







Article

Microbiome Markers of Pancreatic Cancer Based on Bacteria-Derived Extracellular Vesicles Acquired from Blood Samples: A Retrospective Propensity Score Matching Analysis

Jae Ri Kim ^{1,2,†} , Kyulhee Han ^{3,†} , Youngmin Han ¹ , Nayeon Kang ³, Tae-Seop Shin ⁴, Hyeon Ju Park ⁴, Hongbeom Kim ¹, Wooil Kwon ¹, Seungyeoun Lee ⁵ , Yoon-Keun Kim ⁴, Taesung Park ^{3,6,*}  and Jin-Young Jang ^{1,*} 

¹ Department of Surgery, Seoul National University Hospital, Seoul 03080, Korea; jaripo@gmail.com (J.R.K.); vickijoa@naver.com (Y.H.); surgeonkh@gmail.com (H.K.); willdoc78@gmail.com (W.K.)

² Department of Surgery, Gyeongsang National University Changwon Hospital, Changwon 51472, Korea

³ Interdisciplinary Program in Bioinformatics, Seoul National University, Seoul 08826, Korea; hgh97617@gmail.com (K.H.); thisisnykang@gmail.com (N.K.)

⁴ MD Healthcare Inc., Seoul 03923, Korea; tsshin@mdhc.kr (T.-S.S.); hjpark@mdhc.kr (H.J.P.); ykkim@mdhc.kr (Y.-K.K.)

⁵ Department of Mathematics and Statistics, Sejong University, Seoul 05006, Korea; leesy@sejong.ac.kr

⁶ Department of Statistics, Seoul National University, Seoul 08826, Korea

* Correspondence: tspark@stats.snu.ac.kr (T.P.); jangjy4@gmail.com (J.-Y.J.); Tel.: +82-2-880-8924 (T.P.); Fax: +82-2-880-6144 (T.P.); Tel./Fax: +82-2-2072-2194 (J.-Y.J.)

† J.R.K. and K.H. contributed equally as first authors.

‡ T.P. and J.-Y.J. contributed equally to the writing of this manuscript.



Citation: Kim, J.R.; Han, K.; Han, Y.; Kang, N.; Shin, T.-S.; Park, H.J.; Kim, H.; Kwon, W.; Lee, S.; Kim, Y.-K.; et al. Microbiome Markers of Pancreatic Cancer Based on Bacteria-Derived Extracellular Vesicles Acquired from Blood Samples: A Retrospective Propensity Score Matching Analysis. *Biology* **2021**, *10*, 219. <https://doi.org/10.3390/biology10030219>

Academic Editor: Arnold Spiek

Received: 18 February 2021

Accepted: 8 March 2021

Published: 13 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Simple Summary: Although tremendous advances in diagnosis and treatment, pancreatic cancer still remains one of the lethal diseases with an overall survival rate of 10~15%. Early detection and diagnosis of pancreatic cancer is very important in improving the prognosis of patients. The aim of our study was to find new biomarkers, using microbiomes based on bacteria-derived extracellular vesicles, extracted from blood serum. With 38 patients with pancreatic cancer and 52 healthy controls with no history of pancreatic disease, we identified several compositional differences of microbiome between them. Using various combinations of the metagenomic markers which made the compositional differences, we also built a pancreatic cancer prediction model with high area under the receiver operating characteristic curve (0.966 at the phylum level and 1.000 at the genus level). These microbiome markers, based on bacteria-derived extracellular vesicles acquired from blood, show demonstrate the potential of candidate biomarkers for early diagnosis of pancreatic cancer.

Abstract: Novel biomarkers for early diagnosis of pancreatic cancer (PC) are necessary to improve prognosis. We aimed to discover candidate biomarkers by identifying compositional differences of microbiome between patients with PC ($n = 38$) and healthy controls ($n = 52$), using microbial extracellular vesicles (EVs) acquired from blood samples. Composition analysis was performed using 16S rRNA gene analysis and bacteria-derived EVs. Statistically significant differences in microbial compositions were used to construct PC prediction models after propensity score matching analysis to reduce other possible biases. Between-group differences in microbial compositions were identified at the phylum and genus levels. At the phylum level, three species (*Verrucomicrobia*, *Deferrribacteres*, and *Bacteroidetes*) were more abundant and one species (*Actinobacteria*) was less abundant in PC patients. At the genus level, four species (*Stenotrophomonas*, *Sphingomonas*, *Propionibacterium*, and *Corynebacterium*) were less abundant and six species (*Ruminococcaceae* UCG-014, *Lachnospiraceae* NK4A136 group, *Akkermansia*, *Turicibacter*, *Ruminiclostridium*, and *Lachnospiraceae* UCG-001) were more abundant in PC patients. Using the best combination of these microbiome markers, we constructed a PC prediction model that yielded a high area under the receiver operating characteristic curve (0.966 and 1.000, at the phylum and genus level, respectively). These microbiome markers, which altered microbial compositions, are therefore candidate biomarkers for early diagnosis of PC.

Keywords: pancreatic cancer; microbial extracellular vesicles; microbiome markers; early diagnosis; propensity score matching

1. Introduction

Pancreatic cancer (PC) is the leading cause of cancer-associated mortality worldwide [1]. Despite advances in diagnosis, surgical technique, perioperative care, and chemoradiotherapy regimens, the 5-year survival rate for PC patients remains 10–15% due to difficulties with early detection. Only 20% of PC patients are considered to have “early-stage” PC at the time of diagnosis, when curative-intent surgery is possible [2]. Research seeking to discover novel biomarkers for early diagnosis is on-going.

To date, several studies have reported that dysbiosis of the human body microenvironment can increase the risk of inflammation or malignant tumor development [3]. The gut microbiota has been reported to affect the risk of adverse health outcomes, including that of liver cirrhosis/hepatocellular carcinoma [4,5], cardiovascular disease [6], breast cancer [7], and colorectal cancer [8]. Moreover, previous studies have suggested an association between PC and the presence of gut bacteria, such as *Helicobacter pylori*, and salivary microbiota have been shown to be associated with chronic pancreatitis or PC [9–11]. However, alterations of blood microbiota in patients with PC remain unknown.

Microbiota produce large quantities of bacteria-derived extracellular vesicles (EVs), which allow interaction with human cells [12]. EVs contain bacteria-derived genomes and are detectable in human blood, urine, bile, and stool samples. They can help evaluate microbiota composition [13–15]. A recent study, aimed to develop a diagnostic model for ovarian cancer, used microbiome profiles from serum bacteria-derived EVs alongside clinicopathologic data [16]. However, to the best of our knowledge, the association between the blood microbiome, assessed with bacteria-derived EVs, and PC has not been previously studied. This is the first study to identify candidate biomarkers for the diagnosis of early-stage PC, by comparing the differences in microbiome compositions using blood microbiota between patients with PC and healthy controls.

2. Results

After quality control of samples, 82 patients with PC and 116 healthy controls were identified as eligible for this study. However, the age and sex of the two groups were different enough to discriminate PC patients from controls with a high area under the ROC curve (AUC) of 0.882. After propensity score matching (PSM), only 38 patients with PC (17 men and 21 women) and 52 controls (17 men and 35 women) were selected. As a result, the covariates could no longer contribute to predicting PC (AUC of 0.571, Appendix A Figure A1).

The age and sex of the patients before and after PSM are described in Table 1A. Table 1B shows the clinicopathologic characteristics of the PC group after PSM. Most PC patients were pathologically confirmed as having advanced stage cancer (68.4% had stage II and 29.0% had stage III/IV, as defined by the seventh American Joint Committee on Cancer). Although all 38 patients underwent surgery, 6 of them received palliative-intent surgery owing to inoperable primary lesions or metastases identified during the procedure. Two patients had histological findings other than adenocarcinoma, including one patient with colloid carcinoma and another with adenosquamous cell carcinoma.

2.1. Microbiome Composition Comparisons

Using 21 phyla and 353 genera based on 90 samples, we assigned 1118 and 978 operational taxonomy units (OTUs) at the phylum and genus levels, respectively. The number of OTUs was 789 and 850 in the PC and control groups, respectively. Figure 1a shows between-group differences in α -diversity (within-sample diversity) based on the Shannon index. Although there were no significant differences at the genus level ($p = 0.14$,

Wilcoxon rank-sum test), the median Shannon index was higher for the PC than for the control group at the phylum level ($p = 0.0084$, Wilcoxon rank-sum test). An overview of each level's composition is shown in Figure 1b. Figure 2 shows the multidimensional scaling plots at the OTU and genus levels, which capture β -diversity (between-sample variability). principal coordinate analysis (PCoA) based on the Aitchison and Bray-Curtis distances revealed dot patterns, suggesting that the PC patients and controls were distinct at both levels.

Table 1. Basic characteristics of patients.

(A) Age and sex of entire patients before and after PSM							
		Before PSM (n = 198)			After PSM (n = 90)		
		Pancreatic Cancer (n = 82)	Controls (n = 116)	p-Value *	Pancreatic Cancer (n = 38)	Controls (n = 52)	p-Value *
Sex	Male	55 (67.1%)	19 (16.4%)	1.131×10^{-12}	17 (44.7%)	17 (32.7%)	0.35
	Female	27 (32.9%)	97 (83.6%)		21 (55.3%)	35 (67.3%)	
Age (mean \pm SD)		63.07 \pm 9.83	49.28 \pm 12.45	1.4×10^{-12}	57.24 \pm 8.38	57.63 \pm 10.50	0.51
(B) Clinicopathologic characteristics of patients with pancreatic cancer after PSM							
		Pancreatic Cancer (n = 38)					
Age, Mean \pm SD		57.2 \pm 8.4					
Sex, M:F		17:21					
CEA > 5		5 (13.2%)					
CA 19-9 > 37		30 (78.9%)					
Neoadjuvant chemotherapy		7 (18.4%)					
Neoadjuvant radiotherapy		5 (13.2%)					
Operation		R0	26 (68.4%)				
		R1	6 (15.8%)				
		R2	6 (15.8%)				
Size, mean \pm SD		3.9 \pm 1.7					
Histology		Adenocarcinoma	36 (94.7%)				
		Others	2 (5.3%)				
Stage (AJCC stage, 7th)		I	1 (2.63%)				
		II	26 (68.4%)				
		III	5 (13.2%)				
		IV	6 (15.8%)				

(A): PSM: propensity score matching, SD: standard deviation. * Chi-square test or Wilcoxon rank-sum test. (B): SD: standard deviation, CEA: carcinoembryonic antigen, CA 19-9: carbohydrate antigen 19-9, AJCC: American Joint Committee on Cancer.

2.2. Biomarker Selection

Depending on the statistical method that was used, between-group differences in microbial composition at the phylum level included *Verrucomicrobia* (significant difference by nine methods), followed by *Deferribacteres* and *Actinobacteria* (significant by seven methods), and *Bacteroidetes* (significant by six methods) (Table 2). At the genus level, using more than six methods, differences in the following communities were found: *Stenotrophomonas/Sphingomonas/Ruminococcaceae UCG-014/Propionibacterium/Lachnospiraceae NK4A136 group* (significant by eight methods), *Akkermansia* (significant by seven methods), and *Turicibacter/Ruminiclostridium/Lachnospiraceae UCG-001/Corynebacterium* (significant by six methods). Figure 3 shows the overall log2 counts of each marker, selected through statistical methods for the PC and control groups. At the phylum level, the abundance of *Verrucomicrobia/Deferribacteres/Bacteroidetes* was more and that of *Actinobacteria* was less in the PC group than in the control group. Moreover, the abundance of *Stenotrophomonas/Sphingomonas/Propionibacterium/Corynebacterium* was decreased and that of the other six microbiota increased in the PC compared with the control group at the genus level.

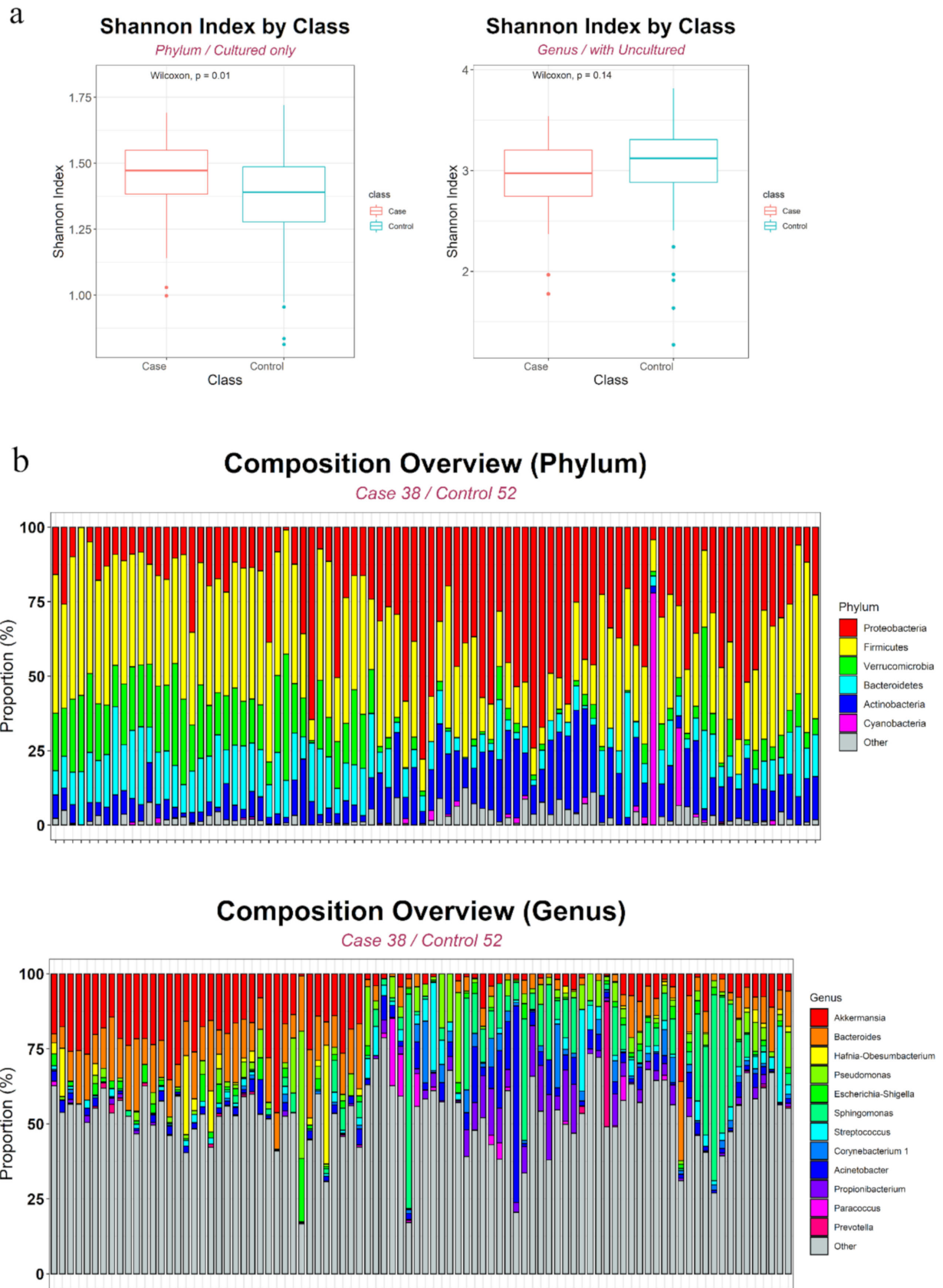


Figure 1. Alpha diversity (a) and overall microbiome composition analysis at phylum and genus level (b).

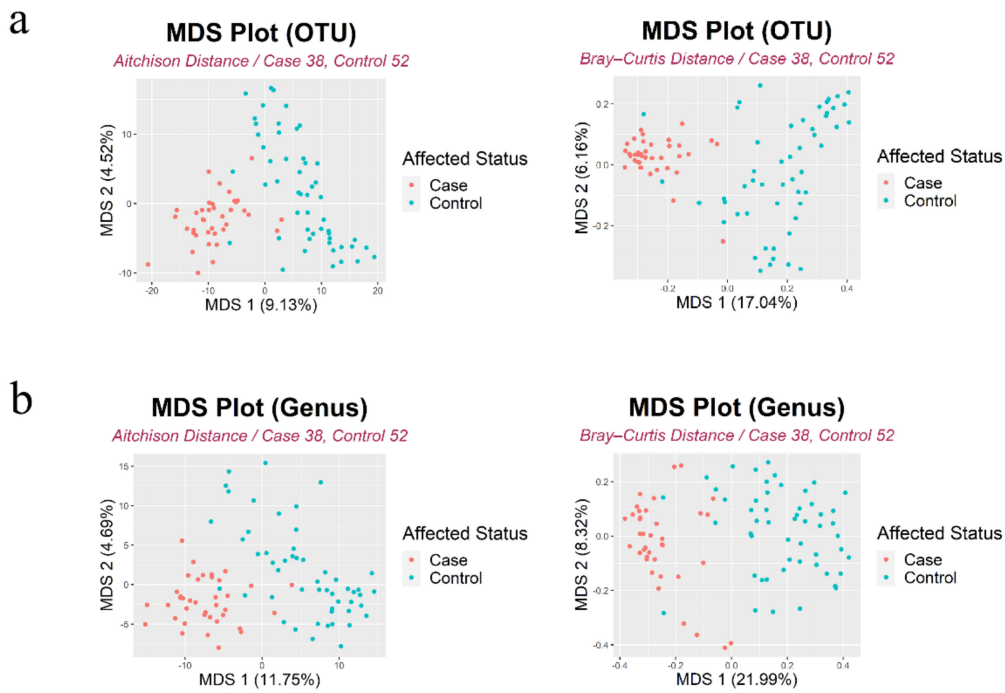


Figure 2. The beta diversity using multidimensional scaling (MDS) plot at OTU (a) and genus level (b).

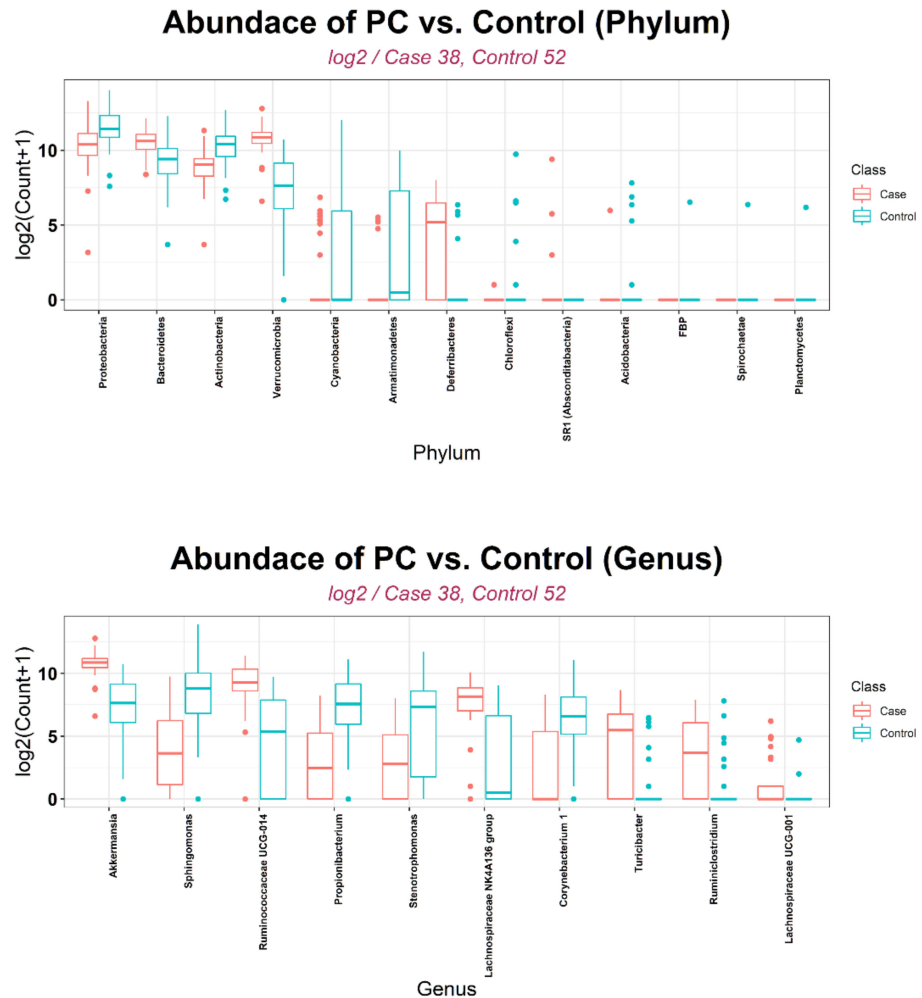


Figure 3. Log₂ counts of abundant OTUs in PC and control group.

Table 2. Selected biomarkers at phylum and genus level using 9 statistical methods. Markers were organized in the order that were significant in many methods.

(A) Phylum (L2) level											
Phylum	CLR_Perm	DESeq2_LRT	DESeq2_Wald	edgeR	Wilcoxon	ZIBSeq	ZIG_Gaussian	ZIG_log_Normal	ANCOM	Freq	Sig.
<i>Verrucomicrobia</i>	0.0000	0.0000	0.0000	0.0010	0.0000	0.0000	0.0000	0.0000	Verrucomicrobia	case	9
<i>Deferribacteres</i>	0.0000	0.0000	0.0000	0.2083	0.0019	0.5871	0.0000	0.0343	Deferribacteres	case	7
<i>Actinobacteria</i>	0.0000	0.0016	0.0011	0.0512	0.0204	0.0001	0.0057	0.5981	Actinobacteria	control	7
<i>Bacteroidetes</i>	0.0588	0.0210	0.0222	0.8176	0.0317	0.0021	0.0087	0.8280	Bacteroidetes	case	6
SR1 (<i>Absconditabacteria</i>)	0.7637	0.0000	0.0000	0.8176	0.4562	0.9999	0.0379	0.7192	-	control	3
<i>Spirochaetae</i>	0.1657	0.0022	0.0018	1.0000	0.5852	0.9999	0.0000	0.7192	-	control	3
<i>Proteobacteria</i>	0.0525	0.0032	0.0023	0.2005	0.1705	0.0013	0.0872	0.7192	-	control	3
<i>Planctomycetes</i>	0.2538	0.0026	0.0021	1.0000	0.5852	0.9999	0.0000	0.7192	-	control	3
FBP	0.1628	0.0019	0.0015	1.0000	0.5852	0.9999	0.0000	0.7192	-	control	3
<i>Cyanobacteria</i>	0.2538	0.0001	0.0000	0.1059	0.4562	0.5871	0.0010	0.5981	-	control	3
<i>Chloroflexi</i>	0.2682	0.0002	0.0001	1.0000	0.9513	0.0763	0.0000	0.7192	-	control	3
<i>Armatimonadetes</i>	0.1050	0.0001	0.0000	1.0000	0.1153	0.9999	0.0005	0.7192	-	control	3
<i>Acidobacteria</i>	0.7637	0.0009	0.0011	1.0000	0.4562	0.9999	0.0379	0.7192	-	control	3
(B) Genus (L6) level											
Genus	CLR_Perm	DESeq2_LRT	DESeq2_Wald	edgeR	Wilcoxon	ZIBSeq	ZIG_Gaussian	ZIG_log_Normal	ANCOM	Freq	Sig.
<i>Stenotrophomonas</i>	0.0159	0.0000	0.0000	1.0000	0.0095	0.0028	0.0000	8.82474E-06	Significant	control	8
<i>Sphingomonas</i>	0.0000	0.0002	0.0000	1.0000	0.0042	0.0020	0.0000	0.000394665	Significant	control	8
<i>Ruminococcaceae</i> UCG-014	0.0159	0.0000	0.0000	1.0000	0.0006	0.0000	0.0001	4.27329E-07	Significant	case	8
<i>Propionibacterium</i>	0.0000	0.0000	0.0000	1.0000	0.0004	0.0251	0.0000	0.000131155	Significant	control	8
<i>Lachnospiraceae</i> NK4A136 group	0.0000	0.0032	0.0030	1.0000	0.0043	0.0001	0.0009	0.000453009	Significant	case	8
<i>Akkermansia</i>	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.278837499	Significant	case	7
<i>Turicibacter</i>	0.0000	0.0000	0.0000	1.0000	0.0052	0.7879	0.0109	0.043444196	-	case	6
<i>Ruminiclostridium</i>	0.0000	0.0015	0.0011	1.0000	0.0306	0.1365	0.0000	0.049630223	-	case	6
<i>Lachnospiraceae</i> UCG-001	0.0422	0.0000	0.0000	1.0000	0.0361	0.9999	0.0004	0.043444196	-	case	6
<i>Corynebacterium</i> 1	0.0159	0.0003	0.0000	1.0000	0.0131	0.8506	0.0126	0.159296954	Significant	control	6

2.3. Development of PC Prediction Model

We developed prediction models using statistically significant microbiome markers (Figure A2). After an exhaustive search that considered all possible combinations using randomly separated model development (MD) and test sets, we arrived at the best models for the phylum and genus levels, respectively (Table A1). Figure 4 shows the overall performance of the model. The selected markers in the best fitting models were *Verrucomicrobia* and *Actinobacteria* at the phylum level and *Sphingomonas*, *Ruminococcaceae UCG-014*, *Propionibacterium*, *Akkermansia*, *Ruminiclostridium*, *Lachnospiraceae UCG-001*, and *Corynebacterium 1* at the genus level. Sensitivity, specificity, and accuracy estimates were average values calculated by 2-fold cross validation. In the test set, the AUC of these best models was 0.966 at the phylum level and 0.913 at the genus level. The receiver operating characteristic (ROC) curves of these models are shown in Figure 5.

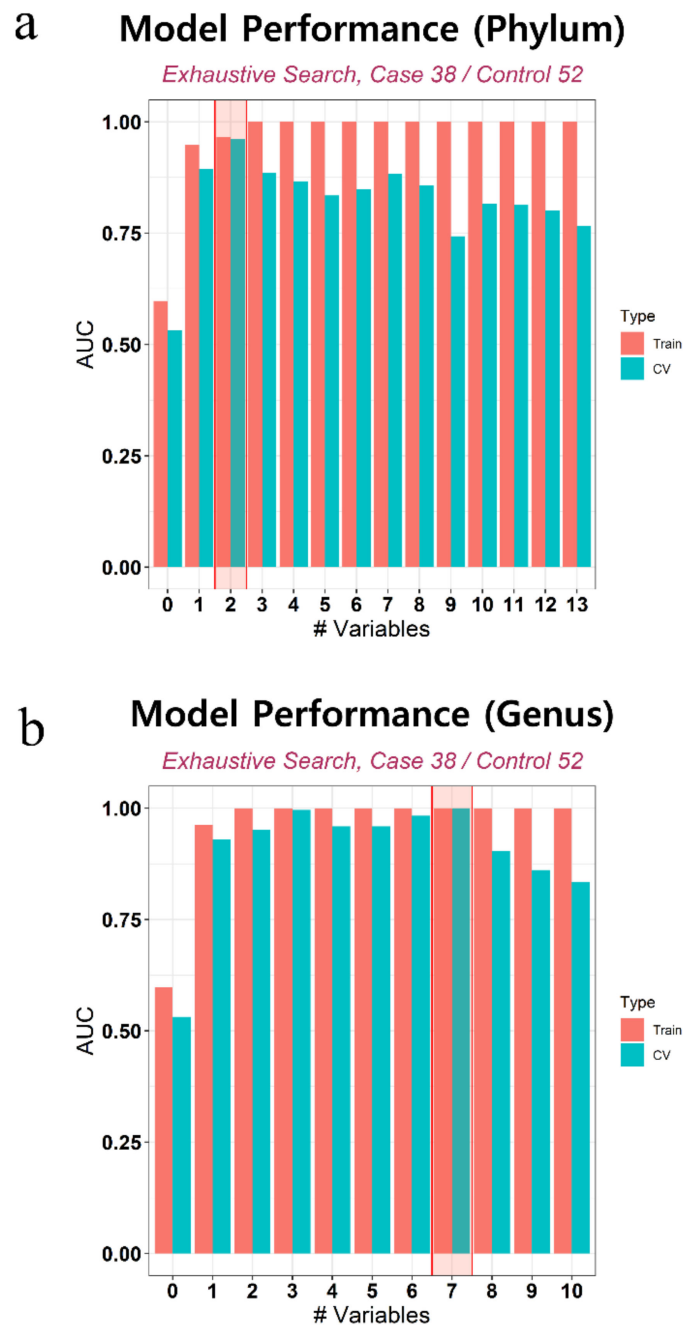


Figure 4. The bar plot describes the overall performance of models at phylum and genus level.

2.4. In Vitro Experiments to Determine the Biological Functions of Bacteria-Derived EVs

As a previous result, we found several microbiome markers that decreased from the PC patients. From this, to elucidate the role of decreased markers from PC patients, we generated and isolated the EV from *C. glutamicum* culture medium that almost same the selected marker bacteria *Corynebacterium 1*. Transmission electron microscopy image (Figure 6a), Dynamic light scattering (Figure 6b) and nanoparticle tracking analysis (Figure 6c) show that the EV from *C. glutamicum* was averagely 133.3 nm size and the 7.85×10^{11} particles were in the 1 mg/mL of the samples. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis was performed to determine protein patterns and protein size distribution (Figure 6d). Furthermore, to evaluate its functional efficacy to inhibit expression of tumor necrosis factor- α (TNF- α), which related the carcinogenesis, we used an EV obtained from *E. coli* as a stimulant to trigger the secretion of TNF- α . As a result, the amount of secreted TNF- α was reduced in all samples prepared with *C. glutamicum* EV at a concentration of 0.1 to 10 $\mu\text{g/mL}$ (Figure 6e). The amount of TNF- α showed a tendency to decrease depending on the concentration of *C. glutamicum* EV.

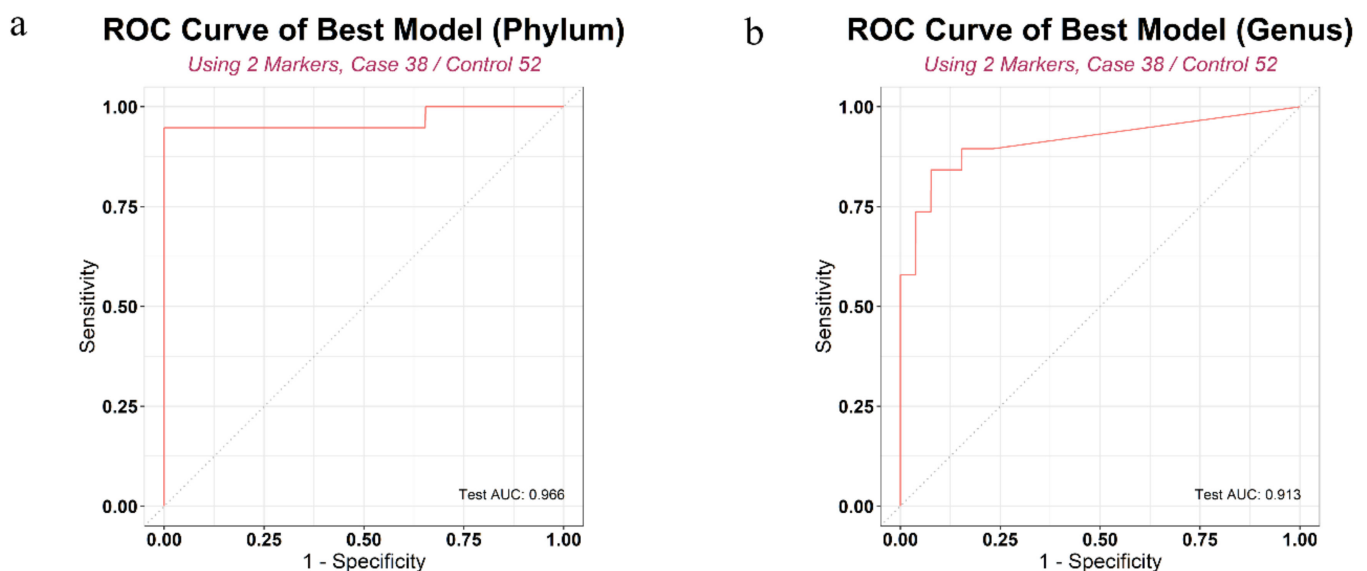


Figure 5. ROC curve of prediction model: (a) phylum level, (b) genus level. Logistic regression model was built using microbiome markers and covariates (age, gender) to distinguish PC and control groups.

2.5. Sensitivity Analysis According to Various Matching Conditions

At the phylum level, the markers (*Verrucomicrobia* and *Actinobacteria*) that were used in the final prediction models were consistently selected regardless of the changes in the calipers (Figure A3). However, among the seven markers used at the genus level, only four (*Akkermansia*, *Sphingomonas*, *Ruminococcaceae UCG-014*, and *Propionibacterium*) were selected consistently according to the changes in the calipers.

To identify the performance of the prediction models under the various calipers, we also analyzed the AUCs of the other best models that were independently selected in each condition of changed calipers, in addition to investigating the changes of the AUCs of the selected best model (with 2 markers at the phylum level and 7 at the genus level) found in this study (Figure A4). As a result, under various conditions, the testing AUCs of the microbial marker models were quite constant compared with those of the covariate model.

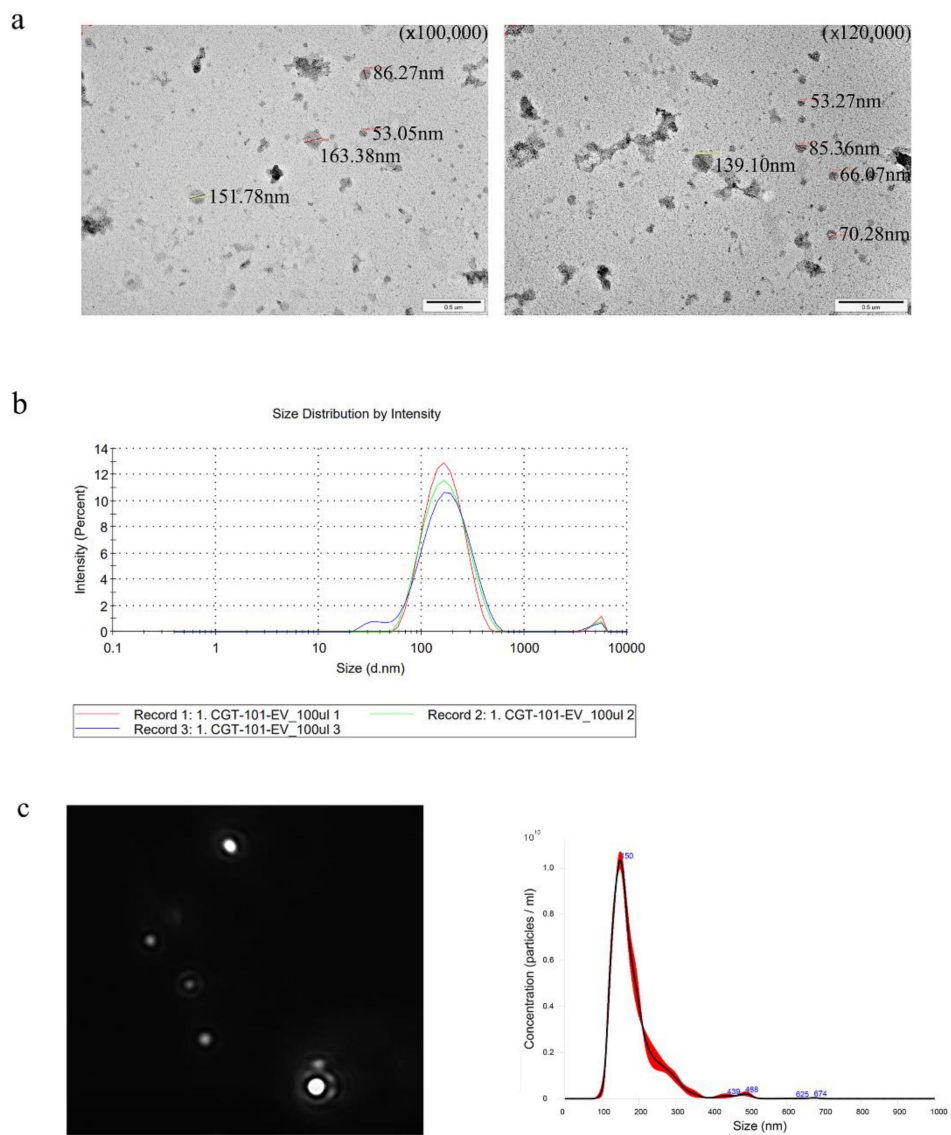


Figure 6. Cont.

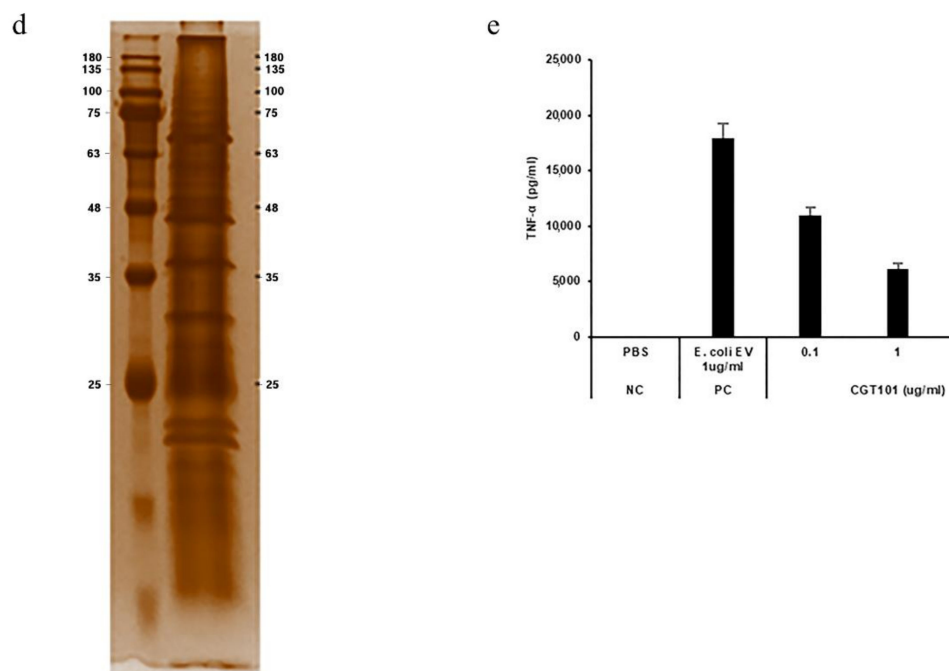


Figure 6. Biologic characteristics of *C. glutamicum* EV through in vitro experiment: (a) TEM, (b) Dynamic Light Scattering (DLS), (c) Nanoparticle Tracking Analysis (NTA), (d) Sodium Dodecyl Sulfate-PolyAcrylamide Gel Electrophoresis (SDS-PAGE), (e) The amount of TNF- α secretion according to the concentration of *C. glutamicum* EV, using *E. coli* EV as stimulant.

3. Discussion

This is the first study to investigate the altered composition of microbiomes in patients with PC and to evaluate the relationship between microbiota and PC, using bacteria-derived EVs. We performed 16S rRNA gene analysis and compared microbiome composition between patients with PC and healthy controls. Some of the compositional differences identified in this study might be candidate biomarkers of early-stage PC; using these candidate biomarkers, we proposed and validated a PC prediction model.

Previous studies have reported on the relationship between microbiota and PC or chronic pancreatitis. Microbial composition in patients with PC is known to be altered at several sites, including the oral cavity, gastrointestinal tract, and pancreatic tissues. Microbial risk factors for PC include oral microbiota in periodontal disease, altered presence of *H. pylori*, and hepatotropic viruses. Periodontitis is a form of chronic gingival inflammation and a common type of oral infection, previously linked to the risk of pancreatic [11,17–20] and other organ [8,21,22] malignancies. Several pathogens, including *P. gingivalis*, *N. elongata*, *Fusobacterium*, and *S. mitis* have been reported. Evading the host immune system, these pathogens can trigger the Toll-like receptor signaling pathways and promote pancreatic carcinogenesis in animal models [23].

H. pylori is a gut microbe that can reach the pancreas through the circulatory system or the pancreatic/biliary duct. However, the relationship between *H. pylori* and PC remains controversial. To date, it has been suggested that this relationship is mediated by specific risk factors, including infection with cytotoxin-associated gene A-negative *H. pylori* strain [24], non-O blood type [25], and smoking [26]. Moreover, hepatotropic viruses, including hepatitis B virus and hepatitis C virus, have been associated with direct and indirect (via pancreatitis) development of PC [27–29].

Previous studies of microbiota composition used fecal, salivary, biliary, or tissue samples. Findings obtained from samples collected of the digestive system might reflect the genome profiles of gut microbiota. Meanwhile, several recent studies reported on the presence of EVs that contain bacterial genome DNA fragments in serum. The size of EVs from either gram-negative or gram-positive bacteria are very small (10 to 300 nm in diameter) [30,31], allowing them to cross intestinal cellular membrane and travel throughout

the blood system [32]. Based on these findings, subsequent microbiome analyses using bacteria-derived EVs extracted from serum were undertaken [5,33].

In the present study, at the phylum level, *Actinobacteria* were less and *Verrucomicrobia* were more abundant in the PC group than in control group. At the genus level, *Akkermansia/Ruminococcaceae UCG-014/Ruminiclostridium* were more abundant and *Sphingomonas/Propionibacterium/Corynebacterium 1* were less abundant in PC patients than in the control group. Some of these findings are consistent with those from previous studies, which involved fecal microbiome analysis of PC patients [34].

Some species of *Actinobacteria* are known to produce butyrate and modulate immune function. Reduction of butyrate levels can promote inflammation, which acts as an anti-inflammatory agent mainly by blocking the activation of nuclear factor kB in intestinal epithelial cells [35]. Consistent with our findings, abundance of *Actinobacteria* in colorectal cancer patients has been reported to be lower than that in the control group [36]. In addition, while the exact mechanism remains unclear, *Akkermansia* (in *Verrucomicrobia* phylum) is known as an immune modulator, likely related to the programmed cell death protein 1 blockade pathway [37]. Finally, some *Sphingomonas* species are known to have the ability to stimulate natural killer T cells, which suppress tumor progression [38]. However, the specific function and exact mechanism associated with carcinogenesis in most taxa remain poorly understood. Future preclinical research is necessary to understand the relationship between these microbiota and PC.

As stated, few previous studies have evaluated bacteria-derived EVs acquired from blood samples; our report is the first such study on PC patients. Although we found some differences in the microbiome profiles of PC patients, we were unable to determine their biological function or behavior. Future metagenomic research should endeavor to elucidate the role of microbiota in the transition from normal to malignant tissue. In *in vitro* experiments, we found that the amount of acute phase inflammatory cytokines tended to decrease depending on the concentration of *C. glutamicum* EV. Therefore, for PC patients, a low level of *C. glutamicum* might be associated with inflammation, which can trigger cancer development.

Alongside its novelty, this study has some limitations. First, the number of samples was relatively small. We performed PSM and sensitivity analysis with various calipers to increase the generalizability of our findings. Second, although we performed internal validation with randomly separated MD and test sets, external validation with a larger cohort is required. Third, we could not explain the composition of gut microbiota directly through this study. Bacteria-derived EVs present in blood samples are believed to mostly originate from the gut microbiota. However, factors such as gut barrier, host immunity, and organ status can alter the composition of microbiota detected in the blood, making it distinct from that concurrently present in the gut. Further studies are required to describe and explain these suspected differences and their implications, alongside the functional interchangeability of gut and blood microbiota.

4. Materials and Methods

This study included patients diagnosed with PC between 2009 and 2015 at the Seoul National University Hospital and healthy controls who received regular checkups at the Seoul National University Boramae Hospital and Inje University Haeundae Paik Hospital. All patients with PC underwent surgical resection, and final pathology reports were confirmed. The control group included healthy adults without any clinical or imaging evidence of pancreatic disease or history of other cancers. Following data collection, including blood samples, the final study population was selected based on PSM analysis to reduce selection bias by equating the groups based on the covariates. This study complied with the principles of the Declaration of Helsinki and was approved by the institutional review board of Seoul National University Hospital (1601-137-739). The informed consent requirement was waived due to the retrospective nature of the study and use of anonymous clinical data.

4.1. Blood Sample Preparation and DNA Extraction

The blood samples were placed into Vacutainer EDTA tubes, and serum was centrifuged at $2000\times g$ for 15 min at $4\text{ }^{\circ}\text{C}$ to remove cell debris. The supernatant was collected and incubated with proteinase K at $56\text{ }^{\circ}\text{C}$ for 30 min. Subsequently, the samples were boiled at $100\text{ }^{\circ}\text{C}$ for 40 min to extract DNA from EVs; afterward, the supernatant was collected by centrifugation at $10,000\times g$ at $4\text{ }^{\circ}\text{C}$. A DNA isolation kit (DNeasy Blood & Tissue Kit, QIAGEN, Hilden, Germany) was used to extract the total DNA from 1 mL of supernatant. The quality and quantity of DNA were measured using the QIAxpert system (QIAGEN, Hilden, Germany).

4.2. Microbiomic Sequencing

Bacterial genomic DNA was amplified with 16S_V3_F (5-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG-3) and 16S_V4_R (5-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC-3) primers, which are specific to V3-V4 hypervariable regions of the 16S rDNA gene [39]. The libraries were prepared using polymerase chain reaction products according to the MiSeq System guide (Illumina, San Diego, CA, USA) and quantified using QIAxpert (QIAGEN, Hilden, Germany). Each amplicon was then quantified and sequenced on MiSeq (Illumina, San Diego, CA, USA) according to the manufacturer's recommendations.

4.3. Taxonomic Assignment and Profiling

Raw pyrosequencing reads from the sequencer were filtered using MiSeq (Illumina, San Diego, CA, USA), according to the barcode and primer sequences. Taxonomic assignment was performed by the profiling program MDx-Pro ver.1 (MD Healthcare, Seoul, Korea). After checking the read length (≥ 300 bp) and quality score (average Phred score ≥ 20), high-quality sequencing reads were selected. OTUs were clustered using sequence clustering algorithms CD-HIT [40].

Subsequently, taxonomic assignment was performed using UCLUST and QIIME against the GREENGENES reference database (gg_13_5_99) [41,42]. OTUs with sequences $< 0.005\%$ of the total were removed from the OTU table; a total of 1134 OTUs were obtained. Samples with a low number of read counts (< 2500) were filtered during quality control process. The resulting OTU table was used for predictive functional analysis with Tax4Fun software (metagenomics package version 0.1.014) [43].

4.4. Propensity Score Matching and Statistical Analysis

We used PSM analysis to minimize the impact of covariates on effect estimates. The propensity score is a probability that a unit with specific characteristic will be assigned to treatment group. PSM is a statistical matching technique that uses the propensity score to estimate the effectiveness of interventions, given particular covariates [44]. For example, PSM was used to adjust non-random drug assignment to determine whether drugs have the effects of protecting infants from apnea [45]. PSM was also applied to reduce the selection bias for estimating breast cancer risk in relation to antidepressant medications [46]. While PSM has been used for the case of large samples, it was shown by a simulation study that PSM performed well for the case of small samples [47]. Recently, PSM was successfully applied to fecal microbiota studies. PSM reduced the influence of lifestyle variables which might attenuate the relevance between fecal bacteria and the risk of gastric cancer [48]. PSM was also used to control the effect of clinical variables on microbiota composition to find the relationship between fecal microbiota and Parkinson's disease [49].

We applied PSM to our case-control data to reduce heterogeneity of age and sex between PC and control groups. We considered age and gender as potential confounders, because those were highly unbalanced between PC and control groups. To evaluate the consistency of estimates under varying matching conditions, sensitivity analysis was performed at various levels of matching, using calipers [50], which are values that express the strictness of covariate matching for a given propensity score. Small caliper values

represent strict matching; in contrast, large caliper values represent models that are close to those observed with random sampling. We measured trends of the selected markers and their performances and built prediction models according to the change in calipers.

All statistical analyses of clinical data were performed in R version 4.0.2 on Windows 10 (Version 4.0.2, <http://www.R-project.org> (accessed on 15 December 2020)). Categorical variables were presented as counts with percentages and compared using the chi-square test. Continuous variables were compared using a Kruskal-Wallis test or one-way analysis of variance (ANOVA). A p -value < 0.05 was considered indicative of a statistically significant difference.

The α -diversity of microbiota for each sample was evaluated using the Shannon index. The Wilcoxon rank-sum test was used to compare α -diversities between groups. Moreover, β -diversity was measured by Aitchison and Bray-Curtis distances with 90 matched samples [51,52]. Based on these distance measures, PCoA was performed, and the cmd-scale function in R was used to assess the statistical significance of separation among groups. When comparing groups and constructing models using calculated relative abundance, OTUs were used without those uncultured or unidentified. However, uncultured or unidentified OTUs were included in the analysis to provide an overview of microbial composition in each group.

4.5. Marker Selection and Prediction Model Development for PC

Using nine statistical methods, including microbiome-specific methods (metagenome-Seq [Gaussian, Log Normal], ZIBSeq, ANCOM, CLR permutation), simple statistical test (Wilcoxon rank-sum test), and conventional methods for identifying differentially expressed genes in RNAseq data (DESeq2 [LRT, Wald], edgeR), microbiome was sorted in order of compositional proportion. The relative differences in OTUs abundances between the PC and control group were investigated to identify candidate cancer biomarkers. We selected the candidate OTUs with an average relative abundance of $>1\%$ and p -value < 0.05 .

To develop a prediction model for PC, we randomly divided our samples into MD and test sets to minimize selection bias. Logistic regression analysis was performed using selected OTUs via an exhaustive search method. All possible combinations of candidate OTUs were tested by repeating 10 times for 2-fold cross validation to find the optimal variable combination to discriminate between the PC and control group. The final model was selected based on the lowest Akaike's information criteria from among the development set; it was subsequently validated using ROC curves and the AUC calculation based on the test set [53,54].

4.6. Additional In Vitro Experiments Using *Corynebacterium glutamicum* Strain

To identify the basic characteristics of bacteria-derived EVs, we performed additional in vitro experiments using *Corynebacterium glutamicum* strain (which was the same as the *Corynebacterium 1*, reported at the genus level), which was less in abundance in PC patients than in the controls. *C. glutamicum* KCTC 9097 was cultured in LB medium for 15 h at 37 °C. The cultured solution was collected and centrifuged at $10,000 \times g$ for 15 min. After filtration (0.22 μm) of the supernatant separated from the cells, EVs were extracted using ultrafiltration and ultracentrifugation methods. Five microliters of diluted EVs were dropped on Formvar-carbon coated EM grids and left aside to allow membranes adsorb for 2 min. The vesicle-coat grids were fixed with 0.25% glutaraldehyde for four minutes and washed twice with distilled water for one minute each. The grids were stained with 2% uranyl acetate at pH 7 for 5 min and viewed using a H-7650 transmission electron microscope (Hitachi, Tokyo, Japan) at a voltage of 80 kV. To identify the size and distribution of the particles of *C. glutamicum* EV, DLS and NTA were performed.

5. Conclusions

This study revealed compositional differences of microbiome between patients with PC and healthy controls, following covariate matching that reduced the impact of selection

bias. Among altered microbial communities, we identified candidate biomarkers such as *Verrucomicrobia* and *Actinobacteria* at the phylum and *Sphingomonas*, *Ruminococcaceae* UCG-014, *Propiobacterium*, *Akkermansia*, *Ruminiclostridium*, *Lachnospiraceae* UCG-001, and *Corynebacterium* at the genus level, while developing prediction models for PC. Further studies with larger cohorts are necessary to validate the present findings. Moreover, research is required into rare microbial strains whose roles in the host immune system function or in carcinogenesis remains unclear.

Author Contributions: Conceptualization, T.P. and J.-Y.J.; methodology, J.R.K., K.H., T.-S.S., H.J.P., N.K., S.L. and T.P.; formal analysis, J.R.K., K.H., T.-S.S., H.J.P., Y.H., N.K., S.L., Y.-K.K., T.P. and J.-Y.J.; investigation, J.R.K., K.H., T.-S.S., H.J.P., Y.H., H.K., W.K., Y.-K.K. and J.-Y.J.; resources, H.K., W.K., and J.-Y.J.; data curation, J.R.K., K.H., Y.H., T.-S.S., H.J.P., H.K., W.K. and Y.-K.K.; writing—original draft preparation, J.R.K., H.K., T.P. and J.-Y.J.; writing—review and editing, J.R.K., K.H., T.-S.S., H.J.P., Y.H., N.K., H.K., W.K., S.L., Y.-K.K., T.P. and J.-Y.J.; funding acquisition, T.P. and J.-Y.J. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HI16C2037).

Institutional Review Board Statement: This study protocols were approved by the Institutional Review Board at Seoul National University Hospital (IRB No. 1601-137-739).

Informed Consent Statement: The informed consent was waived under IRB approval. This study is a retrospective designed study using only data from blood samples that have already been obtained.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to patients' privacy.

Acknowledgments: The biospecimens and data used in this study were provided by the Biobank of Seoul National University Hospital, a member of Korea Biobank Network.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

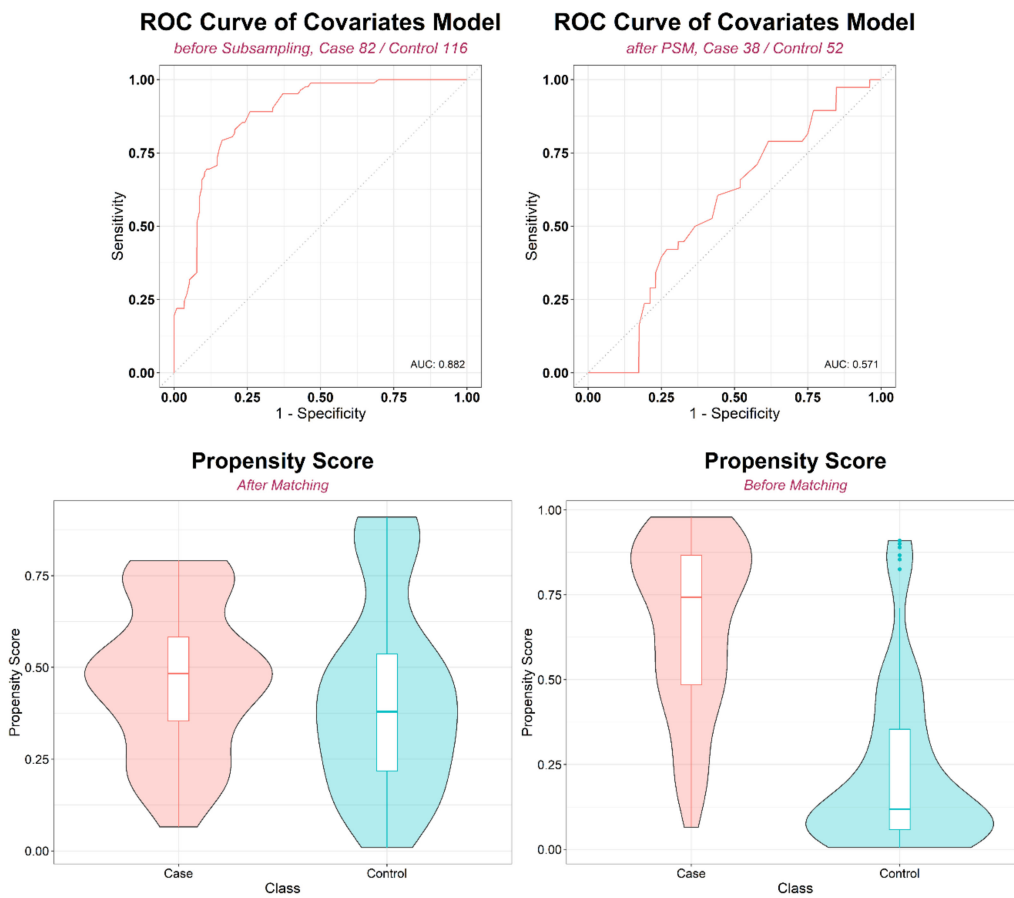


Figure A1. AUC curves and violin plots before and after covariate matching (PSM using age and sex). Based on these plots, we concluded that the balance between the two groups was achieved and matching was successful.

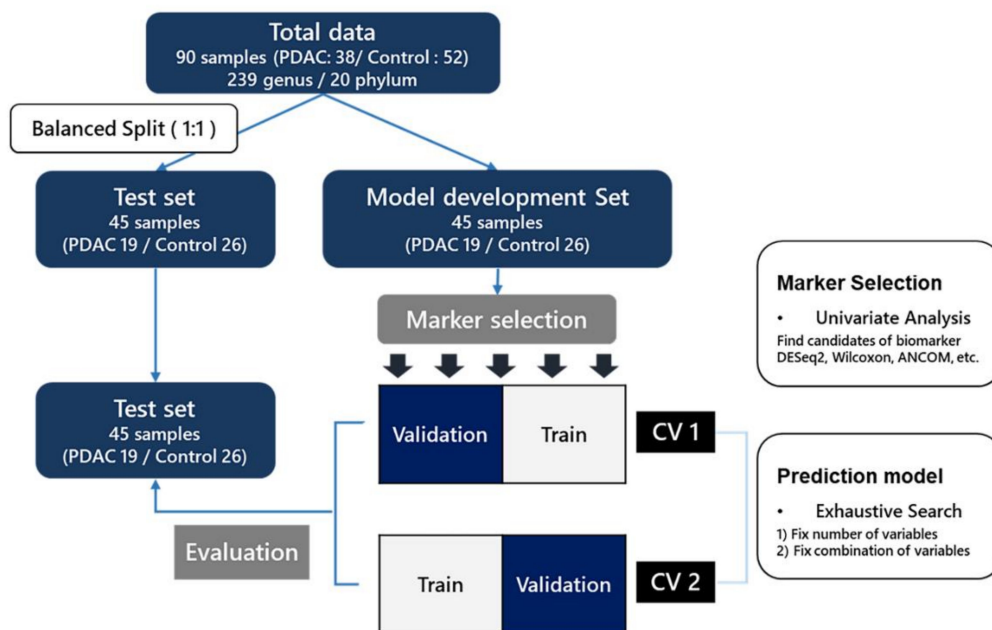


Figure A2. The scheme of prediction models. We built the models based on the statistically significant microbiome markers.

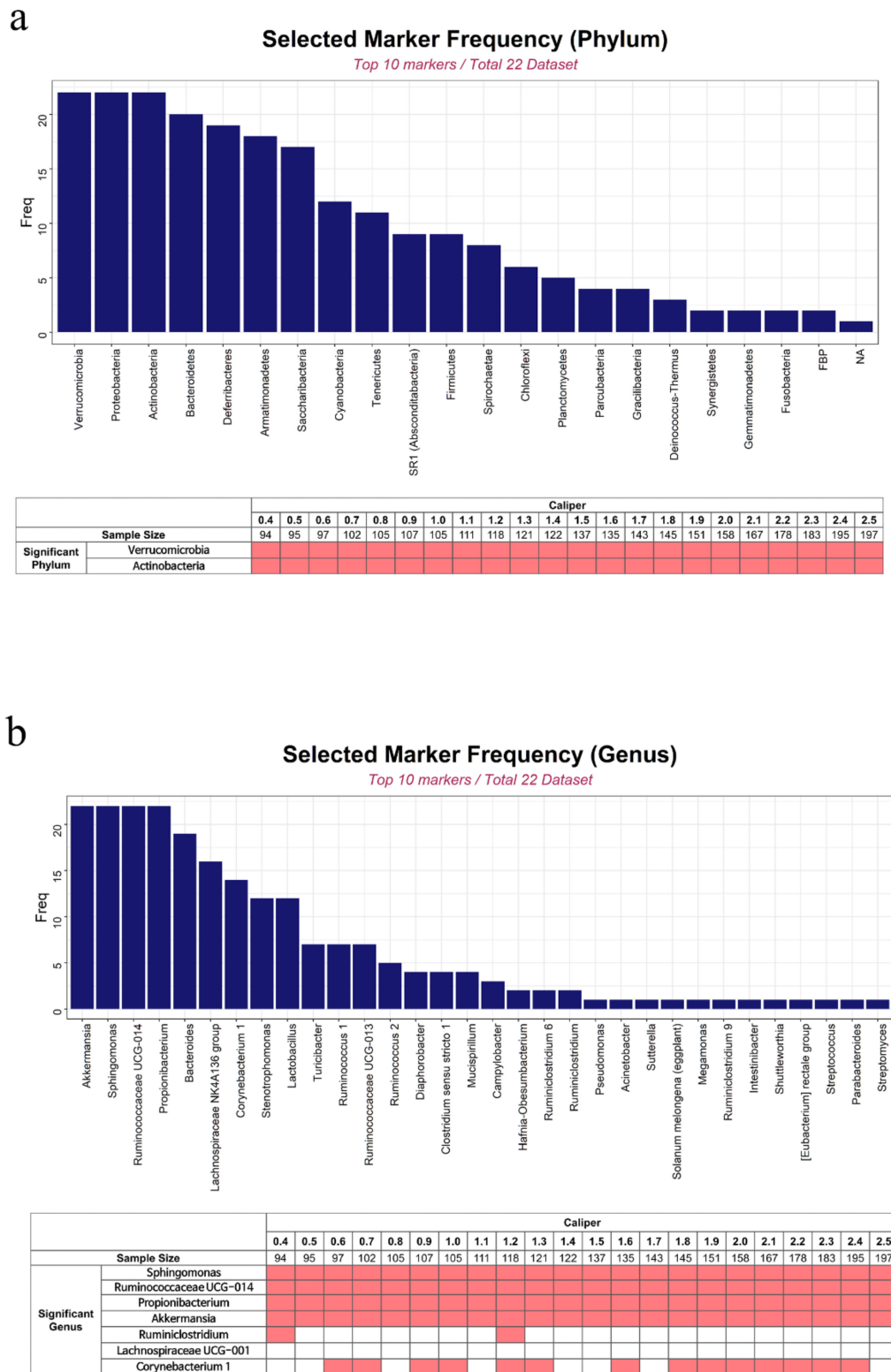


Figure A3. Selected markers by sensitivity analysis at phylum level (a) and genus level (b), using various calipers. The bar plot shows the selected frequency of each marker as the top 10 significant markers when using various calipers. The table shows whether each marker was included in the list of top 10 significant markers when using each caliper. Under the given caliper, the pink colored cells represent the significant markers, while the white cells do non-significant makers. Thus, for each marker, the more pink cells, the more robust to the change of caliper.

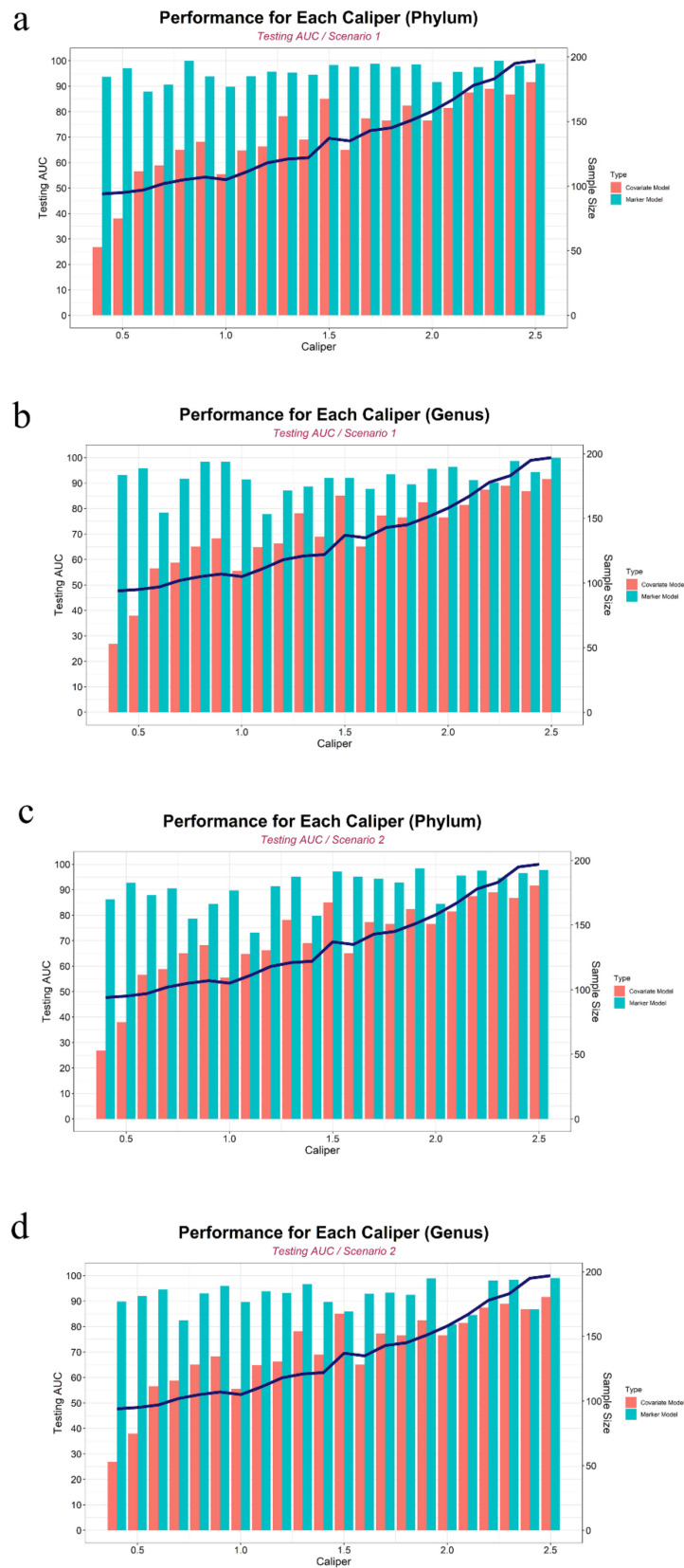


Figure A4. Testing AUCs of the prediction models under various conditions. Testing AUC of the discovered best model (using 2 markers) according to every caliper at phylum level (a) and genus level (b). Testing AUC using different models which are selected as best according to every caliper at phylum level (c) and genus level (d).

Table A1. Combinations of top prediction models and their diagnostic performance in Train and CV set.

(A) Phylum (L2) level							
Number of Variables	Model	Train Set			Cross-Validation (CV) Set		
		Sensitivity	Specificity	AUC	Sensitivity	Specificity	AUC
0	class~age + sex	0.311	0.846	0.598	0.206	0.846	0.531
1	class~age + sex + p16	0.944	0.846	0.953	0.839	0.846	0.886
2	class~age + sex + p16 + p6	1.000	0.846	0.966	1.000	0.846	0.962
3	class~age + sex + p16 + p6 + p2	1.000	0.846	1.000	0.783	0.885	0.863
4	class~age + sex + p16 + p6 + p2 + p11	1.000	0.846	1.000	0.833	0.885	0.897
5	class~age + sex + p16 + p6 + p2 + p11 + p19	1.000	0.846	1.000	0.733	0.846	0.855
6	class~age + sex + p16 + p6 + p2 + p11 + p19 + p18	1.000	0.846	1.000	0.400	0.923	0.844
7	class~age + sex + p16 + p20 + p6 + p2 + p11 + p19 + p18	1.000	0.846	1.000	0.400	0.923	0.867
8	class~age + sex + p16 + p20 + p6 + p2 + p13 + p11 + p19 + p18	1.000	0.846	1.000	0.578	0.846	0.833
9	class~age + sex + p16 + p20 + p6 + p2 + p13 + p11 + p15 + p19 + p18	1.000	0.846	1.000	0.200	0.923	0.797
10	class~age + sex + p16 + p20 + p6 + p9 + p2 + p13 + p11 + p15 + p19 + p18	1.000	0.846	1.000	0.222	0.923	0.616
11	class~age + sex + p16 + p20 + p6 + p9 + p22 + p2 + p13 + p11 + p15 + p19 + p18	1.000	0.846	1.000	0.100	0.923	0.658
12	class~age + sex + p16 + p20 + p6 + p9 + p22 + p10 + p2 + p13 + p11 + p15 + p19 + p18	1.000	0.846	1.000	0.389	0.923	0.684
13	class~age + sex + p16 + p20 + p6 + p9 + p22 + p10 + p2 + p14 + p13 + p11 + p15 + p19 + p18	1.000	0.846	1.000	0.000	1.000	0.654

Table A1. Cont.

(B) Genus (L6) level							
Number of Variables	Model	Train Set			Cross-Validation (CV) Set		
		Sensitivity	Specificity	AUC	Sensitivity	Specificity	AUC
0	class~age + sex	0.3111	0.8462	0.5976	0.2056	0.8462	0.5314
1	class~age + sex + g150	0.9444	0.8462	0.9615	0.9444	0.8462	0.9299
2	class~age + sex + g150 + g64	1.0000	0.8462	1.0000	0.9500	0.9231	0.9515
3	class~age + sex + g150 + g64 + g22	1.0000	0.8462	1.0000	1.0000	0.9231	0.9957
4	class~age + sex + g15 + g150 + g64+g22	1.0000	0.8462	1.0000	0.9444	0.9231	0.9594
5	class~age + sex + g15 + g150 + g64 + g225 + g22	1.0000	0.8462	1.0000	0.9444	0.9231	0.9594
6	class~age + sex + g210 + g23 + g150 + g64 + g225 + g22	1.0000	0.8462	1.0000	1.0000	0.9231	0.9829
7	class~age + sex + g49 + g15 + g210 + g150 + g64 + g225 + g22	1.0000	0.8462	1.0000	1.0000	0.9231	1.0000
8	class~age + sex + g19 + g49 + g15 + g210 + g150 + g64 + g225 + g22	1.0000	0.8462	1.0000	0.8444	0.8846	0.9034
9	class~age + sex + g19 + g49 + g15 + g210 + g23 + g150 + g64 + g225 + g22	1.0000	0.8462	1.0000	0.7944	0.8462	0.8613
10	class~age + sex + g19 + g49 + g15 + g210 + g23 + g150 + g59 + g64 + g225 + g22	1.0000	0.8462	1.0000	0.3500	0.9231	0.8338

The finally selected model was highlighted as bold. The OTU symbols are as follows: p16, *Verrucomicrobia*; p6, *Actinobacteria*; g49, *Sphingomonas*; g15, *Ruminococcaceae UCG-014*; g210, *Propionibacterium*; g150, *Akkermansia*; g64, *Ruminiclostridium*; g225, *Lachnospiraceae UCG-001*; g22, *Corynebacterium 1*.

References

- Rahib, L.; Smith, B.D.; Aizenberg, R.; Rosenzweig, A.B.; Fleshman, J.M.; Matrisian, L.M. Projecting cancer incidence and deaths to 2030: The unexpected burden of thyroid, liver, and pancreas cancers in the United States. *Cancer Res.* **2014**, *74*, 2913–2921. [[CrossRef](#)] [[PubMed](#)]
- Riall, T.S.; Lillemo, K.D. Underutilization of surgical resection in patients with localized pancreatic cancer. *Ann. Surg.* **2007**, *246*, 181–182. [[CrossRef](#)]
- Barteneva, N.S.; Baiken, Y.; Fasler-Kan, E.; Alibek, K.; Wang, S.; Maltsev, N.; Ponomarev, E.D.; Sautbayeva, Z.; Kauanova, S.; Moore, A.; et al. Extracellular vesicles in gastrointestinal cancer in conjunction with microbiota: On the border of Kingdoms. *Biochim. Biophys. Acta Rev. Cancer* **2017**, *1868*, 372–393. [[CrossRef](#)] [[PubMed](#)]
- Chen, Y.; Yang, F.; Lu, H.; Wang, B.; Chen, Y.; Lei, D.; Wang, Y.; Zhu, B.; Li, L. Characterization of fecal microbial communities in patients with liver cirrhosis. *Hepatology* **2011**, *54*, 562–572. [[CrossRef](#)]
- Cho, E.J.; Leem, S.; Kim, S.A.; Yang, J.; Lee, Y.B.; Kim, S.S.; Cheong, J.Y.; Cho, S.W.; Kim, J.W.; Kim, S.M.; et al. Circulating Microbiota-Based Metagenomic Signature for Detection of Hepatocellular Carcinoma. *Sci. Rep.* **2019**, *9*, 1–8. [[CrossRef](#)] [[PubMed](#)]
- Wang, Z.; Klipfell, E.; Bennett, B.J.; Koeth, R.; Levison, B.S.; Dugar, B.; Feldstein, A.E.; Britt, E.B.; Fu, X.; Chung, Y.M.; et al. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* **2011**, *472*, 57–63. [[CrossRef](#)] [[PubMed](#)]
- Zhu, J.; Liao, M.; Yao, Z.; Liang, W.; Li, Q.; Liu, J.; Yang, H.; Ji, Y.; Wei, W.; Tan, A.; et al. Breast cancer in postmenopausal women is associated with an altered gut metagenome. *Microbiome* **2018**, *6*, 1–13. [[CrossRef](#)] [[PubMed](#)]

8. Thomas, A.M.; Manghi, P.; Asnicar, F.; Pasolli, E.; Armanini, F.; Zolfo, M.; Beghini, F.; Manara, S.; Karcher, N.; Pozzi, C.; et al. Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation. *Nat. Med.* **2019**, *25*, 667–678. [[CrossRef](#)]
9. Yu, G.; Murphy, G.; Michel, A.; Weinstein, S.J.; Männistö, S.; Albanes, D.; Pawlita, M.; Stolzenberg-Solomon, R.Z. Seropositivity to *Helicobacter pylori* and risk of pancreatic cancer. *Cancer Epidemiol. Biomark. Prev.* **2013**, *22*, 2416–2419. [[CrossRef](#)]
10. Ahn, J.; Segers, S.; Hayes, R.B. Periodontal disease, *Porphyromonas gingivalis* serum antibody levels and orodigestive cancer mortality. *Carcinogenesis* **2012**, *33*, 1055–1058. [[CrossRef](#)]
11. Farrell, J.J.; Zhang, L.; Zhou, H.; Chia, D.; Elashoff, D.; Akin, D.; Paster, B.J.; Joshipura, K.; Wong, D.T. Variations of oral microbiota are associated with pancreatic diseases including pancreatic cancer. *Gut* **2012**, *61*, 582–588. [[CrossRef](#)] [[PubMed](#)]
12. Charbonneau, M.R.; Blanton, L.V.; DiGiulio, D.B.; Relman, D.A.; Lebrilla, C.B.; Mills, D.A.; Gordon, J.I. A microbial perspective of human developmental biology. *Nature* **2016**, *535*, 48–55. [[CrossRef](#)]
13. Kang, C.S.; Ban, M.; Choi, E.J.; Moon, H.G.; Jeon, J.S.; Kim, D.K.; Park, S.K.; Jeon, S.G.; Roh, T.Y.; Myung, S.J.; et al. Extracellular vesicles derived from gut microbiota, especially *Akkermansia muciniphila*, protect the progression of dextran sulfate sodium-induced colitis. *PLoS ONE* **2013**, *8*, e76520. [[CrossRef](#)] [[PubMed](#)]
14. Yoo, J.Y.; Rho, M.; You, Y.A.; Kwon, E.J.; Kim, M.H.; Kym, S.; Jee, Y.K.; Kim, Y.K.; Kim, Y.J. 16S rRNA gene-based metagenomic analysis reveals differences in bacteria-derived extracellular vesicles in the urine of pregnant and non-pregnant women. *Exp. Mol. Med.* **2016**, *48*, e208. [[CrossRef](#)] [[PubMed](#)]
15. Andaloussi, S.E.L.; Mäger, I.; Breakefield, X.O.; Wood, M.J. Extracellular vesicles: Biology and emerging therapeutic opportunities. *Nat. Rev. Drug Discov.* **2013**, *12*, 347–357. [[CrossRef](#)]
16. Kim, S.I.; Kang, N.; Leem, S.; Yang, J.; Jo, H.; Lee, M.; Kim, H.S.; Dhanasekaran, D.N.; Kim, Y.K.; Park, T.; et al. Metagenomic Analysis of Serum Microbe-Derived Extracellular Vesicles and Diagnostic Models to Differentiate Ovarian Cancer and Benign Ovarian Tumor. *Cancers* **2020**, *12*, 1309. [[CrossRef](#)]
17. Michaud, D.S.; Izard, J.; Wilhelm-Benartzi, C.S.; You, D.H.; Grote, V.A.; Tjønneland, A.; Dahm, C.C.; Overvad, K.; Jenab, M.; Fedirko, V.; et al. Plasma antibodies to oral bacteria and risk of pancreatic cancer in a large European prospective cohort study. *Gut* **2013**, *62*, 1764–1770. [[CrossRef](#)] [[PubMed](#)]
18. Bracci, P.M. Oral Health and the Oral Microbiome in Pancreatic Cancer: An Overview of Epidemiological Studies. *Cancer J.* **2017**, *23*, 310–314. [[CrossRef](#)] [[PubMed](#)]
19. Fan, X.; Alekseyenko, A.V.; Wu, J.; Peters, B.A.; Jacobs, E.J.; Gapstur, S.M.; Purdue, M.P.; Abnet, C.C.; Stolzenberg-Solomon, R.; Miller, G.; et al. Human oral microbiome and prospective risk for pancreatic cancer: A population-based nested case-control study. *Gut* **2018**, *67*, 120–127. [[CrossRef](#)]
20. Wei, M.Y.; Shi, S.; Liang, C.; Meng, Q.C.; Hua, J.; Zhang, Y.Y.; Liu, J.; Zhang, B.; Xu, J.; Yu, X.J.; et al. The microbiota and microbiome in pancreatic cancer: More influential than expected. *Mol. Cancer* **2019**, *18*, 1–15. [[CrossRef](#)]
21. Flemer, B.; Warren, R.D.; Barrett, M.P.; Cisek, K.; Das, A.; Jeffery, I.B.; Hurley, E.; O’Riordain, M.; Shanahan, F.; O’Toole, P.W. The oral microbiota in colorectal cancer is distinctive and predictive. *Gut* **2018**, *67*, 1454–1463. [[CrossRef](#)]
22. Tezal, M.; Sullivan, M.A.; Hyland, A.; Marshall, J.R.; Stoler, D.; Reid, M.E.; Loree, T.R.; Rigual, N.R.; Merzianu, M.; Hauck, L. Chronic periodontitis and the incidence of head and neck squamous cell carcinoma. *Cancer Epidemiol. Biomark. Prev.* **2009**, *18*, 2406–2412. [[CrossRef](#)] [[PubMed](#)]
23. Zambirinis, C.P.; Levie, E.; Nguy, S.; Avanzi, A.; Barilla, R.; Xu, Y.; Seifert, L.; Daley, D.; Greco, S.H.; Deutsch, M.; et al. TLR9 ligation in pancreatic stellate cells promotes tumorigenesis. *J. Exp. Med.* **2015**, *212*, 2077–2094. [[CrossRef](#)] [[PubMed](#)]
24. Xiao, M.J.; Wang, Y.M.; Gao, Y. Association between *Helicobacter pylori* Infection and Pancreatic Cancer Development: A Meta-Analysis. *PLoS ONE.* **2013**, *8*, e75559. [[CrossRef](#)] [[PubMed](#)]
25. Risch, H.A. Pancreatic cancer: *Helicobacter pylori* colonization, N-nitrosamine exposures, and ABO blood group. *Mol. Carcinog.* **2012**, *51*, 109–118. [[CrossRef](#)] [[PubMed](#)]
26. Lindkvist, B.; Johansen, D.; Borgström, A.; Manjer, J. A prospective study of *Helicobacter pylori* in relation to the risk for pancreatic cancer. *BMC Cancer* **2008**, *8*, 1–9. [[CrossRef](#)]
27. Chen, Y.W.; Bai, X.; Zhang, Q.; Wen, L.; Su, W.; Fu, Q.; Sun, X.; Lou, Y.; Yang, J.; Zhang, J.; et al. The hepatitis B virus X protein promotes pancreatic cancer through modulation of the PI3K/AKT signaling pathway. *Cancer Lett.* **2016**, *380*, 98–105. [[CrossRef](#)] [[PubMed](#)]
28. Jin, Y.; Gao, H.; Chen, H.; Wang, J.; Chen, M.; Li, G.; Wang, L.; Gu, J.; Tu, H. Identification and impact of hepatitis B virus DNA and antigens in pancreatic cancer tissues and adjacent non-cancerous tissues. *Cancer Lett.* **2013**, *335*, 447–454. [[CrossRef](#)] [[PubMed](#)]
29. Hassan, M.M.; Li, D.; El-Deeb, A.S.; Wolff, R.A.; Bondy, M.L.; Davila, M.; Abbruzzese, J.L. Association between hepatitis B virus and pancreatic cancer. *J. Clin. Oncol.* **2008**, *26*, 4557–4562. [[CrossRef](#)] [[PubMed](#)]
30. Kulp, A.; Kuehn, M.J. Biological functions and biogenesis of secreted bacterial outer membrane vesicles. *Ann. Rev. Microbiol.* **2010**, *64*, 163–184. [[CrossRef](#)] [[PubMed](#)]
31. Brown, L.; Wolf, J.M.; Prados-Rosales, R.; Casadevall, A. Through the wall: Extracellular vesicles in Gram-positive bacteria, mycobacteria and fungi. *Nat. Rev. Microbiol.* **2015**, *13*, 620–630. [[CrossRef](#)] [[PubMed](#)]
32. Jang, S.C.; Kim, S.R.; Yoon, Y.J.; Park, K.S.; Kim, J.H.; Lee, J.; Kim, O.Y.; Choi, E.J.; Kim, D.K.; Choi, D.S. In vivo kinetic biodistribution of nano-sized outer membrane vesicles derived from bacteria. *Small* **2015**, *11*, 456–461. [[CrossRef](#)] [[PubMed](#)]

33. Lee, H.; Lee, H.K.; Min, S.K.; Lee, W.H. 16S rDNA microbiome composition pattern analysis as a diagnostic biomarker for biliary tract cancer. *World J. Surg. Oncol.* **2020**, *18*, 19. [[CrossRef](#)]
34. Half, E.; Keren, N.; Reshef, L.; Dorfman, T.; Lachter, I.; Kluger, Y.; Reshef, N.; Knobler, H.; Maor, Y.; Stein, A.; et al. Fecal microbiome signatures of pancreatic cancer patients. *Sci. Rep.* **2019**, *9*, 1–12. [[CrossRef](#)] [[PubMed](#)]
35. Inan, M.S.; Rasoulpour, R.J.; Yin, L.; Hubbard, A.K.; Rosenberg, D.W.; Giardina, C. The luminal short-chain fatty acid butyrate modulates NF-kappaB activity in a human colonic epithelial cell line. *Gastroenterology* **2000**, *118*, 724–734. [[CrossRef](#)]
36. Saus, E.; Iraola-Guzmán, S.; Willis, J.R.; Brunet-Vega, A.; Gabaldón, T. Microbiome and colorectal cancer: Roles in carcinogenesis and clinical potential. *Mol. Aspects Med.* **2019**, *69*, 93–106. [[CrossRef](#)]
37. Routy, B.; Le Chatelier, E.; Derosa, L.; Duong, C.P.M.; Alou, M.T.; Daillère, R.; Fluckiger, A.; Messaoudene, M.; Rauber, C.; Roberti, M.P.; et al. Gut microbiome influences efficacy of PD-1-based immunotherapy against epithelial tumors. *Science* **2018**, *359*, 91–97. [[CrossRef](#)]
38. Khan, M.A.; Aljarbou, A.N.; Aldebasi, Y.H.; Alorainy, M.S.; Rahmani, A.H.; Younus, H.; Khan, A. Liposomal formulation of glycosphingolipids from *Sphingomonas paucimobilis* induces antitumour immunity in mice. *J. Drug Target.* **2018**, *26*, 709–719. [[CrossRef](#)]
39. Nossa, C.W.; Oberdorf, W.E.; Yang, L.; Aas, J.A.; Paster, B.J.; Desantis, T.Z.; Brodie, E.L.; Malamud, D.; Poles, M.A.; Pei, Z. Design of 16S rRNA gene primers for 454 pyrosequencing of the human foregut microbiome. *World J. Gastroenterol.* **2010**, *16*, 4135–4144. [[CrossRef](#)]
40. Li, W.; Godzik, A. Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **2006**, *22*, 1658–1659. [[CrossRef](#)]
41. Edgar, R.C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **2010**, *26*, 2460–2461. [[CrossRef](#)] [[PubMed](#)]
42. Caporaso, J.G.; Kuczynski, J.; Stombaugh, J.; Bittinger, K.; Bushman, F.D.; Costello, E.K.; Fierer, N.; Peña, A.G.; Goodrich, J.K.; Gordon, J.I.; et al. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **2010**, *7*, 335–336. [[CrossRef](#)] [[PubMed](#)]
43. Asshauer, K.P.; Wemheuer, B.; Daniel, R.; Meinicke, P. Tax4Fun: Predicting functional profiles from metagenomic 16S rRNA data. *Bioinformatics* **2015**, *31*, 2882–2884. [[CrossRef](#)] [[PubMed](#)]
44. Rosenbaum, P.R.; Rubin, D.B. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* **1983**, *70*, 41–55. [[CrossRef](#)]
45. Matthew, C.W.; Amy, T.D.; Polly, A.N.; Ronald, G.; Mari, P. Using propensity scores to reduce case-control selection bias. *Epidemiology* **2012**, *23*, 772.
46. Paul, W.; Lucas, S.; Mark, L.; Kim, T.; Ben, M.; Stephen, J.R. A matched case control study with propensity score balancing examining the protective effect of paracetamol against parentally reported apnoea in infants. *Resuscitation* **2012**, *83*, 440–446.
47. Romain, P.; Matthieu, R.; Sylvie, C. Evaluation of the propensity score methods for estimating marginal odds ratios in case of small sample size. *BMC Med. Res. Methodol.* **2012**, *12*, 1–10.
48. Juan, W.; Cong, Z.; Shuo, X.; Chunjie, X.; Ruiping, W.; Dongqing, Y.; Bin, L.; Liyun, S.; Ruimin, T.; Yuhao, T.; et al. Fecal Microbiome Alteration May Be a Potential Marker for Gastric Cancer. *Dis. Markers* **2020**, *2020*, 3461315.
49. Cosma-Grigorov, A.; Meixner, H.; Mrochen, A.; Wirtz, S.; Winkler, J.; Marxreiter, F. Changes in Gastrointestinal Microbiome Composition in PD: A Pivotal Role of Covariates. *Front. Neurol.* **2020**, *11*, 1041. [[CrossRef](#)]
50. Althausen, R.P.; Rubin, D. Computerized Construction of a Matched Sample. *Am. J. Sociol.* **1970**, *76*, 325–346. [[CrossRef](#)]
51. Aitchison, J.; Barceló-Vidal, C.; Martín-Fernández, J.A.; Pawlowsky-Glahn, V. Logratio analysis and compositional distance. *Math. Geol.* **2000**, *32*, 271–275. [[CrossRef](#)]
52. Bray, J.R.; Curtis, J.T. An Ordination of the Upland Forest Communities of Southern Wisconsin. *Ecol. Monographs* **1957**, *27*, 325–349. [[CrossRef](#)]
53. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Contr.* **1974**, *19*, 716–723. [[CrossRef](#)]
54. Green, D.M.; Swets, J.A. *Signal Detection Theory and Psychophysics*, 1st ed.; John Wiley: Oxford, UK, 1966; p. 455.