Supplementary Methods

DNA isolation. Samples were digested in a cocktail of lysozyme (10 mg/ml, Sigma-Aldrich) and mutanolysin (25 KU/ml, Sigma-Aldrich) for 1 hr at 37 ºC. Next, 20% SDS and phenol:chloroform:isoamyl alcohol (25:24:1) were added and samples were homogenized by bead beating using FastPrep™ Lysing Matrix B. DNA was precipitated and resuspended in a final volume of 50 µl of DNA-free water (Molzym).

16S rRNA gene library generation. The V6 hypervariable region of the 16S rRNA gene was amplified using a two-step PCR strategy as previously described with slight modification to the PCR primers.[1] The first PCR step was conducted using primers that contain (from 5' to 3' ends) the Illumina paired-end sequencing adapters → 4-6 nucleotides barcode (allowing for multiplexing of 10 samples at a time) → universal V6 priming sequences V6-F 5'-CAACGCGWRGAACCTTACC-3' and V6-R 5'-CRRCACGAGCTGACGAC-3'.[1] The second PCR uses the PCRFWD1/PCRRVS1 primers that are complementary to the flow cell primer at the 5' end and Illumina paired-end sequencing adapter at the 3' end.[1]

Sequence analysis. Raw paired-end reads were merged into consensus sequences using FLASH requiring a minimum 20 bp overlap and a 5% maximum mismatch density, and subsequently filtered for quality (targeting an error rate < 0.1%) and length (minimum 60 bp) using Trimmomatic and QIIME.[2] Passing sequences were then trimmed of primers, evaluated for chimeras with UCLUST[3] (*de novo* mode), and filtered for host-associated contaminant using Bowtie2[4] searches of NCBI Homo sapiens Annotation Release 106. Additionally chloroplast and mitochondrial contaminants were detected and filtered using the RDP classifier with a confidence threshold of 50%. High-quality clean 16S sequences were then subjected to high-resolution taxonomic assessment using Resphera Insight.[5-7]

Sequence contaminant filtering. Negative controls (n=16) were evaluated for contaminating taxa associated with sample preparation. Several taxa were consistently well represented throughout the negative control set including: *Undibacterium oligocarboniphilum*, *Methylobacterium spp*, *Leuconostoc spp*. and *Propionibacterium acnes*. However, we found

evidence of low frequency cross-contamination from the urine pellet samples into the negative controls through barcode assignment errors. To compensate for this phenomenon, we designed the following filtering policy to accurately remove contaminant organisms from the dataset: Given the i^{th} OTU, O_i, which is represented with a maximum relative abundance of X_i in negative control sample set, if O_i has a relative abundance greater than $\delta * X_i$% in at least one urine sample, then it is assumed to be a real taxon; otherwise it is classified as a putative contaminant from the DNA extraction process. Assessments of δ values [1.5, 2, 3, 4, 5] suggested minimal variation in the resulting contaminant set, and a final value of 3 was selected for a final contaminant list prior to downstream analysis.

Inflammation analysis. The H&E stained slides from 8 biopsy cores (representing the right and left apex and mid prostate) from 58 of the men with benign prostate biopsies were analyzed for inflammation. Inflammation analyses were limited to the benign samples because the presence of cancer could not be blinded to the pathologist. Acute and chronic inflammation were scored using the consensus scoring system from Nickel et al.[8] For the purposes of this study, the total grade of chronic inflammation (1/mild, 2/moderate, or 3/severe) scored for each biopsy was averaged for each case and categorized as a degree of chronic inflammation: low (0 – 0.25), medium (0.5-0.75), or high (≥ 1). Acute inflammation was categorized as absent or present.

***P. acnes,* and *T. vaginalis* real-time PCR**. *P. acnes* qPCR was performed with species-specific primers Pacnes-F 5'-GCGTGAGTGACGGTAATGGGTA-3' and Pacnes-R 5'-TTCCGACGCGATCAACCA-3'. *Trichomonas vaginalis* (*T. vaginalis*) qPCR was performed with the BTUB9 and BTUB2 primer set as previously described.[9]

Statistical analysis. After contaminant removal, we performed random subsampling to 5,000 sequences per sample to provide even coverage prior to downstream statistical comparisons (rationale for subsampling described in depth in[10]). Differential abundance analysis was performed using the negative binomial test implemented in the DESeq R package. P-values were adjusted for multiple hypothesis testing using the False Discovery Rate (FDR). Beta-diversity analysis, including Bray-Curtis and UniFrac distance computation and principal coordinates analysis (PCoA), was performed in QIIME.

References

1. Arthur, J. C., Perez-Chanona, E., Mühlbauer, M. et al.: Intestinal Inflammation Targets Cancer-Inducing Activity of the Microbiota. Science, **338:** 120, 2012
2. Caporaso, J. G., Kuczynski, J., Stombaugh, J. et al.: QIIME allows analysis of high-throughput community sequencing data. Nat Methods, **7:** 335, 2010
3. Edgar, R. C.: Search and clustering orders of magnitude faster than BLAST. Bioinformatics, **26:** 2460, 2010
4. Langmead, B., Salzberg, S. L.: Fast gapped-read alignment with Bowtie 2. Nat Methods, **9:** 357, 2012
5. Abernethy, M. G., Rosenfeld, A., White, J. R. et al.: Urinary microbiome and cytokine levels in women with interstitial cystitis. Obstet Gynecol, **129:** 500, 2017
6. Ottesen, A., Ramachandran, P., Reed, E. et al.: Enrichment dynamics of Listeria monocytogenes and the associated microbiome from naturally contaminated ice cream linked to a listeriosis outbreak. BMC Microbiol, **16:** 275, 2016
7. Daquigan, N., Grim, C. J., White, J. R. et al.: Early recovery of Salmonella from food using a 6-hour non-selective pre-enrichment and reformulation of tetrathionate broth. Frontiers in Microbiology, **7:** 2103, 2016
8. Nickel, J. C., True, L. D., Krieger, J. N. et al.: Consensus development of a histopathological classification system for chronic prostatic inflammation. BJU International, **87:** 797, 2001
9. Sfanos, K. S., Sauvageot, J., Fedor, H. L. et al.: A molecular analysis of prokaryotic and viral DNA sequences in prostate tissue from patients with prostate cancer indicates the presence of multiple and diverse microorganisms. Prostate, **68:** 306, 2008
10. Weiss, S., Xu, Z. Z., Peddada, S. et al.: Normalization and microbial differential abundance strategies depend upon data characteristics. Microbiome, **5:** 27, 2017