



HAL
open science

Women reproductive microbiota varies with HLA

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.. Women reproductive microbiota varies with HLA. 2023. hal-04246492

HAL Id: hal-04246492

<https://hal.science/hal-04246492>

Preprint submitted on 17 Oct 2023

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.

Women reproductive microbiota varies with HLA

Sarah LECLAIRE¹, Mandar BANDEKAR², Melissah ROWE³, Jarmo RITARI⁴, Annalaura JOKINIEMI², Jukka PARTANEN⁴, Pia ALLINEN⁵, Liisa KUUSIPALO⁶, Jukka KEKÄLÄINEN²

¹Laboratoire Evolution et Diversité Biologique (EDB), UMR5174, Université Toulouse 3 Paul Sabatier, CNRS, IRD, 118 rte de Narbonne, 31062 Toulouse, France

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

³Department of Animal Ecology, Netherlands Institute of Ecology (NIOO-KNAW), Wageningen, The Netherlands

⁴Finnish Red Cross Blood Service, Research and Development, Haartmaninkatu 8, 00290 Helsinki, Finland

⁵Ovumia Kuopio, Ajurinkatu 16, 70110 Kuopio, Finland

⁶North Karelia Central Hospital, Tikkamäentie 16, 80210 Joensuu, Finland

Introduction

In mammals, follicular fluid is the liquid filling the follicular antrum and thus forms the microenvironment of the oocyte during its development and maturation. It is composed of several biologically active molecules that acts as major oocyte nutrients or important mediators of the interaction between male and female gametes (e.g., Fitzpatrick et al. 2020; Jokiniemi et al. 2020). Follicular fluid composition is thus a key element to the success of fertilization and the development of viable embryos (Kim et al. 2006; Dumesic et al. 2015; Ciepiela et al. 2019).

Recently, a few studies have shown that many bacteria colonize the follicular fluid, even in the absence of symptomatic infection (Einenkel et al. 2019; Salary et al. 2020; Usman et al. 2021; Wu et al. 2023). One of the plausible way of follicular fluid colonization is the ascension of bacteria along the women reproductive tract. Therefore, in humans, bacteria commonly detected in follicular fluid include for instance *Lactobacillus*, *Propionibacterium* and *Gardnerella* (Pelzer, Allan, et al. 2013; Wu et al. 2023), which are normal bacteria of the women's lower reproductive tract (Chen et al. 2017; Pelzer, Willner, Buttini, Hafner, et al. 2018). Follicular fluid microbiota also contain bacteria commonly detected in the gastrointestinal tract, skin or oral mucosa (Pelzer, Allan, et al. 2013), which may spread via hematogenous dissemination.

Several studies have shown that the microbiota of follicular fluid may be a fundamental modulator of host reproduction (Pelzer et al. 2011; Pelzer, Allan, et al. 2013; Dong et al. 2023; Wu et al. 2023). For instance, in women, the presence of *Lactobacillus* in follicular fluids has been associated with higher pregnancy rates (Pelzer, Allan, et al. 2013; Wu et al. 2023) and, in mouse and pigs, the presence of several microorganisms compromise oocyte quality or maturation (Sluss and Reichert 1983; Pelzer, Harris, et al. 2013). Mechanisms include deleterious effects of bacteria on DNA fragmentation within oocytes (Pelzer, Harris, et al. 2013; Usman et al. 2021). Therefore, unraveling the factors that governs follicular fluid microbiota may be crucial for the ethiology of infertility.

The microbiota is shaped by multiple environmental and host factors, including diet, habitat, sex, age and host genetics (Grice and Segre 2011; Kers et al. 2018; Rothschild et al. 2018). Among the host genetic factors, several immune genes may play a critical role, because they encode for proteins involved with detecting specific bacteria products (Marietta et al. 2015). The major histocompatibility complex (MHC; called HLA in humans), one of the most polymorphic group of genes in vertebrates, is crucial for the recognition of self and non-self protein antigens by T-cells. Consequently, MHC genes are also involved in defence against bacterial pathogens and tolerance towards symbiotic bacteria (Roland et al. 2020). As shown in several vertebrates including humans, inter-individual variation in MHC genotype may thus drive inter-individual difference in microbiota (Toivanen et al. 2001; Bolnick

et al. 2014a; Marietta et al. 2015; Khan et al. 2019; Andeweg et al. 2021a). However, although this association has been studied for the human gut microbiota (Andeweg et al. 2021a), the association between the MHC and the human reproductive tract microbiota remains unknown.

Here we investigated the association between HLA profiles and follicular fluid microbiota in women. First, because individuals with higher MHC diversity are expected to recognize a broader range of antigens (Pierini and Lenz 2018), we expected microbiota diversity to decrease with MHC diversity, as shown in human gut, fish gut and bird feathers (Bolnick et al. 2014b; Leclaire et al. 2019; Andeweg et al. 2021a). Second, because two individuals with the same HLA genotypes are hypothesized to have similar immune control of their microbiota, we expected similarity in microbiota composition between individuals to covary with HLA similarity.

Material and methods

Sample collection

The participants are 27 women, with mean age 31.7 ± 2.7 s.e. years, who underwent transvaginal follicular aspiration for *in vitro* fertilization. Among these women, 3 were sampled twice at several months interval. Women were recruited via three fertility clinics: North Karelia Central Hospital (Joensuu, Finland, $n = 20$), ISLAB (Kuopio, Finland, $n = 4$) and InOva (Kuopio, Finland, $n = 10$) between April 2017 and December 2020. All the participants were Caucasian except one who was Asian. However, since the inclusion of this Asian woman did not affect our results, we included this sample in our final dataset. All results remained similar if we exclude the Asian woman from the analysis. Thus, we report results including her. Transvaginal follicular punctures were performed under local anesthesia, using ultrasound guidance. Prior to follicular fluid collection, follicle maturation was hyperstimulated with follicle-stimulating hormone, and premature ovulation was prevented using a gonadotrophin-releasing hormone antagonist. When the diameter of the largest follicle reached 18–20 mm, human chorion gonadotrophin was administered. After collection, the follicular fluid was centrifuged at $500 \times g$ for 10 min. The supernatant was aliquoted and stored in liquid nitrogen for later use. Nine of the women had been diagnosed with female factor infertility (i.e., anovulation, polycystic ovarian syndrome, or endometriosis; $n = 9$ women), whereas 18 women were not diagnosed with female factor infertility problem (e.g., unexplained couple infertility or pair infertility due to the man; $n = 18$ women). All the women donated a blood sample for genetic analyses.

Because the follicular fluid is obtained transvaginally, and it is thus prone to bacterial contamination from the lower genital tract, we collected additional follicular fluid samples and vaginal

swabs from four women. Before follicular fluid collection, a sterile cotton swab (FLOQSwabs, COPAN, California) was inserted in the vagina to collect mucus from the area of the transvaginal puncture. Swabs were stored in sterile 5 ml cryotubes.

Microbiota analyses

To characterize the bacterial community present in each sample, we performed 16S rRNA amplicon high-throughput sequencing. PCR amplification, Miseq library preparation and amplicon sequencing were performed at Genome Quebec, Montréal, Canada. Staggered primers 515F (Caporaso et al. 2011) and 806R (Caporaso et al. 2011) were used to amplify the variable region V4 of 16S rRNA gene from bacteria. We included four DNA extraction blank controls and three PCR blank controls to detect and withdraw potential contaminants. Follicular fluid DNA was extracted in five different batches. To control for batch effect, one sample was repeatedly extracted in each batch. In addition, to control for repeatability in extraction, PCR and sequencing, DNA extraction was repeated twice for six follicular fluid samples.

We first analysed the sequence reads using the OBITools package (Boyer et al., 2016). Briefly, after assembly of paired-end reads, we assigned reads to their respective samples and excluded reads with low assembly scores or reads that contained ambiguous bases (i.e., “N”). We dereplicated strictly identical reads using the “obiuniq” algorithm (Boyer et al., 2016). We then analysed the sequences using the FROGS 4.1 pipeline (Escudié et al., 2018). Briefly, ASVs were formed using the swarm denoising method with the fastidious method with $d = 1$ (Mahé et al. 2022). Chimeras were removed using VSEARCH (Rognes et al., 2016). Taxonomic assignment was then performed using blast affiliation and the SILVA v138 database (Quast et al., 2012). ASVs sharing same taxonomy with at least 99% identity and 99% alignment coverage were aggregated together. We removed ASVs that were taxonomically classified as mitochondria, chloroplasta or Archaea. We also removed ASVs that had fewer than 10 reads ($n = 50$ ASVs) or that were detected in a single sample ($n = 22$ ASVs). Finally, we used the isContaminant function with the prevalence option (package decontam in R) to remove potential contaminants from the data set (Davis et al. 2018). 14 ASVs were found to be contaminants. They belonged to the genus *Escherichia-Shigella* or the family *Burkholderiaceae*. We found that the number of reads of the remaining OTUs belonging to *Escherichia-Shigella* were similar in follicular fluid samples and extraction controls (t-test: $t_{13} = -1.5$, $P = 0.10$). It has been suggested that reads attributed to *Escherichia-Shigella* may represent remnant DNA from the *Escherichia coli* used to produce *Taq* DNA polymerase (Stinson et al. 2019). Thus, we decided to exclude all ASVs belonging to *Escherichia-Shigella* ($n = 5$ additional ASVs). In addition, using the one sample that was extracted at least once in each

extraction batch, we observed high variation in the relative abundance of *Acinetobacter* between replicates (Fig. S1). When considering all follicular fluid samples, the relative abundance of *Acinetobacter* was found to vary with the DNA extraction batch ($\chi^2_4 = 29$, $P < 0.00001$; Fig. S1). *Acinetobacter* are frequent cleanroom contaminants because they resist sterilization procedures and disinfectants (Ridgway and Olson 1982; Hendrickson et al. 2021). In a conservative way, we decided therefore to exclude all OTUs belonging to *Acinetobacter* from the analyses ($n = 53$ ASVs). We obtained mean \pm SE per follicular fluid sample: $51\,518 \pm 11\,512$ sequences (median: 20 388 sequences; range: 1 480 – 255 176 sequences). DNA extraction replicates were repeatable (Permanova: $F_{6,10} = 5.35$, $r^2 = 0.76$, $P < 0.001$; Fig. S2) and were therefore averaged.

To analyse α -diversity of follicular fluid microbiota, we used the weighed phylogenetic diversity (Faith's PD) based on the nonrooted phylogenetic tree (function "pd" in package picante in R). Faith's PD is the sum of total phylogenetic length of OTUs in each sample. Prior calculating Faith's diversity, we rarefied the dataset to the smallest library size (i.e., 1480 sequences). While there have been considerable debates concerning rarefaction in microbiome datasets (McMurdie and Holmes 2014; Weiss et al. 2017; Cameron et al. 2021), the wide variability in our library sizes, ranging from 1 480 to 255 176 sequences in follicular fluid samples, justifies our decision to rarefy. However, because rarefaction can result in a substantial loss of reads (McMurdie and Holmes 2014) and can introduce a stochastic component to the analysis because the rarefied taxa count table depends on the specific rarefaction replicate, we used multiple rarefactions, which has been proposed to recover some of the information lost through a single rarefaction (Cameron et al. 2021).

To analyse β -diversity, we used the phylogenetic isometric log-ratio transform (PhILR) which allows analysis of compositional data where the parts can be related through a phylogenetic tree (Silverman et al., 2017). PhILR has been suggested to be a replacement for the weighted UniFrac distances when accounting for the compositional nature of the data is needed (Gloor et al., 2017). OTU sequences were aligned using Decipher (Wright, 2015). Then a maximum-likelihood tree was constructed using the package phangorn (Schliep, 2011). This tree was used as a reference from which the Euclidean distances were calculated on phylogenetic isometric log-ratio transformed OTU abundance (i.e., the PhILR transform) between samples (Egozcue et al., 2003; Gloor et al., 2017) using the package PhILR (Silverman et al., 2017). In the PhILR transform, we added a pseudocount of 0.1 reads. Rarefaction is unnecessary when using PhILR transformation (Silverman et al. 2017).

HLA genotyping

DNA of all the women was extracted from EDTA blood using automatized Qiasymphony SP instrument (Qiagen) or PureLink® Genomic DNA Kit (Invitrogen). Extracted DNA samples were genotyped on Illumina GlobalScreeningArray-24v2-0 + Multi-Disease beadchip at the Institute for Molecular Medicine Finland (FIMM). All genotypes were named with GenomeStudio 2.0.3 software, and a subset was checked manually based on pre-determined selection criteria such as low call rates, bad cluster separation, low signal intensity, quality scores, and heterozygote excess. The parameters used in the quality control of X chromosomal and autosomal markers were: call rate < 97%; cluster separation < 0.3; AB R Mean (low intensity SNPs) < 0.3; AB T Mean (SNPs where two clusters are close to each other) < 0.2 or > 0.8; heterozygote excess < -0.3 (not X chromosome) or > 0.2; AA or BB T Dev > 0.05; and AB T Dev (too many or CNV like clusters) > 0.07. PLINK v1.90b6.6 (Chang et al. 2015) (www.cog-genomics.org/plink/1.9/) was used to manage and filter the genotyping data. For genotyping quality control, the following settings were applied: missing genotype frequency (geno) < 0.05, *P*-value for deviation from Hardy–Weinberg equilibrium > 1e-5; minor allele frequency > 0.01; individual missing genotype rate (mind) < 0.05.

Imputation of seven classical HLA genes (*HLA-A*, *-B*, *-C*, *-DRB1*, *-DQA1*, *-DQB1*, and *-DPB1*) at four-digit (i.e., protein level) resolution was carried out in R v3.4.4 (R Core Team 2019), using the package HIBAG v1.18.1 (Zheng et al. 2014) with default settings and the Finnish reference panel (Ritari et al. 2020) (<https://github.com/FRCBS/HLA-imputation>) for the human genome build GRCh37. The one Asian individual was imputed as described above except using the HIBAG Asian reference panel (https://hibag.s3.amazonaws.com/download/hlares_param/Asian-HLA4.html) instead of the Finnish panel.

HLA diversity of the women was calculated as (1) the number of heterozygous sites and (2) the mean Sandberg divergence over the seven imputed HLA genes per women. The Sandberg divergence is the pairwise Euclidean distance between the two protein sequences represented by the five physico-chemical z-descriptors (Sandberg et al. 1998). For HLA genes, the Sandberg distance has been found to correlate with the total number of peptides bound by the alleles (Pierini and Lenz 2018). The Sandberg divergence at each HLA gene was calculated using the DistCalc function in the MHCTools package in R (Roved et al. 2022). Number of heterozygous sites and Sandberg divergence were correlated (Spearman correlation test: $S = 837$, $r = 0.74$, $P < 0.0001$). Dissimilarity between each women pair was estimated (1) on a pure shared/non-shared allele basis and (2) by mean the Sandberg distance between HLA alleles. HLA dissimilarity based on a pure shared/non-shared allele basis (hereafter called allelic dissimilarity) was calculated as $100 \times (1 - (\text{the number of shared alleles} / \text{the total number of alleles}))$. For each HLA gene, Sandberg distance was calculated by constructing a UPGMA dendrogram

of alleles (function `hclust` in R), which represents clusters of functionally similar HLA sequences. Then HLA distance between two women was estimated using the tree distance (function `treedist` in package *vegan*), which is found by combining species in a common dendrogram and seeing how much of the dendrogram height is shared and how much is unique. The allelic dissimilarity was correlated to the Sandberg dissimilarity (Mantel test: $r = 0.91$, $P = 0.001$).

Whole genome diversity and similarity analysis

The whole-genome similarity measure between each woman-woman combination was calculated as the total sum of the number of shared genotypes for all genotyped and quality-filtered bi-allelic single nucleotide polymorphisms. In other words, genotyped variants were averaged in a single value, which prevented any single variants from dominating the genome-wide similarity estimate, which in turn can be expected to increase the accuracy of the overall estimate. Genome wide heterozygosity was not correlated with HLA diversity (Spearman rank correlation test: $S = 2514$, $r = 0.23$, $P = 0.24$ for the number of heterozygous HLA sites, and $S = 2194$, $r = 0.33$, $P = 0.09$ for the Sandberg divergence, $n = 27$ women). In contrast, whole genome dissimilarity was positively associated with HLA dissimilarity (Mantel test: $r = 0.30$, $P = 0.002$ for allelic dissimilarity and $r = 0.24$, $P = 0.024$ for Sandberg distance).

Statistical analyses

To determine whether HLA diversity was associated with microbiome α -diversity, we used linear mixed models. The DNA extraction batch was included as a random variable. For each analysis, 500 tests were run. Each test corresponded to a new rarefaction round (Cameron et al. 2021). In addition, for each of the three women who were sampled twice, only one of their two samples was randomly sampled in each rarefaction round to avoid pseudoreplication. In the statistical section, we reported the mean of the statistic parameters over the 500 tests. The DNA extraction batch was not included as a covariate, because it had no effect on the microbiome diversity ($F_{3,23} = 0.91$, $P = 0.46$). We characterized the bacterial species preferentially associated with variations in HLA diversity using ANCOM-BC on the non-rarefied dataset (`ancombc2` function in the ANCOM package) (Lin and Peddada 2020).

To determine whether women who were more similar at HLA sites had more similar microbiome composition, we used partial Mantel test. A binomial matrix of DNA extraction batch similarity (0: the two samples were extracted in the same batch and 1: the two samples were not extracted in the same batch) was included as a covariate, because women whose DNA was extracted in the same extraction batch had more similar microbiome composition (mantel test: $r = 0.14$, $P =$

0.016). To avoid pseudoreplication, for each analysis, we ran 8 different tests, each test including a unique combination of one sample of the females sampled twice. In the statistical section, we reported the mean of the statistic parameters over the 8 tests.

We performed all statistical analyses using the R statistical software (R Core Team, 2020).

Results

Characterization of the follicular fluid microbiome

Considering the 34 follicular fluid samples, the most abundant species belonged to the genus *Lactobacillus* and included *Lactobacillus iners*, *Lactobacillus jensenii*, *Lactobacillus acidophilus* / *L. crispatus*, *Lactobacillus gasseri* / *L. johnsonii*. Other abundant species were *Gardnerella vaginalis*, *Sneathia amnii*, *Prevotella sp*, *Atopobium vaginae*, *Kocuria sp.* and *Alishewanella sp.* (Fig. S3). ANCOM-BC analyses at the species level showed that follicular fluid microbiomes with higher Faith's diversity were characterized by higher abundance of *Gardnerella vaginalis*, *Atopobium vaginae*, *Sneathia amnii*, *Prevotella sp*, *Megasphaera_sp* and lower abundance of *Lactobacillus jensenii*, *Lactobacillus acidophilus* / *L. crispatus* and *Lactobacillus gasseri* / *L. johnsonii* (Fig. S3). The composition and α -diversity of follicular fluid microbiome were not affected by the female factor infertility diagnoses (permanova: $F_{1,22} = 0.58$, $r^2 = 0.02$, $P = 0.87$ and Wilcoxon rank sum test: $W = 61$, $P = 0.33$, respectively).

When focusing on the four women for whose we had both follicular fluid and vaginal swab samples, we found no differences in the composition or diversity of microbiome between sample types (PERMANOVA: $F_{1,7} = 613$, $r^2 = 0.09$ $P = 0.77$ and paired t-test: $t_3 = 0.58$, $P = 0.60$, respectively ; Fig. S4). When investigating the OTUs that were present in follicular fluid samples but absent in vaginal swabs, we found however that OTUs belonging to the genera *Alishewanella*, *Lawsonella* and *Kocuria* were absent from all vaginal swab samples but were present in all (100%), 3 (75%) or 1 (25%) follicular fluid samples respectively, albeit in low relative abundance (0.26 ± 0.19 %, 0.34 ± 0.07 % and 0.54 % respectively, when considering the samples where the OTUs were present). When considering the 30 follicular fluid samples for which we have HLA genotype, we found that *Alishewanella* and *Lawsonella* were present in all samples, and *Kocuria* in all except one sample, and represented (5.01 ± 1.4 %, 1.33 ± 0.39 % and 1.99 ± 0.95 %, respectively).

Microbiome α -diversity and HLA diversity

Women who had a higher heterozygosity or Sandberg divergence at HLA genes had a higher α -diversity in follicular fluid microbiome ($\chi^2_{24} = 8.9$, $P = 0.007$ and $\chi^2_{24} = 9.93$, $P = 0.005$; Fig. S5). This relationship was also significant when considering HLA class I genes only ($\chi^2_{24} = 11.53$, $P = 0.003$ and $\chi^2_{24} = 4.97$, $P = 0.038$; Fig. 2). In contrast, when considering HLA class II genes only, the association between HLA diversity and microbiome diversity was significant only for Sandberg divergence (heterozygosity: $\chi^2_{24} = 3.73$, $P = 0.072$ and Sandberg divergence: $\chi^2_{24} = 7.95$, $P = 0.011$; Fig. S5). ANCOM-BC analyses carried out at the bacterial species level showed that women who were heterozygous at the three class I genes had lower relative abundance of *Lactobacillus jensenii* and *Lactobacillus acidophilus* / *L. crispatus*, but higher relative abundance of *Sneathia amnii*, *Atopobium vaginae*, *Megasphaera sp* and *Prevotella bivia* than women who had at least one homozygous class I genes (Fig. 3). There was no association between follicular fluid microbiome α -diversity and genome-wide heterozygosity ($\chi^2_{24} = 0.01$, $P = 0.94$; Fig. S6).

Microbiome similarity and HLA similarity

Women who were more distant at HLA genes did not have more dissimilar follicular fluid microbiome (allelic distance: $r = 0.08$, $P = 0.22$ and Sandberg distance: $r = 0.07$, $P = 0.22$; Fig. S7). However, when we focused on the women with no female factor infertility ($n = 18$ women; 9 women excluded), we found that women with more distant HLA genes had more dissimilar follicular fluid microbiome (allelic distance: $r = 0.34$, $P = 0.0058$ and Sandberg distance: $r = 0.33$, $P = 0.0079$; Fig. 4). This effect was significant for both HLA Class I genes ($r = 0.30$, $P = 0.022$ and $r = 0.25$, $P = 0.043$ for allelic and Sandberg distance respectively) and HLA class II genes ($r = 0.27$, $P = 0.012$ and $r = 0.30$, $P = 0.0081$ for allelic and Sandberg distance respectively). Women with a higher genome-wide similarity did not have more similar follicular fluid microbiome ($r = 0.09$, $P = 0.36$ when considering all women, and $r = -0.02$, $P = 0.61$ when focusing on the women with no female factor infertility).

Discussion

Using high-throughput sequencing of 16S rRNA genes, we characterized the follicular fluid microbiome of 27 Finnish women visiting infertility clinics. The most abundant OTUs detected in follicular fluid belonged to the genera *Lactobacillus*, *Gardnerella*, *Prevotella*, *Atopobium* and *Sneathia*. This result contrasts with the few other studies that characterized the follicular fluid microbiota of women. However, these studies showed no consensus on the core microbial composition. For instance, a study using 16S sequencing found that the follicular fluid microbiome of infertile women was dominated by *Achromobacter*, and that *Lactobacillus* was not the dominant genus (Wu et al. 2023). *Achromobacter*

is an environmental contaminant widely distributed in soils and waters and commonly detected in hospitals (Amoureux et al. 2013). Other studies using cultivation methods detected *Lactobacillus*, but also *Actinomyces*, *Propionobacterium*, *Bifidobacterium*, *Staphylococcus*, *Streptococcus* and *Enterococcus* (Pelzer, Allan, et al. 2013; Usman et al. 2021; Kim et al. 2022), which are normal bacteria of the skin, the oral mucosa and the gastrointestinal tract (Stackebrandt et al. 2006). Recently, several studies have suggested that low-biomass microbial niches, such as the human fetus, the placenta or the endometrium, are sensitive to contamination and data misinterpretation (Olomu et al. 2020; Molina et al. 2021; Kennedy et al. 2023). The follicular fluid, like other sites of the upper reproductive tract, have low microbial biomass, and clearly, the original source of the bacteria detected in our study and others can be questioned. The major bacteria that we detected in follicular fluid (e.g., *Lactobacillus*, *Gardnerella*, *Prevotella*, *Atopobium* and *Sneathia*) are common inhabitants of the woman vagina (Gajer et al. 2012; Tabatabaei et al. 2019; France et al. 2022), and we did not detect differences between the composition of follicular fluid microbiome and vaginal microbiome in the four women who were sampled at both sites. We cannot exclude therefore that the follicular fluid has been contaminated with vaginal bacterial strains at the time of puncture for transvaginal oocyte aspiration. For instance, *Lactobacillus* abundance is known to be increased in samples collected transcervically or transvaginally compared to samples collected using hysterectomy, laparoscopy or cesarean section (Winters et al. 2019; Canha-Gouveia et al. 2023). However, we observed that *Alishewanella*, *Lawsonella* and *Kocuria* were present in follicular fluid samples, but in none of the vaginal swab samples. These genera are thus not contaminants from the vagina, and might be true inhabitants of follicular fluid. Similarly, a previous culture-based study detected *Kocuria* in follicular fluid, but not in vagina (Kim et al. 2022), and another study found that *Kocuria* were increased in endometrium compared to endocervix (Pelzer, Willner, Buttini, and Huygens 2018). *Alishewanella* and *Lawsonella* have previously been detected in the women reproductive tract, including the vagina and the endometrium (Hong et al. 2016; Hernandez et al. 2020; Kuznetsova et al. 2021; Liu et al. 2022). However, clearly, bacterial contamination is a big challenge in the studies of follicular fluid microbiome.

Whatever the true origin of the microbiome studied (vaginal vs. follicular fluid), we report the first evidence that the women reproductive microbiome varied with HLA diversity (heterozygosity and divergence). We found that women with higher HLA diversity had a higher α -diversity in follicular fluid microbiome. This result contrasts with the widespread hypothesis that a higher diversity of MHC molecules should lead to a larger number of foreign peptides recognized and thus to a larger number of bacterial species eliminated. Accordingly, individuals with higher HLA class II genes diversity have a broader antibody repertoire against common microbial pathogens (Khan et al. 2022a), and microbiome diversity is negatively associated with MHC diversity in the human gut (Andeweg et al.

2021b), fish gut (Bolnick et al. 2014a) and bird feathers (Leclaire et al. 2019). However, MHC molecules may also promote benign host-bacteria interactions (Roland et al. 2020). For instance, in the gut of mice, the presence of MHC-II on B-cells promote the proliferation of beneficial bacteria and increased microbiota diversity (Alexeev 2021). Therefore, a positive association between MHC and microbiome diversity may also arise if a higher diversity of MHC molecules facilitate the proliferation of a more diverse range of symbiotic microbes (Schubert et al. 2021). Similar to our results, microbiota diversity increased with MHC diversity in the skin of two amphibians (Hernández-Gómez et al. 2018; Belasen et al. 2021).

We found that women with higher HLA diversity have higher abundance of *Sneathia amnii*, *Prevotella bivia*, *Atopobium vaginae*, and *Megasphaera* and lower abundance of *Lactobacillus jensenii* and *Lactobacillus acidophilus* / *L. crispatus*. In the vagina, this highly diverse microbial composition, described as “CST-IV”, is commonly associated with bacterial vaginosis (reviewed in Smith and Ravel 2017), a disorder often causing vaginal discharge and foul-smelling odor. Bacterial vaginosis is associated with sexually transmitted infections and adverse reproductive outcome, including pre-term birth, low birthweight and miscarriage (reviewed in Chen et al. 2021). Bacterial vaginosis plays also a role in the acquisition of human papillomavirus and the development of associated cervical neoplasia (Gillet et al. 2012). Interestingly, the positive association between HLA diversity and a CST-IV microbiome, might therefore explain the surprising positive association between HLA diversity and increased risk of cervical cancer (Wang et al. 2023).

Although high reproductive microbiome diversity is considered deleterious for fitness, high HLA diversity is conversely expected to be beneficial for fitness, because it confers enhanced resistance to several pathogens (Planelles et al. 2006; Khan et al. 2022b) and reduces risks of several cancers (Wang et al. 2023). The positive association between HLA diversity and follicular fluid microbiome diversity in women raises therefore a potential conflict between the fitness benefits of a highly diverse HLA and of a poorly diverse reproductive microbiome. However, the effect of the CST-IV microbial state on reproduction is mixed and, in some women, this microbial state is asymptomatic (David A Eschenbach et al. 1988). Some studies suggest that CST-IV state increases the risk of adverse outcome only in interaction with a specific genotype (i.e., gene-environment interaction). For instance, a study showed that, bacterial vaginosis per se is not associated with an increased risk of spontaneous preterm delivery, but its association with susceptible genotypes at several immune genes (i.e., interleukin 6, protein kinase C alpha and fms-like tyrosine kinase 1) increases the risk (Gómez et al. 2010). We suggest that further studies should investigate whether the negative outcome associated with a CST-IV microbial communities varies with HLA diversity. For instance, HLA-B heterozygotes have higher tolerance to HIV, with high viral load but minimal disease progression (Regoes et al. 2014). The

potential evolutionary conflict between HLA diversity and reproductive microbiota diversity might be solved if HLA diversity confers higher tolerance to bacterial vaginosis.

We also found that women who are more similar at HLA genes host more similar follicular fluid microbiome. Similarly, humans with similar HLA genes have similar gut microbiota (Andeweg et al. 2021a). However, the association between similarity at HLA gene and similarity in microbiome composition was significant only for women who had no female factor infertility. Although we found no differences in the microbiome composition of women with female factor infertility and that of healthy women, previous studies have detected that polycystic ovary syndrome or endometriosis alters the reproductive microbiota (Hong et al. 2020; Jiang et al. 2021). Our result may suggest that, compared to HLA genes, other factors might be stronger determinants of the reproductive microbiome of women with female factor infertility.

In conclusion, our study reveals that the HLA is an important genetic factor driving differences in reproductive microbiotas between women. Unraveling whether the HLA effects on microbiota drive variations in reproductive outcome would be an important next step.

Acknowledgements

We thank Ulrica Candolin for hosting SL during her stay at the University of Helsinki. This study was funded by grants from the Fondation Fyssen (to SL) and Academy of Finland (grant 308485 to JK).

References

Alexeev S. 2021. B-Cell-Intrinsic MHCII Signaling Shapes Microbiota Composition.

Amoureux L, Bador J, Fardeheb S, Mabilhe C, Couchot C, Massip C, Salignon A-L, Berlie G, Varin V, Neuwirth C. 2013. Detection of *Achromobacter xylosoxidans* in hospital, domestic, and outdoor environmental samples and comparison with human clinical isolates. *Appl Environ Microbiol*. 79(23):7142–7149.

Andeweg SP, Keşmir C, Dutilh BE. 2021a. Quantifying the Impact of Human Leukocyte Antigen on the Human Gut Microbiota. *mSphere*. 6(4):e00476-21.

Andeweg SP, Keşmir C, Dutilh BE. 2021b. Quantifying the Impact of Human Leukocyte Antigen on the Human Gut Microbiota. *mSphere*. 6(4):e00476-21.

Belasen AM, Riolo MA, Bletz MC, Lyra ML, Toledo LF, James TY. 2021. Geography, Host Genetics, and Cross-Domain Microbial Networks Structure the Skin Microbiota of Fragmented Brazilian Atlantic Forest Frog Populations. *Ecol Evol*. 11(14):9293–9307. doi:10.1002/ece3.7594.

Bolnick DI, Snowberg LK, Caporaso JG, Lauber C, Knight R, Stutz WE. 2014a. Major histocompatibility complex class IIb polymorphism influences gut microbiota composition and diversity. *Mol Ecol.* 23(19):4831–4845.

Bolnick DI, Snowberg LK, Caporaso JG, Lauber C, Knight R, Stutz WE. 2014b. Major histocompatibility complex class IIb polymorphism influences gut microbiota composition and diversity. *Mol Ecol.* 23(19):4831–4845.

Cameron ES, Schmidt PJ, Tremblay BJ-M, Emelko MB, Müller KM. 2021. Enhancing diversity analysis by repeatedly rarefying next generation sequencing data describing microbial communities. *Sci Rep.* 11(1):22302. doi:10.1038/s41598-021-01636-1.

Canha-Gouveia A, Pérez-Prieto I, Rodríguez CM, Escamez T, Leonés-Baños I, Salas-Espejo E, Prieto-Sánchez MT, Sánchez-Ferrer ML, Coy P, Altmäe S. 2023. The female upper reproductive tract harbors endogenous microbial profiles. *Front Endocrinol.* 14. [accessed 2023 Sep 20]. <https://www.frontiersin.org/articles/10.3389/fendo.2023.1096050>.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, Fierer N, Knight R. 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci.* 108(supplement_1):4516–4522.

Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience.* 4(1):s13742-015.

Chen C, Song X, Wei W, Zhong H, Dai J, Lan Z, Li F, Yu X, Feng Q, Wang Z, et al. 2017. The microbiota continuum along the female reproductive tract and its relation to uterine-related diseases. *Nat Commun.* 8(1):875. doi:10.1038/s41467-017-00901-0.

Chen X, Lu Y, Chen T, Li R. 2021. The Female Vaginal Microbiome in Health and Bacterial Vaginosis. *Front Cell Infect Microbiol.* 11. [accessed 2023 Sep 15]. <https://www.frontiersin.org/articles/10.3389/fcimb.2021.631972>.

Ciepiela P, Dulęba AJ, Kario A, Chełstowski K, Branecka-Woźniak D, Kurzawa R. 2019. Oocyte matched follicular fluid anti-Müllerian hormone is an excellent predictor of live birth after fresh single embryo transfer. *Hum Reprod.* 34(11):2244–2253.

David A Eschenbach, Hillier S, Critchlow C, Stevens C, DeRouen T, Holmes KK. 1988. Diagnosis and clinical manifestations of bacterial vaginosis. *Am J Obstet Gynecol.* 158(4):819–828.

Davis NM, Proctor DM, Holmes SP, Relman DA, Callahan BJ. 2018. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome.* 6(1):226. doi:10.1186/s40168-018-0605-2.

Dong Y-H, Fu Z, Zhang N-N, Shao J-Y, Shen J, Yang E, Sun S-Y, Zhao Z-M, Xiao A, Liu C-J, et al. 2023. Urogenital tract and rectal microbiota composition and its influence on reproductive outcomes in infertile patients. *Front Microbiol.* 14. doi:10.3389/fmicb.2023.1051437. [accessed 2023 Jun 15]. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9950574/>.

Dumesic DA, Meldrum DR, Katz-Jaffe MG, Krisher RL, Schoolcraft WB. 2015. Oocyte environment: follicular fluid and cumulus cells are critical for oocyte health. *Fertil Steril.* 103(2):303–316.

- Einenkel R, Zygmunt M, Muzzio DO. 2019. Microorganisms in the healthy upper reproductive tract: from denial to beneficial assignments for reproductive biology. *Reprod Biol.* 19(2):113–118. doi:10.1016/j.repbio.2019.04.001.
- Fitzpatrick JL, Willis C, Devigili A, Young A, Carroll M, Hunter HR, Brison DR. 2020. Chemical signals from eggs facilitate cryptic female choice in humans. *Proc R Soc B Biol Sci.* 287(1928):20200805. doi:10.1098/rspb.2020.0805.
- France M, Alizadeh M, Brown S, Ma B, Ravel J. 2022. Towards a deeper understanding of the vaginal microbiota. *Nat Microbiol.* 7(3):367–378. doi:10.1038/s41564-022-01083-2.
- Gajer P, Brotman RM, Bai G, Sakamoto J, Schütte UME, Zhong X, Koenig SSK, Fu L, Ma Z, Zhou X, et al. 2012. Temporal Dynamics of the Human Vaginal Microbiota. *Sci Transl Med.* 4(132):132ra52. doi:10.1126/scitranslmed.3003605.
- Gillet E, Meys JF, Verstraelen H, Verhelst R, De Sutter P, Temmerman M, Broeck DV. 2012. Association between bacterial vaginosis and cervical intraepithelial neoplasia: systematic review and meta-analysis. [accessed 2023 Sep 26]. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0045201>.
- Gómez LM, Sammel MD, Appleby DH, Elovitz MA, Baldwin DA, Jeffcoat MK, Macones GA, Parry S. 2010. Evidence of a gene-environment interaction that predisposes to spontaneous preterm birth: a role for asymptomatic bacterial vaginosis and DNA variants in genes that control the inflammatory response. *Am J Obstet Gynecol.* 202(4):386.e1-386.e6. doi:10.1016/j.ajog.2010.01.042.
- Grice EA, Segre JA. 2011. The skin microbiome. *Nat Rev Microbiol.* 9(4):244–253.
- Hendrickson R, Urbaniak C, Minich JJ, Aronson HS, Martino C, Stepanauskas R, Knight R, Venkateswaran K. 2021. Clean room microbiome complexity impacts planetary protection bioburden. *Microbiome.* 9(1):238. doi:10.1186/s40168-021-01159-x.
- Hernandes C, Silveira P, Rodrigues Sereia AF, Christoff AP, Mendes H, Valter de Oliveira LF, Podgaec S. 2020. Microbiome Profile of Deep Endometriosis Patients: Comparison of Vaginal Fluid, Endometrium and Lesion. *Diagnostics.* 10(3):163. doi:10.3390/diagnostics10030163.
- Hernández-Gómez O, Briggler JT, Williams RN. 2018. Influence of immunogenetics, sex and body condition on the cutaneous microbial communities of two giant salamanders. *Mol Ecol.* 27(8):1915–1929. doi:10.1111/mec.14500.
- Hong KH, Hong SK, Im Cho S, Ra E, Han KH, Kang SB, Kim E-C, Park SS, Seong M-W. 2016. Analysis of the vaginal microbiome by next-generation sequencing and evaluation of its performance as a clinical diagnostic tool in vaginitis. *Ann Lab Med.* 36(5):441.
- Hong X, Qin P, Huang K, Ding X, Ma J, Xuan Y, Zhu X, Peng D, Wang B. 2020. Association between polycystic ovary syndrome and the vaginal microbiome: A case-control study. *Clin Endocrinol (Oxf).* 93(1):52–60. doi:10.1111/cen.14198.
- Jiang I, Yong PJ, Allaire C, Bedaiwy MA. 2021. Intricate Connections between the Microbiota and Endometriosis. *Int J Mol Sci.* 22(11):5644. doi:10.3390/ijms22115644.
- Jokiniemi A, Magris M, Ritari J, Kuusipalo L, Lundgren T, Partanen J, Kekäläinen J. 2020. Post-copulatory genetic matchmaking: HLA-dependent effects of cervical mucus on human sperm function. *Proc R Soc B Biol Sci.* 287(1933):20201682. doi:10.1098/rspb.2020.1682.

Kennedy KM, de Goffau MC, Perez-Muñoz ME, Arrieta M-C, Bäckhed F, Bork P, Braun T, Bushman FD, Dore J, de Vos WM, et al. 2023. Questioning the fetal microbiome illustrates pitfalls of low-biomass microbial studies. *Nature*. 613(7945):639–649. doi:10.1038/s41586-022-05546-8.

Kers JG, Velkers FC, Fischer EAJ, Hermes GDA, Stegeman JA, Smidt H. 2018. Host and Environmental Factors Affecting the Intestinal Microbiota in Chickens. *Front Microbiol*. 9. [accessed 2023 Jun 14]. <https://www.frontiersin.org/articles/10.3389/fmicb.2018.00235>.

Khan MW, Stephens WZ, Mohammed AD, Round JL, Kubinak JL. 2019. Does MHC heterozygosity influence microbiota form and function? *PLoS One*. 14(5):e0215946.

Khan T, Rahman M, Ahmed I, Al Ali F, Jithesh PV, Marr N. 2022a. Human leukocyte antigen class II gene diversity tunes antibody repertoires to common pathogens. *Front Immunol*. 13. [accessed 2023 Sep 20]. <https://www.frontiersin.org/articles/10.3389/fimmu.2022.856497>.

Khan T, Rahman M, Ahmed I, Al Ali F, Jithesh PV, Marr N. 2022b. Human leukocyte antigen class II gene diversity tunes antibody repertoires to common pathogens. *Front Immunol*.:4283.

Kim SM, Won KH, Hong YH, Kim SK, Lee JR, Jee BC, Suh CS. 2022. Microbiology of Human Follicular Fluid and the Vagina and Its Impact on in Vitro Fertilization Outcomes. *Yonsei Med J*. 63(10):941–947.

Kim Y-S, Kim M-S, Lee S-H, Choi B-C, Lim J-M, Cha KY, Baek K-H. 2006. Proteomic analysis of recurrent spontaneous abortion: identification of an inadequately expressed set of proteins in human follicular fluid. *Proteomics*. 6(11):3445–3454.

Kuznetsova NB, Борисовна КН, Bushtyreva Ю, Олеговна БИ, Дубова VS, Сергеевна ДВ, Varinova VV, Владиславовна БВ, Polev DE, Евгеньевич ПД, et al. 2021. VAGINAL MICROBIOME IN PREGNANT WOMEN WITH PRETERM PRELABOR RUPTURE OF MEMBRANES AT 22-28 WEEKS' GESTATION. *Obstet Gynecol*.(1):94–102. doi:10.18565/aig.2021.1.94-102.

Leclaire S, Strandh M, Dell'Araccia G, Gabirot M, Westerdahl H, Bonadonna F. 2019. Plumage microbiota covaries with the major histocompatibility complex in blue petrels. *Mol Ecol*. 28(4):833–846.

Lin H, Peddada SD. 2020. Analysis of compositions of microbiomes with bias correction. *Nat Commun*. 11(1):3514. doi:10.1038/s41467-020-17041-7.

Liu C-J, Xiao W-Y, Fang J-F, Dong Y-H, Ye K-F, He M-P, Wang Y-S, Li X, Zhao Z-M, Yuan T. 2022. Genital microbiota of women from six ethnic groups with and without human papillomavirus infection in shangri-la, China. *Front Cell Infect Microbiol*. 12:935068.

Mahé F, Czech L, Stamatakis A, Quince C, de Vargas C, Dunthorn M, Rognes T. 2022. Swarm v3: towards tera-scale amplicon clustering. *Bioinformatics*. 38(1):267–269.

Marietta E, Rishi A, Taneja V. 2015. Immunogenetic control of the intestinal microbiota. *Immunology*. 145(3):313–322. doi:10.1111/imm.12474.

McMurdie PJ, Holmes S. 2014. Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Comput Biol*. 10(4):e1003531.

Molina NM, Sola-Leyva A, Haahr T, Aghajanova L, Laudanski P, Castilla JA, Altmäe S. 2021. Analysing endometrial microbiome: methodological considerations and recommendations for good practice. *Hum Reprod*. 36(4):859–879. doi:10.1093/humrep/deab009.

- Olomu IN, Pena-Cortes LC, Long RA, Vyas A, Krichevskiy O, Luellwitz R, Singh P, Mulks MH. 2020. Elimination of “kitome” and “splashome” contamination results in lack of detection of a unique placental microbiome. *BMC Microbiol.* 20(1):157. doi:10.1186/s12866-020-01839-y.
- Pelzer ES, Allan JA, Cunningham K, Mengersen K, Allan JM, Launchbury T, Beagley K, Knox CL. 2011. Microbial colonization of follicular fluid: alterations in cytokine expression and adverse assisted reproduction technology outcomes. *Hum Reprod.* 26(7):1799–1812. doi:10.1093/humrep/der108.
- Pelzer ES, Allan JA, Waterhouse MA, Ross T, Beagley KW, Knox CL. 2013. Microorganisms within Human Follicular Fluid: Effects on IVF. *PLOS ONE.* 8(3):e59062. doi:10.1371/journal.pone.0059062.
- Pelzer ES, Harris JE, Allan JA, Waterhouse MA, Ross T, Beagley KW, Knox CL. 2013. TUNEL analysis of DNA fragmentation in mouse unfertilized oocytes: The effect of microorganisms within human follicular fluid collected during IVF cycles. *J Reprod Immunol.* 99(1):69–79. doi:10.1016/j.jri.2013.07.004.
- Pelzer ES, Willner D, Buttini M, Hafner LM, Theodoropoulos C, Huygens F. 2018. The fallopian tube microbiome: implications for reproductive health. *Oncotarget.* 9(30):21541–21551. doi:10.18632/oncotarget.25059.
- Pelzer ES, Willner D, Buttini M, Huygens F. 2018. A role for the endometrial microbiome in dysfunctional menstrual bleeding. *Antonie Van Leeuwenhoek.* 111(6):933–943. doi:10.1007/s10482-017-0992-6.
- Pierini F, Lenz TL. 2018. Divergent Allele Advantage at Human MHC Genes: Signatures of Past and Ongoing Selection. *Mol Biol Evol.* 35(9):2145–2158. doi:10.1093/molbev/msy116.
- Planelles D, Nagore E, Moret A, Botella-Estrada R, Vila E, Guillén C, Montoro JA. 2006. HLA class II polymorphisms in Spanish melanoma patients: homozygosity for HLA-DQA1 locus can be a potential melanoma risk factor. *Br J Dermatol.* 154(2):261–266. doi:10.1111/j.1365-2133.2005.06896.x.
- R Core Team. 2019. R: A language and environment for statistical computing.
- Regoes RR, McLaren PJ, Battegay M, Bernasconi E, Calmy A, Günthard HF, Hoffmann M, Rauch A, Telenti A, Fellay J. 2014. Disentangling human tolerance and resistance against HIV. *PLoS Biol.* 12(9):e1001951.
- Ridgway HF, Olson BH. 1982. Chlorine resistance patterns of bacteria from two drinking water distribution systems. *Appl Environ Microbiol.* 44(4):972–987. doi:10.1128/aem.44.4.972-987.1982.
- Ritari J, Hyvärinen K, Clancy J, FinnGen, Partanen J, Koskela S. 2020. Increasing accuracy of HLA imputation by a population-specific reference panel in a FinnGen biobank cohort. *NAR Genomics Bioinforma.* 2(2):lqaa030. doi:10.1093/nargab/lqaa030.
- Roland MM, Mohammed AD, Kubinak JL. 2020. How MHCII signaling promotes benign host-microbiota interactions. *PLOS Pathog.* 16(6):e1008558. doi:10.1371/journal.ppat.1008558.
- Rothschild D, Weissbrod O, Barkan E, Kurilshikov A, Korem T, Zeevi D, Costea PI, Godneva A, Kalka IN, Bar N, et al. 2018. Environment dominates over host genetics in shaping human gut microbiota. *Nature.* 555(7695):210–215. doi:10.1038/nature25973.

- Roved J, Hansson B, Stervander M, Hasselquist D, Westerdahl H. 2022. MHCtools—an R package for MHC high-throughput sequencing data: Genotyping, haplotype and supertype inference, and downstream genetic analyses in non-model organisms. *Mol Ecol Resour.* 22(7):2775–2792.
- Salary A, Kafi M, Derakhshandeh A, Moezzi MS. 2020. Detection of bacteria in bovine ovarian follicular fluid. *Lett Appl Microbiol.* 70(3):137–142.
- Sandberg M, Eriksson L, Jonsson J, Sjöström M, Wold S. 1998. New chemical descriptors relevant for the design of biologically active peptides. A multivariate characterization of 87 amino acids. *J Med Chem.* 41(14):2481–2491.
- Schubert N, Nichols HJ, Winternitz JC. 2021. How can the MHC mediate social odor via the microbiota community? A deep dive into mechanisms. *Behav Ecol.* 32(3):359–373.
- Silverman JD, Washburne AD, Mukherjee S, David LA. 2017. A phylogenetic transform enhances analysis of compositional microbiota data. *Elife.* 6:e21887.
- Sluss PM, Reichert LE. 1983. Presence of Bacteria in Porcine Follicular Fluid and Their Ability to Generate an Inhibitor of Follicle-Stimulating Hormone Binding to Receptor. *Biol Reprod.* 29(2):335–341. doi:10.1095/biolreprod29.2.335.
- Smith SB, Ravel J. 2017. The vaginal microbiota, host defence and reproductive physiology. *J Physiol.* 595(2):451–463. doi:10.1113/JP271694.
- Stackebrandt E, Cummins CS, Johnson JL. 2006. Family propionibacteriaceae: the genus *Propionibacterium*. *The prokaryotes.* 3:400–418.
- Stinson LF, Keelan JA, Payne MS. 2019. Identification and removal of contaminating microbial DNA from PCR reagents: impact on low-biomass microbiome analyses. *Lett Appl Microbiol.* 68(1):2–8. doi:10.1111/lam.13091.
- Tabatabaei N, Eren A, Barreiro L, Yotova V, Dumaine A, Allard C, Fraser W. 2019. Vaginal microbiome in early pregnancy and subsequent risk of spontaneous preterm birth: a case–control study. *BJOG Int J Obstet Gynaecol.* 126(3):349–358. doi:10.1111/1471-0528.15299.
- Toivanen P, Vahtovuori J, Eerola E. 2001. Influence of major histocompatibility complex on bacterial composition of fecal flora. *Infect Immun.* 69:2372–2377.
- Usman SF, Shuaibu IR, Durojaiye K, Medugu N, Iregbu KC. 2021. The presence of microorganisms in follicular fluid and its effect on the outcome of in vitro fertilization-embryo transfer (IVF-ET) treatment cycles. *PLOS ONE.* 16(2):e0246644. doi:10.1371/journal.pone.0246644.
- Wang Q-L, Wang T-M, Deng C-M, Zhang W-L, He Y-Q, Xue W-Q, Liao Y, Yang D-W, Zheng M-Q, Jia W-H. 2023. Association of HLA diversity with the risk of 25 cancers in the UK Biobank. *eBioMedicine.* 92. doi:10.1016/j.ebiom.2023.104588. [accessed 2023 Sep 26]. [https://www.thelancet.com/journals/ebiom/article/PIIS2352-3964\(23\)00153-6/fulltext](https://www.thelancet.com/journals/ebiom/article/PIIS2352-3964(23)00153-6/fulltext).
- Weiss S, Xu ZZ, Peddada S, Amir A, Bittinger K, Gonzalez A, Lozupone C, Zaneveld JR, Vázquez-Baeza Y, Birmingham A, et al. 2017. Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome.* 5(1):27. doi:10.1186/s40168-017-0237-y.

Winters AD, Romero R, Gervasi MT, Gomez-Lopez N, Tran MR, Garcia-Flores V, Pacora P, Jung E, Hassan SS, Hsu C-D, et al. 2019. Does the endometrial cavity have a molecular microbial signature? *Sci Rep.* 9(1):9905. doi:10.1038/s41598-019-46173-0.

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. Microbiological composition of follicular fluid in patients undergoing IVF and its association with infertility. *Am J Reprod Immunol.* 89(3):e13652.

Zheng X, Shen J, Cox C, Wakefield JC, Ehm MG, Nelson MR, Weir BS. 2014. HIBAG—HLA genotype imputation with attribute bagging. *Pharmacogenomics J.* 14(2):192–200.

Figure 1: Relative abundance of the most abundant bacterial species in follicular fluid. Each bar corresponds to one of the 27 women sampled.

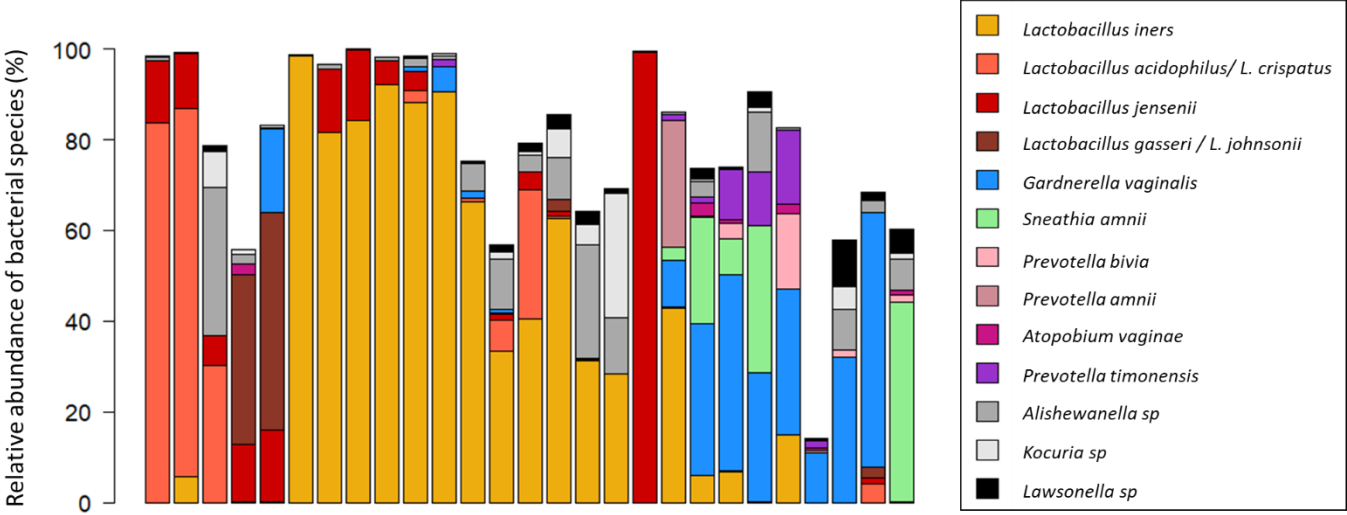


Figure 2: Microbiota diversity (Faith's PD) in follicular fluid according to (a) the number of heterozygous HLA Class_I sites and (b) the Sandberg divergence in HLA Class I sites. The red, green and blue dots correspond to the three women sampled twice. Only one of their two samples was considered in the statistical analyses during each of the 5000 rarefaction rounds. For (a), similar results were found when pooling all the women with at least one homozygous site. The black line and grey shadow in (b) represent the linear trend and confidence interval.

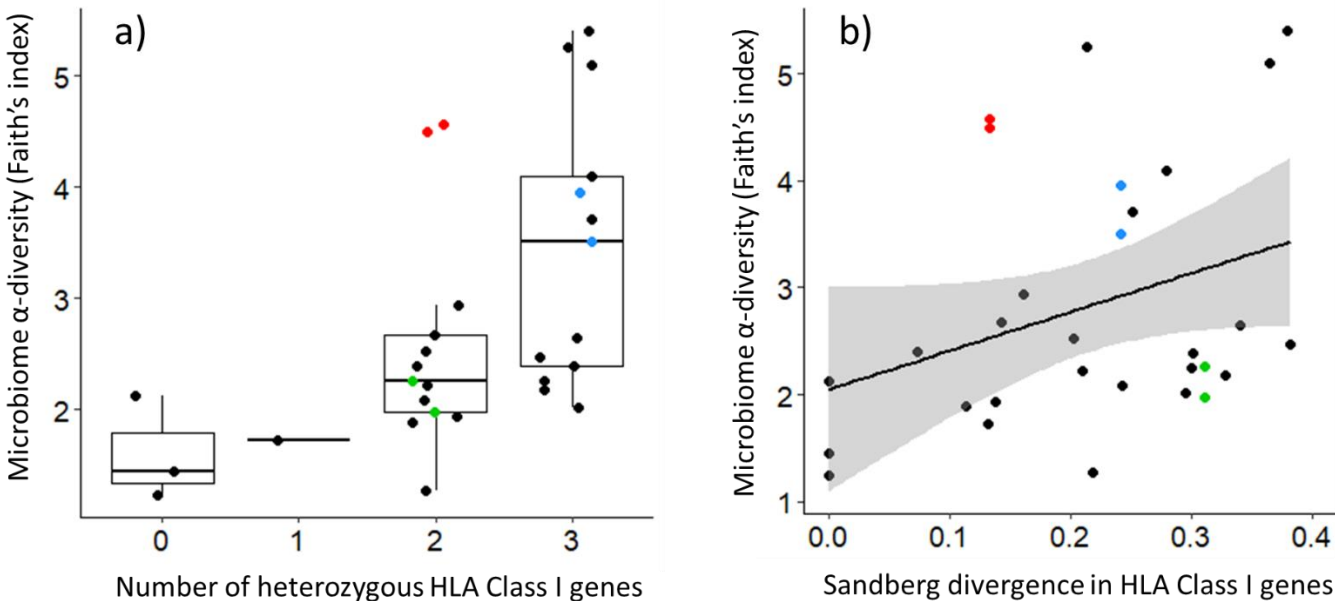


Figure 3: Results of ANCOMBC analyses, as shown by the log fold change of the 6 bacterial species that were differently abundant between women who were heterozygous at the three HLA class I genes and women who were heterozygous at < 3 HLA class I genes.

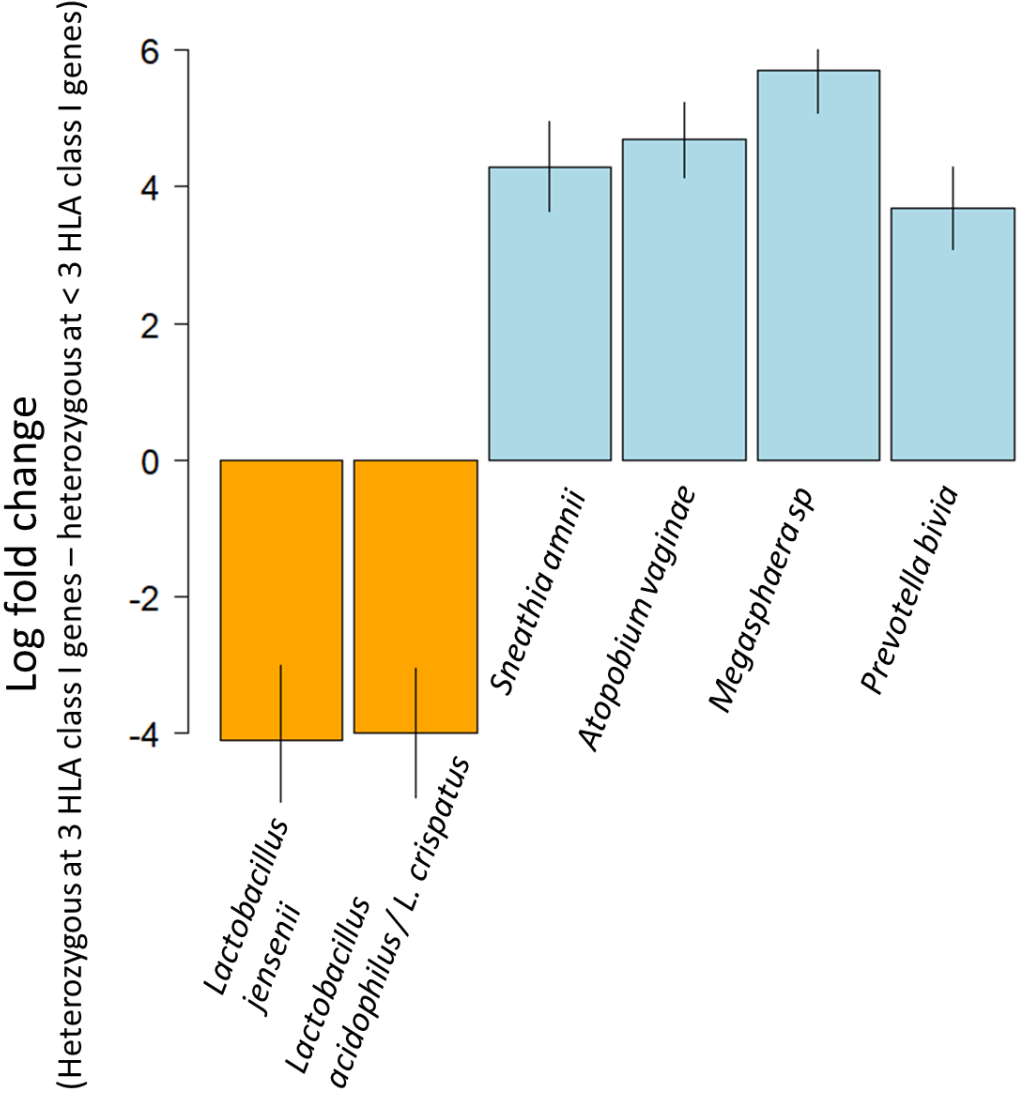


Figure 4 : Dissimilarity in microbiome composition in women pairs with no infertility diagnoses according to their dissimilarity in HLA class I and II genes. Dissimilarity in HLA genes was estimated (a) on a pure shared-non shared allele basis and (b) using Sandberg dissimilarity. The black line and grey shadow represent the linear trend and confidence interval. The dot at 0 HLA dissimilarity corresponds to two women who had the same HLA genotype.

