD, Kubinak JL. 2020 How MHCII signaling promotes benign host-  
544 microbiota interactions. *PLoS Pathog.* **16**, e1008558. (doi:10.1371/journal.ppat.1008558)
- 545 33. Kato LM, Kawamoto S, Maruya M, Fagarasan S. 2014 Gut TFH and IgA: key players for  
546 regulation of bacterial communities and immune homeostasis. *Immunol. Cell Biol.* **92**, 49–56.  
547 (doi:10.1038/icb.2013.54)
- 548 34. Kubinak JL, Round JL. 2016 Do antibodies select a healthy microbiota? *Nat. Rev. Immunol.* **16**,  
549 767–774.
- 550 35. Chen K, Magri G, Grasset EK, Cerutti A. 2020 Rethinking mucosal antibody responses: IgM, IgG  
551 and IgD join IgA. *Nat. Rev. Immunol.* **20**, 427–441. (doi:10.1038/s41577-019-0261-1)
- 552 36. Toivanen P, Vaahtovuori J, Eerola E. 2001 Influence of major histocompatibility complex on  
553 bacterial composition of fecal flora. *Infect. Immun.* **69**, 2372–2377.
- 554 37. Kubinak JL *et al.* 2015 MHC variation sculpts individualized microbial communities that control  
555 susceptibility to enteric infection. *Nat. Commun.* **6**, 8642.
- 556 38. Khan MW, Stephens WZ, Mohammed AD, Round JL, Kubinak JL. 2019 Does MHC heterozygosity  
557 influence microbiota form and function? *PLoS One* **14**, e0215946.
- 558 39. Leclaire S, Strandh M, Dell’Ariccia G, Gabirot M, Westerdahl H, Bonadonna F. 2019 Plumage  
559 microbiota covaries with the major histocompatibility complex in blue petrels. *Mol. Ecol.* **28**,  
560 833–846.
- 561 40. Andeweg SP, Keşmir C, Dutilh BE. 2021 Quantifying the Impact of Human Leukocyte Antigen on  
562 the Human Gut Microbiota. *mSphere* **6**, e00476-21.
- 563 41. Bolnick DI, Snowberg LK, Caporaso JG, Lauber C, Knight R, Stutz WE. 2014 Major  
564 histocompatibility complex class IIb polymorphism influences gut microbiota composition and  
565 diversity. *Mol. Ecol.* **23**, 4831–4845.
- 566 42. Montero BK *et al.* 2021 Evidence of MHC class I and II influencing viral and helminth infection  
567 via the microbiome in a non-human primate. *PLoS Pathog.* **17**, e1009675.  
568 (doi:10.1371/journal.ppat.1009675)
- 569 43. Fleischer R, Schmid DW, Brändel SD, Rasche A, Corman VM, Drosten C, Tschapka M, Sommer S.  
570 2022 Interaction between MHC diversity and constitution, gut microbiota and Astrovirus  
571 infections in a neotropical bat. *Mol. Ecol.* **31**, 3342–3359.
- 572 44. Pelzer ES, Allan JA, Waterhouse MA, Ross T, Beagley KW, Knox CL. 2013 Microorganisms within  
573 Human Follicular Fluid: Effects on IVF. *PLOS ONE* **8**, e59062.  
574 (doi:10.1371/journal.pone.0059062)
- 575 45. Ou S, Liao M, Cui L, Du Y, Zhao L, Peng C, Jiang L, Liu B. 2023 Associations between microbial  
576 presence in follicular fluid with IVF outcomes: a systematic review and meta-analysis. *J. Assist.*  
577 *Reprod. Genet.* **40**, 2501–2511. (doi:10.1007/s10815-023-02912-x)

- 578 46. Wei W *et al.* 2023 Characterization of the follicular fluid microbiota based on culturomics and  
579 sequencing analysis. *J. Med. Microbiol.* **72**, 001741. (doi:10.1099/jmm.0.001741)
- 580 47. Wu Y-R, Dong Y-H, Liu C-J, Tang X-D, Zhang N-N, Shen J, Wu Z, Li X-R, Shao J-Y. 2023  
581 Microbiological composition of follicular fluid in patients undergoing IVF and its association  
582 with infertility. *Am. J. Reprod. Immunol.* **89**, e13652. (doi:10.1111/aji.13652)
- 583 48. Usman SF, Shuaibu IR, Durojaiye K, Medugu N, Iregbu KC. 2021 The presence of  
584 microorganisms in follicular fluid and its effect on the outcome of in vitro fertilization-embryo  
585 transfer (IVF-ET) treatment cycles. *PLOS ONE* **16**, e0246644.  
586 (doi:10.1371/journal.pone.0246644)
- 587 49. Kim SM, Won KH, Hong YH, Kim SK, Lee JR, Jee BC, Suh CS. 2022 Microbiology of Human  
588 Follicular Fluid and the Vagina and Its Impact on in Vitro Fertilization Outcomes. *Yonsei Med. J.*  
589 **63**, 941–947.
- 590 50. Glassing A, Dowd SE, Galandiuk S, Davis B, Chiodini RJ. 2016 Inherent bacterial DNA  
591 contamination of extraction and sequencing reagents may affect interpretation of microbiota in  
592 low bacterial biomass samples. *Gut Pathog.* **8**, 24. (doi:10.1186/s13099-016-0103-7)
- 593 51. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, Fierer N,  
594 Knight R. 2011 Global patterns of 16S rRNA diversity at a depth of millions of sequences per  
595 sample. *Proc. Natl. Acad. Sci.* **108**, 4516–4522.
- 596 52. Rognes T, Flouri T, Nichols B, Quince C, Mahé F. 2016 VSEARCH: a versatile open source tool for  
597 metagenomics. *PeerJ* **4**, e2584.
- 598 53. Martin M. 2011 Cutadapt removes adapter sequences from high-throughput sequencing reads.  
599 *EMBnet J.* **17**, 10–12.
- 600 54. Mahé F, Czech L, Stamatakis A, Quince C, de Vargas C, Dunthorn M, Rognes T. 2022 Swarm v3:  
601 towards tera-scale amplicon clustering. *Bioinformatics* **38**, 267–269.
- 602 55. Bokulich NA, Subramanian S, Faith JJ, Gevers D, Gordon JI, Knight R, Mills DA, Caporaso JG.  
603 2013 Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing.  
604 *Nat. Methods* **10**, 57–59.
- 605 56. Davis NM, Proctor DM, Holmes SP, Relman DA, Callahan BJ. 2018 Simple statistical  
606 identification and removal of contaminant sequences in marker-gene and metagenomics data.  
607 *Microbiome* **6**, 226. (doi:10.1186/s40168-018-0605-2)
- 608 57. Stinson LF, Keelan JA, Payne MS. 2019 Identification and removal of contaminating microbial  
609 DNA from PCR reagents: impact on low-biomass microbiome analyses. *Lett. Appl. Microbiol.* **68**,  
610 2–8. (doi:10.1111/lam.13091)
- 611 58. Ridgway HF, Olson BH. 1982 Chlorine resistance patterns of bacteria from two drinking water  
612 distribution systems. *Appl. Environ. Microbiol.* **44**, 972–987. (doi:10.1128/aem.44.4.972-  
613 987.1982)
- 614 59. Hendrickson R, Urbaniak C, Minich JJ, Aronson HS, Martino C, Stepanauskas R, Knight R,  
615 Venkateswaran K. 2021 Clean room microbiome complexity impacts planetary protection  
616 bioburden. *Microbiome* **9**, 238. (doi:10.1186/s40168-021-01159-x)

- 617 60. Özer O, Lenz TL. 2021 Unique Pathogen Peptidomes Facilitate Pathogen-Specific Selection and  
618 Specialization of MHC Alleles. *Mol. Biol. Evol.* **38**, 4376–4387. (doi:10.1093/molbev/msab176)
- 619 61. McCoy CO, Matsen FA. 2013 Abundance-weighted phylogenetic diversity measures distinguish  
620 microbial community states and are robust to sampling depth. *PeerJ* **1**, e157.  
621 (doi:10.7717/peerj.157)
- 622 62. Lin H, Peddada SD. 2020 Analysis of compositions of microbiomes with bias correction. *Nat.*  
623 *Commun.* **11**, 3514. (doi:10.1038/s41467-020-17041-7)
- 624 63. McMurdie PJ, Holmes S. 2014 Waste not, want not: why rarefying microbiome data is  
625 inadmissible. *PLoS Comput. Biol.* **10**, e1003531.
- 626 64. Weiss S *et al.* 2017 Normalization and microbial differential abundance strategies depend upon  
627 data characteristics. *Microbiome* **5**, 27. (doi:10.1186/s40168-017-0237-y)
- 628 65. Cameron ES, Schmidt PJ, Tremblay BJ-M, Emelko MB, Müller KM. 2021 Enhancing diversity  
629 analysis by repeatedly rarefying next generation sequencing data describing microbial  
630 communities. *Sci. Rep.* **11**, 22302. (doi:10.1038/s41598-021-01636-1)
- 631 66. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. 2015 Second-generation PLINK:  
632 rising to the challenge of larger and richer datasets. *Gigascience* **4**, s13742-015.
- 633 67. Sandberg M, Eriksson L, Jonsson J, Sjöström M, Wold S. 1998 New chemical descriptors relevant  
634 for the design of biologically active peptides. A multivariate characterization of 87 amino acids.  
635 *J. Med. Chem.* **41**, 2481–2491.
- 636 68. Pierini F, Lenz TL. 2018 Divergent Allele Advantage at Human MHC Genes: Signatures of Past  
637 and Ongoing Selection. *Mol. Biol. Evol.* **35**, 2145–2158. (doi:10.1093/molbev/msy116)
- 638 69. Roved J, Hansson B, Stervander M, Hasselquist D, Westerdahl H. 2022 MHCtools—an R package  
639 for MHC high-throughput sequencing data: Genotyping, haplotype and supertype inference,  
640 and downstream genetic analyses in non-model organisms. *Mol. Ecol. Resour.* **22**, 2775–2792.
- 641 70. Oksanen J, Kindt R, Legendre P, O’Hara B, Stevens MHH, Oksanen MJ, Suggests M. 2007 The  
642 vegan package. *Community Ecol. Package* **10**, 719.
- 643 71. Zaneveld JR, McMinds R, Thurber RV. 2017 Stress and stability: applying the Anna Karenina  
644 principle to animal microbiomes. *Nat. Microbiol.* **2**, 1–8.
- 645 72. Schubert N, Nichols HJ, Winternitz JC. 2021 How can the MHC mediate social odor via the  
646 microbiota community? A deep dive into mechanisms. *Behav. Ecol.* **32**, 359–373.
- 647 73. Hernández-Gómez O, Briggler JT, Williams RN. 2018 Influence of immunogenetics, sex and body  
648 condition on the cutaneous microbial communities of two giant salamanders. *Mol. Ecol.* **27**,  
649 1915–1929. (doi:10.1111/mec.14500)
- 650 74. Belasen AM, Riolo MA, Bletz MC, Lyra ML, Toledo LF, James TY. 2021 Geography, Host Genetics,  
651 and Cross-Domain Microbial Networks Structure the Skin Microbiota of Fragmented Brazilian  
652 Atlantic Forest Frog Populations. *Ecol. Evol.* **11**, 9293–9307. (doi:10.1002/ece3.7594)

- 653 75. Davies CS, Worsley SF, Maher KH, Komdeur J, Burke T, Dugdale HL, Richardson DS. 2022  
654 Immunogenetic variation shapes the gut microbiome in a natural vertebrate population.  
655 *Microbiome* **10**, 41. (doi:10.1186/s40168-022-01233-y)
- 656 76. Shao L, Kamalu O, Mayer L. 2005 Non-classical MHC class I molecules on intestinal epithelial  
657 cells: mediators of mucosal crosstalk. *Immunol. Rev.* **206**, 160–176. (doi:10.1111/j.0105-  
658 2896.2005.00295.x)
- 659 77. Tanoue T *et al.* 2019 A defined commensal consortium elicits CD8 T cells and anti-cancer  
660 immunity. *Nature* **565**, 600–605.
- 661 78. Smith SB, Ravel J. 2017 The vaginal microbiota, host defence and reproductive physiology. *J.*  
662 *Physiol.* **595**, 451–463. (doi:10.1113/JP271694)
- 663 79. Chen X, Lu Y, Chen T, Li R. 2021 The Female Vaginal Microbiome in Health and Bacterial  
664 Vaginosis. *Front. Cell. Infect. Microbiol.* **11**.
- 665 80. Gillet E, Meys JF, Verstraelen H, Verhelst R, De Sutter P, Temmerman M, Broeck DV. 2012  
666 Association between bacterial vaginosis and cervical intraepithelial neoplasia: systematic  
667 review and meta-analysis.
- 668 81. Champer M, Wong A, Champer J, Brito I, Messer P, Hou J, Wright J. 2018 The role of the vaginal  
669 microbiome in gynaecological cancer. *BJOG Int. J. Obstet. Gynaecol.* **125**, 309–315.  
670 (doi:10.1111/1471-0528.14631)
- 671 82. Wang Q-L *et al.* 2023 Association of HLA diversity with the risk of 25 cancers in the UK Biobank.  
672 *eBioMedicine* **92**. (doi:10.1016/j.ebiom.2023.104588)
- 673 83. Planelles D, Nagore E, Moret A, Botella-Estrada R, Vila E, Guillén C, Montoro JA. 2006 HLA class  
674 II polymorphisms in Spanish melanoma patients: homozygosity for HLA-DQA1 locus can be a  
675 potential melanoma risk factor. *Br. J. Dermatol.* **154**, 261–266. (doi:10.1111/j.1365-  
676 2133.2005.06896.x)
- 677 84. Eschenbach DA, Hilier S, Critchlow C, Stevens C, DeRouen T, Holmes KK. 1988 Diagnosis and  
678 clinical manifestations of bacterial vaginosis. *Am. J. Obstet. Gynecol.* **158**, 819–828.
- 679 85. Gómez LM, Sammel MD, Appleby DH, Elovitz MA, Baldwin DA, Jeffcoat MK, Macones GA, Parry  
680 S. 2010 Evidence of a gene-environment interaction that predisposes to spontaneous preterm  
681 birth: a role for asymptomatic bacterial vaginosis and DNA variants in genes that control the  
682 inflammatory response. *Am. J. Obstet. Gynecol.* **202**, 386.e1-386.e6.  
683 (doi:10.1016/j.ajog.2010.01.042)
- 684 86. Regoes RR *et al.* 2014 Disentangling human tolerance and resistance against HIV. *PLoS Biol.* **12**,  
685 e1001951.
- 686 87. Liversedge NH, Turner A, Horner PJ, Keay SD, Jenkins JM, Hull MGR. 1999 The influence of  
687 bacterial vaginosis on in-vitro fertilization and embryo implantation during assisted  
688 reproduction treatment. *Hum. Reprod.* **14**, 2411–2415. (doi:10.1093/humrep/14.9.2411)
- 689 88. Jiang I, Yong PJ, Allaire C, Bedaiwy MA. 2021 Intricate Connections between the Microbiota and  
690 Endometriosis. *Int. J. Mol. Sci.* **22**, 5644. (doi:10.3390/ijms22115644)

- 691 89. Matarese G, Placido GD, Nikas Y, Alviggi C. 2003 Pathogenesis of endometriosis: natural  
692 immunity dysfunction or autoimmune disease? *Trends Mol. Med.* **9**, 223–228.  
693 (doi:10.1016/S1471-4914(03)00051-0)
- 694 90. Horne AW, Stock SJ, King AE. 2008 Innate immunity and disorders of the female reproductive  
695 tract. *Reproduction* **135**, 739.
- 696 91. Xiao H, Yin T, Diao L, Zhang Y, Huang C. 2023 Association between immunity and different  
697 clinical symptoms in patients with polycystic ovary syndrome. *Am. J. Reprod. Immunol.* **90**,  
698 e13780. (doi:10.1111/aji.13780)
- 699 92. Olomu IN, Pena-Cortes LC, Long RA, Vyas A, Krichevskiy O, Luellwitz R, Singh P, Mulks MH. 2020  
700 Elimination of “kitome” and “splashome” contamination results in lack of detection of a unique  
701 placental microbiome. *BMC Microbiol.* **20**, 157. (doi:10.1186/s12866-020-01839-y)
- 702 93. Molina NM, Sola-Leyva A, Haahr T, Aghajanova L, Laudanski P, Castilla JA, Altmäe S. 2021  
703 Analysing endometrial microbiome: methodological considerations and recommendations for  
704 good practice. *Hum. Reprod.* **36**, 859–879. (doi:10.1093/humrep/deab009)
- 705 94. Kennedy KM *et al.* 2023 Questioning the fetal microbiome illustrates pitfalls of low-biomass  
706 microbial studies. *Nature* **613**, 639–649. (doi:10.1038/s41586-022-05546-8)
- 707 95. Gajer P *et al.* 2012 Temporal Dynamics of the Human Vaginal Microbiota. *Sci. Transl. Med.* **4**,  
708 132ra52. (doi:10.1126/scitranslmed.3003605)
- 709 96. Tabatabaei N, Eren A, Barreiro L, Yotova V, Dumaine A, Allard C, Fraser W. 2019 Vaginal  
710 microbiome in early pregnancy and subsequent risk of spontaneous preterm birth: a case–  
711 control study. *BJOG Int. J. Obstet. Gynaecol.* **126**, 349–358. (doi:10.1111/1471-0528.15299)
- 712 97. Ravel J *et al.* 2011 Vaginal microbiome of reproductive-age women. *Proc. Natl. Acad. Sci.* **108**,  
713 4680–4687. (doi:10.1073/pnas.1002611107)
- 714 98. Winters AD *et al.* 2019 Does the endometrial cavity have a molecular microbial signature? *Sci.*  
715 *Rep.* **9**, 9905. (doi:10.1038/s41598-019-46173-0)
- 716 99. Fernandez N, Cooper J, Sprinks M, AbdElrahman M, Fiszer D, Kurpisz M, Dealtry G. 1999 A  
717 critical review of the role of the major histocompatibility complex in fertilization,  
718 preimplantation development and feto-maternal interactions. *Hum. Reprod. Update* **5**, 234–  
719 248.
- 720 100. Skarstein F, Folstad I, Liljedal S, Grahn M. 2005 MHC and fertilization success in the Arctic charr  
721 (*Salvelinus alpinus*). *Behav. Ecol. Sociobiol.* **57**, 374–380.
- 722 101. Marques PI *et al.* 2019 Semen quality is affected by HLA class I alleles together with sexually  
723 transmitted diseases. *Andrology* **7**, 867–877. (doi:10.1111/andr.12625)

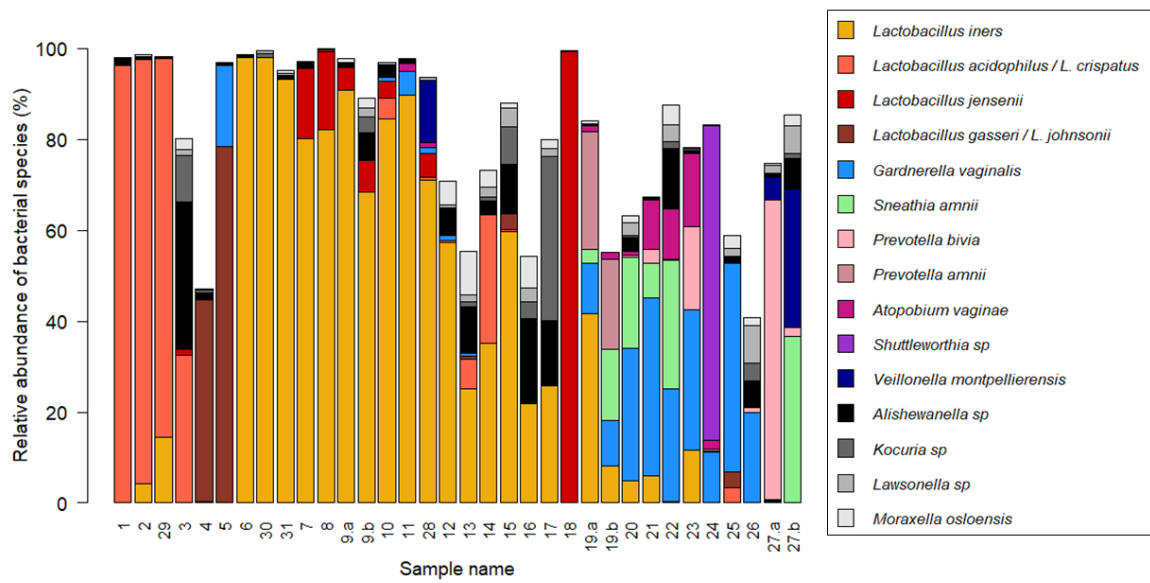
724

725

726

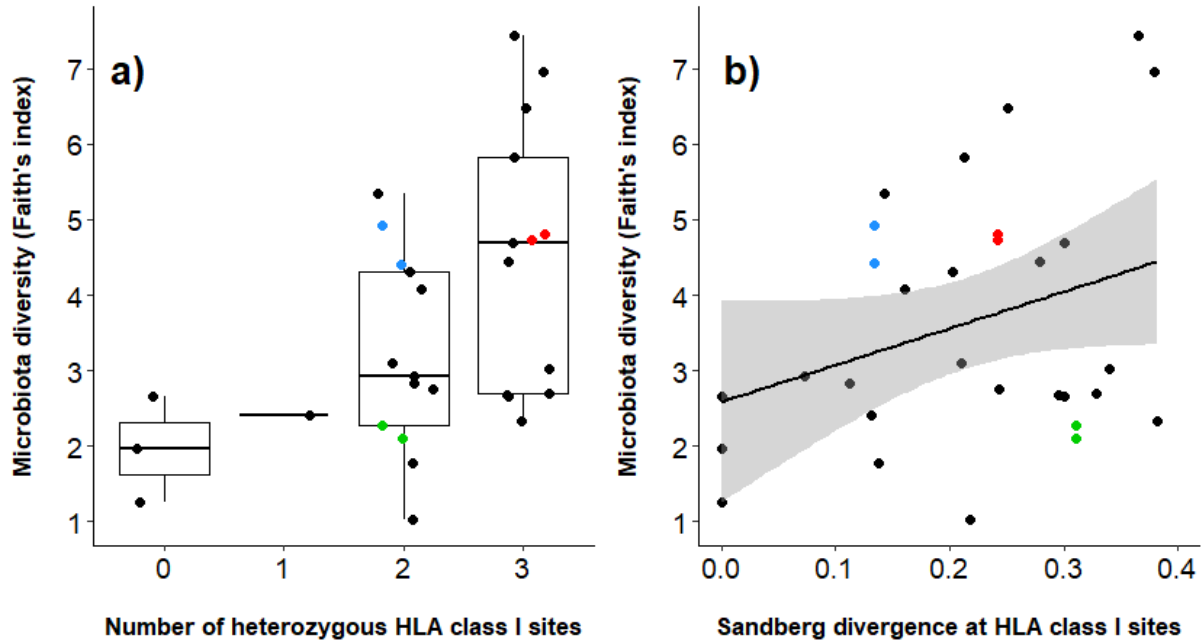
727

728 **Figure 1:** Relative abundance of the most abundant bacterial species in follicular fluid. Each bar  
 729 corresponds to one of the sample. The samples of the three women sampled twice are named with  
 730 “.a” and “.b”. The four women who were not genotyped at HLA correspond to samples 28-31. The  
 731 Asian woman is sample 26. Bars are ordered based on community-state like type. In the vagina, CST-I  
 732 is dominated by *Lactobacillus crispatus*, CST-II is dominated by *L. gasseri*, CST-III is dominated by *L.*  
 733 *iners*, CST-V is dominated by *L. jensenii* and CST-IV is composed of a polymicrobial mixture of strict  
 734 and facultative anaerobes including species of the genera *Gardnerella*, *Atopobium*, *Prevotella* [97].  
 735 Only the 15 most abundant bacterial species are shown.





738 **Figure 2:** Microbiota diversity (Faith's PD) in follicular fluid according to (a) the number of  
 739 heterozygous HLA class I sites and (b) the Sandberg divergence at HLA class I sites. The red, green and  
 740 blue dots correspond to the three women sampled twice. Only one of their two samples was  
 741 considered in the statistical analyses during each of the 500 rarefaction rounds. For (a), similar  
 742 results were found when pooling all the women with at least one homozygous site. The black line  
 743 and grey shadow in (b) represent the linear trend and confidence interval.

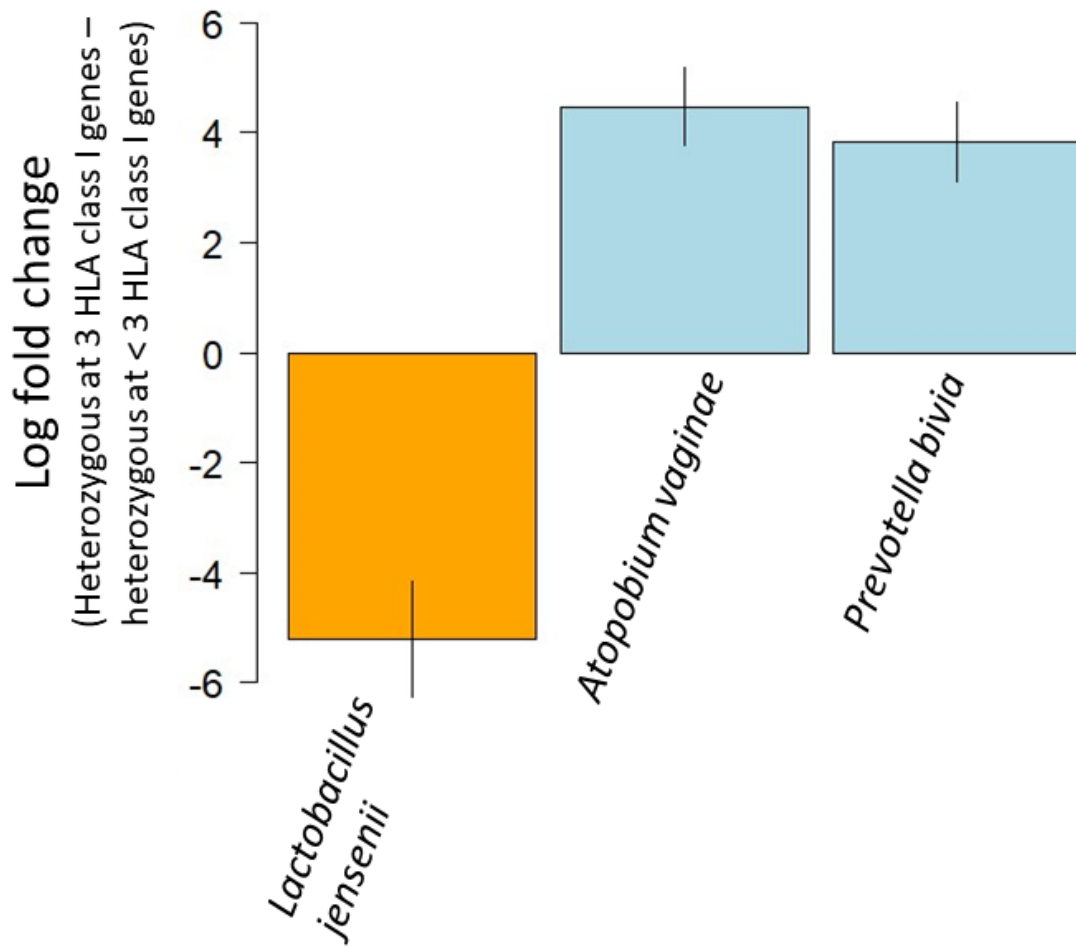


744

745

746

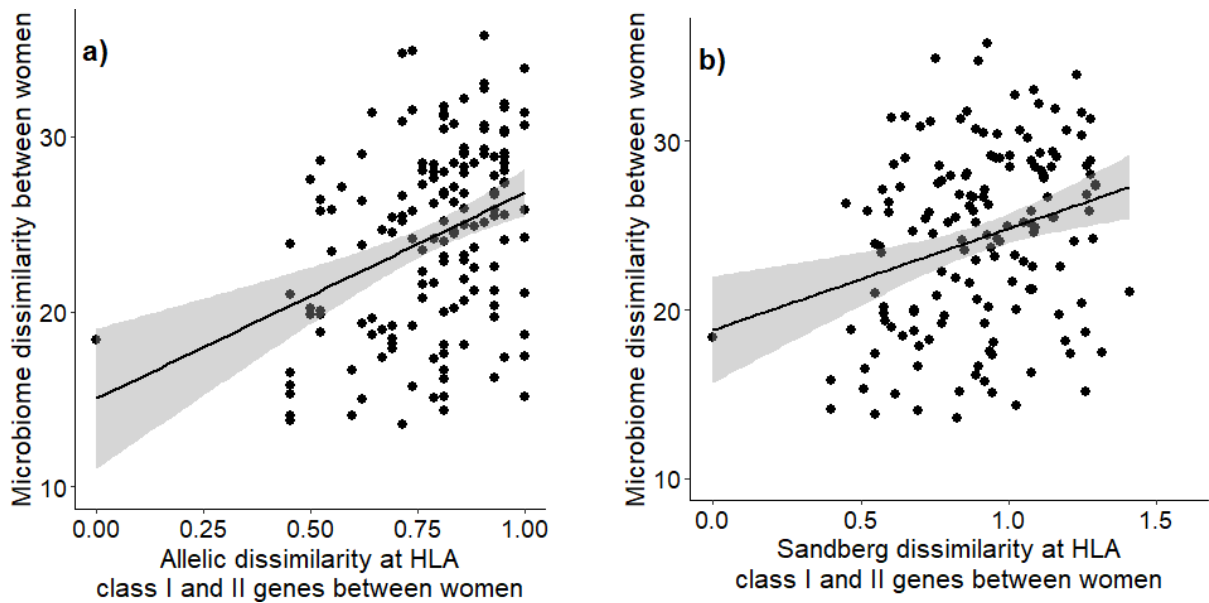
747 **Figure 3:** Results of ANCOMBC analyses, as shown by the log fold change of the 3 bacterial species  
748 that were differently abundant between women who were heterozygous at the three HLA class I  
749 genes and women who were heterozygous at < 3 HLA class I genes.



750

751

752 **Figure 4** : Dissimilarity in microbiome composition in pairs of women with no female factor infertility  
753 diagnoses (n = 18 women, 153 pairs) according to their dissimilarity in HLA class I and class II genes.  
754 Dissimilarity in HLA class II genes was estimated (a) on a pure shared-non shared allele basis and (b)  
755 using Sandberg dissimilarity. The black line and grey shadow represent the linear trend and  
756 confidence interval. The dot at 0 HLA dissimilarity corresponds to two women who had the same HLA  
757 genotype. When excluding these two women from the dataset (i.e., 33 pairs excluded), results  
758 remained significant for allelic dissimilarity (r = 0.19, P = 0.050), but not for Sandberg dissimilarity (r =  
759 0.12, P = 0.17).



760