PNAS Nexus, 2024, 3, pgae419
https://doi.org/10.1093/pnasnexus/pgae419
Advance access publication 19 September 2024
Research Report

DEFB119 stratifies dysbiosis with distorted networks in the seminal microbiome associated with male infertility

Jing Jin[a,1], Howard Chi Ho Yim[b,*,1], Hsiao Mei Ellie Chang[a], Yiwei Wang[a], Kathleen Hoi Kei Choy[a], Sze Yan Chan[a], Odai A. M. Alqawasmeh[c], Jinyue Liao[a], Xiao-Tao Jiang[b,*], David Yiu Leung Chan[c,*] and Ellis Kin Lam Fok[a,d,e,*]

[a]Faculty of Medicine, School of Biomedical Sciences, The Chinese University of Hong Kong, Hong Kong SAR
[b]Faculty of Medicine and Health, Microbiome Research Centre, St George and Sutherland Campus, School of Clinical Medicine, The University of New South Wales, Sydney 2217, Australia
[c]Department of Obstetrics and Gynecology, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong SAR
[d]School of Biomedical Sciences Core Laboratory, Shenzhen Research Institute, The Chinese University of Hong Kong, Shenzhen 518057, PR China
[e]Sichuan University-The Chinese University of Hong Kong Joint Laboratory for Reproductive Medicine, West China Second University Hospital, Chengdu, PR China
*To whom correspondence should be addressed: Email: ellisfok@cuhk.edu.hk (E.K.L.F.); drdcyl16@cuhk.edu.hk (D.Y.L.C.); xiaotao.jiang@unsw.edu.au (X.-T.J.); c.yim@unsw.edu.au (H.C.H.Y.)
[1]J.J. and H.C.H.Y. contributed equally to this work.
Edited By Christopher Dupont

Abstract

Infertility is associated with the alteration of the seminal microbiome. However, the onset of dysbiosis remains controversial and the involvement of host factors remains elusive. This study investigates the alterations of the seminal microbiome in male infertility and examines the association and function of DEFB119, a reproductive-tract-specific host antimicrobial peptide, on the seminal microbiome and male fertility. While we observed comparable genera, diversity and evenness of bacterial communities, a marked decrease in the modularity of the metacommunities was observed in patients with abnormal spermiogram ($n = 57$) as compared to the control ($n = 30$). A marked elevation of DEFB119 was observed in a subpopulation of male infertile patients ($n = 5$). Elevated seminal DEFB119 was associated with a decrease in the observed genera, diversity and evenness of bacterial communities, and further distortion of the metacommunities. Mediation analysis suggests the involvement of elevated DEFB119 and dysbiosis of the seminal microbiome in mediating the abnormalities in the spermiogram. Functional experiments showed that recombinant DEFB119 significantly decrease the progressive motility of sperm in patients with abnormal spermiogram. Moreover, DEFB119 demonstrated species-specific antimicrobial activity against common seminal and nonseminal species. Our work identifies an important host factor that mediates the host–microbiome interaction and stratifies the seminal microbiome associated with male infertility. These results may lead to a new diagnostic method for male infertility and regimens for formulating the microbiome in the reproductive tract and other organ systems.

Keywords: β-defensin, bacterial networks, host–microbe interaction, microbiota, infertility

Significance Statement

Infertility is associated with the alteration of the seminal microbiome. However, the onset of dysbiosis remains controversial and the involvement of host factors remains elusive. Our study reveals the association of distorted bacterial networks in male infertility ($n = 57$) and the elevation of the DEFB119 level in seminal plasma that stratifies male infertile patients with dysbiosis of the seminal microbiome ($n = 5$). These findings uncover an important host factor that mediates the host–microbiome interaction and stratifies the seminal microbiome associated with male infertility. These results may lead to a new diagnostic method for male infertility and regimens for formulating the microbiome in the reproductive tract and other organ systems.

Introduction

The microbiomes of the reproductive tracts play pivotal roles in fecundity (1–5). Defects of which represent an important cause of infertility that affects 15% of couples worldwide (6, 7). The microbiome of the semen, herein referred to as the seminal microbiome, has been shown to be modulated by both physiological, environmental and genetic factors (2). Previous studies have shown that the major bacterial phyla in the seminal microbiome are Firmicutes, Proteobacteria, Actinobacteria, and Bacteroidetes (2). This balanced composition of the seminal microbiome is regulated by the innate immune response and secretions of various glands in the male reproductive tract (MRT). Alteration in the healthy composition of the seminal microbiome is associated with sperm abnormalities such as motility and DNA damage (8, 9).

Competing Interest: The authors declare no competing interests.
Received: February 9, 2024. **Accepted:** September 6, 2024

Apart from the involvement in regulating sperm functions and qualities, the seminal microbiome also alters the microenvironment of the female reproductive tract (FRT) after deposition. The seminal microbiome can be shared between heterosexual couples, giving complementary semino-vaginal microbiomes that affect the sperm migration along the FRT, the conception and embryo development in the oviduct, and the receptivity of the uterus for implantation (10, 11). While several phyla of the seminal microbiome have been shown to be associated with sperm abnormalities, the onset of dysbiosis in infertile cases remains controversial. Moreover, the question as to whether a host factor is involved in the host–microbes interactions remains unknown.

The β-defensin family is a group of small antimicrobial peptides mainly expressed by the epithelial cells (12). Ubiquitously expressed β-defensins such as HBD1-4 play major roles in host defense and innate immune responses (13). Intriguingly, a number of β-defensin family members are specifically expressed in the MRT. These reproductive-tract enriched/specific β-defensins play important functions in sperm functions and infertility (14, 15). Surprisingly, although β-defensins are present in semen and possess antimicrobial activity, their physiological functions and involvement in modulating the seminal microbiome have not been explored.

DEFB119 is a reproductive tract-specific β-defensin highly expressed in the testis and the epididymis (16). Single-cell sequencing analysis of testicular cells isolated from nonobstructive azoospermia men demonstrates an elevated expression of DEFB119 in the Sertoli cells (17). Male mice lacking the mouse ortholog DEFB19 demonstrated subfertility in both sexes (18, 19). DEFB119 is also expressed in sperm and is involved in sperm chemotaxis (19). However, the antimicrobial activity of DEFB119 and its potential involvement in regulating the reproductive tract microbiome has not been reported.

In view of the antimicrobial activity conserved among members of the β-defensin family, we hypothesize that reproductive-tract β-defensin serves as a host factor in maintaining the seminal microbiome. Defects of which may lead to abnormal spermiogram and contribute to male infertility. In this study, we investigate the seminal microbiome of infertile couples with normal or abnormal spermiogram parameters. We also investigate the involvement and function of host factor DEFB119 on the dysbiosis of the seminal microbiome.

Results

Distorted bacterial network in male-factor infertility

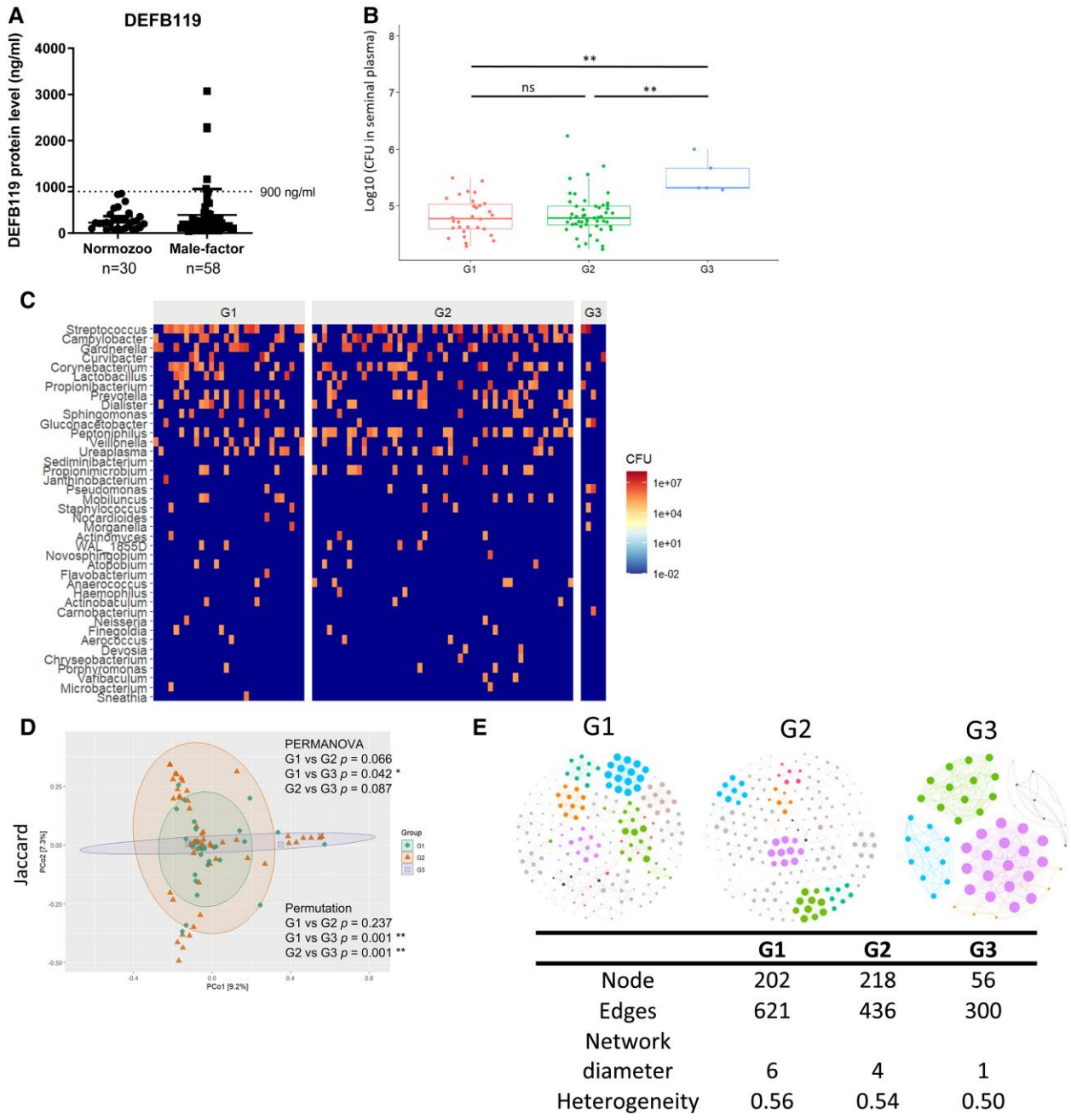
To compare the seminal microbiome in men with normal or abnormal spermiogram, we recruited 88 patients seeking assisted reproduction with either abnormal (male-factor infertility, $n = 58$) or normal spermiogram (partners of female-factor infertility or idiopathic infertility, $n = 30$) (Table S1). Patients with abnormal spermiogram include oligozoospermia, asthenozoospermia, teratozoospermia, or combined cases. The partners of 32 out of 58 patients with abnormal spermiogram were diagnosed as female-factor infertility. The age of the patient, the age of the partner, the treatment outcome, and semen volume were comparable. However, the sperm concentration, total motility, progressive motility (PR), and morphology were significantly lowered in the male-factor infertile group.

We then performed 16S rRNA gene sequencing on seminal plasma samples from this cohort of patients. After sequencing, a sequence curation pipeline optimized for analyses of amplicon libraries was performed for quality control with a low sequencing error rate (20). In total, 2,591 amplicon sequence variants (ASVs) were identified across all the seminal plasma samples. We adopted quantitative microbiome profiling (QMP) in our analysis. After removing 179 contaminants in the negative controls, 2,412 ASVs were maintained. We observed no significant difference in the bacterial load between normozoospermia and male-factor infertile patients (Fig. S1A). Consistent with previous reports (2), Firmicutes, Proteobacteria, Actinobacteria, and Bacteroidetes were the major phyla in the seminal microbiome with highest abundance among the eight phyla (Fig. S1B). We then compared the richness and evenness of the bacterial community using multiple indices. In our cohort, we observed no significant difference in α- and β-diversity indices between the normozoospermia and male-factor infertile group (Figs. S2 and S3).

To further investigate the structure and interaction of seminal microbial communities, we performed community network analysis using the Weighted Correlation Network Analysis (WGCNA) algorithm (21), which has not been applied in the seminal microbiome. In this analysis, the nodes represent ASVs and the edges that connect these nodes represent correlations between ASVs. Notably, while the number of nodes was comparable, we observed a 14% decrease in the number of edges (Fig. S1C). In microbial community networks, the ASVs clustered into independent modules, a property known as modularity, with a small group of ASVs serving as module connectors. We observed the absence of a module hub and increased abundance of within-module connectivity of Proteobacteria and Firmicutes in male-factor infertile patients. Besides, the disappearance of peripheral node in Fusobacteria was also observed in male-factor infertile patients. These results attribute to the shrinkage of network diameter and exacerbate the heterogeneity (Fig. S4). These data suggest that the bacterial network is distorted in patients with abnormal spermiogram despite the comparable richness and evenness of the metacommunities.

Elevation of DEFB119 stratifies the dysbiosis in male-factor infertility

Previous studies have utilized sperm parameters as stratification factors in the analysis of seminal microbiomes. We speculated that a host factor potentially involved in the host–microbe interactions would provide a better stratification of the seminal microbiome and a higher resolution of the dysbiosis associated with male infertility. Therefore, we explored the involvement of β-defensins in regulating the seminal microbiome, we focused on DEFB119, a β-defensin that plays pivotal roles in sperm production and functions (18, 19). We determined the protein level of DEFB119 in seminal plasma by enzyme-linked immunoassay with an antibody against the C terminus of the protein. We observed a range of DEFB119 levels in subjects with normal spermiogram (median 227.6, interquartile range [IQR] 108.0–367.2 ng/mL). Patients with abnormal spermiogram expressed comparable level of DEFB119 in the seminal plasma (median 203.1, IQR 128.6–324.6 ng/mL). Interestingly, a subgroup of patients ($n = 5$) demonstrated a marked elevation of DEFB119 in seminal plasma above the 100th centile (>900 ng/mL) of subjects with normal spermiogram (Fig. 1A and Table 1). Since there was a biased in the weighting of male infertile patients in our cohort, we have further evaluated the level of DEFB119 in an additional cohort of 21 normozoospermia patients. Notably, we observed a significant association of the elevation of DEFB119 with male infertility (χ^2, $P = 0.03$, Table S2). The elevation of DEFB119 thus represents 4.5% of the male infertile patients in our cohort.



	G1	G2	G3
Node	202	218	56
Edges	621	436	300
Network diameter	6	4	1
Heterogeneity	0.56	0.54	0.50

Fig. 1. Elevation of DEFB119 is associated with male infertility and dysbiosis of the seminal microbiome. A) Comparison of the protein level of DEFB119 in the seminal plasma collected from normozoospermia and male-factor infertile patients. A threshold of 900 ng/mL was determined to surpass the maximum level observed in normozoospermic subjects. B) Quantification of bacterial load (CFU) in patients of G1 (normozoospermia, normal DEFB119, $n = 30$), G2 (male infertility, normal DEFB119, $n = 52$), and G3 (male infertility, elevated DEFB119, $n = 5$). C) Heat map showing the comparison of the top 40 most abundant genera identified in the three groups of patients in both relative and QMP. D) Principal co-ordinate analysis plot showing the comparison of β-diversity indices Jaccard distance between G1, 2, and 3. E) WGCNA analysis showing the microbial communities networks in G1, 2, and 3. Corresponding information is shown in the bottom panel. Data represent median ± IQR.

To examine the metacommunity structure associated with elevated DEFB119, we categorized the patients according to the following grouping: G1—Normal spermiogram and low DEFB119 level ($n = 30$); G2—Abnormal spermiogram and low DEFB119 ($n = 52$); and G3—Abnormal spermiogram and elevated level of DEFB119 ($n = 5$). The spermiogram parameters were comparable in male-factor infertile patients with normal (G2) or elevated levels of DEFB119 (G3) (Table 1). Notably, a significantly higher bacterial load was observed in G3 (Fig. 1B). The dominant phylotypes of G3 were Firmicutes and Proteobacteria with Actinobacteria and Bacteroidetes present in smaller proportions (Fig. S5). Comparing the top 40 most abundant genera, the abundance of *Bactobacillus*, *Gardnerella*, and *Prevotella* decreased (Fig. 1C). We have also compared the richness and evenness of the bacterial community among the three groups using multiple indices as in our previous analysis. There was no statistically significant difference in the α-diversity in patients with elevated levels of DEFB119 as compared to G1 and G2 groups (Fig. S6). Notably, the β-diversity as measured by Jaccard distance and unweighted Unifrac distance showed that the bacterial community in patients with high levels of DEFB119 (G3) was significantly different from those in G1 and G2 groups (PERMANOVA $P < 0.05$, Figs. 1D and S7).

Table 1. Clinical characteristics of patients with normozoospermia (G1) and male-factor infertility with (G3) or without (G2) elevation of DEFB119 levels.

	G1—normozoo, normal DEFB119 ($n = 30$)	G2—male-factor, normal DEFB119 ($n = 52$)	G3—male-factor, elevated DEFB119 ($n = 5$)	*P*-value
Age, years (mean, CI)	36.63 (35.68–37.59)	35.96 (34.94–36.98)	34.20 (29.20–39.20)	0.28
Partner age, year (mean, CI)	39.93 (36.64–43.22)	39.25 (37.41–41.08)	35.80 (32.59–39.01)	0.29
Diagnosis				
Idiopathic infertility	6	0	0	—
Female-factor infertility	24	29	3	—
Male-factor infertility	0	52	5	—
Spermiogram	—	—	—	—
Semen volume, mL (median, IQR)	2.8 (2.0–3.6)	2.5 (1.7–3.0) vs. G1, 0.91	3.8 (2.8–5.9) vs. G1, 0.40 vs. G2, 0.11	0.09
Sperm concentration, millions/mL (median, IQR)	62.0 (35.1–78.1)	20.0 (10.4–50.8) vs. G1, <0.001	4.8 (2.5–24.0) vs. G1, <0.01 vs. G2, 0.31	<0.0001
Total motility, % (median, IQR)	57.0 (48.5–61.5)	48.5 (37.2–57.7) vs G1, <0.05	37.0 (36.0–53.0) vs G1, 0.07 vs. G2, 0.99	<0.01
Progressive motility, % (median, IQR)	52.0 (44.0–57.70) — —	44.0 (29.7–52.0) vs. G1, < 0.01 —	35.0 (28.53–46.5) vs. G1, 0.06 vs. G2, 0.99	<0.01 — —
Morphology, % (median, IQR)	5.0 (4.4–5.4) — —	2 (1.3–3.0) vs. G1, <0.0001 —	1.0 (1.0–2.3) vs. G1, <0.0001 vs. G2, 0.97	<0.0001 — —
ART regimen and outcome	—	—	—	
Convention IVF, *n*	24	21	1	N.A.
ICSI, *n*	5	32	4	N.A.
Fertilization rate, % (median, IQR)	75 (62.6–93.6)	73.2 (57.1–85.7)	70.0 (55.0–100.0)	0.57
Good quality embryo rate, % (median, IQR)	44.0 (6.5–80.0)	50.0 (40.0–100.0)	100.0 (29.0–100.0)	0.33
DEFB119	—	—	—	—
Protein level in seminal plasma, ng/mL (median, IQR)	227.6 (108.0–367.2) — —	203.1 (128.6–324.6) 193.3 (122.4–298.2) vs. G1, 0.99	2261.0 (1063.0–2685.0) vs. G1, <0.01 vs. G2, <0.001	0.22 <0.001 —

Data represent mean ± 95% CI or median ± IQR. *P*-value indicates analysis by Kruskal–Wallis test with Dunn's multiple comparison.
N.A., not available; ART, assisted reproductive technologies; IVF, in vitro fertilization; ICSI, intracytoplasmic sperm injection.

Differential abundance analysis at the genus level revealed a significant increases in four genera in G3 ($P < 0.05$), including *AF12*, *Enterobacter*, *Anaerobacillus*, and *Carnobacterium*, despite their rare occurrence and low relative abundance (Fig. S8). WGCNA community network analysis showed that patients with elevated levels of DEFB119 (G3) showed a marked decrease in the numbers of both nodes and edges (Fig. 1E). The network diameter and heterogeneity were further diminished in G3 as compared to G2. These data suggest that the elevated level of DEFB119 is associated with dysbiosis of the seminal microbiome and severe distortion of bacterial networks in male-factor infertile patients.

Mediators of abnormal sperm parameters in male infertility

Members of the β-defensin family are known to regulate sperm functions required for the migration in the FRT and successful fertilization (14, 15). While our data showed that the elevated level of DEFB119 was associated with the dysbiosis of the seminal microbiome in male infertility, the infertile outcome could be attributed to the effect of DEFB119 on sperm functions per se, the indirect effect from the dysbiosis of the seminal microbiome or both. To study this, we examined the correlation of DEFB119 level with various sperm parameters in normozoospermia and male-factor infertile patients. Among the sperm parameters, a significant negative correlation was observed in PR (Fig. S9, $P < 0.05$). In line with this, redundancy analysis (RDA) also revealed a negative correlation of DEFB119 with sperm motility and PR (Fig. 2A).

Next, we set out to test if the abnormal sperm parameter was mediated by the elevation of DEFB119 and the dysbiosis of the seminal microbiome. Mediation analysis revealed a notable increase in the number of significant mediations in the male-factor infertile patients but not the normozoospermia control cohort (Fig. 2B). Significant mediations were observed in five out of six tested models, regardless of the initiators and mediators. Further analysis on the proportion of mediation attributed to bacterial species revealed *Streptococcus infantis*, *Streptococcus anginosus*, and *Streptococcus agalactiae* as candidate species (Fig. 2C). These results suggest that the elevation of DEFB119 and the dysbiosis of the seminal microbiome could be the cause or the effect and vice versa. Nonetheless, the RDA and mediation analysis suggests the possible involvement of elevated DEFB119 and dysbiotic seminal microbiome in mediating the abnormal sperm parameters.

Elevated DEFB119 decreases sperm motility in male infertility

To validate the mediation analysis, we set out to examine the effect of elevated DEFB119 on sperm parameters. We mimicked the elevated level of DEFB119, as observed in G3, by recombinant DEFB119 (rDEFB119) treatment in a separate cohort of patients with normal or abnormal spermiogram profiles and normal level

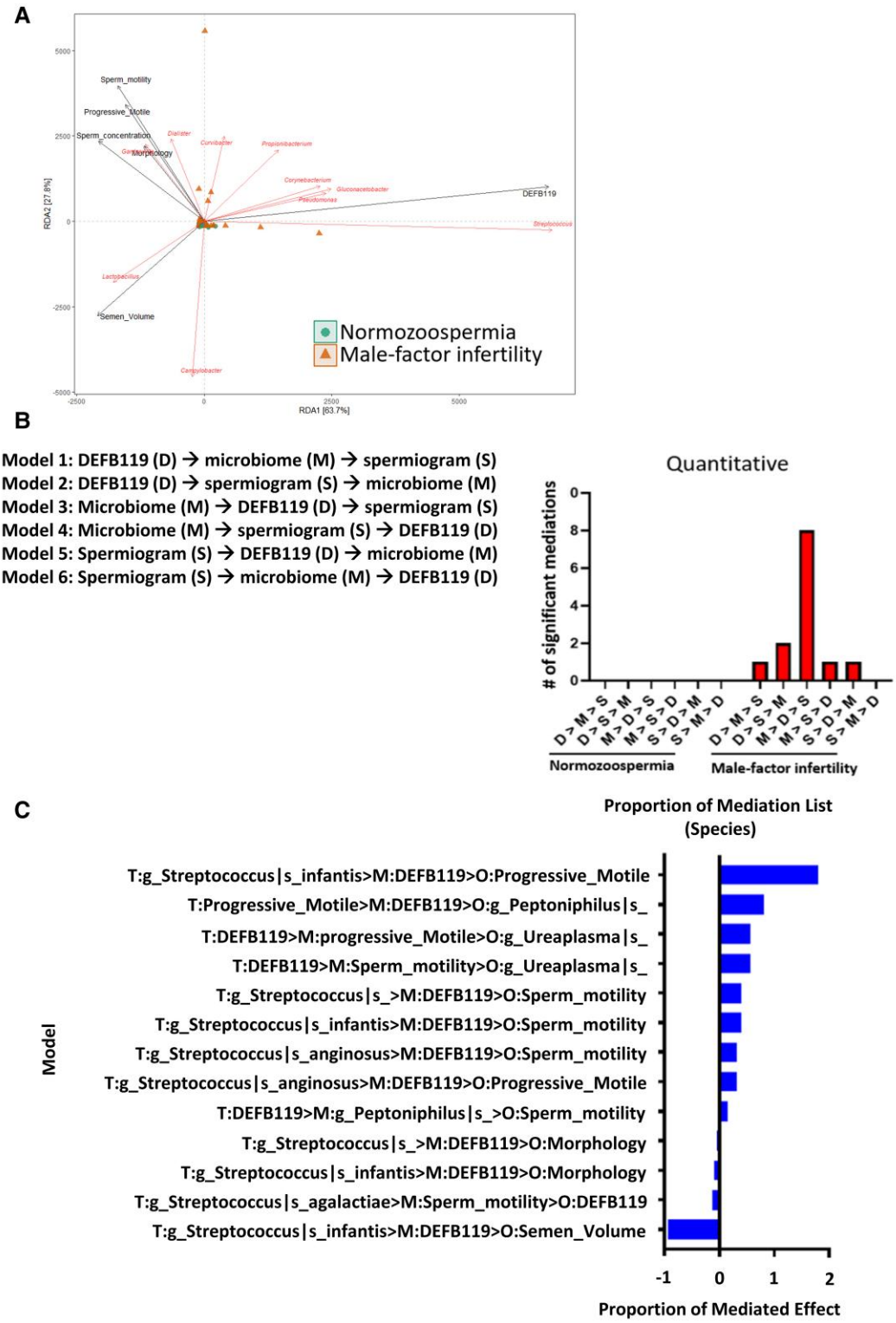


Fig. 2. Dysbiosis of the seminal microbiome and elevation of DEFB119 are potential mediators of abnormal spermiogram in male infertility. A) RDA showing the correlation of DEFB119 levels and indicated genera with semen volume and various sperm parameters, including concentration, morphology, motility and PR in QMP. B) A bar chart showing the number of significant mediations of DEFB119 levels (D) and dysbiosis of the seminal microbiome (M) in the alteration in sperm parameters (S). Two-way ANOVA, $^{*}P < 0.05$. C) A bar chart showing the proportion of mediated effect at species level in quantitative analysis.

of DEFB119 in the seminal plasma i.e. being classified as G1 or G2. Since the decrease in sperm count and morphology are mainly attributed to defects in spermatogenesis and sperm maturation that would not be altered in a short incubation period of rDEFB119 treatment theoretically and both PR and total motility demonstrated a negative correlation with DEFB119 in RDA (Fig. 2A), we examined if the elevated DEFB119 affects the motility of sperm. Our results showed that rDEFB119 significantly decreased both

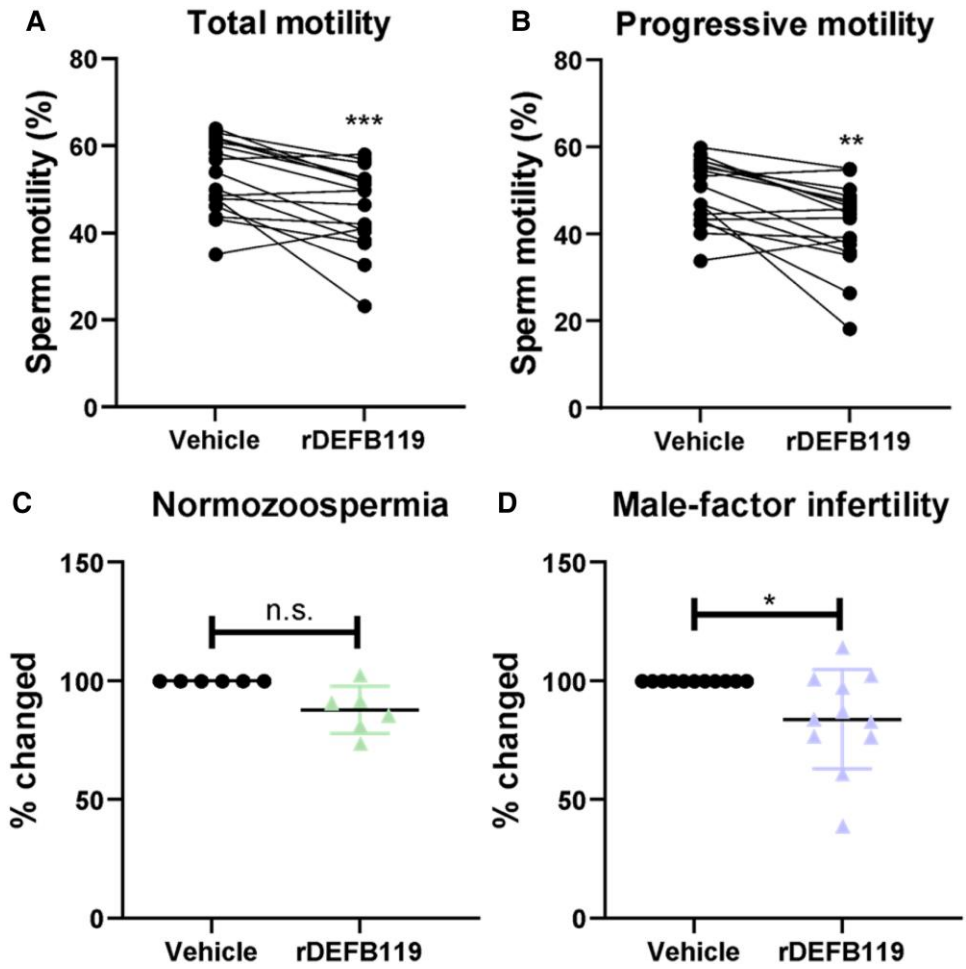


Fig. 3. Effect of elevated DEFB119 level on sperm motility. A, B) Sperm total and PR showing the effect of elevated DEFB119 level on sperm from both normozoospermia and male-factor infertile individuals ($n = 18$). C, D) Sperm PR in normozoospermic ($n = 6$) or male-factor infertile patients ($n = 12$), respectively. $^{*}P < 0.05$, $^{***}P < 0.001$.

total and PR when compared with the vehicle control (Fig. 3A and B, $P < 0.001$). Interestingly, when the sample cohort was further categorized into normozoospermic and male-factor infertile groups based on the spermiogram profile, the decrease in PR was only observed only in male-factor infertile patients (Fig. 3C and D). These results, in corroboration with the mediation analysis, suggest that elevated DEFB119 level has detrimental effects on sperm motility, and may contribute to the infertile outcome of the male-factor infertile patients with disturbed seminal microbial networks.

Species-specific antimicrobial activity of DEFB119 shapes the seminal microbiome

It is well established that the β-defensins possess antimicrobial activity that contributes to the host defense against pathogens (22–26). Of note, reproductive-tract-specific β-defensins play dual roles in host defense and sperm functions in the reproductive tract, their expressions are known to be induced by both physiological and pathological stimuli including hormone and bacterial toxin lipopolysaccharides (27, 28). Therefore, the elevated DEFB119 in seminal plasma could be a cause or an effect of the dysbiosis of the seminal microbiome. To investigate if DEFB119 plays an active role in shaping the seminal microbiome and provoking dysbiosis, we performed bactericidal assay against two dominant genera observed in G1 and G2, *Prevotella* and *Streptococcus*, and the genus found in G3, *Pseudomonas* (Fig. 4A–F). We include species that can be propagated in vitro and have either been reported in the semen, including *Prevotella copri*, *Streptococcus mitis*, *S. anginosus*, and *Pseudomonas aeruginosa* (29–32), or found in other organ systems such as *Prevotella stercorea* in the gut microbiome and *Streptococcus intermedius* in the central nervous system (33, 34). *Streptococcus anginosus* was also the species identified in our mediation analysis (Fig. 4E).

We treated the bacterial culture with various doses of rDEFB119 and observed the bacterial growth after culture. We observed a dose-dependent decrease in the amount of *P. stercorea*, *S. mitis*, *S. Anginosus*, and *S. intermedius* after rDEFB119 treatments as compared to the vehicle controls, suggesting the bactericidal effect of DEFB119 on these species (Fig. 4D–F). Intriguingly, rDEFB119 significantly promoted the growth of *S. mitis* at 3 μg/mL while higher doses demonstrate significant inhibitory effects. This result suggests that reproductive tract β-defensins can exert both promoting and inhibitory effects on specific species in a dose-dependent manner. rDEFB119 exerted a negligible effect on the growth of *Pse. aeruginosa* and *P. copri* regardless of the dosage used (Fig. 4A and C), suggesting that these bacteria were resistant to DEFB119. Taken together, our results suggest that the elevation of DEFB119 in semen plays an active role in provoking the dysbiosis of the seminal microbiome.

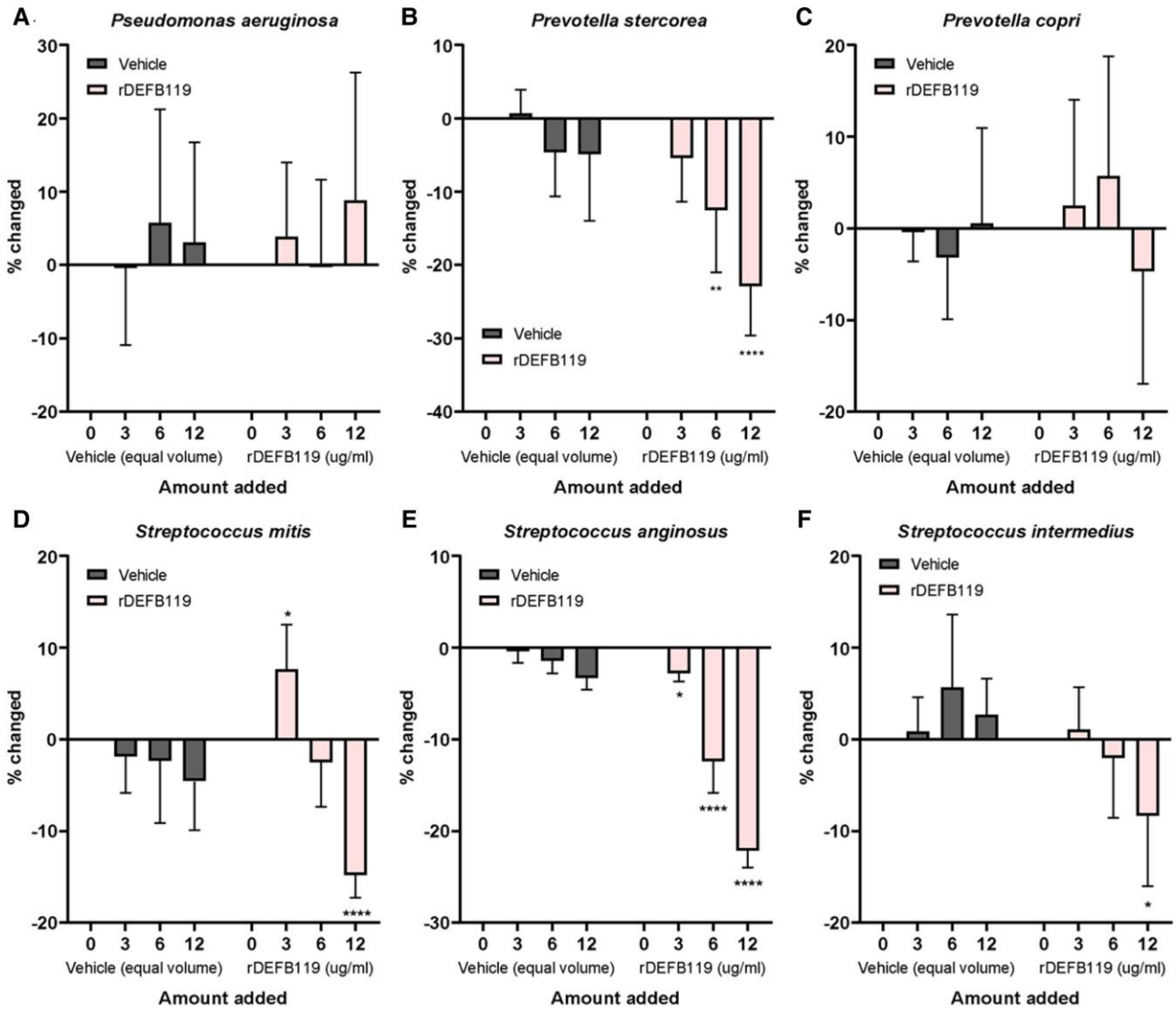


Fig. 4. Species-specific bactericidal effect of DEFB119. A–F) Bactericidal assay of representative species from *Pseudomonas* (A), *Prevotella* (B, C), and *Streptococcus* (D–F). Bacterial cultures were treated with the indicated amount of rDEFB119 or an equal amount of vehicle control. The percentage change in optical density was calculated with reference to the corresponding 0 μg/mL groups. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$.

Discussion

The present study is the first to characterize host–microbiome interaction in the seminal plasma of infertile patients via a reproductive tract-specific β-defensin. Members of the β-defensin family are involved in host defense and sperm functions, e.g. sperm motility and sperm–egg interaction, in the male and female reproductive tracts. Disorder in these β-defensins expressions, caused either by mutations or decreased expression, is associated with male infertility. Notably, our results showed that an elevated level of DEFB119 was significantly associated with male infertility. More importantly, we revealed a significant correlation between the elevated level of DEFB119 and the decrease in the abundance of genera, diversities, and community networks of the seminal microbiome in male-factor infertile patients. In contrast, the elevated level of DEFB119 was not associated with the spermiogram parameters tested. Further functional analysis revealed that elevated DEFB119 decreases the PR of sperm in male infertile patients but not normozoospermic individuals. These results suggest that both the elevated level of DEFB119 and the dysbiosis of the seminal microbiome contribute to the infertile outcome. It should be noted that the spermiogram analysis does not include a complete profile of sperm functions, particularly those to be triggered in the FRT. Furthermore, the seminal microbiome can alter the microbiome in the female tract which affects fertilization and embryo development. Therefore, it is possible that the altered seminal microbiome stemming from elevated levels of DEFB119 may lead to deregulation in the event that occurs in the FRT that contributes to the infertile outcome. Nonetheless, the present study opens an unexplored domain of research on the etiology of male infertility attributed to the interplay between β-defensin and the seminal microbiome.

Previous studies by several groups have shown that specific phylotypes of the seminal microbiome are positively or negatively associated with abnormal spermiogram i.e. male infertility. For example, the *Lactobacillus*-predominant phylotype is observed in the normozoospermic patient while *Pseudomonas*-predominant and *Prevotella*-predominant phylotype is associated with abnormal spermiogram (8, 32). In line with these findings, in the subgroup of patients with elevated levels of DEFB119, we also observed a decrease in *Lactobacillus* and an increase in *Pseudomonas*. At the species level, we observed both promoting

or inhibiting roles of DEFB119 on *Streptococcus* and *Prevotella* but not *Pseudomonas* in a dose-dependent manner, suggesting that DEFB119 provokes the dysbiosis of the seminal microbiome. To this end, it is noteworthy that although we did not identify a phylotype associated with abnormal spermiogram in our cohort of male-factor infertility patients with normal levels of DEFB119, a decrease in the diameter of the microbial community networks and heterogeneity was observed. These findings suggest that distorted communities network could be an event independent of the elevation of DEFB119. More importantly, dysbiosis was developed in patients with elevated DEFB119 (G3) but not those with normal levels of DEFB119 (G1 and 2). Therefore, our results have identified a previously unknown host factor that stratifies the dysbiosis from distorted microbial networks in male infertility. These findings suggest that the elevated level of DEFB119 may be a predictive marker of male infertility associated with dysbiosis of seminal microbiome. Notwithstanding the significant association of elevated DEFB119 with male infertility, it should be noted that the number of cases remains relatively small ($n = 5$). The small sample size in G3 may pose an over- or under-representation of genera in the seminal microbiome of this group. Further investigation of the stratification by a combination of host factors and sperm parameters in a larger cohorts of patients will provide a better resolution on the phylotype that leads to the infertile outcome.

The human microbiota, particularly those in the gut, oral cavity, and skin are known to play important roles in health and disease. The endogenously secreted antimicrobial peptides, including several members of the β-defensin family as well as other defensins, are known to be correlated with unique phylotypes in various organ systems. Our results showed that β-defensin DEFB119 specifically expressed in the reproductive tract caused a significant decrease in Actinobacteria, Bacteroidetes, Tenericutes, and Fusobacteria phyla but an increase in Proteobacteria and Firmicutes. The increased phyla demonstrate resistance to the antimicrobial activity of DEFB119. Similar resistance is also observed at the species level against *Pse. aeruginosa*. Moreover, at the species level, we observed the antimicrobial activity of DEFB119 against *P. stercorea* and *S. intermedius*, both of which have not been found in the semen. These results suggest species-specific antimicrobial activity against bacteria commonly found in the semen, as well as those in other organ systems. In view of the strong expression of a plethora of β-defensin family members in the reproductive tract and the specificity of the antimicrobial activity of individual β-defensin, we postulate that the reproductive tract-specific β-defensins represent a valuable immunocompatible way to formulate desirable microbiome phylotypes in other organ systems so as to restore normal tissue homeostasis.

Taken together, our data indicate that male infertility is associated with distorted bacterial networks. The elevation of DEFB119 in seminal plasma lower sperm motility in male infertile patients and provoke the dysbiosis of the seminal microbiome with a decrease in the number of observed genera, diversity, evenness, and community networks. Our work has provided novel insight into the host–microbiome interaction via reproductive-tract-specific antimicrobial peptides, which shed light on the etiology of male infertility and may provide a valuable tool for formulating microbiomes in other organ systems.

Materials and methods

Study design and participants

Infertile men aged 25–44 years who attended the assisted reproduction clinic at the Prince of Wales Hospital of The Chinese University of Hong Kong were recruited. All participants provided written informed consent and local ethics approval was obtained from the Institutional Review Board, the Joint Chinese University of Hong Kong—New Territories East Cluster Clinical Research Ethics Committee (approval number: CREC2016.499). All participants were in good health and none of them was under antibiotic therapy during sampling. Semen samples were collected according to a standardized protocol after 2–7 days of abstinence. Specifically, men were instructed to void, wash their hands with soap and water, and cleanse the glans penis with water before collecting the sample. Samples were collected via masturbation into a sterile container, without the use of saliva or lubrication, and immediately provided to clinical staff for processing and sample storage.

For diagnosis of male-factor infertility, semen analysis was performed according to the World Health Organization (WHO) 2010 Guidelines (35). Briefly, semen samples were allowed to liquefy for up to 30 min at 37 °C. The samples were manually evaluated by specialists in our andrology laboratory that abided by the international inter-laboratory standard following the National External Quality Assessment Scheme (NEQAS, UK). The samples were first evaluated for volume and then assessed using optical microscopy for concentration, percentage of total motility, percentage of PR, total motile sperm count, and percentage of normal morphology. Briefly, sperm concentration was counted in a hemocytometer chamber. At least 200 spermatozoa were counted per replicate. Sperm motility was evaluated as follows: PR—actively moving spermatozoa either linearly or in a large circle; nonprogressive motility (NP)—patterns of motility with an absence of progression; immotility—no movement. Total motility included both PR and NP. Sperm morphology was assessed by the Diff-Quik rapid staining method. Spermiogram parameters below the 5th centiles were diagnosed as male infertility with the following classification: Oligozoospermia—sperm concentration below 15×10^6 cells/mL; Asthenozoospermia—sperm PR below 32%; Teratozoospermia—sperm morphology below 4%. Patients with spermiogram parameters above 5th centiles were classified as normozoospermia.

For diagnosis of female-infertility, we followed the NICE Guideline (National Institute for Health and Care Excellence) for infertility determination. Briefly, the workup procedures include patient history, physical examination followed by investigation, basal hormonal assessment (FSH, LH, and E2), ovarian functions and pathology (cysts), antral follicle count, anti-mullerian hormone level, tubal patency, uterus anomaly by ultrasound (e.g. polyps and fibroids), hydrosalpinx, pelvic adhesion, anovulation, polycystic ovarian syndrome, and endometriosis. Infertile couples without a known cause of infertility were included as idiopathic infertility.

To examine the association between DEFB119 levels and male infertility, the levels of DEFB119 in seminal plasma were determined by ELISA (LS Bio, LS-F13148, see below). Infertile men were categorized according to the level of DEFB119. The seminal microbiome in the categorized patients was determined by 16S rRNA gene sequencing. The relationship between the level of DEFB119 and the phylotypes in the seminal microbiome of participants was analyzed.

DEFB119 ELISA analysis

Seminal fluid was collected from the supernatant of semen after centrifugation at 2,000 *g* for 15 min and the assay was conducted according to the manufacturer's instruction (Human DEFB119 ELISA Kit—LS-F13148). Briefly, the seminal fluid was diluted at

1:100 with sample diluent, and then the mixture was added to a coated strip plate and incubated for 2 h at 37 °C. After washing, detection reagent A was added and incubated for 1 h at 37 °C. After washing, detection reagent B was added and incubated for 1 h at 37 °C. Finally, the TMB substrate was added and incubated for 20 min at 37 °C in the dark at 37 °C. The reaction was then quenched with the stop solution and the absorbance was measured at 450 nm. The raw data were calculated using the online tool available at: https://elisaanalysis.com.

DNA extraction and library preparation

Seminal plasma DNA was extracted from the seminal fluid using the QIAamp DNA mini kit (Qiagen, Cat. No. #51306). Briefly, A total of 300 µL semen was centrifuged at 167,000 *g* for 15 min, 180 µL Buffer ATL and 40 µL proteinase K were added into the pellet and incubated at 56 °C overnight with agitation at 700 rpm. Then, Buffer AL was added to the mixture and incubated at 70 °C for 30 min. The DNA was precipitated with 100% ethanol and loaded into QIAamp Spin Column. After washing with Buffer AW1 and AW2, the DNA was eluted with a total of 40 µL of prewarmed water. DNA concentration and quality were measured by Nanodrop spectrophotometer.

We adopted a nested PCR protocol to generate amplicon products. First, the full length of the 16S rRNA gene was amplified in primary PCR (20 cycles) using forward primer 5′-AGMGTTYGATYMTGGCTCAG-3′, reverse primer 5′-TACGGYTACCTTGTTACGACTT-3′, and KAPA HiFi HotStart Readymix (Roche, Cat. No. 7958935001). The V3 and V4 hypervariable regions of the 16S rRNA were amplified by internal primers in secondary PCR (25 cycles) using forward primer 5′-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG-3′, reverse primer 5′-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC-3′, and the same Readymix reagents. The amplicons were purified by AMPure XP beads (Backman Coulter, Product No. A63880) to remove unbound primers and dimers according to the manufacturer's manual. The purified amplicons were attached to dual indices and Illumina sequencing adapters using the Nextera XT Index Kit (Illumina, Cat. No. FC-131-2001). The indexed library was validated by Qubit dsDNA HS Assay Kit (Thermo Scientific, Cat. No. Q32851) and Bioanalyzer Agilent DNA 1000 Kit (Agilent, Cat. No. 5067-1505). Three negative controls, that were empty collection tubes, were undergone the same procedures of DNA extraction, nested PCR, and library preparation.

Colony formation unit determination

The colony formation unit (CFU) of patient samples was determined by comparing the amount of DNA with that of serially diluted cultures of a nonpathogenic strain of *Escherichia coli* with determined CFUs after nested PCR, as stated in the previous section. Briefly, DH5α *E. coli* cells were inoculated in 3 mL LB broth and incubated at 37 °C with 250 rpm shaking overnight. The following day, a 10-fold serial dilution of the *E. coli* culture was carried out to 10^8-fold dilution. Then, these diluted *E. coli* were inoculated to LB agar at 37 °C overnight to produce countable plates (usually 25–300 bacterial colonies/plate). Colonies were counted and CFU/mL was calculated after colonies were formed. For qPCR, 1 mL of each diluted sample was spun down at 6,000 *g* for 5 min to collect the bacteria for DNA extraction and nested PCR as the procedure mentioned above. DNA concentrations were measured, and the DNA amount of these samples was calculated. The standard curve relating bacterial DNA amount and colony-forming unit per milliliter was plotted and regression analysis was performed. CFU of patients samples were calculated from the equation of the linear regression.

Next-generation sequencing

The library was sequenced by the Illumina HiSeq system with ~31.93 M paired-end reads. Only 80 paired-end reads were generated in one of the negative controls after sequencing and so this negative control sample was removed from the downstream analysis. The raw data were processed using an in-house pipeline (https://github.com/biofuture/microbiome_singulairty_version/; https://github.com/YimHCLab/SemenMicrobiome/; and https://figshare.com/articles/software/MRC_microbiome_16S_analysis_pipeline_singularity_version/24784080) developed by the Microbiome Research Centre of The University of New South Wales that was wrapped with Quantitative Insights into Microbial Ecology (QIIME2, v2020.8) (36). Overall sequence quality was assessed with Fastp (37), and low quality reads filtering, pair-end reads merging, and chimeric reads detection were performed with QIIME2 plugin DADA2 (38). Then, the denoise-paired function was used to generate ASVs represented sequences (rep-seqs.qza) and abundance table (table.qza). Host decontamination was conducted by mapping the human genome. The ASVs, which allow precise identification of microbes (39), were aligned to reference sequences at a cutoff of ≥ 99% in Greengenes (v13.5) (40) using QIIME2 plugin q2-feature-classifier.

For QMP, it was performed according to the published protocol (41, 42). The R scripts for decontamination, QMP, and mediation analysis were uploaded to https://github.com/YimHCLab/SemenMicrobiome/. Briefly, after running the above QIIME2 pipeline, the abundance of ASVs was corrected using the 16S copy number rrnDB database (version 5.6) using q2-gcn-norm, a QIIME2 plugin (https://forum.qiime2.org/t/q2-gcn-norm-plugin-for-normalizing-sequences-by-16s-rrna-gene-copy-number/12715). One sample did not have the CFU data and thus was removed from the full data set. Moreover, we did not yield any detectable 16S PCR products from our two negative controls, thus we had to perform the decontam step before calculating QMP. Filtering of contaminants was done using the decontam R package (https://github.com/benjjneb/decontam) using a stringent hyperparameter threshold of 0.56 and a prevalence-based contaminant identification. The samples appeared to divide into two groups: one predominantly present in positive samples and the other primarily present in negative controls (Fig. S10), suggesting the contaminant assignment effectively identified those ASVs primarily found in negative controls. One hundred and seventy-nine ASVs were identified as contaminants and were removed from the full data set. The remaining 2,412 ASVs in the full data set were used for the downstream analysis. The QMP was calculated and rarefied from these ASVs using the published R script (https://github.com/raeslab/QMP/blob/master/QMP.R). The QMP value were rounded up to integer for the downstream analysis, which was completed by the R package microeco (v0.3.2) using the absolute quantification function.

rDEFB119 treatment

rDEFB119 was commercially available (Cloud-Clone Corp, Cat. No. RP1653Hu01). The semen samples were liquefied at room temperature for 30 min, and semen analysis was performed as described above. Samples with ≥ 40% total sperm motility were used to examine the effect of rDEFB119 on sperm motility. From each sample, two aliquots of 300 µL raw semen were incubated

with vehicle control or 2 μg rDEFB119 at 37 °C for 1 h. Sperm total and PR were manually evaluated after treatment.

Bactericidal assay

Streptococcus intermedius (JTH08 strain; isolate ID: CC00038), *S. anginosus* (isolate ID: CC00108), *S. mitis* (ATCC49456 strain; isolate ID: CC00095), *Pse. aeruginosa* (SNP0614 strain; isolate ID: CC00104), *P. stercorea* (strain CB35, isolate ID: CC00801), and *P. copri* (strain JCM 13464; isolate ID: CC00987) were kindly provided by Dr Samuel Forster's group at Hudson Institute of Medical Research (43). *Streptococcus* and *Pseudomonas* were cultured on blood agar plates, and colonies were subcultured in BHI broth 24 h prior to the experiment. *Prevotella* was cultured on pre-reduced YCFA plates and subcultured in YCFA broth 48 h prior to the experiment. On the day of the experiment, bacterial culture containing *Pse. aeruginosa* ($1.2-1.4\times10^7$ CFU), *S. mitis* ($2-2.2\times10^4$ CFU), *S. anginosus* ($0.5-1.4\times10^7$ CFU), *S. intermedius* ($2.4-2.8\times10^7$ CFU), *P. stercorea* (1.2×10^5 CFU), and *P. copri* (80 CFU) were mixed with 3–12 μg/mL rDEFB119 or vehicle control in a 96-well plate with a total volume of 200 μL and were incubated under aerobic condition for 24 h for *Pseudomonas* or anaerobic conditions for 24 h for *Streptococcus*, and for 48 h for *Prevotella*. Optical density at 620 nm was measured after the incubation.

Statistical analysis

Statistical analysis of the seminal microbiome was conducted in R 4.0.4 and Rstudio https://github.com/YimHCLab/SemenMicrobiome/. A series of α-diversity analyses were calculated by t test or ANOVA among groups. The PERMANOVA test on β-diversity with 999 permutations was analyzed to compare the significant community dissimilarity among groups. Betadisper analysis was also performed to assess the homogeneity of variance among groups. Demographic characteristics across groups were compared using Mann–Whitney test (two groups) or Kruskal–Wallis test (three groups) with Dunn's multiple comparison test for continuous variables and the χ^2 for categorical variables. Mediation analyses were performed following the published protocol with minor modifications (44). Briefly, significant mediation effects were assessed between sequential pairs of factors along the microbiome (M)–DEFB119 (D)–spermiogram (S) axis in the normal and infertile groups, respectively, using R mediate package, with subject age (A) being covariates. Six models of mediation were tested. For example, mediation analysis for Model 1: DEFB119(D) → microbiome (M, i.e. each ASVs) → spermiogram (S, each spermiogram) was performed by fitting into two linear models, $M=\alpha1+\beta1D+\delta1A+\epsilon i1$ and $S=\alpha2+\beta2D+\gamma1M+\delta2A+\epsilon i2$. The adjusted *P*-values of the mediation models from the Mediate R package were obtained for each model. *P*-value <0.05 represents statistical significance.

Acknowledgments

We would like to thank Dr Samuel Forster (Hudson Institute of Medical Research, Australia) and Prof Georgina Hold (University of New South Wales, Australia) for providing the bacterial species. We would also like to thank Prof Emad El-Omar for the critical comments and suggestions on the manuscript. We would also like to thank the core facilities of the School of Biomedical Sciences for providing technical support.

Supplementary Material

Supplementary material is available at *PNAS Nexus* online.

Funding

This work was supported in part by grants from the Research Grant Council of Hong Kong (T13-602/21-N), the Health and Medical Research Fund, the Food and Health Bureau, The Government of the Hong Kong Special Administrative Region (Ref. no. 17180701), the Direct Grant of CUHK and the Lo Kwee Seong Start-Up Fund to E.K.L.F; Hong Kong Research Matching Fund (Ref. no. RMG01-8601386), the Health and Medical Research Fund, the Food and Health Bureau, The Government of the Hong Kong Special Administrative Region (Ref. no. 16173011) to D.Y.L.C.

Author Contributions

H.C.H.Y. and E.K.L.F. Conceptualization; J.J., H.M.E.C., K.H.K.C., S.Y.C., O.A.A., and D.Y.L.C. Resources, Investigation and Methodology; J.J., H.C.H.Y., H.M.E.C., Y.W., J.L., X.-T.J., and E.K.L.F. Data curation, Formal analysis; J.J., H.C.H.Y., and E.K.L.F. Writing—original draft, Writing—review & editing.

Preprints

This manuscript was posted on a preprint at: https://doi.org/10.21203/rs.3.rs-2716767/v1.

Data Availability

The 16s rRNA gene sequencing data are available in the Sequence Read Archive database with BioProject ID: PRJNA747100. Prepublication access for the reviewer is available at: https://dataview.ncbi.nlm.nih.gov/object/PRJNA747100?reviewer=vvq8rtf3tn7bisaafjtugpapse.

References

1 Lundy SD, *et al.* 2020. The microbiome of the infertile male. *Curr Opin Urol*. 30(3):355–362.
2 Altmäe S, Franasiak JM, Mändar R. 2019. The seminal microbiome in health and disease. *Nat Rev Urol*. 16(12):703–721.
3 Koedooder R, *et al.* 2019. Identification and evaluation of the microbiome in the female and male reproductive tracts. *Hum Reprod Update*. 25(3):298–325.
4 Lundy SD, *et al.* 2021. Functional and taxonomic dysbiosis of the gut, urine, and semen microbiomes in male infertility. *Eur Urol*. 79(6):826 836
5 Alqawasmeh O, *et al.* 2023. The microbiome and male infertility: looking into the past to move forward. *Hum Fertil (Camb)*. 26(3):450–462.
6 Agarwal A, Mulgund A, Hamada A, Chyatte MR. 2015. A unique view on male infertility around the globe. *Reprod Biol Endocrinol*. 13:37.
7 Osadchiy V, Mills JN, Mayer EA, Eleswarapu SV. 2020. The seminal microbiome and male factor infertility. *Curr Sex Health Rep*. 12(3):202–207.
8 Baud D, *et al.* 2019. Sperm microbiota and its impact on semen parameters. *Front Microbiol*. 10:234.
9 Farahani L, *et al.* 2021. The semen microbiome and its impact on sperm function and male fertility: a systematic review and meta-analysis. *Andrology*. 9(1):115–144.
10 Schjenken JE, Robertson SA. 2020. The female response to seminal fluid. *Physiol Rev*. 100(3):1077–1117.

11 Mändar R, *et al.* 2015. Complementary seminovaginal microbiome in couples. *Res Microbiol*. 166(5):440–447.
12 Pazgier M, Hoover DM, Yang D, Lu W, Lubkowski J. 2006. Human beta-defensins. *Cell Mol Life Sci*. 63(11):1294–1313.
13 Holly MK, Diaz K, Smith JG. 2017. Defensins in viral infection and pathogenesis. *Annu Rev Virol*. 4(1):369–391.
14 Tollner TL, *et al.* 2011. A common mutation in the defensin DEFB126 causes impaired sperm function and subfertility. *Sci Transl Med*. 3(92):92ra65.
15 Diao R, *et al.* 2014. Deficient human β-defensin 1 underlies male infertility associated with poor sperm motility and genital tract infection. *Sci Transl Med*. 6(249):249ra108.
16 Radhakrishnan Y, *et al.* 2005. Identification, characterization, and evolution of a primate beta-defensin gene cluster. *Genes Immun*. 6(3):203–210.
17 Wang M, *et al.* 2018. Single-cell RNA sequencing analysis reveals sequential cell fate transition during human spermatogenesis. *Cell Stem Cell*. 23(4):599–614.e4.
18 Jin J, *et al.* 2022. Defb19 regulates the migration of germ cell and is involved in male fertility. *Cell Biosci*. 12(1):188.
19 Li X, *et al.* 2022. β-Defensin 19/119 mediates sperm chemotaxis and is associated with idiopathic infertility. *Cell Rep Med*. 3(12): 100825.
20 Nakatsu G, *et al.* 2015. Gut mucosal microbiome across stages of colorectal carcinogenesis. *Nat Commun*. 6:8727.
21 Langfelder P, Horvath S. 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 9:559.
22 Schutte BC, McCray PB. 2002. [Beta]-defensins in lung host defense. *Annu Rev Physiol*. 64:709–748.
23 Fruitwala S, El-Naccache DW, Chang TL. 2019. Multifaceted immune functions of human defensins and underlying mechanisms. *Semin Cell Dev Biol*. 88:163–172.
24 Yang D, *et al.* 1999. Beta-defensins: linking innate and adaptive immunity through dendritic and T cell CCR6. *Science*. 286(5439): 525–528.
25 Selsted ME, Ouellette AJ. 2005. Mammalian defensins in the antimicrobial immune response. *Nat Immunol*. 6(6):551–557.
26 Li P, *et al.* 2001. An antimicrobial peptide gene found in the male reproductive system of rats. *Science*. 291(5509):1783–1785.
27 Yenugu S, Hamil KG, Radhakrishnan Y, French FS, Hall SH. 2004. The androgen-regulated epididymal sperm-binding protein, human beta-defensin 118 (DEFB118) (formerly ESC42), is an antimicrobial beta-defensin. *Endocrinology*. 145(7):3165–3173.
28 Contreras G, Shirdel I, Braun MS, Wink M. 2020. Defensins: transcriptional regulation and function beyond antimicrobial activity. *Dev Comp Immunol*. 104:103556.
29 Yang H, *et al.* 2020. Potential pathogenic bacteria in seminal microbiota of patients with different types of dysspermatism. *Sci Rep*. 10(1):6876.
30 Mores CR, *et al.* 2021. Genomic relatedness and clinical significance of *Streptococcus mitis* strains isolated from the urogenital tract of sexual partners. *Microb Genom*. 7(3):mgen000535.
31 Moretti E, *et al.* 2009. The presence of bacteria species in semen and sperm quality. *J Assist Reprod Genet*. 26(1):47–56.
32 Weng S-L, *et al.* 2014. Bacterial communities in semen from men of infertile couples: metagenomic sequencing reveals relationships of seminal microbiota to semen quality. *PLoS One*. 9(10): e110152.
33 Yeoh YK, *et al.* 2022. Prevotella species in the human gut is primarily comprised of *Prevotella copri*, *Prevotella stercorea* and related lineages. *Sci Rep*. 12(1):9055.
34 Whiley RA, Beighton D, Winstanley TG, Fraser HY, Hardie JM. 1992. *Streptococcus intermedius*, *Streptococcus constellatus*, and *Streptococcus anginosus* (the *Streptococcus milleri* group): association with different body sites and clinical infections. *J Clin Microbiol*. 30(1):243–244.
35 World Health Organization. 2010. *WHO Laboratory manual for the examination and processing of human semen*. 5th ed. Geneva: World Health Organization.
36 Bolyen E, *et al.* 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol*. 37(8):852–857.
37 Chen S, Zhou Y, Chen Y, Gu J. 2018. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 34(17):i884–i890.
38 Callahan BJ, *et al.* 2016. DADA2: high-resolution sample inference from illumina amplicon data. *Nat Methods*. 13(7):581–583.
39 Callahan BJ, McMurdie PJ, Holmes SP. 2017. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J*. 11(12):2639–2643.
40 DeSantis TZ, *et al.* 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol*. 72(7):5069–5072.
41 Liu C, Cui Y, Li X, Yao M. 2021. Microeco: an R package for data mining in microbial community ecology. *FEMS Microbiol Ecol*. 97(2):fiaa255.
42 Vandeputte D, *et al.* 2017. Quantitative microbiome profiling links gut community variation to microbial load. *Nature*. 551(7681):507–511.
43 Forster SC, *et al.* 2019. A human gut bacterial genome and culture collection for improved metagenomic analyses. *Nat Biotechnol*. 37(2):186–192.
44 Yan Z, *et al.* 2022. Multi-omics analyses of airway host-microbe interactions in chronic obstructive pulmonary disease identify potential therapeutic interventions. *Nat Microbiol*. 7(9):1361–1375.