PAPER

Check for updates

Diet–microbiota interactions influence pregnancy success in females undergoing artificial insemination: insights from the vaginal microbiota and Mediterranean diet

M. Gimeno, [ID] †[a] M. Baizán-Urgell,†[b,c] M. Bernabeu, [ID] *[b] M. C. Collado [ID] ‡[b] and E. Garcia-Verdevio‡[a]

The vaginal microbiota plays a key role in female fertility, yet its interaction with diet and lifestyle remains unclear. In this study, we aimed to assess the influence of the Mediterranean diet (MD) adherence and the vaginal microbial profile on pregnancy outcomes in females undergoing artificial insemination (AI). Vaginal swabs from 104 participants were analysed using amplicon-based sequencing to assess microbial diversity and composition in relation to pregnancy status and MD adherence. The overall pregnancy rate following artificial insemination was 23.07% (8.65% resulted in pregnancy loss before reaching full term). Dominant species clustered by CSTs, with CST II–V showing typical *Lactobacillus* dominance, whereas CST IV-B was enriched in *G. vaginalis* and *A. vaginae*. Importantly, all CST IV-B women failed to achieve pregnancy, and only a minority of CST V women conceived. Random forest modelling using CST, AI number, MD adherence, AI type, BMI, and age achieved moderate predictive performance for pregnancy, with high sensitivity but low specificity. CSTs, pregnancy status, number of IAs and adherence to the MD explained the greatest proportion of variation in the vaginal microbiota structure by RDA. Pregnant women with high MD adherence harbored higher abundance of *Prevotella lymphophilum*, *Anaerotignum massiliense*, and *Micrococcus radiotolerans*, whereas *G. vaginalis* characterized non-pregnant women with low MD adherence. Pregnant women showed lower diversity than non-pregnant women. Among women that got pregnant, those who subsequently miscarried exhibited distinct microbial profiles and reduced diversity. Specific taxa such as *Aerococcus mediterraneensis*, *Streptococcus mitis*, *Peptoniphilus lacrimalis*, *Staphylococcus hominis* were enriched in miscarriage cases, whereas *Winkia neui*, *Bacillus mojavensis*, and a member of Pseudomonota phyla (previously Proteobacteria) were associated with full-term successful pregnancies. These findings highlight the interaction between diet and vaginal microbiota in determining reproductive outcomes, indicating that the MD would affect fertility by modulating microbial communities.

Introduction

The western lifestyle is impacting women's health including their fertility and microbiomes, with potential long-term consequences.[1,2] Factors such as stress levels,[3] unhealthy diets,[4] lack of physical activity,[5] exposure to environmental toxins,[6] and use of antibiotics[7] can disturb the women's microbiota balance of the vaginal, gut, and reproductive tract, all of them essential for reproductive health. Furthermore, the western lifestyle is linked to an increase in the age at first pregnancy, which occurs alongside natural hormonal and fertility changes, and is associated with a higher risk of maternal-neonatal comorbidities and adverse health outcomes.[8]

Maternal age is the most important factor in a woman's reproductive capacity.[9–11] Studies demonstrate that as a woman ages, the number of follicles capable of development and maturation decreases, as does the quality of the eggs.[12] This premise assumes a decrease in birth rates and an increase in embryonic abnormalities rate and defects during fetal development.[13] According to data recently published by the World Health Organization (WHO), up to 48 million couples and 186 million individuals could suffer infertility worldwide,[14]

[a]*Department of Gynecology and Obstetrics, Hospital Universitario Doctor Peset, Valencia, 46017, Spain*
[b]*Department of Biotechnology, Institute of Agrochemistry and Food Technology-National Research Council (IATA-CSIC), Valencia, Spain.*
E-mail: mbernabeu@iata.csic.es
[c]*PhD Programme in Biotechnology, Universitat Politècnica de Valencia, Camí de vera S/N 46022, Valencia, Spain*
†Shared first authorship, equal contribution.
‡Senior authors, equal contribution.

which shows the magnitude of the problem we face and justifies the increase in the number of medical consultations related to these disorders in recent times. Within this framework, we find that in Spain, more than 5% of women aged 18 to 55 have undergone some type of Assisted Reproduction Technique (ART). Among them, artificial insemination (AI) can be the best option for fertility treatment due to its simplicity, affordability, and effectiveness. However, it has a low success rate (7–12%).[15]

It has been suggested that analysing the vaginal microbiome before treatment could help improve the success rate of ARTs. If the microbiome profile is not optimal, treatment could be postponed and adjusted using antibiotics, prebiotics, or probiotics to create a more favourable environment.[16,17] An increasing number of publications are demonstrating the role of the microbiota in maintaining an adequate state of health.[18] Diet is a major determinant of human microbial composition. Evidence shows that adherence to a Mediterranean diet (MD) increases gut microbial diversity and enriches genera such as *Blautia*, *Bifidobacterium*, *Faecalibacterium*, and members of Ruminococcaceae and Lachnospiraceae, largely due to higher intake of plant foods, legumes, and n-3 PUFAs from fish.[19] Although fewer studies have explored dietary effects on the vaginal microbiota, available data indicate that nutritional exposures may modulate vaginal communities. For example, fish-oil or prebiotic supplementation has been associated with reduced Ureaplasma abundance, and healthy dietary patterns rich in fiber or specific micronutrients have been linked to lower risk of bacterial vaginosis.[20] These findings suggest that diet-related metabolites may influence microbial environments relevant to reproductive health and fertility.

The vaginal ecosystem is one of the most changing at different times in a woman's life,[21] greatly influenced by the stages of the hormonal cycle, as well as by pregnancy or menopause.[22] The main bacterial families in this niche have been classified into five groups, the so-called community state types or CSTs, and each of them has different microorganism compositions. However, in four of the five, various *Lactobacillus* species are predominant.[23] The predominance of *Lactobacillus* is hallmark of vaginal health, as these bacteria contribute to pH regulation by producing lactic acid and hydrogen peroxide and thus maintaining an acidic environment that is hostile to the growth and development of infectious agents.[24,25] Moreover, *Lactobacillus* exert additional protective effects by producing a wide range of antimicrobial compounds, including bacteriocins, bacteriocin-like molecules, and biosurfactants, and by inducing host autophagy pathways that promote the clearance of intracellular bacteria, protozoa, and viruses.[26]

Among *Lactobacillus* species, *L. crispatus* has been associated with greater vaginal stability and protection against infections, whereas dominance of *L. iners* is linked to increased susceptibility to colonisation by pathogenic microorganisms.[24,27] These contrasting effects are partly explained by differences in lactic acid isomer production: with the exception of *L. iners*, most *Lactobacillus* species produce both D- and L-lactic acid, with D-lactic acid conferring greater protection against vaginal dysbio-sis.[28] Accordingly, D-lactic acid levels peak in *L. crispatus*-dominated communities and are lowest in microbiota enriched in *L. iners*, *Gardnerella vaginalis*, and other vaginal dysbiosis–associated taxa associated to lower conception rates, including *in vitro* fertilization (IVF) failure and miscarriage.[29]

This study investigates how adherence to the Mediterranean diet modulates the vaginal microbiota and influences reproductive outcomes in women undergoing artificial insemination. By identifying microbial profiles associated with pregnancy success or loss, our aim is to position Mediterranean diet adherence as a modifiable and accessible factor capable of supporting fertility through microbiota-mediated pathways.

Materials and methods

Study participants and design

This study included 104 females aged 18 to 38 years diagnosed with primary infertility, defined by the failure to achieve a pregnancy after 12 months or more of regular unprotected sexual intercourse (https://www.who.int/news-room/fact-sheets/detail/infertility). Participants underwent artificial insemination, with a maximum of 4 cycles per woman when using partner's semen and up to 6 cycles with donor semen. All AI cycles were hormonally stimulated with exogenous follicle-stimulating hormone (FSH) according to the standard ovarian stimulation protocol of the clinic.

Participants were recruited from January 22, 2022, to June 26, 2023, at the Doctor Peset University Hospital in Valencia. Participants were included if they met the following criteria: female sex, aged between 18 and 38 years, first-time consultation at the Infertility Unit of our center, selection of artificial insemination (AI) as the assisted reproductive technique after fulfilling the clinical requirements for undergoing the procedure, and provision of signed informed consent prepared in accordance with the ethical standards of the Declaration of Helsinki. Women with a body mass index (BMI) greater than 40 were not eligible for treatment until they reduced their BMI to at least 35.

No distinction was made based on the type of semen used (partner or donor), the patient's relationship status or partner's gender (male, female, or no partner), or the cause of infertility (female, male, or unexplained). Likewise, no exclusion criteria were applied regarding underlying medical conditions that could contribute to infertility, such as endometriosis, polycystic ovary syndrome (PCOS), or male factor infertility. A total of 108 women were initially recruited; however, the final analysis included 104 participants who underwent a total of 289 AI cycles. Four patients were excluded: three refused the vaginal sample collection on the day of insemination, and one sample was lost during the collection or storage process.

Vaginal swabs were collected before insemination but on the same day. All samples were stored in a freezer at −18 °C and transported in a portable cooler. Dietary habits and adherence to the Mediterranean diet were assessed using the 14-item Mediterranean diet questionnaire originally developed

and validated in the PREDIMED study[30] (Table S1). The questionnaire evaluates the consumption frequency of key components of the Mediterranean diet, such as olive oil, fruits, vegetables, legumes, fish, and red meat. A score below 9 was considered low adherence, while a score of 9 or higher was classified as high adherence.

DNA extraction and targeted 16S rRNA amplicon sequencing

DNA was extracted from the samples using the Maxwell® RSC PureFood GMO and Authentication Kit (Promega, Cat. No. AS1600), following the manufacturer's protocol with additional sample preparation steps to enhance DNA yield and purity.

Prior to extraction, the excess wooden sticks from each swab were trimmed to allow proper insertion into SARSTEDT system extraction tubes (SARSTEDT, Ref. 62.610.201). A volume of 600 µL of sterile phosphate-buffered saline (PBS) was added to each tube, and the samples were vortexed for 10 seconds. Tubes were then left to stand at room temperature for 24 hours to facilitate the release of material from the swab into the solution.

After incubation, tubes were vortexed again and centrifuged at 13 000g for 5 minutes. The swab was carefully removed, the supernatant discarded, and the pellet resuspended. An aliquot of 300 µL was used for DNA extraction using the Maxwell® RSC system. Extracted DNA was quantified using a Qubit® Fluorometer (Thermo Fisher Scientific) and normalized to 15 ng µL^{-1} prior to sequencing.

Sequencing was performed using a paired-end run on an Illumina platform (FISABIO sequencing service, Valencia, Spain), following the manufacturer's instructions.

Computational and statistical analysis

A DADA2 pipeline was used to achieve quality filtering, sequence joining, and chimera removal. ASVs mapping to the human genome (GRCh$_3$8) using the Burrow-Wheeler Aligner[31] in Deconseq v0.4.3[32] were filtered out. Taxonomic assignment was conducted using the Silva v138.1 database. The Bayesian LCA-based Taxonomic Classification Method (BLCA)[33] was used to assign ASVs to species (assignation of 80% confidence). Samples with less than 1000 reads were removed from the study. The full ASV abundance table is available as SI Table S2.

Taxonomic profiles of the vaginal microbial communities were sorted into categories termed community state types (CSTs) and subCSTs through the nearest centroid-based tool VALENCIA (VAginaL community state typE Nearest CentroId clAssifier),[23] which assigns each sample to the nearest predefined community centroid based on taxonomic composition. Prior to classification, ASV tables and taxonomy files were converted to the VALENCIA input format using the official qiime2 conversion script.

We performed a constrained ordination using Redundancy Analysis (RDA) with the rda( ) function from the *vegan* R package. The model assessed the association between microbial community composition (ASV table) and host-related explanatory variables: Community State Type (CST),

pregnancy status, type of artificial insemination, number of insemination attempts, adherence to the Mediterranean diet, body mass index, and maternal age. Model selection was carried out using forward selection (ordiR2step( )), optimizing the adjusted R^2 criterion and testing the significance of additional terms using permutation tests (999 permutations).

To assess the ability to predict pregnancy based on vaginal microbiota and clinical factors, we applied a random forest classification model using the formula: Pregnancy ∼ CST + num_AI + MD_Adh + AI_Type + BMI + Age. Given the limited sample size ($N = 104$), we implemented 5-fold cross-validation to ensure robust performance estimates. Model tuning was conducted to optimize the mtry parameter (*i.e.*, the number of variables randomly sampled at each split), using the area under the ROC curve (AUC) as the primary performance metric.

To assess the effect of Mediterranean diet adherence on microbial composition while adjusting for the confounding influence of CST and pregnancy, we used PERMANOVA *via* the adonis2( ) function in *vegan*.[34] The analysis was based on Bray–Curtis distances and included 999 permutations. To visualise the species classified into each CST, a heatmap was generated using the pheatmap package. The ASV abundance table was filtered to retain the top 25 most prevalent ASVs based on presence across samples. Taxa were agglomerated at the species level using tax_glom( ) and transformed to relative abundance. Genus and species were combined into a single label and samples were ordered by CST. Annotations included CST, pregnancy status, and Mediterranean Diet adherence.

Ecological metrics alpha and beta diversity were employed to characterize microbial composition. Alpha-diversity indexes, richness (Chao1) and evenness indices (Shannon), were obtained using the "phyloseq" R package 1.48.0.[35] The analysis of the composition of microbiomes with bias correction (ANCOM-BC2)[36] was used for the composition of microbiomes in pregnancy status and Mediterranean diet adhesion and the results were plotted as a heatmap. Only log fold change (LFC) of differentially significant abundant taxa with $q < 0.05$ were represented. This analysis was also conducted in a subset of all pregnant women, to compare differences in the vaginal microbiota abundance between women that experienced miscarriage and those with full term pregnancies.

All statistical analyses were performed with R version 4.4.1, and figures were drawn with the "ggplot2" R package.[37] For all methods, *p*-values were adjusted for multiple comparisons using False Discovery Rate (FDR) based on Benjamini – Hochberg (BH). Freeman-Tukey test was used for categorical variables for calculating statistical significance. P values of <0.1 were considered significant.

Results

Characteristics of the population

Vaginal swabs were collected from 104 women undergoing AI as part of fertility treatment, with a mean age of 34.5 (SD = 3.32). Most participants (67.31%) had a normal body mass

index (BMI = 18.5–24.9) at the time of treatment. The majority of participants (71.15%) reported high adherence to the Mediterranean diet (MD) (average = 9.21, SD = 1.47), which was the predominant dietary pattern in the cohort. Among women aged ≤35 (n = 50), the pregnancy rate was 40% (20/50), whereas in women aged > 35 years (n = 54), the rate was 24% (13/54). However, the difference in pregnancy rates between age groups was not statistically significant (Fisher's exact test, p = 0.095, odds ratio = 2.09, 95% CI: 0.84–5.36) (Table 1).

Factors contributing to microbiota variation

A redundancy analysis (RDA) was performed to explore the association between microbiota composition and clinical vari-

Table 1 Characteristics of the study population

Characteristics	N. participants	Percentage (%)
Total	104	100
Age, years		
≤35 years	50	48.08
>35 years	54	51.92
Average age	34.53	
SD	3.32	
BMI		
Normal weight, 18.5–24.9	70	67.31
Overweight, 25–29.9	25	24.04
Obese, >30	9	8.65
Race		
Caucasian	90	86.54
Arab	2	1.92
Latin	12	11.54
Mediterranean diet adherence		
Low, <9 points	30	28.85
High, ≥9 points	74	71.15
Type of insemination		
Conjugal artificial insemination (CAI)	74	71.15
Donor artificial insemination (DAI)	30	28.85
Average MD adherence score	9.21	
SD (MD adherence)	1.47	
Gestation		
Not achieved	71	68.27
≤35 years	28	39.44
>35 years	43	60.56
Average age	34.23	
SD (age)	3.45	
High MD adherence	32	45.97
Low MD adherence	39	54.93
Average MD adherence score	8.22	
SD (MD adherence)	1.55	
Achieved	33	31.73
Miscarriage	9	8.65
Alive newborn	24	23.07
≤35 years	20	40
>35 years	13	24.07
Average age	34.24	
SD (age)	3.01	
High MD adherence	25	75.76
Low MD adherence	8	24.24
Average MD adherence score	9.21	
SD (MD adherence)	1.31	

ables. The model showed that CST (Community State Type) was the variable that explained the greatest proportion of variation in microbiota structure (adjusted R^2 = 2.7%, p = 0.001 ***) (Fig. 1a). In addition, pregnancy status contributed a further 0.8% of explained variance and was also statistically significant (p = 0.003 **). Other variables did not significantly improve the model once CST and pregnancy were accounted for. To assess the impact of adherence to the Mediterranean diet while adjusting for CST and pregnancy, a PERMANOVA was performed using Bray–Curtis distances. The model yielded an R^2 = 0.10364 and was statistically significant (p.adj = 0.001 ***), suggesting that diet adherence is associated with microbiota composition beyond the effects of CST and pregnancy.

A random forest model to predict pregnancy based on vaginal microbiota and clinical factors identified mtry = 2 as the optimal value. This implies that at each decision split in the forest, the model randomly considers two predictors out of the six available variables (CST, number of artificial inseminations, adherence to the Mediterranean diet, AI type, BMI, and age). The number and type of AI were the most influential predictors in this model. The model achieved an average ROC AUC of approximately 0.74 across the cross-validation folds (Fig. 1b). Importantly, it showed high sensitivity (∼0.90) but low specificity (∼0.39), indicating that it was more effective at identifying women who became pregnant than those who did not.

The most abundant species identified across the cohort included *L. crispatus*, *L. iners*, *G. vaginalis*, *L. jensenii*, *L. gasseri*, *L. crispatus*, *P. bivia*, among others (Fig. 1c). The distribution of these dominant species clustered based on CSTs. For example, we observed that most samples classified as CST II by the VALENCIA tool were dominated by *L. gasseri*, most samples classified as CST III were dominated by *L. iners* (and then *L. crispatus*), most samples classified as CST V were dominated by *L. crispatus* and *L. jensenii*, samples classified as CST IV-C were dominated by *L. crispatus* and samples classified as CST IV-B did not have a high relative abundance of lactobacilli but had a high relative abundance of *G. vaginalis*, *A. vaginae* and *P. bivia* among others. Also, all samples classified as CST IV-B corresponded to women which did not achieve pregnancy after AI, and only 3/13 of women classified as CST V achieved pregnancy (Fisher test p = 0.338).
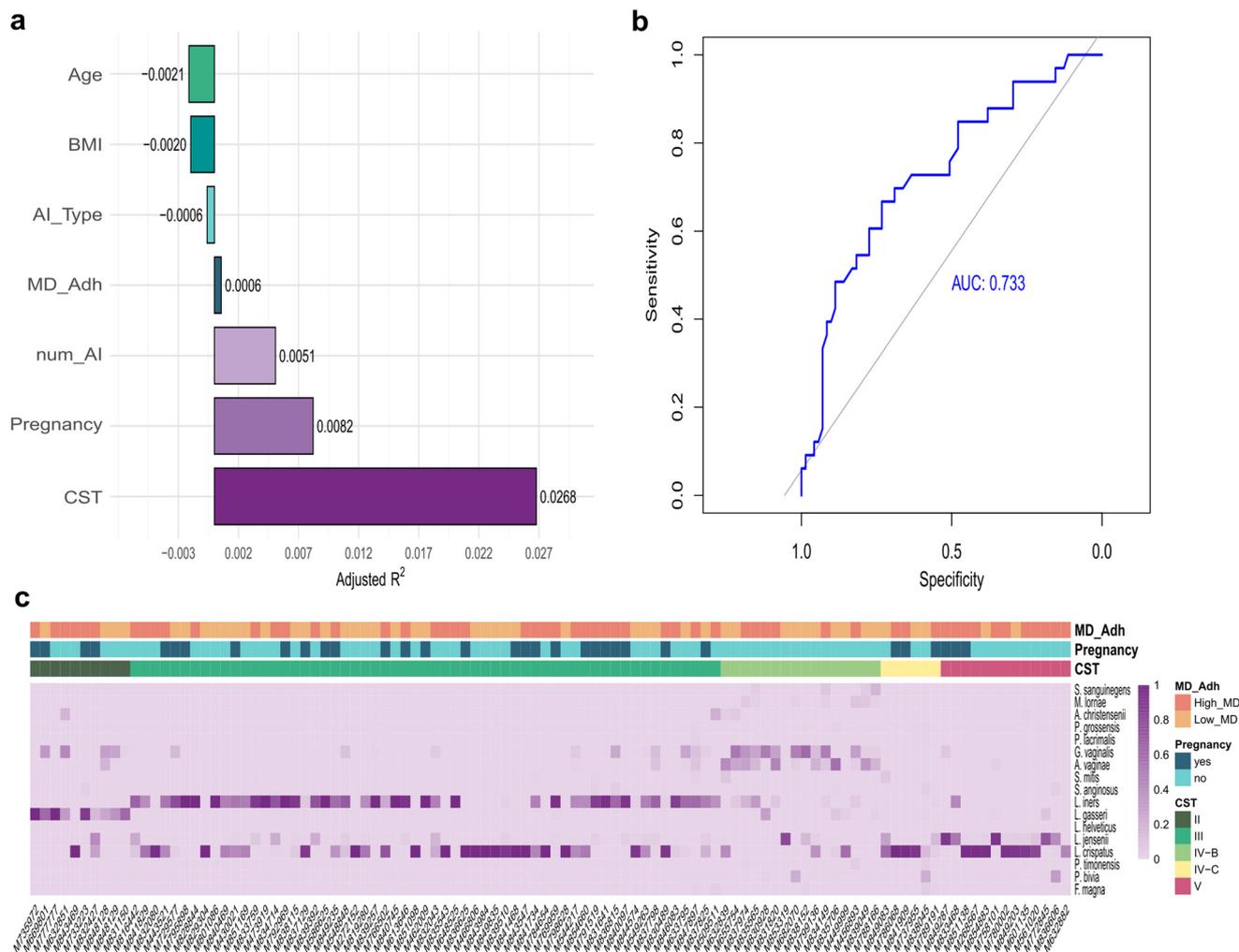
Does a woman's diet influence the vaginal microbiota and fertility outcomes?

Microbial diversity tended to be lower in non-pregnant women with low MD adherence compared to non-pregnant women with high MD adherence (Shannon index, q = 0.053), while richness remained similar between groups (Fig. 2a and b). The composition of the vaginal microbiota also varied according to pregnancy status and MD adherence, particularly in the relative abundance of *Lactobacillus* genus and *Gardnerella vaginalis*, with non-pregnant women showing a more diverse microbiota. Pregnant women displayed a higher relative abundance of *L. iners* (not statistically significant by ANCOM-BC analysis

Fig. 1 Factors shaping the vaginal microbiota of the studied population. (a) Variables included in the redundancy analysis (RDA) model with R^2 adjusted values explaining the proportion of variation in microbiota structure. (b) ROC curve of Random Forest model predicting pregnancy outcome using 5-fold cross-validation. The blue line shows model performance on out-of-fold predictions; the grey line represents random classification (AUC = 0.733). (c) Heatmap of the 25 most prevalent species and samples ordered based on CSTs, pregnancy and Mediterranean diet adherence.

at $q < 0.05$) and significantly decreased abundance of *G. vaginalis* compared to non-pregnant women, independent of MD adherence (Fig. 2c).

We represented the log-fold changes of taxa showing significant differences ($q < 0.05$) between groups based on pregnancy status and MD adherence, following analysis with ANCOM-BC. From greatest to least differences, *P. lymphophilum*, *A. massiliense*, *M. radiotolerans*, and *Acidibacter* genus were significantly more abundant in pregnant women with high MD adherence compared to non-pregnant women with low MD adherence. Among pregnant women, 13 taxa showed significant differences associated with MD adherence. Specifically, within the high MD adherence group, pregnant women exhibited higher abundances of *A. massiliense* and *P. lymphophilum*, and lower abundance of *L. crispatus* and *G. vaginalis*, among others. Remarkably, non-pregnant women with low MD adherence have the highest

abundance of *Gardnerella vaginalis*, even when compared to non-pregnant women with high MD adherence (Fig. 2d).

Distinct microbial shifts in term pregnancies compared to miscarriages

Significant differences were identified among pregnant women when comparing those who carried the pregnancy to term and those who experienced a miscarriage, defined as the spontaneous loss of pregnancy before 20 weeks of gestation. The miscarriage group had a lower diversity compared with the term pregnancy group (Shannon index, $q = 0.092$) although richness was similar (Fig. 3a and b). Significant differences in beta diversity were observed between the studied groups (PERMANOVA UniFrac unweighted, $R^2 = 0.050$, $q = 0.021$) (Fig. 3d–f). Several microbial taxa exhibited differential abundance between women who experienced miscarriage and those who did not. Term pregnancy women had a greater relative abundance of
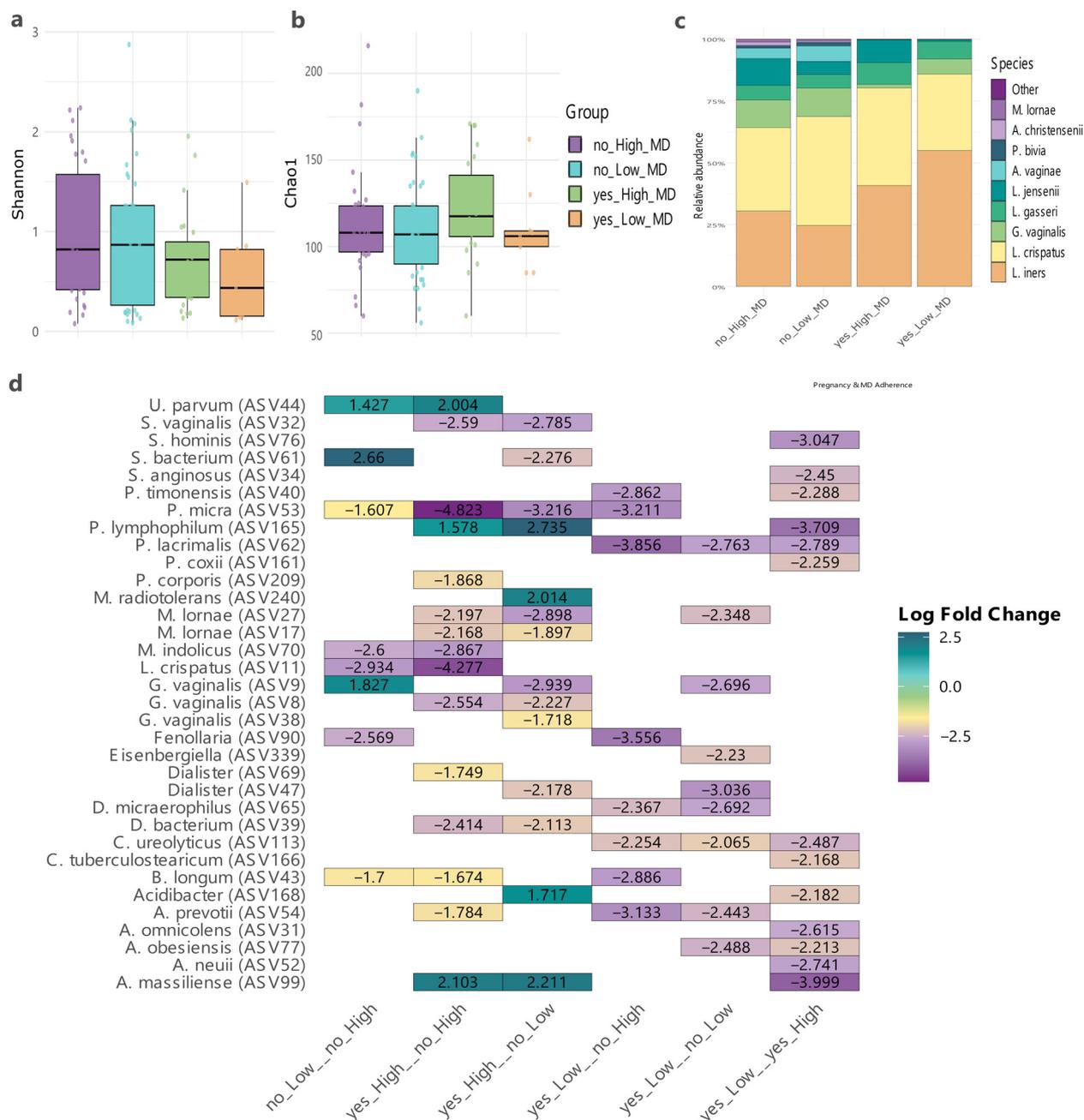
Fig. 2 Characteristics of the vaginal microbiota concerning adherence to MD and pregnancy status. (a) Shannon index. (b) Chao1 index. (c) Top 10 most abundant species. (d) Log fold change (LFC) heatmap of ASVs differentially abundant between all groups *versus* all groups. Pregnancy status is indicated by yes/no and Mediterranean Diet Adherence by High/Low. Only LFCs with *q* < 0.05 are represented.

L. crispatus, whereas women who experienced miscarriage showed an increased abundance of *L. iners*, *L. jensenii* and *L. gasseri* (Fig. 3c), although these differences were not statistically significant according to ANCOM-BC analysis at *q* < 0.05.

Specifically, *Winkia neui*, *Bacillus mojavensis*, and a member of *Pseudomonota* phyla (previously *Proteobacteria*) were enriched in women who carried their pregnancies to term, while *Aerococcus mediterraneensis*, *Streptococcus mitis*, *Peptoniphilus lacrimalis*, *Staphylococcus hominis*, and *Fenollaria* were more abundant in women with miscarriage (Fig. 3g).

These results suggest potential microbial signatures associated with early pregnancy loss.

Discussion

Our study highlights the importance of studying the vaginal microbiota composition in women following artificial insemination treatment and their diet. Among the 104 women included, 71.15% (74/104) showed high adherence to
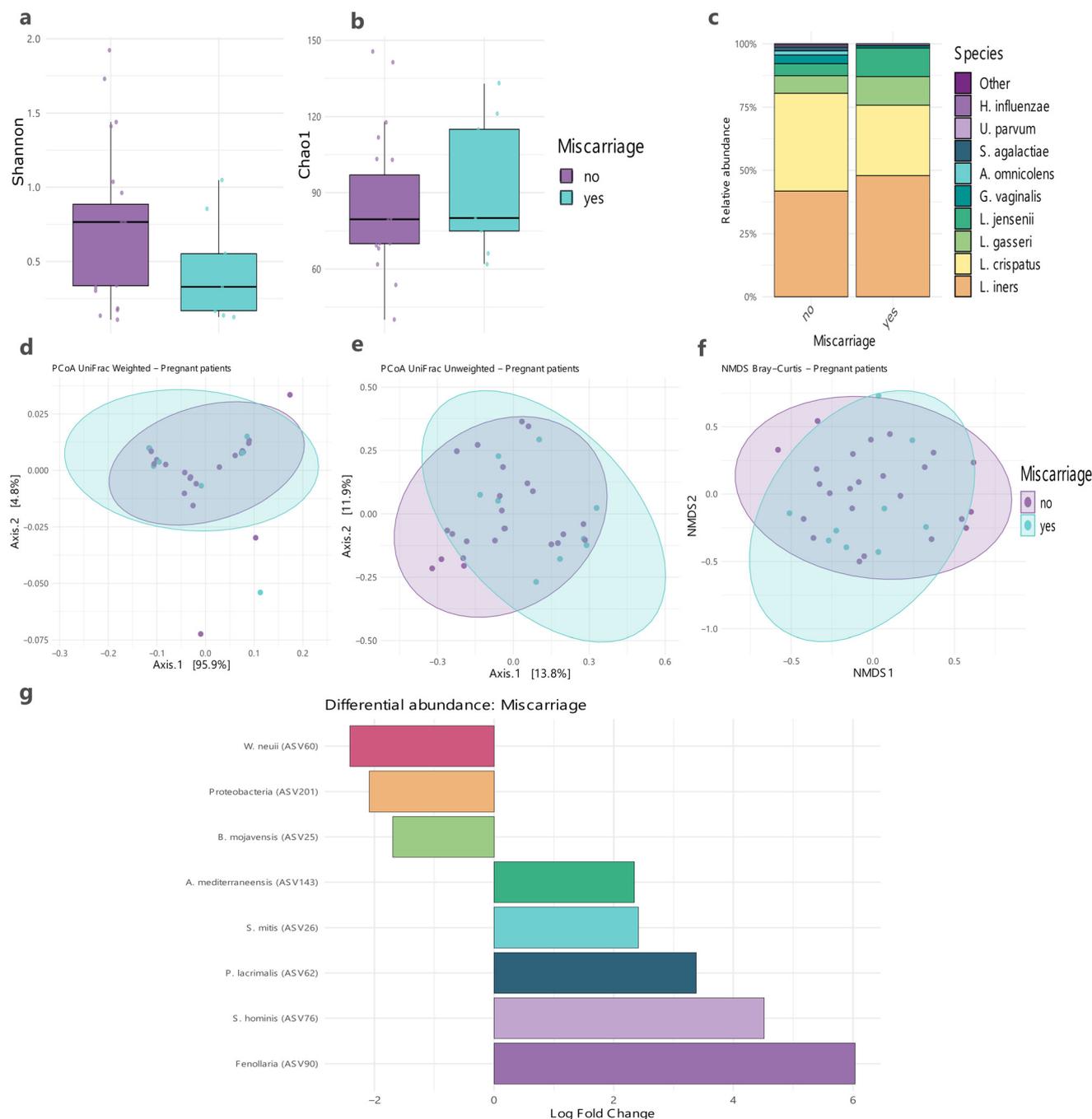
Fig. 3 Vaginal microbiota comparing term pregnancies and miscarriage. (*a*) Shannon index. (*b*) Chao1 index. (*c*) Top 10 species found in term pregnancies and miscarriage. (d) Weighted Unifrac distance Principal Component Analysis (PCoA). (e) Unweighted Unifrac distance PCoA. (f) Bray–Curtis NMDS. (g) Log fold change (LFC) barplot of ASVs differentially abundant in miscarriage compared to term pregnancies. Only LFCs with *q* < 0.05 are represented.

the Mediterranean diet, a significantly higher proportion than the 60% reported for the general Spanish female population (p = 0.013; one-sample proportion test), according to data published by the Spanish Ministry of Health (https://www.sanidad.gob.es/estadEstudios/sanidadDatos/tablas/tabla10.htm). Obesity prevalence in our cohort was 8.65% (9/104), lower than the 14% reported in national data, although this

difference did not reach statistical significance (p = 0.076; one-sample proportion test). A possible explanation for these differences is that women facing infertility may be more aware of the importance of maintaining a healthy lifestyle, particularly through proper nutrition.

Women included in this study exhibited a higher success rate after AI (31.73%) compared to data reported by the

Spanish Society of Fertility, which indicates a pregnancy rate of 15% overall and 10% for women aged 40 and older (https://www.registrosef.com/public/docs/sef2022_IAFIV.pdf). This discrepancy may be attributed to the study's selection criteria, which excluded women over 39 years and those with a BMI greater than 40, unless they reduced their BMI to at least 35 points prior to treatment. The observed differences in pregnancy rates by age group – 40% in women aged ≤35 years and 24% inwomen over 35 – are consistent with previous findings in this field.[38,39]

The random forest model showed good performance in terms of ROC AUC (∼0.74) indicating moderate predictive power of pregnancy. The selected value of mtry = 2 suggests that using a smaller subset of predictors at each tree node improved the model's discrimination ability, potentially by reducing overfitting in this small dataset. The high sensitivity (∼0.90) reflects the model's ability to correctly identify most pregnancies, which could be valuable in clinical screening. However, the low specificity (∼0.39) implies a relatively high rate of false positives, i.e., non-pregnant individuals predicted as pregnant. This pattern may arise due to a possible class imbalance or intrinsic biases in the model structure that favor the positive class ("yes"). Accordingly, the model should be primarily used as an exploratory tool to identify relevant predictors rather than as a clinically deployable classifier. Further model development with a larger dataset and improved class balance may help enhance predictive accuracy and clinical utility.

Our cohort showed overall concordance clustering on CSTs but with few distinctions compared to the VALENCIA reference dataset, where CST I is L. crispatus-dominated, CST II is L. gasseri-dominated, CST III is L. iners-dominated, and CST V is L. jensenii-dominated.[23] Notably, CST V communities were not exclusively dominated by L. jensenii but frequently included a strong contribution from L. crispatus, suggesting overlap or intermediate states between canonical CSTs. Moreover, the enrichment of L. crispatus within CST IV-C in our dataset contrasts with the VALENCIA description of IV-C as diverse and depleted of Lactobacillus spp. and major bacterial vaginosis-associated taxa. Importantly, our findings also suggest that certain CSTs, particularly IV-B and V, may carry prognostic relevance for reproductive outcomes in the context of assisted reproduction. All women classified as CST IV-B failed to achieve pregnancy, and only a minority of women with CST V conceived. These associations should, however, be interpreted cautiously given the study limitations. While prior research has not explicitly linked CST IV-B or V to pregnancy outcomes in artificial insemination specifically, CST IV has been associated with adverse reproductive outcomes, including reduced clinical pregnancy rates and increased early pregnancy loss in IVF patients.[40] A recent IVF study evaluating community state types during embryo transfer found that individuals assigned to CST IV had the lowest pregnancy rates (∼25%) compared to Lactobacillus-dominated CSTs, whereas none of those classified as CST V became pregnant.[41]

It is now well recognized that pregnancy induces physiological and immunological shifts that impact the composition of the vaginal microbiota.[42] In our study, pregnant women displayed a distinct microbial structure compared to non-pregnant women (PERMANOVA R^2 = 0.023, q = 0.001), characterized by lower alpha diversity and the enrichment of specific taxa such as Prevotella lymphophilum, Anaerotignum massiliense, and Micrococcus radiotolerans among those with high adherence to the Mediterranean Diet (MD). These changes likely reflect hormonal influences—particularly increased estrogen levels during pregnancy—that enhance epithelial glycogen deposition and promote the growth of lactic acid-producing bacteria, thus contributing to a protective vaginal environment.[43,44] Lower vaginal microbial diversity in pregnancy, contrary to what is observed in the gut, is not necessarily indicative of dysbiosis. Rather, it is often associated with stability and health, especially when dominated by Lactobacillus species.[45]

The observed differences in microbial composition between pregnant and non-pregnant women in our cohort are particularly relevant in the context of current reproductive health challenges. The rising rates of infertility and sterility observed in the Western world are not solely due to delayed motherhood or professional lifestyle changes. Increasing evidence highlights the role of modifiable lifestyle factors—most notably nutrition—in shaping reproductive potential.[46] The Mediterranean diet is characterized by a high intake of plant-based foods, rich in fiber, antioxidants, and polyphenols, healthy fats, and a low consumption of processed and animal-based products. This dietary pattern has been associated with enhanced fertility outcomes,[47,48] including higher clinical pregnancy and live birth rates among women undergoing assisted reproductive technologies.[49,50]

Studies have shown that diets high in glycemic index carbohydrates, trans fats, and animal proteins are associated with adverse fertility outcomes, whereas plant-based proteins, dietary fiber, omega-3 fatty acids, and dairy products are linked to improved reproductive function.[51,52] Furthermore, micronutrients commonly present in MD-aligned foods—such as vitamins A, C, D, and E, beta-carotene, calcium, and zinc—appear to exert a protective role against bacterial vaginosis,[53] the most common vaginal disorder in women of reproductive age, characterized by a shift in the vaginal microbiota from a Lactobacillus-dominated community to one enriched in anaerobic bacteria.[54]

In our cohort, Gardnerella vaginalis was more prevalent among non-pregnant women with low MD adherence. This species, a hallmark of bacterial vaginosis, has been repeatedly associated with negative reproductive outcomes, including lower implantation rates and increased infertility.[55] Importantly, among pregnant participants, those who later experienced a miscarriage displayed distinct microbial signatures (PERMANOVA R^2 = 0.050, q = 0.021), with a trend toward lower alpha diversity (q = 0.092). Such patterns align with growing evidence that early pregnancy miscarriage is associated with altered vaginal microbial profiles marked by Lactobacillus depletion.[56,57] However, some studies also associate miscarriage with higher bacterial diversity.[58]

Our observation that Aerococcus mediterraneensis, Streptococcus mitis, Peptoniphilus lacrimalis, and Staphylococcus

hominis were enriched in miscarriage cases aligns with existing evidence that vaginal dysbiosis characterized by loss of *Lactobacillus* dominance is associated with early pregnancy loss.[56,57,59] The presence of these taxa has been associated with opportunistic pathogenicity, and their detection suggests a shift toward a pro-inflammatory vaginal microbiota implicated in adverse reproductive outcomes. Notably, *P. lacrimalis* exemplifies this pattern, as it has been associated with cervical dysplasia.[60,61] In contrast, *Winkia neuii*, *Bacillus mojavensis*, and a *Pseudomonota* (formerly *Proteobacteria*) member were more abundant in women with successful pregnancies. *B. mojavensis* has considerable probiotic potential with antibacterial activity against pathogenic bacteria.[62] However, *W. neuii* has been associated with persistent human papillomavirus (HPV) infection,[63] and enrichment in *Pseudomonas* genera have been observed in the vaginal microbiota of women with a short cervix and been associated with recurrent miscarriages.[64,65] These results are not surprising giving that, although these women achieved pregnancy after AI treatment, they were all diagnosed with primary infertility.

Cumulative evidence underscores the interconnected role of nutrition, microbiota composition, and fertility. Given the demonstrated benefits of the Mediterranean diet on microbial balance and reproductive outcomes, adopting healthy dietary patterns may serve as a modifiable factor to enhance the success of fertility treatments. This is particularly relevant in light of current demographic trends, including delayed motherhood and the rising prevalence of obesity and overweight. Promoting nutritional awareness among women of reproductive age could represent a key strategy to improve fertility outcomes and support reproductive health.

Limitations and further investigations

Our work faced some limitations, including the number of women included in each group studied, the lack of comparison with other reproduction methods or natural pregnancy, the lack of inclusion of confounding factors such as recent antibiotic and probiotic use, and the absence of longitudinal follow-up on the impact of diet over time before performing artificial insemination. Collecting the information about diet patterns has been possible thanks to the responses obtained through a questionnaire answered by the study participants, which is subject to the information, memory bias and psychological reactivity.

Despite the mentioned limitations, important results have been obtained that highlight the importance of a healthy diet in relation to the vaginal microbiota and fertility, demonstrating differences in alpha diversity and relative abundance of specific species.

Conclusions

The Mediterranean diet characterized by a high intake of fruits, vegetables, whole grains, vegetable oils and fish,

appears to modulate the vaginal microbiota composition. Our findings suggest that the success of artificial insemination is related to specific microbiota profiles, with pregnant women exhibiting lower alpha diversity and a reduced presence of *Gardnerella* and other anaerobic bacteria. Furthermore, distinct microbial patterns were observed between women who carried pregnancies to term and those who experienced miscarriage, underscoring the potential role of vaginal microbiota as a biomarker for reproductive outcomes.

Author contributions

Conceptualization: MC and EGV; collection of samples and questionnaires: MG; data curation and formal analysis: MB-U and MB; writing – original draft: MB-U, MG, MB; review and editing: MB and MC.

Conflicts of interest

The authors declare no competing interests.

Ethical declaration

This is a descriptive, observational and retrospective study for which an action protocol was drawn up, and approved prior to its implementation by the Ethics and Clinical Research Committee of the Doctor Peset University Hospital (CEIm code 86/21, 19-07-2021).

Data availability

The sequencing data reported in this paper are available from NCBI Sequence Read Archive (SRA) under BioProject ID PRJNA1234600.

Table S1. 14-item Questionnaire of Mediterranean diet adherence. Developed and validated in the PREDIMED study. Table S2. ASV abundance matrix of counts used for downstream analysis. Rows correspond to taxonomic features and columns to samples. Supplementary information (SI) is available. See DOI: https://doi.org/10.1039/d5fo04208a.

Acknowledgements

References

1 E. Silvestris, D. Lovero and R. Palmirotta, *Front. Endocrinol.*, 2019, **10**, 346.

2 J. A. Grieger, L. E. Grzeskowiak, T. Bianco-Miotto, T. Jankovic-Karasoulos, L. J. Moran, R. L. Wilson, S. Y. Leemaqz, L. Poston, L. McCowan, L. C. Kenny, J. Myers, J. J. Walker, R. J. Norman, G. A. Dekker and C. T. Roberts, *Hum. Reprod.*, 2018, **33**, 1063–1070.

3 K. L. Rooney and A. D. Domar, *Dialogues Clin. Neurosci.*, 2018, **20**, 41–47.

4 K. Skoracka, A. E. Ratajczak, A. M. Rychter, A. Dobrowolska and I. Krela-Kaźmierczak, *Adv. Nutr.*, 2021, **12**, 2372–2386.

5 A.-M. Foucaut, C. Faure, C. Julia, S. Czernichow, R. Levy and C. Dupont, *PLoS One*, 2019, **14**, e0210770.

6 W. Rodprasert, J. Toppari and H. E. Virtanen, *Best Pract. Res. Clin. Obstet. Gynaecol.*, 2023, **86**, 102298.

7 B. T. Mayer, S. Srinivasan, T. L. Fiedler, J. M. Marrazzo, D. N. Fredricks and J. T. Schiffer, *J. Infect. Dis.*, 2015, **212**, 793–802.

8 I. Glick, E. Kadish and M. Rottenstreich, *Int. J. Womens Health*, 2023, **15**, 1621–1622.

9 J. Zhao, B. Huang, N. Li, X. Wang, B. Xu and Y. Li, *Aging*, 2023, **15**, 2460–2472.

10 A. Šťastná, T. Fait, J. Kocourková and E. Waldaufová, *Int. J. Environ. Res. Public Health*, 2022, **20**, 668.

11 A. Bouzaglou, I. Aubenas, H. Abbou, S. Rouanet, M. Carbonnel, P. Pirtea and J. M. B. Ayoubi, *Front. Med.*, 2020, **7**, 208.

12 K. George and M. Kamath, *J. Hum. Reprod. Sci.*, 2010, **3**, 121.

13 B. Pethő, S. Váncsa, A. Váradi, G. Agócs, Á. Mátrai, F. Zászkaliczky-Iker, Z. Balogh, F. Bánhidy, P. Hegyi and N. Ács, *Am. J. Obstet. Gynecol.*, 2024, **231**, 490–500.

14 P. Njagi, W. Groot, J. Arsenijevic, S. Dyer, G. Mburu and J. Kiarie, *Hum. Reprod. Open*, 2023, (2), DOI: 10.1093/hropen/hoad007.

15 I. Alon and J. Pinilla, *Int. J. Equity Health*, 2021, **20**, 156.

16 M. Singer, M. Borg, S. Ouburg and S. A. Morré, *J. Gynecol. Obstet. Hum. Reprod.*, 2019, **48**, 223–229.

17 R. Pagar, S. Deshkar, J. Mahore, V. Patole, H. Deshpande, N. Gandham, S. Mirza, M. Junnarkar and N. Nawani, *Microbiol. Res.*, 2024, **286**, 127787.

18 K. Hou, Z.-X. Wu, X.-Y. Chen, J.-Q. Wang, D. Zhang, C. Xiao, D. Zhu, J. B. Koya, L. Wei, J. Li and Z.-S. Chen, *Signal Transduction Targeted Ther.*, 2022, **7**, 135.

19 M. Selma-Royo, F. Crispi, S. Castro-Barquero, I. Casas, M. Larroya, M. Genero, C. Paules, L. Benitez, L. Youssef, R. Pascal, N. Encabo, A. Nakaki, A. Martín-Asuero, M. T. Oller-Guzmán, A. Arranz, E. Vieta, R. Casas, R. Estruch, E. Gratacos, M. C. Collado and F. Crovetto, *Am. J. Clin. Nutr.*, 2025, **122**, 1121–1133.

20 N. Houttu, K. Mokkala, W. T. Saleem, S. Virtanen, J. Juhila, E. Koivuniemi, O. Pellonperä, K. Tertti, P. Luokola, T. Sorsa, A. Salonen, L. Lahti and K. Laitinen, *Biomed. Pharmacother.*, 2022, **149**, 112841.

21 B. Huang, J. M. Fettweis, J. P. Brooks, K. K. Jefferson and G. A. Buck, *Clin. Lab. Med.*, 2014, **34**, 747–761.

22 Q. Zeng, H. Shu, H. Pan, Y. Zhang, L. Fan, Y. Huang and L. Ling, *Front. Cell. Infect. Microbiol.*, 2024, **14**, DOI: 10.3389/fcimb.2024.1402389.

23 M. T. France, B. Ma, P. Gajer, S. Brown, M. S. Humphrys, J. B. Holm, L. E. Waetjen, R. M. Brotman and J. Ravel, *Microbiome*, 2020, **8**, 166.

24 A. Shafquat, R. Joice, S. L. Simmons and C. Huttenhower, *Trends Microbiol.*, 2014, **22**, 261–266.

25 N. Kalia, J. Singh and M. Kaur, *Ann. Clin. Microbiol. Antimicrob.*, 2020, **19**, 5.

26 S. Jena, D. Lawore, L. N. Green and D. K. Brubaker, *Biofilm*, 2025, **10**, 100330.

27 H. Verstraelen, R. Verhelst, G. Claeys, E. De Backer, M. Temmerman and M. Vaneechoutte, *BMC Microbiol.*, 2009, **9**, 116.

28 S. S. Witkin, H. Mendes-Soares, I. M. Linhares, A. Jayaram, W. J. Ledger and L. J. Forney, *mBio*, 2013, **4**, DOI: 10.1128/mBio.00460-13.

29 L. O. Eckert, D. E. Moore, D. L. Patton, K. J. Agnew and D. A. Eschenbach, *Infect. Dis. Obstet. Gynecol.*, 2003, **11**, 11–17.

30 M. A. Martínez-González, A. García-Arellano, E. Toledo, J. Salas-Salvadó, P. Buil-Cosiales, D. Corella, M. I. Covas, H. Schröder, F. Arós, E. Gómez-Gracia, M. Fiol, V. Ruiz-Gutiérrez, J. Lapetra, R. M. Lamuela-Raventos, L. Serra-Majem, X. Pintó, M. A. Muñoz, J. Wärnberg, E. Ros and R. Estruch, *PLoS One*, 2012, **7**, e43134.

31 H. Li and R. Durbin, *Bioinformatics*, 2009, **25**, 1754–1760.

32 R. Schmieder and R. Edwards, *Bioinformatics*, 2011, **27**, 863–864.

33 X. Gao, H. Lin, K. Revanna and Q. Dong, *BMC Bioinf.*, 2017, **18**, 247.

34 J. Oksanen, G. L. Simpson, F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. B. O'Hara, P. Solymos, M. H. H. Stevens, E. Szoecs, H. Wagner, M. Barbour, M. Bedward, B. Bolker, D. Borcard, G. Carvalho,

M. Chirico, M. De Caceres, S. Durand, H. B. A. Evangelista, R. FitzJohn, M. Friendly, B. Furneaux, G. Hannigan, M. O. Hill, L. Lahti, D. McGlinn, M.-H. Ouellette, E. Ribeiro Cunha, T. Smith, A. Stier, C. J. F. Ter Braak, J. Weedon and T. Borman, 2001, https://CRAN.R-project.org/package=vegan.

35 P. J. McMurdie and S. Holmes, *PLoS One*, 2013, **8**, e61217.

36 S. Mandal, W. Van Treuren, R. A. White, M. Eggesbø, R. Knight and S. D. Peddada, *Microb. Ecol. Health Dis.*, 2015, **26**, DOI: 10.3402/mehd.v26.27663.

37 H. Wickham, W. Chang, L. Henry, T. L. Pedersen, K. Takahashi, C. Wilke, K. Woo, H. Yutani, D. Dunnington and T. van den Brand, 2007, https://CRAN.R-project.org/package=ggplot2.

38 D. B. Seifer, V. L. Baker and B. Leader, *Fertil. Steril.*, 2011, **95**, 747–750.

39 H. Leridon, *Hum. Reprod.*, 2004, **19**, 1548–1553.

40 M. Maksimovic Celicanin, T. Haahr, P. Humaidan and A. Skafte-Holm, *Curr. Opin. Obstet. Gynecol.*, 2024, **36**, 155–164.

41 O. Bar, S. Vagios, O. Barkai, J. Elshirbini, I. Souter, J. Xu, K. James, C. Bormann, M. Mitsunami, J. E. Chavarro, P. Foessleitner, D. S. Kwon, M. Yassour and C. Mitchell, *NPJ Biofilms Microbiomes*, 2025, **11**, 95.

42 R. Romero, S. S. Hassan, P. Gajer, A. L. Tarca, D. W. Fadrosh, L. Nikita, M. Galuppi, R. F. Lamont, P. Chaemsaithong, J. Miranda, T. Chaiworapongsa and J. Ravel, *Microbiome*, 2014, **2**, 4.

43 D. A. MacIntyre, M. Chandiramani, Y. S. Lee, L. Kindinger, A. Smith, N. Angelopoulos, B. Lehne, S. Arulkumaran, R. Brown, T. G. Teoh, E. Holmes, J. K. Nicoholson, J. R. Marchesi and P. R. Bennett, *Sci. Rep.*, 2015, **5**, 8988.

44 S. S. Witkin, H. Mendes-Soares, I. M. Linhares, A. Jayaram, W. J. Ledger and L. J. Forney, *mBio*, 2013, **4**, DOI: 10.1128/mBio.00460-13.

45 J. Ravel, P. Gajer, Z. Abdo, G. M. Schneider, S. S. K. Koenig, S. L. McCulle, S. Karlebach, R. Gorle, J. Russell, C. O. Tacket, R. M. Brotman, C. C. Davis, K. Ault, L. Peralta and L. J. Forney, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 4680–4687.

46 N. Panth, A. Gavarkovs, M. Tamez and J. Mattei, *Front. Public Health*, 2018, **6**, DOI: 10.3389/fpubh.2018.00211.

47 L. Martín-Manchado, A. M. Moya-Yeste, M. Sánchez-Sansegundo, J. A. Hurtado-Sánchez, R. A. Gil-Miralles, J. Tuells and A. Zaragoza-Martí, *Front. Nutr.*, 2024, **11**, DOI: 10.3389/fnut.2024.1476784.

48 M. Cristodoro, E. Zambella, I. Fietta, A. Inversetti and N. Di Simone, *Biology*, 2024, **13**, 131.

49 N. J. Kellow, J. Le Cerf, F. Horta, A. L. Dordevic and C. J. Bennett, *Adv. Nutr.*, 2022, **13**, 857–874.

50 M. Vujkovic, J. H. de Vries, J. Lindemans, N. S. Macklon, P. J. van der Spek, E. A. P. Steegers and R. P. M. Steegers-Theunissen, *Fertil. Steril.*, 2010, **94**, 2096–2101.

51 F. Sofi, C. Macchi, R. Abbate, G. F. Gensini and A. Casini, *Public Health Nutr.*, 2014, **17**, 2769–2782.

52 A. J. Gaskins, F. L. Nassan, Y.-H. Chiu, M. Arvizu, P. L. Williams, M. G. Keller, I. Souter, R. Hauser and J. E. Chavarro, *Am. J. Obstet. Gynecol.*, 2019, **220**, 567.

53 Y. H. Neggers, T. R. Nansel, W. W. Andrews, J. R. Schwebke, K. Yu, R. L. Goldenberg and M. A. Klebanoff, *J. Nutr.*, 2007, **137**, 2128–2133.

54 E. H. Koumans, M. Sternberg, C. Bruce, G. McQuillan, J. Kendrick, M. Sutton and L. E. Markowitz, *Sex. Transm. Dis.*, 2007, **34**, 864–869.

55 M. A. Venneri, E. Franceschini, F. Sciarra, E. Rosato, G. D'Ettorre and A. Lenzi, *J. Endocrinol. Invest.*, 2022, **45**, 1151–1160.

56 X. Liu, Y. Cao, X. Xie, X. Qin, X. He, C. Shi, W. Zeng, Y. Guo and Y. Lin, *Comp. Immunol., Microbiol. Infect. Dis.*, 2021, **77**, 101669.

57 M. Al-Memar, S. Bobdiwala, H. Fourie, R. Mannino, Y. Lee, A. Smith, J. Marchesi, D. Timmerman, T. Bourne, P. Bennett and D. MacIntyre, *BJOG*, 2020, **127**, 264–274.

58 T. Fan, X.-M. Zhong, X.-C. Wei, Z.-L. Miao, S.-Y. Luo, H. Cheng and Q. Xiao, *Medicine*, 2020, **99**, e23558.

59 K. Grewal, Y. S. Lee, A. Smith, J. J. Brosens, T. Bourne, M. Al-Memar, S. Kundu, D. A. MacIntyre and P. R. Bennett, *BMC Med.*, 2022, **20**, 38.

60 A. Mollin, M. Katta, J. D. Sobel and R. A. Akins, *PLoS One*, 2022, **17**, e0272012.

61 J. Norenhag, G. Edfeldt, K. Stålberg, F. Garcia, L. W. Hugerth, L. Engstrand, E. Fransson, J. Du, I. Schuppe-Koistinen and M. Olovsson, *Sci. Rep.*, 2024, **14**, 11183.

62 B. Jasim, S. Sreelakshmi, J. Mathew and E. K. Radhakrishnan, *3 Biotech*, 2016, **6**, 187.

63 Z. Liu, Y. Wu, C. Wang, N. Hou and X. Liu, *Sci. Rep.*, 2025, **15**, 34996.

64 A. Silvano, N. Meriggi, S. Renzi, V. Seravalli, M. G. Torcia, D. Cavalieri and M. Di Tommaso, *Nutrients*, 2023, **15**, 2173.

65 F. Zhao, Y. Chen, J. Gao, M. Wu, C. Li, Z. Wang, N. Huang, L. Cui, M. Du and C. Ying, *Front. Cell. Infect. Microbiol.*, 2021, **11**, DOI: 10.3389/fcimb.2021.680643.