Gut microbiota signatures associate with prostate cancer risk

Sofia Kalinen1,2, B.M., Teemu Kallonen2,3, Ph.D., Marianne Gunell2,3, Ph.D., Otto Ettala4, M.D., Ph.D., Ivan Jambor5, M.D., Ph.D., Juha Knaapila4, M.D., Ph.D., Kari T. Syvänen4, M.D., Ph.D., Pekka Taimen6,7, M.D., Ph.D., Matti Poutanen6,8,9, Ph.D., Claes Ohlsson9, M.D., Ph.D., Hannu J. Aronen5, M.D., Ph.D., Helena Ollila10, M.Sc, Sami Pietilä11, Ph.D., Laura L. Elo6,11, Ph.D., Tarja Lamminen9, Ph.D., Antti J. Hakanen2,3, M.D., Ph.D., Evelina Munukka3, Ph.D., Peter J. Boström4, M.D., Ph.D., and the Multi-IMPROD Study group*

1 Research Center for Infections and Immunity, Institute of Biomedicine, University of Turku, Turku, Finland
2 Department of Clinical Microbiology, Turku University Hospital, Turku, Finland
3 Microbiome Biobank, University of Turku, Turku, Finland
4 Department of Urology, University of Turku and Turku University Hospital, Turku, Finland
5 Department of Radiology, University of Turku and Turku University Hospital Turku, Finland
6 Institute of Biomedicine, University of Turku, Turku, Finland
7 Department of Pathology, Turku University Hospital, Turku, Finland
8 Centre for Integrative Physiology and Pharmacology, University of Turku, Turku, Finland
9 Department of Internal Medicine and Clinical Nutrition, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden
10 Turk University Clinical Research Centre, Turku University Hospital, Turku, Finland
11 Turku Bioscience Centre, University of Turku and Åbo Akademi University, Turku, Finland

*Multi-improd study group:

From the multi-improd study group, included in author list: Ivan Jambor, Otto Ettala, Juha Knaapila, Pekka Taimen, Kari T. Syvänen, Tarja Lamminen, Hannu Aronen, Peter J. Boström

Additionally the study group includes following investigators:
Janne Verho, Department of radiology, University of Turku, Turku Finland
Aida Kiviniemi, Department of radiology, University of Turku, Turku Finland
Esa Kähkönen, Department of Urology, Turku University Hospital, Turku, Finland
Iileana Montoya Perez, Department of Computing, University of Turku, Turku Finland
Marjo Seppänen, Department of Surgery, Satakunta Central Hospital, Pori, Finland
Antti Rannikko, Department of Urology, Helsinki University, and Helsinki University Hospital, Helsinki Turku, Finland
Outi Oksanen, Department of Radiology, Helsinki University Hospital, Helsinki, Finland
Jarno Riikonen, Department of Urology, Tampere University Hospital, Tampere, Finland
Sanna-Mari Vimpeli, Department of Radiology, Tampere University Hospital, Tampere, Finland
Harri Merisaari, Department of Radiology, University of Turku
Markku Kallajoki, Department of Pathology, Turku University Hospital
Tuomas Mirtt, Department of Pathology, University of Helsinki, Helsinki, Finland
Jani Saunavaara, Department of radiology, University of Turku, Turku Finland

Keywords: 16S sequencing, 5α-reductase; dihydrotestosterone, copper chaperone, Escherichia, Gut microbiota, Prevotella, Prostate cancer

Corresponding author:
Peter J. Boström MD, PhD
Department of Urology
Turku University Hospital and University of Turku, Turku, Finland
Kiinamyllynkatu 4-8
20520 Turku
Finland
Tel. +358-2-3135925
Fax. +358-2-3133243

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
ABSTRACT

Background: Prostate cancer (PC) is the most common cancer in men worldwide. The incidence of PC varies significantly geographically, which might result e.g. from genetic factors and local discrepancies in screening policies, but also from differences in lifestyle such as diet. Novel environmental factor, namely gut microbiota (GM) has been recently associated with many pathological processes including tumor progression within human body but its role in PC is disputable.

Methods: Within a clinical prospective single center trial, the GM profiles were assessed from 181 men with clinical suspicion of PC utilizing 16S rRNA gene sequencing (Illumina). Sequences were assigned to operational taxonomic units (OTUs) after which differential abundance analysis, α- and β-diversities, and predictive functional (PICRUST) analysis were performed. Further, plasma steroid hormone levels correlated to predicted microbiota functions.

Results: PC was diagnosed in 60% (108/181). Apart for less smoking among subjects with PC, there were no life-style differences between the groups. The GM profiles of men with PC differed significantly from those without cancer, e.g. Prevotella 9, members of family Erysipelotrichaceae and potentially pathogenic Escherichia-Shigella were increased, and e.g. Jonquetella, Moryella, Anaeroglobus, Corynebacterium and CAG-352 were reduced in PC cases. Predictive functional analyses revealed increased 5-α-reductase activity (5-AR), copper absorption and retinal metabolism as functional results of different microbiota. Plasma testosterone negatively correlated with predicted microbial 5-AR activity (Wilcoxon rank sum p=0.057) and in a subgroup of men taking 5-AR inhibitors (n=17), plasma estrone
(p=0.027), and estradiol (p=0.054) levels were higher in men with predicted increased microbial 5-AR function.

**Conclusions:** Certain bacterial and functional features within GM composition are associated with PC risk and altered androgen, copper and retinol metabolism are potential mechanisms. Findings could explain the previously reported association of life-style effects and geographical differences observed in PC.

**Trial registration number:** NCT02241122
INTRODUCTION

Prostate cancer (PC), despite its high incidence, has undiscovered details of etiology and pathogenesis.\textsuperscript{1,2} PC is known to be highly heritable, but also lifestyle, socioeconomic, and environmental factors are suggested to affect PC incidence. Diet is one of the most widely studied life-style factor, and various nutrients and food products have been reported to either increase or decrease PC risk.\textsuperscript{3,4,5}

PC incidence differs markedly between geographical locations, being lowest in Asia and the highest in the Western lifestyle countries.\textsuperscript{2} Differences in ethnicity, genetics, as well as health care-related factors such as intensity of prostate-specific antigen (PSA) screening have a significant effect on the reported variability of global PC incidence. However, potentially other explanatory factors might exist as well. It is of great interest that high rate of incidental PC has also been widely reported in the geographical areas of low clinical PC incidence.\textsuperscript{6} Additionally, studies conducted in immigration populations suggest that non-genetic, individual life-style factors may significantly affect PC risk.\textsuperscript{7} Based on these observations, we might expect that prostate carcinogenesis affect significant portion, if not all aging men, and individual lifestyle and environmental factors may either stimulate or inhibit the neoplastic process in the prostate, explaining the geographical differences in clinical PC. Furthermore, differences in gut microbiota could be one of the main consequences of the lifestyle related to decreased or increased PC risk.

Gut microbiota, \textit{i.e.} a collection of all microbes in the gastrointestinal tract, is acknowledged to affect many metabolic pathways and pathogenetic processes in the human body.\textsuperscript{8} Further, the state of gut dysbiosis \textit{i.e.} disequilibrium of the microbiota leading to low-grade inflammation has been linked to many cancers, even in organs distant to intestines.\textsuperscript{9} Chronic
inflammation, production of superoxide radicals, growth factors and bacterial genotoxins are proposed mechanisms. Majority of the PC studies including microbiological aspects have investigated either prostate tissue or urinary tract microbiota with conflicting results. To date, the effect of gut microbiota on prostate carcinogenesis is poorly documented though there is some evidence that it is associated with PC. Fecal microbiota and PC has been studied in few studies, most including very small sample size. The largest study to date, reported differences in overall gut microbiota composition between PC and non-PC cases as well as enrichment of Bacteroides, and Streptococcus spp. in PC group. In addition, study also suggested mechanisms including altered folate and arginine metabolisms based on microbiome analysis.

To assess the fecal microbiota profiles of PC patients compared to their benign counterparts, we conducted a substudy within a prospective clinical trial (NCT02241122) where microbiological swab samples from 181 men with suspicion of PC were 16S sequenced, and a predictive functional analysis (PICRUST) was performed. Based on our knowledge, this is the largest and most detailed clinical trial studying the gut microbiota of prostate cancer patients.

METHODS

TRIAL DESIGN

The study cohort has been previously reported in detail. Between February 2015, and March 2017, men with a clinical suspicion of PC were prospectively enrolled at four Finnish hospitals for a prospective, open-label, non-randomized trial investigating MRI and biomarkers in PC diagnosis (NCT02241122). The study was academic and investigator
initiated. All study subjects signed the informed consent. The study protocol, the patient information sheet and the informed consent were approved by the Ethics Committee of the Hospital district of Southwest Finland (No 6/2014). Ethical principles of the Declaration of Helsinki guiding physicians and medical research involving human subjects (59th World Medical Association General Assembly, Seoul, Korea, 2008) were complied.

PATIENTS

All men included had clinical suspicion of PC with PSA ranging from 2.5 to 20.0 μg/l, and/or an abnormal finding in digital rectal examination (DRE). Exclusion criteria included previous prostate biopsy, previous prostate surgery, previous diagnosis of PC, acute prostatitis, or contraindications for MRI-imaging.

ASSESSMENTS

After MRI, systematic 12-core prostate and two targeted biopsies from up to two suspected lesions in the MRI (Likert score 3-5) were collected from all subjects. Prophylactic antibiotics were given according to the institutional guidelines and have been previously reported in detail. No enema was administered prior to biopsy. Immediately prior to transrectal biopsies, microbiome samples were collected utilizing sterile rectal swabs (Copan FLOQSwab™, Copan diagnostics Inc., Murrieta, CA, USA) and immediately stored to -20°C. Besides clinical specimens, a detailed questionnaire including family history of PC, general medical, travel, and smoking history was collected prospectively prior to biopsies. Blood samples were collected for genomic DNA sequencing (whole blood) and steroid assays (plasma).
MICROBIOLOGICAL SAMPLE PREPARATION FOR 16S RNA SEQUENCING

Stool samples were diluted in 1 ml of sterile saline after which 500 µl were used in the DNA extraction with Semi-Automatic GXT Stool Extraction Kit VER 2.0 and GenoXtract unit (Hain Lifescience GmbH, Nehren, Germany) as previously described. DNA concentrations were measured with Qubit dsDNA HS Assay Kit and Qubit 2.0 fluorometer (Life Technologies, Carlsbad, USA). Extracted DNAs were stored in -80°C until used. Bacterial V4 gene regions of 16S rRNA were sequenced in three batches with MiSeq® (Illumina, San Diego, California, USA), where negative and positive controls were accompanied. This in-house method has been previously described in detail.

DATA ANALYSIS AND BIOINFORMATICS

Statistical analyses were performed with IBM Statistical Package for the Social Sciences (SPSS®) for Windows, version 26, 64-bit (IBM Corp., Armonk, NY, USA) and R version 4.0.3. Continuous variables were summarized with median and quartiles. Kolmogorov-Smirnov was used to test the data normality. PC cases were grouped by two means both utilizing the ISUP Grade, namely benign cases versus all cancer cases, and benign cases versus ISUP Grade group 1, 2-3, and 4-5 cancers, respectively. Factors potentially impacting either on cancer status or gut microbiota composition, such as age, body mass index (BMI), recent use of antibiotics, inflammatory bowel diseases (IBD), recent high-risk travel history and smoking, were also analyzed with Wilcoxon rank sum, Kruskal-Wallis, Pearson Chi square or Fisher’s exact tests.
Microbial analyses were performed with CLC Genomics Workbench Microbial Genomics module v. 12 (QIAGEN Digital Insights, Aarhus, Denmark). Sequences were assigned to operational taxonomic units (OTU) according to CLC Microbial genomics module workflow. Quality and ambiguous trims were performed with default settings and minimum number of nucleotides was set to 150. SILVA 16S v132 97% was used as the reference database. \( \alpha \)-diversity measures Chao1 and Shannon describing overall bacterial profiles were calculated to evaluate community richness, diversity and evenness with 10527 rarefaction level. \( \beta \)-diversity measure, Bray-Curtis dissimilarity, was calculated and PERMANOVA test with 99,999 permutations was used to calculate the p-values. Sequencing batch effect was controlled with \( \beta \)-diversity measures (Supplementary Figure S1).

Differential abundance analysis with generalized linear model that applies negative binomial distribution was also performed. Sequencing batch and the earlier mentioned potentially gut microbiota impacting factors were corrected in the analysis. Age was not considered in the analysis because of collinearity. The obtained results were filtered with combined abundance in all samples of more than 100 and prevalence of 10 per cent or more. P-values were corrected with false discovery rate (FDR). Statistical significance limit was set to p-value<0.05.

Predictive analysis of functional bacterial genes was constructed with utilizing PICRUS\textsuperscript{t} v.1.0 (Phylogenetic Investigation of Communities by Reconstruction of Unobserved States) tool.\textsuperscript{12} PICRUS\textsuperscript{t} was performed with Qiime v. 1.9\textsuperscript{20} to create KEGG orthologs and pathways\textsuperscript{21} with the reference data of Greengenes v.13.8.\textsuperscript{22} As microbiota and predicted gene products were not normally distributed, data was analyzed with nonparametric Wilcoxon rank sum and Kruskal-Wallis tests and described with medians and 95% confidence intervals.
PLASMA STEROIDS ASSAY

The steroid analyses of plasma samples were conducted using a LC-MS/MS according to an established method. Measured steroids included androstenedione, dehydroepiandrosterone (DHEA), dihydrotestosterone (DHT), estradiol, estrone, progesterone, 17-alpha hydroxyprogesterone (17-alpha-OHP) and testosterone (T). In addition, DHT/T ratio was calculated for the analyses. Wilcoxon rank sum tests and Pearson and Spearman correlations were used to investigate the association of predicted microbial steroid hormone biosynthesis to plasma steroid status.

RESULTS

STUDY SUBJECTS

A total of 364 men were consented for the trial. The flow-chart of the study subjects is presented in Figure 1. After excluding patients withdrawing consent (n=24), with artifacts in MRI at the time of the biopsy (n=2), without providing enough stool material in rectal swabs (n=149) and samples that failed the laboratory quality standards (n=8) (Supplementary Figure S2), a total of 181 men were included in the analyses out of which 167 complete questionnaires were available. Plasma samples were available from 169 cases of which 4 were excluded due to unknown status of 5α-reductase inhibitor (5-ARI) medication. Users and non-users of 5-ARI medication (n=17, 10%) were analyzed separately.

The basic clinicopathological characteristics and the most significant confounding factors potentially affecting microbiota composition are presented in Table 1. Age, PSA, PSA-density, and prostate volume were significantly associated with cancer status and ISUP...
cancer grade. Potentially confounding factors were similar between groups except that PC group contained less active smokers (7%) compared to benign group (13%). However, smoking was not associated with cancer grade (Table 1).

**ALPHA AND BETA DIVERSITY**

Chao1 and Shannon indices did not show statistically significant difference between benign and cancer groups (Supplementary Figure S3A-C). However, Bray-Curtis dissimilarity in PERMANOVA analysis was statistically significant between PC cases and men without PC (p=0.039, Figure 2A), but not with the different ISUP Grade groups (Supplementary Figure S4). The principal coordinate 1 (PCo 1) in Bray-Curtis appears to consist of the presence of *Prevotella 9* (10%), (Figure 2B), which was the most abundant single genera of the microbiome and associated to PC in differential abundance analysis.

**OTU CLUSTERING AND DIFFERENTIAL ABUNDANCE ANALYSIS**

The 181 successfully sequenced samples contained bacterial taxa from 21 different phyla. At the family level, OTU clustering showed differences in *Prevotellaceae* being gradually more abundant in the cancer groups (Figure 2C). Furthermore, differential abundance analysis showed that *Alloprevotella, Prevotella 2, and Prevotella 9* within family *Prevotellaceae* were more abundant in cancer group. Other significantly abundant genera belonged to families *Acidaminococcaceae, Christensenellaceae, Clostridiales vadinBB60, Corynebacteriaceae, Enterobacteriaceae, Erysipelotrichaceae, Lachnospiraceae, Muribaculaceae, Ruminococcaceae, Synergistaceae* and *Veillonellaceae*. Genera increased and reduced in PC are presented in Table 2 and Figures 2D (Supplementary Tables S1-2). Differentially abundant bacteria between IGG can be found in supplements (Supplementary Figures S5A-
D, Table S3). All the microbial taxas that were significantly differentially abundant between cancer and benign subjects, did not correlate with the different PC ISUP grades. E.g., some microbes were especially present in low grade cancers (e.g. Acidaminococcus) , while others were increased only in men with higher grade cancers (e.g. UBA1819, [Clostridium] innocuum).

PREDICTIVE ANALYSIS OF BACTERIAL FUNCTIONAL GENES (PICRUSt)
Three KEGG Pathways were significantly different between the cancer and benign groups; both copper chaperone activity (p=0.008) and steroid hormone biosynthesis (5-α-reductase enzyme, 5-AR, p=0.022) as well as retinol metabolism (p=0.042) were significantly increased in PC group compared to men without cancer (Table 3, Supplementary Table S4).

PLASMA STEROIDS
In order to investigate the possible systemic effects of the steroid hormone metabolism suggested by the KEGG pathways (PICRUSt), plasma steroid concentrations were measured and correlated to the predicted increased 5-AR activity. Firstly, as expected among 5-ARI users (n=17), DHT and DHT/T ratio was significantly lower compared to non-users (Median non-users 380 [239-486], Median users: 85 [49-203], p<0.0001). When values from men with predicted increase in 5-AR activity was compared to rest of the men without adjusting for cancer status, plasma testosterone levels negatively correlated with predicted microbial 5-AR activity (Spearman correlation -0.13774 p=0.095, Wilcoxon rank sum p=0.057). Within the subgroup of men taking 5-ARI medication (n=17), predicted microbial 5-AR was associated with increased estrone (p=0.027) and estradiol (p=0.054) levels.
DISCUSSION

This study assessed the existence of the potential link between gut microbiota and PC. The study included gut microbiota profiles of 181 patients with suspected PC. We found that various gut microbiota signatures differed between benign and cancer groups. Interestingly, predictive functional analyses (PICRUSt) suggests that differences in the steroid hormone and copper metabolism and retinol pathways between the study groups could potentially play a role in the prostate cancer development. Furthermore, altered levels in the plasma steroid concentrations were observed between men with predicted alteration in gut steroid hormone metabolism and those without, as defined by a predicted increase in microbial 5-AR activity in the former group.

According to α-diversity indices, no significant differences were found in community richness nor evenness. On the other hand, Bray-Curtis measure (β-diversity) differed significantly when PC cases were compared to men without PC but did not correlate with the cancer grade. Furthermore, various genera were differentially abundant between benign and cancer patients. Between IGG were seen potential cancer initiator bacteria and bacteria that may be fed by cancer progression, potentially supporting driver-passerger hypothesis proposed earlier in colorectal cancer (CRC).\(^{23}\)

In more detailed analyses, the abundance of family *Prevotellaceae* within phylum Bacteroidetes gradually increased among cancer cases with increasing ISUP grade.

Members of *Prevotellaceae* family are known to belong to a healthy core microbiota within gut and has been shown to be affected especially by diet since they are able to degrade complex plant polysaccharides.\(^{25}\) The role of these gram-negative species in a dysbiotic,
disease-promoting intestinal microbiota is still unclear as conflicting interpretations have been reported.\textsuperscript{25,26} *Prevotella* spp. have recently been shown to be enriched in both in colorectal tumours and mucosal microbiota of CRC patients\textsuperscript{27,28} thus implying a possible role of these species in other cancers. Interestingly, *Prevotella* 9 was the most abundant bacterial genus detected in the whole study cohort but to date the biological significance of *Prevotella* 9 in the context of microbiota and host health has not been elucidated.

Besides *Prevotellaceae*, differential abundance analysis revealed other interesting bacterial families enriched in cancer group. \textit{E.g. Escherichia-Shigella} of *Enterobacteriaceae* family, including the uropathogen *E. coli*, has cancer promoting properties and has previously been reported in prostate cancer patients.\textsuperscript{10} In addition, *Erysipelotrichaceae* has been proposed to play a role in the pathogenesis of several diseases, such as colorectal cancer (*Coprobacillus*) and opportunistic infections ([*Clostridium*] *innocuum*).\textsuperscript{29-31} Decreased abundance of *Erysipelotrichaceae* was found when PC patients receiving oral androgen receptor targeted therapies (bicalutamide, enzalutamide, abiraterone acetate) were compared to PC patients without androgen therapy.\textsuperscript{32}

It is of great interest that PICRUST analysis indicated enriched steroid hormone metabolism pathways in the gut microbiota of the cancer cases. The enzyme 5-AR converts testosterone to dihydrotestosterone (DHT). In the cancer cases the higher abundance of gut bacteria producing 5-AR may contribute to the production of dihydrotestosterone and promote prostate carcinogenesis. In line with our studies, Sfanos et al.\textsuperscript{32} discovered increased steroid hormone biosynthesis in microbiota of patients with androgen deprivation therapy (n=9) when compared to men without therapy (n=16), but our study suggests that microbiota
enriched 5-AR activity has more profound effects in prostate carcinogenesis. We also
demonstrated that the systemic steroid hormone levels correlated with altered predicted
microbial 5-AR activity. For testosterone-DHT axis the effect was not statistically significant
(Wilcoxon rank sum p=0.057), but in the subpopulation of men taking 5-AR inhibitors,
predicted increase in the 5-AR activity in the microbiota was significantly associated with
increased estrone and estradiol levels. The potential mechanistic explanation for altered
steroid hormone metabolism caused by gut microbiota clearly warrants further studies.

The other potential metabolic pathways noted in our study were copper and retinol
metabolism that are less characterized in PC. Copper chaperone is a protein that ferries
copper to cellular organelles and interestingly, accumulation of copper has been reported in
tumor cells, including PC cells. Androgen receptor (AR) activation enhances copper uptake,
again suggesting that hormonal pathways are involved in altered gut microbiota
metabolism. In line with our findings, increased retinol (vitamin-A derivate) concentrations
have been associated with increased PC risk and genetic variants in retinol pathways have
been associated with PC.

The main limitation of our study is the lack of sample size estimations and power
calculations, but this would have been impossible to carry out without prior studies.
Furthermore, low stool content in the sampling swabs, resulting in unsatisfactory amount of
bacterial DNA for NGS in 149 of 338 samples, could have been avoided by fecal sampling
method. Furthermore, our study did not include a microbial follow-up which would be
needed to investigate the actual progression of microbiota and prostate cancer. The
PICRUST analysis was performed with Greengenes database, which may differ from the
results acquired with SILVA. One should also note that the study was limited with subjects with relatively low risk cancer suspicion, which potentially may dilute the results as cases with advanced or metastatic tumors were not included. The strengths of our study relates to the prospective design, detailed prospective data collection and that this is to our knowledge the largest reported study on the subject.

CONCLUSIONS

In this study we discovered previously unreported differences of gut microbiota components between prostate cancer patients and benign subjects. In a predictive analysis steroid hormone metabolism, mineral absorption and retinol metabolism were potential carcinogenetic pathways.
REFERENCES


Figure 1. Study flowchart. 364 patients were enrolled in the study and after exclusions, 181 were included in the analyses.
**Figure 2.**

**A.** Bray-Curtis PCo plot between benign and cancer patients. Benign and cancer microbiota differs significantly in PERMANOVA analysis. **B.** *Prevotella* 9 in the Principal coordinate scatter plot PCo 1 (10% of the calculated differences). *Prevotella* 9 is also significantly increased in PC. Upper bound is Q3 (27806) of cancer cases. **C.** OTU table of the core microbiome at the family level. Family *Prevotellaceae* increases with the cancer severity. **D.** Log2 fold changes of microbial genera associated with PC. Genera increased (red) includes known pathogens as *Escherichia-Shigella* and genera that have been formerly associated with prostate cancer, such as *Erysipelotrichaceae* members (*Clostridium* *innocuum*, *Coprobacillus* and *Dielma*). Reduced genera (green) include CAG-352, *Anaerglobus* and *Corynebacterium*. 
Table 1. The clinicopathological characteristics of subjects. Age, BMI, prostate cancer markers, 5-α-reductase inhibitor medication, recent antibiotic treatments, active smoking, special diet, recent travel and inflammatory bowel diseases were analyzed.

<table>
<thead>
<tr>
<th></th>
<th>Total N</th>
<th>Total (n=181)</th>
<th>Benign (n=72)</th>
<th>ISUP Grade 1 (n=25)</th>
<th>ISUP Grade 2-3 (n=49)</th>
<th>ISUP Grade 4-5 (n=34)</th>
<th>Across the grade groups vs Benign p-value$^{28}$</th>
<th>All cancers (n=108)</th>
<th>All cancers vs Benign p-value$^{28}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median</td>
<td>Q1-Q3</td>
<td>Median</td>
<td>Q1-Q3</td>
<td>Median</td>
<td>Q1-Q3</td>
<td>Median</td>
<td>Q1-Q3</td>
<td>Median</td>
</tr>
<tr>
<td>Age (years)</td>
<td>167</td>
<td>66 50-69</td>
<td>62</td>
<td>55-67</td>
<td>63</td>
<td>58-68</td>
<td>68</td>
<td>65-73</td>
<td>69</td>
</tr>
<tr>
<td>BMI (kg/m$^2$)</td>
<td>164</td>
<td>26.7 24.8-29.0</td>
<td>26.0</td>
<td>24.1-29.5</td>
<td>26.9</td>
<td>24.7-28.9</td>
<td>26.6</td>
<td>24.1-29.2</td>
<td>27.5</td>
</tr>
<tr>
<td>PSA (μg/l)</td>
<td>181</td>
<td>7.2 5.3-9.4</td>
<td>6.1</td>
<td>4.3-8.0</td>
<td>6.0</td>
<td>4.1-9.9</td>
<td>8.4</td>
<td>7.0-10.0</td>
<td>8.0</td>
</tr>
<tr>
<td>PSA-density (ng/ml$^2$)</td>
<td>175</td>
<td>0.17 0.11-0.27</td>
<td>0.12</td>
<td>0.10-0.20</td>
<td>0.14</td>
<td>0.10-0.26</td>
<td>0.24</td>
<td>0.16-0.32</td>
<td>0.23</td>
</tr>
<tr>
<td>free-PSA ratio (%)</td>
<td>155</td>
<td>14 11-19</td>
<td>16</td>
<td>11-20</td>
<td>16</td>
<td>11-21</td>
<td>14</td>
<td>11-19</td>
<td>12</td>
</tr>
<tr>
<td>Prostate volume (ml)</td>
<td>175</td>
<td>40 30-55</td>
<td>44.00</td>
<td>33-62</td>
<td>39</td>
<td>31-49</td>
<td>35</td>
<td>28-52</td>
<td>36</td>
</tr>
<tr>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>5-α-reductase medication$^a$</td>
<td>167</td>
<td>11</td>
<td>12</td>
<td>8</td>
<td>4</td>
<td>18</td>
<td>0.18$^3$</td>
<td>9</td>
<td>0.46$^2$</td>
</tr>
<tr>
<td>Antibiotic treatments within a year</td>
<td>165</td>
<td>30</td>
<td>32</td>
<td>30</td>
<td>20</td>
<td>39</td>
<td>0.30$^2$</td>
<td>28</td>
<td>0.55$^1$</td>
</tr>
<tr>
<td>Active smoking</td>
<td>165</td>
<td>13</td>
<td>20</td>
<td>13</td>
<td>9</td>
<td>7</td>
<td>0.20$^1$</td>
<td>9</td>
<td>0.04$^2$</td>
</tr>
<tr>
<td>Special diet</td>
<td>165</td>
<td>2</td>
<td>2</td>
<td>8</td>
<td>2</td>
<td>0</td>
<td>0.26$^1$</td>
<td>3</td>
<td>1.00$^5$</td>
</tr>
<tr>
<td>Traveling to far off countries within a year$^b$</td>
<td>167</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>9</td>
<td>3</td>
<td>0.83$^5$</td>
<td>6</td>
<td>1.00$^5$</td>
</tr>
<tr>
<td>Inflammatory bowel disease</td>
<td>167</td>
<td>2</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.34$^5$</td>
<td>0</td>
<td>0.06$^5$</td>
</tr>
</tbody>
</table>

Kruskal wallis test

$^a$ Wilcoxon Rank sum test

$^b$ Pearson Chi square

$^c$ Fisher's exact test

$^d$ finasteride/dutasteride

$^e$ Asia, South America and Africa
Table 2 Differential abundance analysis. Statistically significant (FDR-p<0.05) altered microbial genera and their taxonomy, combined abundance, prevalence in samples, median abundance in present samples taxa and fold changes.

<table>
<thead>
<tr>
<th>Family</th>
<th>Genus</th>
<th>Combined Abundance</th>
<th>Prevalence in samples (%)</th>
<th>Median of combined abundance in present samples</th>
<th>Max group mean</th>
<th>Log, fold change</th>
<th>Fold change</th>
<th>P-value</th>
<th>FDR P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methanosarcinaceae</td>
<td>Methanosphaera</td>
<td>504</td>
<td>13</td>
<td>6</td>
<td>0.1</td>
<td>2.5</td>
<td>5.5</td>
<td>&lt;0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>Corynebacteriaceae</td>
<td>Corynebacterium</td>
<td>1359</td>
<td>35</td>
<td>4</td>
<td>13</td>
<td>-2.9</td>
<td>-7.5</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Barnesiellaceae</td>
<td>uncultured-08</td>
<td>2685</td>
<td>17</td>
<td>39</td>
<td>20</td>
<td>2.1</td>
<td>4.2</td>
<td>0.001</td>
<td>0.009</td>
</tr>
<tr>
<td>Marinilaceae</td>
<td>Sanguibacteroides</td>
<td>1124</td>
<td>12</td>
<td>25</td>
<td>6</td>
<td>2.1</td>
<td>4.2</td>
<td>0.001</td>
<td>0.009</td>
</tr>
<tr>
<td>Muribaculaceae</td>
<td>CAG-873</td>
<td>92781</td>
<td>24</td>
<td>4</td>
<td>678</td>
<td>2.3</td>
<td>4.8</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>uncultured organ-02</td>
<td>976</td>
<td>13</td>
<td>1</td>
<td>6</td>
<td>2.9</td>
<td>7.4</td>
<td>&lt;0.001</td>
<td>0.002</td>
</tr>
<tr>
<td>Prevotellaceae</td>
<td>Alfiprevotella</td>
<td>573293</td>
<td>50</td>
<td>8</td>
<td>4625</td>
<td>5.5</td>
<td>-45.6</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Prevotella 2</td>
<td>15028</td>
<td>13</td>
<td>3</td>
<td>136</td>
<td>2.2</td>
<td>4.4</td>
<td>0.001</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>Prevotella 9</td>
<td>4410629</td>
<td>100</td>
<td>244</td>
<td>3068</td>
<td>1.5</td>
<td>2.8</td>
<td>0.003</td>
<td>0.030</td>
</tr>
<tr>
<td>Lentimicrobiaceae</td>
<td>uncultured bacterium-09</td>
<td>11402</td>
<td>24</td>
<td>63</td>
<td>88</td>
<td>2.2</td>
<td>4.6</td>
<td>0.001</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>Ambiguous_taxa-07</td>
<td>4671</td>
<td>16</td>
<td>12</td>
<td>34</td>
<td>2.5</td>
<td>5.8</td>
<td>&lt;0.001</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>gut metagenome-02</td>
<td>9133</td>
<td>11</td>
<td>20</td>
<td>71</td>
<td>-2.4</td>
<td>-5.3</td>
<td>&lt;0.001</td>
<td>0.008</td>
</tr>
<tr>
<td>Streptococcaceae</td>
<td>Lactococcus</td>
<td>2422</td>
<td>60</td>
<td>5</td>
<td>17</td>
<td>1.3</td>
<td>2.5</td>
<td>0.001</td>
<td>0.020</td>
</tr>
<tr>
<td>Clostridiobacteriaceae</td>
<td>Clostridium</td>
<td>406</td>
<td>23</td>
<td>1</td>
<td>17</td>
<td>1.7</td>
<td>3.2</td>
<td>0.001</td>
<td>0.010</td>
</tr>
<tr>
<td>Christensenellaceae</td>
<td>Christensenellaes R-7 group</td>
<td>171030</td>
<td>98</td>
<td>175</td>
<td>1274</td>
<td>1.0</td>
<td>2.1</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>Clostridiales vadB660 group</td>
<td>Ambiguous_taxa-13</td>
<td>33714</td>
<td>80</td>
<td>20</td>
<td>295</td>
<td>1.6</td>
<td>3.1</td>
<td>&lt;0.001</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>uncultured organism-04</td>
<td>518</td>
<td>27</td>
<td>7</td>
<td>4</td>
<td>1.6</td>
<td>3.0</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>uncultured Thermoanaerobacteriales bacterium</td>
<td>2338</td>
<td>43</td>
<td>10</td>
<td>18</td>
<td>1.7</td>
<td>3.2</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>Lachnospiraceae</td>
<td>Butyvibrio</td>
<td>204311</td>
<td>50</td>
<td>4</td>
<td>1584</td>
<td>2.1</td>
<td>4.3</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>Lachnospiraceae FCS020 group</td>
<td>7020</td>
<td>84</td>
<td>17</td>
<td>58</td>
<td>1.0</td>
<td>2.0</td>
<td>0.003</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>Lachnospiraceae UCG-003</td>
<td>16360</td>
<td>31</td>
<td>21</td>
<td>106</td>
<td>2.4</td>
<td>5.4</td>
<td>&lt;0.001</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>Moryella</td>
<td>6991</td>
<td>85</td>
<td>13</td>
<td>61</td>
<td>-1.5</td>
<td>-2.7</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Tyzzerella 4</td>
<td>27950</td>
<td>28</td>
<td>8</td>
<td>237</td>
<td>1.8</td>
<td>3.5</td>
<td>0.004</td>
<td>0.030</td>
</tr>
<tr>
<td>Ruminococcaceae</td>
<td>CAG-352</td>
<td>14571</td>
<td>19</td>
<td>7</td>
<td>125</td>
<td>-3.3</td>
<td>-9.8</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Hydrowedinanaerobacterium</td>
<td>2547</td>
<td>43</td>
<td>7</td>
<td>22</td>
<td>2.0</td>
<td>4.0</td>
<td>&lt;0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>Oscillibacteriaceae</td>
<td>1155</td>
<td>22</td>
<td>4</td>
<td>7</td>
<td>1.5</td>
<td>2.8</td>
<td>0.006</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>UBA1819</td>
<td>27091</td>
<td>92</td>
<td>18</td>
<td>193</td>
<td>1.1</td>
<td>2.2</td>
<td>0.005</td>
<td>0.040</td>
</tr>
<tr>
<td>Erysipelotrichaceae</td>
<td>[Clostridium] limosum group</td>
<td>466</td>
<td>23</td>
<td>3</td>
<td>4</td>
<td>2.1</td>
<td>4.2</td>
<td>&lt;0.001</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>Coprobacillus</td>
<td>1284</td>
<td>17</td>
<td>6</td>
<td>10</td>
<td>4.4</td>
<td>20.7</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Dietera</td>
<td>1167</td>
<td>21</td>
<td>3</td>
<td>1</td>
<td>1.5</td>
<td>2.9</td>
<td>0.001</td>
<td>0.010</td>
</tr>
<tr>
<td>Acidaminococcaceae</td>
<td>Acidaminococcus</td>
<td>112181</td>
<td>43</td>
<td>7</td>
<td>903</td>
<td>2.5</td>
<td>5.6</td>
<td>&lt;0.001</td>
<td>0.005</td>
</tr>
<tr>
<td>Veillonellaceae</td>
<td>Anaeroglobin</td>
<td>12425</td>
<td>44</td>
<td>7</td>
<td>141</td>
<td>-3.0</td>
<td>-7.8</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>vadB667</td>
<td>uncultured rumen bacterium-04</td>
<td>528</td>
<td>22</td>
<td>8</td>
<td>4</td>
<td>1.5</td>
<td>2.8</td>
<td>0.004</td>
<td>0.040</td>
</tr>
<tr>
<td>Burkholderiaceae</td>
<td>Alcaligenes</td>
<td>700</td>
<td>54</td>
<td>3</td>
<td>4</td>
<td>-1.2</td>
<td>-2.2</td>
<td>0.003</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>Paracuttermella</td>
<td>77911</td>
<td>90</td>
<td>44</td>
<td>495</td>
<td>-1.3</td>
<td>-2.5</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>Nissleriaceae</td>
<td>uncultured-29</td>
<td>1206</td>
<td>12</td>
<td>17</td>
<td>10</td>
<td>2.5</td>
<td>5.5</td>
<td>&lt;0.001</td>
<td>0.002</td>
</tr>
<tr>
<td>Enterobacteriaceae</td>
<td>Ambiguous_taxa-29</td>
<td>139</td>
<td>27</td>
<td>2</td>
<td>1</td>
<td>1.3</td>
<td>2.4</td>
<td>0.003</td>
<td>0.030</td>
</tr>
<tr>
<td></td>
<td>Citrobacter</td>
<td>533</td>
<td>13</td>
<td>7</td>
<td>4</td>
<td>1.7</td>
<td>3.4</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>Escherichia-Shigella</td>
<td>551481</td>
<td>98</td>
<td>229</td>
<td>3314</td>
<td>1.3</td>
<td>2.5</td>
<td>0.003</td>
<td>0.003</td>
</tr>
<tr>
<td>Synergistaceae</td>
<td>Jonquelia</td>
<td>27709</td>
<td>40</td>
<td>14</td>
<td>203</td>
<td>-2.4</td>
<td>-5.2</td>
<td>&lt;0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>Puniceococcaceae</td>
<td>uncultured-32</td>
<td>4177</td>
<td>77</td>
<td>10</td>
<td>35</td>
<td>2.4</td>
<td>5.3</td>
<td>&lt;0.001</td>
<td>0.002</td>
</tr>
</tbody>
</table>
**Table 3.** Statistically significant (without FDR-correction) predicted bacterial gene products potentially associated with PC.

<table>
<thead>
<tr>
<th>KEGG Pathway</th>
<th>KEGG Orthologs</th>
<th>Benign</th>
<th>Cancer</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Median</td>
<td>95% CI [LB; UB]</td>
<td>Median</td>
</tr>
<tr>
<td>Mineral absorption</td>
<td>Copper chaperone (95%) Heme oxygenase 1 (5%)</td>
<td>105</td>
<td>[68; 155]</td>
<td>168</td>
</tr>
<tr>
<td>Steroid hormone biosynthesis</td>
<td>5-α-reductase 1 (45%) 5-α-reductase 2 (54%) 3-alpha hydroxysteroid hydrogenase (1%)</td>
<td>352</td>
<td>[188; 10062]</td>
<td>11717</td>
</tr>
<tr>
<td>Retinol metabolism</td>
<td>Alcohol dehydrogenase (67%) Alcohol dehydrogenase propanol preferring (5%) All-trans-retinol 13,14-reductase (20%)</td>
<td>21605</td>
<td>[18660; 23784]</td>
<td>23174</td>
</tr>
</tbody>
</table>